

ABSORBING MARKOV DECISION PROCESSES

FRANÇOIS DUFOUR¹ AND TOMÁS PRIETO-RUMEAU^{2,*}

Abstract. In this paper, we study discrete-time absorbing Markov Decision Processes (MDP) with measurable state space and Borel action space with a given initial distribution. For such models, solutions to the characteristic equation that are not occupation measures may exist. Several necessary and sufficient conditions are provided to guarantee that any solution to the characteristic equation is an occupation measure. Under the so-called continuity-compactness conditions, we first show that a measure is precisely an occupation measure if and only if it satisfies the characteristic equation and an additional absolute continuity condition. Secondly, it is shown that the set of occupation measures is compact in the weak-strong topology if and only if the model is uniformly absorbing. Several examples are provided to illustrate our results.

Mathematics Subject Classification. 91A10, 91A15.

Received September 19, 2023. Accepted December 30, 2023.

1. INTRODUCTION

In this work, we consider a discrete-time absorbing Markov Decision Process (MDP) with measurable state space \mathbf{X} , Borel action space \mathbf{A} , with a given initial distribution denoted by η . We consider a general measurable state space to cover the models studied in game theory. Indeed, most models studied in nonzero-sum games theory assume that the state space is a general measurable space; see for example [1, 2] and the references therein. Recently, nonzero-sum Markov games have been studied for absorbing models in [3], which has motivated the work presented in this paper.

An absorbing MDP is a special type of MDP where some measurable subset Δ of the state space is considered as *absorbing* and leads to the stopping of the process. This means that once the system enters in Δ , it stays there indefinitely with zero reward or cost. The underlying assumption is that, for any policy, the expected entrance time to Δ is finite. For more details on this type of MDP, we refer to the following papers and books [4–7] and their references. It can be easily shown that a discounted model can be reformulated as an absorbing model by adding an additional state and modifying the transition and performance function to make the singleton associated with that state the absorbing set. For more details on this construction, we refer to [4], p. 137 as well as [6], p. 132. However, the notion of an absorbing model is more general than that of a discounted model, in the sense that there exist absorbing models that cannot be transformed into discounted models.

Keywords and phrases: Markov decision processes, absorbing model, occupation measures, characteristic equation, phantom measures, compactness of the set of occupation measures.

¹ Institut Polytechnique de Bordeaux; INRIA Bordeaux Sud Ouest, Team: ASTRAL; IMB, Institut de Mathématiques de Bordeaux, Université de Bordeaux, France.

² Statistics Department, UNED, Madrid, Spain.

* Corresponding author: tprieto@ccia.uned.es

The objective of this paper is to study the properties of the solutions of the characteristic equation and of the set of occupation measures. The notion of occupation measure is particularly important and plays a central role in solving constrained Markov decision processes (see for example the references [4, 8–11]). It describes the expected amount of time spent by the state and action processes in any measurable subset of $\mathbf{X} \times \mathbf{A}$ under any policy. For an absorbing model, any occupation measure satisfies the so-called characteristic equation which is of the form

$$\mu^{\mathbf{X}}(\cdot) = \eta_{\Delta^c}(\cdot) + \int_{\mathbf{X} \times \mathbf{A}} Q\mathbb{I}_{\Delta^c}(\cdot|x, a)\mu(dx, da) \quad (1.1)$$

(see Eq. (3.1) for a more precise statement) where η_{Δ^c} is the restriction of the measure η to Δ^c , $\mu^{\mathbf{X}}$ is the marginal measure of μ on \mathbf{X} , and $Q\mathbb{I}_{\Delta^c}$ is the product of the transition kernel Q of the MDP and \mathbb{I}_{Δ^c} , the indicator kernel of the set Δ^c . However, there may exist solutions to the characteristic equation (1.1) that are not occupation measures (see Ex. 3.5). These are called phantom solutions to the characteristic equation (or simply phantom measures). This is a notable difference from the discounted case, for which there are no phantom measures for the associated characteristic equation, which implies that a solution of the characteristic equation is necessarily an occupation measure. In order to avoid this pathological phenomenon for absorbing models, we will provide in Theorem 3.7 several necessary and sufficient conditions to ensure that there are no phantom measures, thereby guaranteeing that any solution to the characteristic equation is an occupation measure.

By strengthening our hypotheses, that is, assuming the so-called continuity-compactness conditions introduced by Schäl in [12] (see Condition (S) below), we will show in our second main result (see Thm. 4.6) that a measure is an occupation measure if and only if it satisfies the characteristic equation (1.1) and an additional condition. Let us emphasize that a measure satisfies this last condition if its marginal on \mathbf{X} is absolutely continuous with respect to a reference probability measure. The construction of this reference measure is essentially based on Condition (S). An interesting property that plays a central role in the proof of our last result (see Thm. 4.10) is that if the initial distribution η of the MDP is replaced by this reference probability measure, then the MDP remains absorbing.

Finally, it will be shown in our last main result (see Thm. 4.10) that under the continuity-compactness conditions the set of occupation measures is compact in the ws -topology (see [13] for more details on the weak-strong topology) if and only if the model is uniformly absorbing (the notion of a uniformly absorbing MDP was introduced in [5], Def. 3.6). In fact, to be more precise, our result also shows that the set of occupation measures is relatively compact in the ws -topology if and only if the model is uniformly absorbing. In particular, it highlights the fact that the assumption that the MDP is absorbing is not sufficient to guarantee the (relative) compactness of the set of occupation measures. This last point is illustrated by an example. Observe that a discounted MDP is uniformly absorbing and so we recover the classical result that, under the continuity-compactness conditions, the set of occupation measures of a discounted model is compact. The question of the compactness of the set of occupation measures is central to solving constrained optimal control problems using linear programming.

The main results of this paper appear to be more general than those established in the literature, in particular the characterization of phantom and occupation measures (Thms. 3.7 and 4.10), as well as the compactness of occupation measures (Thm. 4.10), notably because these characterizations are made by means of necessary and sufficient conditions.

The rest of the paper is organized as follows. In the remaining of this section we introduce some notation and recall a standard result that will be useful in the sequel. In Section 2, we introduce the control model under consideration and provide some basic definitions. Section 3 is devoted to the analysis of the characteristic equation. In the last section, it is shown that the occupation measures are characterized by the characteristic equation and an additional condition. We also study the compactness property of the set of occupation measures.

Notation and terminology. The power set of a set S (the set of all subsets of S) is denoted by 2^S . On a measurable space (Ω, \mathcal{F}) we will consider the set of finite signed measures $\mathcal{M}(\Omega)$, the set of finite nonnegative

measures $\mathcal{M}^+(\Omega)$, and the set of probability measures $\mathcal{P}(\Omega)$. For a set $\Gamma \in \mathcal{F}$, we denote by $\mathbf{I}_\Gamma : \Omega \rightarrow \{0, 1\}$ the indicator function of the set Γ , that is, $\mathbf{I}_\Gamma(\omega) = 1$ if and only if $\omega \in \Gamma$. For $\omega \in \Omega$, we write $\delta_{\{\omega\}}$ for the Dirac probability measure at ω defined on (Ω, \mathcal{F}) by $\delta_{\{\omega\}}(B) = \mathbf{I}_B(\omega)$ for any $B \in \mathcal{F}$. If $\mu \in \mathcal{M}(\Omega)$ and $\Gamma \in \mathcal{F}$, we denote by μ_Γ the restriction of the measure μ to Γ , that is, the measure on (Ω, \mathcal{F}) defined by $\mu_\Gamma(B) = \mu(\Gamma \cap B)$ for $B \in \mathcal{F}$. The trace σ -algebra of a set $\Gamma \subseteq \Omega$ is denoted by \mathcal{F}_Γ . On $\mathcal{P}(\Omega)$, the s -topology is the coarsest topology that makes $\mu \mapsto \mu(D)$ continuous for every $D \in \mathcal{F}$.

Given a measurable space (Ω, \mathcal{F}) and $\lambda \in \mathcal{P}(\Omega)$, we will denote by $L^1(\Omega, \mathcal{F}, \lambda)$ the family of measurable functions (identifying those which are λ -a.s. equal) $f : \Omega \rightarrow \mathbb{R}$ which are λ -integrable, *i.e.*, $\|f\|_1 = \int_\Omega |f(\omega)| \lambda(d\omega) < \infty$. We note that, throughout this paper, the real numbers set \mathbb{R} will be always endowed with its Borel σ -algebra.

Let (Ω, \mathcal{F}) and $(\tilde{\Omega}, \tilde{\mathcal{F}})$ be two measurable spaces. A kernel on $\tilde{\Omega}$ given Ω is a mapping $Q : \Omega \times \tilde{\mathcal{F}} \rightarrow \mathbb{R}^+$ such that $\omega \mapsto Q(B|\omega)$ is measurable on (Ω, \mathcal{F}) for every $B \in \tilde{\mathcal{F}}$, and $B \mapsto Q(B|\omega)$ is in $\mathcal{M}^+(\tilde{\Omega})$ for every $\omega \in \Omega$. If $Q(\tilde{\Omega}|\omega) = 1$ for all $\omega \in \Omega$ then we say that Q is a *stochastic kernel*. We write \mathbb{I}_Γ for the kernel on Ω given Ω defined by $\mathbb{I}_\Gamma(B|\omega) = \mathbf{I}_\Gamma(\omega) \delta_{\{\omega\}}(B)$ for $\omega \in \Omega$ and $B \in \mathcal{F}$. Let Q be a stochastic kernel on $\tilde{\Omega}$ given Ω . For a bounded measurable function $f : \tilde{\Omega} \rightarrow \mathbb{R}$, we will denote by $Qf : \Omega \rightarrow \mathbb{R}$ the measurable function

$$Qf(\omega) = \int_{\tilde{\Omega}} f(\tilde{\omega}) Q(d\tilde{\omega}|\omega) \quad \text{for } \omega \in \Omega.$$

For a measure $\mu \in \mathcal{M}^+(\Omega)$, we denote by μQ the finite measure on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ given by

$$B \mapsto \mu Q(B) = \int_\Omega Q(B|\omega) \mu(d\omega) \quad \text{for } B \in \tilde{\mathcal{F}}.$$

