BULLETIN OF THE BRAZILIAN MATHEMATICAL SOCIETY

Bull Braz Math Soc, New Series 41(3), 321-338

Weak KAM methods and ergodic optimal problems for countable Markov shifts

Rodrigo Bissacot and Eduardo Garibaldi

Abstract. Let $\sigma : \mathbf{\Sigma} \to \mathbf{\Sigma}$ be the left shift acting on $\mathbf{\Sigma}$, a one-sided Markov subshift on a countable alphabet. Our intention is to guarantee the existence of σ-invariant Borel probabilities that maximize the integral of a given locally Hölder continuous potential $A : \mathbf{\Sigma} \to \mathbb{R}$. Under certain conditions, we are able to show not only that A-maximizing probabilities do exist, but also that they are characterized by the fact their support lies actually in a particular Markov subshift on a finite alphabet. To that end, we make use of objects dual to maximizing measures, the so-called sub-actions (concept analogous to subsolutions of the Hamilton-Jacobi equation), and specially the calibrated sub-actions (notion similar to weak KAM solutions).

Keywords: weak KAM methods, countable Markov shifts, ergodic optimization, maximizimg measures, sub-actions.

Mathematical subject classification: 37A05, 37A60, 37B10.

1 Introduction

The development of the study of maximizing probabilities has given place to a new and exciting field in ergodic theory. Growing in the intersection of topological dynamical systems and optimization theory, this fresh theorical branch is known nowadays as *ergodic optimization*. Many results were already obtained for dynamics defined by a continuous map $T : X \to X$ of a compact metric space X assuming T has some hyperbolicity (see, for instance, [1, 3, 5, 7, 11]). Although ergodic optimal problems in the context of noncompact dynamical systems have been much less discussed, interesting works can be found in the literature (see, for example, [9, 10]).

The principal purpose of this article is to take into account ergodic optimal problems for a class of noncompact symbolic dynamics: topological Markov

Received 3 February 2010.

shifts with a countable number of states. Let then $\mathbf{\Sigma}$ denote a one-sided Markov subshift on a countable alphabet, and $\sigma\colon \mathbf{\Sigma} \to \mathbf{\Sigma}$ the left shift map. If $A\colon \mathbf{\Sigma} \to \mathbb{R}$ is continuous and bounded above, one would like to determine and describe the σ-invariant Borel probability measures μ that maximize the average value $\int A\, d\mu$. In general, such a maximizing probability does not even exist, since $\mathbf{\Sigma}$ may be noncompact. We show that this is not the case when the potential A is sufficiently regular and verifies a coercive condition. In reality, our main theorem (see theorem 1) states that, for one of these specific potentials, its maximizing probabilities have in common the fact of being supported in a certain compact σ-invariant subset that is actually contained in a Markov subshift on a finite alphabet.

A second objective of this paper is to point out that weak KAM methods (or viscosity solutions technics) can be adapted and employed also in noncompact ergodic optimization. Tools of the theory of viscosity solutions have been successfully used in Lagrangian mechanics (see, for instance, [2, 4]). Ergodic optimization on compact spaces has witnessed the usefulness of these methods, specially when ergodic optimal problems are interpreted as questions of variational dynamics (see, for example, [3, 5, 11]). We adopt the same spirit and strategy here.

2 Basic concepts and main result

Our dynamical setting will be special topologically mixing Markov subshifts on a countable alphabet: the *primitive* ones. Let us introduce them precisely.

For the sake of definiteness, the countably infinite alphabet will always be the set of nonnegative integers $\mathbb{Z}_+$. Let thus $\mathbf{M}\colon \mathbb{Z}_+ \times \mathbb{Z}_+ \to \{0, 1\}$ be a transition matrix. Consider the following sets of symbols given in an inductive way by

$$\mathcal{B}_0 = \left\{ i \in \mathbb{Z}_+ \colon \mathbf{M}(i, j) = 1 \text{ for some } j \in \mathbb{Z}_+ \right\} \text{ and}$$

$$\mathcal{B}_n = \left\{ i \in \mathbb{Z}_+ \colon \mathbf{M}(i, j) = 1 \text{ for some } j \in \mathcal{B}_{n-1} \right\}, \text{ for } n > 0.$$

We say that the transition matrix $\mathbf{M}$ is *primitive* if there exist a (possibly countable) subset $\mathbb{F} \subseteq \mathbb{Z}_+$ and an integer $K_0 \geq 0$ such that, for any pair of symbols $i, j \in \bigcap_{n \geq 0} \mathcal{B}_n$, one can find $\ell_1, \ell_2, \ldots, \ell_{K_0} \in \mathbb{F}$ satisfying

$$\mathbf{M}(i, \ell_1)\mathbf{M}(\ell_1, \ell_2) \cdots \mathbf{M}(\ell_{K_0}, j) = 1.$$

In particular, we say that $\mathbf{M}$ is *finitely primitive* when $\mathbb{F}$ is finite.

Consider then the associated Markov subshift

$$\mathbf{\Sigma} = \left\{ \mathbf{x} = (x_0, x_1, \ldots) \in {\mathbb{Z}_+}^{\mathbb{Z}_+} \colon \mathbf{M}(x_j, x_{j+1}) = 1 \right\}.$$

Fixed $\lambda \in (0, 1)$, we equip $\mathbf{\Sigma}$ with the complete metric $d(\mathbf{x}, \mathbf{y}) = \lambda^k$, where $\mathbf{x} = (x_0, x_1, \ldots), \mathbf{y} = (y_0, y_1, \ldots) \in \mathbf{\Sigma}$ and $k = \min\{j : x_j \neq y_j\}$. It is easy to see that $\mathbf{\Sigma}$ is compact if, and only if, $\bigcap_{n\geq 0} \mathcal{B}_n$ is finite[1]. Let $\sigma : \mathbf{\Sigma} \to \mathbf{\Sigma}$ be the shift map, namely, $\sigma(x_0, x_1, x_2, \ldots) = (x_1, x_2, \ldots)$. We will also say that the dynamics $(\mathbf{\Sigma}, \sigma)$ is (finitely) primitive. Since $\mathbf{M}$ is primitive, clearly $(\mathbf{\Sigma}, \sigma)$ is a topologically mixing dynamical system.

Denote by $\mathcal{M}_\sigma$ the σ-invariant Borel probability measures. Let $C^0(\mathbf{\Sigma})$ indicate the space of continuous real-valued functions on $\mathbf{\Sigma}$, equipped with the topology of uniform convergence on compact subsets. We remind then central concepts in the ergodic optimization theory.

Definition 1. If the potential $A \in C^0(\mathbf{\Sigma})$ is bounded above, we define the ergodic maximizing value by

$$\beta_A = \sup_{\mu \in \mathcal{M}_\sigma} \int A \, d\mu.$$

Any σ-invariant probability achieving this supremum is called maximizing (or, if precision is required, A-maximizing).

We are particularly interested in ergodic optimal results for *locally Hölder continuous* potentials.

