On the Continuity of the Projection Mapping from Strategic Measures to Occupation Measures in Absorbing Markov Decision Processes

Piunovskiy, Alexey; Zhang, Yi

Document Version
Publisher's PDF, also known as Version of record

Link to publication on Research at Birmingham portal

On the Continuity of the Projection Mapping from Strategic Measures to Occupation Measures in Absorbing Markov Decision Processes

Alexey Piunovskiy[1] · Yi Zhang[2]

Abstract

In this paper, we prove the following assertion for an absorbing Markov decision process (MDP) with the given initial distribution, which is also assumed to be semi-continuous: the continuity of the projection mapping from the space of strategic measures to the space of occupation measures, both endowed with their weak topologies, is equivalent to the MDP model being uniformly absorbing. An example demonstrates, among other interesting scenarios, that for an absorbing (but not uniformly absorbing) semi-continuous MDP with the given initial distribution, the space of occupation measures can fail to be compact in the weak topology.

Keywords Markov decision processes · Absorbing model · Continuity · Projection mapping

Mathematics Subject Classification 90C40 · 60J05

1 Introduction

In this paper, we consider a Markov decision process (MDP) with a Borel state space $\mathbf{X}$ and a Borel action space $\mathbf{A}$, both being endowed with their Borel σ-algebras. If there is an isolated absorbing state say 0 in the state space, then the MDP model is called absorbing for a given initial distribution P_0 if under each strategy, the expected hitting time to 0 is finite.

✉ Yi Zhang
   y.zhang.29@bham.ac.uk

   Alexey Piunovskiy
   piunov@liv.ac.uk

[1] Department of Mathematical Sciences, University of Liverpool, Liverpool, UK

[2] School of Mathematics, University of Birmingham, Birmingham B15 2TT, UK

Springer

In terms of occupation measures, understood as the total state-action frequencies on $(\mathbf{X} \setminus \{0\}) \times \mathbf{A}$, endowed with the product σ-algebra, an MDP is absorbing for the given initial distribution if the occupation measure of each strategy is a finite measure. Occupation measures are important for the study of optimal control problems of MDPs with total cost criteria because the performance measure of a strategy can be written as an integral of the cost function with respect to the occupation measure of the strategy. This turns the original MDP problem to a static optimization problem in the space $\mathcal{D}$ of occupation measures. Since for absorbing MDP models, $\mathcal{D}$ contains only finite measures, it is natural, as we do in this paper, to endow it with the weak topology, which is metrizable.

The wonderful and insightful paper [11], with many new ideas, intended to develop a rich theory for absorbing MDPs, in particular, for the occupation measures in such MDP models. A key property was used there, which asserts that if the MDP model is absorbing and semi-continuous, then the space of occupation measures is compact with respect to the weak topology, see [11, Lemma 4.7]. Here an MDP model is called semi-continuous if the action space $\mathbf{A}$ is compact, and the transition kernel $p(dy|x, a)$ is either set-wise continuous in $a \in \mathbf{A}$ for each $x \in \mathbf{X}$, or it is continuous with respect to the weak topology in $(x, a) \in \mathbf{X} \times \mathbf{A}$.

In order to prove this result, Feinberg and Rothblum [11] exploited the following fact: for semi-continuous MDP models, the space of strategic measures is compact in the weak topology, as established in [19], and the projection mapping O, carrying the strategic measure of a strategy to the occupation measure of the same strategy, is continuous.

In the present paper, by means of an example, see Example 2 below, we demonstrate that the aforementioned assertion in [11, Lemma 4.7] is inaccurate. This is due to the fact that in general, the projection mapping O may be not continuous for absorbing semi-continuous MDP models, see Theorem 1.

Our second contribution is as follows. We show that for an absorbing MDP model, provided that it is semi-continuous, the projection mapping O is continuous if and only if the MDP model is uniformly absorbing. The latter requires that, for the given initial distribution, the expected hitting time to 0 converges uniformly with respect to all strategies. In fact, the sufficiency part follows from the same reasoning as in the proof of [11, Lemma 4.7], by avoiding the minor error therein. This characterization of the continuity of the projection mapping O shows that the gap in [11] can be naturally closed if one further requires the MDP model there to be uniformly absorbing.

To the best of our knowledge, this definition of uniformly absorbing MDPs for a given initial distribution seems to appear firstly in [12]. In that paper, which focused on non-atomic MDP models, its equivalence to the continuity of the projection mapping O was not discussed.

A trivial example of uniformly absorbing MDP models is given by those discounted MDP models with a nonnegative discount factor strictly smaller than 1.

If the initial distribution is concentrated on a single state, and the state space is countable or finite, endowed with the discrete topology, then the MDP model, assumed to be semi-continuous, is uniformly absorbing if it possesses a uniform Lyapunov function, see Definition 6 below. MDP models with a uniform Lyapunov function were studied intensively in e.g., [1, 5, 14]. In particular, for such MDP models, for

every initial state, the MDP model is uniformly absorbing, see Proposition 1. However, if one considers a fixed initial distribution, then the existence of a uniform Lyapunov function does not imply the model to be uniformly absorbing or even absorbing. This is demonstrated in Example 1, and is sometimes overlooked. Similar models for MDPs with Borel state and action spaces were considered in e.g., [15] and [13, Chapter 9], which assumed that the total value of occupation measures is bounded or w-bounded over all initial states and strategies, and focused on the optimality equation.

Let us mention that, apart from providing an absorbing semi-continuous MDP model with the space of occupation measures being not compact, Example 2 incidentally demonstrates several other scenarios of interest. E.g., it also demonstrates that an absorbing MDP model for a fixed initial state may not possess a uniform Lyapunov function, and that some known solvability conditions are important.

The rest of this paper is organized as follows. We describe the MDP model in Sect. 2, and recall some known facts in Sect. 3. The main results are presented in Sect. 4. The main example, i.e., Example 2, is formulated in Sect. 5, where we also discuss it in the context of optimal control problems of MDPs. The proofs of all the statements (except those known ones) are in Sect. 6. Finally, this paper is finished with a conclusion in Sect. 7.

2 Model Description

Before describing the MDP model, let us fix some notation and conventions used throughout this paper. If a space $\mathbf{Y}$ is discrete (with the discrete topology) and μ is a measure on $\mathcal{B}(\mathbf{Y})$, then, for singletons, we use notation $\mu(y)$, not $\mu(\{y\})$. Integral of a function $c(\cdot)$ with respect to a measure μ is written as $\int_{\mathbf{Y}} c(y)d\mu(y)$ or $\int_{\mathbf{Y}} c(y)\mu(dy)$. In $\mathbb{R} = (-\infty, \infty)$, unless stated otherwise, the usual Euclidean topology is fixed. $\delta_a(dx)$ is the Dirac measure concentrated at the point a, and $\mathbb{I}\{\cdot\}$ is the indicator function.

We fix the following primitives of an MDP model.

- $\mathbf{X}$ and $\mathbf{A}$ are the state and action spaces, both assumed to be nonempty topological Borel spaces. Here a topological Borel space is a Borel subset of some complete metrizable separable space. We endow $\mathbf{X}$ and $\mathbf{A}$ with their Borel σ-algebras $\mathcal{B}(\mathbf{X})$ and $\mathcal{B}(\mathbf{A})$, respectively.
- The transition probability $p(dy|x, a)$ is a (measurable) stochastic kernel on $\mathcal{B}(\mathbf{X})$ given $\mathbf{X} \times \mathbf{A}$.

Given the collection $\{\mathbf{X}, \mathbf{A}, p\}$, for each strategy (see Definition 1 below) and initial distribution on $\mathbf{X}$, one can construct a probability space, and define the controlled (state) process $\{X_t\}_{t=0}^{\infty}$ and controlling (action) process $\{A_t\}_{t=1}^{\infty}$ thereon. Derman in his classic book [6, p.4] termed the bivariate process $\{(X_t, A_{t+1})\}_{t=0}^{\infty}$ a Markov decision process. This is why we refer to $\{\mathbf{X}, \mathbf{A}, p\}$ as the MDP model. The aforementioned construction is described next.

The space of trajectories (or say histories) is the following countable product

$$\mathbf{H} := \mathbf{X} \times (\mathbf{A} \times \mathbf{X})^{\infty}.$$

The generic notation for an element of $\mathbf{H}$ is $\omega = (x_0, a_1, x_1, \ldots) \in \mathbf{H}$. We endow $\mathbf{H}$ with the product topology and the product σ-algebra, which is also the Borel σ-algebra on it. All the random variables, like X_t and A_{t+1}, are just measurable mappings defined on $\mathbf{H}$:

$$X_t(\omega) = x_t, A_{t+1}(\omega) = a_{t+1}, \ \forall \, t \geq 0.$$

Definition 1 A control strategy, or simply say a strategy, $\pi = \{\pi_t\}_{t=1}^{\infty}$ is a sequence of stochastic kernels $\pi_t(da|x_0, a_1, x_1, \ldots, a_{t-1}, x_{t-1})$ on $\mathcal{B}(\mathbf{A})$ given $\mathbf{H}_{t-1} := (\mathbf{X} \times \mathbf{A})^{t-1} \times \mathbf{X}$. A strategy π is called Markov, if for each $t \geq 1$, there is a stochastic kernel π_t^M on $\mathcal{B}(\mathbf{A})$ given $\mathbf{X}$ such that $\pi_t(da|x_0, a_1, x_1, \ldots, a_{t-1}, x_{t-1}) = \pi_t^M(da|x_{t-1})$. If, for some stochastic kernel π^s on $\mathcal{B}(\mathbf{A})$ given $\mathbf{X}$, $\pi_t(da|x_0, a_1, x_1, \ldots, a_{t-1}, x_{t-1}) = \pi^s(da|x_{t-1})$ for any $t = 1, 2, \ldots$, then the strategy $\pi = \{\pi_t\}_{t=1}^{\infty}$ is called stationary, and is identified with and denoted as π^s. If a stationary strategy π^s takes the form $\pi^s(da|x) = \delta_{\varphi(x)}(da)$ for some measurable mapping φ from $\mathbf{X}$ to $\mathbf{A}$, where $\delta_{\varphi(x)}(da)$ denotes the Dirac measure concentrated on the singleton $\{\varphi(x)\}$, then the strategy π^s is called deterministic stationary, and is identified with and denoted by the underlying measurable mapping φ.