The product of the σ -algebras \mathcal{F} and $\tilde{\mathcal{F}}$ is denoted by $\mathcal{F} \otimes \tilde{\mathcal{F}}$ and consists of the σ -algebra generated by the measurable rectangles, that is, the sets of the form $\Gamma \times \tilde{\Gamma}$ for $\Gamma \in \mathcal{F}$ and $\tilde{\Gamma} \in \tilde{\mathcal{F}}$. We denote by $\mu \otimes Q$ the unique finite measure on the product space $(\Omega \times \tilde{\Omega}, \mathcal{F} \otimes \tilde{\mathcal{F}})$ satisfying

$$(\mu \otimes Q)(\Gamma \times \tilde{\Gamma}) = \int_\Omega Q(\tilde{\Gamma}|\omega) \mu(d\omega) \quad \text{for } \Gamma \in \mathcal{F} \text{ and } \tilde{\Gamma} \in \tilde{\mathcal{F}},$$

see Proposition III-2-1 in [14] for a proof of the existence and uniqueness of such measure. Let $(\bar{\Omega}, \bar{\mathcal{F}})$ be a third measurable space and R be a stochastic kernel on $\bar{\Omega}$ given $\tilde{\Omega}$. Then we will denote by QR the stochastic kernel on $\bar{\Omega}$ given Ω defined as

$$QR(\Gamma|\omega) = \int_{\tilde{\Omega}} R(\Gamma|\tilde{\omega}) Q(d\tilde{\omega}|\omega) \quad \text{for } \Gamma \in \bar{\mathcal{F}} \text{ and } \omega \in \Omega.$$

Given $\mu \in \mathcal{M}(\Omega \times \tilde{\Omega})$, the marginal measures are $\mu^\Omega \in \mathcal{M}(\Omega)$ and $\mu^{\tilde{\Omega}} \in \mathcal{M}(\tilde{\Omega})$ defined by $\mu^\Omega(\cdot) = \mu(\cdot \times \tilde{\Omega})$ and $\mu^{\tilde{\Omega}}(\cdot) = \mu(\Omega \times \cdot)$. If π is a kernel on $\tilde{\Omega} \times \bar{\Omega}$ given Ω the marginal kernels are $\pi^{\tilde{\Omega}}$ and $\pi^{\bar{\Omega}}$, respectively defined by $\pi^{\tilde{\Omega}} = \pi(\cdot \times \bar{\Omega}|\omega)$ and $\pi^{\bar{\Omega}} = \pi(\tilde{\Omega} \times \cdot|\omega)$ for $\omega \in \Omega$.

Given metric spaces \mathbf{S} and \mathbf{S}' (endowed with their respective Borel σ -algebras), we say that $f : \Omega \times \mathbf{S} \rightarrow \mathbf{S}'$ is a *Carathéodory function* if $f(\cdot, s)$ is measurable on Ω for every $s \in \mathbf{S}$ and $f(\omega, \cdot)$ is continuous on \mathbf{S} for every $\omega \in \Omega$. The family of the so-defined Carathéodory functions is denoted by $\text{Car}(\Omega \times \mathbf{S}, \mathbf{S}')$. The family of Carathéodory functions which, in addition, are bounded is denoted by $\text{Car}_b(\Omega \times \mathbf{S}, \mathbf{S}')$. When the metric space \mathbf{S} is separable then any $f \in \text{Car}(\Omega \times \mathbf{S}, \mathbf{S}')$ is a jointly measurable function on $(\Omega \times \mathbf{S}, \mathcal{F} \otimes \mathfrak{B}(\mathbf{S}))$; see [15], Lemma 4.51.

If \mathbf{S} is a Polish space (a complete and separable metric space), on $\mathcal{M}(\Omega \times \mathbf{S})$ we will consider the ws -topology (weak-strong topology) which is the coarsest topology for which the mappings

$$\mu \mapsto \int_{\Omega \times \mathbf{S}} f(\omega, s) \mu(d\omega, ds)$$

for $f \in \mathcal{C}ar_b(\Omega \times \mathbf{S}, \mathbb{R})$ are continuous. There are other equivalent definitions of this topology as discussed, for instance, in [16], Section 3.3.

The next disintegration lemma will be useful in the forthcoming (see Thm. 1 in [17]).

Lemma 1.1 (Disintegration lemma). *Let (Ω, \mathcal{F}) be a measurable space and let \mathbf{S} be a Polish space. Let $\Phi : \Omega \rightarrow 2^{\mathbf{S}}$ be a weakly measurable correspondence with nonempty closed values, and let \mathbf{K} be the graph of the correspondence. For every $\mu \in \mathcal{M}^+(\Omega \times \mathbf{S})$ such that $\mu(\mathbf{K}^c) = 0$ there exists a stochastic kernel Q on \mathbf{S} given Ω such that*

$$\mu = \mu^\Omega \otimes Q \tag{1.2}$$

and such that $Q(\Phi(\omega)|\omega) = 1$ for each $\omega \in \Omega$. Moreover, Q is unique μ^Ω -almost surely, meaning that if Q and Q' are two stochastic kernels that satisfy (1.2) then for all ω in a set of μ^Ω -probability one, the probability measures $Q(\cdot|\omega)$ and $Q'(\cdot|\omega)$ coincide.

2. THE ABSORBING CONTROL MODEL

The main goal of this section is to introduce the parameters defining the model with a brief presentation of the construction of the controlled process. We also describe the notions of absorbing and uniformly absorbing MDP, and give the definition of an occupation measure, providing a first elementary property (see Lem. 2.4).

2.1. The control model

We consider a stationary Markov controlled process $(\mathbf{X}, \mathbf{A}, \{\mathbf{A}(x) : x \in \mathbf{X}\}, Q, \eta)$ consisting of:

- A measurable state space \mathbf{X} endowed with a σ -algebra \mathfrak{X} .
- A Borel space \mathbf{A} , representing the action space.
- A family of nonempty measurable sets $\mathbf{A}(x) \subseteq \mathbf{A}$ for $x \in \mathbf{X}$. The set $\mathbf{A}(x)$ gives the available actions in state x . Let $\mathbf{K} = \{(x, a) \in \mathbf{X} \times \mathbf{A} : a \in \mathbf{A}(x)\}$ be the family of feasible state-action pairs. We assume that \mathbf{K} is a measurable subset of $\mathbf{X} \times \mathbf{A}$ endowed with the σ -algebra $\mathfrak{X} \otimes \mathfrak{B}(\mathbf{A})$. We also assume that \mathbf{K} contains the graph of a measurable selector (*i.e.*, the graph of a measurable $f : \mathbf{X} \rightarrow \mathbf{A}$ such that $f(x) \in \mathbf{A}(x)$ for every $x \in \mathbf{X}$).
- A stochastic kernel Q on \mathbf{X} given $\mathbf{X} \times \mathbf{A}$, which stands for the transition probability function.
- An initial distribution given by $\eta \in \mathcal{P}(\mathbf{X})$.

Additionally, we assume that a measurable set $\Delta \in \mathfrak{X}$ is given. As we shall see later in Definition 2.2, the set Δ will play the role of the *absorbing set*. The so-defined control model is denoted by $\mathfrak{M}(\eta, \Delta)$ where we make explicit the dependence on the initial distribution and the absorbing set.

The space of admissible histories of the controlled process up to time $n \in \mathbb{N}$ is denoted by \mathbf{H}_n . It is defined recursively by

$$\mathbf{H}_0 = \mathbf{X} \quad \text{and} \quad \mathbf{H}_n = \mathbf{H}_{n-1} \times \mathbf{A} \times \mathbf{X} \quad \text{for } n \geq 1,$$

all endowed with their corresponding product σ -algebras. A control policy π is a sequence $\{\pi_n\}_{n \geq 0}$ of stochastic kernels on \mathbf{A} given \mathbf{H}_n , denoted by $\pi_n(da|h_n)$, such that

$$\pi_n(\mathbf{A}(x_n)|h_n) = 1 \quad \text{for each } n \geq 0 \text{ and } h_n = (x_0, a_0, \dots, x_n) \in \mathbf{H}_n.$$

The set of all policies is denoted by $\mathbf{\Pi}$.

Let us denote by \mathbf{M} the set of stochastic kernels σ on \mathbf{A} given \mathbf{X} satisfying $\sigma(\mathbf{A}(x)|x) = 1$ for any $x \in \mathbf{X}$. A policy $\pi = \{\pi_n\}_{n \in \mathbb{N}} \in \mathbf{\Pi}$ is called a *stationary randomized policy* if there exists $\sigma \in \mathbf{M}$ satisfying

$$\pi_n(\cdot|h_n) = \sigma(\cdot|x_n) \quad \text{for any } h_n = (x_0, a_0, \dots, x_n) \in \mathbf{H}_n \text{ and } n \in \mathbb{N}.$$

In such a case, we will write σ instead of π to emphasize that the corresponding stationary randomized policy π is generated by σ . Since the set of stationary randomized policies can be identified with \mathbf{M} , we refer to it as \mathbf{M} in a slight abuse of notation.

For each $\sigma \in \mathbf{M}$, we denote by Q_σ the stochastic kernel on \mathbf{X} given \mathbf{X} defined by

$$Q_\sigma(D|x) = \int_{\mathbf{A}} Q(D|x, a)\sigma(da|x) \quad \text{for } x \in \mathbf{X} \text{ and } D \in \mathfrak{X}.$$

The compositions of Q_σ with itself are denoted by Q_σ^t for any $t \geq 0$, with the convention that $Q_\sigma^0(\cdot|x)$ is the Dirac probability measure concentrated at x .

The canonical space of all possible sample paths of the state-action process is $\mathbf{\Omega} = (\mathbf{X} \times \mathbf{A})^\infty$ endowed with the product σ -algebra \mathcal{F} . The coordinate projection functions from $\mathbf{\Omega}$ to the state space \mathbf{X} , the action space \mathbf{A} , and \mathbf{H}_n for $n \geq 0$ are respectively denoted by X_n , A_n , and H_n . We will refer to $\{X_n\}_{n \in \mathbb{N}}$ as to state process and $\{A_n\}_{n \in \mathbb{N}}$ as the action process. It is a well known result that for every policy $\pi \in \mathbf{\Pi}$ and any initial probability measure λ on $(\mathbf{X}, \mathfrak{X})$ there exists a unique probability measure $\mathbb{P}_{\lambda, \pi}$ on $(\mathbf{\Omega}, \mathcal{F})$ such that $\mathbb{P}_{\lambda, \pi}(\mathbf{K}^\infty) = 1$ and such that for every $n \in \mathbb{N}$, $\Gamma \in \mathfrak{X}$, and $\Lambda \in \mathfrak{B}(\mathbf{A})$ we have $\mathbb{P}_{\lambda, \pi}(X_0 \in \Gamma) = \lambda(\Gamma)$,

$$\mathbb{P}_{\lambda, \pi}(X_{n+1} \in \Gamma | H_n, A_n) = Q(\Gamma | X_n, A_n) \quad \text{and} \quad \mathbb{P}_{\lambda, \pi}(A_n \in \Lambda | H_n) = \pi_n(\Lambda | H_n)$$

with $\mathbb{P}_{\lambda, \pi}$ -probability one. The expectation with respect to $\mathbb{P}_{\lambda, \pi}$ is denoted by $\mathbb{E}_{\lambda, \pi}$.