Definition 2. A potential $A : \mathbf{\Sigma} \to \mathbb{R}$ is called locally Hölder continuous when there exists a constant $H_A > 0$ such that, for all integer $k \geq 1$, we have

$$\mathrm{Var}_k(A) := \sup_{\mathbf{x},\mathbf{y}\in\mathbf{\Sigma},\ d(\mathbf{x},\mathbf{y})\leq\lambda^k} \big[A(\mathbf{x}) - A(\mathbf{y})\big] \leq H_A \lambda^k.$$

Such a regularity condition only means that the k-th variation $\mathrm{Var}_k(A)$ decays exponentially fast to zero when $k \to \infty$. We could focus on more general regularity assumptions, like summability of variations. Recall that $A : \mathbf{\Sigma} \to \mathbb{R}$ has summable variations if

$$\mathrm{Var}(A) := \sum_{k=1}^{\infty} \mathrm{Var}_k(A) < \infty.$$

Yet one of our main goals here is to provide examples of the applicability of the weak KAM technics. We believe local Hölder continuity is sufficient for this end.

[1] Since the compact situation is well studied, the interesting case occurs naturally when $\bigcap_{n\geq 0} \mathcal{B}_n$ is countable. The reader is thus invited to assume this hypothesis without hesitation.

Notice that nothing is required from $\mathrm{Var}_0(A) := \sup_{\mathbf{x},\mathbf{y}\in\mathbf{\Sigma}}[A(\mathbf{x}) - A(\mathbf{y})]$, which means that a locally Hölder continuous potential, despide its uniform continuity, may be unbounded. So a common assumption[2] in this article will be

$$\inf A|_{\bigcup_{i\in\mathbb{F}}[i]} > -\infty,$$

where $[i]$ just indicates the cylinder set $\{\mathbf{x} = (x_0, x_1, \ldots) \in \mathbf{\Sigma} : x_0 = i\}$.

Under this hypothesis, we will obtain in the next section a dual formula

$$\beta_A = \inf_{f\in C^0(\mathbf{\Sigma})} \sup_{\mathbf{x}\in\mathbf{\Sigma}} (A + f - f \circ \sigma)(\mathbf{x}).$$

This expression raises the natural question about the existence of functions achieving the above infimum, which motivates the following definition.

Definition 3. Suppose $A : \mathbf{\Sigma} \to \mathbb{R}$ is continuous and bounded above. A sub-action (for the potential A) is a function $u \in C^0(\mathbf{\Sigma})$ verifying

$$(A + u - u \circ \sigma)(\mathbf{x}) \leq \beta_A, \quad \forall\, \mathbf{x} \in \mathbf{\Sigma}.$$

We will see in section 4 that it is possible to construct locally Hölder continuous sub-actions for potentials with the same regularity (see proposition 4). This result is completely new as far as we know.

In the context of a noncompact dynamical system, given an arbitrary bounded above continuous potential, the existence of maximizing probabilities is a non-trivial question. However, we will be able to use the existence of sub-actions as well as their properties in order to guarantee there exist maximizing probabilities when we are taking into account coercive potentials.

Definition 4. A continuous potential $A : \mathbf{\Sigma} \to \mathbb{R}$ is said coercive when

$$\lim_{i\to+\infty} \sup A|_{[i]} = -\infty.$$

In Aubry-Mather theory for Lagrangian systems, superlinearity is the usual coercive hypothesis (see, for instance, [2, 4]). The coercive condition is not strange to the countable Markov shift framework. On the contrary, it is an essential theorical piece (in general implicitly) in several studies of the thermodynamic formalism generalized to a finitely primitive Markov subshift on

[2]Note this assumption is trivially verified when $\mathbb{F}$ is finite. Indeed, choosing a point $\mathbf{x}^i \in [i]$ for each $i \in \mathbb{F}$, if $\mathbf{x} \in [j]$, $j \in \mathbb{F}$, then $A(\mathbf{x}) > A(\mathbf{x}^j) - \mathrm{Var}_1(A)$ obviously implies $\inf A|_{\bigcup_{i\in\mathbb{F}}[i]} > \min_{i\in\mathbb{F}} A(\mathbf{x}^i) - \mathrm{Var}_1(A) > -\infty$.

a countable alphabet. Coerciveness obviously follows from the imposition $\sum_i \exp(\sup A|_{[i]}) < \infty$. This summability condition is equivalent to the finiteness of the topological pressure when the potential A is, for example, locally Hölder continuous. This summability condition also allows to define the Ruelle operator $\mathcal{L}_A f(\mathbf{x}) := \sum_{\sigma(\mathbf{y})=\mathbf{x}} e^{A(\mathbf{y})} f(\mathbf{y})$ for a bounded continuous function $f\colon \mathbf{\Sigma} \to \mathbb{R}$. For more details, we refer the reader to the book of R.D. Mauldin and M. Urbański (see [12]). Furthermore, when $(\mathbf{\Sigma}, \sigma)$ is finitely primitive and A is locally Hölder continuous, it is not difficult to show the hypothesis $\|\mathcal{L}_A 1\|_\infty < \infty$ (omnipresent in the work of O. Sarig [15]) implies coerciveness too.

Given a nonnegative integer I, denote by

$$\Sigma_I = \Big\{ \mathbf{x} = (x_0, x_1, \ldots) \in \{0, \ldots, I\}^{\mathbb{Z}_+} : \mathbf{M}(x_j, x_{j+1}) = 1 \Big\}$$

the Markov subshift on the finite alphabet $\{\iota_1, \ldots, \iota_{r_I}\} := \{0, \ldots, I\} \cap \big(\bigcap_{n\geq 0} \mathcal{B}_n\big)$ associated to the transition matrix $\mathbf{M}|_{\{0,\ldots,I\}\times\{0,\ldots,I\}}$. Obviously Σ_I is a compact σ-invariant subset of $\mathbf{\Sigma}$. So we simply denote $\sigma|_{\Sigma_I}$ by σ.

When $\mathbf{M}$ is finitely primitive, let

$$I_{\mathbb{F}} := \max\{i : i \in \mathbb{F}\}.$$

Our main result concerning the existence of maximizing probabilities can be stated as follows.

Theorem 1. *Suppose $(\mathbf{\Sigma}, \sigma)$ is a finitely primitive Markov subshift on a countable alphabet. Let $A\colon \mathbf{\Sigma} \to \mathbb{R}$ be a bounded above, coercive and locally Hölder continuous potential. Then there exists an integer $\hat{I} > I_{\mathbb{F}}$ such that*

$$\beta_A = \max_{\substack{\mu\in\mathcal{M}_\sigma \\ \text{supp}\mu\subseteq\Sigma_{\hat{I}}}} \int A \, d\mu.$$

In particular, maximizing measures do exist. Furthermore, there exists a compact σ-invariant set $\Omega \subseteq \Sigma_{\hat{I}}$ such that $\mu \in \mathcal{M}_\sigma$ is an A-maximizing probability if, and only if, μ is supported in Ω.

