SIAM J. CONTROL OPTIM.
Vol. 60, No. 4, pp. 2488–2513

MARKOV DECISION PROCESSES WITH INCOMPLETE INFORMATION AND SEMIUNIFORM FELLER TRANSITION PROBABILITIES*

EUGENE A. FEINBERG†, PAVLO O. KASYANOV‡, AND MICHAEL Z. ZGUROVSKY§

Abstract. This paper deals with control of partially observable discrete-time stochastic systems. It introduces and studies Markov Decision Processes with Incomplete Information and with semiuniform Feller transition probabilities. The important feature of these models is that their classic reduction to Completely Observable Markov Decision Processes with belief states preserves semiuniform Feller continuity of transition probabilities. Under mild assumptions on cost functions, optimal policies exist, optimality equations hold, and value iterations converge to optimal values for these models. In particular, for Partially Observable Markov Decision Processes the results of this paper imply new and generalize several known sufficient conditions on transition and observation probabilities for weak continuity of transition probabilities for Markov Decision Processes with belief states, the existence of optimal policies, validity of optimality equations defining optimal policies, and convergence of value iterations to optimal values.

Key words. Markov decision process, incomplete information, semiuniform Feller transition probabilities, value iterations, optimality equations

MSC codes. Primary, 90C40; Secondary, 90C39

DOI. 10.1137/21M1442152

1. Introduction. In many control problems the state of a controlled system is not known, and decision makers know only some information about the state. This takes place in many applications including signal processing, robotics, artificial intelligence, and medicine. Except for lucky exceptions, and Kalman's filtering is among them, problems with incomplete information are known to be difficult [30]. The general approach to solving such problems was identified long ago in [1, 2, 9, 41], and it is based on constructing a controlled system whose states are posterior state distributions for the original system. These posterior distributions are often called belief probabilities or belief states. Finding an optimal policy for a problem with incomplete state observation consists of two steps: (i) finding an optimal policy for the problem with belief states, and (ii) deriving from this policy an optimal policy for the original problem. This approach was introduced in [1, 2, 9, 41] for problems with finite state, observation, and action sets, and it holds for problems with Borel state, observation, and action sets [34, 46]. If there is no optimal policy for the problem with belief states, then there is no optimal policy for the original problem.

This paper deals with optimization of expected total discounted costs for discrete-

*Received by the editors August 24, 2021; accepted for publication (in revised form) April 5, 2022; published electronically August 22, 2022. An extended abstract presenting most of the results of this submission has been accepted to *Proceedings of the* 60*th IEEE Conference on Decision and Control*, 2021, *Austin, Texas*. This extended abstract does not contain proofs.
https://doi.org/10.1137/21M1442152
Funding: The research of the second and the third authors was partially supported by the National Research Foundation of Ukraine, grant 2020.01/0283.
†Department of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, NY 11794-3600 USA (eugene.feinberg@sunysb.edu, http://www.ams.sunysb.edu/~feinberg/).
‡Institute for Applied System Analysis, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute," Peremogy ave., 37, build, 35, 03056, Kyiv, Ukraine (kasyanov@i.ua).
§National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute," Peremogy ave., 37, build, 1, 03056, Kyiv, Ukraine (mzz@kpi.ua).

time models. We describe a large class of problems, for which optimal policies exist, satisfy optimality equations, which define optimal policies, and can be found by value iterations. In particular, this paper provides sufficient conditions for weak continuity of transition probabilities for models with belief states. For a particular model of Partially Observable Markov Decision Process (POMDP), called POMDP_2 in this paper, the related studies are [19, 24, 28, 37]. As known for a long time, weak continuity of transition and observation probabilities for problems with incomplete information does not imply weak continuity of transition probabilities after the reduction to belief states. Examples are provided in [19].

Weak continuity of transition probabilities for models with belief states is an important property because these models are Markov Decision Processes (MDPs) with infinite state spaces. Optimal policies minimizing expected total discounted and undiscounted costs may not exist for such MDPs. According to [15, Theorem 2], for MDPs with nonnegative costs and, if the discount factor is less than 1, with bounded below costs, weak continuity of transition probabilities and $\mathbb{K}$-inf-compactness of cost functions imply the existence of Markov optimal policies for finite-horizon problems and the existence of stationary optimal policies for infinite-horizon problems. Under the mentioned two conditions, optimal policies satisfy optimality equations, and they can be found by value iteration starting from a zero value. For MDPs with belief states, $\mathbb{K}$-inf-compactness of cost functions follows from $\mathbb{K}$-inf-compactness of original cost functions [19, Theorem 3.3], and verifying weak continuity of transition probabilities is a nontrivial matter.

There are several models of controlled systems with incomplete state observations in the literature. Here we mostly consider a contemporary version of the original model introduced in [1, 2, 9, 41] and called a Markov Decision Process with Incomplete Information (MDPII). In this model the transitions are defined by transition probabilities $P(dw_{t+1}, dy_{t+1}|w_t, y_t, a_t)$, where vectors (w_t, y_t) represent states of the system at times $t = 0, 1, \dots,$ w_t and y_t are unobservable and observable components of the state (w_t, y_t), and a_t are actions. In more contemporary studies the research focus switched to POMDPs. As was observed in [33], there are two different POMDP models in the literature, which we call POMDP_1 and POMDP_2. For problems with finite state, observation, and control states, Platzman [33] introduced a "plant" model, which we adapt to problems with general state, observation, and control spaces and call Platzman's model. This model is more general than POMDP_1 and POMDP_2; see Figure 1.1.