Definition 2.1. The hitting time T_Δ of the set Δ is given by $T_\Delta : \mathbf{\Omega} \rightarrow \mathbb{N} \cup \{\infty\}$ defined as

$$T_\Delta(x_0, a_0, x_1, a_1, \dots) = \min\{n \geq 0 : x_n \in \Delta\},$$

where the min over the empty set is defined as $+\infty$.

Next we propose the definition of an absorbing control model.

Definition 2.2. Given an initial distribution $\lambda \in \mathcal{P}(\mathbf{X})$ and $\Delta \in \mathfrak{X}$, we say that the control model $\mathfrak{M}(\lambda, \Delta)$ is absorbing if the conditions (a) and (b) below are satisfied, and we say that it is uniformly absorbing if, additionally, condition (c) holds.

- (a). For every $(x, a) \in \Delta \times \mathbf{A}$ we have $Q(\Delta|x, a) = 1$.
- (b). For any $\pi \in \mathbf{\Pi}$ the expected hitting time $\mathbb{E}_{\lambda, \pi}[T_\Delta]$ is finite.
- (c). The following limit holds:

$$\lim_{n \rightarrow \infty} \sup_{\pi \in \mathbf{M}} \sum_{t=n}^{\infty} \mathbb{P}_{\lambda, \pi}\{T_\Delta > t\} = 0.$$

We define now the occupation measures of an absorbing control model $\mathfrak{M}(\lambda, \Delta)$.

Definition 2.3. The occupation measure $\mu_{\lambda,\pi}$ of the control policy $\pi \in \mathbf{\Pi}$ is

$$\mu_{\lambda,\pi}(\Gamma) = \mathbb{E}_{\lambda,\pi} \left[\sum_{t=0}^{\infty} \mathbf{I}_{\{T_{\Delta} > t\}} \cdot \mathbf{I}_{\{(X_t, A_t) \in \Gamma\}} \right] \quad \text{for } \Gamma \in \mathfrak{X} \otimes \mathfrak{B}(\mathbf{A}).$$

We note that the occupation measure $\mu_{\lambda,\pi}$ takes into account the state-action process up to time T_{Δ} and it does not count the time spent in Δ : indeed, $\mu_{\lambda,\pi}(\Delta \times \mathbf{A}) = 0$. We will consider the following sets of occupation measures $\mathcal{O}_{\lambda} = \{\mu_{\lambda,\pi} : \pi \in \mathbf{\Pi}\}$ and also $\mathcal{O}_{\lambda}^s = \{\mu_{\lambda,\pi} : \pi \in \mathbf{M}\}$.

Lemma 2.4. For an absorbing control model $\mathfrak{M}(\lambda, \Delta)$, the set \mathcal{O}_{λ} is included in $\mathcal{M}^+(\mathbf{X} \times \mathbf{A})$ and it is bounded.

Proof. Note that the total mass of the occupation measure $\mu_{\lambda,\pi}$ is

$$\mu_{\lambda,\pi}(\mathbf{X} \times \mathbf{A}) = \mathbb{E}_{\lambda,\pi} \left[\sum_{t=0}^{\infty} \mathbf{I}_{\{T_{\Delta} > t\}} \right] = \sum_{t=0}^{\infty} \mathbb{P}_{\lambda,\pi}\{T_{\Delta} > t\} = \mathbb{E}_{\lambda,\pi}[T_{\Delta}],$$

which is finite as a consequence of item (b) in Definition 2.2. We also have $\sup_{\pi \in \mathbf{\Pi}} \mathbb{E}_{\lambda,\pi}[T_{\Delta}] < \infty$ according to [18], Sections 4.4 and 5.5 for the special case of a Borel state space or Proposition 2.4(i) in [3] for the general case of a measurable state space. This shows that the set \mathcal{O}_{λ} is bounded. \square

3. THE CHARACTERISTIC EQUATIONS

In this section, we introduce the notion of phantom measure and show that, for an absorbing MDP, phantom measures may exist. In order to avoid this pathological phenomenon, we establish our first main result (see Thm. 3.7) that provides several necessary and sufficient conditions to ensure that there are no phantom measures.

Definition 3.1. Given the absorbing control model $\mathfrak{M}(\eta, \Delta)$, we say $\mu \in \mathcal{M}^+(\mathbf{X} \times \mathbf{A})$ is a solution of the characteristic equations when

$$\mu(\mathbf{K}^c) = 0 \quad \text{and} \quad \mu^{\mathbf{X}} = (\eta + \mu Q)\mathbb{I}_{\Delta^c}. \quad (3.1)$$

We denote by \mathcal{C}_{η} the family of all solutions of the characteristic equations.

Lemma 3.2. Suppose that the control model $\mathfrak{M}(\eta, \Delta)$ is absorbing. Given any $\pi \in \mathbf{\Pi}$ and $\sigma \in \mathbf{M}$, we have

$$\mathbb{P}_{x,\sigma}\{T_{\Delta} = \infty\} = 0 \quad \text{for } \mu_{\eta,\pi}^{\mathbf{X}}\text{-almost all } x \in \mathbf{X}.$$

Proof. For notational convenience, let us introduce the functions h_t and h on \mathbf{X} taking values in $[0, 1]$ as $h_t(x) = \mathbb{P}_{x,\sigma}\{T_{\Delta} > t\}$ and $h(x) = \mathbb{P}_{x,\sigma}\{T_{\Delta} = \infty\}$ respectively. Note that h_t vanishes on Δ and also that h_t is a decreasing sequence of measurable functions which converges pointwise to $h(x) = \mathbb{P}_{x,\sigma}\{T_{\Delta} = \infty\}$. To prove the result, we proceed by contradiction and suppose that $\int h d\mu_{\eta,\pi}^{\mathbf{X}} > 0$. By definition of the occupation measure $\mu_{\eta,\pi}$ and since h vanishes on Δ , this implies that there exists some $s \geq 0$ with $\mathbb{E}_{\eta,\pi}[h(X_s)] > 0$ and, in particular, since $h_t \downarrow h$,

$$\sum_{t=0}^{\infty} \mathbb{E}_{\eta,\pi}[h_t(X_s)] = \infty. \quad (3.2)$$

Define the policy $\gamma \in \mathbf{\Pi}$ as follows: $\gamma_k = \pi_k$ for $0 \leq k \leq s-1$ and $\gamma_k = \sigma$ for every $k \geq s$ (in case that $s = 0$, this definition reduces to $\gamma = \sigma$). Consequently, the distribution of X_s is the same under $\mathbb{P}_{\eta,\pi}$ and

$\mathbb{P}_{\eta,\gamma}$, and we have

$$\mathbb{E}_{\eta,\pi}[h_t(X_s)] = \mathbb{E}_{\eta,\gamma}[h_t(X_s)] = \mathbb{P}_{\eta,\gamma}\{T_\Delta > t + s\}.$$

Recalling (3.2), this implies that the series

$$\sum_{t=0}^{\infty} \mathbb{P}_{\eta,\gamma}\{T_\Delta > t + s\} = \infty,$$

in contradiction with $\mathbb{E}_{\eta,\gamma}[T_\Delta]$ being finite. This gives the result. \square

Proposition 3.3. *Suppose that the control model $\mathfrak{M}(\eta, \Delta)$ is absorbing.*

- (i). *Given any $\pi \in \mathbf{\Pi}$, its occupation measure $\mu_{\eta,\pi}$ satisfies the characteristic equations. Therefore, we have the inclusion $\mathcal{O}_\eta \subseteq \mathcal{C}_\eta$.*
- (ii). *Given any $\pi \in \mathbf{\Pi}$, there exists $\sigma \in \mathbf{M}$ such that $\mu_{\eta,\pi} = \mu_{\eta,\sigma}$. Hence, $\mathcal{O}_\eta^s = \mathcal{O}_\eta$.*

Proof. (i). We have already shown that $\mu_{\eta,\pi}$ is a finite measure on $\mathbf{X} \times \mathbf{A}$. To prove the stated result, note that for any $B \in \mathfrak{X}$ we have

$$\mu^{\mathbf{X}}(B) = \sum_{t=0}^{\infty} \mathbb{P}_{\eta,\pi}\{T_\Delta > t, X_t \in B\} = \eta(B - \Delta) + \sum_{t=1}^{\infty} \mathbb{E}_{\eta,\pi}[\mathbb{P}_{\eta,\pi}\{T_\Delta > t, X_t \in B \mid H_{t-1}, A_{t-1}\}].$$

Observe now that for each $t \geq 1$ the conditional probability within brackets vanishes on the set $\{T_\Delta \leq t - 1\}$, and so

$$\begin{aligned} \mu^{\mathbf{X}}(B) &= \eta(B - \Delta) + \sum_{t=1}^{\infty} \mathbb{E}_{\eta,\pi}[Q(B - \Delta \mid X_{t-1}, A_{t-1}) \cdot \mathbf{I}_{\{T_\Delta > t-1\}}] \\ &= \eta(B - \Delta) + \int_{\mathbf{X} \times \mathbf{A}} Q(B - \Delta \mid x, a) \mu(dx, da), \end{aligned}$$

which can be equivalently written precisely as $\mu^{\mathbf{X}} = (\eta + \mu Q)\mathbb{I}_{\Delta^c}$. By construction of the state-action process, it is clear that $\mu(\mathbf{K}^c) = 0$. This shows that $\mu_{\eta,\pi}$ indeed verifies the characteristic equations.