Its proof is discussed in section 4 and exploits the analogy with Aubry-Mather theory in symbolic dynamics. The existence of a bounded continuous sub-action for the potential A will tell us where to seek maximizing probabilities. Nevertheless, a key step to the demonstration is to analyse first the problem for the compact situation (Σ_I, σ), using the uniform oscillatory behavior of some special sub-actions, that are called calibrated, and should be understood as corresponding to Fathi's weak KAM solutions or viscosity solutions of the Hamilton-Jacobi equation.

Theorem 1 clarifies previous results. For instance, in the special case where $\sum_i \exp(\sup A|_{[i]}) < \infty$, the identity $\beta_A = \max_{\mu\in\mathcal{M}_\sigma,\,\mathrm{supp}\mu\subseteq\Sigma_{\hat{I}}} \int A\, d\mu$ is implicitly present in the work of I. D. Morris. Indeed, in the proof of lemma 3.5 of [13], one obtains that, if $\{\mu_t\}_{t>1}$ is the family of equilibrium states of tA, then there is $\hat{I} \in \mathbb{Z}^+$ such that $\mu_t([i]) \to 0$ as $t \to \infty$ for all $i > \hat{I}$. Since this family of probabilities is uniformly tight and any accumulation measure is maximizing, this shows that an A-maximizing probability exists and is supported in $\Sigma_{\hat{I}}$. Concerning the description of all A-maximizing probabilities, in [9] the authors obtained in a more general context a not so precise characterization for their supports (see remark 6).

3 Characterizations of the ergodic maximizing value

We will present other expressions which one could choose in order to introduce the constant β_A for our particular situation. In this section, we will consider a larger class of potentials: the uniformly continuous ones. Remind that $A\colon \mathbf{\Sigma} \to \mathbb{R}$ is uniformly continuous if $\lim_{k\to\infty} \mathrm{Var}_k(A) = 0$. Notice we are still dealing with functions which may be unbounded.

Given $A \in C^0(\mathbf{\Sigma})$, as usual let $S_k A = \sum_{j=0}^{k-1} A \circ \sigma^j$ and $S_0 A = 0$. Hence, the following result identifies the ergodic maximizing value with a maximum ergodic time average.

Proposition 2. *Let* $(\mathbf{\Sigma}, \sigma)$ *be a primitive Markov subshift on a countable alphabet. Assume the uniformly continuous potential* $A\colon \mathbf{\Sigma} \to \mathbb{R}$ *is bounded above and satisfies* $\inf A|_{\bigcup_{i\in\mathbb{F}}[i]} > -\infty$. *Then we verify*

$$\beta_A = \lim_{k\to\infty} \sup_{\mathbf{x}\in\mathbf{\Sigma}} \frac{1}{k} S_k A(\mathbf{x}) = \inf_{k\geq 1} \sup_{\mathbf{x}\in\mathbf{\Sigma}} \frac{1}{k} S_k A(\mathbf{x}).$$

Proof. Note that $\{\sup_{\mathbf{x}\in\mathbf{\Sigma}} S_k A(\mathbf{x})\}_{k\geq 1}$ is a subadditive sequence of real numbers. Therefore, the limit $\lim_{k\to\infty} \sup_{\mathbf{x}\in\mathbf{\Sigma}} \frac{1}{k} S_k A(\mathbf{x})$ exists and is in fact equal to $\inf_{k\geq 1} \sup_{\mathbf{x}\in\mathbf{\Sigma}} \frac{1}{k} S_k A(\mathbf{x})$.

Given a positive integer k, take a point $\mathbf{x}^k \in \mathbf{\Sigma}$ satisfying

$$\sup_{\mathbf{x}\in\mathbf{\Sigma}} \frac{1}{k} S_k A(\mathbf{x}) - \frac{1}{2^k} < \frac{1}{k} S_k A(\mathbf{x}^k).$$

Since $(\boldsymbol{\Sigma}, \sigma)$ is a primitive Markov subshift, for all sufficiently large k, we can find a periodic point $\mathbf{y}^k = (y_0^k, y_1^k, \dots) \in \boldsymbol{\Sigma}$ of period k, with $y_j^k \in \mathbb{F}$ for each $j \in \{k - K_0, \dots, k-1\}$, such that $d(\mathbf{x}^k, \mathbf{y}^k) \le \lambda^{k-K_0}$. From the immediate inequality

$$\frac{1}{k} S_k A(\mathbf{y}^k) \le \beta_A,$$

we obtain

$$\frac{1}{k} S_k A(\mathbf{x}^k) \le \frac{1}{k} S_k A(\mathbf{x}^k) - \frac{1}{k} S_k A(\mathbf{y}^k) + \beta_A$$

$$\le \frac{1}{k} \left[\mathrm{Var}_{k-K_0}(A) + \ldots + \mathrm{Var}_1(A) + K_0 \left(\sup A - \inf A|_{\bigcup_{i \in \mathbb{F}}[i]} \right) \right] + \beta_A.$$

For k large enough, we thus have

$$\sup_{\mathbf{x} \in \boldsymbol{\Sigma}} \frac{1}{k} S_k A(\mathbf{x}) - \frac{1}{2^k} < \frac{1}{k} \left[\sum_{j=1}^{k} \mathrm{Var}_j(A) + K_0 \left(\sup A - \inf A|_{\bigcup_{i \in \mathbb{F}}[i]} \right) \right] + \beta_A.$$

So $\lim_{k \to \infty} \sup_{\mathbf{x} \in \boldsymbol{\Sigma}} \frac{1}{k} S_k A(\mathbf{x}) \le \beta_A$.

In order to show the equality does hold, take a probability $\mu \in \mathcal{M}_\sigma$ such that $A \in L^1(\mu)$. For any $k > 0$, we clearly have

$$\int A \, d\mu = \int \frac{1}{k} S_k A \, d\mu \le \sup_{\mathbf{x} \in \boldsymbol{\Sigma}} \frac{1}{k} S_k A(\mathbf{x}).$$

Taking the infimum over k and then the supremum over μ, we finish the proof. $\square$

We remark that, for a noncompact dynamical system, in general we have

$$\beta_A \le \limsup_{k \to \infty} \sup_{\mathbf{x} \in \boldsymbol{\Sigma}} \frac{1}{k} S_k A(\mathbf{x}).$$

We refer the reader to [10] for a discussion on such a topic.

We present now a dual characterization of β_A.

