Ergod. Th. & Dynam. Sys., (2026), **46**, 263–290 © The Author(s), 2025. Published by Cambridge University Press.
doi:10.1017/etds.2025.10206

Ergodic optimization for continuous functions on non-Markov shifts

MAO SHINODA†, HIROKI TAKAHASI‡ and KENICHIRO YAMAMOTO§

† *Department of Mathematics, Ochanomizu University, 2-1-1 Otsuka, Bunkyo-ku, Tokyo 112-8610, Japan*
(e-mail: shinoda.mao@ocha.ac.jp)
‡ *Department of Mathematics, Keio University, Yokohama 223-8522, Japan*
(e-mail: hiroki@math.keio.ac.jp)
§ *Department of General Education, Nagaoka University of Technology, Nagaoka 940-2188, Japan*
(e-mail: k_yamamoto@vos.nagaokaut.ac.jp)

(*Received* 8 *June* 2024 *and accepted in revised form* 11 *June* 2025)

Abstract. Ergodic optimization aims to describe dynamically invariant probability measures that maximize the integral of a given function. For a wide class of intrinsically ergodic subshifts over a finite alphabet, we show that the space of continuous functions on the shift space contains two disjoint subsets: one is a dense G_δ set for which all maximizing measures have 'relatively small' entropy; the other is the set of functions having uncountably many, fully supported ergodic maximizing measures with 'relatively large' entropy. This result generalizes and unifies the results of Morris [*Discrete Contin. Dyn. Syst.* **27** (2010), 383–388] and Shinoda [*Nonlinearity* **31** (2018), 2192–2200] on symbolic dynamics, and applies to a wide class of intrinsically ergodic non-Markov symbolic dynamics without the Bowen specification property, including any transitive piecewise monotonic interval map, some coded shifts, and multidimensional β-transformations. Along with these examples of application, we provide an example of an intrinsically ergodic subshift with positive obstruction entropy to specification.

Key words: ergodic optimization, subshift, thermodynamic formalism, specification
2020 Mathematics Subject Classification: 37B10 (Primary); 37D35 (Secondary)

1. *Introduction*

Ergodic optimization aims to describe properties of dynamically invariant maximizing measures. In its most basic form, the main constituent components are: a continuous map T of a compact metric space X; the space $\mathcal{M}(X, T)$ of T-invariant Borel probability measures endowed with the weak* topology; a continuous function $f : X \to \mathbb{R}$. Elements

https://doi.org/10.1017/etds.2025.10206 Published online by Cambridge University Press

of $\mathcal{M}(X, T)$ that attain the supremum

$$\Lambda_T(f) = \sup \left\{ \int f \, d\mu \colon \mu \in \mathcal{M}(X, T) \right\} \tag{1.1}$$

are called *f-maximizing measures*. The set of f-maximizing measures, denoted by $\mathcal{M}_{\max}(f)$, is non-empty and contains ergodic measures. For a given dynamical system (X, T), we aim to establish properties of elements of $\mathcal{M}_{\max}(f)$ for a 'typical' f in a Banach space of real-valued functions on X.

Earlier results in ergodic optimization, especially those by Hunt and Ott [24, 25] and Bousch [4], suggested the possibility of maximizing measures for typical functions being unique and supported on periodic orbits. It has been conjectured that for typical dynamical systems and typical functions in suitable Banach spaces, the maximizing measure is unique and supported on a periodic orbit. For a more precise statement known as the TPO (typically periodic optimization) conjecture, see [27, Conjecture 4.11], [28, Conjecture 7.3], and [45, Conjecture 1]. For results toward a resolution of this conjecture in a variety of settings, see [3, 27, 28] and the references therein.

In ergodic optimization, the regularity of functions is crucial. For (X, T) with some expanding or hyperbolic behavior and a Hölder continuous f, the Mañé–Conze–Guivarc'h lemma characterizes f-maximizing measures via their supports [4, 5, 18, 33]. The analysis of functions in the space $C(X)$ of real-valued continuous functions on X endowed with the supremum norm $\| \cdot \|_{C^0}$ is completely different: the Mañé–Conze–Guivarc'h lemma is not valid, but duality arguments are available.

Recall that dense G_δ sets are countable intersections of open dense subsets and a property that holds for a dense G_δ set is said to be generic. For an expanding map of the circle, Bousch and Jenkinson [6] proved that the maximizing measure is unique and fully supported (charging any non-empty open set) for generic continuous functions. Brémont [8] proved that if (X, T) has the property that measures supported on periodic orbits (closed orbit measures) are dense in $\mathcal{M}(X, T)$ and the entropy function is upper semicontinuous, then the maximizing measure is unique and has zero entropy for generic continuous functions. For expanding maps of the circle, the Bowen specification property [7] holds and so closed orbit measures are dense [40]. Moreover, the entropy function is upper semicontinuous. Hence, the result of Brémont [8] applies together with that of Bousch and Jenkinson [6]: the maximizing measure is unique, fully supported, has zero entropy, and is not strongly mixing for generic continuous functions.

Morris [34] unified the result of Bousch and Jenkinson [6] and that of Brémont [8], proving that for (X, T) with the Bowen specification property and for a generic continuous function, the maximizing measure is unique, fully supported, has zero entropy, and is not strongly mixing [34, Corollary 2]. In contrast to Morris's result, Shinoda [37, Theorem A] proved that for a dense set of continuous functions on a topologically mixing Markov shift, there exist uncountably many, fully supported ergodic maximizing measures with positive entropy, which are actually Bernoulli and hence strongly mixing. For an analogous result on expanding Markov interval maps with holes, see [38].

https://doi.org/10.1017/etds.2025.10206 Published online by Cambridge University Press

1.1. *Statements of the results.* In the context of symbolic dynamics, our main result below generalizes and unifies Morris's result [34, Corollary 2] and Shinoda's one [37, Theorem A] concerning entropies and supports of maximizing measures for continuous functions, to include a wide class of non-Markov shifts without the Bowen specification property. For a subshift Σ, let $h_{\mathrm{top}}(\Sigma)$ denote its topological entropy, and for a shift-invariant Borel probability measure μ on Σ, let $h(\mu)$ denote its measure-theoretic entropy. Let $h^{\perp}_{\mathrm{spec}}(\Sigma)$ denote *the obstruction entropy to specification* on Σ (see §2.3).

THEOREM A. *Let Σ be a subshift that satisfies $h^{\perp}_{\mathrm{spec}}(\Sigma) < h_{\mathrm{top}}(\Sigma)$. For any $H \in [h^{\perp}_{\mathrm{spec}}(\Sigma), h_{\mathrm{top}}(\Sigma))$, the following statements hold.*

(a) *The set $\mathscr{R}_H = \{f \in C(\Sigma) : h(\mu) \leq H \text{ for all} \mu \in \mathcal{M}_{\max}(f)\}$ is dense G_δ.*

(b) *For any $f \in C(\Sigma) \setminus \mathscr{R}_H$ and any neighborhood U of f in $C(\Sigma)$, there exists $g \in U$ such that the set $\{\mu \in \mathcal{M}_{\max}(g) : h(\mu) > H\}$ contains uncountably many, fully supported ergodic measures.*

Several remarks are in order on the meaning and the scope of Theorem A. First of all, the condition $h^{\perp}_{\mathrm{spec}}(\Sigma) < h_{\mathrm{top}}(\Sigma)$ together with the obstruction entropy to specification was introduced by Climenhaga and Thompson in the series of papers [14–16] on the thermodynamic formalism for dynamical systems with non-uniform specification. It implies that for any $H \in [h^{\perp}_{\mathrm{spec}}(\Sigma), h_{\mathrm{top}}(\Sigma))$, the language $\mathcal{L}(\Sigma)$ of the shift space admits a decomposition $\mathcal{L}(\Sigma) = \mathcal{C}^p\mathcal{G}\mathcal{C}^s$ into a prefix $\mathcal{C}^p$, a 'good' core $\mathcal{G}$ and a suffix $\mathcal{C}^s$, such that on a 'fattened' $\mathcal{G}$, there is a specification and the exponential growth rate of the cardinality of $\mathcal{C}^p \cup \mathcal{C}^s$ per word length does not exceed H. See §2.2 for a formal definition. The condition $h^{\perp}_{\mathrm{spec}}(\Sigma) < h_{\mathrm{top}}(\Sigma)$ implies the intrinsic ergodicity of Σ [15, Theorem C], that is, the uniqueness of the measure of maximal entropy.

In the case where $h^{\perp}_{\mathrm{spec}}(\Sigma) = 0$, one can take $H = 0$. By Theorem A(a), the set $\mathscr{R}_0 = \{f \in C(\Sigma) : h(\mu) = 0 \text{ for all} \mu \in \mathcal{M}_{\max}(f)\}$ is dense G_δ, which in part generalizes Morris's result [34, Corollary 2]. By Theorem A(b), any $f \in C(\Sigma) \setminus \mathscr{R}_0$ is approximated by another for which there exist uncountably many, fully supported ergodic measures with positive entropy. This in part generalizes Shinoda's result [37, Theorem A]. By virtue of the constant H, analogous statements hold in the case $h^{\perp}_{\mathrm{spec}}(\Sigma) > 0$ as stated in Theorem A.

Theorem A does not assert the non-mixingness or Bernoulliness of the maximizing measures. Combining an argument in the proof of [34, Corollary 2] with [36, Theorem 3.4], one can show that if closed orbit measures are dense in the space of ergodic measures, then any maximizing measure of a generic continuous function is not strongly mixing. We do not know if the uncountably many maximizing measures in Theorem A(b) are Bernoulli.

As a consequence of the variational principle and the result of Jenkinson [27, Theorem 3.7], the set $C(\Sigma) \setminus \mathscr{R}_H$ is non-empty. It is difficult to exhibit a pathological continuous function as in Theorem A(b) for which there exist uncountably many maximizing measures that are fully supported and have positive entropy. Trivial examples with this property are functions that are homologous to a constant. The Mañé–Conze–Guivarc'h lemma implies that any continuous function that is not homologous to a constant and has a fully supported maximizing measure cannot be Hölder continuous.

https://doi.org/10.1017/etds.2025.10206 Published online by Cambridge University Press

The next theorem provides a sufficient condition for the density of such pathological continuous functions. Let Σ be a subshift and let σ denote the left shift acting on Σ. We say *ergodic measures on* Σ *are entropy dense* if for any non-ergodic measure $\mu \in \mathcal{M}(\Sigma, \sigma)$, any $\varepsilon > 0$, and any neighborhood U of μ in $\mathcal{M}(\Sigma, \sigma)$, there exists an ergodic measure $\nu \in U$ such that $h(\nu) > h(\mu) - \varepsilon$. Clearly, the entropy density implies the density of ergodic measures.

THEOREM B. *Let* Σ *be a subshift with positive topological entropy and* $h^{\perp}_{\mathrm{spec}}(\Sigma) = 0$. *Suppose that ergodic measures on* Σ *are entropy dense. Then, there exists a dense subset* $\mathscr{D}$ *of* $C(\Sigma)$ *such that for any* $f \in \mathscr{D}$, $\mathcal{M}_{\max}(f)$ *contains uncountably many fully supported ergodic measures with positive entropy.*

There is a wealth of examples of symbolic dynamics to which Theorem A or Theorem B applies: piecewise monotonic maps [19], some coded shifts including S-gap shifts [32], multidimensional β-transformations [9, 10], and so on. See §4 for more details on these applications.

We will provide two examples of intrinsically ergodic subshifts to which Theorem A applies but Theorem B does not. One is a subshift without the entropy density (see §4.2). The other example is a subshift with positive obstruction entropy to specification, which is stated as follows.

THEOREM C. *For any integer* $N \geq 3$, *there exists a subshift* Σ *on* N *symbols such that* $(1/6)\log(N-1) \leq h^{\perp}_{\mathrm{spec}}(\Sigma) \leq \log(N-1) < h_{\mathrm{top}}(\Sigma)$.

Theorem C has an independent interest. Any transitive Markov shift is intrinsically ergodic [35], and its obstruction entropy to specification is 0. As in §4, for example, there are a number of intrinsically ergodic non-Markov shifts for which the obstruction entropy to specification is 0. Intrinsically ergodic subshifts with positive obstruction entropy to specification have received little attention and should be investigated further. Under the hypothesis of Theorem A, measures whose entropy does not exceed the obstruction entropy to specification play no role in search of measures of maximal entropy, but they do play roles in problems concerning rare events not captured by measures of maximal entropy, such as large deviations and multifractal analysis.

1.2. *Structure of the paper.* For proofs of parts (a) and (b) in Theorem A, we develop ideas of Morris [34, Theorem 1.1, Corollary 2] and Shinoda [37, Theorem A], respectively, both related to approximations of ergodic measures. On the one hand, a key observation is that the assumption of the Bowen specification property in [34, Corollary 2] can actually be weakened to the property that any ergodic measure can be approximated by another with arbitrarily small entropy. On the other hand, a key ingredient in the proof of [37, Theorem A] is the construction of paths of fully supported ergodic measures of positive entropy in arbitrarily small neighborhoods of maximizing measures. This construction relies on Sigmund's result [42].

The rest of this paper consists of four sections. The most critical component is Proposition 2.4 that provides two different kinds of approximations of ergodic measures in

https://doi.org/10.1017/etds.2025.10206 Published online by Cambridge University Press

the weak* topology. The proof of Proposition 2.4 relies on the thermodynamic formalism for subshifts with non-uniform specification [**15**] and the zero temperature limit [**1**] summarized in §2. In §3, we combine results in §2 with part of the proof of [**34**, Theorem 1.1] and a duality argument in [**26**] to complete the proofs of Theorems A and B. In §4, we provide examples of application of these two theorems: any transitive piecewise monotonic interval map; some coded shifts including S-gap shifts; some multidimensional β-transformations. In §5, we prove Theorem C by 'fattening' S-gap shifts.

2. *Preliminaries*

This section contains preliminary results needed for the proof of Theorem A. In §2.1, we begin by introducing basic terminologies in symbolic dynamics. In §2.2, we recall the result of Climenhaga and Thompson [**15**] on the thermodynamic formalism for subshifts with non-uniform specification. In §2.3, we introduce the obstruction entropy to specification and in §2.4, recall the notion of ground states in zero temperature limit. Section 2.5 states and proves Proposition 2.4 that is a key ingredient in the proofs of Theorems A and B.

2.1. *Subshift, word, language.*

Let $\mathcal{A}$ be a non-empty finite discrete set, called an *alphabet*. A finite string $w = w_1 w_2 \cdots w_n$ of elements of $\mathcal{A}$ is called a *word* of length n from $\mathcal{A}$. For convenience, we introduce an empty word $\emptyset$ by the rules $\emptyset w = w\emptyset = w$ for any word w from $\mathcal{A}$ and $\emptyset\emptyset = \emptyset$. The word length of w from $\mathcal{A}$ is denoted by $|w|$ and the word length of the empty word is set to be 0. Let $\mathbb{N}$ denote the set of positive integers and set $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.