(ii). We will show that for any $\pi \in \mathbf{\Pi}$ there is some $\sigma \in \mathbf{M}$ with $\mu_{\eta,\pi} = \mu_{\eta,\sigma}$. Since $\mu_{\eta,\pi}$ is in $\mathcal{M}^+(\mathbf{X} \times \mathbf{A})$ with $\mu_{\eta,\pi}(\mathbf{K}^c) = 0$, the measure $\mu_{\eta,\pi}$ can be disintegrated as

$$\mu_{\eta,\pi} = \mu_{\eta,\pi}^{\mathbf{X}} \otimes \sigma \quad \text{for some } \sigma \in \mathbf{M}. \quad (3.3)$$

It follows that the characteristic equation can be written as

$$\mu_{\eta,\pi}^{\mathbf{X}} = (\eta + \mu_{\eta,\pi} Q)\mathbb{I}_{\Delta^c} = (\eta + \mu_{\eta,\pi}^{\mathbf{X}} Q_\sigma)\mathbb{I}_{\Delta^c}.$$

Iterating this equation we obtain that for any $t \geq 0$

$$\mu_{\eta,\pi}^{\mathbf{X}} = \eta \mathbb{I}_{\Delta^c} + \sum_{k=0}^{t-1} (Q_\sigma \mathbb{I}_{\Delta^c})^k + \mu_{\eta,\pi}^{\mathbf{X}} (Q_\sigma \mathbb{I}_{\Delta^c})^t. \quad (3.4)$$

We have, for any $x \in \mathbf{X}$ and $t \geq 0$,

$$(Q_\sigma \mathbb{I}_{\Delta^c})^t(\mathbf{X}|x) = Q_\sigma^t(\Delta^c|x) = \mathbb{P}_{x,\sigma}\{T_\Delta > t\}$$

and $\mathbb{P}_{x,\sigma}\{T_\Delta > t\} \downarrow \mathbb{P}_{x,\sigma}\{T_\Delta = \infty\}$ as $t \rightarrow \infty$ for any $x \in \mathbf{X}$. Applying Lemma 3.2, we obtain that $\lim_{t \rightarrow \infty} \mu_{\eta,\pi}^{\mathbf{X}}(Q_\sigma \mathbb{I}_{\Delta^c})^t(\mathbf{X}) = 0$. Taking the limit as $t \rightarrow \infty$ in (3.4) we obtain

$$\mu_{\eta,\pi}^{\mathbf{X}} = \eta \mathbb{I}_{\Delta^c} \sum_{k=0}^{\infty} (Q_\sigma \mathbb{I}_{\Delta^c})^k = \mu_{\eta,\sigma}^{\mathbf{X}}.$$

Once we have shown that the \mathbf{X} -marginals of the occupation measures of $\mu_{\eta,\pi}$ and $\mu_{\eta,\sigma}$ coincide, we conclude the result from (3.3). \square

Definition 3.4. A measure $\mu \in \mathcal{M}^+(\mathbf{X} \times \mathbf{A})$ is called a phantom solution to the characteristic equation or simply a phantom measure if it satisfies the characteristic equations (3.1) but it is not given by the occupation measure of any policy in $\mathbf{\Pi}$, that is, $\mu \in \mathcal{C}_\eta$ but $\mu \notin \mathcal{O}_\eta$.

The following simple example illustrates the existence of phantom measures and, in particular, the existence of such measures shows that [6], Lemma 4.3 is not stated in a precise way.

Example 3.5. Let us consider the model defined by $\mathbf{X} = \{0, 1, 2\}$, $\mathbf{A} = \{a\}$, $\Delta = \{0\}$, $\eta = \delta_{\{1\}}$, and

$$Q(\{0\}|0, a) = Q(\{0\}|1, a) = Q(\{2\}|2, a) = 1.$$

A measure ν on \mathbf{X} is a solution of (3.1) if and only if

$$\nu(\{0\}) = 0, \quad \nu(\{1\}) = 1, \quad \text{and} \quad \nu(\{2\}) = K,$$

where K is an arbitrary nonnegative real number. The unique occupation measure of this model corresponds to $K = 0$ while, for $K > 0$, we obtain phantom measures. It will be shown in Theorem 4.6 that the occupation measures are characterized by (3.1) and an additional condition. This is the first notable difference with the discounted model, where the occupation measures are essentially characterized by a single equation of type (3.1), that is, $\mu(\mathbf{K}^c) = 0$ and $\mu^{\mathbf{X}} = \eta + \alpha \mu Q$ (where α is the discount factor), thus excluding the existence of phantom measures.

Definition 3.6. Let $\mathfrak{M}(\eta, \Delta)$ be an absorbing control model. Given a measure $\vartheta \in \mathcal{M}^+(\mathbf{X} \times \mathbf{A})$ we say that ϑ is $Q \mathbb{I}_{\Delta^c}$ -invariant when $\vartheta(\mathbf{K}^c) = 0$ and $\vartheta^{\mathbf{X}} = \vartheta Q \mathbb{I}_{\Delta^c}$.

Let us make some comments on this definition. Note that we are not excluding that ϑ is the null measure. We call such a measure invariant because, by disintegration, there exists $\sigma \in \mathbf{M}$ satisfying $\vartheta^{\mathbf{X}} = \vartheta^{\mathbf{X}} Q_\sigma \mathbb{I}_{\Delta^c}$ which can be written as

$$\vartheta^{\mathbf{X}}(\Delta) = 0 \quad \text{and} \quad \vartheta^{\mathbf{X}}(B) = \int_{\mathbf{X}} Q_\sigma(B|x) \vartheta^{\mathbf{X}}(dx)$$

for measurable $B \subseteq \Delta^c$, and so $\vartheta^{\mathbf{X}}$ is an invariant measure for the substochastic kernel Q_σ on Δ^c .

Theorem 3.7. Let $\mathfrak{M}(\eta, \Delta)$ be absorbing.

- (i). A measure is in \mathcal{C}_η if and only if it can be decomposed as the sum of an occupation measure $\mu_{\eta,\sigma}$ in \mathcal{O}_η^s and a $Q \mathbb{I}_{\Delta^c}$ -invariant measure $\nu \otimes \sigma$, with $\sigma \in \mathbf{M}$ and $\nu \in \mathcal{M}^+(\mathbf{X})$.
- (ii). The following statements are equivalent.
 - (a). The unique $Q \mathbb{I}_{\Delta^c}$ -invariant measure is the null measure on $\mathbf{X} \times \mathbf{A}$.

- (b). $\mathcal{C}_\eta = \mathcal{O}_\eta$, i.e., there are no phantom measures.
 (c). The set \mathcal{C}_η is bounded.

Proof. (i). Suppose that $\mu \in \mathcal{M}^+(\mathbf{X} \times \mathbf{A})$ is a solution of the characteristic equations. Proceeding as in the proof of Proposition 3.3, we derive the existence of $\sigma \in \mathbf{M}$ such that $\mu = \mu^{\mathbf{X}} \otimes \sigma$ and

$$\mu^{\mathbf{X}} = \eta \mathbb{I}_{\Delta^c} \sum_{k=0}^{t-1} (Q_\sigma \mathbb{I}_{\Delta^c})^k + \mu^{\mathbf{X}} (Q_\sigma \mathbb{I}_{\Delta^c})^t \quad (3.5)$$

for any $t \geq 1$. It follows by taking the limit as $t \rightarrow \infty$ that

$$\nu = \lim_{t \rightarrow \infty} \mu^{\mathbf{X}} (Q_\sigma \mathbb{I}_{\Delta^c})^t = \mu^{\mathbf{X}} - \mu_{\eta, \sigma}^{\mathbf{X}}$$

defines a measure in $\mathcal{M}^+(\mathbf{X} \times \mathbf{A})$ satisfying $\mu^{\mathbf{X}} = \mu_{\eta, \sigma}^{\mathbf{X}} + \nu$. Recalling that

$$\mu^{\mathbf{X}} = (\eta + \mu^{\mathbf{X}} Q_\sigma) \mathbb{I}_{\Delta^c} \quad \text{and} \quad \mu_{\eta, \sigma}^{\mathbf{X}} = (\eta + \mu_{\eta, \sigma}^{\mathbf{X}} Q_\sigma) \mathbb{I}_{\Delta^c},$$

we get that $\nu = \nu Q_\sigma \mathbb{I}_{\Delta^c}$. The measure $\vartheta = \nu \otimes \sigma$ satisfies the conditions in the statement of this proposition: $\mu = \mu_{\eta, \sigma} + \vartheta$.

Conversely, it is straightforward to check that the sum of a $Q \mathbb{I}_{\Delta^c}$ -invariant measure and an occupation measure (which satisfies the characteristic equations) satisfies itself the characteristic equations and, hence, belongs to \mathcal{C}_η . In fact, the sum of any measure in \mathcal{C}_η and a $Q \mathbb{I}_{\Delta^c}$ -invariant measure lies in \mathcal{C}_η .

(ii). The implication (a) \Rightarrow (b) follows directly from item (i), while (b) \Rightarrow (c) is derived from Lemma 2.4. We prove (c) \Rightarrow (a) by contradiction and, hence, we suppose that there exists a non-null $Q \mathbb{I}_{\Delta^c}$ -invariant measure ϑ . For any $K > 0$ we have that $K\vartheta$ is also a $Q \mathbb{I}_{\Delta^c}$ -invariant measure and so for any $\mu \in \mathcal{C}_\eta$ and every $K > 0$ we have $\mu + K\vartheta \in \mathcal{C}_\eta$, which is not compatible with \mathcal{C}_η being bounded. \square

Our next result gives some insight on the behavior of phantom measures. In particular, the first part of the result establishes, loosely speaking, that the state process $\{X_t\}$ under any policy $\pi \in \mathbf{\Pi}$ never visits the support of a non-null $Q \mathbb{I}_{\Delta^c}$ -invariant measure. At this point, recall that a $Q \mathbb{I}_{\Delta^c}$ -invariant measure is in $\mathcal{M}^+(\mathbf{X} \times \mathbf{A})$ by definition, and so it is a finite measure.