Platzman's model is a particular case of an MDPII when the transition probability does not depend on observations. In other words, the transition probability in Platzman's model is $P(dw_{t+1}, dy_{t+1}|w_t, a_t)$. POMDP_i, $i = 1, 2$, are Platzman's models whose transition probabilities have special structural properties. These properties are $P(dw_{t+1}, dy_{t+1}|w_t, a_t) = Q_1(dy_{t+1}|w_t, a_t)\ P_1(dw_{t+1}|w_t, a_t)$ for POMDP_1 and $P(dw_{t+1}, dy_{t+1}|w_t, a_t) = Q_2(dy_{t+1}|a_t, w_{t+1})P_2(dw_{t+1}|w_t, a_t)$ for POMDP_2, where P_i and Q_i, $i = 1, 2$, are transition and observation kernels, respectively. Figure 1.1 illustrates the relations between definitions of these four models based on the generality of the transition probabilities $P(dw_{t+1}, dy_{t+1}|w_t, y_t, a_t)$. In particular, references [29, 43, 44] considered POMDP_1, and references [19, 24, 28] considered POMDP_2.

Belief-MDPs for MDPIIs are called Markov Decision Processes with Complete Information (MDPCIs) in this paper. As mentioned above, the reduction of an MDPII with Borel state, action, and observation sets to an MDPCI was introduced in [34, 46]. The reduction of a POMDP_2 to a completely observable belief-MDP is described in [24, Chapter 4]. The reduction of an MDPII to a POMDP_2 described in [19, section

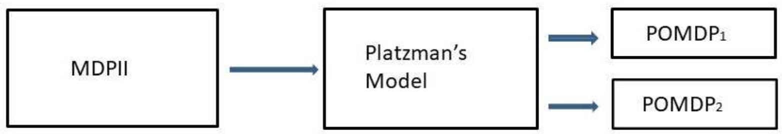

Fig. 1.1. *Relations between models of partially observable controlled Markov processes. Platzman's model is defined as a particular case of an MDPII.* POMDP_1 *and* POMDP_2 *are defined as particular cases of Platzman's model.*

8.3] and the reduction of a POMDP_2 to a completely observable belief-MDP described in [24, Chapter 4] also imply the reduction of an MDPII to an MDPCI.

This paper introduces the class of MDPIIs with semiuniform Feller transition probabilities. Theorem 6.2 states that an MDPII has a transition probability from this class if and only if the transition probability of the corresponding MDPCI also belongs to this class. Theorem 6.1 states similar results under more general conditions, which imply weaker continuity properties of value functions than the properties described in Theorem 6.2. In view of Lemma 4.2, semiuniform Feller transition probabilities are weakly continuous. In addition, under mild conditions on cost functions described in section 5, there are optimal policies for MDPs with semiuniform Feller transition probabilities. This paper provides several sufficient conditions for the existence of optimal policies, validity of optimality equations, and convergence of value iterations. In particular, the general theory implies the following sufficient conditions for weak continuity of transition probabilities for completely observable belief-MDPs corresponding to POMDPs: (i) P_i is weakly continuous and Q_i is continuous in total variation for a POMDP_i, $i = 1, 2$ (for $i = 2$ this result was established in [19]); (ii) P_2 is continuous in total variation and Q_2 is continuous in total variation in the control parameter; sufficiency of continuity of P_2 in total variation was established in [28] for uncontrolled observation kernels, that is, $Q_2(y_{t+1}|a_t, w_{t+1}) = Q_2(y_{t+1}|w_{t+1})$.

Section 2 describes MDPIIs with expected total costs, and section 3 describes their classic reduction to an MDPCI. Section 4 introduces semiuniform Feller stochastic kernels and it provides the properties of semiuniform Feller stochastic kernels. In particular, Lemma 4.2 states that semiuniform Feller stochastic kernels are weakly continuous. Semiuniform Feller stochastic kernels were introduced and studied in [21], and some of the statements of section 4 are taken from there. The basic known facts regarding the reduction of MDPIIs to MDPCIs are that this reduction preserves Borel measurability of transition probabilities [34, 46], but it does not preserve weak continuity of transition probabilities [19, Examples 4.1 and 4.3]. Section 5 describes the theory of MDPs with the expected total costs and semiuniform Feller transition probabilities. Theorem 5.3 establishes the validity of optimality equations, convergence of value iterations to optimal values, existence of Markov optimal policies for finite horizon problems, and existence of stationary optimal policies for infinite-horizon problems. Related facts for MDPs with weakly and setwise continuous transition probabilities are [15, Theorem 2] and [13, Theorem 3.1], respectively. MDPs with weakly and setwise continuous transition probabilities and with compact action sets were introduced and studied by Schäl [38, 39, 40]. Balder [3] described a common approach to these models. MDPs with weakly and setwise continuous transition probabilities and possibly noncompact action sets were studied in [15] and [13, 25], respectively. Weak continuity of transition probabilities is broadly used for problems with incomplete information, as described in this paper, and for inventory control [12]. Section 6 describes the results on the validity of optimality equations, convergence of value iterations to optimal values, and the existence of optimal policies for belief-MDPs corresponding to MDPIIs, Platzman's model, and POMDPs. Proofs of several statements are presented in Appendix A.

Platzman's model in [33], references [19, 24, 43, 44] on POMDPs, and some papers on MDPIIs including [34] considered one-step costs depending only on the unobservable states and actions. References [10, 19, 46] studied MDPIIs with one-step costs depending on unobservable states, observations, and actions. In this paper we consider one-step costs depending on unobservable states, observations, and actions. Because of this, we consider in this paper more general POMDP models than are usually considered in the literature. However, as shown in section 6, if one-step costs do not depend on observations, our results imply the known and new results for the classic Platzman's model [33] and POMDPs [19, 24, 43, 44] with belief-MDPs having smaller state spaces $\mathbb{P}(\mathbb{W})$ than state spaces $\mathbb{P}(\mathbb{W}) \times \mathbb{Y}$ for MDPCIs corresponding to Platzman's models, to POMDPs with one-step costs depending on observations, and to MDPIIs. In general, costs may depend on observations in applications. For example, for healthcare decisions during pandemics, costs depend not only on the health conditions of all the members of the population, which may be unknown, but also on the numbers of people with detected infections and on their conditions.