Proposition 3. *Let $(\boldsymbol{\Sigma}, \sigma)$ be a primitive Markov subshift on a countable alphabet. Suppose the uniformly continuous potential $A \colon \boldsymbol{\Sigma} \to \mathbb{R}$ is bounded above and verifies* $\inf A|_{\bigcup_{i \in \mathbb{F}}[i]} > -\infty$. *Then*

$$\beta_A = \inf_{f \in C^0(\boldsymbol{\Sigma})} \sup_{\mathbf{x} \in \boldsymbol{\Sigma}} (A + f - f \circ \sigma)(\mathbf{x}).$$

Proof. Denote by $C^0_A(\mathbf{\Sigma})$ the set of continuous functions $f\colon \mathbf{\Sigma} \to \mathbb{R}$ satisfying $\sup(A + f - f \circ \sigma) < \infty$. Note that all bounded continuous real-valued functions belong to $C^0_A(\mathbf{\Sigma})$. Moreover, we clearly have

$$\inf_{f \in C^0(\mathbf{\Sigma})} \sup_{\mathbf{x} \in \mathbf{\Sigma}} (A + f - f \circ \sigma)(\mathbf{x}) = \inf_{f \in C^0_A(\mathbf{\Sigma})} \sup_{\mathbf{x} \in \mathbf{\Sigma}} (A + f - f \circ \sigma)(\mathbf{x}) < \infty.$$

By conciseness, write $\varkappa = \inf_{f \in C^0_A(\mathbf{\Sigma})} \sup_{\mathbf{x} \in \mathbf{\Sigma}} (A + f - f \circ \sigma)(\mathbf{x})$. Fix $\epsilon > 0$. Choose a function $f \in C^0_A(\mathbf{\Sigma})$ such that $A + f - f \circ \sigma < \varkappa + \epsilon$. For any $\mu \in \mathcal{M}_\sigma$, we verify

$$\int A \, d\mu = \int (A + f - f \circ \sigma) \, d\mu \leq \varkappa + \epsilon.$$

Hence, $\beta_A \leq \varkappa + \epsilon$. Since $\epsilon > 0$ is arbitrary, we get $\beta_A \leq \varkappa$.

Consider then $f_k = -\frac{1}{k} \sum_{j=1}^{k} S_j A \in C^0(\mathbf{\Sigma})$. The identity

$$A = \frac{1}{k} S_k(A \circ \sigma) + f_k \circ \sigma - f_k$$

implies $\sup(A + f_k - f_k \circ \sigma) = \sup \frac{1}{k} S_k(A \circ \sigma) \leq \sup A < \infty$, that is, $f_k \in C^0_A(\mathbf{\Sigma})$. Therefore, we obtain

$$\varkappa \leq \inf_{k \geq 1} \sup_{\mathbf{x} \in \mathbf{\Sigma}} \frac{1}{k} S_k(A \circ \sigma)(\mathbf{x}).$$

The result follows thus from the previous proposition. □

In ergodic optimization on compact spaces, a similar dual expression of the corresponding ergodic maximizing value is well known (see, for example, [1]).

4 Sub-actions and maximizing probabilities

A minimal sub-action

We will show the existence of minimal sub-actions for locally Hölder continuous potentials. Similar results have been obtained in the compact situation (see, for example, [3, 5]).

Proposition 4. *Assume* $(\mathbf{\Sigma}, \sigma)$ *is a primitive Markov subshift on a countable alphabet. Let* $A\colon \mathbf{\Sigma} \to \mathbb{R}$ *be a bounded above and locally Hölder continuous potential such that* $\inf A|_{\bigcup_{i \in \mathbb{F}} [i]} > -\infty$. *Then there exists an unique minimal,* *nonnegative, bounded and locally Hölder continuous function* $u_A\colon \mathbf{\Sigma} \to \mathbb{R}_+$ *verifying*

$$A + u_A - u_A \circ \sigma \le \beta_A.$$

The minimality is in the sense that, for any nonnegative sub-action $u \in C^0(\mathbf{\Sigma}, \mathbb{R}_+)$ *(not necessarily locally Hölder continuous), we have* $u_A \le u$.

Proof. Given $\mathbf{x} \in \mathbf{\Sigma}$, define

$$u_A(\mathbf{x}) := \sup \left\{ S_k(A - \beta_A)(\mathbf{y})\colon k \ge 0,\ \mathbf{y} \in \mathbf{\Sigma},\ \sigma^k(\mathbf{y}) = \mathbf{x} \right\}.$$

As $S_0(A - \beta_A) = 0$ by convention, obviously $u_A \ge 0$.

Take an integer $k > K_0$ and a point $\mathbf{y} \in \mathbf{\Sigma}$ verifying $\sigma^k(\mathbf{y}) = \mathbf{x}$. We can thus find a periodic point $\mathbf{y}^k = (y_0^k, y_1^k, \ldots) \in \mathbf{\Sigma}$ of period k, with $y_j^k \in \mathbb{F}$ when $j \in \{k - K_0, \ldots, k-1\}$, such that $d(\mathbf{y}, \mathbf{y}^k) \le \lambda^{k-K_0}$. First notice that

$$S_k A(\mathbf{y}) - S_k A(\mathbf{y}^k) \le \mathrm{Var}_{k-K_0}(A) + \ldots + \mathrm{Var}_1(A) + K_0 \left(\sup A - \inf A|_{\bigcup_{i \in \mathbb{F}}[i]} \right).$$

Since clearly $S_k A(\mathbf{y}^k) \le k\beta_A$, we then obtain

$$S_k(A - \beta_A)(\mathbf{y}) \le \mathrm{Var}(A) + K_0 \left(\sup A - \inf A|_{\bigcup_{i \in \mathbb{F}}[i]} \right), \quad \forall\, k > K_0,$$

which assures that

$$\begin{aligned} 0 &\le u_A(\mathbf{x}) \\ &\le \max \left\{ \mathrm{Var}(A) + K_0 \left(\sup A - \inf A|_{\bigcup_{i \in \mathbb{F}}[i]} \right),\ K_0(\sup A - \beta_A) \right\}. \end{aligned} \tag{4.1}$$

So $u_A\colon \mathbf{\Sigma} \to \mathbb{R}_+$ is a well defined bounded function. Moreover, from the identity $A \circ \sigma^k + S_k(A - \beta_A) = S_{k+1}(A - \beta_A) + \beta_A$ and the definition of u_A, we get $A + u_A \le u_A \circ \sigma + \beta_A$.

Concerning its regularity, u_A is a locally Hölder continuous function. Indeed, let $\mathbf{x} = (x_0, x_1, \ldots)$, $\bar{\mathbf{x}} = (\bar{x}_0, \bar{x}_1, \ldots) \in \mathbf{\Sigma}$ be arbitrary points with $d(\mathbf{x}, \bar{\mathbf{x}}) \le \lambda^k$ for some $k \ge 1$. Given $\epsilon > 0$, take an integer $\bar{k} \ge 0$ and a point $\bar{\mathbf{y}} = (\bar{y}_0, \bar{y}_1, \ldots) \in \mathbf{\Sigma}$, with $\sigma^{\bar{k}}(\bar{\mathbf{y}}) = \bar{\mathbf{x}}$, such that

$$u_A(\bar{\mathbf{x}}) - \epsilon < S_{\bar{k}}(A - \beta_A)(\bar{\mathbf{y}}).$$