The set of all strategies is denoted as Δ^{All}, and the set of Markov strategies is denoted as Δ^{Markov}.

Given a strategy $\pi = \{\pi_t\}_{t=1}^{\infty}$, for a fixed state-action pair $(z, b) \in \mathbf{X} \times \mathbf{A}$, we define its shifted strategy $^{(z,b)}\pi = \{^{(z,b)}\pi_t\}_{t=1}^{\infty}$ by

$$^{(z,b)}\pi_t(da|y_0, a_1, \ldots, a_{t-1}, y_{t-1}) := \pi_{t+1}(da|z, b, y_0, a_1, \ldots, a_{t-1}, y_{t-1}) \quad (1)$$

for $y_i \in \mathbf{X}$ with $i \in \{0, 1, \ldots, t-1\}$ and $a_i \in \mathbf{A}$ with $i \in \{1, 2, \ldots, t-1\}$.

Let the initial distribution P_0 on $(\mathbf{X}, \mathcal{B}(\mathbf{X}))$ be given. If a strategy π is also fixed, then the strategic measure on $(\mathbf{H}, \mathcal{B}(\mathbf{H}))$, constructed in the standard way using the Ionescu-Tulcea Theorem, is denoted as $\mathsf{P}_{\mathsf{P}_0}^{\pi}$. It is the unique probability measure on $(\mathbf{H}, \mathcal{B}(\mathbf{H}))$ such that

$$\mathsf{P}_{\mathsf{P}_0}^{\pi}(X_0 \in dx) = \mathsf{P}_0(dx);$$
$$\mathsf{P}_{\mathsf{P}_0}^{\pi}(A_{t+1} \in da|X_0, A_1, \ldots, X_t) = \pi_{t+1}(da|X_0, A_1, \ldots, X_t);$$
$$\mathsf{P}_{\mathsf{P}_0}^{\pi}(X_{t+1} \in dx|X_0, A_1, \ldots, X_t, A_{t+1}) = p(dx|X_t, A_{t+1}) \ \forall \, t \geq 0.$$

The corresponding mathematical expectation is denoted as $\mathsf{E}_{\mathsf{P}_0}^{\pi}$. If the initial distribution $\mathsf{P}_0(dx) = \delta_{x_0}(dx)$ is degenerate, we use notations $\mathsf{P}_{x_0}^{\pi}$ and $\mathsf{E}_{x_0}^{\pi}$.

We denote by

$$\mathcal{P} := \{\mathsf{P}_{\mathsf{P}_0}^{\pi} : \pi \in \Delta^{\text{All}}\}$$

the space of all strategic measures with the given initial distribution P_0.

An important class of MDP models is the absorbing model defined as follows.

Definition 2 The MDP is called absorbing (at 0) for the given initial distribution P_0 if there is an isolated state, say 0 in $\mathbf{X}$ such that 0 is absorbing, i.e.,

$$p(\{0\}|0, a) \equiv 1,$$

and

$$\mathsf{E}^\pi_{P_0}[T_0] < \infty, \ \forall \, \pi \in \Delta^{\text{All}},$$

where

$$T_0 := \inf\{t \geq 0 : \ X_t = 0\}$$

denotes the hitting time to the state 0 by the controlled process. As usual, $\inf \emptyset := \infty$.

Verbally, an MDP model is absorbing (at 0) for the initial distribution P_0 if under each strategy, the expected hitting time to the isolated absorbing state 0 is finite. The above definition of an absorbing MDP model for the initial distribution P_0 is the same as the one given by Feinberg and Rothblum in [11, p.7]. Given that the state 0 is absorbing, this definition also coincides with the one in [1, Definition 7.1]. In general, in [1, Definition 7.1], for the model to be absorbing for P_0, the state 0 was not required to be absorbing itself, and it was required that $\mathsf{E}^\pi_{P_0}[\tau_0] < \infty, \ \forall \, \pi \in \Delta^{\text{All}}$, where

$$\tau_0 := \inf\{t \geq 1 : \ X_t = 0\}$$

is the return time to state 0.

Thus, for an absorbing MDP model (at state 0) for the initial distribution P_0, for each strategy, the series $\sum_{t=0}^\infty \mathsf{E}^\pi_{P_0}[\mathbb{I}\{T_0 > t\}]$ converges, because $\sum_{t=0}^\infty \mathsf{E}^\pi_{P_0}[\mathbb{I}\{T_0 > t\}] = \mathsf{E}^\pi_{P_0}[T_0] < \infty$. If we require the convergence of the above series to be uniform with respect to all strategies $\pi \in \Delta^{\text{All}}$, then the resulting model will be called uniformly absorbing, formulated in the next definition.

Definition 3 An absorbing (at 0) MDP model for the initial distribution P_0 is called *uniformly absorbing* for P_0 if

$$\lim_{n \to \infty} \sup_{\pi \in \Delta^{\text{All}}} \mathsf{E}^\pi_{P_0} \left[\sum_{t=n}^\infty \mathbb{I}\{t < T_0\} \right] = 0.$$

This definition was given in [12, Definition 3.6]. Since the state 0 was required to be absorbing in Definition 2, the requirement in Definition 3 is the same as

$$\lim_{n \to \infty} \sup_{\pi \in \Delta^{\text{All}}} \mathsf{E}^\pi_{P_0} \left[\sum_{t=n}^\infty \mathbb{I}\{X_t \in (\mathbf{X} \setminus \{0\})\} \right] = 0 \tag{2}$$

because

$$
\mathsf{E}_{\mathsf{P}_0}^{\pi} \left[\sum_{t=n}^{\infty} \mathbb{I}\{t < T_0\} \right] = \mathsf{E}_{\mathsf{P}_0}^{\pi} \left[\sum_{t=n}^{\infty} \mathbb{I}\{X_t \in (\mathbf{X} \setminus \{0\})\} \right], \ \forall \, n \geq 0.
$$

This equality, when specialized to $n = 0$, allows us to formulate the definition of absorbing MDP models in terms of the finiteness of the occupation measures, defined as follows.

Definition 4 Consider an MDP model $\{\mathbf{X}, \mathbf{A}, p\}$ with $0 \in \mathbf{X}$ being an isolated absorbing state, i.e., $p(\{0\}|0, a) \equiv 1$. The occupation measure $\eta_{\mathsf{P}_0}^{\pi}$ of a strategy π for the given initial distribution P_0 is the $[0, \infty]$-valued measure on $((\mathbf{X} \setminus \{0\}) \times \mathbf{A}, \mathcal{B}((\mathbf{X} \setminus \{0\}) \times \mathbf{A}))$ defined by the formula

$$
\eta_{\mathsf{P}_0}^{\pi}(\Gamma^{\mathbf{X}} \times \Gamma^{\mathbf{A}}) := \sum_{t=1}^{\infty} \mathsf{P}_{\mathsf{P}_0}^{\pi}(X_{t-1} \in \Gamma^{\mathbf{X}}, A_t \in \Gamma^{\mathbf{A}}), \ \forall \, \Gamma^{\mathbf{X}} \in \mathcal{B}(\mathbf{X} \setminus \{0\}), \ \Gamma^{\mathbf{A}} \in \mathcal{B}(\mathbf{A}).
$$

If the initial distribution is concentrated on a singleton, say $\{x_0\}$, then we write $\eta_{x_0}^{\pi}$ for the occupation measure of a strategy π. Let

$$
\mathcal{D} := \{\eta_{\mathsf{P}_0}^{\pi} : \pi \in \Delta^{\mathrm{All}}\}
$$

be the space of all occupation measures for the initial distribution P_0.

Now we see that an MDP model with an absorbing state 0 is absorbing (at 0) for the initial distribution P_0 if and only if

$$
\eta_{\mathsf{P}_0}^{\pi}((\mathbf{X} \setminus \{0\}) \times \mathbf{A}) < \infty
$$

for each $\pi \in \Delta^{\mathrm{All}}$.

For brevity, we make the next definition.

Definition 5 An MDP model $\{\mathbf{X}, \mathbf{A}, p\}$ is called semi-continuous (S) if the following two conditions are satisfied:

(a) The action space $\mathbf{A}$ is a compact topological Borel space.
(b) For each $x \in \mathbf{X}$, the function $\int_{\mathbf{X}} u(y) p(dy|x, a)$ is continuous in $a \in \mathbf{A}$ for every bounded measurable function $u(\cdot)$ on $\mathbf{X}$.

An MDP model $\{\mathbf{X}, \mathbf{A}, p\}$ is called semi-continuous (W) if the following two conditions are satisfied:

(c) The action space $\mathbf{A}$ is a compact topological Borel space.
(d) The function $\int_{\mathbf{X}} u(y) p(dy|x, a)$ is continuous in $(x, a) \in \mathbf{X} \times \mathbf{A}$ for every bounded continuous function $u(\cdot)$ on $\mathbf{X}$.

An MDP model $\{\mathbf{X}, \mathbf{A}, p\}$ is called semi-continuous if it is either semi-continuous (S) or semi-continuous (W).

Conditions (a,b) in Definition 5 are respectively called Condition (S) (1,2) in [19, Section 6], whereas conditions (c,d) in Definition 5 are respectively called Condition (W) (1,2) in [19, Section 5]. If $\mathbf{X}$ is countable or finite, and is endowed with the discrete topology, then semi-continuity (S) and semi-continuity (W) mean the same. Admittedly, the term of "semi-continuous MDP models" has other meanings in the literature. Nevertheless, in the present paper, its meaning is unambiguous.

3 Some Relevant Facts

In this section we present some facts about MDP models and optimal control problems of MDPs.

3.1 Facts About MDP Model

3.1.1 MDP with a Uniform Lyapunov Function

An important class of semi-continuous MDP models with a countable or finite state space that are uniformly absorbing (at 0) for a given initial state x_0 is given by those that admit a uniform Lyapunov function, defined as follows.