2. Model description. For a metric space $\mathbb{S} = (\mathbb{S}, \rho_{\mathbb{S}})$, where $\rho_{\mathbb{S}}$ is a metric, let $\tau(\mathbb{S})$ be the topology of $\mathbb{S}$ (the family of all open subsets of $\mathbb{S}$), and let $\mathcal{B}(\mathbb{S})$ be its Borel σ-field, that is, the σ-field generated by all open subsets of the metric space $\mathbb{S}$. For a subset S of $\mathbb{S}$, let $\bar{S}$ denote the *closure of* S and S^o the *interior of* S. Then $S^o \subset S \subset \bar{S}$, S^o is open, and $\bar{S}$ is closed. Let $\partial S := \bar{S} \setminus S^o$ denote the *boundary of* S. We denote by $\mathbb{P}(\mathbb{S})$ the *set of probability measures* on $(\mathbb{S}, \mathcal{B}(\mathbb{S}))$. A sequence of probability measures $\{\mu^{(n)}\}_{n=1,2,\ldots}$ from $\mathbb{P}(\mathbb{S})$ *converges weakly* to $\mu \in \mathbb{P}(\mathbb{S})$ if for every bounded continuous function f on $\mathbb{S}$,

$$\int_{\mathbb{S}} f(s)\mu^{(n)}(ds) \to \int_{\mathbb{S}} f(s)\mu(ds) \qquad \text{as} \quad n \to \infty.$$

A sequence of probability measures $\{\mu^{(n)}\}_{n=1,2,\ldots}$ from $\mathbb{P}(\mathbb{S})$ *converges in total variation* to $\mu \in \mathbb{P}(\mathbb{S})$ if

$$\sup_{C \in \mathcal{B}(\mathbb{S})} |\mu^{(n)}(C) - \mu(C)| \to 0 \qquad \text{as} \quad n \to \infty; \tag{2.1}$$

see [18, 20] for properties of these types of convergence of probability measures. Note that $\mathbb{P}(\mathbb{S})$ is a separable metric space with respect to the topology of weak convergence for probability measures, when $\mathbb{S}$ is a separable metric space; see [32, Chapter II]. Moreover, according to Bogachev [7, Theorem 8.3.2], if the metric space $\mathbb{S}$ is separable, then the topology of weak convergence of probability measures on $(\mathbb{S}, \mathcal{B}(\mathbb{S}))$ coincides with the topology generated by the *Kantorovich–Rubinshtein metric*

$$\begin{aligned} &\rho_{\mathbb{P}(\mathbb{S})}(\mu, \nu) \\ &:= \sup\left\{ \int_{\mathbb{S}} f(s)\mu(ds) - \int_{\mathbb{S}} f(s)\nu(ds) \;\middle|\; f \in \mathrm{Lip}_1(\mathbb{S}),\ \sup_{s\in\mathbb{S}} |f(s)| \le 1 \right\}, \end{aligned} \tag{2.2}$$

$\mu, \nu \in \mathbb{P}(\mathbb{S})$, where

$$\mathrm{Lip}_1(\mathbb{S}) := \{f : \mathbb{S} \to \mathbb{R},\ |f(s_1) - f(s_2)| \le \rho_{\mathbb{S}}(s_1, s_2)\ \forall s_1, s_2 \in \mathbb{S}\}.$$

For a Borel subset S of a metric space $(\mathbb{S}, \rho_{\mathbb{S}})$, we always consider the metric space (S, ρ_S), where $\rho_S := \rho_{\mathbb{S}}\big|_{S\times S}$. A subset B of S is called open (closed) in S if B is open (resp., closed) in (S, ρ_S). Of course, if $S = \mathbb{S}$, we omit "in $\mathbb{S}$." Observe that, in general, an open (closed) set in S may not be open (resp., closed). For $S \in \mathcal{B}(\mathbb{S})$ we denote by $\mathcal{B}(S)$ the Borel σ-field on (S, ρ_S). Observe that $\mathcal{B}(S) = \{S \cap B : B \in \mathcal{B}(\mathbb{S})\}$.

For metric spaces $\mathbb{S}_1$ and $\mathbb{S}_2$, a (Borel measurable) *stochastic kernel* $\Psi(ds_1|s_2)$ on $\mathbb{S}_1$ given $\mathbb{S}_2$ is a mapping $\Psi(\,\cdot\,|\,\cdot\,) : \mathcal{B}(\mathbb{S}_1) \times \mathbb{S}_2 \to [0, 1]$, such that $\Psi(\,\cdot\,|s_2)$ is a probability measure on $\mathbb{S}_1$ for any $s_2 \in \mathbb{S}_2$, and $\Psi(B|\,\cdot\,)$ is a Borel measurable function on $\mathbb{S}_2$ for any Borel set $B \in \mathcal{B}(\mathbb{S}_1)$. Another name for a stochastic kernel is a transition probability. A stochastic kernel $\Psi(ds_1|s_2)$ on $\mathbb{S}_1$ given $\mathbb{S}_2$ defines a Borel measurable mapping $s_2 \mapsto \Psi(\,\cdot\,|s_2)$ of $\mathbb{S}_2$ to the metric space $\mathbb{P}(\mathbb{S}_1)$ endowed with the topology of weak convergence. A stochastic kernel $\Psi(ds_1|s_2)$ on $\mathbb{S}_1$ given $\mathbb{S}_2$ is called *weakly continuous (continuous in total variation)* if $\Psi(\,\cdot\,|s^{(n)})$ converges weakly (in total variation) to $\Psi(\,\cdot\,|s)$ whenever $s^{(n)}$ converges to s in $\mathbb{S}_2$. For one-point sets $\{s_1\} \subset \mathbb{S}_1$, we sometimes write $\Psi(s_1|s_2)$ instead of $\Psi(\{s_1\}|s_2)$. Sometimes a weakly continuous stochastic kernel is called Feller, and a stochastic kernel continuous in total variation is called uniformly Feller [31].