Proposition 3.8. *Suppose that $\mathfrak{M}(\eta, \Delta)$ is absorbing.*

- (i). Let ϑ be a non-null $Q \mathbb{I}_{\Delta^c}$ -invariant measure: we have $\mu_{\eta, \pi}^{\mathbf{X}} \perp \vartheta^{\mathbf{X}}$ for every $\pi \in \mathbf{\Pi}$.
 (ii). For any $\mu \in \mathcal{C}_\eta$, there exist a unique $\mu_{\eta, \sigma} \in \mathcal{O}_\eta^s$ and a unique $Q \mathbb{I}_{\Delta^c}$ -invariant measure ϑ such that

$$\mu = \mu_{\eta, \sigma} + \vartheta.$$

Proof. (i). By disintegration of ϑ there exists $\sigma \in \mathbf{M}$ with $\vartheta^{\mathbf{X}} = \vartheta^{\mathbf{X}} Q_\sigma \mathbb{I}_{\Delta^c}$. Iterating this equation we obtain $\vartheta^{\mathbf{X}} = \vartheta^{\mathbf{X}} Q_\sigma^t \mathbb{I}_{\Delta^c}$ for every $t \geq 1$ and taking the limit as $t \rightarrow \infty$ yields

$$\vartheta^{\mathbf{X}}(\mathbf{X}) = \vartheta^{\mathbf{X}}(\Delta^c) = \int_{\mathbf{X}} \mathbb{P}_{x, \sigma} \{T_\Delta = \infty\} \vartheta^{\mathbf{X}}(dx)$$

and so the set $B = \{x \in \mathbf{X} : \mathbb{P}_{x, \sigma} \{T_\Delta = \infty\} < 1\}$ satisfies $\vartheta^{\mathbf{X}}(B) = 0$. By the Lebesgue decomposition theorem there exist two finite measures $\vartheta_1^{\mathbf{X}}, \vartheta_2^{\mathbf{X}} \in \mathcal{M}^+(\mathbf{X})$ such that

$$\vartheta^{\mathbf{X}} = \vartheta_1^{\mathbf{X}} + \vartheta_2^{\mathbf{X}} \quad \text{with} \quad \vartheta_1^{\mathbf{X}} \ll \mu_{\eta, \pi}^{\mathbf{X}} \quad \text{and} \quad \vartheta_2^{\mathbf{X}} \perp \mu_{\eta, \pi}^{\mathbf{X}}.$$

Therefore,

$$\begin{aligned} \vartheta^{\mathbf{X}}(\mathbf{X}) &= \int_{\mathbf{X}} \mathbb{P}_{x,\sigma}\{T_{\Delta} = \infty\} \vartheta_1^{\mathbf{X}}(dx) + \int_{\mathbf{X}} \mathbb{P}_{x,\sigma}\{T_{\Delta} = \infty\} \vartheta_2^{\mathbf{X}}(dx) \\ &= \int_{\mathbf{X}} \mathbb{P}_{x,\sigma}\{T_{\Delta} = \infty\} \cdot \frac{d\vartheta_1^{\mathbf{X}}}{d\mu_{\eta,\pi}^{\mathbf{X}}} \mu_{\eta,\pi}^{\mathbf{X}}(dx) + \int_{\mathbf{X}} \mathbb{P}_{x,\sigma}\{T_{\Delta} = \infty\} \vartheta_2^{\mathbf{X}}(dx). \end{aligned}$$

Applying Lemma 3.2, we obtain that $\mathbb{P}_{x,\sigma}\{T_{\Delta} = \infty\}$ vanishes $\mu_{\eta,\pi}^{\mathbf{X}}$ -a.s. and so

$$\vartheta^{\mathbf{X}}(\mathbf{X}) = \int_{\mathbf{X}} \mathbb{P}_{x,\sigma}\{T_{\Delta} = \infty\} \vartheta_2^{\mathbf{X}}(dx).$$

Observe that $\vartheta_2^{\mathbf{X}}(B) = 0$ and so

$$\vartheta^{\mathbf{X}}(\mathbf{X}) = \int_{B^c} \mathbb{P}_{x,\sigma}\{T_{\Delta} = \infty\} \vartheta_2^{\mathbf{X}}(dx) = \vartheta_2^{\mathbf{X}}(B^c) = \vartheta_2^{\mathbf{X}}(\mathbf{X}).$$

This shows that the measure $\vartheta_1^{\mathbf{X}}$ is null and, therefore, $\vartheta^{\mathbf{X}} = \vartheta_2^{\mathbf{X}}$, showing the result.

(ii). Consider $\mu \in \mathcal{C}_{\eta}$. Suppose that

$$\mu = \mu_{\eta,\sigma} + \vartheta_{\sigma} = \mu_{\eta,\gamma} + \vartheta_{\gamma} \tag{3.6}$$

for $\mu_{\eta,\sigma} \in \mathcal{O}_{\eta}^s$ and $\mu_{\eta,\gamma} \in \mathcal{O}_{\eta}^s$, and where ϑ_{σ} and ϑ_{γ} are $Q\mathbb{I}_{\Delta^c}$ -invariant measures. By using standard arguments used for the special case of a Borel state space, it can be shown that the set of strategic measures $\{\mathbb{P}_{\eta,\pi}\}_{\pi \in \mathbf{\Pi}}$ is convex for the general case of a measurable state space. This implies the convexity of \mathcal{O}_{η} . Moreover, the set of $Q\mathbb{I}_{\Delta^c}$ -invariant measures is convex by definition. Therefore, by observing that $\mu = \frac{1}{2}(\mu_{\eta,\sigma} + \mu_{\eta,\gamma}) + \frac{1}{2}(\vartheta_{\sigma} + \vartheta_{\gamma})$ and from item (i), there exists a set $\Lambda \in \mathfrak{X}$ such that $\mu_{\eta,\sigma}(\Lambda^c \times \mathbf{A}) = \mu_{\eta,\gamma}(\Lambda^c \times \mathbf{A}) = 0$ and $\vartheta_{\sigma}(\Lambda \times \mathbf{A}) = \vartheta_{\gamma}(\Lambda \times \mathbf{A}) = 0$. Therefore, $\mu_{\eta,\sigma} = \mu_{\eta,\gamma}$ and $\vartheta_{\sigma} = \vartheta_{\gamma}$, showing the result. \square

Remark 3.9. In the previous proposition, the condition that the $Q\mathbb{I}_{\Delta^c}$ -invariant measure is finite is very important. Indeed, it is known that the statement of Proposition 3.8 may be false even for a σ -finite non-null $Q\mathbb{I}_{\Delta^c}$ -invariant measure: see Subsection 2.2.21 in [7].

4. COMPACTNESS OF OCCUPATION MEASURES

We will first show in Theorem 4.6 that a measure is an occupation measure if and only if it satisfies the characteristic equation (3.1) and an additional condition. Roughly speaking, a measure satisfies this additional condition if it is absolutely continuous with respect to a reference probability measure η^{β} which will be introduced in Definition 4.2. An interesting property that plays a central role in the analysis of the compactness of the set occupation measures is that the model $\mathfrak{M}(\eta^{\beta}, \Delta)$ is also absorbing (see Prop. 4.4). Secondly, we will show in Theorem 4.10 that the set of occupation measures is compact in the ws -topology if and only if the model is uniformly absorbing. In particular, this result highlights the fact that the assumption that the MDP is absorbing is not sufficient to guarantee the (relative) compactness of the set of occupation measures. This last point is illustrated by an example, at the end of this section, of an absorbing model which is not uniformly absorbing, and whose set of occupation measures \mathcal{O}_{η} is neither relatively compact nor compact. As a corollary, since a discounted MDP is uniformly absorbing, we recover the classical result stating that, under suitable continuity-compactness conditions, the set of occupation measures of a discounted model is compact.

Condition (S)

(S₁) The action set \mathbf{A} is compact and the correspondence from \mathbf{X} to \mathbf{A} given by $x \mapsto \mathbf{A}(x)$ is weakly measurable with nonempty compact values.

(S₂) For any $x \in \mathbf{X}$ and $\Gamma \in \mathfrak{X}$, the mapping $a \mapsto Q(\Gamma|x, a)$ is continuous on \mathbf{A} .

Lemma 4.1. *Suppose that the Conditions (S₁)-(S₂) are satisfied. There exists $\xi^* \in \mathbf{M}$ such that*

$$Q(\cdot|x, a) \ll Q_{\xi^*}(\cdot|x) \quad \text{for every } (x, a) \in \mathbf{K}. \quad (4.1)$$

Proof. The multifunction from \mathbf{X} to \mathbf{A} defined by $x \mapsto \mathbf{A}(x)$ is weakly measurable. By (S₁) and Corollary 18.15 in [15], we obtain the existence of a sequence $\{\xi_n\}_{n \in \mathbb{N}}$ of measurable selectors for the multifunction $x \mapsto \mathbf{A}(x)$ satisfying

$$\mathbf{A}(x) = \overline{\{\xi_n(x) : n \in \mathbb{N}\}} \quad \text{for any } x \in \mathbf{X}. \quad (4.2)$$

Define now $\xi^* \in \mathbf{M}$ by means of

$$\xi^*(da|x) = \sum_{k \in \mathbb{N}} \frac{1}{2^{k+1}} \delta_{\xi_k(x)}(da) \quad \text{so that} \quad Q_{\xi^*}(dy|x) = \sum_{k \in \mathbb{N}} \frac{1}{2^{k+1}} Q(dy|x, \xi_k(x)).$$

To prove the result, fix arbitrary $(x, a) \in \mathbf{K}$ and $B \in \mathfrak{X}$ such that $Q_{\xi^*}(B|x) = 0$. This implies that $Q(B|x, \xi_k^*(x)) = 0$ for all $k \in \mathbb{N}$. Using (S₂) and (4.2) we obtain that $Q(B|x, a) = 0$. \square

This result leads to the following definition

Definition 4.2. The probability measure $\eta^\beta \in \mathcal{P}(\mathbf{X})$ is

$$\eta^\beta = (1 - \beta) \sum_{k \in \mathbb{N}} \beta^k \eta Q_{\xi^*}^k, \quad (4.3)$$

where β is some fixed parameter with $0 < \beta < 1$.

Let us make some comments on this definition. By construction, the kernel Q_{ξ^*} somehow approximates all possible one-step transitions (recall Lem. 4.1). The measure η^β spreads Q_{ξ^*} over the time horizon. Therefore, η^β “covers” the whole MDP over space and time. That is why, as shown in our next result, η^β dominates any occupation measure.