Consider the point $\mathbf{y} = (\bar{y}_0, \bar{y}_1, \ldots, \bar{y}_{\bar{k}-1}, x_0, x_1, \ldots) \in \mathbf{\Sigma}$ satisfying $\sigma^{\bar{k}}(\mathbf{y}) = \mathbf{x}$. So we have

$$\begin{aligned} u_A(\bar{\mathbf{x}}) - u_A(\mathbf{x}) - \epsilon &< S_{\bar{k}} A(\bar{\mathbf{y}}) - S_{\bar{k}} A(\mathbf{y}) \\ &\le \mathrm{Var}_{k+\bar{k}}(A) + \mathrm{Var}_{k+\bar{k}-1}(A) + \ldots + \mathrm{Var}_k(A) \\ &\le H_A \left(\lambda^{k+\bar{k}} + \lambda^{k+\bar{k}-1} + \ldots + \lambda^k \right) \\ &\le \frac{H_A}{1-\lambda} \lambda^k. \end{aligned}$$

Since ϵ can be considered arbitrarily small, this shows that

$$\mathrm{Var}_k(u_A) \leq \frac{H_A}{1-\lambda}\lambda^k,$$

which means u_A is locally Hölder continuous (with constant $H_{u_A} = \frac{H_A}{1-\lambda}$).

Suppose now that $u \in C^0(\mathbf{\Sigma}, \mathbb{R}_+)$ is a nonnegative sub-action for the potential A. Given $\mathbf{x} \in \mathbf{\Sigma}$, if the point $\mathbf{y} \in \mathbf{\Sigma}$ satisfies $\sigma^k(\mathbf{y}) = \mathbf{x}$ for some $k \geq 0$, it is easy to see that $u(\mathbf{x}) + k\beta_A \geq S_k A(\mathbf{y}) + u(\mathbf{y}) \geq S_k A(\mathbf{y})$. This proves that $u(\mathbf{x}) \geq u_A(\mathbf{x})$. □

Remark 5. If we keep the previous hypotheses when consindering a potential $A \in C^0(\mathbf{\Sigma})$ with summable variations, we still obtain a minimal, non-negative and bounded sub-action $u_A \colon \mathbf{\Sigma} \to \mathbb{R}_+$. Nevertheless, from

$$\mathrm{Var}_k(u_A) \leq \sum_{j \geq k} \mathrm{Var}_j(A),$$

we only assure its uniform continuity.

It is important to notice that the existence of a sub-action as above indicates where we shall look for maximizing probabilities in the coercive case.

Proposition 5. *Let $(\mathbf{\Sigma}, \sigma)$ be a primitive Markov subshift on a countable alphabet. Suppose $u \in C^0(\mathbf{\Sigma})$ is a bounded sub-action for a bounded above and coercive potential $A \in C^0(\mathbf{\Sigma})$. If $\mu \in \mathcal{M}_\sigma$ is an A-maximizing probability, then μ is supported in a Markov subshift on a finite alphabet.*

Proof. Let $\mu \in \mathcal{M}_\sigma$ be an A-maximizing probability. Since $u \in C^0(\mathbf{\Sigma})$ is a sub-action for the potential A, we have

$$A + u - u \circ \sigma - \beta_A \leq 0 \quad \text{and} \quad \int (A + u - u \circ \sigma - \beta_A)\, d\mu = 0.$$

Therefore, the support of μ is a subset of the closed set $(A + u - u \circ \sigma - \beta_A)^{-1}(0)$.

Let $\eta > 0$ be a real constant. As A is coercive and u is bounded, there exists $\hat{I} \in \mathbb{Z}_+$ such that

$$\sup(A + u - u \circ \sigma - \beta_A)|_{\bigcup_{i > \hat{I}}[i]} < -\eta. \tag{4.2}$$

In particular, we obtain $\mu(\bigcup_{i > \hat{I}}[i]) = 0$, or in a more useful way $\mathrm{supp}(\mu) \subseteq \bigcup_{i \leq \hat{I}}[i]$.

Being $\mathrm{supp}(\mu)$ a σ-invariant set, we get $\mathrm{supp}(\mu) \subseteq \bigcap_{k \geq 0} \sigma^{-k}\left(\bigcup_{i \leq \hat{I}}[i]\right) = \Sigma_{\hat{I}}$, which ends the proof. □

Remark 6. In [9], when considering a primitive subshift on a countable alphabet, the authors showed there exist invariant probabilities that maximize the integral of a bounded above and coercive potential A with summable variations and satisfying $\inf A|_{\bigcup_{i \in \mathbb{F}}[i]} > -\infty$. They characterized them by the fact that their support lies in a compact subset of $\mathbf{\Sigma}$. Remark 5 and proposition 5 go beyond guaranteeing that those A-maximizing probabilities are actually supported in a Markov subshift on a finite alphabet.

Results for compact approximations

In the context of a transitive expanding transformation defined on a compact metric space, the theory of ergodic optimization has received special attention, which has yielded a more detailed theorical picture when the potential is sufficiently regular as, let us say, Lipschitz continuous (see, for instance, [1, 3, 5, 7]). To demonstrate theorem 1, we will take advantage of results concerning ergodic optimal problems for the compact approximations (Σ_I, σ).

We suppose henceforth that $(\mathbf{\Sigma}, \sigma)$ is a finitely primitive and $A \colon \mathbf{\Sigma} \to \mathbb{R}$ is a bounded above and locally Hölder continuous potential. Recall (from footnote 2) that in this case $\inf A|_{\bigcup_{i \in \mathbb{F}}[i]} > -\infty$.

For $I \geq I_{\mathbb{F}}$, we will need to consider the following ergodic constants

$$\beta_A(I) := \max_{\substack{\mu \in \mathcal{M}_\sigma \\ \text{supp}\mu \subseteq \Sigma_I}} \int_{\Sigma_I} A \, d\mu.$$

Each one corresponds to the ergodic maximizing value associated to the Lipschitz continuous potential $A|_{\Sigma_I}$ defined on the compact metric space Σ_I. Recall Σ_I is the Markov subshift on the finite alphabet

$$\{\iota_1, \ldots, \iota_{r_I}\} := \{0, \ldots, I\} \cap \left(\bigcap_{n \geq 0} \mathcal{B}_n \right)$$

associated to the transition matrix $\mathbf{M}|_{\{0,\ldots,I\} \times \{0,\ldots,I\}}$. If $I \geq I_{\mathbb{F}}$, then obviously $\mathbb{F} \subset \{\iota_1, \ldots, \iota_{r_I}\}$ and (Σ_I, σ) is a topologically mixing dynamical system.

Remember that, in ergodic optimization on compact spaces, we call *sub-action* for the potential $A|_{\Sigma_I}$ any function $u \in C^0(\Sigma_I)$ satisfying, for each point $\mathbf{x} \in \Sigma_I$, $A(\mathbf{x}) + u(\mathbf{x}) - u \circ \sigma(\mathbf{x}) \leq \beta_A(I)$. Besides, a sub-action $u \in C^0(\Sigma_I)$ is said to be *calibrated* when, for every $\mathbf{x} \in \Sigma_I$, one can find a point $\bar{\mathbf{x}} \in \Sigma_I$, with $\sigma(\bar{\mathbf{x}}) = \mathbf{x}$, such that

$$A(\bar{\mathbf{x}}) + u(\bar{\mathbf{x}}) - u(\mathbf{x}) = \beta_A(I).$$

Main properties of calibrated sub-actions are discussed, for instance, in [3, 5, 7].

