LARGE DEVIATION PRINCIPLE BY PERIODIC ORBITS FOR SUBSHIFTS BEYOND SPECIFICATION

KENICHIRO YAMAMOTO

Abstract. We study a level-2 large deviation principle (LDP) by periodic orbits for subshifts beyond the classical specification framework used by Kifer. We provide two criteria for periodic-orbit LDP beyond specification. The first is formulated in terms of the periodic almost specification property introduced by Thompson. The second is based on countable Markov extensions and entropy-dense liftable invariant measures. The first criterion applies to several classes of subshifts beyond specification, including β-shifts, S-gap shifts, and the non-intrinsically ergodic examples of Kwietniak–Oprocha–Rams. The second applies to certain classes of symbolic coding spaces of piecewise monotonic interval maps, including those associated with (α, β)-transformations and generalized β-transformations.

Keywords: large deviation principle; periodic orbits; subshifts; periodic almost specification property; countable Markov extensions; symbolic dynamics; β-shifts; S-gap shifts; non-intrinsically ergodic subshifts; (α, β)-shifts; generalized β-shifts.

1. Introduction

Large deviation theory provides a quantitative description of the exponential rarity of atypical statistical behavior of orbits in dynamical systems. We briefly recall the general notion of a large deviation principle. Let X be a compact metrizable space and $\mathcal{M}(X)$ be the set of Borel probability measures on X with the weak-topology. We say that a sequence of Borel probability measures $\{\Omega_n\}_{n=1}^{\infty}$ on $\mathcal{M}(X)$ satisfies a *level-2 large deviation principle with rate function* $q : \mathcal{M}(X) \to [-\infty, 0]$ if q is upper semicontinuous, for every closed subset $\mathcal{K} \subset \mathcal{M}(X)$ we have

$$\limsup_{n\to\infty} \frac{1}{n} \log \Omega_n(\mathcal{K}) \ \leq \ \sup_{\mu\in\mathcal{K}} q(\mu),$$

and if for every open subset $\mathcal{G} \subset \mathcal{M}(X)$ we have

$$\liminf_{n\to\infty} \frac{1}{n} \log \Omega_n(\mathcal{G}) \ \geq \ \sup_{\mu\in\mathcal{G}} q(\mu).$$

For expansive systems satisfying the specification property, level-2 LDPs for Birkhoff averages were established by Takahashi [34] and Young [42].

2020 *Mathematics Subject Classification.* Primary 37A50, 37B10, 37E05; Secondary 37C30, 60F10.

These results show that the fluctuations of orbit averages with respect to the equilibrium measure satisfy an LDP with the thermodynamic rate function

$$q(\mu) = h(\mu) + \int \varphi \, d\mu - P(\varphi),$$

where $h(\mu)$ denotes the metric entropy of the invariant probability measure μ, and $P(\varphi)$ is the topological pressure of the continuous observable φ. This formula reveals a deep connection between large deviations and the variational principle for pressure.

Shortly afterward, Kifer [21] proved a periodic-orbit formulation of the LDP under the classical specification property introduced by Bowen [2]. This viewpoint shows that the thermodynamic behavior of a system can often be described in terms of its finite periodic structures.

Recently, this periodic-orbit perspective has been broadened to various symbolic and smooth contexts; see Pollicott–Sridharan [28], Gelfert–Wolf [13], Comman-Rivera-Letelier [8], and Takahasi [36]. Nevertheless, beyond the specification setting, the existing general theory remains rather limited. A canonical example is the β-shift introduced by Rényi [29], which fails specification but still exhibits strong statistical regularity. However, even for such systems, including the β-shift, a general large deviation principle by periodic orbits has not yet been established.

Motivated by these developments, we first work in the symbolic setting, focusing on subshifts over finite alphabets. Building on this line of research, we establish large deviation principles for two broad classes of subshifts that extend Kifer's periodic-orbit LDP framework beyond systems with the specification property. In particular, these classes include not only β-shifts but also a wider range of subshifts lacking the specification property. The two classes are characterized by different structural features.

The first class of subshifts relies on the periodic almost specification property, a natural weakening of specification introduced by Thompson [38], which still retains sufficient orbit-gluing structure to support thermodynamic analysis. This allows us to treat systems such as β-shifts, topologically mixing S-gap shifts [24], and examples of Kwietniak–Oprocha–Rams [23] that lack intrinsic ergodicity.

The second class of subshifts considered in this paper consists of systems admitting a topologically mixing countable Markov extension with entropy-dense liftable measures; that is, every invariant measure can be approximated, both weakly and in entropy, by invariant measures admitting lifts to the Markov extension (see subsections 2.1 and 2.3 for the precise definitions). Important examples arise from the symbolic coding spaces of piecewise monotonic interval maps, including not only β-shifts but also their natural generalizations, namely (α, β)-shifts and generalized β-shifts, which are the coding spaces of (α, β)-transformations [26] and generalized

β-transformations [14], respectively. Thus, our results apply not only to subshifts studied in their own right but also to symbolic coding spaces arising from certain classes of piecewise monotonic interval maps.

To formulate our result, we introduce the following terminology. For a subshift X over a finite alphabet and a continuous potential $\varphi\colon X \to \mathbb{R}$, we say that the pair (X, φ) satisfies a *level-2 large deviation principle by periodic orbits* if the associated sequence of probability measures

$$\Omega_n^{\varphi} = \frac{1}{\sum_{x \in P_n(X)} e^{S_n\varphi(x)}} \sum_{x \in P_n(X)} e^{S_n\varphi(x)}\, \delta_{\mathcal{E}_n(x)},$$

on $\mathcal{M}(X)$ satisfies the level-2 LDP stated above. Here $P_n(X)$ denotes the set of periodic points of period n, $S_n\varphi(x) = \sum_{k=0}^{n-1} \varphi(\sigma^k x)$ is the n-th Birkhoff sum, $\mathcal{E}_n(x) = \frac{1}{n}\sum_{j=0}^{n-1} \delta_{\sigma^j x}$ is the empirical measure associated with x, σ is the left shift map, and δ_y is the Dirac measure at y. The measure Ω_n^{φ} thus represents the distribution of periodic points weighted by the potential φ. Note that earlier results mostly established this principle for Hölder potentials, whereas we show that it holds for every continuous potential.

We now state our two main theorems; see Section 2 for the necessary definitions.

Theorem A. *Let X be a subshift over a finite alphabet. If X satisfies the periodic almost specification property, then for every continuous potential $\varphi\colon X \to \mathbb{R}$, (X, φ) satisfies a level-2 large deviation principle by periodic orbits with rate function $q^{\varphi}\colon \mathcal{M}(X) \to [-\infty, 0]$ expressed by*

$$q^{\varphi}(\mu) = \begin{cases} h(\mu) + \int \varphi\, d\mu - P(\varphi) & (\mu \text{ is invariant}), \\ -\infty & (\text{otherwise}). \end{cases} \tag{1.1}$$

This applies to β-shifts, topologically mixing S-gap shifts, and examples of Kwietniak-Oprocha–Rams lacking intrinsic ergodicity.

Theorem B. *Let X be a subshift over a finite alphabet. Suppose that the following conditions hold:*

- (I) *There exist a topologically mixing countable Markov shift Σ and a countable-to-one continuous surjection $\pi\colon \Sigma \to X$ such that $\pi \circ \sigma = \sigma \circ \pi$.*
- (II) *The set of liftable measures on X via π is entropy-dense in the set of invariant probability measures on X.*

Then for every continuous potential $\varphi\colon X \to \mathbb{R}$, (X, φ) satisfies a level-2 large deviation principle by periodic orbits with rate function $q^{\varphi}\colon \mathcal{M}(X) \to [-\infty, 0]$ given by (1.1).

In this paper, we apply Theorem B to coding spaces arising from piecewise monotonic interval maps, including β-shifts, (α, β)-shifts, and generalized β-shifts. However, condition (I) is relatively mild: every one-sided subshift admits a countable-to-one extension by a countable Markov shift via the Markov follower set construction ([16]). Moreover, approximation

properties of the type described in condition (II) have been studied in settings beyond piecewise monotonic interval maps ([20]). We therefore expect that Theorem B may also be applicable to classes of subshifts beyond the coding spaces considered in this paper.

Remark 1.1. Level–2 large deviation principles for Birkhoff averages beyond the specification setting have been established in several works (e.g. [5, 7, 27, 31]). These results, however, do not imply the periodic–orbit LDP established in this paper, as the two formulations are of a different nature. For instance, a system may satisfy a Birkhoff–average LDP while the periodic–orbit LDP fails (see Remark 4.2).

Most existing results on the periodic–orbit LDP rely on a functional approach and typically assume that the potential admits a unique equilibrium state. In contrast, our topological approach applies not only beyond the specification property but also to arbitrary continuous potentials, possibly allowing multiple equilibrium states.