Let $\mathbb{S}_1, \mathbb{S}_2$, and $\mathbb{S}_3$ be Borel subsets of Polish spaces (a Polish space is a complete separable metric space), and let Ψ on $\mathbb{S}_1 \times \mathbb{S}_2$ given $\mathbb{S}_3$ be a stochastic kernel. For each $A \in \mathcal{B}(\mathbb{S}_1)$, $B \in \mathcal{B}(\mathbb{S}_2)$, and $s_3 \in \mathbb{S}_3$, let

$$\Psi(A, B|s_3) := \Psi(A \times B|s_3). \tag{2.3}$$

In particular, we consider *marginal* stochastic kernels $\Psi(\mathbb{S}_1, \,\cdot\,|\,\cdot\,)$ on $\mathbb{S}_2$ given $\mathbb{S}_3$ and $\Psi(\,\cdot\,, \mathbb{S}_2|\,\cdot\,)$ on $\mathbb{S}_1$ given $\mathbb{S}_3$.

A *Markov Decision Process with Incomplete Information (MDPII)* (Dynkin and Yushkevich [10, Chapter 8], Rhenius [34], Yushkevich [46]; see also Rieder [35] and Bäuerle and Rieder [4] for a version of this model with transition probabilities having densities) is specified by a tuple $(\mathbb{W} \times \mathbb{Y}, \mathbb{A}, P, c)$, where

(i) $\mathbb{W} \times \mathbb{Y}$ is the *state space*, where $\mathbb{W}$ and $\mathbb{Y}$ are Borel subsets of Polish spaces, and for $(w, y) \in \mathbb{W} \times \mathbb{Y}$ the unobservable component of the state (w, y) is w, and the observable component is y;

(ii) $\mathbb{A}$ is the *action space*, which is assumed to be a Borel subset of a Polish space;

(iii) P is a stochastic kernel on $\mathbb{W} \times \mathbb{Y}$ given $\mathbb{W} \times \mathbb{Y} \times \mathbb{A}$, which determines the distribution $P(\,\cdot\,|w, y, a)$ on $\mathbb{W} \times \mathbb{Y}$ of the new state if $(w, y) \in \mathbb{W} \times \mathbb{Y}$ is the current state, and if $a \in \mathbb{A}$ is the current action, and it is assumed that the stochastic kernel P on $\mathbb{W} \times \mathbb{Y}$ given $\mathbb{W} \times \mathbb{Y} \times \mathbb{A}$ is weakly continuous in $(w, y, a) \in \mathbb{W} \times \mathbb{Y} \times \mathbb{A}$;

(iv) $P_0(\,\cdot\,|w)$ is a stochastic kernel on $\mathbb{Y}$ given $\mathbb{W}$, which determines the distribution of the observable part y_0 of the initial state, which may depend on the value of unobservable component $w_0 = w$ of the initial state;

(v) $c : \mathbb{W} \times \mathbb{Y} \times \mathbb{A} \to \overline{\mathbb{R}}_+ = [0, +\infty]$ is a Borel measurable *one-step cost function*.

The Markov decision process with incomplete information *evolves* as follows. At time $t = 0$, the unobservable component w_0 of the initial state has a given prior distribution $p \in \mathbb{P}(\mathbb{W})$. Let y_0 be the observable part of the initial state. At each time epoch $t = 0, 1, \ldots,$ if the state of the system is $(w_t, y_t) \in \mathbb{W} \times \mathbb{Y}$ and the decision-maker chooses an action $a_t \in \mathbb{A}$, then the cost $c(w_t, y_t, a_t)$ is incurred and the system moves to state (w_{t+1}, y_{t+1}) according to the transition law $P(\,\cdot\,|w_t, y_t, a_t)$.

Define the *observable histories*: $h_0 := y_0 \in \mathbb{H}_0$ and $h_t := (y_0, a_0, y_1, a_1, \ldots, y_{t-1}, a_{t-1}, y_t) \in \mathbb{H}_t$ for all $t = 1, 2, \ldots,$ where $\mathbb{H}_0 := \mathbb{Y}$ and $\mathbb{H}_t := \mathbb{H}_{t-1} \times \mathbb{A} \times \mathbb{Y}$ if

$t = 1, 2, \ldots$. Then a *policy* for the MDPII is defined as a sequence $\pi = \{\pi_t\}$ such that, for each $t = 0, 1, \ldots, \pi_t$ is a transition kernel on $\mathbb{A}$ given $\mathbb{H}_t$. Moreover, π is called *nonrandomized* if each probability measure $\pi_t(\,\cdot\,|h_t)$ is concentrated at one point. The *set of all policies* is denoted by Π. The Ionescu Tulcea theorem (Bertsekas and Shreve [5, pp. 140–141] or Hernández-Lerma and Lasserre [26, p. 178]) implies that a policy $\pi \in \Pi$, initial distribution $p \in \mathbb{P}(\mathbb{W})$, and initial state y_0 together with the transition kernel P determine a unique probability measure P_p^π on the set of all trajectories $\mathbb{H}_\infty = (\mathbb{W} \times \mathbb{Y} \times \mathbb{A})^\infty$ endowed with the product σ-field defined by Borel σ-fields of $\mathbb{W}$, $\mathbb{Y}$, and $\mathbb{A}$, respectively. The expectation with respect to this probability measure is denoted by $\mathbb{E}_p^\pi$.