Definition 6 Consider a semi-continuous MDP model with a countable or finite state space $\mathbf{X}$ (endowed with the discrete topology) with the isolated absorbing state 0. A $[1, \infty)$-valued function $\mu(\cdot)$ on $\mathbf{X}$ is said to be a uniform Lyapunov function if the following conditions are satisfied:

(a) $1 + \sum_{y \in \mathbf{X} \setminus \{0\}} p(y|x, a)\mu(y) \leq \mu(x), \ \forall \ x \in \mathbf{X}, a \in \mathbf{A}$.
(b) For each $x \in \mathbf{X}$, the mapping $a \in \mathbf{A} \to \sum_{y \in \mathbf{X} \setminus \{0\}} p(y|x, a)\mu(y)$ is continuous.
(c) For each $x \in \mathbf{X}$ and each deterministic stationary strategy φ, $\lim_{t \to \infty} \mathsf{E}_x^\varphi[\mu(X_t) \times \mathbb{I}\{\tau_0 > t\}] = 0$, where we recall that $\tau_0 := \min\{t \geq 1 : X_t = 0\}$.

The above definition of a uniform Lyapunov function was taken from [1, Definition 7.4], see also [5, Definition 4.2]. In both of these two references, this definition is ascribed to [14]. Many characterizations and consequences of a semi-continuous MDP model with a uniform Lyapunov function can be found in [1, 5], among which is the following one, whose proof can be found on p.107 of [1].

Proposition 1 *Consider a semi-continuous MDP model with a countable or finite state space $\mathbf{X}$ (endowed with the discrete topology) with the isolated absorbing state 0. If there exists a uniform Lyapunov function $\mu(\cdot)$, then the MDP model is uniformly absorbing (at 0) for each initial state x_0.*

For a semi-continuous MDP model with a countable or finite state space $\mathbf{X}$ (endowed with the discrete topology) with the isolated absorbing state 0, if there exists a uniform Lyapunov function $\mu(\cdot)$, it can happen that for some initial distribution P_0, the MDP model is not absorbing (at 0) for P_0. Sometimes, this is overlooked. We present an example to illustrate this.

Example 1 Consider an MDP model $\{\mathbf{X}, \mathbf{A}, p\}$ with $\mathbf{X} = \{0, 1, \dots \}$, endowed with the discrete topology, and $\mathbf{A}$ being a singleton, so that we will omit the argument $a \in \mathbf{A}$ everywhere in this example, and $p(0|x) = \frac{1}{2^x} = 1 - p(x|x)$ for all $x \in \{1, 2, \dots \}$ and $p(0|0) = 1$. There is only one strategy say π in this model. Then for each $x_0 \in \{1, 2, \dots \}$,

$$\mathsf{E}_{x_0}^{\pi}[T_0] = 2^{x_0}.$$

hence, this MDP model is absorbing, in fact uniformly absorbing (at 0) for every initial state x_0. It has a uniform Lyapunov function given by $\mu(0) = 1$ and $\mu(x) = 2^x$ for $x \geq 1$. Indeed, (a,b) in Definition 6 can be checked to be satisfied by $\mu(\cdot)$, and for (c), note for each $x \geq 1$ that

$$\mathsf{E}_x^{\pi}[\mu(X_n)\mathbb{I}\{\tau_0 > n\}] = \mathsf{E}_x^{\pi}[2^x \mathbb{I}\{X_n = x\}] = 2^x \mathsf{P}_x^{\pi}(X_n = x) = 2^x \left(1 - \frac{1}{2^x}\right)^n \to 0$$

as $n \to \infty$, and the convergence also takes place for $x = 0$. Now if we take the initial distribution as the geometric distribution on $\{1, 2, \dots \}$ with parameter $\frac{2}{5}$, then $\mathsf{E}_{\mathsf{P}_0}[T_0] = \sum_{i=1}^{\infty} \frac{2}{5}(\frac{3}{5})^{i-1} 2^i = \infty$.

3.1.2 Projection Mapping and the Topologies on $\mathcal{D}$ and $\mathcal{P}$

Consider an absorbing (at 0) MDP model for the initial distribution P_0. Let the projection mapping O from $\mathcal{P}$ to $\mathcal{D}$ be defined for each $\mathsf{P} \in \mathcal{P}$ by

$$O(\mathsf{P}) := \sum_{n=0}^{\infty} \int_{\mathbf{H}} \mathbb{I}\{(x_n, a_{n+1}) \in dx \times da\} \mathsf{P}(d\omega)$$

on $\mathcal{B}((\mathbf{X} \setminus \{0\}) \times \mathbf{A})$. In particular,

$$O(\mathsf{P}_{\mathsf{P}_0}^{\pi}) = \eta_{\mathsf{P}_0}^{\pi}. \tag{3}$$

The main results of this paper concern the continuity of the projection mapping O from $\mathcal{P}$ to $\mathcal{D}$. Therefore, we endow $\mathcal{D}$ and $\mathcal{P}$ with suitable topologies, described as follows.

Consider an absorbing (at 0) MDP model for the initial distribution P_0. Then $\mathcal{D}$ is a subset of the space $\mathcal{M}((\mathbf{X} \setminus \{0\}) \times \mathbf{A})$ of finite measures on $((\mathbf{X} \setminus \{0\}) \times \mathbf{A}, \mathcal{B}((\mathbf{X} \setminus \{0\}) \times \mathbf{A}))$. We endow $\mathcal{M}((\mathbf{X} \setminus \{0\}) \times \mathbf{A})$ with the weak topology generated by the set of bounded continuous functions defined on $(\mathbf{X} \setminus \{0\}) \times \mathbf{A}$, i.e., the coarsest topology on $\mathcal{M}((\mathbf{X} \setminus \{0\}) \times \mathbf{A})$ with respect to which, for each bounded continuous function $g(\cdot)$ on $(\mathbf{X} \setminus \{0\}) \times \mathbf{A}$, $\int_{(\mathbf{X} \times \{0\}) \times \mathbf{A}} g(x, a)\eta(dx \times da)$ is continuous in $\eta \in \mathcal{M}((\mathbf{X} \setminus \{0\}) \times \mathbf{A})$. Since $\mathbf{X} \setminus \{0\}$ and $\mathbf{A}$ are Borel spaces, according to Theorem 8.9.4 (i) of [3] and [4, p.49], when endowed with the weak topology, $\mathcal{M}((\mathbf{X} \setminus \{0\}) \times \mathbf{A})$ is metrizable, and its topological subspace $\mathcal{D}$ is metrizable, too. Indeed, a popular topology on $\mathcal{D}$ is this weak topology, when $\mathcal{D} \subseteq \mathcal{M}((\mathbf{X} \setminus \{0\}) \times \mathbf{A})$.

Recall also that the so called ws-topology on $\mathcal{M}((\mathbf{X} \setminus \{0\}) \times \mathbf{A})$ is the coarsest topology with respect to which, for each bounded Carathéodory function $g(\cdot)$ on $(\mathbf{X} \setminus \{0\}) \times \mathbf{A}$ (meaning that $g(x, \cdot)$ is continuous for each fixed $x \in \mathbf{X} \setminus \{0\}$ and $g(\cdot, a)$ is measurable for each fixed $a \in \mathbf{A}$), $\int_{(\mathbf{X} \setminus \{0\}) \times \mathbf{A}} g(x, a)\eta(dx \times da)$ is continuous in $\eta \in \mathcal{M}((\mathbf{X} \setminus \{0\}) \times \mathbf{A})$. Clearly, the weak topology is coarser than the ws-topology.

On the space $\mathcal{P}$ (for the fixed initial distribution P_0), one can consider the so called ws$^\infty$-topology of Schäl, introduced in [19], see p.359 therein. This is the coarsest topology on $\mathcal{P}$, with respect to which, for each integer $0 \leq T < \infty$ and each bounded measurable function $f(h_T) = f(x_0, a_1, x_1, \ldots, a_T, x_T)$ continuous in $(a_1, a_2, \ldots, a_T)$ under arbitrarily fixed $x_0, x_1, \ldots, x_T \in \mathbf{X}$, the mapping $\mathsf{P} \in \mathcal{P} \to \int_{\mathbf{H}_T} f(h_T)\mathsf{P}_T(dh_T)$ is continuous. Here $\mathbf{H}_0 = \mathbf{X}$, and $\mathbf{H}_T = (\mathbf{X} \times \mathbf{A})^T \times \mathbf{X}$, and P_T denotes the marginal of $\mathsf{P} \in \mathcal{P}$ on $\mathbf{H}_T$, $T \geq 0$ being an integer.

The ws$^\infty$-topology works particularly well with semi-continuous (S) MDP models. One of the main results in [19] is the following compactness result, whereas for semi-continuous (W) MDP models, the weak topology on $\mathcal{P}$ is more convenient. What actually happens is that for semi-continuous (S) MDP models, on $\mathcal{P}$, the ws$^\infty$-topology is the same as the weak topology, see Proposition 3 below.

Proposition 2 *Consider the MDP model* $\{\mathbf{X}, \mathbf{A}, p\}$ *with some initial distribution* P_0. *If the MDP model is semi-continuous (S), then* $\mathcal{P}$ *endowed with the ws$^\infty$-topology is compact, whereas if the MDP model is semi-continuous (W), then* $\mathcal{P}$ *endowed with the weak topology is compact.*

Proof See [19, Theorem 6.6] for the first assertion, and [19, Theorem 5.6] for the second assertion. □

Nowak further studied the ws$^\infty$-topology on $\mathcal{P}$ in [17], and proved the following useful result.

Proposition 3 *Suppose that the MDP model* $\{\mathbf{X}, \mathbf{A}, p\}$ *is semi-continuous (S). Let some initial distribution* P_0 *be given. Then on* $\mathcal{P}$, *the ws$^\infty$-topology coincides with the weak topology. Consequently, the space* $\mathcal{P}$ *endowed with the ws$^\infty$-topology is metrizable, and compact in the weak topology, in view of Proposition 2.*

Proof This follows from [17, Theorem 1]. A simpler proof of [17, Theorem 1] was given in [16]. □

In view of Propositions 2 and 3, we will always consider the weak topology on $\mathcal{P}$ when we discuss semi-continuous MDP models, for which, $\mathcal{P}$ is always compact.