Remark 1.2. A preliminary version of part of this work appeared in an earlier non-refereed note by the present author [41], where the periodic-orbit LDP was established only for β-shifts. The present paper extends that non-refereed note by developing two general approaches to large deviation principles by periodic orbits (Theorems A and B) and applying them to a broader class of dynamical systems, including S-gap shifts, (α, β)-shifts, generalized β-shifts, and non-intrinsically ergodic subshifts in the sense of Kwietniak–Oprocha–Rams. These examples illustrate that the present results extend large deviation principles to systems beyond the specification setting.

The remainder of this paper is organized as follows. Section 2 recalls fundamental definitions and known results from symbolic dynamics and the theory of piecewise monotonic interval maps, including the notions of periodic almost specification, liftability via countable Markov extensions, and Hofbauer's Markov diagram construction. Section 3 provides the proofs of the two main theorems. Subsection 3.1 establishes the level-2 large deviation principle by periodic orbits for subshifts with the periodic almost specification property, thereby proving Theorem A. Subsection 3.2 gives the proof of Theorem B. Section 4 discusses concrete applications of these results, including β-shifts, S-gap shifts, non-intrinsically ergodic examples of Kwietniak–Oprocha–Rams, and the symbolic coding spaces of (α, β)- and generalized β-transformations.

2. Definitions and known facts

2.1. **Symbolic dynamics.** Let $\mathbb{N}_0$ denote the set of nonnegative integers and $\mathbb{Z}$ the set of all integers. For a finite or countable alphabet A, we write $A^{\mathbb{N}_0}$ and $A^{\mathbb{Z}}$ for the one-sided and two-sided full shifts over A, respectively, each endowed with the product topology. Here the alphabet A is equipped

with the discrete topology. The (left) shift map σ is defined by $(\sigma(x))_i = x_{i+1}$ for $x = (x_i)_{i\geq 0}$ and $i \geq 0$ in the one-sided case, and for $x = (x_i)_{i\in\mathbb{Z}}$ and $i \in \mathbb{Z}$ in the two-sided case. A closed, shift-invariant subset $\Sigma \subset A^{\mathbb{N}_0}$ (resp. $\Sigma \subset A^{\mathbb{Z}}$) is called a one-sided (resp. two-sided) subshift. Since the two-sided case can be handled analogously up to notational modifications, we shall primarily work with one-sided subshifts in what follows. Unless otherwise stated, all definitions, propositions, lemmas, and proofs are given for the one-sided case, and the two-sided versions are completely analogous.

A subshift Σ is called a *Markov shift* (or *topological Markov shift*) if there exists a matrix $M = (M_{a,b})_{(a,b)\in A^2}$ with entries in $\{0, 1\}$ such that

$$\Sigma = \{(x_i)_{i\geq 0} \in A^{\mathbb{N}_0} : M_{x_i,x_{i+1}} = 1, \quad \text{for all } i \geq 0\}.$$

In this case, we also say that Σ is a Markov shift with an *adjacency matrix* M.

For a subshift Σ, its *language* is defined by

$$\mathcal{L}(\Sigma) := \{x_0x_1 \cdots x_{n-1} : (x_i)_i \in \Sigma,\ n \geq 1\},$$

and for each $w \in \mathcal{L}(\Sigma)$, the associated *cylinder* is

$$[w] := \{x \in \Sigma : x_0 \cdots x_{|w|-1} = w\},$$

where $|w|$ denotes the length of w. For $n \geq 1$, we write

$$\mathcal{L}_n(\Sigma) := \{w \in \mathcal{L}(\Sigma) : |w| = n\}$$

for the set of all admissible words of length n. For $u, v \in \mathcal{L}(\Sigma)$, we use the concatenation uv to denote the word obtained by joining them. We say that Σ is *topologically transitive* if for any $u, v \in \mathcal{L}(\Sigma)$ there exists $w \in \mathcal{L}(\Sigma)$ such that $uwv \in \mathcal{L}(\Sigma)$. We say that Σ is *topologically mixing* if for any $u, v \in \mathcal{L}(\Sigma)$, there exists an integer $N \geq 1$ such that for any $n \geq N$, there exists $w \in \mathcal{L}(\Sigma)$ with $|w| = n$ such that $uwv \in \mathcal{L}(\Sigma)$.

We denote by $\mathcal{M}(\Sigma)$ the set of all Borel probability measures on Σ endowed with the weak topology, by $\mathcal{M}_\sigma(\Sigma) \subset \mathcal{M}(\Sigma)$ the set of σ-invariant ones, and by $\mathcal{M}^e_\sigma(\Sigma) \subset \mathcal{M}_\sigma(\Sigma)$ the set of ergodic ones. Throughout this paper, we denote by $h(\mu)$ the metric entropy of $\mu \in \mathcal{M}_\sigma(\Sigma)$. The following notion will play an important role throughout the paper and appears in the condition (II) stated in Theorem B. Let $\mathcal{M} \subset \mathcal{M}_\sigma(\Sigma)$. We say that $\mathcal{M}$ is entropy-dense in $M_\sigma(\Sigma)$ if for every $\mu \in \mathcal{M}_\sigma(\Sigma)$, every neighborhood $\mathcal{G}$ of μ in $\mathcal{M}_\sigma(\Sigma)$, and every $\varepsilon > 0$, there exists $\nu \in \mathcal{M} \cap \mathcal{G}$ such that

$$h(\nu) \geq h(\mu) - \varepsilon.$$

Hereafter in this subsection, we restrict our attention to subshifts X over a finite alphabet A.

For $n \geq 1$, let

$$P_n(X) := \{x \in X : \sigma^n x = x\}$$

be the set of all periodic points of period n. For $x \in X$, we define its *empirical measure* by

$$\mathcal{E}_n(x) := \frac{1}{n}\sum_{j=0}^{n-1} \delta_{\sigma^j x},$$

where δ_y denotes the *Dirac measure* at y. For convenience, throughout this paper we regard $\mathcal{E}_n$ as a map from X to $\mathcal{M}(X)$, and use the notation

$$\mathcal{E}_n^{-1}(\mathcal{A}) := \{x \in X : \mathcal{E}_n(x) \in \mathcal{A}\}$$

for $n \geq 1$ and $\mathcal{A} \subset \mathcal{M}(X)$. If $x \in P_n(X)$ for some $n \geq 1$, then $\mathcal{E}_n(x)$ is called the *periodic measure*, and we denote by $\mathcal{M}_\sigma^p(X)$ the set of all such measures.

Denote by $C(X)$ the set of real-valued continuous functions on X endowed with the supremum norm $\|\varphi\| := \sup\{|\varphi(x)| : x \in X\}$. Fix a countable dense subset $\{\varphi_n\}_{n\geq 1} \subset C(X)$ such that $\|\varphi_n\| \neq 0$ for every $n \geq 1$, and define a compatible metric D on $\mathcal{M}(X)$ by

$$D(\mu,\nu) := \sum_{n=1}^{\infty} \frac{1}{2^{n+1}\,\|\varphi_n\|} \left| \int \varphi_n \, d\mu - \int \varphi_n \, d\nu \right|.$$

This metric generates the weak topology on $\mathcal{M}(X)$, and $(\mathcal{M}(X), D)$ is compact since X is a subshift over a finite alphabet.

For $\varphi \in C(X)$ and $n \geq 1$, we write

$$S_n\varphi(x) := \sum_{j=0}^{n-1} \varphi(\sigma^j x).$$

The *topological pressure* for $\varphi \in C(X)$ is defined by

$$P(\varphi) := \limsup_{n\to\infty} \frac{1}{n} \log \sum_{w\in\mathcal{L}_n(X)} \exp\left(\sup_{x\in[w]} S_n\varphi(x) \right). \tag{2.1}$$

The topological pressure satisfies the usual *variational principle*:

$$P(\varphi) = \sup_{\mu\in\mathcal{M}_\sigma(X)} \left(h(\mu) + \int \varphi \, d\mu \right). \tag{2.2}$$

An invariant measure $\mu_\varphi \in \mathcal{M}_\sigma(X)$ is called an *equilibrium state* for φ if it attains the above supremum. In particular, when $\varphi \equiv 0$, the pressure reduces to the *topological entropy*, $P(0) = h_{\mathrm{top}}(X)$, and the equilibrium states coincide with the *measures of maximal entropy*.

2.2. **Almost specification property.** In this subsection we recall Bowen's classical specification property, which forms the foundation of many results in thermodynamic formalism. We then employ weakened versions of the specification property—specifically, the almost specification property introduced by Thompson and its periodic analogue—which play a central role in the proof of Theorem A. Throughout this subsection, let X be a subshift

over a finite alphabet. The classical specification property was introduced by Bowen [2], and we recall the definition below.