Given a collection $\mathcal{L}$ of words from $\mathcal{A}$ and the empty word, we consider its decomposition: collections of words $\mathcal{C}^p$, $\mathcal{G}$, $\mathcal{C}^s \subset \mathcal{L}$ such that $\mathcal{C}^p\mathcal{G}\mathcal{C}^s = \mathcal{L}$. This means that every word in $\mathcal{L}$ can be written as a concatenation of a prefix (from $\mathcal{C}^p$), a 'good' core (from $\mathcal{G}$), and a suffix (from $\mathcal{C}^s$). Given such a decomposition, we consider for every $M \in \mathbb{N}$, the following 'fattened' set of good words:

$$\mathcal{G}^M = \{uvw \in \mathcal{L} \colon u \in \mathcal{C}^p, v \in \mathcal{G}, w \in \mathcal{C}^s,\ |u|, |w| \le M\}.$$

Note that $\mathcal{L} = \bigcup_{M\in\mathbb{N}} \mathcal{G}^M$. This gives a filtration of the language.

Let $\mathcal{A}^{\mathbb{N}}$ and $\mathcal{A}^{\mathbb{Z}}$ denote the one-sided and the two-sided Cartesian product topological spaces of $\mathcal{A}$, respectively. The topologies of these two spaces are metrizable with the Hamming metric d: for distinct points $x = (x_i)_i$, $y = (y_i)_i$, $d(x, y) = 2^{-\min\{|i| \colon x_i \neq y_i\}}$. The left shift acts continuously on these spaces. A shift-invariant closed subset of $\mathcal{A}^{\mathbb{N}}$ or $\mathcal{A}^{\mathbb{Z}}$ is called a *subshift* over the alphabet $\mathcal{A}$, or a subshift on $\#\mathcal{A}$ symbols. A *language* of a subshift Σ, denoted by $\mathcal{L}(\Sigma)$, is the set of the empty word $\emptyset$ and words from $\mathcal{A}$ that appear in some elements of Σ. For a subshift Σ and $n \in \mathbb{N}_0$, let $\mathcal{L}_n(\Sigma)$ denote the set of elements of $\mathcal{L}(\Sigma)$ with word length n. For a subset $\mathcal{D}$ of $\mathcal{L}(\Sigma)$, put

$$\mathcal{D}_n = \mathcal{D} \cap \mathcal{L}_n(\Sigma).$$

Note that $\mathcal{L}_0(\Sigma) = \mathcal{D}_0 = \{\emptyset\}$, the singleton consisting of the empty word.

https://doi.org/10.1017/etds.2025.10206 Published online by Cambridge University Press

2.2. *Thermodynamic formalism with non-uniform specification.* Let Σ be a subshift and let σ denote the left shift acting on Σ. For $n \in \mathbb{N}$ and $w = w_1 \cdots w_n \in \mathcal{L}_n(\Sigma)$, define an *n-cylinder*

$$[w] = \{(x_i)_i \in \Sigma : x_i = w_i \text{ for } i = 1, \ldots, n\}.$$

Let $\phi \in C(\Sigma)$ and let $\mathcal{D} \subset \mathcal{L}(\Sigma)$. For $n \in \mathbb{N}$, define

$$\Lambda_n(\mathcal{D}, \phi) = \sum_{w \in \mathcal{D}_n} \sup_{[w]} \exp S_n\phi,$$

where $S_n\phi = \sum_{i=0}^{n-1} \phi \circ \sigma^i$, and put

$$h(\mathcal{D}, \phi) = \limsup_{n\to\infty} \frac{\log \Lambda_n(\mathcal{D}, \phi)}{n}.$$

In the case where $\phi \equiv 0$, let us abbreviate $h(\mathcal{D}) = h(\mathcal{D}, \phi)$. Note that $h(\mathcal{L}(\Sigma))$ equals $h_{\text{top}}(\Sigma)$. The quantity $h(\mathcal{L}(\Sigma), \phi)$ is called the topological pressure and denoted by $P(\phi)$. The variational principle [43] asserts that

$$P(\phi) = \sup \left\{ h(\mu) + \int \phi \, d\mu : \mu \in \mathcal{M}(\Sigma, \sigma) \right\}.$$

Elements of $\mathcal{M}(\Sigma, \sigma)$ that attain this supremum are called *equilibrium states* for the potential ϕ. Since (Σ, σ) is expansive, the entropy function $\mu \in \mathcal{M}(\Sigma, \sigma) \mapsto h(\mu)$ is upper semicontinuous. So, (the minus of) the free energy $\mu \in \mathcal{M}(\Sigma, \sigma) \mapsto h(\mu) + \int \phi \, d\mu$ is upper semicontinuous and equilibrium states always exist. The uniqueness is an issue.

The Bowen specification property allows us to glue pieces of orbits together to form one orbit. This property can be used to show the uniqueness of equilibrium states. Since the Bowen specification property does not always hold for non-Markov shifts, a number of weaker conditions have been introduced and their consequences have been investigated (see [31] and the references therein). Climenhaga and Thompson [15] introduced the following condition and, under it, proved that the measure of maximal entropy is unique and satisfies a certain Gibbs property. We say a collection of words $\mathcal{G} \subset \mathcal{L}(\Sigma)$ has *(W)-specification* if there exists an integer $t \geq 0$, called a *gap size*, such that for all $m \geq 2$ and all $v^{(1)}, \ldots, v^{(m)} \in \mathcal{G}$, there exist $w^{(1)}, \ldots, w^{(m-1)} \in \mathcal{L}(\Sigma)$ such that $v^{(1)}w^{(1)}v^{(2)}w^{(2)} \cdots w^{(m-1)}v^{(m)} \in \mathcal{L}(\Sigma)$ and $|w^{(i)}| \leq t$ for $i = 1, \ldots, m-1$.

THEOREM 2.1. [15, Theorem C] *Let Σ be a subshift and let $\phi \in C(\Sigma)$ be Hölder continuous. Suppose there exist collections of words $\mathcal{C}^p$, $\mathcal{G}$, $\mathcal{C}^s \subset \mathcal{L}(\Sigma)$ such that $\mathcal{L}(\Sigma) = \mathcal{C}^p\mathcal{G}\mathcal{C}^s$, and the following two conditions hold:*

(a) *$\mathcal{G}^M$ has (W)-specification for all $M \in \mathbb{N}$;*

(b) *the collections $\mathcal{C}^p$, $\mathcal{C}^s$ satisfy*

$$\sum_{n=1}^{\infty} \Lambda_n(\mathcal{C}^p \cup \mathcal{C}^s, \phi) e^{-P(\phi)n} < \infty.$$

Then, the equilibrium state for the potential ϕ is unique and has the weak Gibbs property: there exists $K > 0$ and for each $M \in \mathbb{N}$, there exists $K_M > 0$ such that for all $n \in \mathbb{N}$ and

https://doi.org/10.1017/etds.2025.10206 Published online by Cambridge University Press

all $w \in \mathcal{G}^M \cap \mathcal{L}_n(\Sigma)$*, we have*

$$K_M \leq \frac{\mu_\phi([w])}{\exp(-P(\phi)n + \sup_{[w]} S_n\phi)} \leq K. \tag{2.1}$$

Remark 2.2. More general potential functions than those in Theorem 2.1 were treated in [**15**, Theorem C], but we will not need them in this paper.

2.3. *Obstruction entropy to specification.* For a subshift Σ, we define

$$h^{\perp}_{\mathrm{spec}}(\Sigma) = \inf \left\{ h(\mathcal{C}^p \cup \mathcal{C}^s) \colon \begin{array}{l} \mathcal{L}(\Sigma) = \mathcal{C}^p\mathcal{G}\mathcal{C}^s \text{ and} \\ \mathcal{G}^M \text{ has (W)-specification for all } M \in \mathbb{N} \end{array} \right\}, \tag{2.2}$$

and call this number *the obstruction entropy to specification* on Σ. Note that $h^{\perp}_{\mathrm{spec}}(\Sigma) \leq h_{\mathrm{top}}(\Sigma)$. By Theorem 2.1 with $\phi \equiv 0$, the strict inequality $h^{\perp}_{\mathrm{spec}}(\Sigma) < h_{\mathrm{top}}(\Sigma)$ implies the intrinsic ergodicity of Σ.

As is evident from equation (2.2), the lower estimate of $h^{\perp}_{\mathrm{spec}}(\Sigma)$ is much harder than the upper one: to show $h^{\perp}_{\mathrm{spec}}(\Sigma) = 0$, we only have to exhibit for any $\varepsilon > 0$, one decomposition for which $h(\mathcal{C}^p \cup \mathcal{C}^s) < \varepsilon$, whereas to show $h^{\perp}_{\mathrm{spec}}(\Sigma) \geq c$ for some $c > 0$, we must verify $h(\mathcal{C}^p \cup \mathcal{C}^s) \geq c$ for *any* decomposition $\mathcal{L}(\Sigma) = \mathcal{C}^p\mathcal{G}\mathcal{C}^s$ for which $\mathcal{G}^M$ has (W)-specification for all $M \in \mathbb{N}$.

2.4. *Entropy in zero temperature limit and continuity of equilibrium states.* Let Σ be a subshift and let $f \in C(\Sigma)$. We say $\mu \in \mathcal{M}(\Sigma, \sigma)$ is a *ground state* for the potential f if there exists a net $(\mu_\beta)_{\beta \geq 0}$ in $\mathcal{M}(\Sigma, \sigma)$ such that:

- μ_β is an equilibrium state for the potential βf for all $\beta \geq 0$;
- $\mu_\beta \to \mu$ in the weak* topology as $\beta \to \infty$.

The term 'ground state' has been borrowed from statistical mechanics as an analogy: since the parameter β may be viewed as the inverse of temperature, taking the limit $\beta \to \infty$ is called the zero temperature limit [**1**].

For the reader's convenience, we include a proof of the next lemma on the zero temperature limit and the continuity of equilibrium states, which can be found in [**1**], for example.

LEMMA 2.3. *Let* Σ *be a subshift and let* $\mu \in \mathcal{M}(\Sigma, \sigma)$ *be a ground state for the potential* $f \in C(\Sigma)$*. Let* $(\mu_\beta)_{\beta \geq 0}$ *be a net in* $\mathcal{M}(\Sigma, \sigma)$ *such that* μ_β *is an equilibrium state for the potential* βf *for all* $\beta \geq 0$*, and* $\mu_\beta \to \mu$ *in the weak* topology as* $\beta \to \infty$*. Then,* μ *is f-maximizing, and satisfies*

$$h(\mu) = \sup\{h(\nu) \colon \nu \in \mathcal{M}_{\max}(f)\} = \max\{h(\nu) \colon \nu \in \mathcal{M}_{\max}(f)\}$$

and

$$\lim_{\beta \to \infty} h(\mu_\beta) = h(\mu).$$

Moreover, if μ_β *is the unique equilibrium state for the potential* βf *for all* $\beta \geq 0$*, then* $\beta \mapsto \mu_\beta$ *is continuous.*

https://doi.org/10.1017/etds.2025.10206 Published online by Cambridge University Press

Proof. For all $\beta \geq 0$, we have $h(\mu_\beta) + \beta \int f \, d\mu_\beta \geq h(\mu) + \beta \int f \, d\mu$. Dividing both sides by $\beta > 0$ and letting $\beta \to \infty$ yields $\int f \, d\mu_\beta \to \int f \, d\mu = \Lambda_\sigma(f)$. Hence, μ is f-maximizing. Since $\mathcal{M}_{\max}(f)$ is closed, the supremum in the statement of the lemma is attained.

The upper semicontinuity of the entropy function gives $\limsup_{\beta\to\infty} h(\mu_\beta) \leq h(\mu)$. If there were $\nu \in \mathcal{M}_{\max}(f)$ such that $h(\mu) < h(\nu)$, then for all sufficiently large $\beta > 0$, we would have

$$h(\nu) + \beta \int f \, d\nu = h(\nu) + \beta \int f \, d\mu > h(\mu_\beta) + \beta \int f \, d\mu_\beta,$$

in contradiction to the assumption that μ_β is an equilibrium state for the potential βf. Finally, we claim $\liminf_{\beta\to\infty} h(\mu_\beta) \geq h(\mu)$, for otherwise, we would reach a contradiction by applying the same argument to a subsequence of $(\mu_\beta)_{\beta\geq 0}$.

Let $\beta_0 \geq 0$. Let $(\mu_{\beta_n})_{n\in\mathbb{N}}$ be any convergent subsequence of $(\mu_\beta)_{\beta\geq 0}$ such that $\beta_n \to \beta_0$ as $n \to \infty$. Let μ denote the limit measure. Since the pressure continuously depends on the potential [43] and the entropy function is upper semicontinuous, we have

$$P(\beta_0 f) = \lim_{n\to\infty} P(\beta_n f) = \lim_{n\to\infty} \left(h(\mu_{\beta_n}) + \beta_n \int f \, d\mu_{\beta_n} \right) \leq h(\mu) + \beta \int f \, d\mu$$

and, therefore, $\mu = \mu_{\beta_0}$, which completes the proof of the continuity. □

2.5. *Two approximations of ergodic measures.* The next proposition is a key ingredient in the proof of Theorem A.