Lemma 6. *Assume* $(\mathbf{\Sigma}, \sigma)$ *is a finitely primitive Markov subshift on a countable alphabet. Let* $A\colon \mathbf{\Sigma} \to \mathbb{R}$ *be a bounded above and locally Hölder continuous potential. Consider an integer* $I \geq I_{\mathbb{F}}$. *If* $u \in C^0(\Sigma_I)$ *is a calibrated sub-action for the Lipschitz continuous potential* $A|_{\Sigma_I}$, *then*

$$\operatorname{osc}(u) := \max_{\mathbf{x},\mathbf{y}\in\Sigma_I} [u(\mathbf{x}) - u(\mathbf{y})] \leq \operatorname{Var}(A) + K_0 \left(\sup A - \inf_{i\in\mathbb{F}} A|_{[i]} \right).$$

Proof. Take arbitrary points $\mathbf{x}, \mathbf{y} \in \Sigma_I$. As u is a calibrated sub-action, we define inductively a sequence $\{\mathbf{x}^k = (x_0^k, x_1^k, \ldots)\} \subseteq \Sigma_I$ by choosing $\mathbf{x}^0 := \mathbf{x}$ and, for all $k \geq 0$, demanding $\sigma(\mathbf{x}^{k+1}) = \mathbf{x}^k$ with $u(\mathbf{x}^k) = u(\mathbf{x}^{k+1}) + A(\mathbf{x}^{k+1}) - \beta_A(I)$.

Write $\mathbf{y}^0 := \mathbf{y} = (y_0, y_1, \ldots)$. Since$(\mathbf{\Sigma}, \sigma)$ is finitely primitive and $I \geq I_{\mathbb{F}}$, there exists a word $(w_1, w_2, \ldots, w_{K_0}) \in \mathbb{F}^{K_0}$, with $\mathbf{M}(w_j, w_{j+1}) = 1$, such that $\mathbf{M}(x_0^{K_0+1}, w_1) = 1 = \mathbf{M}(w_{K_0}, y_0)$. So we may consider the point $\mathbf{y}^k \in \Sigma_I$ defined by

$$\mathbf{y}^k = \begin{cases} (w_{K_0-k+1}, \ldots, w_{K_0}, y_0, y_1, \ldots) & \text{if } 1 \leq k \leq K_0 \\ (x_0^k, \ldots, x_0^{K_0+1}, w_1, w_2, \ldots, w_{K_0}, y_0, y_1, \ldots) & \text{if } k > K_0 \end{cases}.$$

Clearly, $\sigma(\mathbf{y}^{k+1}) = \mathbf{y}^k$ and $u(\mathbf{y}^k) \geq u(\mathbf{y}^{k+1}) + A(\mathbf{y}^{k+1}) - \beta_A(I)$.

Then notice that

$$\begin{aligned} u(\mathbf{x}) - u(\mathbf{y}) &\leq u(\mathbf{x}^1) - u(\mathbf{y}^1) + A(\mathbf{x}^1) - A(\mathbf{y}^1) \\ &\leq u(\mathbf{x}^2) - u(\mathbf{y}^2) + A(\mathbf{x}^1) - A(\mathbf{y}^1) + A(\mathbf{x}^2) - A(\mathbf{y}^2) \\ &\vdots \\ &\leq u(\mathbf{x}^k) - u(\mathbf{y}^k) + \sum_{j=1}^{k} [A(\mathbf{x}^j) - A(\mathbf{y}^j)]. \end{aligned}$$

As $d(\mathbf{x}^k, \mathbf{y}^k) = \lambda^{k-K_0-1} d(\mathbf{x}^{K_0+1}, \mathbf{y}^{K_0+1})$ for $k > K_0$, the continuity of u implies $\lim_{k\to\infty}[u(\mathbf{y}^k) - u(\mathbf{x}^k)] = 0$. Hence, we obtain

$$\begin{aligned} u(\mathbf{y}) - u(\mathbf{x}) &\leq \sum_{j=1}^{\infty} [A(\mathbf{x}^j) - A(\mathbf{y}^j)] \\ &= \sum_{j=1}^{K_0} [A(\mathbf{x}^j) - A(\mathbf{y}^j)] + \sum_{j=K_0+1}^{\infty} [A(\mathbf{x}^j) - A(\mathbf{y}^j)] \\ &\leq K_0 \left(\sup A - \inf A|_{\bigcup_{i\in\mathbb{F}}[i]}\right) + \operatorname{Var}(A), \end{aligned}$$

from which the statement follows immediately. □

It is necessary to recall other central notions and facts of ergodic optimization on compact spaces. A point $\mathbf{x} \in \Sigma_I$ is said to be non-wandering with respect to the Lipschitz continuous potential $A|_{\Sigma_I}$ if, for all $\epsilon > 0$, one can find a point $\mathbf{y} \in \Sigma_I$ and an integer $n > 0$ such that

$$d(\mathbf{x}, \mathbf{y}) < \epsilon,\ \ d(\mathbf{x}, \sigma^n(\mathbf{y})) < \epsilon \text{ and } |S_n(A - \beta_A(I))(\mathbf{y})| < \epsilon.$$

Let $\Omega(A, I) \subseteq \Sigma_I$ denote the set of non-wandering points with respect to $A|_{\Sigma_I}$.

This set is a compact σ-invariant subset of Σ_I. For any sub-action $u \in C^0(\Sigma_I)$,

$$\Omega(A, I) \subseteq \{\mathbf{x} \in \Sigma_I : (A + u - u \circ \sigma - \beta_A(I))(\mathbf{x}) = 0\}. \tag{4.3}$$

Furthermore, $\Omega(A, I)$ characterizes the maximizing probabilities in the sense that, for $\mu \in \mathcal{M}_\sigma$ with $\operatorname{supp}(\mu) \subseteq \Sigma_I$, one has

$$\int_{\Sigma_I} A\, d\mu = \beta_A(I) \ \Leftrightarrow\ \operatorname{supp}(\mu) \subseteq \Omega(A, I). \tag{4.4}$$

The demonstrations of these properties and more details on the non-wandering set with respect to a Lipschitz continuous potential may be found, for instance, in [3, 5, 6, 11].

Since $(\Sigma_{I_{\mathbb{F}}}, \sigma)$ is a topologically mixing dynamical system, we may consider a probability measure $\mu_{\mathbb{F}} \in \mathcal{M}_\sigma$ whose support is a periodic orbit in $\Sigma_{I_{\mathbb{F}}} \cap \mathbb{F}^{\mathbb{Z}_+}$. In particular, for all $I \geq I_{\mathbb{F}}$, notice that

$$\beta_A(I) \geq \int_{\Sigma_I} A\, d\mu_{\mathbb{F}} \geq \inf A|_{\bigcup_{i \in \mathbb{F}}[i]}. \tag{4.5}$$

Let us assume in addition that the potential $A\colon \mathbf{\Sigma} \to \mathbb{R}$ is coercive. A fundamental inequality is thus the following one.