Definition 2.1. We say that X satisfies the *specification property* if there exists an integer $N \geq 0$ such that for any finite collection of words $w^{(1)}$, $w^{(2)}, \ldots, w^{(k)} \in \mathcal{L}(X)$, there exist words $v^{(1)}, v^{(2)}, \cdots, v^{(k-1)} \in \mathcal{L}_N(X)$ such that

$$v = w^{(1)}v^{(1)}w^{(2)}v^{(2)} \cdots w^{(k-1)}v^{(k-1)}w^{(k)} \in \mathcal{L}(X).$$

If the cylinder set $[v]$ contains a periodic point with period $|v| + N$, then we say that X satisfies the *periodic specification property*.

We say that a subshift X is *sofic* if X is a topological factor of a Markov shift over a finite alphabet; that is, there exists a continuous surjection $\pi \colon Y \to X$ with $\sigma \circ \pi = \pi \circ \sigma$, where Y is such a Markov shift. It is known that every topologically mixing sofic shift satisfies the periodic specification property (see, for example, [10, Propositions (21.2) and (21.4)]). Subsequently, Thompson [38] introduced the notion of the almost specification property. This notion generalizes the classical specification property by allowing a vanishing proportion of "mistakes." To quantify such deviations, we use the Hamming distance between finite words: for two words $u = u_0u_1 \cdots u_{n-1}$ and $v = v_0v_1 \cdots v_{n-1}$ of the same length n, let

$$d_H(u, v) := \#\{\, 0 \leq i < n : u_i \neq v_i \,\}.$$

We now recall the definition of the almost specification property.

Definition 2.2. We say that X satisfies the *almost specification property* if there exists a mistake function $g : \mathbb{N} \to \mathbb{N}$ such that $g(n)/n \to 0$ as $n \to \infty$ and the following holds: for any finite collection of words $w^{(1)}, w^{(2)}, \ldots, w^{(k)} \in \mathcal{L}(X)$, there exist words $v^{(1)}, v^{(2)}, \ldots, v^{(k)} \in \mathcal{L}(X)$ such that

- $|w^{(i)}| = |v^{(i)}|$ for every i;
- $d_H(w^{(i)}, v^{(i)}) \leq g(|w^{(i)}|)$ for all i; and
- the concatenation $v = v^{(1)}v^{(2)} \cdots v^{(k)}$ belongs to $\mathcal{L}(X)$.

If the cylinder set $[v]$ contains a periodic point with period $|v|$, then we say that X satisfies the *periodic almost specification property*.

It is straightforward to see that the (resp. periodic) specification property implies the (resp. periodic) almost specification property.

In the remainder of this subsection, we recall several known results that will be needed for the proof of Theorem A.

Proposition 2.3. ([27, Proposition 2.1]). Let $\mu \in \mathcal{M}^e_\sigma(X)$ and $\epsilon > 0$. Then there exists $\delta > 0$ and an integer N such that for every neighborhood $\mathcal{G} \subset \mathcal{M}(X)$ of μ and every $n \geq N$, there exists a finite subset $\Gamma = \Gamma_n \subset \mathcal{E}_n^{-1}(\mathcal{G})$ satisfying:

- $\#\Gamma \geq e^{n(h(\mu)-\epsilon)}$;
- for any distinct $(x_i)_i, (y_i)_i \in \Gamma$, $d_H(x_0x_1 \cdots x_{n-1}, y_0y_1 \cdots y_{n-1}) > \delta n$.

The following proposition is derived from a small modification of [40, Theorem 1.2].

Theorem 2.4. ([40, Theorem 1.2]). Suppose that X satisfies the periodic almost specification property. Then for any $\varphi \in C(X)$,

$$P(\varphi) = \lim_{n\to\infty} \frac{1}{n} \log \sum_{x\in P_n(X)} e^{S_n\varphi(x)}$$

holds.

Theorem 2.5. ([27, Theorem 2.1]). Suppose that X satisfies the almost specification property. Then $\mathcal{M}^e_\sigma(X)$ is entropy-dense in $\mathcal{M}_\sigma(X)$.

2.3. **Countable Markov shifts and liftability.** In this subsection, we recall several properties of countable Markov shifts and introduce liftable measures. We then establish an entropy-density property that will play a key role in the proof of Theorem B. Let Σ be a Markov shift over a countable alphabet A. For a subset $F \subset A$, define $\Sigma_F := \Sigma \cap F^{\mathbb{N}_0}$. If F is finite, then Σ_F is a subshift of finite type.

The following theorem due to Takahasi provides an entropy approximation property for countable Markov shifts.

Theorem 2.6. ([35, Main Theorem]) Let Σ be a topologically transitive Markov shift over a countable alphabet A. Then for every $\varepsilon > 0$, every invariant probability measure $\bar{\mu} \in \mathcal{M}_\sigma(\Sigma)$ with finite entropy, and every neighborhood $\mathcal{G}$ of $\bar{\mu}$ in $\mathcal{M}(\Sigma)$, there exist a finite set $F \subset A$ and an ergodic measure $\bar{\nu} \in \mathcal{G}$ such that $h(\bar{\nu}) \geq h(\bar{\mu}) - \varepsilon$ and $\bar{\nu}(\Sigma_F) = 1$.

In particular, $\mathcal{M}^e_\sigma(\Sigma)$ is entropy-dense in $\mathcal{M}_\sigma(\Sigma)$. Moreover, the approximating ergodic measures can be chosen to be supported on finite-alphabet subshifts of Σ.

Let X be a subshift over a finite alphabet and let $\pi : \Sigma \to X$ be a factor map. An invariant probability measure $\mu \in \mathcal{M}_\sigma(X)$ is called *liftable (via π)* if there exists an invariant probability measure $\bar{\mu} \in \mathcal{M}_\sigma(\Sigma)$ satisfying $\mu = \bar{\mu} \circ \pi^{-1}$. In such case, we say that $\bar{\mu}$ is a *lift* of μ.

Under the conditions (I) and (II) stated in Theorem B, ergodic measures are entropy-dense among invariant measures.

Proposition 2.7. Let X be a subshift on a finite alphabet and suppose that the conditions (I) and (II) stated in Theorem B hold. Then $\mathcal{M}^e_\sigma(X)$ is entropy-dense in $\mathcal{M}_\sigma(X)$.

Proof. Let $\mu \in \mathcal{M}_\sigma(X)$, let $\mathcal{G}$ be a neighborhood of μ, and let $\varepsilon > 0$. By the conditions (I) and (II) stated in Theorem B, there exist a topologically mixing countable Markov shift Σ and a countable-to-one factor map $\pi\colon \Sigma \to X$. Moreover, the set of liftable measures is entropy-dense in $\mathcal{M}_\sigma(X)$.

By entropy-density of liftable measures, there exists a liftable measure $\nu \in \mathcal{G}$ such that $h(\nu) \geq h(\mu) - \varepsilon$.

Choose a lift $\bar{\nu} \in \mathcal{M}_\sigma(\Sigma)$ satisfying $\nu = \bar{\nu} \circ \pi^{-1}$. Since π is countable-to-one, we have $h(\bar{\nu}) = h(\nu)$. Applying Theorem 2.6, we obtain an ergodic measure $\bar{\rho} \in \mathcal{M}_\sigma(\Sigma)$ such that $h(\bar{\rho}) \geq h(\bar{\nu}) - \varepsilon$, and $\bar{\rho}$ belongs to a sufficiently small neighborhood of $\bar{\nu}$.

Define $\rho := \bar{\rho} \circ \pi^{-1}$. Since factor maps preserve ergodicity, ρ is ergodic. Moreover, $h(\rho) = h(\bar{\rho})$ by entropy preservation of countable-to-one factor maps. Therefore $h(\rho) \geq h(\nu) - \varepsilon \geq h(\mu) - 2\varepsilon$.

Since the map $\bar{\eta} \in \mathcal{M}(\Sigma) \mapsto \bar{\eta} \circ \pi^{-1} \in \mathcal{M}(X)$ is continuous, by choosing the neighborhood of $\bar{\nu}$ sufficiently small we also obtain $\rho \in \mathcal{G}$. This proves the proposition. □

2.4. **Markov diagram for piecewise monotonic maps.** As mentioned in the Introduction, the main applications of Theorem B concern coding spaces associated with piecewise monotonic maps. We therefore recall Hofbauer's Markov diagram construction, which provides a countable-to-one factor map from a countable Markov shift onto the coding space, and establish a sufficient condition ensuring that the coding space satisfies the conditions (I) and (II) stated in Theorem B. We say that $T : [0,1] \to [0,1]$ is a *piecewise monotonic map* if there exist $k \geq 1$ and critical points

$$0 = c_0 < c_1 < \cdots < c_k < c_{k+1} = 1$$

such that $T|_{(c_j, c_{j+1})}$ is strictly monotonic and continuous for each $0 \leq j \leq k$. For such a map T, we define

$$X_T := \bigcap_{n \geq 0} T^{-n} \left(\bigcup_{j=0}^{k} (c_j, c_{j+1}) \right),$$

and the *coding map* $\mathcal{I} : X_T \to \{0, \ldots, k\}^{\mathbb{N}_0}$ by

$$(\mathcal{I}(x))_n = j \quad \text{iff} \quad T^n(x) \in (c_j, c_{j+1}).$$

We denote by Σ_T the closure of $\mathcal{I}(X_T)$ in $\{0, \ldots, k\}^{\mathbb{N}_0}$, and call Σ_T the *coding space* of T.