3.2 Facts About Optimal Control Problems

Now we add an additional element to the MDP model, which is the cost function $c(\cdot)$. This is a measurable function from $(\mathbf{X} \times \mathbf{A}, \mathcal{B}(\mathbf{X} \times \mathbf{A}))$ to $[-\infty, \infty]$ in general.

Define for each strategy π and an initial distribution P_0

$$v^\pi(\mathsf{P}_0) := \mathsf{E}_{P_0}^\pi \left[\sum_{t=1}^\infty c(X_{t-1}, A_t) \right]$$

$$:= \mathsf{E}_{P_0}^\pi \left[\sum_{t=1}^\infty c^+(X_{t-1}, A_t) \right]$$

$$- \mathsf{E}_{P_0}^\pi \left[\sum_{t=1}^\infty c^-(X_{t-1}, A_t) \right],$$

where $c^+(x, a) := \max\{c(x, a), 0\}$, $c^-(x, a) = \max\{-c(x, a), 0\}$, and we accept that $\infty - \infty := \infty$. When $\mathsf{P}_0(dx) = \delta_{x_0}(dx)$ for some $x_0 \in \mathbf{X}$, then we write $v^\pi(x_0)$ instead of $v^\pi(\mathsf{P}_0)$.

For future reference, let

$$v^*(x) := \inf_{\pi \in \Delta^{\text{All}}} v^\pi(x) \tag{4}$$

A strategy π^* is called optimal for P_0 if it solves the following optimal control problem:

$$\text{Minimize over } \pi \in \Delta^{\text{All}} : v^\pi(\mathsf{P}_0), \tag{5}$$

i.e.,

$$v^{\pi^*}(\mathsf{P}_0) = \min_{\pi \in \Delta^{\text{All}}} v^\pi(\mathsf{P}_0).$$

In the next statement, sufficient conditions for the existence of an optimal strategy for P_0 are given. See also [7] and the references therein.

Proposition 4 *Suppose that the MDP model is semi-continuous (S), and $c(\cdot)$ is bounded below and $(-\infty, \infty]$-valued on $\mathbf{X} \times \mathbf{A}$ such that $c(x, a)$ is lower semicontinous in $a \in \mathbf{A}$ for each $x \in \mathbf{X}$. If for the given initial distribution P_0,*

$$\mathsf{E}_{P_0}^\pi \left[\sum_{t=1}^\infty c^-(X_{t-1}, A_t) \right] < \infty \ \forall \ \pi \in \Delta^{\text{All}}, \tag{6}$$

and

$$\inf_{N \ge n} \inf_{\pi \in \Delta^{\text{All}}} \sum_{t=n+1}^N \mathsf{E}_{P_0}^\pi \left[c(X_{t-1}, A_t) \right] \to 0 \ \text{ as } n \to \infty, \tag{7}$$

then there exists an optimal strategy for P_0. The same assertion holds if the MDP model is semi-continuous (W), and $c(\cdot)$ is a bounded below $(-\infty, \infty]$-valued lower semicontinous function on $\mathbf{X} \times \mathbf{A}$.

Proof The proof of the first assertion follows from [19, Theorems 2.2 and 6.6]. The last assertion holds by [19, Theorems 2.2 and 5.6]. □

In [19], condition (7) is called (C), and condition (6) is referred to as "General Assumption" (GA).

For an absorbing (at 0) MDP model for P_0, we will take $c(\cdot)$ such that $c(0, a) \equiv 0$ at the absorbing isolated point 0. Then 0 will be referred to as the costless cemetery. If Condition (GA) is satisfied, then the optimal control problem (5) can be rewritten in terms of occupation measures in the following way:

$$\text{Minimize over } \eta \in \mathcal{D}\text{:} \int_{\mathbf{X}\setminus\{\mathbf{0}\}\times\mathbf{A}} c(y, a)\eta(dy \times da),$$

This is one of the main reasons for studying $\mathcal{D}$ in the literature of MDPs.

4 Main Results

In the wonderful and insightful paper [11], a rich theory for absorbing MDP models was developed. In that paper, the following assertion was claimed, see the proof of [11, Lemma 4.7]:

- If the MDP model $\{\mathbf{X}, \mathbf{A}, p\}$ is semi-continuous, and is absorbing (at 0) for the given initial distribution P_0, $\mathcal{P}$ is endowed with the weak topology (see Proposition 3 and the paragraph below it), and $\mathcal{D}$ is endowed with the weak topology generated by bounded continuous functions on $(\mathbf{X} \setminus \{0\}) \times \mathbf{A}$, then the projection mapping O defined in (3) is continuous.

If the above claim was true, then by Propositions 2 and 3 as well as the fact $\mathcal{D} = O(\mathcal{P})$, it would then follow that $\mathcal{D}$ is compact, provided that the MDP model $\{\mathbf{X}, \mathbf{A}, p\}$ is semi-continuous, and is absorbing (at 0) for P_0. This assertion was formulated as [11, Lemma 4.7], and was used several times therein.

The first main result of this paper is that we show by means of an example that the above claims regarding the continuity of O and the compactness of $\mathcal{D}$ are false.

Theorem 1 *There is an MDP model $\{\mathbf{X}, \mathbf{A}, p\}$ such that the following assertions hold for a fixed initial distribution P_0:*

(a) The MDP model is semi-continuous. Consequently, the space of all strategic measures $\mathcal{P}$ for the initial distribution P_0 is compact in the weak topology.
(b) The MDP model is absorbing (at 0) for P_0.
(c) The MDP model is not uniformly absorbing (at 0) for P_0.
(d) $\mathcal{D}$ is not compact with respect to the weak topology.
(e) The mapping O defined by (3) is not continuous between $\mathcal{P}$ and $\mathcal{D}$, both of which are endowed with their weak topologies.
(f) The MDP model does not possess a uniform Lyapunov function.

The proof of this theorem is given in Subsection 6.1.

Since the continuity of the projection mapping O defined by (3) from $\mathcal{P}$ (endowed with the weak topology) to $\mathcal{D}$ (endowed with the weak topology) played an important role in the reasoning in [11], and Theorem 1 shows that this is not guaranteed if the
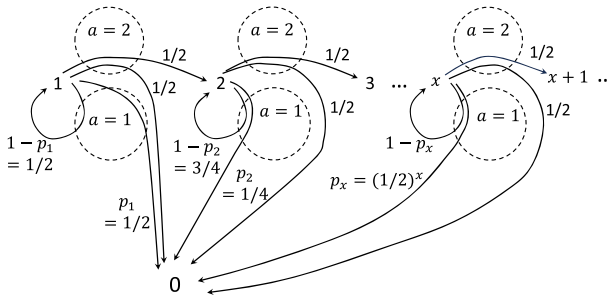
Fig. 1 Graphical representation of the MDP

MDP model is semi-continuous and absorbing (at 0) for P_0, we next investigate when it holds. Our second main result asserts that for a semi-continuous MDP model, provided that it is absorbing for P_0, the aforementioned continuity of the projection mapping O holds if and only if the MDP model is uniformly absorbing for P_0.

Theorem 2 *Consider a semi-continuous MDP model with the fixed initial distribution* P_0. *Suppose that the MDP model is absorbing (at* 0*) for* P_0. *Then the projection mapping O defined by (3) from $\mathcal{P}$ (endowed with the weak topology) to $\mathcal{D}$ (endowed with the weak topology) is continuous if and only if the MDP model is uniformly absorbing (at* 0*) for* P_0.

The proof is given in Subsection 6.2.

5 Example

Theorem 1 will be proved by means of the following MDP model.

Example 2 The elements of the MDP model $\{\mathbf{X}, \mathbf{A}, p\}$ are as follows.

- The state and action spaces are $\mathbf{X} := \{0, 1, 2, \ldots\}$ and $\mathbf{A} := \{1, 2\}$, which are equipped with the discrete topology.
- 0 is the isolated state, which is absorbing: $p(0|0, a) \equiv 0$.
- For $x = 1, 2, \ldots$, $p(0|x, 1) = p_x = \left(\frac{1}{2}\right)^x$ and $p(x|x, 1) = 1 - p_x$.
- For $x = 1, 2, \ldots$, $p(0|x, 2) = \frac{1}{2}$ and $p(x + 1|x, 2) = \frac{1}{2}$.
- Other transition probabilities equal zero.

The transition diagram of the MDP model in Example 2 is given in Fig. 1.

The MDP model in Example 2 is explored in the proof of Theorem 1. Incidentally, it can also be used to demonstrate that condition (7) is important for the solvability, see the next proposition.

Proposition 5 *Consider the MDP model in Example 2 with the initial distribution* $\mathsf{P}_0(dx) = \delta_1(dx)$, *and the optimal control problem (5) for the cost function $c(\cdot)$ defined by $c(x, a) \equiv -1$ for all $x \geq 1$, and $c(0, a) \equiv 0$. Then the following assertions hold.*

(a) *There is no optimal strategy for the given* P_0.

(b) *Condition (7) is not satisfied, whereas all the other conditions in Proposition 4 are satisfied.*

The proof of this Proposition is given in Subsection 6.3.

6 Proof of the Statements

In this section we provide the detailed proofs of Theorem 1, Theorem 2 and Proposition 5.

6.1 Proof of Theorem 1

Proof of Theorem 1 Throughout this proof we consider the MDP model in Example 2. Unless stated otherwise, we fix the initial distribution P_0 concentrated on the singleton $\{1\}$, i.e.,

$$\mathsf{P}_0(1) = 1.$$

(a) This MDP is vacuously semi-continuous, and the space of all strategic measures $\mathcal{P} = \{\mathsf{P}_{\mathsf{P}_0}^\pi, \ \pi \in \Delta^{\text{All}}\}$ is compact in the weak topology by Propositions 2 and 3.

(b) Consider the function $w^*(\cdot)$ on $\mathbf{X} = \{0, 1, \dots\}$ defined by

$$w^*(x) := \sup_{\pi \in \Delta^{\text{All}}} \mathsf{E}_x^\pi[T_0] = \sup_{\pi \in \Delta^{\text{All}}} \mathsf{E}_x^\pi \left[\sum_{t=0}^\infty \mathbb{I}\{X_t \in \{1, 2, \dots\}\} \right] \qquad (8)$$
$$x \in \{0, 1, \dots\}.$$

We shall show that $w^*(x) < \infty$ for each $x \in \{0, 1, \dots\}$. This would in particular, imply that $\mathsf{E}_1^\pi[T_0] < \infty$ for all $\pi \in \Delta^{\text{All}}$, i.e., the MDP model is absorbing (at 0) for the fixed P_0 concentrated on $\{1\}$.