PROPOSITION 2.4. *Let Σ be a subshift that has a language decomposition $\mathcal{L}(\Sigma) = \mathcal{C}^p\mathcal{G}\mathcal{C}^s$ such that $h(\mathcal{C}^p \cup \mathcal{C}^s) < h_{\text{top}}(\Sigma)$ and $\mathcal{G}^M$ has (W)-specification for all $M \in \mathbb{N}$. For each $\mu \in \mathcal{M}(\Sigma, \sigma)$ that is ergodic and satisfies $h(\mu) > h(\mathcal{C}^p \cup \mathcal{C}^s)$, the following statements hold.*

(a) *For any $\delta > 0$ and any open subset U of $\mathcal{M}(\Sigma, \sigma)$ that contains μ, there exists an ergodic measure $\nu \in U$ such that $h(\nu) < \delta$.*

(b) *For any open subset V of $\mathcal{M}(\Sigma, \sigma)$ that contains μ, there exists a Lipschitz continuous function $f : \Sigma \to \mathbb{R}$ such that:*
   (i) *for any $\beta \geq 0$, there exists a unique equilibrium state for the potential βf, which has the weak Gibbs property, denoted by $\mu_{\beta f}$;*
   (ii) *$\beta \in [0, \infty) \mapsto \mu_{\beta f} \in \mathcal{M}(\Sigma, \sigma)$ is continuous and injective;*
   (iii) *for all sufficiently large $\beta > 0$, we have $\mu_{\beta f} \in V$ and $h(\mu_{\beta f})$ converges as $\beta \to \infty$ to a constant in $(h(\mathcal{C}^p \cup \mathcal{C}^s), h_{\text{top}}(\Sigma)]$.*

Proposition 2.4 provides two different kinds of approximations of ergodic measures with entropy greater than $h(\mathcal{C}^p \cup \mathcal{C}^s)$. In statement (a), the measure μ is approximated in the weak* topology by measures with arbitrarily small entropy. In statement (b), μ is approximated in the weak* topology by a ground state with positive entropy in the zero temperature limit.

https://doi.org/10.1017/etds.2025.10206 Published online by Cambridge University Press

COROLLARY 2.5. *Let Σ be a subshift that satisfies $h^{\perp}_{\mathrm{spec}}(\Sigma) < h_{\mathrm{top}}(\Sigma)$. If $h^{\perp}_{\mathrm{spec}}(\Sigma) \le H < h_{\mathrm{top}}(\Sigma)$, then for any $\mu \in \mathcal{M}(\Sigma, \sigma)$ that is ergodic and satisfies $h(\mu) > H$, the following statements hold.*

(a) *For any open subset U of $\mathcal{M}(\Sigma, \sigma)$ that contains μ and for any $\delta > 0$, there exists $\nu \in U$ that is ergodic and satisfies $h(\nu) < \delta$.*

(b) *For any open subset V of $\mathcal{M}(\Sigma, \sigma)$ that contains μ, there exists a homeomorphism $t \in [0, 1] \mapsto \nu_t \in V$ onto its image such that for any $t \in [0, 1]$, ν_t is ergodic, fully supported, and satisfies $h(\nu_t) > H$.*

Proof. Part (a) follows from Proposition 2.4(a). Since measures with the weak Gibbs property in equation (2.1) are fully supported, part (b) follows from Proposition 2.4(b). □

Corollary 2.5(a) will be used in the proof of Theorem A(a) to show the density of ergodic measures with arbitrary small entropy. Corollary 2.5(b) will be used in the proof of Theorem A(b) to deduce the abundance of paths of fully supported ergodic measures with positive entropy.

The rest of this subsection is entirely dedicated to a proof of Proposition 2.4. We begin by approximating ergodic measures with sufficiently large entropy with a collection of cylinders.

LEMMA 2.6. *Let Σ be a subshift, let $\mu \in \mathcal{M}(\Sigma, \sigma)$ be ergodic with positive entropy, and let $f_1, \ldots, f_m$ be a finite collection of elements of $C(\Sigma)$. For any $\varepsilon \in (0, h(\mu))$, there exists $N \in \mathbb{N}$ such that for each integer $k_0 \ge N$, there exists a subset $\mathcal{P}_{k_0} = \mathcal{P}_{k_0}(\mu, f_1, \ldots, f_m, \varepsilon)$ of $\mathcal{L}_{k_0}(\Sigma)$ such that for every $w \in \mathcal{P}_{k_0}$, we have*

$$e^{-k_0(h(\mu)+\varepsilon)} \le \mu([w]) \le e^{-k_0(h(\mu)-\varepsilon)} \tag{2.3}$$

and

$$\max_{x \in [w]} \max_{1 \le \ell \le m} \left| \frac{S_{k_0} f_\ell(x)}{k_0} - \int f_\ell \, d\mu \right| < \varepsilon. \tag{2.4}$$

Moreover, we have

$$\tfrac{1}{2} e^{(h(\mu)-\varepsilon)k_0} \le \#\mathcal{P}_{k_0} \le e^{(h(\mu)+\varepsilon)k_0}. \tag{2.5}$$

Furthermore, if $w_1 \cdots w_{k_0+1} \in \mathcal{P}_{k_0+1}$, then $w_1 \cdots w_{k_0} \in \mathcal{P}_{k_0}$.

Proof. Since the collection of 1-cylinders in Σ generates the Borel sigma-algebra of Σ, it follows from the Shannon–McMillan–Breiman theorem and the Birkhoff ergodic theorem that for μ-almost every $x \in \Sigma$, we have

$$\lim_{k_0 \to \infty} \frac{\log \mu([x_1 \cdots x_{k_0}])}{k_0} = -h(\mu)$$

and

$$\lim_{k_0 \to \infty} \frac{S_{k_0} f_\ell(x)}{k_0} = \int f_\ell \, d\mu \quad \text{for } 1 \le \ell \le m.$$

https://doi.org/10.1017/etds.2025.10206 Published online by Cambridge University Press

For each $x \in \Sigma$ for which the above two equalities hold, let $N(x) \in \mathbb{N}$ denote the minimal integer such that

$$\sup_{k_0 \geq N(x)} \left| \frac{\log \mu([x_1 \cdots x_{k_0}])}{k_0} + h(\mu) \right| < \varepsilon$$

and

$$\sup_{k_0 \geq N(x)} \max_{1 \leq \ell \leq m} \left| \frac{S_{k_0} f_\ell(x)}{k_0} - \int f_\ell \, d\mu \right| < \frac{\varepsilon}{2}.$$

There exist a Borel set $\Sigma_0 \subset \Sigma$ and $N \in \mathbb{N}$ such that $\mu(\Sigma_0) = 1/2$ and $N(x) = N$ for all $x \in \Sigma_0$. For an integer $k_0 \geq N$, we set

$$\mathcal{P}_{k_0} = \{w \in \mathcal{L}_{k_0}(\Sigma) \colon [w] \cap \Sigma_0 \neq \emptyset\}. \tag{2.6}$$

For each $w \in \mathcal{P}_{k_0}$, we have equations (2.3) and (2.4). From equation (2.3) we obtain equation (2.5). The last assertion of Lemma 2.6 is a direct consequence of the definition of $\mathcal{P}_{k_0}$ in equation (2.6). □

Given a subshift Σ that has a language decomposition $\mathcal{L}(\Sigma) = \mathcal{C}^p \mathcal{G} \mathcal{C}^s$ with $h(\mathcal{C}^p \cup \mathcal{C}^s) < h_{\text{top}}(\Sigma)$ and an ergodic measure $\mu \in \mathcal{M}(\Sigma, \sigma)$ with $h(\mu) > h(\mathcal{C}^p \cup \mathcal{C}^s)$, we would like to apply Lemma 2.6 and connect elements of $\mathcal{P}_{k_0}$ with a uniform gap size to construct a subshift, and implement the thermodynamic formalism in §2.2. However, elements of $\mathcal{P}_{k_0}$ cannot be connected since they may be out of the good core $\mathcal{G}$. To fix this point, we use the next lemma that extracts words in $\mathcal{G}$ of length approximately k_0 from many elements of $\mathcal{P}_{k_0}$.

LEMMA 2.7. *Let Σ be a subshift that has a language decomposition $\mathcal{L}(\Sigma) = \mathcal{C}^p \mathcal{G} \mathcal{C}^s$ with $h(\mathcal{C}^p \cup \mathcal{C}^s) < h_{\text{top}}(\Sigma)$. Let $\mu \in \mathcal{M}(\Sigma, \sigma)$ be ergodic with $h(\mu) > h(\mathcal{C}^p \cup \mathcal{C}^s)$. Let $f_1, \ldots, f_m$ be a finite collection of elements of $C(\Sigma)$. If $\varepsilon \in (0, h(\mu))$ is sufficiently small, then for any $M \geq 1$, there exist an integer $k \geq M$ and a subset $\mathcal{Q}_k = \mathcal{Q}_k(\mu, f_1, \ldots, f_m, \varepsilon)$ of $\mathcal{G}_k$ such that*

$$\frac{\log \#\mathcal{Q}_k}{k} > h(\mu) - \varepsilon \tag{2.7}$$

and

$$\max_{x \in [v]} \max_{1 \leq \ell \leq m} \left| \frac{S_k f_\ell(x)}{k} - \int f_\ell \, d\mu \right| < \varepsilon \quad \text{for all } v \in \mathcal{Q}_k. \tag{2.8}$$

Proof. Set $\underline{h}(\mu) = (h(\mu) - h(\mathcal{C}^p \cup \mathcal{C}^s))/2$. Let $M \geq 1$ and

$$\varepsilon \in (0, \min\{\underline{h}(\mu), \sqrt{\underline{h}(\mu)}/2, \sqrt[3]{\underline{h}(\mu)}, 1/3\}).$$

Let $N \in \mathbb{N}$ be given by Lemma 2.6. For each $k_0 \geq N$, let $\mathcal{P}_{k_0} = \mathcal{P}_{k_0}(\mu, f_1, \ldots, f_m, \varepsilon^3)$ be the subset of $\mathcal{L}_{k_0}(\Sigma)$ given by Lemma 2.6. We may assume N is sufficiently large so that

$$\#(\mathcal{C}_n^p \cup \mathcal{C}_n^s) \leq e^{(h(\mathcal{C}^p \cup \mathcal{C}^s) + \varepsilon^3)n} \quad \text{for } n \geq N. \tag{2.9}$$

https://doi.org/10.1017/etds.2025.10206 Published online by Cambridge University Press

For each $w \in \mathcal{P}_{k_0}$, we fix a decomposition

$$w = p(w)c(w)s(w), \quad p(w) \in \mathcal{C}^p, c(w) \in \mathcal{G}, s(w) \in \mathcal{C}^s.$$

These three words can be the empty word. For $a, b \in \mathbb{N}_0$ with $a + b \le k_0$, define

$$\mathcal{Q}^p_{k_0}(a) = \{w \in \mathcal{P}_{k_0} \colon |p(w)| = a\} \quad \text{and} \quad \mathcal{Q}^s_{k_0}(b) = \{w \in \mathcal{P}_{k_0} \colon |s(w)| = b\}.$$

We set

$$\mathcal{Q}_{k_0}(a, b) = \mathcal{Q}^p_{k_0}(a) \cap \mathcal{Q}^s_{k_0}(b).$$

The language decomposition $\mathcal{L}(\Sigma) = \mathcal{C}^p\mathcal{G}\mathcal{C}^s$ implies $\mathcal{P}_{k_0} = \bigsqcup_{a,b\in\mathbb{N}_0, a+b\le k_0} \mathcal{Q}_{k_0}(a, b)$.

We construct a subset of $\mathcal{P}_{k_0}$ that has a definite fraction in exponential scale and consists of words with relatively short prefixes in $\mathcal{C}^p$ and suffixes in $\mathcal{C}^s$. By equation (2.9), for $a = N, \dots, k_0$, we have

$$\#\{p(w) \colon w \in \mathcal{P}_{k_0},\ |p(w)| = a\} \le e^{h(\mathcal{C}^p\cup\mathcal{C}^s)+\varepsilon^3)a}.$$

By equation (2.3) with ε replaced by ε^3, for each $w \in \mathcal{P}_{k_0}$ with $|p(w)| = a$,

$$\begin{aligned}\#\{v \in \mathcal{L}_{k_0-a}(\Sigma) \colon p(w)v \in \mathcal{P}_{k_0}\} &\le \frac{\mu([p(w)])}{\min\{\mu([p(w)v]) \colon v \in \mathcal{L}_{k_0-a}(\Sigma) \colon p(w)v \in \mathcal{P}_{k_0}\}} \\ &\le \frac{e^{-a(h(\mu)-\varepsilon^3)}}{e^{-k_0(h(\mu)+\varepsilon^3)}} \le e^{(k_0-a)h(\mu)+2\varepsilon^3 k_0}.\end{aligned}$$

Multiplying these two estimates yield

$$\begin{aligned}\#\mathcal{Q}^p_{k_0}(a) &\le e^{(h(\mathcal{C}^p\cup\mathcal{C}^s)+\varepsilon^3)a} e^{(k_0-a)h(\mu)+2\varepsilon^3 k_0} \\ &\le e^{(h(\mu)+3\varepsilon^3)k_0} e^{-2\underline{h}(\mu)a} \quad \text{for } a = N, \dots, k_0.\end{aligned}$$

Since $\varepsilon < \min\{\sqrt{\underline{h}(\mu)}, 1/3\}$, for all sufficiently large k_0, we have

$$\sum_{a=\lfloor\varepsilon^2 k_0/\underline{h}(\mu)\rfloor}^{k_0} \#\mathcal{Q}^p_{k_0}(a) \le e^{(h(\mu)+3\varepsilon^3)k_0} \sum_{a=\lfloor\varepsilon^2 k_0/\underline{h}(\mu)\rfloor}^{k_0} e^{-2\underline{h}(\mu)a} \le e^{(h(\mu)-\varepsilon^2)k_0}. \tag{2.10}$$

Similarly, by the last assertion of Lemma 2.6 and equation (2.5) with ε replaced by ε^3, for $b = 0, \dots, k_0 - N$, we have

$$\#\{p(w)c(w) \colon w \in \mathcal{P}_{k_0},\ |s(w)| = b\} \le \#\mathcal{P}_{k_0-b} \le e^{(h(\mu)+\varepsilon^3)(k_0-b)}.$$

By equation (2.9), for each $b = N, \dots, k_0$ and each $w \in \mathcal{P}_{k_0}$ with $|s(w)| = b$,

$$\#\{v \in \mathcal{C}^s_b \colon p(w)c(w)v \in \mathcal{P}_{k_0}\} \le e^{(h(\mathcal{C}^p\cup\mathcal{C}^s)+\varepsilon^3)b}.$$

If $k_0 \ge 2N$, then multiplying these two estimates yield

$$\begin{aligned}\#\mathcal{Q}^s_{k_0}(b) &\le e^{(h(\mu)+\varepsilon^3)(k_0-b)} e^{(h(\mathcal{C}^p\cup\mathcal{C}^s)+\varepsilon^3)b} \\ &= e^{(h(\mu)+\varepsilon^3)k_0} e^{-2b\underline{h}(\mu)} \quad \text{for } b = N, \dots, k_0 - N.\end{aligned}$$

https://doi.org/10.1017/etds.2025.10206 Published online by Cambridge University Press

Since $\varepsilon < \min\{\sqrt{\underline{h}(\mu)}, 1/3\}$, for all sufficiently large $k_0 \geq 2N$, we have

$$\sum_{b=\lfloor \varepsilon^2 k_0/\underline{h}(\mu)\rfloor}^{k_0-N} \#\mathcal{Q}^s_{k_0}(b) \leq e^{(h(\mu)+\varepsilon^3)k_0} \sum_{b=\lfloor \varepsilon^2 k_0/\underline{h}(\mu)\rfloor}^{k_0} e^{-2b\underline{h}(\mu)} \leq e^{(h(\mu)-\varepsilon^2)k_0}. \tag{2.11}$$

Let $\mathcal{A}$ be the alphabet of Σ. By equation (2.9),

$$\begin{aligned}\sum_{b=k_0-N+1}^{k_0} \#\mathcal{Q}^s_{k_0}(b) &\leq \sum_{b=k_0-N+1}^{k_0} (\#\mathcal{A})^{k_0-b} e^{(h(\mathcal{C}^p\cup\mathcal{C}^s)+\varepsilon^3)b} \\ &\leq (\#\mathcal{A})^N \sum_{b=k_0-N+1}^{k_0} e^{h(\mathcal{C}^p\cup\mathcal{C}^s)b} \leq N(\#\mathcal{A})^N e^{h(\mathcal{C}^p\cup\mathcal{C}^s)k_0}. \end{aligned}\tag{2.12}$$