Following Hofbauer, we recall the construction of the *Markov diagram*. Let $C \subset \Sigma_T$ be a closed subset contained in the cylinder $[j]$ for some $0 \leq j \leq k$. A non-empty closed subset $D \subset \Sigma_T$ is said to be a *successor* of C if $D = [l] \cap \sigma(C)$ for some $0 \leq l \leq k$. We write $C \to D$ when D is a successor of C.

Define inductively a collection of vertices:

$$\mathcal{D}_0 := \{[0], \ldots, [k]\}, \quad \mathcal{D}_{n+1} := \mathcal{D}_n \cup \{D : C \to D \text{ for some } C \in \mathcal{D}_n\},$$

and let $\mathcal{D}_T := \bigcup_{n \geq 0} \mathcal{D}_n$. Since each $\mathcal{D}_n$ is finite, $\mathcal{D}_T$ is at most countable. The oriented graph $(\mathcal{D}_T, \to)$ is called the *Markov diagram* of T. We write $\mathcal{D}$ for $\mathcal{D}_T$ when no confusion arises.

We define the adjacency matrix $M(\mathcal{D}) = (M_{C,D})_{(C,D)\in\mathcal{D}^2}$ by

$$M_{C,D} = \begin{cases} 1 & (C \to D), \\ 0 & (\text{otherwise}). \end{cases}$$

Then

$$\Sigma_{\mathcal{D}} := \{(C_i)_{i\geq 0} \in \mathcal{D}^{\mathbb{N}_0} : C_i \to C_{i+1},\ i \in \mathbb{N}_0\}$$

is a countable-state Markov shift with an adjacency matrix $M(\mathcal{D})$. We say that $\mathcal{D}$ is *irreducible* if for any $C, D \in \mathcal{D}$, there exist an integer $n \geq 1$ and $C_0, \cdots, C_n \in \mathcal{D}$ such that $C_0 = C$, $C_n = D$ and $C_i \to C_{i+1}$ holds for $0 \leq i \leq n-1$. In this case, $\Sigma_{\mathcal{D}}$ is topologically transitive. If, in addition, there exists $C \in \mathcal{D}$ such that $C \to C$, then $\Sigma_{\mathcal{D}}$ is topologically mixing.

Define a projection

$$\Psi : \Sigma_{\mathcal{D}} \to \{0, \ldots, k\}^{\mathbb{N}_0}, \qquad \Psi((C_i)_i) = (x_i)_i,$$

where x_i is the unique integer such that $C_i \subset [x_i]$. Then it is straightforward to see that Ψ is continuous, countable-to-one, and satisfies $\Psi(\Sigma_{\mathcal{D}}) = \Sigma_T$ and $\sigma \circ \Psi = \Psi \circ \sigma$.

Although Ψ is surjective, not every invariant measure is liftable to $\Sigma_{\mathcal{D}}$ in general. Hofbauer gave a sufficient condition for liftability, which we recall here.

Lemma 2.8. ([15, Lemma 3]). Let $\mu \in \mathcal{M}^e_\sigma(\Sigma_T)$. If μ has positive metric entropy or if μ is a periodic measure, then μ is liftable via Ψ.

The following proposition provides a convenient criterion for verifying the conditions (I) and (II) stated in Theorem B in the setting of piecewise monotonic maps.

Proposition 2.9. Let $T\colon [0,1] \to [0,1]$ with critical points $0 = c_0 < c_1 < \cdots < c_k < c_{k+1}$, Σ_T, $\Sigma_{\mathcal{D}}$ and Ψ be as above. Assume the following conditions:

- **Exactness (EX).** For any non-empty open interval $I \subset [0,1]$, there exist a positive integer τ, integers $1 \leq i_0, \ldots, i_{\tau-1} \leq k$, and an open subinterval $L \subset I$ such that $T^j(L) \subset (c_{i_j-1}, c_{i_j})$ for all $0 \leq j \leq \tau-1$ and $T^\tau(L) = (0,1)$.
- **Density of Periodic measures (DP).** $\mathcal{M}^p_\sigma(\Sigma_T)$ is dense in $\mathcal{M}^e_\sigma(\Sigma_T)$.

Then the conditions (I) and (II) stated in Theorem B hold with $(X, \Sigma, \pi) = (\Sigma_T, \Sigma_{\mathcal{D}}, \Psi)$.

Proof. By (EX), it follows from [5, Lemma 4.1] that $\mathcal{D}$ is irreducible. Again by (EX), $T((c_{i_{\tau-1}-1}, c_{i_{\tau-1}})) = (0,1)$, which implies that $[i_{\tau-1}] \to [i_{\tau-1}]$. Hence $\Sigma_{\mathcal{D}}$ is topologically mixing, and the condition (I) holds.

Let $\mu \in \mathcal{M}_\sigma(\Sigma_T)$, $\mathcal{G}$ be a neighborhood of μ, and $\varepsilon > 0$. By the ergodic decomposition theorem and Jacob's theorem, there exists a convex combination of ergodic measures $\rho = \sum_{i=1}^p a_i \mu_i$ such that $\rho \in \mathcal{G}$ and $|h(\mu) - h(\rho)| \leq \varepsilon$.

For $1 \le i \le p$, define an ergodic measure ν_i by $\nu_i = \mu_i$ if $h(\mu_i) > 0$, and choose $\nu_i \in \mathcal{M}_\sigma^p(\Sigma_T)$ if $h(\mu_i) = 0$. Let $\nu = \sum_{i=1}^p a_i \nu_i$. Since every periodic measure has zero entropy, we have $h(\nu_i) = h(\mu_i)$ for every $1 \le i \le p$, and hence $h(\nu) = h(\rho)$.

By (DP), we can choose the periodic measures ν_i so that $\nu \in \mathcal{G}$. By Lemma 2.8, each ν_i is liftable. Let $\bar{\nu}_i$ be a lift of ν_i and set $\bar{\nu} = \sum_{i=1}^p a_i \bar{\nu}_i$. Then $\bar{\nu}$ is a lift of ν, and hence ν is liftable.

Therefore, for every $\mu \in \mathcal{M}_\sigma(\Sigma_T)$, every neighborhood $\mathcal{G}$ of μ, and every $\varepsilon > 0$, there exists a liftable measure $\nu \in \mathcal{G}$ such that $h(\nu) \ge h(\mu) - \varepsilon$. This shows that the set of liftable measures is entropy-dense in $\mathcal{M}_\sigma(\Sigma_T)$, and hence the condition (II) holds. □

3. Proofs

In this section we prove Theorems A and B. The following known results will be used repeatedly in their proofs.

Theorem 3.1. ([13, Theorem 6]). Let X be a subshift on a finite alphabet. Then for any $\varphi \in C(X)$ and any closed subset $\mathcal{K} \subset \mathcal{M}_\sigma(X)$, we have

$$\limsup_{n\to\infty} \frac{1}{n} \log \sum_{x \in P_n(X) \cap \mathcal{E}_n^{-1}(\mathcal{K})} e^{S_n\varphi(x)} \le \sup_{\mu \in \mathcal{K}} \left(h(\mu) + \int \varphi d\mu \right).$$

3.1. **Proof of Theorem A.** In this subsection we present the proof of Theorem A.

Proposition 3.2. Let X be a subshift over a finite alphabet and let $\varphi \in C(X)$. Suppose that X satisfies the periodic almost specification property. Then for any $\mu \in \mathcal{M}_\sigma^e(X)$ and any open neighborhood $\mathcal{G}$ of μ, we have

$$\liminf_{n\to\infty} \frac{1}{n} \log \sum_{x \in P_n(X) \cap \mathcal{E}_n^{-1}(\mathcal{G})} e^{S_n\varphi(x)} \ge h(\mu) + \int \varphi d\mu.$$

Proof. Given any $\epsilon > 0$, choose $\gamma > 0$ sufficiently small so that $\mathbb{B}(\mu, 2\gamma) \subset \mathcal{G}$ and for all $\nu \in \mathbb{B}(\mu, 2\gamma)$,

$$\left| \int \varphi d\mu - \int \varphi d\nu \right| \le \epsilon,$$

where $\mathbb{B}(\mu, 2\gamma) := \{\nu \in \mathcal{M}(X) : D(\mu, \nu) \le 2\gamma\}$.