Notation 7. The coerciveness of the potential allows us to determine an integer $\hat{I} > I_{\mathbb{F}}$ satisfying

$$\sup A|_{\bigcup_{i > \hat{I}}[i]} < \inf A|_{\bigcup_{i \in \mathbb{F}}[i]} - \left[\operatorname{Var}(A) + K_0 \left(\sup A - \inf A|_{\bigcup_{i \in \mathbb{F}}[i]}\right)\right]. \tag{4.6}$$

So we have an important lemma.

Lemma 7. *Suppose $(\mathbf{\Sigma}, \sigma)$ is a finitely primitive Markov subshift on a countable alphabet. Let $A\colon \mathbf{\Sigma} \to \mathbb{R}$ be a bounded above, coercive and locally Hölder continuous potential. Then*

$$\beta_A(I) = \beta_A(\hat{I}) \quad \forall\ I \geq \hat{I},$$

where the positive integer $\hat{I}$ is defined by (4.6). Furthermore, given an integer $I \geq \hat{I}$, only $(A|_{\Sigma_{\hat{I}}})$-maximizing probabilities maximize the integral of $A|_{\Sigma_I}$ among σ-invariant probabilities supported in Σ_I.

Proof. Clearly $\beta_A(\hat{I}) \leq \beta_A(I)$ whenever $I \geq \hat{I}$. In order to obtain the equality, it is enough to show that every $(A|_{\Sigma_I})$-maximizing probability is actually supported in $\Sigma_{\hat{I}}$.

Suppose on the contrary the existence of a probability measure $\mu \in \mathcal{M}_\sigma$, with $\text{supp}(\mu) \subseteq \Sigma_I$ and $\int_{\Sigma_I} A\, d\mu = \beta_A(I)$, such that $\text{supp}(\mu) - \Sigma_{\hat{I}} \neq \emptyset$.

Take then $\mathbf{x} = (x_0, x_1, \ldots) \in \text{supp}(\mu) - \Sigma_{\hat{I}}$. We may assume $x_0 > \hat{I}$. Therefore, from (4.5) and (4.6), it follows

$$\begin{aligned} A(\mathbf{x}) - \beta_A(I) &\leq \sup A|_{\bigcup_{i>\hat{I}}[i]} - \inf A|_{\bigcup_{i\in\mathbb{F}}[i]} \\ &< -\left[\text{Var}(A) + K_0\left(\sup A - \inf A|_{\bigcup_{i\in\mathbb{F}}[i]}\right)\right]. \end{aligned}$$

Let $u \in C^0(\Sigma_I)$ be a calibrated sub-action for the Lipschitz continuous potential $A|_{\Sigma_I}$. Thanks to (4.4) and (4.3), we have $A(\mathbf{x}) + u(\mathbf{x}) - u\circ\sigma(\mathbf{x}) - \beta_A(I) = 0$, which then yields

$$u(\mathbf{x}) - u\circ\sigma(\mathbf{x}) > \text{Var}(A) + K_0\left(\sup A - \inf A|_{\bigcup_{i\in\mathbb{F}}[i]}\right).$$

However, this inequality contradicts lemma 6 which assures that

$$\text{osc}(u) \leq \text{Var}(A) + K_0\left(\sup A - \inf A|_{\bigcup_{i\in\mathbb{F}}[i]}\right).$$

Hence, necessarily $\text{supp}(\mu) \subseteq \Sigma_{\hat{I}}$ whenever $\mu \in \mathcal{M}_\sigma$ maximizes the integral of $A|_{\Sigma_I}$ among the σ-invariant probabilities supported in Σ_I. □

Proof of Theorem 1

Our strategy is to extend the statement of lemma 7 to the noncompact dynamical system $(\mathbf{\Sigma}, \sigma)$. More precisely, we will show that

$$\int A\, d\mu \leq \beta_A(\hat{I}), \quad \forall\, \mu \in \mathcal{M}_\sigma. \tag{4.7}$$

Clearly it will follow $\beta_A = \beta_A(\hat{I})$, guaranteeing the existence of maximizing probabilities. Propositions 4 and 5 and lemma 7 will then assure that only $A|_{\Sigma_{\hat{I}}}$-maximizing probabilities maximize the integral of the potential A among all σ-invariant Borel probability measures. Besides, from (4.4), the compact σ-invariant subset of $\Sigma_{\hat{I}}$ in the statement of theorem 1 will immediately be $\Omega = \Omega(A, \hat{I})$.

So we just need to demonstrate (4.7). As a matter of fact, this inequality is a consequence of the denseness of probabilities whose support is a pediodic orbit (see, for instance, [14]) and lemma 7. For the sake of completeness, we discuss its proof carefully.

Notice first that, thanks to the ergodic decomposition theorem, it is enough to suppose $\mu \in \mathcal{M}_\sigma$ ergodic. It seems convenient to recall that as usual we are considering the space of bounded real-valued functions on $\mathbf{\Sigma}$ and its subspaces equipped with the uniform norm. We take then a dense sequence $\{f_j\}_{j\geq 0}$ of bounded uniformly continuous real-valued functions on $\mathbf{\Sigma}$. Let $\Lambda_j \subseteq \mathbf{\Sigma}$ denote the set of points for which the Birkhoff's ergodic theorem holds for f_j as a μ-integrable function. Take then a point $\mathbf{z} \in \bigcap_{j\geq 0} \Lambda_j$. It is not difficult to see that the sequence of Borel probability measures

$$\nu_k := \frac{1}{k}\sum_{j=0}^{k-1} \delta_{\sigma^j(\mathbf{z})}$$

converges in the weak topology to μ.

Since $(\mathbf{\Sigma}, \sigma)$ is a finitely primitive Markov subshift, for every integer $k > K_0$, let $\mathbf{y}^k = (y_0^k, y_1^k, \ldots) \in \mathbf{\Sigma}$ be a periodic point of period k, with $y_j^k \in \mathbb{F}$ whenever $j \in \{k - K_0, \ldots, k-1\}$, such that $d(\mathbf{z}, \mathbf{y}^k) \leq \lambda^{k-K_0}$. Consider then the σ-invariant Borel probability measure

$$\mu_k := \frac{1}{k}\sum_{j=0}^{k-1} \delta_{\sigma^j(\mathbf{y}^k)} \in \mathcal{M}_\sigma .$$

Let $f\colon \mathbf{\Sigma} \to \mathbb{R}$ be a bounded function dependending on n coordinates, that is, satisfying $\mathrm{Var}_n(f) = 0$. Notice that (supposing $k > K_0 + n$)

$$\begin{aligned}\left|\int f\, d\mu_k - \int f\, d\nu_k\right| &= \frac{1}{k}\left|S_k f(\mathbf{y}^k) - S_k f(\mathbf{z})\right| \\ &\leq \frac{2}{k}(K_0 + n)\|f\|_\infty \to 0 \text{ as } k \to \infty.\end{aligned}$$

As functions depending on finitely many coordinates are dense among bounded uniformly continuous real-valued functions on $\mathbf{\Sigma}$, we conclude that the sequences $\{\mu_k\}$ and $\{\nu_k\}$ have the same weak limit μ. However, lemma 7 assures that, for each index k,

$$\int A\, d\mu_k \leq \beta_A(\hat{I}).$$

Thus, (4.7) follows just by passing to the limit.