Combining equations (2.10), (2.11), (2.12) and the first inequality in equation (2.5) with ε replaced by ε^3, and then using $\varepsilon < \min\{\sqrt[3]{\underline{h}(\mu)}, 1\}$, for all sufficiently large $k_0 \geq 2N$, we have

$$\begin{aligned}\sum_{a,b=0}^{\lfloor \varepsilon^2 k_0/\underline{h}(\mu)\rfloor-1} \#\mathcal{Q}_{k_0}(a,b) &\geq \#\mathcal{P}_{k_0} - \sum_{a=\lfloor \varepsilon^2 k_0/\underline{h}(\mu)\rfloor}^{k_0} \#\mathcal{Q}^p_{k_0}(a) - \sum_{b=\lfloor \varepsilon^2 k_0/\underline{h}(\mu)\rfloor}^{k_0} \#\mathcal{Q}^s_{k_0}(b) \\ &\geq \frac{1}{2} e^{(h(\mu)-\varepsilon^3)k_0} - 2e^{(h(\mu)-\varepsilon^2)k_0} - N(\#\mathcal{A})^N e^{h(\mathcal{C}^p\cup\mathcal{C}^s)k_0} \\ &\geq e^{(h(\mu)-2\varepsilon^3)k_0}. \end{aligned}$$

Moreover, since $\varepsilon < \sqrt{\underline{h}(\mu)}/2$ and $k_0 \geq 2$, there exist a_0, $b_0 \in \mathbb{N}_0$ such that $a_0 + b_0 \leq 2\varepsilon^2 k_0/\underline{h}(\mu) \leq k_0/2 \leq k_0 - 1$ and

$$\#\mathcal{Q}_{k_0}(a_0, b_0) \geq \frac{e^{(h(\mu)-2\varepsilon^3)k_0}}{(\varepsilon^2 k_0/\underline{h}(\mu))^2}. \tag{2.13}$$

We now set

$$k = k_0 - a_0 - b_0 \quad \text{and} \quad \mathcal{Q}_k = \{c(w)\colon w \in \mathcal{Q}_{k_0}(a_0, b_0)\}.$$

For all sufficiently large $k_0 \geq 2N$, by $\varepsilon < \min\{\underline{h}(\mu), 1/2\}$, we have

$$M \leq \left(1 - \frac{\varepsilon}{\underline{h}(\mu)}\right)k_0 < \left(1 - \frac{2\varepsilon^2}{\underline{h}(\mu)}\right)k_0 \leq k \leq k_0, \tag{2.14}$$

and from equations (2.9) and (2.13),

$$\#\mathcal{Q}_k \geq \frac{\#\mathcal{Q}_{k_0}(a_0, b_0)}{\#\mathcal{C}^p_{a_0}\#\mathcal{C}^s_{b_0}} > e^{(h(\mu)-\varepsilon)k_0}.$$

Since $k \leq k_0$, rearranging this inequality yields equation (2.7).

Let $f_1, \ldots, f_m$ be a finite collection of elements of $C(\Sigma)$. For any $v \in \mathcal{Q}_k$, there exist $u \in \mathcal{C}^p_{a_0}$, $w \in \mathcal{C}^s_{b_0}$ such that $uvw \in \mathcal{Q}_{k_0}(a_0, b_0)$. Fix $y \in [uvw]$. For all $1 \leq \ell \leq m$, we have

$$\left|\frac{S_k f_\ell(\sigma^{a_0} y)}{k} - \frac{S_{k_0} f_\ell(y)}{k}\right| \leq \frac{a_0 + b_0}{k}\|f_\ell\|_{C^0} \leq \frac{2\varepsilon^2 k_0}{\underline{h}(\mu)k}\|f_\ell\|_{C^0},$$

https://doi.org/10.1017/etds.2025.10206 Published online by Cambridge University Press

and by equation (2.4) with ε replaced by ε^3,

$$\left|\frac{S_{k_0} f_\ell(y)}{k} - \int f_\ell \, d\mu\right| \le \frac{k_0\varepsilon^3}{k} + \frac{k_0 - k}{k}\|f_\ell\|_{C^0}.$$

Combining these two inequalities and using $k/k_0 \ge 1 - 2\varepsilon^2/\underline{h}(\mu)$ in equation (2.14), we have

$$\left|\frac{S_k f_\ell(\sigma^{a_0} y)}{k} - \int f_\ell \, d\mu\right| < \frac{k_0\varepsilon^3}{k} + \frac{k_0 - k}{k}\|f_\ell\|_{C^0} + \frac{2\varepsilon^2 k_0}{\underline{h}(\mu)k}\|f_\ell\|_{C^0} < \frac{\varepsilon}{2} \tag{2.15}$$

provided ε is sufficiently small depending on μ and $f_1, \ldots, f_m$. Since each f_ℓ is uniformly continuous, if k_0 is sufficiently large, then for all $x \in [v]$ and $1 \le \ell \le m$, we have

$$\left|\frac{S_k f_\ell(x)}{k} - \frac{S_k f_\ell(\sigma^{a_0} y)}{k}\right| < \frac{\varepsilon}{2}. \tag{2.16}$$

Combining equations (2.15) and (2.16), we obtain equation (2.8). □

We are in a position to complete the proof of Proposition 2.4.

Proof of Proposition 2.4(a). Let $\mu \in \mathcal{M}(\Sigma, \sigma)$ be ergodic with $h(\mu) > h(\mathcal{C}^p \cup \mathcal{C}^s)$. Let U be an open subset of $\mathcal{M}(\Sigma, \sigma)$ that contains μ. Let $f_1, \ldots, f_m$ be a finite collection of elements of $C(\Sigma)$ and let $\varepsilon \in (0, h(\mu))$ be such that

$$\left\{\nu \in \mathcal{M}(\Sigma, \sigma) \colon \max_{1\le\ell\le m}\left|\int f_\ell \, d\nu - \int f_\ell \, d\mu\right| < 3\varepsilon\right\} \subset U. \tag{2.17}$$

By Lemma 2.7, if ε is sufficiently small, then there exists an arbitrarily large integer k and a subset $\mathcal{Q}_k = \mathcal{Q}_k(\mu, f_1, \ldots, f_m, \varepsilon)$ of $\mathcal{G}_k$ for which equations (2.7) and (2.8) hold. Pick $v \in \mathcal{Q}_k$. If Σ is one-sided, define

$$\Omega = \{vw^{(1)}vw^{(2)}vw^{(3)} \cdots \in \Sigma \colon |w^{(j)}| \le t \ \text{ for } j \in \mathbb{N}\}.$$

If Σ is two-sided, define

$$\Omega = \{x = \cdots vw^{(-2)}vw^{(-1)}vw^{(0)}vw^{(1)}v \cdots \in \Sigma \colon |w^{(j)}| \le t \ \text{ for } j \in \mathbb{Z}, [x_0] \supset [v]\}.$$

In other words, Ω is the set of points in Σ whose symbol sequences are alternate concatenations of v and elements of $\mathcal{L}(\Sigma)$ with word lengths not exceeding t. Since the gap size t for the specification of $\mathcal{G}$ is independent of k, the topological entropy of the subshift $\tilde{\Omega} = \overline{\bigcup_{n=0}^{\infty} \sigma^n \Omega}$ converges to 0 as $k \to \infty$.

Let $x \in \Omega$. We have $x = v^{(1)}w^{(1)}v^{(2)}w^{(2)} \cdots$ or $x = \cdots w^{(-1)}v^{(0)}w^{(0)}v^{(1)} \cdots$, and $[x_0] \supset [v_0]$ according to whether Σ is one-sided or two-sided, where $v^{(j)} \in \mathcal{Q}_k$ and $|w^{(j)}| \le t$. From equation (2.8), for all j, we have

$$\begin{aligned}\max_{x\in[v^{(j)}w^{(j)}]} \max_{1\le\ell\le m}\left|S_{|v^{(j)}w^{(j)}|} f_\ell(x) - |v^{(j)}w^{(j)}|\int f_\ell \, d\mu\right| &\le |v^{(j)}|\varepsilon + 2|w^{(j)}| \cdot \|f_\ell\|_{C^0} \\ &\le 2|v^{(j)}w^{(j)}|\varepsilon.\end{aligned}$$

https://doi.org/10.1017/etds.2025.10206 Published online by Cambridge University Press

The last inequality holds for all sufficiently large k. It follows that there exists $N \geq 1$ that is independent of $x \in \Gamma$ such that for all $n \geq N$, we have

$$\max_{1\leq \ell \leq m} \left| \frac{S_n f_\ell(x)}{n} - \int f_\ell \, d\mu \right| < 3\varepsilon.$$

We obtain the same estimate for all $x \in \tilde{\Omega}$ and all sufficiently large $n \geq N$. Consequently, for any measure $\nu \in \mathcal{M}(\tilde{\Omega}, \sigma|_{\tilde{\Omega}})$, we have

$$\max_{1\leq \ell \leq m} \left| \int f_\ell \, d\nu - \int f_\ell \, d\mu \right| < 3\varepsilon.$$

By the Krein–Milman theorem, the same inequality remains valid for all $\nu \in \mathcal{M}(\tilde{\Omega}, \sigma|_{\tilde{\Omega}})$. Together with equation (2.17), we obtain $\mathcal{M}(\tilde{\Omega}, \sigma|_{\tilde{\Omega}}) \subset U$. It follows that U contains a shift-invariant ergodic measure with arbitrarily small entropy. The proof of Proposition 2.4(a) is complete. □

Proof of Proposition 2.4(b). Let $\mu \in \mathcal{M}(\Sigma, \sigma)$ be ergodic with $h(\mu) > h(\mathcal{C}^p \cup \mathcal{C}^s)$. Let V be an open subset of $\mathcal{M}(\Sigma, \sigma)$ that contains μ. Let $f_1, \ldots, f_m$ be a finite collection of elements of $C(\Sigma)$ and let $\varepsilon \in (0, h(\mu)/3)$ be such that

$$\left\{ \nu \in \mathcal{M}(\Sigma, \sigma) \colon \max_{1\leq \ell \leq m} \left| \int f_\ell \, d\nu - \int f_\ell \, d\mu \right| < 3\varepsilon \right\} \subset V. \tag{2.18}$$

By Lemma 2.7, if ε is sufficiently small, then there exists an arbitrarily large integer k and a subset $\mathcal{Q}_k = \mathcal{Q}_k(\mu, f_1, \ldots, f_m, \varepsilon)$ of $\mathcal{G}_k$ for which equations (2.7) and (2.8) hold. We take a subset $\mathcal{Q}'_k$ of $\mathcal{Q}_k$ such that $h(\mu) - 3\varepsilon < (1/k) \log \#\mathcal{Q}'_k < h_{\mathrm{top}}(\Sigma) - \varepsilon$. Let $t \geq 0$ be a gap size for the (W)-specification of $\mathcal{G}$. If Σ is one-sided, then define

$$\Gamma = \{v^{(1)} w^{(1)} v^{(2)} w^{(2)} v^{(3)} w^{(3)} \cdots \in \Sigma \colon v^{(j)} \in \mathcal{Q}'_k, |w^{(j)}| \leq t \ \text{ for } j \in \mathbb{N}\}.$$

If Σ is two-sided, then define

$$\Gamma = \{x = \cdots w^{(-1)} v^{(0)} w^{(0)} v^{(1)} \cdots \in \Sigma \colon v^{(j)} \in \mathcal{Q}'_k, |w^{(j)}| \leq t \ \text{ for } j \in \mathbb{Z}, x \in [v_0]\}.$$

In other words, Γ is the set of points in Σ whose symbol sequences are alternate concatenations of elements of $\mathcal{Q}'_k$ and that of $\mathcal{L}(\Sigma)$ with word lengths not exceeding t. Since the topological entropy of the subshift $\tilde{\Gamma} = \overline{\bigcup_{n=0}^{\infty} \sigma^n \Gamma}$ is strictly smaller than $h_{\mathrm{top}}(\Sigma)$, $\tilde{\Gamma}$ is a proper subset of Σ. Define $f \colon \Sigma \to \mathbb{R}$ by

$$f(x) = -\min\{d(x, y) \colon y \in \tilde{\Gamma}\},$$

where d denotes the Hamming metric on Σ (see §2.1).

We claim that f is Lipschitz continuous. Indeed, for any pair x, y of distinct points in Σ, take $i_0, i_1, i_2 \in \mathbb{N}$ with $d(x, y) = 2^{-i_0}$, $f(x) = -2^{-i_1}$, $f(y) = -2^{-i_2}$. If $i_1 \leq i_0$ or $i_2 \leq i_0$, then $i_1 = i_2$ and so $f(x) - f(y) = 0$. Otherwise, $|f(x) - f(y)| = |2^{-i_1} - 2^{-i_2}| < 2^{-i_0+1} = 2d(x, y)$. The Lipschitz continuity of f implies

$$\sup_{n\in\mathbb{N}} \max\{S_n f(x) - S_n f(y) \colon x_i = y_i \ \text{ for } i = 1, \ldots, n\} < \infty. \tag{2.19}$$

https://doi.org/10.1017/etds.2025.10206 Published online by Cambridge University Press

By virtue of equation (2.18), the same argument as in the proof of Proposition 2.4(a) shows $\mathcal{M}(\tilde{\Gamma}, \sigma|_{\tilde{\Gamma}}) \subset V$. Clearly, $\Lambda_\sigma(f) = 0$ and, hence, any f-maximizing measure belongs to $\mathcal{M}(\tilde{\Gamma}, \sigma|_{\tilde{\Gamma}})$. If ε is sufficiently small and k is sufficiently large, then

$$h_{\text{top}}(\tilde{\Gamma}) \geq h(\mu) - 4\varepsilon > h(\mathcal{C}^p \cup \mathcal{C}^s). \tag{2.20}$$

Let $\nu_0 \in \mathcal{M}(\tilde{\Gamma}, \sigma|_{\tilde{\Gamma}})$ be an ergodic measure of maximal entropy for the subshift $\tilde{\Gamma}$. Note that

$$\limsup_{n\to\infty} \sup_{x\in\Sigma} \frac{S_n f(x)}{n} = \sup_{x\in\Sigma} \limsup_{n\to\infty} \frac{S_n f(x)}{n} = \Lambda_\sigma(f) = 0.$$

Combining this with equation (2.20), for all $\beta \geq 0$, we get

$$\begin{aligned} h(\mathcal{C}^p \cup \mathcal{C}^s) + \beta \cdot \limsup_{n\to\infty} \sup_{x\in\Sigma} \frac{S_n f(x)}{n} &= h(\mathcal{C}^p \cup \mathcal{C}^s) < h_{\text{top}}(\tilde{\Gamma}) \\ &= h(\nu_0) = h(\nu_0) + \int \beta f \, d\nu_0 \leq P(\beta f). \end{aligned}$$

This estimate implies

$$\sum_{n=1}^{\infty} \Lambda_n(\mathcal{C}^p \cup \mathcal{C}^s, \beta f) e^{-P(\beta f)n} < \infty.$$

By Theorem 2.1, the equilibrium state for the potential βf is unique, denoted by $\mu_{\beta f}$, and has the weak Gibbs property in equation (2.1). This verifies condition (i).