By Proposition 2.3, there exist $\delta > 0$ and $N \in \mathbb{N}$ such that for any $n \ge N$, there exists a finite subset $\Gamma = \Gamma_n \subset \mathcal{E}^{-1}(\mathbb{B}(\mu, 2\gamma))$ satisfying $\#\Gamma \ge e^{n(h(\mu)-\epsilon)}$ and $d_H(x_0x_1\cdots x_{n-1}, y_0y_1\cdots y_{n-1}) > \delta n$ for any distinct $(x_i)_i,\ (y_i)_i \in \Gamma$.

Let $g\colon \mathbb{N} \to \mathbb{N}$ be a mistake function as in the definition of the periodic almost specification property. By [7, Lemma 2.8], further enlarging N if necessary, for any $n \ge N$, we have $g(n) \le \delta n$, and for any $x = (x_i)_i,\ y = (y_i)_i \in X$, $D(\mathcal{E}_n(x), \mathcal{E}_n(y)) \le \gamma$ and $|S_n\varphi(x) - S_n\varphi(y)| \le \epsilon n$ whenever

$$d_H(x_0x_1\cdots x_{n-1}, y_0y_1\cdots y_{n-1}) \le g(n).$$

By the periodic almost specification property, for each $z = (z_i)_i \in \Gamma$, there exists a periodic point $z^* = (z_i^*)_i \in P_n(X)$ such that

$$d_H(z_0 z_1 \cdots z_{n-1}, z_0^* z_1^* \cdots z_{n-1}^*) \leq g(n). \tag{3.1}$$

Then for each $z \in \Gamma$, we have

$$D(\mu, \mathcal{E}_n(z^*)) \leq D(\mu, \mathcal{E}_n(z)) + D(\mathcal{E}_n(z), \mathcal{E}_n(z^*)) \leq 2\gamma. \tag{3.2}$$

and hence $z^* \in \mathcal{E}^{-1}(\mathbb{B}(\mu, 2\gamma)) \subset \mathcal{E}_n^{-1}(\mathcal{G})$.

By the equations (3.1) and (3.2), and the property of Γ, the map

$$z \in \Gamma \mapsto z^* \in P_n(X) \cap \mathcal{E}_n^{-1}(\mathcal{G})$$

is injective and $S_n\varphi(z^*) \geq S_n\varphi(z) - \epsilon n$ holds for any $z \in \Gamma$. Therefore we obtain

$$\begin{aligned}
\sum_{x \in P_n(X) \cap \mathcal{E}_n^{-1}(\mathcal{G})} e^{S_n\varphi(x)} &\geq \sum_{z \in \Gamma} e^{S_n\varphi(z^*)} \\
&\geq \sum_{z \in \Gamma} e^{S_n\varphi(z) - \epsilon n} \\
&\geq \sum_{z \in \Gamma} e^{n(\int \varphi d\mu - 2\epsilon)} \\
&= \#\Gamma e^{n(\int \varphi d\mu - 2\epsilon)} \\
&\geq e^{n(h(\mu) + \int \varphi d\mu - 3\epsilon)}.
\end{aligned}$$

This completes the proof of Proposition 3.2. □

We now prove Theorem A. Let (X, φ) be as in Theorem A. By Theorems 2.4 and 3.1, we have

$$\begin{aligned}
\limsup_{n\to\infty} \frac{1}{n} \log \Omega_n^{\varphi}(\mathcal{K}) &= \limsup_{n\to\infty} \frac{1}{n} \log \sum_{x \in P_n(X) \cap \mathcal{E}_n^{-1}(\mathcal{K})} e^{S_n\varphi(x)} - P(\varphi) \\
&\leq \sup_{\mu \in \mathcal{K}} \left(h(\mu) + \int \varphi d\mu \right) - P(\varphi) \\
&= \sup_{\mu \in \mathcal{K}} q^{\varphi}(\mu)
\end{aligned}$$

for any closed set $\mathcal{K} \subset \mathcal{M}$.

Let $\mathcal{G} \subset \mathcal{M}(X)$ be an open subset and fix $\mu \in \mathcal{G}$. In what follows, we show that

$$\liminf_{n\to\infty} \frac{1}{n} \log \Omega_n^{\varphi}(\mathcal{G}) \geq q^{\varphi}(\mu). \tag{3.3}$$

If μ is not σ-invariant, then $q^{\varphi}(\mu) = -\infty$ so we may, without loss of generality, assume that $\mu \in \mathcal{M}_{\sigma}(X)$. Given any $\epsilon > 0$, choose a neighborhood $\mathcal{G}^{\dagger} \subset \mathcal{G}$ of μ such that $|\int \varphi d\mu - \int \varphi d\nu| \leq \epsilon$ for any $\nu \in \mathcal{G}^{\dagger}$. Since X satisfies the almost specification property, Theorem 2.5 guarantees the existence of an ergodic measure $\nu \in \mathcal{G}^{\dagger}$ such that $h(\nu) \geq h(\mu) - \epsilon$.

Hence, by Theorem 2.4 and Proposition 3.2,

$$\begin{aligned}\liminf_{n\to\infty}\frac{1}{n}\log\Omega_n^{\varphi}(\mathcal{G}) &= \liminf_{n\to\infty}\frac{1}{n}\log\sum_{x\in P_n(X)\cap\mathcal{E}_n^{-1}(\mathcal{G})} e^{S_n\varphi(x)}-P(\varphi)\\ &\geq h(\nu)+\int\varphi d\nu-P(\varphi)\\ &\geq h(\mu)+\int\varphi d\mu-P(\varphi)-2\epsilon.\end{aligned}$$

Since $\epsilon>0$ is arbitrary, inequality (3.3) follows, completing the proof of Theorem A.

3.2. **Proof of Theorem B.** In this subsection we prove Theorem B. We begin with the following proposition.

Proposition 3.3. Let X be a subshift on a finite alphabet and suppose that the conditions (I) and (II) stated in Theorem B hold. Let $\varphi\in C(X)$. Then for any $\epsilon>0$, any $\mu\in\mathcal{M}_\sigma^e(X)$, and any neighborhood $\mathcal{G}$ of μ, there exists a subshift $Y\subset X$ such that

- Y satisfies the periodic specification property;
- $P(\varphi|_Y)\geq P(\varphi)-3\epsilon$, where $\varphi|_Y$ denotes the restriction of φ with respect to Y;
- there exists an ergodic measure $\nu\in\mathcal{G}$ such that $h(\nu)\geq h(\mu)-2\epsilon$ and $\nu(Y)=1$.

Proof. By the condition (I) stated in Theorem B, there exist a topologically mixing countable Markov shift Σ over a countable alphabet A, and a countable-to-one factor map $\pi\colon\Sigma\to X$. Since σ is expansive, there is an ergodic equilibrium state m such that $P(\varphi)=h(m)+\int\varphi dm$. Choose $\delta>0$ so small that $D(m,\rho)\leq 2\delta$ implies $|\int\varphi dm-\int\varphi d\rho|\leq\epsilon$. By the condition (II), we can find $m^\diamond\in\mathcal{M}_\sigma(X)$ such that $m^\diamond$ is liftable, $D(m,m^\diamond)\leq\delta$ and $h(m^\diamond)\geq h(m)-\epsilon$. Since $m^\diamond$ is liftable, there is a $\bar{m^\diamond}\in\mathcal{M}_\sigma^e(\Sigma)$ such that $m^\diamond=\bar{m^\diamond}\circ\pi^{-1}$. Since Σ is a topologically mixing (and hence transitive) countable Markov shift, it follows from Theorem 2.6 that there exist a finite set $F_1\subset A$ and an ergodic measure $\bar{m^*}$ on Σ_{F_1} such that denoting $m^*:=\bar{m^*}\circ\pi^{-1}$, we have $D(m^\diamond,m^*)\leq\delta$ (hence $D(m,m^*)\leq 2\delta$), $h(m^*)\geq h(m^\diamond)-\epsilon=h(m)-2\epsilon$ and $m^*(\pi(\Sigma_{F_1}))=1$.

Similarly, we may find a finite subset $F_2\subset A$ and an ergodic $\nu\in\mathcal{G}$ such that $h(\nu)\geq h(\mu)-2\epsilon$ and $\nu(\pi(\Sigma_{F_2}))=1$. Since Σ is topologically mixing and both F_1 and F_2 are finite, there exists a finite subset $F\subset A$ which contains $F_1\cup F_2$ such that Σ_F is a topologically mixing subshift of finite type. Then $Y:=\pi(\Sigma_F)$ is a topologically mixing sofic shift, which implies that Y satisfies the periodic specification property. Since $m^*(Y)\geq m^*(\pi(\Sigma_{F_1}))=1$, we have $m^*(Y)=1$. Hence the classical Variational Principle (2.2) implies that

$$P(\varphi|_Y)\geq h(m^*)+\int\varphi dm^*\geq h(m)+\int\varphi dm-3\epsilon=P(\varphi)-3\epsilon.$$

Similarly, we have $\nu(Y) = 1$, which proves the proposition. □

We note that Propositions 3.4 and 3.5 are analogous to Theorem 2.4 and Proposition 3.2, which were used in the proof of Theorem A. Here, they serve as the corresponding ingredients for Theorem B.