According to [2, Propositions 9.8 and 9.10], $w^*(\cdot)$ is the minimal nonnegative solution to the Bellman equation

$$w^*(0) = 0;$$
$$w^*(x) = 1 + \max \left\{ (1 - p_x)w^*(x), \ \frac{1}{2}w^*(x+1) \right\}, \quad x = 1, 2, \dots. \qquad (9)$$

We claim that $w^*(0) = 0$ and

$$w^*(x) \leq 2 + 2^x \ \forall \ x \in \{1, 2, \dots\}. \qquad (10)$$

To see this, it suffices to note that the function $v(\cdot)$ on $\{0, 1, \dots\}$ given by $v(x) := 2 + 2^x$ and $v(0) = 0$ solves the Bellman equation (9), and this is true because, $v(0) = 0$

trivially satisfies the first equality in (9), whereas for $x \in \{1, 2, \dots\}$,

$$
1 + \max \left\{ \left[1 - \left(\frac{1}{2} \right)^x \right] [2 + 2^x]; \ \frac{1}{2} (2 + 2^{x+1}) \right\}
$$

$$
= 1 + \max \left\{ 1 + 2^x - \left(\frac{1}{2} \right)^{x-1}; \ 1 + 2^x \right\} = 2 + 2^x.
$$

Therefore, $w^*(0) = 0$ and $w^*(x) \le 2 + 2^x$ for $x \in \{1, 2, \dots\}$. Thus, (b) is proved.

Incidentally, from the calculations similar to those in the proof of (c), we will see actually that

$$
w^*(x) = 2 + 2^x \ \forall \, x \in \{1, 2, \dots\}.
$$

This equality will be used in the proof of a subsequent statement, and we formulate it as Lemma 1 below, and will prove it there. Nevertheless, for the purpose here, the validity of the inequality $\le$ is sufficient.

(c) We have seen in part (b) that this MDP model is absorbing (at 0) for the given initial distribution P_0 concentrated on $\{1\}$. We now verify that it is not uniformly absorbing (at 0) for P_0.

To show this, consider the deterministic stationary strategies

$$
\varphi^n(x) := 2 \cdot \mathbb{I}\{x \le n\} + \mathbb{I}\{x > n\} \ \forall \, n = 0, 1, 2, \dots. \tag{11}
$$

With slight abuse of notation, let us denote by

$$
\eta^{\pi}_{\mathsf{P}_0}(x) := \eta^{\pi}_{\mathsf{P}_0}(\{x\} \times \mathbf{A})
$$

the marginal on $\mathbf{X} \setminus \{0\} = \{1, 2, \dots\}$ of the occupation measure of the strategy π for the initial distribution P_0.

For each $n \in \{0, 1, \dots\}$, $\eta^{\varphi^n}_1$ is given by

$$
\eta^{\varphi^n}_1(x) = \begin{cases} \left(\frac{1}{2} \right)^{x-1}, & \text{if } x < n + 1; \\ \left(\frac{1}{2} \right)^n \frac{1}{p_{n+1}} = 2, & \text{if } x = n + 1; \\ 0, & \text{if } x > n + 1. \end{cases} \tag{12}
$$

Indeed, under the deterministic stationary strategy φ^n any state $x < n + 1$ is reached with probability $\left(\frac{1}{2} \right)^{x-1}$, and given that it is reached, the controlled process spends exactly one time unit on it; any state $x > n + 1$ is never reached. This justifies the first and the third equalities in (12). For the state $x = n + 1$, note that $\left(\frac{1}{2} \right)^n$ is the probability that $X_n = n + 1$. (The complementary probability $1 - \left(\frac{1}{2} \right)^n$ is for $X_n = 0$; the state $n + 1$ cannot appear at the steps $t < n$.) After that $X_{n+t} = n + 1$ is realized with probability $(1 - p_{n+1})^t$, leading to

$$\eta_1^{\varphi^n}(n+1) = \left(\frac{1}{2}\right)^n \left[1 + (1-p_{n+1}) + (1-p_{n+1})^2 - \ldots\right]$$

$$= \left(\frac{1}{2}\right)^n \frac{1}{p_{n+1}} = 2. \tag{13}$$

For each $n \geq 0$,

$$\sup_{\pi \in \Delta^{\text{All}}} \mathsf{E}_1^\pi \left[\sum_{t=n}^\infty \mathbb{I}\{X_t \in (\mathbf{X} \setminus \{0\})\}\right] \geq \mathsf{E}_1^{\varphi^n} \left[\sum_{t=n}^\infty \mathbb{I}\{X_t = n+1\}\right]$$

$$= \eta_1^{\varphi^n}(n+1) = 2, \tag{14}$$

so that $\sup_{\pi \in \Delta^{\text{All}}} \mathsf{E}_1^\pi \left[\sum_{t=n}^\infty \mathbb{I}\{X_t \in (\mathbf{X} \setminus \{0\})\}\right]$ does not converge to 0 as $n \to \infty$. In view of (2), it follows that this MDP model is not uniformly absorbing at 0 for the given P_0.

(d) Now we fix the weak topology on $\mathcal{D}$, as described in Sect. 3. Accordingly, in this proof, the notions of compactness and convergence of sequences in $\mathcal{D}$ are understood with respect to it, and this will not be signified repeatedly. The target here is to show that the space $\mathcal{D}$ is not compact. This is equivalent to showing that $\mathcal{D}$ is not sequentially compact, because the weak topology on $\mathcal{D}$ is metrizable, as mentioned in Sect. 3. Consequently, it suffices to show that the set $\{\eta_1^{\varphi^n}, \ n = 0, 1, 2, \ldots\}$ of occupation measures on $\mathcal{B}((\mathbf{X} \setminus \{0\}) \times \mathbf{A})$ has no accumulation points in $\mathcal{D}$, where the deterministic stationary strategies φ^n are defined by (11) in the proof of (c).

Suppose for contradiction that $\eta_1^{\varphi^{n_i}} \to \eta_1^\pi \in \mathcal{D}$ for some subsequence $n_i \to \infty$ as $i \to \infty$, and for some strategy π. For a fixed $j = 1, 2, \ldots$ take $d^j(x, a) := \mathbb{I}\{x = j\}$. Then

$$\lim_{i \to \infty} \sum_{(x,a) \in (\mathbf{X} \setminus \{0\}) \times \mathbf{A}} d^j(x, a) \eta_1^{\varphi^{n_i}}(x, a) = \left(\frac{1}{2}\right)^{j-1} \tag{15}$$

according to (12), because $n_i \to \infty$ as $i \to \infty$.

Since $d^j(\cdot)$ is bounded continuous on $(\mathbf{X} \setminus \{0\}) \times \mathbf{A}$, from $\eta_1^{\varphi^{n_i}} \to \eta_1^\pi \in \mathcal{D}$ we see that

$$\lim_{i \to \infty} \sum_{(x,a) \in (\mathbf{X} \setminus \{0\}) \times \mathbf{A}} d^j(x, a) \eta_1^{\varphi^{n_i}}(x, a) = \sum_{(x,a) \in (\mathbf{X} \setminus \{0\}) \times \mathbf{A}} d^j(x, a) \eta_1^\pi(x, a), \ \forall \, j \geq 1. \tag{16}$$

Applying (15) and (16) to $j = 1$, we see that

$$\sum_{(x,a) \in (\mathbf{X} \setminus \{0\}) \times \mathbf{A}} d^1(x, a) \eta_1^\pi(x, a) = \left(\frac{1}{2}\right)^0 = 1.$$

On the other hand, if $\pi_1(1|1) = \varepsilon \geq 0$ and $\pi_1(2|1) = 1 - \varepsilon$, then for $j = 1$, we have

$$\sum_{(x,a)\in(\mathbf{X}\setminus\{0\})\times\mathbf{A}} d^1(x,a)\eta_1^\pi(x,a) \geq \mathsf{E}_1^\pi\left[\mathbb{I}\{X_0 = 1\} + \mathbb{I}\{X_1 = 1\}\right] = 1 + \varepsilon(1 - p_1),$$

so that, it is necessary that $\varepsilon = 0$.

The similar calculation for $j = 2$, $d^2(x,a) = \mathbb{I}\{x = 2\}$, and $\pi_2(a|x_0, a_1, x_1)$ with $x_0 = 1, a_1 = 2$ and $x_1 = 2$ leads to equality $\pi_2(2|1, 2, 2) = 1$, and so on. Thus, π has the same occupation measure as the deterministic stationary strategy given by $\varphi(x) \equiv 2$, i.e., $\eta_1^\pi = \eta_1^\varphi$.

However, for the bounded continuous function $d(\cdot)$ on $(\mathbf{X} \setminus \{0\}) \times \mathbf{A}$ given by $d(x,a) \equiv 1$,

$$\sum_{(x,a)\in(\mathbf{X}\setminus\{0\})\times\mathbf{A}} d(x,a)\eta_1^\varphi(x,a) = 1 + \frac{1}{2} + \frac{1}{4} + \ldots = 2 \tag{17}$$

whereas

$$\sum_{(x,a)\in(\mathbf{X}\setminus\{0\})\times\mathbf{A}} d(x,a)\eta_1^{\varphi^{n_i}}(x,a) = \sum_{x=1}^{n_i}\left(\frac{1}{2}\right)^{x-1} + 2$$

$$= 4 - \frac{1}{2^{n_i-1}} \to 4 \text{ as } i \to \infty. \tag{18}$$

The previous two equalities yield the desired contradiction against the assumption $\eta_1^{\varphi^{n_i}} \to \eta_1^\pi$, because $d(\cdot)$ is bounded and continuous on $(\mathbf{X} \setminus \{0\}) \times \mathbf{A}$. Thus, the space $\mathcal{D}$ is not compact with respect to the weak topology, as required.