The continuity of the map $\beta \in [0, \infty) \mapsto \mu_{\beta f} \in \mathcal{M}(\Sigma, \sigma)$ follows from Lemma 2.3. To prove the injectivity of the map, let $0 \leq \beta_1 < \beta_2$ and suppose $\mu_{\beta_1 f} = \mu_{\beta_2 f}$ by contradiction. Since f is non-positive and non-constant, we have $P(\beta_1 f) > P(\beta_2 f)$. By the weak Gibbs property of $\mu_{\beta_1 f}$, $\mu_{\beta_2 f}$, and equation (2.19), there exists a constant $K \geq 1$ such that for all $n \geq 1$, $w \in \mathcal{G}_n$, and $x \in [w]$, we have

$$K^{-1} \leq \frac{\exp(-P(\beta_1 f)n + \beta_1 S_n f(x))}{\exp(-P(\beta_2 f)n + \beta_2 S_n f(x))} \leq K.$$

This is equivalent to

$$|(\beta_1 - \beta_2) S_n f(x) - n(P(\beta_1 f) - P(\beta_2 f))| \leq C \tag{2.21}$$

for some constant $C > 0$ independent of n, w, x. By the ergodicity of ν_0, equation (2.20), and $\int f \, d\nu_0 = 0$, one can use the set $\mathcal{Q}_k(\nu_0, f, \delta)$ given by Lemma 2.7 for a sufficiently small $\delta > 0$ and a sufficiently large $k \in \mathbb{N}$ to show that for all sufficiently large $n \in \mathbb{N}$, there exists $w \in \mathcal{G}_n$ such that $\sup_{x\in[w]} |(1/n) S_n f(x)| < \delta$. This and equation (2.21) together imply $P(\beta_1 f) = P(\beta_2 f)$, which is a contradiction, and thus verifies condition (ii).

By Lemma 2.3, any ground state for the potential f belongs to $\mathcal{M}_{\max}(f)$ and so belongs to $\mathcal{M}(\tilde{\Gamma}, \sigma|_{\tilde{\Gamma}})$. By Lemma 2.3 and equation (2.20), we obtain $\lim_{\beta\to\infty} h(\mu_{\beta f}) = h(\nu_0) = h_{\text{top}}(\tilde{\Gamma}) > h(\mathcal{C}^p \cup \mathcal{C}^s)$, which verifies condition (iii). The proof of Proposition 2.4(b) is complete. □

https://doi.org/10.1017/etds.2025.10206 Published online by Cambridge University Press

3. *On the proofs of Theorems A and B*

We are almost ready to proceed to proving the main results on ergodic optimization. After recalling some functional analytic ingredients in §3.1, we complete the proof of Theorem A in §3.2. In §3.3, we complete the proof of Theorem B.

3.1. *Functional analysis.*

The proof of Theorem A(b) relies on the result of Israel [**26**, §V], who proved an approximation theorem about tangent functionals to convex functions and used it for lattice systems in statistical mechanics to prove the existence of a dense set of continuous interactions for which there exist uncountably many ergodic equilibrium states. Below, we recall his result and some other auxiliary ones.

For a Banach space V with a norm $\|\cdot\|$, let V^* denote the set of real-valued bounded functionals on V. For each $\mu \in V^*$, let $\|\mu\|$ denote the norm

$$\|\mu\| = \sup\{|\mu(f)| \colon f \in V,\ \|f\| = 1\}.$$

Let $\Lambda, \mu \in V^*$. We say:

- μ is *tangent* to Λ at $f \in V$ if $\mu(f) \le \Lambda(f+g) - \Lambda(f)$ holds for all $g \in V$;
- μ is *bounded* by Λ if $\mu(f) \le \Lambda(f)$ holds for all $f \in V$;
- Λ is *convex* if $\Lambda(tf + (1-t)g) \le t\Lambda(f) + (1-t)\Lambda(g)$ holds for all $f, g \in V$, and $t \in [0,1]$.

The following is a version of the theorem of Bishop and Phelps [**2**].

THEOREM 3.1. [**26**, Theorem V.1.1] *Let V be a Banach space, and let $\Lambda \in V^*$ be convex and continuous. For any $\mu_0 \in V^*$ that is bounded by Λ, any $f_0 \in V$, and any $\varepsilon > 0$, there exist $\mu \in V^*$ and $f \in V$ such that μ is tangent to Λ at f, and*

$$\|\mu - \mu_0\| \le \varepsilon \quad \text{and} \quad \|f - f_0\| \le \frac{1}{\varepsilon}(\Lambda(f_0) - \mu_0(f_0) + s),$$

where $s = \sup\{\mu_0(g) - \Lambda(g) \colon g \in V\} \le 0$.

For a continuous map T of a compact metric space X, the functional $\Lambda_T \in C(X)^*$ given by equation (1.1) is convex and continuous. The next lemma characterizes maximizing measures in terms of Λ_T. Recall that the set of real-valued bounded linear functionals on $C(X)$ can be identified with the set of (finite) signed Borel measures on X by the Riesz representation theorem.

LEMMA 3.2. [**8**, Lemma 2.3] *Let T be a continuous map of a compact metric space X and let $f \in C(X)$. A bounded linear functional $\mu \in C(X)^*$ is tangent to Λ_T at f if and only if μ belongs to $\mathcal{M}(X,T)$ and is f-maximizing.*

If T is a continuous map of X, then for any $\mu \in \mathcal{M}(X,T)$, there exists a unique Borel probability measure b_μ on $\mathcal{M}(X,T)$ such that $b_\mu(\mathcal{M}^{\mathrm{e}}(X,T)) = 1$ and $\mu = \int_{\mathcal{M}^{\mathrm{e}}(X,T)} \nu \, db_\mu(\nu)$, where $\mathcal{M}^{\mathrm{e}}(X,T)$ denotes the set of elements of $\mathcal{M}(X,T)$ that are ergodic. We call b_μ the *barycentre* of μ. Put

$$\operatorname{supp}(b_\mu) = \bigcap \{F \colon F \subset \mathcal{M}(X,T), \text{ closed}, b_\mu(F) = 1\}.$$

Since $\mathcal{M}(X,T)$ has a countable base, we have $b_\mu(\operatorname{supp}(b_\mu)) = 1$.

https://doi.org/10.1017/etds.2025.10206 Published online by Cambridge University Press

LEMMA 3.3. *Let T be a continuous map of a compact metric space X.*

(a) *If there exists a constant $C \geq 0$ such that $h(\nu) \leq C$ for all $\nu \in \mathrm{supp}(b_\mu)$, then $h(\mu) \leq C$.*

(b) *Let $f \in C(X)$, $\mu \in \mathcal{M}_{\max}(f)$. Then, $\mathrm{supp}(b_\mu)$ is contained in $\mathcal{M}_{\max}(f)$.*

Proof. Part (a) is a consequence of the Jacobs theorem [43]. For a proof of part (b), see [37, Proposition 6], for example. □

The next lemma asserts that the barycentre map $\mu \mapsto b_\mu$ from $\mathcal{M}(X, T)$ to the set of Borel probability measures on $\mathcal{M}(X, T)$ is isometric.

LEMMA 3.4. [26, Corollary IV.4.2] *Let T be a continuous map of a compact metric space X. For all μ_1, $\mu_2 \in \mathcal{M}(X, T)$, we have*

$$\|b_{\mu_1} - b_{\mu_2}\| = \|\mu_1 - \mu_2\|.$$

3.2. *Proof of Theorem A.* Let Σ be a subshift with $h^{\perp}_{\mathrm{spec}}(\Sigma) < h_{\mathrm{top}}(\Sigma)$ and suppose $h^{\perp}_{\mathrm{spec}}(\Sigma) \leq H < h_{\mathrm{top}}(\Sigma)$. For each $n \in \mathbb{N}$, define

$$O_n = \left\{ \mu \in \overline{\mathcal{M}^{\mathrm{e}}(\Sigma, \sigma)} : 0 \leq h(\mu) < H + \frac{1}{n} \right\}$$

and

$$U_n = \{ f \in C(\Sigma) : \overline{\mathcal{M}^{\mathrm{e}}(\Sigma, \sigma)} \cap \mathcal{M}_{\max}(f) \subset O_n \}.$$

Clearly, the set $\mathscr{R}_H$ in Theorem A is contained in $\bigcap_{n=1}^{\infty} U_n$. Conversely, let $f \in \bigcap_{n=1}^{\infty} U_n$. The entropy of any ergodic measure in $\mathcal{M}_{\max}(f)$ does not exceed H and by Lemma 3.3, the entropy of any non-ergodic measure in $\mathcal{M}_{\max}(f)$ does not exceed H either. Hence, we obtain $\mathscr{R}_H = \bigcap_{n=1}^{\infty} U_n$.

By the upper semicontinuity of the entropy function, O_n is an open subset of $\overline{\mathcal{M}^{\mathrm{e}}(\Sigma, \sigma)}$. From Corollary 2.5(a), O_n is a dense subset of $\overline{\mathcal{M}^{\mathrm{e}}(\Sigma, \sigma)}$. By [34, Theorem 1.1], U_n is an open and dense subset of $C(\Sigma)$. Therefore, $\mathscr{R}_H$ is dense G_δ as required in Theorem A(a).

To prove Theorem A(b), pick $f_0 \in C(\Sigma) \setminus \mathscr{R}_H$. By Lemma 3.3(a), there exists an ergodic measure $\lambda \in \mathcal{M}_{\max}(f_0)$ with $h(\lambda) > H$. Let $\varepsilon \in (0, 1/2)$ and put

$$R = \left\{ \nu \in \mathcal{M}(\Sigma, \sigma) : \int f_0 \, d\nu \geq \Lambda_\sigma(f_0) - \varepsilon^2 \right\}.$$

By Corollary 2.5(b) applied to λ, there exists a homeomorphism $t \in [0, 1] \mapsto \nu_t \in R \cap \mathcal{M}^{\mathrm{e}}(\Sigma, \sigma)$ onto its image such that ν_t is fully supported and satisfies $h(\nu_t) > H$ for all $t \in [0, 1]$. Let m denote the Lebesgue measure on $[0, 1]$ and define a Borel probability measure $\hat{m}$ on $\mathcal{M}^{\mathrm{e}}(\Sigma, \sigma)$ by

$$\hat{m}(\cdot) = m\{t \in [0, 1] : \nu_t \in \cdot\}.$$

We have $\hat{m}(R \cap \mathcal{M}^{\mathrm{e}}(\Sigma, \sigma)) = 1$. Put $\mu_0 = \int_{\mathcal{M}^{\mathrm{e}}(\Sigma,\sigma)} \nu \, d\hat{m}(\nu)$. Then, μ_0 belongs to R and is bounded by Λ_σ as an element of $C(\Sigma)^*$. Note that $b_{\mu_0} = \hat{m}$. By Theorem 3.1, there exist

https://doi.org/10.1017/etds.2025.10206 Published online by Cambridge University Press

$f \in C(\Sigma)$ and $\mu \in C(\Sigma)^*$ such that μ is tangent to Λ_σ at f, and

$$\|\mu - \mu_0\| \le \varepsilon \quad \text{and} \quad \|f - f_0\|_{C^0} \le \frac{1}{\varepsilon}\left(\Lambda_\sigma(f_0) - \int f_0 \, d\mu_0\right) \le \varepsilon. \tag{3.1}$$

By Lemma 3.2, μ belongs to $\mathcal{M}(\Sigma, \sigma)$ and is f-maximizing.

Take an open subset U of $\mathcal{M}(\Sigma, \sigma)$ such that $\operatorname{supp}(b_\mu) \subset U$ and $b_{\mu_0}(U \setminus \operatorname{supp}(b_\mu)) < \varepsilon$. Since $\mathcal{M}^{\mathrm{e}}(\Sigma, \sigma)$ is a metric space, it is a normal space. By the Urysohn lemma, there exists a function $g \in C(\mathcal{M}(\Sigma, \sigma))$ such that $\|g\|_{C^0} = 1$, $g \equiv 0$ on $\mathcal{M}(\Sigma, \sigma) \setminus U$, and $g \equiv 1$ on $\operatorname{supp}(b_\mu)$. We have

$$\begin{aligned} b_{\mu_0}(\operatorname{supp}(b_\mu)) &> b_{\mu_0}(U) - \varepsilon > b_{\mu_0}(g) - \varepsilon \ge b_\mu(g) - 2\varepsilon \\ &\ge b_\mu(\operatorname{supp}(b_\mu)) - 2\varepsilon = 1 - 2\varepsilon > 0. \end{aligned}$$

To deduce the third inequality, we have used $\|b_\mu - b_{\mu_0}\| = \|\mu - \mu_0\| \le \varepsilon$ from Lemma 3.4 and the first inequality in equation (3.1). Since b_{μ_0} is non-atomic, it follows that the set $\{\nu_t : t \in [0, 1],\ \nu_t \in \operatorname{supp}(b_\mu)\}$ contains uncountably many elements, which belong to $\mathcal{M}_{\max}(f)$ by Lemma 3.3. Since $f_0 \in C(\Sigma)$ and $\varepsilon \in (0, 1/2)$ are arbitrary, the proof of Theorem A(b) is complete.

3.3. *Proof of Theorem B.* Let Σ be a subshift that satisfies $h^{\perp}_{\mathrm{spec}}(\Sigma) = 0$ and suppose that ergodic measures on Σ are entropy dense. Recall that

$$\mathscr{R}_0 = \{f \in C(\Sigma) : h(\mu) = 0 \text{ for all } \mu \in \mathcal{M}_{\max}(f)\},$$

and let $f \in \mathscr{R}_0$. There exists an ergodic measure $\mu \in \mathcal{M}_{\max}(f)$ with $h(\mu) = 0$. From the entropy density and the affinity of the entropy function, μ is approximated in the weak* topology by ergodic measures with positive entropy. Taking these approximating measures and repeating the argument in the proof of Theorem A(b) in §3.2, one can show that f is in the closure of functions for which there exists a maximizing measure with positive entropy. Hence, $\mathscr{R}_0$ has empty interior. Then, the desired statement is a direct consequence of Theorem A(b).

4. *Examples and applications*

In this section, we provide examples of non-Markov symbolic dynamical systems to which Theorem A or Theorem B applies.