Proposition 3.4. Let X be a subshift on a finite alphabet and suppose that the conditions (I) and (II) stated in Theorem B hold. Then for any $\varphi \in C(X)$, we have

$$P(\varphi) = \lim_{n\to\infty} \frac{1}{n} \log \sum_{x\in P_n(X)} e^{S_n\varphi(x)}.$$

Proof. The inequality $\limsup_{n\to\infty} \frac{1}{n} \log \sum_{x\in P_n(X)} e^{S_n\varphi(x)} \leq P(\varphi)$ easily follows from the definition (2.1) of $P(\varphi)$.

On the other hand, by Proposition 3.3 for any $\epsilon > 0$, we can find a subshift $Y \subset X$ such that Y satisfies the periodic specification property and $P(\varphi|_Y) \geq P(\varphi) - 3\epsilon$. Since Y satisfies the periodic specification property, by Proposition 2.4, we have

$$P(\varphi|_Y) = \lim_{n\to\infty} \frac{1}{n} \log \sum_{x\in P_n(Y)} e^{S_n\varphi(x)}.$$

Hence

$$\liminf_{n\to\infty} \frac{1}{n} \log \sum_{x\in P_n(X)} e^{S_n\varphi(x)} \geq \lim_{n\to\infty} \frac{1}{n} \log \sum_{x\in P_n(Y)} e^{S_n\varphi(x)} = P(\varphi|_Y) \geq P(\varphi) - 3\epsilon,$$

which proves the proposition. □

Proposition 3.5. Let X be a subshift on a finite alphabet and suppose that the conditions (I) and (II) stated in Theorem B hold. Then for any $\mu \in \mathcal{M}^e_\sigma(X)$ and any open neighborhood $\mathcal{G} \subset \mathcal{M}(X)$ of μ, we have

$$\liminf_{n\to\infty} \frac{1}{n} \log \sum_{x\in P_n(X)\cap \mathcal{E}_n^{-1}(\mathcal{G})} e^{S_n\varphi(x)} \geq h(\mu) + \int \varphi d\mu.$$

Proof. By Proposition 3.3, there exist a subshift $Y \subset X$ and an ergodic measure $\nu \in \mathcal{G}$ such that Y satisfies the periodic specification property, $P(\varphi|_Y) \geq P(\varphi) - 3\epsilon$, $h(\nu) \geq h(\mu) - 2\epsilon$ and $\nu(Y) = 1$. Since Y satisfies the periodic specification property, Proposition 3.2 implies that

$$\liminf_{n\to\infty} \frac{1}{n} \log \sum_{x\in P_n(Y)\cap \mathcal{E}_n^{-1}(\mathcal{G})} e^{S_n\varphi(x)} \geq h(\nu) + \int \varphi d\nu.$$

Hence we have

$$\begin{aligned}\liminf_{n\to\infty}\frac{1}{n}\log\sum_{x\in P_n(X)\cap\mathcal{E}_n^{-1}(\mathcal{G})}e^{S_n\varphi(x)} &\geq \liminf_{n\to\infty}\frac{1}{n}\log\sum_{x\in P_n(Y)\cap\mathcal{E}_n^{-1}(\mathcal{G})}e^{S_n\varphi(x)}\\ &\geq h(\nu)+\int\varphi d\nu\\ &\geq h(\mu)+\int\varphi d\mu-3\epsilon,\end{aligned}$$

which proves the proposition. □

Now we give a proof of Theorem B. Let X be a subshift on a finite alphabet and suppose that the conditions (I) and (II) stated in Theorem B hold. The upper bound

$$\limsup_{n\to\infty}\frac{1}{n}\log\Omega_n^{\varphi}(\mathcal{K})\leq\sup_{\mu\in\mathcal{K}}q^{\varphi}(\mu)\qquad(\mathcal{K}\subset\mathcal{M}(X):\ \text{closed})$$

follows by the same argument as in the proof of Theorem A, using Theorem 3.1 and Proposition 3.4.

Let $\mathcal{G}\subset\mathcal{M}(X)$ be an open subset and fix $\mu\in\mathcal{G}$. In what follows, we show that

$$\liminf_{n\to\infty}\frac{1}{n}\log\Omega_n^{\varphi}(\mathcal{G})\geq q^{\varphi}(\mu).\tag{3.4}$$

If μ is not σ-invariant, then $q^{\varphi}(\mu)=-\infty$ so we may, without loss of generality, assume that $\mu\in\mathcal{M}_{\sigma}(X)$. Given any $\epsilon>0$, choose a neighborhood $\mathcal{G}^{\dagger}\subset\mathcal{G}$ of μ such that $|\int\varphi d\mu-\int\varphi d\nu|\leq\epsilon$ for any $\nu\in\mathcal{G}^{\dagger}$. By the conditions (I) and (II) stated in Theorem B, Proposition 2.7 implies that there exists an ergodic measure $\nu\in\mathcal{G}^{\dagger}$ such that $h(\nu)\geq h(\mu)-\epsilon$. Thus it follows from Propositions 3.4 and 3.5 that

$$\begin{aligned}\liminf_{n\to\infty}\frac{1}{n}\log\Omega_n^{\varphi}(\mathcal{G}) &= \liminf_{n\to\infty}\frac{1}{n}\log\sum_{x\in P_n(X)\cap\mathcal{E}_n^{-1}(\mathcal{G})}e^{S_n\varphi(x)}-P(\varphi)\\ &\geq h(\nu)+\int\varphi d\nu-P(\varphi)\\ &\geq h(\mu)+\int\varphi d\mu-P(\varphi)-2\epsilon.\end{aligned}$$

Since $\epsilon>0$ is arbitrary, the desired inequality (3.4) follows, completing the proof of Theorem B.

4. Applications

In this section, we apply Theorems A and B to several concrete systems. We begin with the β-shift, a prototypical example beyond specification to which both Theorems A and B apply. This example illustrates the overlap in the applicability of the two theorems, while the following subsections focus on applications of Theorem A and Theorem B, respectively.

Example 4.1. (β-shifts). For $\beta > 1$, the *β-transformation* $T_\beta : [0,1] \to [0,1]$ is defined by

$$T_\beta(x) = \begin{cases} \beta x \ (\mathrm{mod} 1), & (x \neq 1), \\ \lim\limits_{y \to 1-0} (\beta y \ (\mathrm{mod} 1)), & (x = 1), \end{cases}$$

and was introduced by Rényi [29]. The associated symbolic coding space Σ_β is called the *β-shift*.

It is well known that Σ_β fails to satisfy the specification property for Lebesgue-almost every $\beta > 1$ (see, e.g., [30]). Nevertheless, Thompson [38] showed that every β-shift satisfies the *periodic almost specification property*. Hence Theorem A applies to all β-shifts and yields a level-2 large deviation principle by periodic orbits.

Let $k \geq 1$ be the largest integer less than β. To verify the conditions (I) and (II) stated in Theorem B, it suffices to check conditions (EX) and (DP) given in Proposition 2.9.

For condition (EX), one can show that for any open interval $I \subset [0,1]$, there exist a positive integer n and an open subinterval $L \subset I$ such that $T_\beta^n | L$ is strictly monotonic and continuous, and $T_\beta^n(L) = (0,1)$.

Moreover, by a classical result of Sigmund [32, Theorem (ii)], the set of periodic measures $\mathcal{M}_\sigma^p(\Sigma_\beta)$ is dense in the set of ergodic measures $\mathcal{M}_\sigma^e(\Sigma_\beta)$. Thus condition (DP) holds.

Therefore, Proposition 2.9 implies that the conditions (I) and (II) stated in Theorem B hold and Theorem B applies to β-shifts as well.

Remark 4.1. Although the almost specification property for the systems considered here is proved in Thompson [38], a careful reading of the proofs therein shows that the periodic almost specification property also holds. We therefore omit a detailed proof of this fact. The same observation applies to the subsequent examples as well.

4.1. **Applications of Theorem A.** In this subsection, we further illustrate Theorem A through symbolic dynamical systems satisfying the periodic almost specification property. Beyond the β-shift discussed in Example 4.1, we consider representative examples such as the S-gap shifts and other related subshifts with non-unique equilibrium states. Such systems provide important models for analyzing non-uniform specification behavior and the non-uniqueness of equilibrium states in symbolic dynamics (see, for example, [1, 6, 9, 19, 22, 23]).