Let us specify the performance criterion. For a finite horizon $T = 0, 1, \ldots,$ and for a policy $\pi \in \Pi$, let the *expected total discounted costs* be

$$v_{T,\alpha}^\pi(p) := \mathbb{E}_p^\pi \sum_{t=0}^{T-1} \alpha^t c(w_t, y_t, a_t), \quad p \in \mathbb{P}(\mathbb{W}), \tag{2.4}$$

where $\alpha \geq 0$ is the discount factor, $v_{0,\alpha}^\pi(p) = 0$.

When $T = \infty$, (2.4) defines an *infinite horizon expected total discounted cost*, and we denote it by $v_\alpha^\pi(p)$. For any function $g^\pi(p)$, including $g^\pi(p) = v_{T,\alpha}^\pi(p)$ and $g^\pi(p) = v_\alpha^\pi(p)$, define the *optimal value* $g(p) := \inf_{\pi\in\Pi} g^\pi(p)$, $p \in \mathbb{P}(\mathbb{W})$. For a given initial distribution $p \in \mathbb{P}(\mathbb{W})$ of the initial unobservable component w_0, a policy π is called *optimal* for the respective criterion if $g^\pi(p) = g(p)$ for all $p \in \mathbb{P}(\mathbb{W})$. A policy is called *$T$-horizon discount-optimal* if $g^\pi = v_{T,\alpha}^\pi$, and it is called *discount-optimal* if $g^\pi = v_\alpha^\pi$.

We remark that the standard assumptions on the discount factor are either $\alpha \in [0, 1)$ or $\alpha \in [0, 1]$. However, since we assume that transition probabilities are weakly continuous and one-step costs are $\mathbb{K}$-inf-compact or satisfy a relaxed version of $\mathbb{K}$-inf-compactness stated in Definition 5.2, the same monotonicity and continuity arguments apply to $\alpha > 0$; see the proof of Theorem 3 in [15]. In addition, if $\alpha \in [0, 1)$, then it is possible to assume that c is bounded from below rather than nonnegative. This remark also applies to MDPs with setwise continuous transition probabilities P and measurable cost functions $c(x, a)$, which are inf-compact in variable a; see [13]. Of course, if $\alpha > 1$, then for many infinite-horizon problems the objective function is equal to $+\infty$. The literature on MDPs with discount factors greater than 1 exists [27]. In particular, discount factors are relevant to opportunity costs and interest rates. Discount factors greater than 1 are relevant to negative interest rates, which are offered by some banks in some countries.

We recall that an MDP is defined by its state space, action space, transition probabilities, and one-step costs. An MDP is a particular case of an MDPII. Formally speaking, an MDP $(\mathbb{X}, \mathbb{A}, P, c)$ is an MDPII $(\mathbb{W} \times \mathbb{Y}, \mathbb{A}, P, c)$ with $\mathbb{W}$ being a singelton and $\mathbb{Y} = \mathbb{X}$, where we follow the convention that $\mathbb{W} \times \mathbb{X} = \mathbb{X}$ in this case. In addition, for an MDP an initial state is observable. For an MDP we consider an initial state x instead of the initial pair (P_0, p), where p is the probability concentrated on a single point of which $\mathbb{W}$ consists. For an MDP, a nonrandomized policy is called *Markov* if all decisions depend only on the current state and time. A Markov policy is called *stationary* if all decisions depend only on current states.

3. Reduction of MDPIIs to MDPCIs. In this section we formulate the well-known reduction of an MDPII $(\mathbb{W} \times \mathbb{Y}, \mathbb{A}, P, c)$ to a belief-MDP ([5, 10, 26, 34, 46]), which is called an MDPCI. For epoch $t = 0, 1, \ldots$ consider the joint conditional probability $R(dw_{t+1}dy_{t+1}|z_t,y_t,a_t)$ on the next state (w_{t+1},y_{t+1}) given the current state (z_t,y_t) and the current control action a_t defined by

$$R(B\times C|z,y,a):=\int_{\mathbb{W}}P(B\times C|w,y,a)z(dw), \tag{3.1}$$

$B\in\mathcal{B}(\mathbb{W})$, $C\in\mathcal{B}(\mathbb{Y})$, $(z,y,a)\in\mathbb{P}(\mathbb{W})\times\mathbb{Y}\times\mathbb{A}$. According to Bertsekas and Shreve [5, Proposition 7.27], there exists a stochastic kernel $H(z,y,a,y')[\,\cdot\,]=H(\,\cdot\,|z,y,a,y')$ on $\mathbb{W}$ given $\mathbb{P}(\mathbb{W})\times\mathbb{Y}\times\mathbb{A}\times\mathbb{Y}$ such that

$$R(B\times C|z,y,a)=\int_C H(B|z,y,a,y')R(\mathbb{W},dy'|z,y,a), \tag{3.2}$$

$B\in\mathcal{B}(\mathbb{W})$, $C\in\mathcal{B}(\mathbb{Y})$, $(z,y,a)\in\mathbb{P}(\mathbb{W})\times\mathbb{Y}\times\mathbb{A}$. The stochastic kernel $H(\,\cdot\,|z,y,a,y')$ introduced in (3.2) defines a measurable mapping $H:\mathbb{P}(\mathbb{W})\times\mathbb{Y}\times\mathbb{A}\times\mathbb{Y}\to\mathbb{P}(\mathbb{W})$. Moreover, the mapping $y'\mapsto H(z,y,a,y')$ is defined $R(\mathbb{W},\,\cdot\,|z,y,a)$-a.s. uniquely for each triple $(z,y,a)\in\mathbb{P}(\mathbb{W})\times\mathbb{Y}\times\mathbb{A}$.