4.1. *Piecewise monotonic maps and their coding spaces.* We say $T : [0, 1] \to [0, 1]$ is a *piecewise monotonic map* if there exist finitely many non-degenerate closed subintervals $I_1, \ldots, I_k$ of $[0, 1]$ with disjoint interiors such that $\bigcup_{j=1}^{k} I_j = [0, 1]$, and the restriction of T to the interior of I_j is strictly monotone and continuous for every $1 \le j \le k$. The intervals I_j are called *monotonic pieces* of T. Given a piecewise monotonic map T with k monotonic pieces $I_1, \ldots, I_k$, let $X_T = \bigcap_{n=0}^{\infty} T^{-n}(\bigcup_{j=1}^{k}(I_j \setminus \partial I_j))$ and define $\pi : X_T \to \{1, \ldots, k\}^{\mathbb{N}}$ by $T^n(\pi(x)) \in I_{(\pi(x))_{n+1}} \setminus \partial I_{(\pi(x))_{n+1}}$ for every $n \ge 0$. The closure of $\pi(X_T)$ in the full shift space $\{1, \ldots, k\}^{\mathbb{N}}$ is a subshift, called the *coding space* of T and denoted by Σ_T. Notice that $\pi \circ T|_{X_T} = \sigma \circ \pi$, where σ denotes the left shift acting on Σ_T. We say T is *transitive* if there is $x \in [0, 1]$ such that $\{T^n(x) : n \ge 0\}$ is dense in $[0, 1]$. It is easy

https://doi.org/10.1017/etds.2025.10206 Published online by Cambridge University Press

to see that if T is transitive, then π is injective. If, moreover, T has positive topological entropy, then $h_{\text{top}}(\Sigma_T) > 0$.

Continuous topologically mixing piecewise monotonic maps and their coding spaces have the Bowen specification property (see e.g. [9]), while piecewise monotonic maps with discontinuities and their coding spaces rarely have that. Typical examples of such maps are as follows:

- (the (α, β)-transformation) $T_{\alpha,\beta}(x) = \beta x + \alpha - \lfloor \beta x + \alpha \rfloor$ ($\alpha \in [0, 1)$, $\beta > 1$);
- (the $(-\beta)$-transformation) $T_{-\beta}(x) = -\beta x + \lfloor \beta x \rfloor + 1$ ($\beta > 1$).

The map $T_{0,\beta}$ is called the *β-transformation*. From the general result of Buzzi [9, Theorem 1.5] it follows that: for each fixed $\alpha \in [0, 1)$ the set of $\beta > 1$ for which $\Sigma_{T_{\alpha,\beta}}$ has Bowen's specification property is of zero Lebesgue measure; the set of $\beta > 1$ for which $\Sigma_{T_{-\beta}}$ has Bowen's specification property is of zero Lebesgue measure.

PROPOSITION 4.1. *If $T\colon [0, 1] \to [0, 1]$ is a transitive piecewise monotonic map with positive topological entropy, then $h^{\perp}_{\text{spec}}(\Sigma_T) = 0$.*

The original idea in our proof of Proposition 4.1 is due to Climenhaga [12]. We will use a Markov diagram [19], a directed graph whose vertices are closed subsets of the coding space. For the rest of this subsection, let $T\colon [0, 1] \to [0, 1]$ be a transitive piecewise monotonic map with positive topological entropy and k monotonic pieces. In what follows, we use the notation $[\cdot]$ to denote cylinders in the coding space Σ_T. Let C be a non-empty closed subset of Σ_T such that $C \subset [i]$ holds for some $i \in \{1, \ldots, k\}$. We say a non-empty closed subset D of Σ_T is a *successor* of C if there exists $j \in \{1, \ldots, k\}$ such that $D = [j] \cap \sigma C$. If D is a successor of C, we write $C \to D$. We set $\mathcal{D}_0 = \{[1], \ldots, [k]\}$, and define $\mathcal{D}_1, \mathcal{D}_2, \ldots$ by the recursion formula

$$\mathcal{D}_{n+1} = \mathcal{D}_n \cup \{D\colon D \text{ is a successor of some } C \in \mathcal{D}_n\}.$$

We set

$$\mathcal{D} = \bigcup_{n=0}^{\infty} \mathcal{D}_n.$$

The directed graph $(\mathcal{D}, \to)$ is called a *Markov diagram* associated with Σ_T.

For a set $\mathcal{C} \subset \mathcal{D}$, let $\Sigma_{\mathcal{C}}$ denote the set of paths $C_1 \to C_2 \to \cdots$ that contains infinitely many edges and all whose vertices are contained in $\mathcal{C}$. We define a map $\Psi\colon \Sigma_{\mathcal{D}} \to \{1, \ldots, k\}^{\mathbb{N}}$ as follows. For each vertex C_1 or each path $C_1 \to \cdots \to C_n$ in the diagram $(\mathcal{D}, \to)$, define

$$(C_1 \cdots C_n) = x_1 \cdots x_n \in \{1, \ldots, k\}^n,$$

where $x_i \in \{1, \ldots, k\}$ and $C_i \subset [x_i]$ for $i = 1, \ldots, n$. For $(C_n)_{n=1}^{\infty} \in \{1, \ldots, k\}^{\mathbb{N}}$, define

$$\Psi((C_n)_{n=1}^{\infty}) \in \bigcap_{n=1}^{\infty} [(C_1 \cdots C_n)].$$

https://doi.org/10.1017/etds.2025.10206 Published online by Cambridge University Press

Note that $\Sigma_{\mathcal{D}}$ is a Markov shift over the countably infinite alphabet $\mathcal{D}$, $\Psi(\Sigma_{\mathcal{D}}) = \Sigma_T$, and Ψ is a semiconjugacy between the left shift on $\Sigma_{\mathcal{D}}$ and $\sigma\colon \Sigma_T \to \Sigma_T$. The next lemma follows from [20, Theorems 10, 11].

LEMMA 4.2. [23, pp. 224, 226] *There exist a subset $\mathcal{C}$ of $\mathcal{D}$ and a finite subset $\mathcal{F}$ of $\mathcal{C}$ such that:*

(a) *for any pair (C, C') of vertices in $\mathcal{C}$, there is a path that leads from C to C';*
(b) *$C \in \mathcal{C}$ and $C \to D$ imply $D \in \mathcal{C}$;*
(c) $\Psi(\{(C_n)_{n=1}^{\infty} \in \Sigma_{\mathcal{C}}\colon C_1 \in \mathcal{F}\}) = \Psi(\Sigma_{\mathcal{C}}) = \Sigma_T$.

Proof of Proposition 4.1. Let $\mathcal{C}$ be a subset of $\mathcal{D}$ and $\mathcal{F} \subset \mathcal{C}$ be a finite set for which the conclusion of Lemma 4.2 holds. Let $H > 0$ and let $N > (4/H)\log 2$ be an integer with $\mathcal{F} \subset \mathcal{D}_N$. We now define $\mathcal{C}^p, \mathcal{G}, \mathcal{C}^s \subset \mathcal{L}(\Sigma_T)$ by

$$\mathcal{C}^p = \{\emptyset\},$$
$$\mathcal{G} = \{\emptyset\} \cup \bigcup_{n=1}^{\infty} \{(C_1 \cdots C_n)\colon C_1 \to \cdots \to C_n,\ C_1, C_n \in \mathcal{D}_N \cap \mathcal{C}\},$$
$$\mathcal{C}^s = \{\emptyset\} \cup \bigcup_{n=1}^{\infty} \left\{ \begin{matrix} (C_1 \cdots C_n)\colon C_1 \to \cdots \to C_n \\ C_1 \in \mathcal{D}_{N+1},\ C_i \in \mathcal{C} \setminus \mathcal{D}_N \text{ for } 1 \le i \le n \end{matrix} \right\}.$$

Let $w \in \mathcal{L}(\Sigma_T) \setminus \{\emptyset\}$. By Lemma 4.2(c), there exists $(C_n)_{n=1}^{\infty} \in \Sigma_{\mathcal{C}}$ such that $\Psi((C_n)_{n=1}^{\infty}) \in [w]$ and $C_1 \in \mathcal{F}$. Set $i_0 = \sup\{1 \le i \le |w|\colon C_i \in \mathcal{D}_N\}$. Since $C_1 \in \mathcal{F} \subset \mathcal{D}_N$, we have $i_0 \ge 1$. Set $u = (C_1 \cdots C_{i_0})$ and $v = (C_{i_0+1} \cdots C_{|w|})$. By Lemma 4.2(b), we have $u \in \mathcal{G}$, $v \in \mathcal{C}^s$ and $w = uv$, which implies $\mathcal{L}(\Sigma_T) \subset \mathcal{C}^p\mathcal{G}\mathcal{C}^s = \mathcal{G}\mathcal{C}^s$. The reverse inclusion is obvious.

For each $M \in \mathbb{N}$, by the finiteness of $\#\mathcal{D}_{M+N}$ and Lemma 4.2(a), there exists $t_M \in \mathbb{N}$ such that for any pair of vertices in $\mathcal{D}_{M+N} \cap \mathcal{C}$, there is a path not containing more than t_M edges and leading from one vertex to the other. Let $u, v \in \mathcal{G}^M$. By the above definition of $\mathcal{G}$, there exist two paths $C_1 \to \cdots \to C_{|u|}$ and $D_1 \to \cdots \to D_{|v|}$ such that $C_{|u|} \in \mathcal{D}_{M+N} \cap \mathcal{C}$ and $D_1 \in \mathcal{D}_N \cap \mathcal{C}$. Hence, there is a path $C_{|u|} \to \cdots \to D_1$ not containing more than t_M edges, which implies that $\mathcal{G}^M$ has (W)-specification with gap size t_M.

Notice that for all $q \in \mathbb{N}$, the number of paths starting at each vertex in $\mathcal{D} \setminus \mathcal{D}_N$ not containing more than qN edges and all whose vertices are contained in $\mathcal{D} \setminus \mathcal{D}_N$ does not exceed 2^q (see the proof of [20, Corollary 1(i)]). Hence, for each integer $n = (q-1)N + r$ with $q, r \in \mathbb{N}$, $q \ge 2$, $0 \le r \le N-1$, we have $(q-1)N \le n < qN$ and $\#\mathcal{C}_n^s \le 2^q \#\mathcal{D}_{N+1}$. If n is sufficiently large, then

$$\frac{\log \#\mathcal{C}_n^s}{n} \le \frac{q \log 2}{n} + \frac{\log \#\mathcal{D}_{N+1}}{n} \le \frac{2 \log 2}{N} + \frac{H}{2} < H.$$

Since $H > 0$ is arbitrary, $h_{\mathrm{spec}}^{\perp}(\Sigma_T) = 0$ holds. □

By Proposition 4.1, if T is a transitive piecewise monotonic interval map with positive topological entropy, then $h_{\mathrm{spec}}^{\perp}(\Sigma_T) = 0$. The entropy density for Σ_T follows from the

https://doi.org/10.1017/etds.2025.10206 Published online by Cambridge University Press

density of shift-invariant closed orbit measures in the space of ergodic measures [39, Proposition 3.1]. The latter condition is equivalent to the density of T-invariant closed orbit measures in the space of ergodic measures [44, Theorem A]. The density of closed orbit measures has been verified for a wide class of transitive piecewise monotonic maps with positive topological entropy (see e.g. [21–23, 39, 41]). Hence, Theorem B applies to the coding spaces of these piecewise monotonic maps. It has been conjectured that the density of closed orbit measures holds for any transitive piecewise monotonic map with positive topological entropy [23].

4.2. *Coded shifts.* Let $\mathcal{A}$ be a finite subset of $\mathbb{N}$. Let $\mathcal{A}^0$ denote the singleton consisting of the empty word $\emptyset$. For each $n \in \mathbb{N}$, let $\mathcal{A}^n$ denote the set of words from $\mathcal{A}$ with word length n. Put $\mathcal{A}^* = \bigcup_{n=0}^{\infty} \mathcal{A}^n$. Let $\mathcal{H}$ be a countable subset of $\mathcal{A}^*$ that contains the empty word. The *coded shift* $\Sigma_{\mathcal{H}}$ generated by $\mathcal{H}$ is the two-sided subshift whose language is

$$\overline{\{w^{(1)}w^{(2)}\cdots w^{(k)}\colon w^{(1)},\dots,w^{(k)} \in \mathcal{H}, k \in \mathbb{N}\}},$$

where $\overline{\{\cdot\}}$ denotes the closure under the operation of passing to subwords.

The coded shift $\Sigma_{\mathcal{H}}$ has the following natural language decomposition $\mathcal{L}(\Sigma_{\mathcal{H}}) = \mathcal{C}^p\mathcal{G}\mathcal{C}^s$ for which $\mathcal{G}^M$ has (W)-specification for all $M \in \mathbb{N}$:

$$\begin{aligned} \mathcal{C}^p &= \bigcup_{n=1}^{\infty} i_n^s(\mathcal{H}),\\ \mathcal{G} &= \bigcup_{k=1}^{\infty}\{w^{(1)}\cdots w^{(k)}\colon w^{(1)},\dots,w^{(k)} \in \mathcal{H}\},\\ \mathcal{C}^s &= \bigcup_{n=1}^{\infty} i_n^p(\mathcal{H}), \end{aligned} \tag{4.1}$$

where i_n^s (respectively i_n^p) denotes the map that extracts the last (respectively the first) n symbols of a word $w \in \mathcal{A}^*$, and satisfies $i_n^s(w) = \emptyset$ if $|w| < n$ (respectively $i_n^p(w) = \emptyset$ if $|w| < n$). When $\mathcal{H}$ is a finite set, clearly $h(\mathcal{C}^p \cup \mathcal{C}^s) = 0$ holds, and $\Sigma_{\mathcal{H}}$ is a sofic shift [30]. Every transitive sofic shift is a coded shift [32, §13.5].

Coded shifts with countably infinite generators are sources of interesting examples. Well-known examples are S-gap shifts [32] and the Dyck shift [29]. As shown below, there exist intrinsically ergodic coded shifts for which the entropy density fails. For more details on coded shifts, see [14, 30].

4.2.1. *S-gap shifts.* Let S be an infinite subset of $\mathbb{N}_0$. An *S-gap shift* [32] is a two-sided subshift on $\{1, 2\}$ defined by the rule that the number of 2 terms between consecutive 1 terms is an integer in S. To be more precise, for $a \in \mathcal{A}$ and $n \in \mathbb{N}$, let a^n denote the n-fold concatenation of a and put $a^0 = \emptyset$, the empty word. The S-gap shift is the coded shift generated by $\{2^n 1\colon n \in S\} \cup \{\emptyset\}$. Generic S-gap shifts do not have the Bowen specification property [14, §3.3]. The topological entropy of the S-gap shift is $\log \lambda$, where λ is the unique solution to the equation $1 = \sum_{n\in S} x^{-n-1}$ (see [32, Exercise 4.3.7]). The natural decomposition equation (4.1) immediately gives $\#\mathcal{C}_n^p = 1 = \#\mathcal{C}_n^s$ for all $n \in \mathbb{N}$, and so the

https://doi.org/10.1017/etds.2025.10206 Published online by Cambridge University Press

obstruction entropy to specification is 0. Hence, Theorem A applies to any S-gap shift. Since ergodic measures on any S-gap shift are entropy dense [17], Theorem B applies to any S-gap shift.