(e) As mentioned in Sect. 4, if the projection mapping O defined by (3) was continuous from $\mathcal{P}$ endowed with the weak topology to $\mathcal{D}$ endowed with the weak topology, then $\mathcal{D}$ would have been compact, because $\mathcal{P}$ is compact by Proposition 2. Since $\mathcal{D}$ is not compact as shown in (d), O is not continuous.

Alternatively, we may deduce it more explicitly, as follows. Consider the strategies φ^n and φ as in the proof of (d). Let us show that $\mathsf{P}_1^{\varphi^n} \to \mathsf{P}_1^\varphi$ as $n \to \infty$ in the weak topology. By Proposition 3, it is sufficient to show that $\mathsf{P}_1^{\varphi^n} \to \mathsf{P}_1^\varphi$ as $n \to \infty$ in the ws$^\infty$-topology. For a fixed $1 \leq T < \infty$, if $n \geq T$, then only the sequences $h_{T-1} = (x_0, a_1, x_1, \ldots, x_{T-1})$ of the form

$$(1, 2, 0, 2, 0, 2, 0, 2, 0, \ldots, 0)$$
$$(1, 2, 2, 2, 0, 2, 0, 2, 0, \ldots, 0)$$
$$(1, 2, 2, 2, 3, 2, 0, 2, 0, \ldots, 0)$$
$$(1, 2, 2, 2, 3, 2, 4, 2, 0, \ldots, 0)$$
$$(1, 2, 2, 2, 3, 2, 4, 2, 5, \ldots, 0)$$
$$\vdots$$
$$(1, 2, 2, 2, 3, 2, 4, 2, 5, \ldots, 0)$$

$$(1, 2, 2, 2, 3, 2, 4, 2, 5, \ldots, T)$$

can be realized with positive $\mathsf{P}_1^{\varphi^n}$-probabilities $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{32}, \ldots, \left(\frac{1}{2}\right)^{T-1}, \left(\frac{1}{2}\right)^{T-1}$, correspondingly. (Note, the last two probabilities coincide.) The same holds true for P_1^{φ}. Consequently, for each $T \geq 1$, $\mathsf{P}^n := \mathsf{P}_1^{\varphi^n}$ and $\mathsf{P} := \mathsf{P}_1^{\varphi}$,

$$\lim_{n \to \infty} \int_{\mathbf{H}_{T-1}} f(h_{T-1}) \mathsf{P}_{T-1}^n(dh_{T-1}) = \int_{\mathbf{H}_{T-1}} f(h_{T-1}) \mathsf{P}_{T-1}(dh_{T-1})$$

for each bounded function $f(\cdot)$ on $\mathbf{H}_{T-1}$, which is trivially continuous. Therefore, $\mathsf{P}_1^{\varphi^n} \to \mathsf{P}_1^{\varphi}$ as $n \to \infty$ in the ws$^\infty$-topology, and thus in the weak topology.

However, in the proof of (d), we have seen that $O(\mathsf{P}_1^{\varphi^n}) = \eta_1^{\varphi^n}$ does not converge to $O(\mathsf{P}_1^{\varphi}) = \eta_1^{\varphi}$ in the weak topology, see (17) and (18). This shows the claimed discontinuity of the mapping O.

(f) Since this MDP model is not uniformly absorbing (at 0) for the initial state $x_0 = 1$, it follows from Proposition 1 that there cannot be uniform Lyapunov functions. In fact, conditions (a,b) in Definition 6 can be satisfied, but any function satisfying them violates condition (c) in Definition 6. We demonstrate this fact explicitly as follows. First, the function defined by $\mu(0) = 1$ and $\mu(x) = 2 + 2^x$ for $x \geq 1$ satisfies condition (a) in Definition 6 because $1 + \sum_{y=1}^{\infty} \mu(y)p(y|0, a) \equiv 1 = \mu(0)$, and for each $x \geq 1$, it holds that

$$1 + \sum_{y=1}^{\infty} \mu(y)p(y|x, 1) = 1 + (1 - p_x)\mu(x)$$

$$= 1 + \left(1 - \frac{1}{2^x}\right)(2 + 2^x) = 1 + 2 + 2^x - \frac{1}{2^{x-1}} - 1$$

$$= 2 + 2^x - \frac{1}{2^{x-1}} \leq 2 + 2^x = \mu(x)$$

and

$$1 + \sum_{y=1}^{\infty} \mu(y)p(y|x, 2) = 1 + \frac{1}{2}\mu(x + 1)$$

$$= 1 + 1 + 2^x = 2 + 2^x = \mu(x).$$

Condition (b) in Definition 6 is trivially satisfied.

Next, let us show that, for each function $\mu(\cdot)$ satisfying condition (a) in Definition 6, condition (c) therein is violated. Indeed, condition (a) in Definition 6 implies that

$$1 + \max \left\{ (1 - p_x)\mu(x); \frac{1}{2}\mu(x + 1) \right\} \leq \mu(x), \ \forall \, x > 0,$$

leading to, for each $x \geq 1$,

$$1 + (1 - p_x)\mu(x) \leq \mu(x) \implies \mu(x) \geq \frac{1}{p_x} = 2^x;$$

$$\mu(x) \geq 1 + \frac{1}{2}\mu(x+1) \geq 1 + \frac{1}{2} \cdot 2^{x+1} = 1 + 2^x.$$

Now, for the deterministic stationary strategy $\varphi(x) \equiv 2$, we have for each $x \geq 1$ that

$$\mathsf{E}_x^\varphi[\mu(X_t) \cdot \mathbb{I}\{T_0 > t\}] \geq \left(\frac{1}{2}\right)^t [1 + 2^{x+t}] = \left(\frac{1}{2}\right)^t + 2^x,$$

where $\left(\frac{1}{2}\right)^t = \mathsf{P}_x^\varphi(T_0 > t)$ and, if $T_0 > t$, then $X_t = x + t$ with probability 1. Since the above expression does not converge to 0 as $t \to \infty$, we see that condition (c) in Definition 6 is not satisfied by $\mu(\cdot)$. $\qquad\square$

6.2 Proof of Theorem 2

Proof of Theorem 2. Suppose the MDP model is semi-continuous. First, we prove the 'if' part. This is done by mimicking the reasoning in the proof of [11, Lemma 4.7], avoiding the minor inaccuracy therein. Suppose $\mathsf{P}_{\mathsf{P}_0}^{\pi^i} \to \mathsf{P}_{\mathsf{P}_0}^{\hat{\pi}}$ as $i \to \infty$ in the weak topology, and fix an arbitrary $\varepsilon > 0$ and an arbitrary bounded continuous function $g(\cdot)$ on $\mathbf{X} \times \mathbf{A}$ such that $g(0, a) \equiv 0$. Let

$$\bar{g} := \sup_{(x,a)\in(\mathbf{X}\setminus\{0\})\times\mathbf{A}} |g(x, a)|.$$

Let $N \geq 0$ be such that, for all $\pi \in \Delta^{\mathrm{All}}$,

$$\bar{g} \sup_{\pi \in \Delta^{\mathrm{All}}} \mathsf{E}_{\mathsf{P}_0}^\pi \left[\sum_{t=N}^\infty \mathbb{I}\{X_t \in (\mathbf{X} \setminus \{0\})\} \right] < \varepsilon/3.$$

Such an $N \geq 0$ exists because the MDP model is assumed to be uniformly absorbing (at 0) for P_0 here. Then let K be such that, for all $i \geq K$,

$$\left| \mathsf{E}_{\mathsf{P}_0}^{\pi^i} \left[\sum_{t=1}^{N-1} g(X_{t-1}, A_t) \right] - \mathsf{E}_{\mathsf{P}_0}^{\hat{\pi}} \left[\sum_{t=1}^{N-1} g(X_{t-1}, A_t) \right] \right| < \varepsilon/3.$$

Such a K exists because by assumption, $\mathsf{P}_{\mathsf{P}_0}^{\pi^i} \to \mathsf{P}_{\mathsf{P}_0}^{\hat{\pi}}$ as $i \to \infty$ in the weak topology.
Now, for all $i \geq K$,

$$\left| \int_{(\mathbf{X}\setminus\{0\})\times\mathbf{A}} g(x, a)\eta_{\mathsf{P}_0}^{\pi^i}(dx \times da) - \int_{(\mathbf{X}\setminus\{0\})\times\mathbf{A}} g(x, a)\eta_{\mathsf{P}_0}^{\hat{\pi}}(dx \times da) \right|$$

$$= \left| \mathsf{E}_{\mathsf{P}_0}^{\pi^i} \left[\sum_{t=1}^{\infty} g(X_{t-1}, A_t) \right] - \mathsf{E}_{\mathsf{P}_0}^{\hat{\pi}} \left[\sum_{t=1}^{\infty} g(X_{t-1}, A_t) \right] \right|$$

$$\leq \left| \mathsf{E}_{\mathsf{P}_0}^{\pi^i} \left[\sum_{t=1}^{N-1} g(X_{t-1}, A_t) \right] - \mathsf{E}_{\mathsf{P}_0}^{\hat{\pi}} \left[\sum_{t=1}^{N-1} g(X_{t-1}, A_t) \right] \right|$$

$$+ \mathsf{E}_{\mathsf{P}_0}^{\pi^i} \left[\sum_{t=N}^{\infty} |g(X_{t-1}, A_t)| \right] + \mathsf{E}_{\mathsf{P}_0}^{\hat{\pi}} \left[\sum_{t=N}^{\infty} |g(X_{t-1}, A_t)| \right] < \varepsilon,$$

implying that $\eta^{\pi^i} = O(\mathsf{P}_{\mathsf{P}_0}^{\pi^i}) \to \eta^{\hat{\pi}} = O(\mathsf{P}_{\mathsf{P}_0}^{\hat{\pi}})$ as $i \to \infty$. Since the MDP model is semi-continuous, $\mathcal{P}$ is metrizable by Proposition 3. It follows that the mapping O defined by (3) is continuous.

Second, we prove the 'only if' part. Suppose the MDP model is semi-continuous and absorbing (at 0) for P_0, but not uniformly absorbing (at 0) for P_0. Then there is $\varepsilon > 0$ such that, for each $k \geq 1$, there is $n > k$ such that $\sup_{\pi \in \Delta^{\text{All}}} \mathsf{E}_{\mathsf{P}_0}^{\pi} \left[\sum_{t=n}^{\infty} \mathbb{I}\{X_t \neq 0\} \right] > \varepsilon$. Below, such $\varepsilon > 0$ is fixed.