Let $\mathbf{I}B$ denote the *indicator of an event* B. The MDPCI is defined as an MDP with parameters $(\mathbb{P}(\mathbb{W})\times\mathbb{Y},\mathbb{A},q,\bar{c})$, where

(i) $\mathbb{P}(\mathbb{W})\times\mathbb{Y}$ is the state space;

(ii) $\mathbb{A}$ is the action set available at all states $(z,y)\in\mathbb{P}(\mathbb{W})\times\mathbb{Y}$;

(iii) the one-step cost function $\bar{c}:\mathbb{P}(\mathbb{W})\times\mathbb{Y}\times\mathbb{A}\to\overline{\mathbb{R}}$ is defined as

$$\bar{c}(z,y,a):=\int_{\mathbb{W}}c(w,y,a)z(dw),\quad z\in\mathbb{P}(\mathbb{W}),\ y\in\mathbb{Y},\ a\in\mathbb{A}; \tag{3.3}$$

(iv) q on $\mathbb{P}(\mathbb{W})\times\mathbb{Y}$ given $\mathbb{P}(\mathbb{W})\times\mathbb{Y}\times\mathbb{A}$ is a stochastic kernel which determines the distribution of the new state as follows: for $(z,y,a)\in\mathbb{P}(\mathbb{W})\times\mathbb{Y}\times\mathbb{A}$ and for $D\in\mathcal{B}(\mathbb{P}(\mathbb{W}))$ and $C\in\mathcal{B}(\mathbb{Y})$,

$$q(D\times C|z,y,a):=\int_C \mathbf{I}\{H(z,y,a,y')\in D\}R(\mathbb{W},dy'|z,y,a); \tag{3.4}$$

see Yushkevich [46], Bertsekas and Shreve [5, Corollary 7.27.1, p. 139], or Dynkin and Yushkevich [10, p. 215] for details. Note that a particular measurable choice of a stochastic kernel H from (3.2) does not affect the definition of q in (3.4).

There is a correspondence between the policies for an MDPII $(\mathbb{W}\times\mathbb{Y},\mathbb{A},P,c)$ and for the corresponding MDPCI $(\mathbb{P}(\mathbb{W})\times\mathbb{Y},\mathbb{A},q,\bar{c})$ in the sense that for a policy in one of these models there exists a policy in another model with the same expected total costs; see [34, 46] or [24, section 4.3]. In section 6 we provide sufficient conditions for the existence of an optimal policy in the MDPCI $(\mathbb{P}(\mathbb{W})\times\mathbb{Y},\mathbb{A},q,\bar{c})$ in terms of the assumptions on the initial MDPII $(\mathbb{W}\times\mathbb{Y},\mathbb{A},P,c)$ and apply the results to Platzman's model and POMDPs. In particular, under natural conditions the existence of optimal policies and validity of optimality equations and value iterations for MDPCIs follow from Theorem 5.3. For problems with finite and infinite horizons, if ϕ is a Markov optimal policy for the MDPCI, then an optimal policy π for the MDPII can be defined as $a_t=\pi_t(h_t)=\phi_t(z_t,y_t)$, where z_t is the posterior distribution of the unobservable component w_t of the state x_t given the observations $h_t=(y_0,a_0,\dots,y_{t-1},a_{t-1},y_t)$, the initial distribution p of w_0, and $t>0$. As discussed in section 6, for Platzman's models and, in particular, for POMDPs, the values of $\phi_t(z_t,y_t)$ can be selected independent of y_t if one-step costs do not depend on observations. For infinite-horizon MDPs usually there exist stationary optimal policies, and the described scheme applies to them since stationary policies are Markov.

4. Semiuniform Feller stochastic kernels and their properties. In this section we formulate the semiuniform Feller property for stochastic kernels and describe its basic properties. In particular, Theorem 4.6 provides its equivalent definitions. Theorem 4.8 establishes a necessary and sufficient condition for a stochastic kernel to be semiuniform Feller. This condition is Assumption 4.7, whose stronger version was introduced in [18, Theorem 4.4]. Theorem 4.9 describes the preservation of semiuniform Fellerness under the integration operation.

Let $\mathbb{S}_1$, $\mathbb{S}_2$, and $\mathbb{S}_3$ be Borel subsets of Polish spaces, and let Ψ on $\mathbb{S}_1\times\mathbb{S}_2$ given $\mathbb{S}_3$ be a stochastic kernel.

Definition 4.1 (Feinberg, Kasyanov, and Zgurovsky [21]). *A stochastic kernel Ψ on $\mathbb{S}_1\times\mathbb{S}_2$ given $\mathbb{S}_3$ is semiuniform Feller if, for each sequence $\{s_3^{(n)}\}_{n=1,2,\dots}\subset\mathbb{S}_3$ that converges to s_3 in $\mathbb{S}_3$ and for each bounded continuous function f on $\mathbb{S}_1$,*

$$\lim_{n\to\infty}\sup_{B\in\mathcal{B}(\mathbb{S}_2)}\left|\int_{\mathbb{S}_1} f(s_1)\Psi(ds_1,B|s_3^{(n)})-\int_{\mathbb{S}_1} f(s_1)\Psi(ds_1,B|s_3)\right|=0. \tag{4.1}$$

We recall that the marginal measure $\Psi(ds_1,B|s_3)$, $s_3\in\mathbb{S}_3$, is defined in (2.3). The term "semiuniform" is used in Definition 4.1 because the uniform property holds in (4.1) only with respect to the second coordinate. If the uniform property holds with respect to both coordinates, then the stochastic kernel Ψ on $\mathbb{S}_1\times\mathbb{S}_2$ given $\mathbb{S}_3$ is continuous in total variation, and it is sometimes called uniformly Feller [31].