4.2.2. *Coded shifts without the entropy density.* Kucherenko *et al* [30, Example 1] considered a coded shift Σ on $\{1, 2\}$ generated by $\{1^i 2^i : i \in I\} \cup \{\emptyset\}$, where I is an infinite subset of $\mathbb{N}$. The natural decomposition equation (4.1) implies $h(\mathcal{C}^p \cup \mathcal{C}^s) = 0$ and $0 = h^{\perp}_{\mathrm{spec}}(\Sigma) < h_{\mathrm{top}}(\Sigma)$, and so Theorem A applies. Theorem B does not apply since ergodic measures are not dense in $\mathcal{M}(\Sigma, \sigma)$ and, consequently, the entropy density fails. Indeed, if both 1 and 2 appear in $x \in \Sigma$ infinitely many times, then we have

$$\lim_{n\to\infty} \frac{1}{n} \#\{0 \le i \le n-1 : x_i = 1\} = \frac{1}{2},$$

provided the limit exists. For $a \in \{1, 2\}$, define $a^\infty \in \{1, 2\}^{\mathbb{Z}}$ by $(a^\infty)_i = a$ for all $i \in \mathbb{Z}$, and for $x \in \Sigma$, let δ_x denote the unit point mass at x. From the above observation and the Birkhoff ergodic theorem, it follows that any non-ergodic measure of the form $(1 - t) \cdot \delta_{1^\infty} + t \cdot \delta_{2^\infty}$, $t \in (0, 1/2) \cup (1/2, 1)$ cannot be approximated by ergodic ones.

4.3. *Multidimensional β-transformations.* Let Σ be a subshift. We say $w = w_1 \cdots w_{|w|} \in \mathcal{L}(\Sigma)$ is a *left constraint* if there exists $v \in \mathcal{L}(\Sigma)$ such that $w_2 \cdots w_{|w|} v \in \mathcal{L}(\Sigma)$ and $wv \notin \mathcal{L}(\Sigma)$. A *right constraint* is defined analogously. Let $\mathcal{C}^\ell$, $\mathcal{C}^r$ denote the collections of left constraints, right constraints, respectively. If Σ is topologically transitive, then $\mathcal{C}^\ell$, $\mathcal{C}^r$ form a complete list of obstructions to specification [11, Theorem 1.6(1)]. Combining [11, Theorems 1.1, 1.4, 1.6] yields the following statement.

PROPOSITION 4.3. *Let Σ be a subshift.*

(a) *Suppose Σ is topologically transitive. If $h(\mathcal{C}^\ell \cup \mathcal{C}^r) < h_{\mathrm{top}}(\Sigma)$, then $h^{\perp}_{\mathrm{spec}}(\Sigma) < h_{\mathrm{top}}(\Sigma)$ holds and Theorem A applies.*

(b) *Suppose Σ is one-sided and topologically exact: for any $w \in \mathcal{L}(\Sigma) \setminus \{\emptyset\}$, there exists $N \in \mathbb{N}$ such that $\sigma^N[w] = \Sigma$. If $h(\mathcal{C}^\ell) < h_{\mathrm{top}}(\Sigma)$, then $h^{\perp}_{\mathrm{spec}}(\Sigma) < h_{\mathrm{top}}(\Sigma)$ holds and Theorem A applies.*

Proposition 4.3(b) applies to many of the piecewise affine transformations investigated by Buzzi [9]. These are maps on $[0, 1)^d$, $d \in \mathbb{N}$ given by $T(x) = Bx + b \bmod \mathbb{Z}^d$, where $B : \mathbb{R}^d \to \mathbb{R}^d$ is an expanding linear map and $b \in \mathbb{R}^d$. The coding space Σ_T of T is defined as in the case of piecewise monotonic maps in §4.1, using the collection of maximal open subsets of $(0, 1)^d$ on which $T(x) - Bx$ is constant. Buzzi showed that $h(\mathcal{C}^\ell) < h_{\mathrm{top}}(\Sigma_T)$ and that Σ_T is topologically exact if either: (i) all eigenvalues of B exceed $1 + \sqrt{d}$ in absolute value; or (ii) B, b have all integer entries. Since the maps with $d = 1$ are almost the β-transformations, the maps with $d \ge 2$ are called *multidimensional β-transformations*.

For the coding spaces of multidimensional β-transformations, little is known when the obstruction entropy specification is 0, or when the entropy density holds and Theorem B applies.

https://doi.org/10.1017/etds.2025.10206 Published online by Cambridge University Press

5. *Subshifts with positive obstruction entropy to specification*

This last section is devoted to Theorem C. In §5.1, we introduce fat S-gap shifts by modifying the definition of S-gap shifts. In §5.2, we complete the proof of Theorem C by showing that any fat S-gap shift on N symbols, $N \geq 3$ has positive obstruction entropy to specification.

5.1. *Fat S-gap shifts.*

Let S be an infinite subset of $\mathbb{N}_0$. A *fat S gap shift* is a two-sided subshift on $\{1, \ldots, N\}$, $N \geq 2$ defined by the rule that the lengths of the words between consecutive 1 terms are integers in S. More precisely, the fat S-gap shift $\Sigma_{S,N}$ on N symbols is the coded shift generated by

$$\mathcal{H} = \{w1 \colon w \in \{2, \ldots, N\}^n,\ n \in S\} \cup \{\emptyset\}. \tag{5.1}$$

Note that $\Sigma_{S,2}$ is nothing but an S-gap shift, whose obstruction entropy to specification is 0.

For all $S \subset \mathbb{N}_0$ and $N \geq 2$, we clearly have $\Sigma_{S,N} \supset \{2, \ldots, N\}^{\mathbb{Z}}$ and, thus, $h_{\text{top}}(\Sigma_{S,N}) \geq \log(N-1)$. The next lemma asserts that this inequality is strict if $N \geq 3$.

LEMMA 5.1. *For any infinite set $S \subset \mathbb{N}$ and any integer $N \geq 3$, we have*

$$h_{\text{top}}(\Sigma_{S,N}) > \log(N-1).$$

Proof. The original idea of the proof of this lemma is due to Climenhaga [13]. For ease of notation, we set $l_n = \#\mathcal{L}_n(\Sigma_{S,N})$. The radius of convergence of the power series $F(x) = \sum_{n=1}^{\infty} l_n x^n$ equals $e^{-h_{\text{top}}(\Sigma_{S,N})}$. For $k, n \in \mathbb{N}$, we set

$$A_n^k = \#\{w^{(1)} \cdots w^{(k)} \colon w^{(1)}, \ldots, w^{(k)} \in \mathcal{H} \setminus \{\emptyset\},\ |w^{(1)}| + \cdots + |w^{(k)}| = n\}.$$

Then, it is not difficult to prove the following properties:

(i) $\sum_{k=1}^{\infty} A_n^k \leq l_n$;
(ii) $A_n^1 = (N-1)^{n-1}$ if $n-1 \in S$ and $A_n^1 = 0$ otherwise;
(iii) $A_n^{k+\ell} = \sum_{m=1}^{n-1} A_{n-m}^k A_m^\ell$ for $k, \ell \in \mathbb{N}$.

We introduce a power series $F_k(x) = \sum_{n=1}^{\infty} A_n^k x^n$. Then property (iii) gives $F_k(x) = F_{k-1}(x)F_1(x) = \cdots = F_1(x)^k$. For all $x \geq 0$, we have

$$\sum_{k=1}^{\infty} F_1(x)^k = \sum_{k=1}^{\infty} F_k(x) = \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} A_n^k x^n \leq F(x),$$

where the last inequality follows from property (i). Let x_0 denote the unique positive solution to the equation

$$N - 1 = \sum_{n \in S} ((N-1)x)^{n+1} \quad \text{for } 0 < x < \frac{1}{N-1}.$$

By property (ii), we have

$$F_1(x_0) = \sum_{n=1}^{\infty} A_n^1 x_0^n = \frac{1}{N-1} \sum_{n \in S} ((N-1)x_0)^{n+1} = 1,$$

https://doi.org/10.1017/etds.2025.10206 Published online by Cambridge University Press

which implies that $F(x_0)$ diverges. Hence, the radius of convergence of $F(x)$ does not exceed x_0 and we obtain $h_{\mathrm{top}}(\Sigma_{S,N}) \geq \log x_0^{-1} > \log(N-1)$, as required. □

5.2. *Proof of Theorem C.* Let $S = \{2^n : n \in \mathbb{N}\}$ and let $N \geq 3$ be an integer. We consider the fat S-gap shift $\Sigma = \Sigma_{S,N}$. Lemma 5.1 gives $h_{\mathrm{top}}(\Sigma) > \log(N-1)$. Applying the natural decomposition of coded shifts in equation (4.1), we immediately get $h(\mathcal{C}^p \cup \mathcal{C}^s) = \log(N-1)$ and, hence, $h_{\mathrm{top}}(\Sigma) > \log(N-1) \geq h^{\perp}_{\mathrm{spec}}(\Sigma)$.

Let $\mathcal{L}(\Sigma) = \tilde{\mathcal{C}}^p \tilde{\mathcal{G}} \tilde{\mathcal{C}}^s$ be an arbitrary decomposition such that $\tilde{\mathcal{G}}^M$ has (W)-specification for all $M \in \mathbb{N}$. The definition of Σ gives $w1w \in \mathcal{L}(\Sigma)$ for all $w \in \{2, \ldots, N\}^*$. For each $w \in \{2, \ldots, N\}^* \setminus \{\emptyset\}$, we fix a decomposition

$$w1w = p(w1w) \cdot c(w1w) \cdot s(w1w), \quad p(w1w) \in \tilde{\mathcal{C}}^p, c(w1w) \in \tilde{\mathcal{G}}, s(w1w) \in \tilde{\mathcal{C}}^s.$$

Since 1 appears in $w1w$ only once, it appears in only one of the three words $p(w1w)$, $c(w1w)$, $s(w1w)$, which can be the empty word $\emptyset$. We will use the following simple estimates on word lengths for $r \in \{p, s\}$:

$$0 \leq |r(w1w)| \leq |w| \text{ if 1 does not appear in } r(w1w); \tag{5.2}$$

$$|w| + 1 \leq |r(w1w)| \leq 2|w| + 1 \text{ if 1 appears in } r(w1w). \tag{5.3}$$

We estimate $h(\tilde{\mathcal{C}}^p \cup \tilde{\mathcal{C}}^s)$ from below by constructing a countably infinite family of finite-to-one maps that take values in $\tilde{\mathcal{C}}^p \cup \tilde{\mathcal{C}}^s$. We need some notation introduced in §4.2.

Case 1: $\tilde{\mathcal{G}} \subset \{2, \ldots, N\}^*$. For each $k \in \mathbb{N}$, we define a map $\Phi_k \colon \{2, \ldots, N\}^{2^k} \mapsto \bigcup_{n=2^k+1}^{2^{k+1}+1} (\tilde{\mathcal{C}}^p_n \cup \tilde{\mathcal{C}}^s_n)$ by

$$\Phi_k(w) = \begin{cases} s(w1w) & \text{if 1 appears in } s(w1w), \\ p(w1w) & \text{otherwise.} \end{cases} \tag{5.4}$$

For any $w \in \{2, \ldots, N\}^{2^k}$, 1 does not appear in $c(w1w)$ since $\tilde{\mathcal{G}} \subset \{2, \ldots, N\}^*$. By equation (5.3), if 1 appears in $p(w1w)$, then $2^k + 1 \leq |p(w1w)| \leq 2^{k+1} + 1$ and if 1 appears in $s(w1w)$, then $2^k + 1 \leq |s(w1w)| \leq 2^{k+1} + 1$. Hence, Φ_k is well defined.

For any $w \in \{2, \ldots, N\}^{2^k}$, 1 appears in $\Phi_k(w)$ by the definition in equation (5.4). This implies that Φ_k is injective and so

$$\#\Bigg(\bigcup_{n=2^\ell+1}^{2^{k+1}+1} (\tilde{\mathcal{C}}^p_n \cup \tilde{\mathcal{C}}^s_n) \Bigg) \geq \#\Phi_k(\{2, \ldots, N\}^{2^k}) \geq \#\{2, \ldots, N\}^{2^k} = (N-1)^{2^k}.$$

Pick $n_k \in \{2^k + 1, \ldots, 2^{k+1} + 1\}$ such that

$$\#(\tilde{\mathcal{C}}^p_{n_k} \cup \tilde{\mathcal{C}}^s_{n_k}) \geq \frac{(N-1)^{2^k}}{2^k + 1}.$$

Since $\liminf_{k\to\infty} n_k = \infty$, we have

$$\begin{aligned} h(\tilde{\mathcal{C}}^p \cup \tilde{\mathcal{C}}^s) &\geq \limsup_{l\to\infty} \frac{\log \#(\tilde{\mathcal{C}}^p_{n_k} \cup \tilde{\mathcal{C}}^s_{n_k})}{n_k} \\ &\geq \limsup_{l\to\infty} \frac{1}{n_k} \log \frac{(N-1)^{(n_k-1)/2}}{n_k} = \frac{\log(N-1)}{2}. \end{aligned} \tag{5.5}$$

https://doi.org/10.1017/etds.2025.10206 Published online by Cambridge University Press

Case 2: $\tilde{\mathcal{G}} \not\subset \{2, \ldots, N\}^*$. We modify the argument in Case 1. Fix $u \in \tilde{\mathcal{G}} \setminus \{2, \ldots, N\}^*$ for which there exist $m \geq 2$, $u^{(0)}, u^{(m)} \in \{2, \ldots, N\}^*$, $u^{(1)}, \ldots, u^{(m-1)} \in \{2, \ldots, N\}^* \setminus \{\emptyset\}$ such that

$$u = u^{(0)} 1 u^{(1)} 1 \cdots u^{(m-1)} 1 u^{(m)}. \tag{5.6}$$

Let $t \geq 0$ be a gap size for the (W)-specification of $\tilde{\mathcal{G}}$. For each $k \in \mathbb{N}$, let

$$a_k = 2^k - 2t - |u^{(m)}| - 1. \tag{5.7}$$

In what follows, we assume k is sufficiently large so that

$$\frac{2a_k}{3} > 2^{k-1}. \tag{5.8}$$

Define a map $\Psi_k \colon \{2, \ldots, N\}^{a_k} \mapsto \bigcup_{n=a_k-2^{k-1}}^{2a_k+1} (\tilde{\mathcal{C}}_n^p \cup \tilde{\mathcal{C}}_n^s)$ by

$$\Psi_k(w) = \begin{cases} s(w1w) & \text{if 1 appears in } s(w1w), \\ p(w1w) & \text{otherwise.} \end{cases} \tag{5.9}$$

Compare equation (5.9) and equation (5.4). To see that Ψ_k is well defined, let $w \in \{2, \ldots, N\}^{a_k}$. By equation (5.3), if 1 appears in $p(w1w)$, then $a_k + 1 \leq |p(w1w)| \leq 2a_k + 1$ and if 1 appears in $s(w1w)$, then $a_k + 1 \leq |s(w1w)| \leq 2a_k + 1$. If 1 appears in $c(w1w)$, then the definition in equation (5.9) gives $\Psi_k(w) = p(w1w)$ and the lemma below yields $a_k - 2^{k-1} \leq |p(w1w)| \leq a_k$. Hence, Ψ_k is well defined.