Example 4.2. (S-gap shifts). Let $S \subset \mathbb{N}_0$ be an infinite set. The *S-gap shift* is the two-sided subshift $\Sigma_S \subset \{0,1\}^{\mathbb{Z}}$ defined by the rule that the number of 0's between consecutive 1's is an integer in S. Equivalently, the language of Σ_S consists of all subblocks of words of the form

$$\text{“}0^{n_1}10^{n_2}1 \cdots 0^{n_k}1\text{”} \quad (n_i \in S \ (1 \leq i \leq k), \ k \geq 1).$$

The system Σ_S is topologically mixing if and only if

$$\gcd\{\, n+1 : n \in S \,\} = 1,$$

and it satisfies the specification property if and only if

$$\sup_{i\geq 1}(m_{i+1} - m_i) < \infty, \quad \text{where } S = \{m_1 < m_2 < \cdots\}$$

(see [19, Example 3.4] for instance). Thus there exist many choices of S for which Σ_S is topologically mixing yet fails to have the specification property.

In fact, it is known that every topologically mixing S-gap shift possesses the periodic almost specification property (see [7, Remark 5.1] for instance). Hence, by Theorem A, every topologically mixing S-gap shift satisfies a level-2 large deviation principle by periodic orbits.

Remark 4.2. If the S-gap shift Σ_S is not topologically mixing, that is, $p = \gcd\{\, n+1 : n \in S \,\} > 1$, then every periodic orbit of Σ_S has minimal period divisible by p. In this case, the periodic-orbit LDP does not hold, since only multiples of p occur as periods.

Recent work of Takahasi [37] shows that the Dyck shift, a classical example of a subshift failing intrinsic ergodicity, nevertheless satisfies a level-2 large deviation principle by periodic orbits. His result highlights that strong large-deviation behavior can persist even when the system admits multiple measures of maximal entropy. In this context, Theorem A provides a complementary perspective. Indeed, Example 4.3 below yields another non-intrinsically ergodic subshift beyond specification, obtained from the construction of Kwietniak–Oprocha–Rams, and shows that the same periodic-orbit LDP holds in this broader setting.

Example 4.3. (A non–intrinsically ergodic subshift). We recall a construction due to Kwietniak–Oprocha–Rams [23] that yields a class of subshifts without the specification property and with multiple measures of maximal entropy, while still satisfying the periodic almost specification property.

Fix integers $p, q \geq 2$ and set

$$A := \{1, \ldots, p\} \times \{0, \ldots, q-1\} \;\cup\; \{(0,0)\},$$

which will be the alphabet. For a finite block

$$((a_1, b_1), \ldots, (a_n, b_n)) \in A^n$$

we write it in the two-row form

$$\begin{bmatrix} a_1 & \cdots & a_n \\ b_1 & \cdots & b_n \end{bmatrix}.$$

Let $\{R_n\}_{n\geq 1}$ be an increasing sequence with $R_n \subset \{1, \ldots, n\}$. For each $n \geq 1$ define the family F_n of blocks of the form

$$\begin{bmatrix} 0 & a_1 & \cdots & a_n \\ 0 & b_1 & \cdots & b_n \end{bmatrix},$$

where $1 \le a_i \le p$, $0 \le b_1, \dots, b_n \le q-1$, and $b_j = 0$ whenever $j \in R_n$. Put

$$\mathcal{F} := \bigcup_{n \ge 1} F_n \ \cup \ \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}.$$

Next set

$$\mathcal{M} = \left\{ \begin{bmatrix} a & \dots & a \\ b_1 & \dots & b_n \end{bmatrix} : 1 \le a \le p,\ 0 \le b_1, \dots, b_n \le q-1,\ n \ge 1 \right\}$$

and define the language $\mathcal{L}_R$ to be the set of all subblocks of words of the form

$$\text{“}UV_1 \cdots V_k\text{”} \quad (U \in \mathcal{M},\ V_1, \cdots, V_k \in \mathcal{F},\ k \ge 1).$$

Finally, let

$$X_R := \{x \in A^{\mathbb{Z}} : \text{ every finite subblock of } x \text{ belongs to } \mathcal{L}_R\}.$$

Kwietniak–Oprocha–Rams show in [23, Proposition 4.6] that if, for example,

$$R_n = \{j \le n : \ j = k^2 \text{ for some } k \in \mathbb{N}\},$$

then the subshift X_R satisfies the following properties:

(I) X_R does not have the specification property;
(II) X_R is not intrinsically ergodic (that is, there exist multiple ergodic measures of maximal entropy);
(III) X_R satisfies the periodic almost specification property.

Since property (III) provides the periodic almost specification hypothesis required in Theorem A, we may apply Theorem A to X_R.

This yields a nontrivial family of symbolic systems for which a level-2 large-deviation principle by periodic orbits holds despite the failure of both specification and intrinsic ergodicity; and thus provides a complementary example to recent results obtained for other non-intrinsically ergodic systems [37].

Example 4.4. (Full shift with multiple equilibrium states). In [22], Kucherenko, Quas and Wolf constructed continuous (non-Hölder) potentials on the two-sided full shift on two symbols which exhibit phase transitions and, in particular, admit more than one ergodic equilibrium state; see [22, Theorem 1].

Let φ be such a potential. Since the two-sided full shift has the periodic specification property, Theorem A applies to (X, φ). Thus, even though φ has multiple equilibrium states, a level-2 large deviation principle by periodic orbits nevertheless holds.

This shows that the results apply beyond the Hölder setting and remain valid even when the equilibrium state is not unique.

4.2. **Applications of Theorem B.** In this subsection, we illustrate Theorem B through symbolic dynamical systems beyond the classical β-shift. It is known that for Lebesgue almost every parameters $(\alpha, \beta) \in [0, 1) \times (1, \infty)$, the (α, β)-shift fails to satisfy the specification property, and that, except for continuous cases, the generalized β-shift also lacks specification for almost every $\beta > 1$ (see [3, Main Theorem] for details). Nevertheless, these systems satisfy the conditions (I) and (II) stated in Theorem B, and hence a level-2 large deviation principle holds by periodic orbits. These systems are natural extensions of the classical β-shift and thus provide representative examples where Theorem B can be effectively applied. They have also been studied intensively in recent years (see, for example, [4, 11, 12, 25, 33, 36]).

Example 4.5. ((α, β)-shifts). For $\beta > 1$ and $0 \leq \alpha < 1$, the (α, β)-*transformation* $T_{\alpha,\beta} : [0, 1] \to [0, 1]$ is defined by

$$T_{\alpha,\beta}(x) = \begin{cases} \beta x + \alpha \pmod 1, & (x \neq 1), \\ \lim\limits_{y \to 1-0} (\beta y + \alpha \pmod 1), & (x = 1), \end{cases}$$

and was introduced by Parry [26]. Let $\Sigma_{\alpha,\beta}$ denote the symbolic coding space of $T_{\alpha,\beta}$. Clearly, $\Sigma_{0,\beta}$ coincides with the β-shift.

It is known that if $0 \leq \alpha < 1$ and $\beta > 2$, then $T_{\alpha,\beta}$ satisfies condition (EX) (see, for instance, [5, Example 4.1]).

Furthermore, it is known that $\mathcal{M}^p_\sigma(\Sigma_{\alpha,\beta})$ is dense in $\mathcal{M}^e_\sigma(\Sigma_{\alpha,\beta})$. Hence condition (DP) holds ([17]).

Therefore, by Proposition 2.9, the conditions (I) and (II) stated in Theorem B are satisfied. Consequently, Theorem B applies, and the (α, β)-shift $\Sigma_{\alpha,\beta}$ satisfies a level-2 large deviation principle by periodic orbits.

Example 4.6. (Generalized β-shifts). Let $\beta > 1$ and k be a smallest integer, which is greater than or equal to β, and fix $E = (E_1, \ldots, E_k) \in \{+1, -1\}^k$. We partition $[0, 1]$ into k intervals

$$I_1 := [0, 1/\beta),\ I_2 := [1/\beta, 2/\beta), \ldots, I_k := [(k-1)/\beta, 1].$$

The generalized β-transformation $T_{\beta,E} \colon [0, 1] \to [0, 1]$ was introduced by Góra [14] and defined by

$$T_{\beta,E}(x) = \begin{cases} \beta x - i + 1 & (x \in I_i,\ E(i) = +1), \\ -\beta x + i & (x \in I_i,\ E(i) = -1). \end{cases}$$

It is straightforward to verify that $T_{\beta,E}$ is continuous if and only if $E_i \neq E_{i+1}$ for all $1 \leq i < k$. As mentioned at the beginning of this subsection, when $T_{\beta,E}$ is discontinuous, that is, when $E_i = E_{i+1}$ for some i, the system fails to satisfy the specification property for Lebesgue almost every $\beta > 1$. If $E = (-1, \ldots, -1)$, then we call $T_{\beta,E}$ a $(-\beta)$-transformation [18]. Let $\Sigma_{\beta,E}$ denote the coding space of $T_{\beta,E}$ and call it the generalized β-shift.