Lemma 4.2. *A semiuniform Feller stochastic kernel Ψ on $\mathbb{S}_1\times\mathbb{S}_2$ given $\mathbb{S}_3$ is weakly continuous.*

Proof. Definition 4.1 implies that for each sequence $\{s_3^{(n)}\}_{n=1,2,\dots}\subset\mathbb{S}_3$ that converges to s_3 in $\mathbb{S}_3$, for each bounded continuous function f on $\mathbb{S}_1$, and for each $B\in\mathcal{B}(\mathbb{S}_2)$,

$$\lim_{n\to\infty}\int_{\mathbb{S}_1} f(s_1)\Psi(ds_1,B|s_3^{(n)})=\int_{\mathbb{S}_1} f(s_1)\Psi(ds_1,B|s_3),$$

and, in view of Schäl [38, Theorem 3.7(iii,viii)], this property implies weak continuity of Ψ on $\mathbb{S}_1\times\mathbb{S}_2$ given $\mathbb{S}_3$. ∎

Let us consider some basic definitions.

Definition 4.3. *Let $\mathbb{S}$ be a metric space. A function $f:\mathbb{S}\to\mathbb{R}$ is called*

(i) *lower semicontinuous (l.s.c.) at a point $s\in\mathbb{S}$ if $\liminf_{s'\to s} f(s')\ge f(s)$;*
(ii) *upper semicontinuous at $s\in\mathbb{S}$ if $-f$ is l.s.c. at s;*
(iii) *continuous at $s\in\mathbb{S}$ if f is both lower and upper semicontinuous at s;*
(iv) *lower / upper semicontinuous (resp., continuous) (on $\mathbb{S}$) if f is lower/upper semicontinuous (resp., continuous) at each $s\in\mathbb{S}$.*

For a metric space $\mathbb{S}$, let $\mathbb{F}(\mathbb{S})$, $\mathbb{L}(\mathbb{S})$, and $\mathbb{C}(\mathbb{S})$ be the spaces of all real-valued functions, all real-valued l.s.c. functions, and all real-valued continuous functions, respectively, defined on the metric space $\mathbb{S}$. The following definitions are taken from [14].

Definition 4.4. *A family $\mathtt{F}\subset\mathbb{F}(\mathbb{S})$ of real-valued functions on a metric space $\mathbb{S}$ is called*

(i) *lower semiequicontinuous at a point $s\in\mathbb{S}$ if $\liminf_{s'\to s}\inf_{f\in\mathtt{F}}(f(s')-f(s))\ge 0$;*
(ii) *upper semiequicontinuous at a point $s\in\mathbb{S}$ if the family $\{-f : f\in\mathtt{F}\}$ is lower semiequicontinuous at $s\in\mathbb{S}$;*

(iii) *equicontinuous at a point* $s \in \mathbb{S}$, *if* $\mathtt{F}$ *is both lower and upper semiequicontinuous at* $s \in \mathbb{S}$, *that is,* $\lim_{s' \to s} \sup_{f \in \mathtt{F}} |f(s') - f(s)| = 0$;

(iv) *lower/upper semiequicontinuous (resp., equicontinuous) (on* $\mathbb{S}$*) if it is lower/upper semiequicontinuous (resp., equicontinuous) at all* $s \in \mathbb{S}$;

(v) *uniformly bounded (on* $\mathbb{S}$*) if there exists a constant* $M < +\infty$ *such that* $|f(s)| \le M$ *for all* $s \in \mathbb{S}$ *and for all* $f \in \mathtt{F}$.

Obviously, if a family $\mathtt{F} \subset \mathbb{F}(\mathbb{S})$ is lower semiequicontinuous, then $\mathtt{F} \subset \mathbb{L}(\mathbb{S})$. Moreover, if a family $\mathtt{F} \subset \mathbb{F}(\mathbb{S})$ is equicontinuous, then $\mathtt{F} \subset \mathbb{C}(\mathbb{S})$.

4.1. Basic properties of semiuniform Feller stochastic kernels. Let $\mathbb{S}_1$, $\mathbb{S}_2$, and $\mathbb{S}_3$ be Borel subsets of Polish spaces, and let Ψ on $\mathbb{S}_1 \times \mathbb{S}_2$ given $\mathbb{S}_3$ be a stochastic kernel. For each set $A \in \mathcal{B}(\mathbb{S}_1)$ consider the family of functions

$$\mathtt{F}_A^{\Psi} = \{s_3 \mapsto \Psi(A \times B | s_3) : B \in \mathcal{B}(\mathbb{S}_2)\} \tag{4.2}$$

mapping $\mathbb{S}_3$ into $[0,1]$. Consider the following type of continuity for stochastic kernels on $\mathbb{S}_1 \times \mathbb{S}_2$ given $\mathbb{S}_3$.

DEFINITION 4.5. *A stochastic kernel* Ψ *on* $\mathbb{S}_1 \times \mathbb{S}_2$ *given* $\mathbb{S}_3$ *is called WTV-continuous if for each* $\mathcal{O} \in \tau(\mathbb{S}_1)$ *the family of functions* $\mathtt{F}_{\mathcal{O}}^{\Psi}$ *is lower semiequicontinuous on* $\mathbb{S}_3$.

Definition 4.4 directly implies that the stochastic kernel Ψ on $\mathbb{S}_1 \times \mathbb{S}_2$ given $\mathbb{S}_3$ is WTV-continuous if and only if for each $\mathcal{O} \in \tau(\mathbb{S}_1)$

$$\liminf_{n\to\infty} \inf_{B \in \mathcal{B}(\mathbb{S}_2) \setminus \{\emptyset\}} \left(\Psi(\mathcal{O} \times B | s_3^{(n)}) - \Psi(\mathcal{O} \times B | s_3) \right) \ge 0, \tag{4.3}$$

whenever $s_3^{(n)}$ converges to s_3 in $\mathbb{S}_3$.