LEMMA 5.2. *If* $w \in \{2, \ldots, N\}^{a_k}$ *and* 1 *appears in* $c(w1w)$, *then* $a_k - 2^{k-1} \leq |p(w1w)| \leq a_k$.

Proof. The upper bound follows from equation (5.2). For the lower bound, note that there exist $v, v' \in \{2, \ldots, N\}^*$ such that

$$c(w1w) = v1v'.$$

We have $w1w = p(w1w) \cdot v1v' \cdot s(w1w)$, and so v is a subword of w and $|p(w1w)| = |w| - |v|$. It suffices to show that $|v| \leq 2^{k-1}$.

By equation (5.7), we have

$$|v| \leq |w| = a_k < 2^k - 2t - |u^{(m)}|. \tag{5.10}$$

Since $u, c(w1w) \in \tilde{\mathcal{G}}$, there exists $u' \in \mathcal{L}(\Sigma)$ such that $|u'| \leq t$ and $uu'c(w1w) \in \mathcal{L}(\Sigma)$, and hence, $1u^{(m)}u'v1 \in \mathcal{L}(\Sigma)$. If 1 does not appear in u', then it does not appear in $u^{(m)}u'v$ either. Since Σ is the coded shift generated by $\mathcal{H}$ in equation (5.1), it follows that $|u^{(m)}u'v| \in S$. By equation (5.10), we have

$$|v| \leq |u^{(m)}u'v| \leq |u^{(m)}| + |u'| + |w| < 2^k - t \leq 2^k.$$

Since $S \cap [2^{k-1} + 1, 2^k - 1] = \emptyset$, we obtain $|v| \leq 2^{k-1}$ as required. If 1 appears in u', then let $t' \in [1, |u'|]$ be the largest integer such that $i_{t'}^s(u')$ does not contain 1. We have $t' \leq |u'| \leq t$, $1i_{t'}^s(u')v1 \in \mathcal{L}(\Sigma)$, and 1 does not appear in $i_{t'}^s(u')v$. Since Σ is the coded

https://doi.org/10.1017/etds.2025.10206 Published online by Cambridge University Press

shift generated by $\mathcal{H}$ in equation (5.1), this implies $|i^s_{t'}(u')v| = t' + |v| \in S$. Again by equation (5.10), we obtain

$$|v| \le t' + |v| \le t + |w| < 2^k.$$

Since $S \cap [2^{k-1} + 1, 2^k - 1] = \emptyset$, we obtain $|v| \le 2^{k-1}$, as required. □

LEMMA 5.3. *The map Ψ_k is at most $(N-1)^{2^{k-1}}$ to 1.*

Proof. Let $w, w' \in \{2, \dots, N\}^{a_k}$, $w \ne w'$, and suppose $\Psi_k(w) = \Psi_k(w')$. The definition in equation (5.9) implies that 1 appears in neither $s(w1w)$ nor $s(w'1w')$, and $\Psi_k(w) = p(w1w)$, $\Psi_k(w') = p(w'1w')$.

If 1 does not appear in $c(w1w)$ and appears in $c(w'1w')$, then by equations (5.2), (5.3), we get $|p(w1w)| \ge a_k + 1$ and $|p(w'1w')| \le a_k$, and so $\Psi_k(w) \ne \Psi_k(w')$, which is a contradiction. If 1 appears in $c(w1w)$ and does not appear in $c(w'1w')$, then we also obtain a contradiction. If 1 appears in neither $c(w1w)$ nor $c(w'1w')$, then it appears in both $p(w1w)$ and $p(w'1w')$, and so $w = w'$, which is a contradiction. Therefore, 1 appears in both $c(w1w)$ and $c(w'1w')$. By Lemma 5.2, there exists $n \in \{a_k - 2^{k-1}, \dots, a_k\}$ such that $i^p_n(w) = i^p_n(w')$. Therefore, Ψ_k is at most $(N-1)^{2^{k-1}}$ to 1. □

By Lemma 5.3, we have

$$\#\Bigg(\bigcup_{n=a_k-2^{k-1}}^{2a_k+1} (\tilde{\mathcal{C}}^p_n \cup \tilde{\mathcal{C}}^s_n)\Bigg) \ge \#\Psi_k(\{2, \dots, N\}^{a_k}) \ge \#\{2, \dots, N\}^{a_k}(N-1)^{-2^{k-1}} = (N-1)^{a_k - 2^{k-1}}.$$

Hence, there exists $n_k \in \{a_k - 2^{k-1}, \dots, 2a_k + 1\}$ such that

$$\#(\tilde{\mathcal{C}}^p_{n_k} \cup \tilde{\mathcal{C}}^s_{n_k}) \ge \frac{(N-1)^{a_k - 2^{k-1}}}{a_k + 2 + 2^{k-1}} \ge \frac{(N-1)^{a_k/3}}{5a_k/3 + 2} \ge \frac{(N-1)^{(n_k-1)/6}}{5n_k + 2}.$$

For the last inequality, we have used equation (5.8). Since $\lim_{k\to\infty} n_k = \infty$, we have

$$h(\tilde{\mathcal{C}}^p \cup \tilde{\mathcal{C}}^s) \ge \limsup_{k\to\infty} \frac{\log \#(\tilde{\mathcal{C}}^p_{n_k} \cup \tilde{\mathcal{C}}^s_{n_k})}{n_k} \ge \limsup_{k\to\infty} \frac{1}{n_k} \log \frac{(N-1)^{(n_k-1)/6}}{5n_k + 2} \ge \frac{\log(N-1)}{6}. \tag{5.11}$$

Since $\mathcal{L}(\Sigma) = \tilde{\mathcal{C}}^p \tilde{\mathcal{G}} \tilde{\mathcal{C}}^s$ is an arbitrary decomposition such that $\tilde{\mathcal{G}}^M$ has (W)-specification for all $M \in \mathbb{N}$, from equations (5.5) and (5.11), we obtain $h^{\perp}_{\mathrm{spec}}(\Sigma) \ge (1/6)\log(N-1)$. The proof of Theorem C is complete.

Acknowledgments. We thank the referee for their careful reading of the manuscript and giving valuable suggestions. M.S. was supported by the JSPS KAKENHI 21K13816. H.T. was supported by the JSPS KAKENHI 25K21999. K.Y. was supported by the JSPS KAKENHI 21K03321. This work was supported by the Research Institute for Mathematical Sciences, an International Joint Usage/Research Center located in Kyoto University.

https://doi.org/10.1017/etds.2025.10206 Published online by Cambridge University Press

References

[1] A. T. Baraviera, R. Leplaideur and A. O. Lopes. *Ergodic Optimization, Zero Temperature Limits and the Max-Plus Algebra (IMPA Mathematical Publications, 29th Brazilian Mathematics Colloquium)* Instituto Nacional de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, 2013.

[2] E. Bishop and R. R. Phelps. The support functionals of a convex set. *Convexity (AMS Proceedings of Symposia in Pure Mathematics, 7).* Ed. V. L. Klee. American Mathematical Society, Providence, RI, 1963, pp. 27–35.

[3] J. Bochi. Ergodic optimization of Birkhoff averages and Lyapunov exponents. *Proceedings of the International Congress of Mathematicians, Rio de Janeiro, 2018. Volume III. Invited Lectures*. Ed. B. Sirakov, P. Ney de Souza and M. Viana. World Scientific Publishing, Hackensack, NJ, 2018, pp. 1825–1846.

[4] T. Bousch. Le poisson n'a pas d'arêtes. *Ann. Inst. Henri Poincaré Probab. Stat.* **36** (2000), 489–508.

[5] T. Bousch. La condition de Walters , *Ann. Sci. Éc. Norm. Supér. (4)* **34** (2001), 5988–6017.

[6] T. Bousch and O. Jenkinson. Cohomology class of dynamically non-negative C^k functions. *Invent. Math.* **148** (2002), 207–217.

[7] R. Bowen. Entropy for group endomorphisms and homogeneous spaces, *Trans. Amer. Math. Soc.* **153** (1971), 401–414.

[8] J. Brémont. Entropy and maximizing measures of generic continuous functions. *C. R. Math. Acad. Sci.* **346** (2008), 199–201.

[9] J. Buzzi. Specification on the interval. *Trans. Amer. Math. Soc.* **349** (1997), 2737–2754.

[10] J. Buzzi. Subshifts of quasi-finite type. *Invent. Math.* **159** (2005), 369–406.

[11] V. Climenhaga. Specification and towers in shift spaces. *Comm. Math. Phys.* **364** (2018), 441–504.

[12] V. Climenhaga. V. Climenhaga's slides (accessed 21 May 2025). https://www.math.uh.edu/climenha/doc/marseille-specification.pdf.

[13] V. Climenhaga. V. Climenhaga's Math blog (accessed 21 May 2025). https://vaughnclimenhaga.wordpress.com/2014/09/08/entropy-of-s-gap-shifts/.

[14] V. Climenhaga and D. J. Thompson. Intrinsic ergodicity beyond specification: β-shifts, S-gap shifts, and their factors. *Israel J. Math.* **192** (2012), 785–817.

[15] V. Climenhaga and D. J. Thompson. Equilibrium states beyond specification and the Bowen property. *J. Lond. Math. Soc.* **87** (2013), 401–427.

[16] V. Climenhaga and D. J. Thompson. Intrinsic ergodicity via obstruction entropies. *Ergod. Th. & Dynam. Sys.* **34** (2014), 1816–1831.

[17] V. Climenhaga, D. J. Thompson and K. Yamamoto. Large deviations for systems with non-uniform structure. *Trans. Amer. Math. Soc.* **369** (2017), 4167–4192.

[18] G. Contreras, A. O. Lopes and P. Thieullen. Lyapunov minimizing measures for expanding maps of the circle. *Ergod. Th. & Dynam. Sys.* **21** (2001), 1379–1409.

[19] F. Hofbauer. On intrinsic ergodicity of piecewise monotonic transformations with positive entropy II. *Israel J. Math.* **38** (1981), 107–115.

[20] F. Hofbauer. Piecewise invertible dynamical systems. *Probab. Theory Related Fields* **72** (1986), 359–386.

[21] F. Hofbauer. Generic properties of invariant measures for simple piecewise monotonic transformations. *Israel J. Math.* **59** (1987), 64–80.

[22] F. Hofbauer. Generic properties of invariant measures for continuous piecewise monotonic transformations. *Monatsh. Math.* **106** (1988), 301–312.

[23] F. Hofbauer and P. Raith. Density of periodic orbit measures for transformations on the interval with two monotonic pieces. *Fund. Math.* **157** (1998), 221–234; dedicated to the memory of Wiesław Szlenk.

[24] B. Hunt and E. Ott. Optimal periodic orbits of chaotic systems. *Phys. Rev. Lett.* **76** (1996), 2254.

[25] B. Hunt and E. Ott. Optimal periodic orbits of chaotic systems occur at low period. *Phys. Rev. E* **54** (1996), 328.

[26] R. B. Israel. *Convexity in the Theory of Lattice Gases (Princeton Series in Physics)*. Princeton University Press, Princeton, NJ, 1979.

[27] O. Jenkinson. Ergodic optimization. *Discrete Contin. Dyn. Syst. A* **15** (2006), 197–224.

[28] O. Jenkinson. Ergodic optimization in dynamical systems. *Ergod. Th. & Dynam. Sys.* **39** (2019), 2593–2618.

[29] W. Krieger. On the uniqueness of the equilibrium state. *Math. Syst. Theory* **8** (1974/75), 97–104.

[30] T. Kucherenko, M. Schmoll and C. Wolf. Ergodic theory on coded shift spaces. *Adv. Math.* **457** (2024), 109913.

[31] D. Kwietniak, M. Łącka and P. Oprocha. A panorama of specification-like properties and their consequences. *Dynamics and Numbers (Contemporary Mathematics, 669)*. American Mathematical Society, Providence, RI, 2016, pp. 155–186.

https://doi.org/10.1017/etds.2025.10206 Published online by Cambridge University Press

[32] D. Lind and B. Marcus. *An Introduction to Symbolic Dynamics and Coding (Cambridge Mathematical Library)*, 2nd edn. Cambridge University Press, Cambridge, 2021.
[33] I. D. Morris. The Mañé–Conze–Guivarc'h lemma for intermittent maps of the circle. *Ergod. Th. & Dynam. Sys.* **29** (2009), 1603–1611.
[34] I. D. Morris. Ergodic optimization for generic continuous functions. *Discrete Contin. Dyn. Syst.* **27** (2010), 383–388.
[35] W. Parry. Intrinsic Markov chains. *Trans. Amer. Math. Soc.* **112** (1964), 55–65.
[36] K. R. Parthasarathy. On the category of ergodic measures. *Illinois J. Math.* **5** (1961), 648–656.
[37] M. Shinoda. Uncountably many maximizing measures for a dense set of continuous functions. *Nonlinearity* **31** (2018), 2192–2200.
[38] M. Shinoda and H. Takahasi. Lyapunov optimization for non-generic one-dimensional expanding Markov maps. *Ergod. Th. & Dynam. Sys.* **40** (2020), 2571–2592.
[39] M. Shinoda and K. Yamamoto. Density of periodic measures and large deviation principle for generalised mod one transformations. *Nonlinearity* **37** (2023), 025003.
[40] K. Sigmund. Generic properties of invariant measures for Axiom A diffeomorphisms. *Invent. Math.* **11** (1970), 99–109.
[41] K. Sigmund. On the distribution of periodic points for β-shifts. *Monatsh. Math.* **82** (1976), 247–252.
[42] K. Sigmund. On the connectedness of ergodic systems. *Manuscripta Math.* **22** (1977), 27–32.
[43] P. Walters. *An Introduction to Ergodic Theory (Graduate Texts in Mathematics, 79)*. Springer, New York, 1982.
[44] K. Yamamoto. On the density of periodic measures for piecewise monotonic maps and their coding spaces. *Tsukuba J. Math.* **44** (2020), 309–324.
[45] G. Yuan and B. Hunt. Optimal orbits of hyperbolic systems. *Nonlinearity* **12** (1999), 1207–1224.

https://doi.org/10.1017/etds.2025.10206 Published online by Cambridge University Press