Now

- there is $n_1 \geq 1$ such that there exists a strategy π^1 satisfying the inequality

$$\mathsf{E}_{\mathsf{P}_0}^{\pi^1} \left[\sum_{t=n_1}^{\infty} \mathbb{I}\{X_t \neq 0\} \right] > \varepsilon;$$

- there are $n_2 > n_1$ and π^2 such that

$$\mathsf{E}_{\mathsf{P}_0}^{\pi^2} \left[\sum_{t=n_2}^{\infty} \mathbb{I}\{X_t \neq 0\} \right] > \varepsilon;$$

- and so on.

In this manner, we obtain sequences $\{n_i\}_{i=1}^{\infty}$ and $\{\pi^i\}_{i=1}^{\infty}$ such that $n_i \uparrow \infty$ and, for each $i = 1, 2, \ldots,$

$$\mathsf{E}_{\mathsf{P}_0}^{\pi^i} \left[\sum_{t=n_i}^{\infty} \mathbb{I}\{X_t \neq 0\} \right] > \varepsilon.$$

Since $\mathcal{P}$ is compact and metrizable in the ws^{∞}-topology by Propositions 2 and 3, there exists a subsequence $\{\pi^{i_j}\}_{j=1}^{\infty}$ (with the corresponding subsequence $\{n_{i_j}\}_{j=1}^{\infty}$, $\lim_{j \to \infty} n_{i_j} = \infty$) such that $\mathsf{P}_{\mathsf{P}_0}^{\pi^{i_j}} \to \mathsf{P}_{\mathsf{P}_0}^{\hat{\pi}}$ in the weak topology for some strategy $\hat{\pi}$. We will show that the sequence $\{\eta_{\mathsf{P}_0}^{\pi^{i_j}}\}_{j=1}^{\infty}$ does not converge to $\eta_{\mathsf{P}_0}^{\hat{\pi}}$ in the weak topology.

For the constant $\varepsilon > 0$ fixed above, we choose $N > 0$ such that

$$\mathsf{E}_{\mathsf{P}_0}^{\hat{\pi}}\left[\sum_{t=N}^{\infty}\mathbb{I}\{X_{t-1}\neq 0\}\right] = \mathsf{E}_{\mathsf{P}_0}^{\hat{\pi}}\left[\sum_{t=N}^{\infty}\mathbb{I}\{t < T_0\}\right] < \varepsilon/3.$$

This can be done because the MDP is absorbing (at 0) for P_0: $\mathsf{E}_{\mathsf{P}_0}^{\hat{\pi}}[T_0] = \sum_{n=0}^{\infty}\mathsf{P}_{\mathsf{P}_0}^{\hat{\pi}}(T_0 > n) < \infty$.

After that, choose $K > 0$ such that, if $j > K$, then

$$\mathsf{E}_{\mathsf{P}_0}^{\pi^{i_j}}\left[\sum_{t=1}^{N-1}\mathbb{I}\{X_{t-1}\neq 0\}\right] - \mathsf{E}_{\mathsf{P}_0}^{\hat{\pi}}\left[\sum_{t=1}^{N-1}\mathbb{I}\{X_{t-1}\neq 0\}\right] > -\varepsilon/3.$$

This can be done because $\mathsf{P}_{\mathsf{P}_0}^{\pi^{i_j}} \to \mathsf{P}_{\mathsf{P}_0}^{\hat{\pi}}$ in the weak topology. Recall that 0 is an isolated state so that $\mathbb{I}\{x \neq 0\}$ is a bounded continuous function on $\mathbf{X}$.

Now, for $j > K$, $n_{i_j} > N$, i.e., for all big values of j, we have

$$\mathsf{E}_{\mathsf{P}_0}^{\pi^{i_j}}\left[\sum_{t=N}^{\infty}\mathbb{I}\{X_t \neq 0\}\right] \geq \mathsf{E}_{\mathsf{P}_0}^{\pi^{i_j}}\left[\sum_{t=n_{i_j}}^{\infty}\mathbb{I}\{X_t \neq 0\}\right] > \varepsilon$$

and the following relations for the bounded continuous function $g(x,a) \equiv 1$ on $(\mathbf{X}\setminus\{0\})\times\mathbf{A}$:

$$\int_{(\mathbf{X}\setminus\{0\})\times\mathbf{A}} g(x,a)\eta_{\mathsf{P}_0}^{\pi^{i_j}}(dx\times da) - \int_{(\mathbf{X}\setminus\{0\})\times\mathbf{A}} g(x,a)\eta_{\mathsf{P}_0}^{\hat{\pi}}(dx\times da)$$

$$= \mathsf{E}_{\mathsf{P}_0}^{\pi^{i_j}}\left[\sum_{t=1}^{N-1}\mathbb{I}\{X_{t-1}\neq 0\}\right] - \mathsf{E}_{\mathsf{P}_0}^{\hat{\pi}}\left[\sum_{t=1}^{N-1}\mathbb{I}\{X_{t-1}\neq 0\}\right]$$

$$+ \mathsf{E}_{\mathsf{P}_0}^{\pi^{i_j}}\left[\sum_{t=N}^{\infty}\mathbb{I}\{X_t\neq 0\}\right] - \mathsf{E}_{\mathsf{P}_0}^{\hat{\pi}}\left[\sum_{t=N}^{\infty}\mathbb{I}\{X_{t-1}\neq 0\}\right]$$

$$> -\frac{\varepsilon}{3} + \varepsilon - \frac{\varepsilon}{3} = \varepsilon/3.$$

Thus, the sequence $\{\eta_{\mathsf{P}_0}^{\pi^{i_j}}\}_{j=1}^{\infty}$ does not converge to $\eta_{\mathsf{P}_0}^{\hat{\pi}}$ in the weak topology, and the mapping O defined by (3) is not continuous. $\square$

6.3 Proof of Proposition 5

In this subsection, we prove Proposition 5. Firstly, we present a lemma.

Lemma 1 *Consider the MDP model in Example 2, and the function $w^*(\cdot)$ on $\mathbf{X}$ defined in (8). Then*

$$w^*(0) = 0;$$
$$w^*(x) = 2 + 2^x \ \forall \ x \in \{1, 2, \dots\}.$$

Proof The first equality holds vacuously. The rest verifies the second equality. For a fixed $x \geq 1$, consider the following deterministic stationary strategies

$$\varphi^n(y) := 2 \cdot \mathbb{I}\{y \leq x + n\} + \mathbb{I}\{y > x + n\}, \ \ n = 0, 1, \dots.$$

Then similar considerations to those for (12) result in, with slight abuse of notation,

$$\eta_x^{\varphi^n}(y) = \begin{cases} \left(\frac{1}{2}\right)^{y-x}, & \text{if } x \leq y < x + n + 1; \\ \left(\frac{1}{2}\right)^{y-x} \frac{1}{p_{x+n+1}} = 2^x, & \text{if } y = x + n + 1; \\ 0, & \text{if } y > x + n + 1, \text{ or } y < x. \end{cases}$$

for the marginal on $\mathbf{X} \setminus \{0\}$ of the occupation measure of the strategy φ^n for the initial state x. Now

$$\mathsf{E}_x^{\varphi^n}[T_0] = \sum_{y=1}^{\infty} \eta_x^{\varphi^n}(y) = \sum_{y=x}^{x+n} \left(\frac{1}{2}\right)^{y-x} + 2^x = \frac{1 - \left(\frac{1}{2}\right)^{n+1}}{\frac{1}{2}} + 2^x = 2 + 2^x - \frac{1}{2^n},$$

and hence,

$$w^*(x) = \sup_{\pi \in \Delta^{\mathrm{All}}} \mathsf{E}_x^{\pi}[T_0] \geq \sup_{n \geq 0} \mathsf{E}_x^{\varphi^n}[T_0] = \sup_{n \geq 0} \left\{ 2 + 2^x - \frac{1}{2^n} \right\} = 2 + 2^x,$$

where the first equality is by the definition of the function $w^*(\cdot)$, see (8).

On the other hand, it was shown in the proof of (b) of Theorem 1 for each $x \geq 1$ that $w^*(x) \leq 2 + 2^x$, see (10). Combining the previous two inequalities yields $w^*(x) = 2 + 2^x$ for all $x \geq 1$. Thus, the statement is proved. $\qquad \square$

Now we are in position to prove Proposition 5.

Proof of Proposition 5. (a) Recall that

$$v^*(x) := \inf_{\pi \in \Delta^{\mathrm{All}}} v^{\pi}(x)$$

for each $x \in \mathbf{X}$. By definition of the function $w^*(\cdot)$, see (8), $v^*(x) = -w^*(x)$ for each $x \in \mathbf{X}$. Now, according to Lemma 1,

$$v^*(0) = 0,$$
$$v^*(x) = -2 - 2^x, \ \forall \ x \in \{1, 2, \dots\}. \tag{19}$$

Suppose for contradiction that $\pi = \{\pi_n\}_{n=1}^{\infty}$ is an optimal strategy for the initial state 1. If $\pi_1(1|1) = \varepsilon$ and $\pi_1(2|1) = 1 - \varepsilon$, then

$$v^{\pi}(1) = -1 + \varepsilon \cdot \frac{1}{2} \cdot v^{(1,1)\pi}(1) + (1 - \varepsilon) \cdot \frac{1}{2} \cdot v^{(1,2)\pi}(2).$$

where for $j \in \{1, 2\}$ $^{(1,j)}\pi$ is the shifted strategy defined in (1). From (19), we then see

$$v^{\pi}(1) \geq -1 + \varepsilon \cdot \frac{1}{2} \cdot v^*(1) + (1 - \varepsilon) \cdot \frac{1}{2} \cdot v^*(2) = -4 + \varepsilon.$$

Hence, in order to obtain $v^{\pi}(1) = v^*(1) = -4$, it is necessary that $\varepsilon = 0$. The similar calculation for $\pi_2(a|x_0, a_1, x_1)$ with $x_0 = 1$, $a_1 = 2$ and $x_1 = 2$ leads to the equality $\pi_2(2|1, 2, 2) = 1$, and so on. Thus,

$$v^{\pi}(1) = v^{\varphi}(1),$$

where φ is the deterministic stationary strategy given by $\varphi(x) \equiv 2$.