Lemma 4.3. *Suppose that $\mathfrak{M}(\eta, \Delta)$ is absorbing and that the Conditions (S₁)-(S₂) are satisfied. For every $\pi \in \mathbf{\Pi}$ we have*

$$\mu_{\eta, \pi}^{\mathbf{X}} \ll \mu_{\eta, \xi^*}^{\mathbf{X}} \sim \eta_{\Delta^c}^\beta \ll \eta^\beta.$$

Proof. By Proposition 3.3(ii), there exists $\sigma \in \mathbf{M}$ such that

$$\mu_{\eta, \pi}^{\mathbf{X}} = \eta \sum_{k=0}^{\infty} Q_{\sigma}^k \mathbb{I}_{\Delta^c}. \quad (4.4)$$

We are going to show that for every $k \geq 1$ and $x \in \mathbf{X}$ we have $Q_{\sigma}^k(\cdot|x) \ll Q_{\xi^*}^k(\cdot|x)$. The proof is by induction.

For the case $k = 1$ suppose that $Q_{\xi^*}(B|x) = 0$. By Lemma 4.1 this implies that $Q(B|x, a) = 0$ for all $a \in \mathbf{A}(x)$ and $Q_\sigma(B|x) = 0$ follows. Assuming the result true for some k and, for $k + 1$, suppose that $Q_{\xi^*}^{k+1}(B|x) = 0$. Note that

$$Q_{\xi^*}^{k+1}(B|x) = \int_{\mathbf{X}} Q_{\xi^*}^k(B|y) Q_{\xi^*}(dy|x) = 0$$

implies that $Q_{\xi^*}^k(B|y) = 0$ for all $y \in C$ where C is such that $Q_{\xi^*}(C|x) = 1$. By the induction hypothesis, we have that $Q_\sigma^k(B|y) = 0$ for all $y \in C$ with $Q_\sigma(C|x) = 1$. This shows that, indeed, $Q_\sigma^{k+1}(B|x) = 0$.

As a direct consequence we have that

$$\eta \sum_{k=0}^{\infty} Q_\sigma^k \ll \eta \sum_{k=0}^{\infty} Q_{\xi^*}^k \sim (1 - \beta) \eta \sum_{k=0}^{\infty} \beta^k Q_{\xi^*}^k = \eta^\beta$$

and the stated result follows from (4.4). \square

Proposition 4.4. *Under Conditions (S₁)-(S₂), if the control model $\mathfrak{M}(\eta, \Delta)$ is absorbing then $\mathfrak{M}(\eta^\beta, \Delta)$ is also absorbing.*

Proof. Consider an arbitrary $\pi \in \mathbf{\Pi}$. We have the following equalities

$$\begin{aligned} \mathbb{E}_{\eta^\beta, \pi}[T_\Delta] &= \sum_{j=0}^{\infty} \mathbb{E}_{\eta^\beta, \pi}[\mathbf{I}_{\Delta^c}(X_j)] \\ &= (1 - \beta) \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \beta^k \mathbb{E}_{\eta Q_{\xi^*}^k, \pi}[\mathbf{I}_{\Delta^c}(X_j)]. \end{aligned}$$

Observe that using the policy π for the initial distribution $\eta Q_{\xi^*}^k$ is equivalent to using the policy $\gamma^k \in \mathbf{\Pi}$ given by

$$\gamma_j^k(da|x_0, a_0, \dots, x_j) = \xi^*(da|x_j) \quad \text{for } 0 \leq j < k$$

and

$$\gamma_j^k(da|x_0, a_0, \dots, x_j) = \pi_{j-k}(da|x_k, a_k, \dots, x_j) \quad \text{for } j \geq k$$

for the initial distribution η just by making a shift of k time units. Therefore,

$$\mathbb{E}_{\eta Q_{\xi^*}^k, \pi}[\mathbf{I}_{\Delta^c}(X_j)] = \mathbb{E}_{\eta, \gamma^k}[\mathbf{I}_{\Delta^c}(X_{j+k})].$$

It follows that

$$\begin{aligned} \mathbb{E}_{\eta^\beta, \pi}[T_\Delta] &= (1 - \beta) \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \beta^k \mathbb{E}_{\eta, \gamma^k}[\mathbf{I}_{\Delta^c}(X_{j+k})] \\ &= (1 - \beta) \sum_{k=0}^{\infty} \beta^k \sum_{j=0}^{\infty} \mathbb{E}_{\eta, \gamma^k}[\mathbf{I}_{\Delta^c}(X_{j+k})] \\ &\leq (1 - \beta) \sum_{k=0}^{\infty} \beta^k \mathbb{E}_{\eta, \gamma^k}[T_\Delta]. \end{aligned}$$

The model $\mathfrak{M}(\eta, \Delta)$ being absorbing, we have that

$$\sup_{\pi \in \Pi} \mathbb{E}_{\eta, \pi}[T_{\Delta}] = \mathbf{c} < \infty$$

and so $\mathbb{E}_{\eta^{\beta}, \pi}[T_{\Delta}] \leq \mathbf{c}$, which shows that $\mathfrak{M}(\eta^{\beta}, \Delta)$ is absorbing as well. \square

Proposition 4.5. *Suppose that $\mathfrak{M}(\eta, \Delta)$ is absorbing and that the Conditions (S₁)-(S₂) are satisfied. Let Γ be an arbitrary subset of \mathbf{M} and let $\{h_{\pi}\}_{\pi \in \Gamma}$ be a family of non-negative functions in $L^1(\mathbf{X}, \mathfrak{X}, \eta^{\beta})$ which are uniformly η^{β} -integrable. Under these conditions,*

$$\limsup_{t \rightarrow \infty} \sup_{\pi \in \Gamma} \int_{\mathbf{X}} Q_{\pi}^t(\Delta^c | x) h_{\pi}(x) \eta^{\beta}(dx) = 0.$$

Proof. Consider a fixed arbitrary $\epsilon > 0$. By the uniform integrability hypothesis, there exists $c_{\epsilon} > 0$ such that

$$\sup_{\pi \in \Gamma} \int_{\{x \in \mathbf{X}: h_{\pi}(x) > c_{\epsilon}\}} h_{\pi}(x) \eta^{\beta}(dx) \leq \epsilon.$$

Therefore, for any $\pi \in \Gamma$ and $t \geq 1$

$$\int_{\mathbf{X}} Q_{\pi}^t(\Delta^c | x) h_{\pi}(x) \eta^{\beta}(dx) \leq \epsilon + c_{\epsilon} \mathbb{P}_{\eta^{\beta}, \pi}\{T_{\Delta} > t\} \leq \epsilon + \frac{c_{\epsilon}}{t} \cdot \mathbb{E}_{\eta^{\beta}, \pi}[T_{\Delta}] \leq \epsilon + \frac{c_{\epsilon}}{t} \sup_{\pi \in \Pi} \mathbb{E}_{\eta^{\beta}, \pi}[T_{\Delta}].$$

By Proposition 4.4 we have that the above supremum is finite. Hence, for t sufficiently large we obtain that

$$\sup_{\pi \in \Gamma} \int_{\mathbf{X}} Q_{\pi}^t(\Delta^c | x) h_{\pi}(x) \eta^{\beta}(dx) < 2\epsilon$$

and the result follows. \square

Our next result gives a characterization of \mathcal{O}_{η} based on the probability measure η^{β} .

Theorem 4.6. *Let $\mathfrak{M}(\eta, \Delta)$ be an absorbing model satisfying the conditions (S₁)-(S₂). Given a measure $\mu \in \mathcal{C}_{\eta}$, we have $\mu \in \mathcal{O}_{\eta}$ if and only if $\mu^{\mathbf{X}} \ll \eta^{\beta}$. Equivalently,*

$$\mathcal{O}_{\eta} = \left\{ \mu \in \mathcal{M}^+(\mathbf{X} \times \mathbf{A}) : \mu(\mathbf{K}^c) = 0, \quad \mu^{\mathbf{X}} = (\eta + \mu^{\mathbf{X}} Q_{\sigma}) \mathbb{I}_{\Delta^c}, \quad \text{and} \quad \mu^{\mathbf{X}} \ll \eta^{\beta} \right\}.$$

Proof. Clearly, if $\mu \in \mathcal{O}_{\eta}$ then $\mu^{\mathbf{X}} \ll \eta^{\beta}$ as a direct consequence of Lemma 4.3.

Conversely, let us assume that $\mu \in \mathcal{C}_{\eta}$ is such that $\mu^{\mathbf{X}} \ll \eta^{\beta}$. Using Theorem 3.7(i), we can find $\sigma \in \mathbf{M}$ and a $Q_{\sigma} \mathbb{I}_{\Delta^c}$ -invariant measure $\vartheta \in \mathcal{M}^+(\mathbf{X} \times \mathbf{A})$ for the kernel $Q_{\sigma} \mathbb{I}_{\Delta^c}$ such that

$$\mu = \mu_{\eta, \sigma} + \vartheta \quad \text{and} \quad \vartheta^{\mathbf{X}} = \vartheta^{\mathbf{X}} Q_{\sigma} \mathbb{I}_{\Delta^c}.$$

This implies that $\vartheta^{\mathbf{X}} \ll \eta^{\beta}$ and so for every $t \geq 1$

$$\vartheta^{\mathbf{X}}(\mathbf{X}) = \vartheta^{\mathbf{X}}(Q_{\sigma} \mathbb{I}_{\Delta^c})^t(\mathbf{X}) = \vartheta^{\mathbf{X}} Q_{\sigma}^t(\Delta^c) = \int_{\mathbf{X}} Q_{\sigma}^t(\Delta^c | x) \frac{d\vartheta^{\mathbf{X}}}{d\eta^{\beta}}(x) \eta^{\beta}(dx).$$

Applying Proposition 4.5 to the set $\Gamma = \{\sigma\}$ and the function $d\vartheta^{\mathbf{X}}/d\eta^{\beta}$, we can take the limit as $t \rightarrow \infty$ in the last expression to obtain that $\vartheta^{\mathbf{X}}(\mathbf{X}) = 0$. This shows that, indeed, $\mu \in \mathcal{O}_{\eta}$. \square

Lemma 4.7. *Assume that the Conditions (S₁)-(S₂) hold. Let $\{\nu_\pi\}_{\pi \in \mathbf{M}}$ be a relatively s -compact subset of $\mathcal{M}^+(\mathbf{X})$. The set $\{\nu_\pi Q_\pi\}_{\pi \in \mathbf{M}}$ is relatively s -compact subset of $\mathcal{M}^+(\mathbf{X})$.*

Proof. Define

$$\mathbf{\Lambda} = \{\nu_\pi \otimes \pi\}_{\pi \in \mathbf{M}} \subseteq \mathcal{M}^+(\mathbf{X} \times \mathbf{A}).$$