A final remark

Notice that, in reality, the coerciveness of the potential was exactly used twice in our arguments. Indeed, the coercive condition was employed just to assure both inequalities (4.2) and (4.6).

Nevertheless, during the construction of the sub-action $u_A \in C^0(\mathbf{\Sigma})$ in the proof of proposition 4, its boundness was made explicit in (4.1). Therefore, one can easily adapted the demonstration of proposition 5 using this information and the fact that $\beta_A \geq \inf A|_{\bigcup_{i \in \mathbb{F}}[i]}$ in order to guarantee the following statement.

Proposition 8. *Let $(\mathbf{\Sigma}, \sigma)$ be a finitely primitive Markov subshift on a countable alphabet. Assume $A \colon \mathbf{\Sigma} \to \mathbb{R}$ is a bounded above and locally Hölder continuous potential. Suppose there exists an integer $\hat{I} > I_{\mathbb{F}}$ such that*

$$\sup A|_{\bigcup_{i>\hat{I}}[i]} < \inf A|_{\bigcup_{i \in \mathbb{F}}[i]} - \left[\operatorname{Var}(A) + K_0 \left(\sup A - \inf A|_{\bigcup_{i \in \mathbb{F}}[i]}\right)\right].$$

Then, $\operatorname{supp}(\mu) \subseteq \Sigma_{\hat{I}}$ *whenever $\mu \in \mathcal{M}_\sigma$ is an A-maximizing probability.*

Since lemma 7 is actually a consequence of inequality (4.6) and not of the coerciveness of the potential, one may now obtain a more general version of theorem 1, without necessarily imposing an asymptotic behavior to $\sup A|_{[i]}$. In fact, we have the following result.

Theorem 9. *Suppose $(\mathbf{\Sigma}, \sigma)$ is a finitely primitive Markov subshift on a countable alphabet. Let $A \colon \mathbf{\Sigma} \to \mathbb{R}$ be a bounded above and locally Hölder continuous potential. Assume the existence of an integer $\hat{I} > I_{\mathbb{F}}$ such that*

$$\sup A|_{\bigcup_{i>\hat{I}}[i]} < \inf A|_{\bigcup_{i \in \mathbb{F}}[i]} - \left[\operatorname{Var}(A) + K_0 \left(\sup A - \inf A|_{\bigcup_{i \in \mathbb{F}}[i]}\right)\right].$$

Then $\beta_A = \beta_A(\hat{I})$. Moreover, $\mu \in \mathcal{M}_\sigma$ is an A-maximizing probability if, and only if, $\operatorname{supp}(\mu) \subseteq \Omega(A, \hat{I})$.

We decided to discuss this generalized result at the end of the paper because the existence of $\hat{I}$ in the above statement seems to be just a technical assumption. Coerciveness, in turn, is compelling, as the works in thermodynamic formalism indicate. Besides, it is important to have in mind that certain maximizing probabilities can be seen as zero temperature limits of Gibbs-equilibrium states (see [8, 13]).

Finally, we would like to point out that inequality (4.6), which has proved to be so fundamental, is quite similar to the oscillation condition proposed in [9] (see definition 5.1 there). It is interesting to refind such a condition as a natural consequence of uniform oscillatory behaviour of calibrated sub-actions defined on compact approximations.

References

[1] J.P. Conze and Y. Guivarc'h. *Croissance des sommes ergodiques et principe variationnel*, manuscript, circa (1993).

[2] G. Contreras and R. Iturriaga. *Global minimizers of autonomous Lagrangians.* 22° Colóquio Brasileiro de Matemática, IMPA, Rio de Janeiro (1999).

[3] G. Contreras, A.O. Lopes and Ph. Thieullen. *Lyapunov minimizing measures for expanding maps of the circle.* Ergodic Theory and Dynamical Systems, **21** (2001), 1379–1409.

[4] A. Fathi. *The weak KAM theorem in Lagrangian dynamics*, book to appear, Cambridge University Press.

[5] E. Garibaldi and A.O. Lopes. *On the Aubry-Mather theory for symbolic dynamics.* Ergodic Theory and Dynamical Systems, **28** (2008), 791–815.

[6] E. Garibaldi, A.O. Lopes and Ph. Thieullen. *On calibrated and separating sub-actions.* Bulletin of the Brazilian Mathematical Society, **40** (2009), 577–602.

[7] O. Jenkinson. Ergodic optimization, *Discrete and Continuous Dynamical Systems, Series A*, **15** (2006), 197–224.

[8] O. Jenkinson, R.D. Mauldin and M. Urbański. *Zero temperature limits of Gibbs-equilibrium states for countable alphabet subshifts of finite type.* Journal of Statistical Physics, **119** (2005), 765–776.

[9] O. Jenkinson, R.D. Mauldin and M. Urbański. *Ergodic optimization for countable alphabet subshifts of finite type.* Ergodic Theory and Dynamical Systems, **26** (2006), 1791–1803.

[10] O. Jenkinson, R.D. Mauldin and M. Urbański. *Ergodic optimization for non-compact dynamical systems.* Dynamical Systems, **22** (2007), 379–388.

[11] A.O. Lopes and Ph. Thieullen. *Sub-actions for Anosov diffeomorfisms.* Astérisque, **287** (2003), 135–146.

[12] R.D. Mauldin and M. Urbański. *Graph directed Markov systems: geometry and dynamics of limit sets.* Cambridge University Press (2003).

[13] I.D. Morris. *Entropy for zero-temperature limits of Gibbs-equilibrium states for countable-alphabet subshifts of finite type.* Journal of Statistical Physics, **126** (2007), 315–324.

[14] K.R. Parthasarathy. *On the category of ergodic measures.* Illinois Journal of Mathematics, **5** (1961), 648–656.

[15] O.M. Sarig. *Thermodynamic formalism for countable Markov shifts.* Ergodic Theory and Dynamical Systems, **19** (1999), 1565–1593.

Rodrigo Bissacot
Departamento Matemática, UFMG
30161-970 Belo Horizonte, MG
BRAZIL

E-mail: rodrigo.bissacot@gmail.com

Eduardo Garibaldi
Departamento de Matemática, UNICAMP
13083-859 Campinas, SP
BRAZIL

E-mail: garibaldi@ime.unicamp.br