Let $\beta > 2$. By an argument similar to the proof of [4, Proposition 3.14], we can verify that condition (EX) holds for $T_{\beta,E}$.

Moreover, it is shown in [31] that the set of periodic measures is dense in the set of ergodic measures of $\Sigma_{\beta,E}$, so that condition (DP) holds.

Therefore, Proposition 2.9 applies, and Theorem B yields a level-2 large deviation principle by periodic orbits for the generalized β-shift $\Sigma_{\beta,E}$.

Acknowledgments. We would like to thank the anonymous referee for careful reading and valuable comments, which significantly improved the presentation of this paper.

Note on AI usage. The author used ChatGPT (GPT-5, OpenAI) solely for English proofreading and language refinement. The use of this AI tool did not influence the originality, mathematical content, or intellectual contribution of this research.

References

[1] Baker, Simon; Ghenciu, Andrei E, Dynamical properties of S-gap shifts and other shift spaces, J. Math. Anal. Appl. 430 (2015), no. 2, 633–647.
[2] Bowen, Rufus, Some systems with unique equilibrium states, Math. Systems Theory 8 (1974/75), no. 3, 193–202.
[3] Buzzi, Jérôme, Specification on the interval, Trans. Amer. Math. Soc. 349 (1997), no. 7, 2737–2754.
[4] Carapezza, Leonard; López, Marco; Robertson, Donald, Unique equilibrium states for some intermediate beta transformations, Stoch. Dyn. 21 (2021), no. 6, Paper No. 2150035, 25 pp.
[5] Chung, Yong, Moo; Yamamoto, Kenichiro, Large deviation principle for piecewise monotonic maps with density of periodic measures, Ergodic Theory Dynam. Systems 43 (2023), no. 3, 861–872.
[6] Climenhaga, Vaughn; Thompson, Daniel J., Intrinsic ergodicity beyond specification: β-shifts, S-gap shifts, and their factors, Israel J. Math. 192 (2012), no. 2, 785–817.
[7] Climenhaga, Vaughn; Thompson, Daniel J.; Yamamoto, Kenichiro, Large deviations for systems with non-uniform structure, Trans. Amer. Math. Soc. 369 (2017), no. 6, 4167–4192.
[8] Comman, Henri; Rivera-Letelier, Juan, Large deviation principles for non-uniformly hyperbolic rational maps, Ergodic Theory Dynam. Systems 31 (2011), no. 2, 321–349.
[9] Dastjerdi, Dawoud Ahmadi; Jangjoo, Somaye, Dynamics and topology of S-gap shifts, Topology Appl. 159 (2012), no. 10-11, 2654–2661.
[10] Denker, Manfred; Grillenberger, Christian; Sigmund, Karl, Ergodic theory on compact spaces, Lecture Notes in Math., Vol. 527 Springer-Verlag, Berlin-New York, 1976, iv+360 pp.
[11] Faller, Bastien, Contribution to the ergodic theory of piecewise monotone continuous map, Ph. D. thesis, École Polytechnique Fédérale de Lausanne, 2008.
[12] Faller, B.; Pfister, C.-E., A point is normal for almost all maps $\beta x + \alpha$ mod1 or generalized β-transformations, Ergodic Theory Dynam. Systems 29 (2009), no. 5, 1529–1547.
[13] Gelfert, Katrin; Wolf, Christian, On the distribution of periodic orbits, Discrete Contin. Dyn. Syst. 26 (2010), no. 3, 949–966.
[14] Góra, Paweł, Invariant densities for generalized β-maps, Ergodic Theory Dynam. Systems, 27 (5) (2007),1583–1598.
[15] Hofbauer, Franz, Intrinsic ergodicity of piecewise monotonic transformations with positive entropy II, Israel J. Math. 38 (1-2) (1981), 107–115.
[16] Hofbauer, Franz, Piecewise invertible dynamical systems, Probab. Theory Relat. Fields 72 (1986), no. 3, 359–386.
[17] Hofbauer, Franz, Generic properties of invariant measures for simple piecewise monotonic transformations, Israel J. Math. 59 (1987), no. 1, 64–80.
[18] Ito, Shunji; Sadahiro, Taizo, Beta-expansions with negative bases, Integers 9 (2009), A22, 239–259.
[19] Jung, Uijin, On the existence of open and bi-continuing codes, Trans. Amer. Math. Soc. 363 (2011), no. 3, 1399–1417.

[20] Keller, Gerhard, Lifting measures to Markov extensions, Monatsh. Math. 108 (1989), no. 2-3, 183–200.
[21] Kifer, Yuri, Large deviations, averaging and periodic orbits of dynamical systems, Comm. Math. Phys. 162 (1994), no. 1, 33–46.
[22] Kucherenko, Tamara; Quas, Anthony; Wolf, Christian, Multiple phase transitions on compact symbolic systems, Adv. Math. 385 (2021), Paper No. 107768, 19 pp.
[23] Kwietniak, Dominik; Oprocha, Piotr; Rams, Michał, On entropy of dynamical systems with almost specification, Israel J. Math. 213 (2016), no. 1, 475–503.
[24] Lind, Douglas; Marcus, Brian, An introduction to symbolic dynamics and coding, Second edition Cambridge Math. Lib. Cambridge University Press, Cambridge, 2021. xix+550 pp.
[25] Oguchi, Mai; Shinoda, Mao, Hausdorff dimension of the parameters for (α, β)-shifts with the specification property, Dyn. Syst. 39 (2024), no. 4, 848–855.
[26] Parry, William, Representations for real numbers, Acta Math. Acad. Sci. Hungar. 15 (1964), 95–105.
[27] Pfister, Charles-Edouard; Sullivan, Wayne G, Large deviations estimates for dynamical systems without the specification property, Applications to the β-shifts. Nonlinearity 18 (2005), no. 1, 237–261.
[28] Pollicott, Mark; Sridharan, Shrihari, Large deviation results for periodic points of a rational map, J. Dyn. Syst. Geom. Theor. 5 (2007), no. 1, 69–77.
[29] Rényi, Alfréd, Representations for real numbers and their ergodic properties, Acta Math. Acad. Sci. Hungar. 8 (1957), 477–493.
[30] Schmeling, Jörg. Symbolic dynamics for β-shifts and self-normal numbers.
[31] Shinoda, Mao; Yamamoto, Kenichiro, Density of periodic measures and large deviation principle for generalised mod one transformations, Nonlinearity 37 (2024), no. 2, Paper No. 025003, 15 pp.
[32] Sigmund, Karl. On the distribution of periodic points for β-shifts. Monatsh. Math. 82 (1976), no. 3, 247–252.
[33] Suzuki, S., Artin–Mazur zeta functions of generalized β-transformations, Kyushu Journal of Mathematics, Vol. 71 (2017), No. 1, 85–103.
[34] Takahashi, Yōichirō, Entropy functional (free energy) for dynamical systems and their random perturbations, Stochastic analysis (Katata/Kyoto, 1982), 437–467, North-Holland Math. Library, 32, North-Holland, Amsterdam, 1984.
[35] Takahasi, Hiroki, Entropy approachability for transitive Markov shifts over infinite alphabet, Proc. Amer. Math. Soc. 148 (2020), no. 9, 3847–3857.
[36] Takahasi, Hiroki, Uniqueness of minimizer for countable Markov shifts and equidistribution of periodic points, J. Stat. Phys. 181 (2020), no. 6, 2415–2431.
[37] Takahasi, Hiroki, Large Deviation Principles for periodic points of the Dyck shift, arXiv:2503.14164, 2025.
[38] Thompson, D. J., Irregular sets, the β-shift and the almost specification property, Trans. Amer. Math. Soc. **364** (2012), 5395–5414.
[39] Thompson, D. J., Generalized β-transformations and the entropy of unimodal maps, Commentarii Mathematici Helvetici, Vol. 92 (2017), No. 4, 777–800.
[40] Yamamoto, Kenichiro, On the weaker forms of the specification property and their applications. Proc. Amer. Math. Soc. 137 (2009), no. 11, 3807–3814.
[41] Yamamoto, Kenichiro, Large deviations and periodic orbits of dynamical systems without the specification property, Shibaura Institute of Technology Research Bulletin, Science and Engineering Series, vol. 56, no. 1 (2012), 43-45 (non-refereed).
[42] Young, Lai-Sang, Some large deviation results for dynamical systems, Trans. Amer. Math. Soc. 318 (1990), no. 2, 525–543.

Department of General Education, Nagaoka University of Technology, Nagaoka 940-2188, JAPAN
Email address: k_yamamoto@vos.nagaokaut.ac.jp