By hypothesis, $\mathbf{\Lambda}^{\mathbf{X}}$ is relatively s -compact and $\mathbf{\Lambda}^{\mathbf{A}}$ is tight since \mathbf{A} is a compact metric space. Therefore, $\mathbf{\Lambda}$ is relatively sequentially ws -compact by using Theorem 2.5 in [13]. So, there exists a sequence $\{\pi_n\}$ in \mathbf{M} such that

$$\{\nu_{\pi_n} \otimes \pi_n\} \xrightarrow{ws} \mu \quad (4.5)$$

for some measure $\mu \in \mathcal{M}^+(\mathbf{X} \times \mathbf{A})$. By the limit (4.5) and since the function $(x, a) \mapsto Q(B|x, a)$ on $\mathbf{X} \times \mathbf{A}$ is a bounded Carathéodory function for every $B \in \mathfrak{X}$, it follows that

$$\nu_{\pi_n} Q_{\pi_n}(B) = \int_{\mathbf{X} \times \mathbf{A}} Q(B|x, a) d(\nu_{\pi_n} \otimes \pi_n) \rightarrow \int_{\mathbf{X} \times \mathbf{A}} Q(B|x, a) d\mu = \mu Q(B),$$

which shows that

$$\nu_{\pi_n} Q_{\pi_n} \xrightarrow{s} \mu Q,$$

establishing that $\{\nu_\pi Q_\pi\}_{\pi \in \mathbf{M}}$ is relatively sequentially s -compact. Finally, [13], Proposition 2.2 or [19], Theorem 2.6 give the result on relative s -compactness. \square

Given the initial distribution $\eta \in \mathcal{P}(\mathbf{X})$, a policy $\pi \in \mathbf{\Pi}$, and $k \in \mathbb{N}$ we define $\mu_{\eta, \pi, k}$ as the distribution of (X_k, A_k) under $\mathbb{P}_{\eta, \pi}$. In symbols, we let

$$\mu_{\eta, \pi, k} = \mathbb{P}_{\eta, \pi} \circ (X_k, A_k)^{-1} \in \mathcal{P}(\mathbf{X} \times \mathbf{A}).$$

We consider the following subsets of measures

$$\mathcal{D}_{\eta, k}^s = \{\mu_{\eta, \pi, k}\}_{\pi \in \mathbf{M}} \quad \text{for } k \in \mathbb{N},$$

which correspond to the families of state-action distributions of the process at time k for stationary policies.

Lemma 4.8. *Assume that the Conditions (S₁)-(S₂) hold. For each $k \in \mathbb{N}$, the sets $\mathcal{D}_{\eta, k}^s$ are relatively ws -compact and*

$$\lim_{n \rightarrow \infty} \sup_{\pi \in \mathbf{M}} \mu_{\eta, \pi, k}^{\mathbf{X}}(\Gamma_n) = 0 \quad (4.6)$$

for any decreasing sequence of sets $\Gamma_n \in \mathfrak{X}$ with $\Gamma_n \downarrow \emptyset$.

Proof. First of all we prove by induction that $(\mathcal{D}_{\eta, k}^s)^{\mathbf{X}}$ is relatively s -compact in $\mathcal{M}^+(\mathbf{X})$ for every $k \in \mathbb{N}$. For $k = 0$ we have $(\mathcal{D}_{\eta, 0}^s)^{\mathbf{X}} = \{\eta\}$ and the result is obvious. Assuming the result is true for some $k \geq 0$, observe that $\mu_{\eta, \pi, k+1}^{\mathbf{X}} = \mu_{\eta, \pi, k}^{\mathbf{X}} Q_\pi$ and we get the claim by applying Lemma 4.7.

Once we know that $\{\mu_{\eta, \pi, k}^{\mathbf{X}}\}_{\pi \in \mathbf{M}}$ is relatively s -compact and \mathbf{A} being a compact metric space, it follows that

$$\{\mu_{\eta, \pi, k}^{\mathbf{X}} \otimes \pi\}_{\pi \in \mathbf{M}} = \{\mu_{\eta, \pi, k}\}_{\pi \in \mathbf{M}} = \mathcal{D}_{\eta, k}^s$$

is relatively ws -compact by Theorem 5.2 in [13], showing the first claim. The last part of the result is obtained by using Lemma 4.6.5 and Theorem 4.7.25 in [20]. \square

Lemma 4.9. *Suppose $\mathfrak{M}(\eta, \Delta)$ is absorbing and that the Conditions (S₁)-(S₂) hold. The set \mathcal{O}_η is ws -closed in $\mathcal{M}^+(\mathbf{X} \times \mathbf{A})$.*

Proof. Let us consider a net $\{\nu_\alpha\}_{\alpha \in \mathcal{I}}$ in \mathcal{O}_η converging to some $\nu \in \mathcal{M}^+(\mathbf{X} \times \mathbf{A})$ in the ws -topology. From Theorem 4.6 we have that for any $\alpha \in \mathcal{I}$

$$\nu_\alpha(\mathbf{K}^c) = 0, \quad \nu_\alpha^{\mathbf{X}} = (\eta + \nu_\alpha Q)\mathbb{I}_{\Delta^c}, \quad \text{and} \quad \nu_\alpha^{\mathbf{X}} \ll \eta^\beta.$$

It easily follows that the limiting measure ν satisfies $\nu^{\mathbf{X}} \ll \eta^\beta$ and the equation $\nu^{\mathbf{X}} = (\eta + \nu Q)\mathbb{I}_{\Delta^c}$. Let us show that we also have $\nu(\mathbf{K}^c) = 0$.

Since the action sets $\mathbf{A}(x)$ are compact, we have that the function $(x, a) \mapsto \mathbf{I}_{\mathbf{K}^c}(x, a)$ is measurable, non-negative and lower semi-continuous on \mathbf{A} for each $x \in \mathbf{X}$. Let us denote by $\tilde{\mathfrak{X}}$ the completion of \mathfrak{X} with respect to λ_β . By [16], Proposition 3.9 there exists a sequence $\{\tilde{\Psi}_n\}_{n \in \mathbb{N}}$ of $\tilde{\mathfrak{X}} \otimes \mathfrak{B}(\mathbf{A})$ -measurable bounded Carathéodory functions such that $\tilde{\Psi}_n \uparrow \mathbf{I}_{\mathbf{K}^c}$ as $n \rightarrow \infty$.

Applying [21], Lemma A-6 to $\tilde{\Psi}_n$ (being both a l.s.c. and a u.s.c. integrand), for every $n \in \mathbb{N}$ there exists a measurable set $N_n \in \mathfrak{X}$ with $\eta^\beta(N_n) = 0$ and a $\mathfrak{X} \otimes \mathfrak{B}(\mathbf{A})$ -measurable bounded Carathéodory function $\tilde{\Psi}_n$ such that

$$\tilde{\Psi}_n(x, a) = \tilde{\Psi}_n(x, a) \quad \text{for all } x \in \mathbf{X} - N_n \text{ and } a \in \mathbf{A}.$$

Letting $N = \cup_{n \in \mathbb{N}} N_n$ we define $\Psi_n(x, a) = \mathbf{I}_{N^c}(x) \cdot \tilde{\Psi}_n(x, a)$ for $(x, a) \in \mathbf{X} \times \mathbf{A}$. It should be clear that each Ψ_n is a $\mathfrak{X} \otimes \mathfrak{B}(\mathbf{A})$ -measurable bounded Carathéodory function with

$$\Psi_n \uparrow \mathbf{I}_{(N^c \times \mathbf{A}) \cap \mathbf{K}^c}.$$

Hence, given $\epsilon > 0$ we can find some n_0 with

$$\nu((N^c \times \mathbf{A}) \cap \mathbf{K}^c) \leq \int_{\mathbf{X} \times \mathbf{A}} \Psi_{n_0} d\nu + \epsilon/2.$$

Since $\nu^{\mathbf{X}} \ll \eta^\beta$ and $\eta^\beta(N) = 0$ it is $\nu^{\mathbf{X}}(N) = 0$ and, in fact, we have

$$\nu(\mathbf{K}^c) \leq \int_{\mathbf{X} \times \mathbf{A}} \Psi_{n_0} d\nu + \epsilon/2.$$

There exists some α_0 such that for every $\alpha \succ \alpha_0$

$$\nu(\mathbf{K}^c) \leq \int_{\mathbf{X} \times \mathbf{A}} \Psi_{n_0} d\nu + \epsilon/2 \leq \int_{\mathbf{X} \times \mathbf{A}} \Psi_{n_0} d\nu_\alpha + \epsilon \leq \nu_\alpha((N^c \times \mathbf{A}) \cap \mathbf{K}^c) + \epsilon.$$

Since the last measure equals zero we have shown that $\nu(\mathbf{K}^c) \leq \epsilon$. As $\epsilon > 0$ was arbitrary, we conclude that $\nu(\mathbf{K}^c) = 0$. This completes the proof that $\nu \in \mathcal{O}_\eta$ and, therefore, \mathcal{O}_η is closed. \square

Theorem 4.10. *Suppose $\mathfrak{M}(\eta, \Delta)$ is absorbing and that the Conditions (S₁)-(S₂) are satisfied. The following statements are equivalent.*

- (a) $\mathfrak{M}(\eta, \Delta)$ is uniformly absorbing.
- (b) The set \mathcal{O}_η is relatively ws -compact.
- (c) The set \mathcal{O}_η is ws -compact.

Proof. (a) \Rightarrow (b). The set of \mathbf{A} -marginal measures of \mathcal{O}_η is tight since \mathbf{A} is compact. Recalling that $\mathcal{O}_\eta = \mathcal{O}_\eta^s$, by Proposition 3.3(ii) and applying [13], Theorem 5.2 we need to show that the set of \mathbf{X} -marginal measures of \mathcal{O}_η^s is relatively s -compact or, equivalently, that

$$\lim_{n \rightarrow \infty} \sup_{\pi \in \mathbf{M}} \mu_{\eta, \pi}^{\mathbf{X}}(\Gamma_n) = 0 \quad (4.7)$$

for any decreasing sequence of sets $\Gamma_n \in \mathfrak{X}$ with $\Gamma_n \downarrow \emptyset$ (see Lem. 4.6.5 and Thm. 4.7.25 in [20]). Now, observe that for every $k \in \mathbb{N}$

$$\sup_{\pi \in \mathbf{M}} \mu_{\eta, \pi}^{\mathbf{X}}(\Gamma_n) \leq \sum_{t=0}^k \sup_{\pi \in \mathbf{M}} \mu_{\eta, \pi, t}^{\mathbf{X}}(\Gamma_n) + \sup_{\pi \in \mathbf{M}} \sum_{t > k} \mathbb{P}_{\eta, \pi, \mu} \{T_\Delta > t\}. \quad (4.8)$$

We have that $\sup_{\pi \in \mathbf{M}} \sum_{t > k} \mathbb{P}_{\eta, \pi, \mu} \{T_\Delta > t\}$ as $k \rightarrow \infty$ since $\mathfrak{M}(\eta, \Delta)$ is uniformly absorbing by hypothesis. Moreover, for each $k \in \mathbb{N}$, the leftmost term on the right-hand side of (4.8) converges to zero as $n \rightarrow \infty$ by using Lemma 4.8 showing that the limit (4.7) holds. This shows that \mathcal{O}_η is relatively compact for the ws -topology.