On the other hand,

$$v^{\varphi}(1) = -1 - \frac{1}{2} - \frac{1}{4} - \ldots = -2 > -4 = v^*(1),$$

and thus $v^{\pi}(1) > v^*(1)$, which is a desired contradiction against that π is optimal for the initial state 1. Consequently, there are no optimal strategies to problem (5) for the MDP model with the given initial state 1 and the given cost function $c(\cdot)$.

(b) The last assertion of this part is trivially true. In particular, by Lemma 1,

$$\sup_{\pi \in \Delta^{\mathrm{All}}} \left\{ \mathsf{E}_1^{\pi} \left[\sum_{t=1}^{\infty} c^-(X_{t-1}, A_t) \right] \right\} = w^*(1) = 4 < \infty,$$

where the first equality is by (8). Thus, it follows from (a) and Proposition 4 that condition (7) cannot be satisfied.

To show explicitly that Condition (7) is violated, consider again the deterministic stationary strategies φ^n given by (11), i.e.,

$$\varphi^n(x) := 2 \cdot \mathbb{I}\{x \leq n\} + \mathbb{I}\{x > n\} \ \forall \ n = 0, 1, 2, \ldots.$$

Now

$$\inf_{N \geq n} \inf_{\pi \in \Delta^{\mathrm{All}}} \sum_{t=n+1}^{N} \mathsf{E}_1^{\pi} \left[c(X_{t-1}, A_t) \right] \leq \mathsf{E}_1^{\varphi^n} \left[c(X_n, A_{n+1}) \right]$$

$$= \left(\frac{1}{2} \right)^n \left[-1 - (1 - p_{n+1}) - (1 - p_{n+1})^2 - \ldots \right]$$

$$= -\left(\frac{1}{2}\right)^n \frac{1}{p_{n+1}} = -2,$$

where the inequality holds because $c(\cdot) \leq 0$, and the equalities hold by the similar argument as in (13). □

Remark 1 A strategy π is called uniformly optimal if $v^\pi(x) = v^*(x)$ for all $x \in \mathbf{X}$. Consider an MDP model with a cost function $c(\cdot)$ such that $\sup_{\pi \in \Delta^{\mathrm{All}}} \mathsf{E}_x^\pi \left[\sum_{n=0}^\infty c^-(X_n, A_{n+1}) \right] < \infty$ and $v^*(x) \in (-\infty, \infty)$ for all $x \in \mathbf{X}$, where $v^*(\cdot)$ is defined by (4). Then according to [20, Theorem 2.2], a deterministic stationary strategy φ is uniformly optimal if and only if

$$v^*(x) = c(x, \varphi(x)) + \int_{\mathbf{X}} v^*(y) p(dy|x, \varphi(x)); \quad \lim_{n \to \infty} \mathsf{E}_x^\varphi \left[v^*(X_n) \right] = 0. \quad (20)$$

The two equalities in (20) are called the Dubins-Savage conditions.

Now, consider the optimal control problem (5) for the MDP model in Example 2 with the cost function $c(x, a) \equiv -1$ for $x \in \{1, 2, \dots\}$ and $c(0, a) \equiv 0$. We have seen from Proposition 5(a) that there is no uniformly optimal strategy. In particular, $\varphi(x) \equiv 2$ is not uniformly optimal. Let us verify this fact again by checking the Dubins-Savage conditions. They are sufficient and necessary for the uniform optimality of φ because $c(\cdot)$ is $(-\infty, 0]$-valued and $v^*(\cdot)$ is finite-valued. Now, we observe that φ actually satisfies the first equality in (20). Nevertheless, the second equality in (20) is violated because

$$\mathsf{E}_1^\varphi[v^*(X_n)] = \left(\frac{1}{2}\right)^n [-2 - 2^{n+1}] = -\left(\frac{1}{2}\right)^{n-1} - 2 \to -2 \text{ as } n \to \infty,$$

here $\left(\frac{1}{2}\right)^n$ is the probability that $X_n = n + 1$.

7 Conclusion

In conclusion, we showed that for a semi-continuous absorbing MDP with a fixed initial distribution, the continuity of the projection mapping O from the space of strategic measures $\mathcal{P}$ to the space of occupation measures $\mathcal{D}$, both of which are endowed with their weak topologies, is equivalent to the MDP model being uniformly absorbing. This is confirmed by an example. Provided that the absorbing MDP model is semi-continuous, the continuity of O is a sufficient condition for the compactness of $\mathcal{D}$ in the weak topology. Whether it is also necessary is an interesting open problem for future studies.

In connection to the last problem, we would like to comment on the very relevant and interesting paper [8], which was brought to our attention during the review of the present paper by one of the referees. Absorbing MDP models were studied in [8], where the state space was allowed to be a general measurable space. Assuming that the model is semi-continuous (S), it was shown in [8, Theorem 4.10] that the absorbing

MDP model is uniformly absorbing if and only if the space of occupation measures $\mathcal{D}$ is compact with respect to the ws-topology. Now let us specialize to a Borel state space. Then this result, together with Theorem 2, shows that the ws-compactness of $\mathcal{D}$ is equivalent to the continuity of the mapping O from the space of strategic measures $\mathcal{P}$ to the space of occupation measures $\mathcal{D}$, both of which are endowed with their weak topologies. In general, the ws-topology on $\mathcal{D}$ is at least as strong as the weak topology on $\mathcal{D}$. If the ws-topology and the weak topology on $\mathcal{D}$ coincide, which is the case if the state space is denumerable and endowed with the discrete topology, then, from the above discussion, we see that for an absorbing semi-continuous (S) MDP model, the continuity of O is indeed equivalent to the compactness of $\mathcal{D}$, which is endowed with the weak topology.

Suppose the ws-topology is equivalent to the weak topology on $\mathcal{D}$, and the absorbing MDP model is semi-continuous (S). Then the necessity part of Theorem 2 would follow directly from [8, Theorem 4.10]: if the MDP model is not uniformly absorbing, then, by [8, Theorem 4.10], $\mathcal{D}$ is not compact in the ws-topology, which is assumed to be equivalent to the weak topology, and O cannot be continuous when $\mathcal{D}$ is endowed with the weak topology.

Finally, we mention that if the MDP model is not absorbing, then $\mathcal{D}$ contains measures that are not (totally) finite. In this case, a different topology was introduced on $\mathcal{D}$ in e.g., [10, 18]. In that topology, the projection mapping O is continuous by definition.

Added after the acceptance of this paper: After the paper was accepted, we learnt that Dufour and Prieto-Rumeau [9] had picked up and solved the problem mentioned in the conclusion of the present paper. More exactly, it was shown in [9, Theorem 3.1] that if the MDP model is semi-continuous (or even satisfies a more general condition) and absorbing at 0 for a given initial distribution P_0, then the compactness of $\mathcal{D}$ in the weak topology is indeed equivalent to the MDP model being uniformly absorbing at 0 for the given initial distribution P_0.

Declarations

References

1. Altman, E.: Constrained Markov Decision Processes. Chapman and Hall/CRC, Boca Raton (1999)
2. Bertsekas, D., Shreve, S.: Stochastic Optimal Control. Academic Press, New York (1978)
3. Bogachev, V.: Measure Theory, vol. 2. Springer, Berlin (2007)
4. Bogachev, V.: Weak Convergence of Measures. American Mathematical Society, Providence (2018)
5. Cavazos-Cadena, R., Hernández-Lerma: Equivalence of Lyapunov stability criteria in a class of Markov decision processes. Appl. Math. Optim. **26**, 113–137 (1992)
6. Derman, C.: Finite State Markovian Decision Processes. Academic Press, New York (1970)
7. Dufour, F., Genadot, A.: On the expected total cost with unbounded returns for Markov decision processes. Appl. Math. Optim. **82**, 433–450 (2020)
8. Dufour, F., Prieto-Rumeau, T.: Absorbing Markov decision processes. ESAIM: COCV **30**, 1–18 (2024)
9. Dufour, F., Prieto-Rumeau, T.: A note on weak compactness of occupation measures for an absorbing Markov decision process. Preprint at arXiv:2402.10672 (2024)
10. Dufour, F., Horiguchi, M., Piunovskiy, A.: The expected total cost criterion for Markov decision processes under constraints: a convex analytic approach. Adv. Appl. Prob. **44**, 774–793 (2012)
11. Feinberg, E.A., Rothblum, U.: Splitting randomized stationary policies in total-reward Markov decision processes. Math. Oper. Res. **37**, 129–153 (2012)
12. Feinberg, E.A., Piunovskiy, A.: Sufficiency of deterministic policies for atomless discounted and uniformly absorbing MDPs with multiple criteria. SIAM J. Control Optim. **57**, 163–191 (2019)
13. Hernández-Lerma, O., Lasserre, J.B.: Further Topics on Discrete-Time Markov Control Processes. Springer, New York (1999)
14. Hordijk, A.: Dynamic Programming and Potential Theory. Mathematish Centrum, Amsterdam (1975)
15. James, H., Collins, E.: An analysis of transient Markov decision processes. J. Appl. Probab. **43**, 603–621 (2006)
16. Kurushima, A., Piunovskiy, A., Zhang, Y.: Nowak's theorem on probability measures induced by strategies revisited. Theory Probab. Appl. **62**, 328–334 (2018)
17. Nowak, A.: On the weak topology on a space of probability measures induced by policies. Bullet. Pol. Acad. Sci. Math. **36**, 181–186 (1988)
18. Piunovskiy, A., Zhang, Y.: Extreme occupation measures in Markov decision processes with an absorbing state. SIAM J. Control Optim. **62**, 65–90 (2023)
19. Schäl, M.: On dynamic programming: compactness of the space of policies. Stoch. Proc. Appl. **3**, 345–364 (1975)
20. Yushkevich, A.A., Chitashvili, R.Ya.: Controllable random sequences and Markov chains. Russ. Math. Surv. **37**, 239–274 (1982)