(b) \Rightarrow (a). Since \mathcal{O}_η is relatively compact for the ws -topology, it follows from [13], Theorem 5.2 that the set of \mathbf{X} -marginal measures of $\mathcal{O}_\eta = \mathcal{O}_\eta^s$ is relatively s -compact. Recalling that $\mu_{\eta, \pi}^{\mathbf{X}} \ll \eta^\beta$ for any $\pi \in \mathbf{M}$, using Lemma 4.3, Proposition 2.2 in [13], and Corollary 2.7 in [19], we get that the family $\{h_\pi\}_{\pi \in \mathbf{M}}$ of density functions $h_\pi = d\mu_{\eta, \pi}^{\mathbf{X}}/d\eta^\beta$ is uniformly η^β -integrable. Now, observe that for $\pi \in \mathbf{M}$

$$\sum_{k=t}^{\infty} \mathbb{P}_{\eta, \pi} \{T_\Delta > k\} = \sum_{j=0}^{\infty} \mathbb{P}_{\eta Q_\pi^j, \pi} \{T_\Delta > t\} = \mu_{\eta, \pi}^{\mathbf{X}} Q_\pi^t(\Delta^c) = \int_{\mathbf{X}} Q_\pi^t(\Delta^c | x) h_\pi(x) \eta^\beta(dx)$$

and by using Proposition 4.5 we can conclude that the rightmost term in the previous equation converges to zero uniformly in $\pi \in \mathbf{M}$ as $t \rightarrow \infty$. This establishes that $\mathfrak{M}(\eta, \Delta)$ is indeed uniformly absorbing.

(b) \Leftrightarrow (c). This is obvious using Lemma 4.9 and recalling that $\mathcal{O}_\eta = \mathcal{O}_\eta^s$ □

Example 4.11. In this example, we show that for a specific model which is not uniformly absorbing, the set of occupation measures \mathcal{O}_η is not relatively compact and so, not compact. Consider the model with state and action spaces

$$\mathbf{X} = \{(i, j) \in \mathbb{N} \times \mathbb{N} : 0 \leq j \leq 2^i - 1\} \cup \{\bar{x}\} \quad \text{and} \quad \mathbf{A} = \{c, s\},$$

with $\mathbf{A}(x) = \{c, s\}$ for $x \in \mathbb{N} \times \{0\}$ and $\mathbf{A}(x) = \{s\}$ otherwise. The transition kernel is given by:

- $Q(\bar{x} | (i, 0), c) = \frac{1}{2}$ and $Q((i+1, 0) | (i, 0), c) = \frac{1}{2}$ for $i \in \mathbb{N}$;
- $Q((i, j+1) | (i, j), s) = 1$ for $i \in \mathbb{N}$ and $0 \leq j < 2^i - 1$;
- $Q(\bar{x} | (i, 2^i - 1), s) = 1$ for $i \in \mathbb{N}$;
- $Q(\bar{x} | \bar{x}, s) = 1$.

In [5], Example 3.13, it has been shown that this model is absorbing to $\Delta = \{\bar{x}\}$ with initial distribution $\eta = \delta_{(0,0)}$ but not uniformly absorbing. Our objective is to illustrate by means of this example that the set of occupation measures \mathcal{O}_η is not compact for the ws -topology, as can be derived from Theorem 4.10.

Let us consider the sequence $\{\gamma_t\}_{t \in \mathbb{N}}$ of nonrandomized stationary policies in \mathbf{M} defined as follows. The policy $\gamma_t(da|x)$ takes the action c for all states $(i, 0)$ for $0 \leq i < t$ and takes the action s for all states $(i, 0)$ with $i \geq t$.

If \mathcal{O}_η were relatively compact, then for any decreasing sequence of sets $\Gamma_n \downarrow \emptyset$ we would have

$$\lim_{n \rightarrow \infty} \sup_{t \in \mathbb{N}} \mu_{\eta, \gamma_t}^{\mathbf{X}}(\Gamma_n) = 0. \quad (4.9)$$

Let us consider

$$\Gamma_n = \{(i, j) \in \mathbf{X} : i \geq n \text{ and } 1 \leq j \leq 2^i - 1\}$$

which indeed satisfies $\Gamma_n \downarrow \emptyset$. Observe that

$$\mathbb{P}_{\eta, \gamma_n}\{X_{n+j} = (n, j)\} = \frac{1}{2^n} \quad \text{for } n \geq 1 \text{ and } j \in \{0, \dots, 2^n - 1\}.$$

Therefore,

$$\sup_{t \in \mathbb{N}} \mu_{\eta, \gamma_t}^{\mathbf{X}}(\Gamma_n) \geq \mu_{\eta, \gamma_n}^{\mathbf{X}}(\Gamma_n) = 2^n \cdot \frac{1}{2^n} = 1,$$

contradicting (4.9), implying that \mathcal{O}_η is not relatively compact. This exhibits that, for this control model which is not uniformly absorbing, the set of occupation measures \mathcal{O}_η is not compact.

Note that, in this example, the weak topology and the *ws*-topology on $\mathcal{P}(\mathbf{X} \times \mathbf{A})$ are the same. In view of this example, the result stated in [6], Lemma 4.7 is not accurate since, in order to have compactness of occupation measures, the absorbing property does not suffice, while the uniformly absorbing condition is precisely the necessary and sufficient condition.

Acknowledgements. Supported by grant PID2021-122442NB-I00 from the Spanish *Ministerio de Ciencia e Innovación*.

REFERENCES

- [1] W. He and Y. Sun, Stationary Markov perfect equilibria in discounted stochastic games. *J. Econom. Theory* **169** (2017) 35–61.
- [2] A.S. Nowak and T. Raghavan, Existence of stationary correlated equilibria with symmetric information for discounted stochastic games. *Math. Oper. Res.* **17** (1992) 519–526.
- [3] F. Dufour and T. Prieto-Rumeau, Nash equilibria for total expected reward absorbing Markov games: the constrained and unconstrained cases. *Appl. Math. Optim.* **89** (2024) 34.
- [4] E. Altman, Constrained Markov Decision Processes. Stochastic Modeling. Chapman & Hall/CRC, Boca Raton, FL (1999).
- [5] E.A. Feinberg and A.B. Piunovskiy, Sufficiency of deterministic policies for atomless discounted and uniformly absorbing MDPs with multiple criteria. *SIAM J. Control Optim.* **57** (2019) 163–191.
- [6] E.A. Feinberg and U.G. Rothblum, Splitting randomized stationary policies in total-reward Markov decision processes. *Math. Oper. Res.* **37** (2012) 129–153.
- [7] A.B. Piunovskiy, Examples in Markov decision processes. Vol. 2 of *Imperial College Press Optimization Series*. Imperial College Press, London (2013).
- [8] V.S. Borkar, Convex analytic methods in Markov decision processes. in Handbook of Markov decision processes. Vol. 40 of *Internat. Ser. Oper. Res. Management Sci.*. Kluwer Acad. Publ., Boston, MA (2002) 347–375.
- [9] O. Hernández-Lerma and J.-B. Lasserre, Discrete-time Markov control processes: basic optimality criteria. Vol. 30 of *Applications of Mathematics*. Springer-Verlag, New York (1996).
- [10] O. Hernández-Lerma and J.-B. Lasserre, Further topics on discrete-time Markov control processes. Vol. 42 of *Applications of Mathematics*. Springer-Verlag, New York (1999).
- [11] A.B. Piunovskiy, Optimal control of random sequences in problems with constraints. Vol. 410 of *Mathematics and its Applications*. Kluwer Academic Publishers, Dordrecht (1997).
- [12] M. Schäl, On dynamic programming: compactness of the space of policies. *Stochastic Processes Appl.* **3** (1975) 345–364.
- [13] E.J. Balder, On *ws*-convergence of product measures. *Math. Oper. Res.* **26** (2001) 494–518.
- [14] J. Neveu, *Bases mathématiques du calcul des probabilités*. Masson et Cie, Éditeurs, Paris (1970). Préface de R. Fortet, Deuxième édition, revue et corrigée.
- [15] C.D. Aliprantis and K.C. Border, Infinite Dimensional Analysis, 3rd edn. Springer, Berlin (2006).

- [16] L.C. Florescu and C. Godet-Thobie, Young Measures and Compactness in Measure Spaces. De Gruyter, Berlin (2012).
- [17] M. Valadier, Désintégration d'une mesure sur un produit. *C. R. Acad. Sci. Paris Sér. A* **276** (1973) 33–35.
- [18] E.B. Dynkin and A.A. Yushkevich, Controlled Markov Processes. Vol. 235 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin (1979).
- [19] P. Gänsler, Compactness and sequential compactness in spaces of measures. *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **17** (1971) 124–146.
- [20] V.I. Bogachev, Measure Theory. Vol. I, II. Springer-Verlag, Berlin (2007).
- [21] E.J. Balder, A general approach to lower semicontinuity and lower closure in optimal control theory. *SIAM J. Control Optim.* **22** (1984) 570–598.



Please help to maintain this journal in open access!

This journal is currently published in open access under the Subscribe to Open model (S2O). We are thankful to our subscribers and supporters for making it possible to publish this journal in open access in the current year, free of charge for authors and readers.

Check with your library that it subscribes to the journal, or consider making a personal donation to the S2O programme by contacting subscribers@edpsciences.org.

More information, including a list of supporters and financial transparency reports, is available at <https://edpsciences.org/en/subscribe-to-open-s2o>.