Since $\emptyset \in \mathcal{B}(\mathbb{S}_2)$, (4.3) holds if and only if

$$\lim_{n\to\infty} \inf_{B \in \mathcal{B}(\mathbb{S}_2)} \left(\Psi(\mathcal{O} \times B | s_3^{(n)}) - \Psi(\mathcal{O} \times B | s_3) \right) = 0. \tag{4.4}$$

WTV-continuity of the stochastic kernel Ψ on $\mathbb{S}_1 \times \mathbb{S}_2$ given $\mathbb{S}_3$ implies continuity in total variation of its marginal kernel $\Psi(\mathbb{S}_1, \cdot | \cdot)$ on $\mathbb{S}_2$ given $\mathbb{S}_3$ because

$$\begin{aligned} &\lim_{n\to\infty} \sup_{B \in \mathcal{B}(\mathbb{S}_2)} \left| \Psi(\mathbb{S}_1 \times B | s_3^{(n)}) - \Psi(\mathbb{S}_1 \times B | s_3) \right| \\ &= \lim_{n\to\infty} \sup_{B \in \mathcal{B}(\mathbb{S}_2)} \left(\Psi(\mathbb{S}_1 \times B | s_3^{(n)}) - \Psi(\mathbb{S}_1 \times B | s_3) \right) = 0, \end{aligned}$$

where the second equality follows from equality (4.4) with $\mathcal{O} := \mathbb{S}_1$ and from $\Psi(\mathbb{S}_1 \times \mathbb{S}_2 | \cdot) = 1$.

Similarly to Parthasarathy [32, Theorem II.6.1], where the necessary and sufficient conditions for weakly convergent probability measures were considered, the following theorem provides several useful equivalent definitions of the semiuniform Feller stochastic kernels.

THEOREM 4.6 (Feinberg, Kasyanov, and Zgurovsky [21, Theorem 3]). *For a stochastic kernel* Ψ *on* $\mathbb{S}_1 \times \mathbb{S}_2$ *given* $\mathbb{S}_3$ *the following conditions are equivalent:*

(a) *the stochastic kernel* Ψ *on* $\mathbb{S}_1 \times \mathbb{S}_2$ *given* $\mathbb{S}_3$ *is semiuniform Feller;*

(b) *the stochastic kernel* Ψ *on* $\mathbb{S}_1 \times \mathbb{S}_2$ *given* $\mathbb{S}_3$ *is WTV-continuous;*

(c) *if* $s_3^{(n)}$ *converges to* s_3 *in* $\mathbb{S}_3$, *then for each closed set* C *in* $\mathbb{S}_1$

$$\lim_{n\to\infty}\sup_{B\in\mathcal{B}(\mathbb{S}_2)}\Big(\Psi(C\times B|s_3^{(n)})-\Psi(C\times B|s_3)\Big)=0; \tag{4.5}$$

(d) *if $s_3^{(n)}$ converges to s_3 in $\mathbb{S}_3$, then for each $A\in\mathcal{B}(\mathbb{S}_1)$ such that $\Psi(\partial A,\mathbb{S}_2|s_3)=0$,*

$$\lim_{n\to\infty}\sup_{B\in\mathcal{B}(\mathbb{S}_2)}|\Psi(A\times B|s_3^{(n)})-\Psi(A\times B|s_3)|=0; \tag{4.6}$$

(e) *if $s_3^{(n)}$ converges to s_3 in $\mathbb{S}_3$, then, for each nonnegative bounded l.s.c. function f on $\mathbb{S}_1$,*

$$\liminf_{n\to\infty}\inf_{B\in\mathcal{B}(\mathbb{S}_2)}\left(\int_{\mathbb{S}_1}f(s_1)\Psi(ds_1,B|s_3^{(n)})-\int_{\mathbb{S}_1}f(s_1)\Psi(ds_1,B|s_3)\right)=0; \tag{4.7}$$

and each of these conditions implies continuity in total variation of the marginal kernel $\Psi(\mathbb{S}_1,\cdot|\cdot)$ on $\mathbb{S}_2$ given $\mathbb{S}_3$.

Note that, since $\emptyset\in\mathcal{B}(\mathbb{S}_2)$, (4.5) holds if and only if

$$\limsup_{n\to\infty}\sup_{B\in\mathcal{B}(\mathbb{S}_2)\setminus\{\emptyset\}}\Big(\Psi(C\times B|s_3^{(n)})-\Psi(C\times B|s_3)\Big)\le 0, \tag{4.8}$$

and similar remarks are applicable to (4.6) and (4.7) with the inequality "$\ge$" taking place in (4.7).

Let us consider the following assumption. According to Feinberg, Kasyanov, and Zgurovsky [21, Example 1], Assumption 4.7 is weaker than combined assumptions (i) and (ii) in [18, Theorem 4.4], where the base $\tau_b^{s_3}(\mathbb{S}_1)$ is the same for all $s_3\in\mathbb{S}_3$.

ASSUMPTION 4.7. *Let Ψ be a stochastic kernel on $\mathbb{S}_1\times\mathbb{S}_2$ given $\mathbb{S}_3$, and let for each $s_3\in\mathbb{S}_3$ the topology on $\mathbb{S}_1$ have a countable base $\tau_b^{s_3}(\mathbb{S}_1)$ such that*

(i) $\mathbb{S}_1\in\tau_b^{s_3}(\mathbb{S}_1)$;

(ii) *for each finite intersection $\mathcal{O}=\cap_{i=1}^k\mathcal{O}_i$, $k=1,2,\dots,$ of sets $\mathcal{O}_i\in\tau_b^{s_3}(\mathbb{S}_1)$, $i=1,2,\dots,k$, the family of functions $\mathtt{F}_{\mathcal{O}}^{\Psi}$, defined in* (4.2)*, is equicontinuous at s_3.*

Note that Assumption 4.7(ii) holds if and only if for each finite intersection $\mathcal{O}=\cap_{i=1}^k\mathcal{O}_i$ of sets $\mathcal{O}_i\in\tau_b^{s_3}(\mathbb{S}_1)$, $i=1,2,\dots,k$,

$$\lim_{n\to\infty}\sup_{B\in\mathcal{B}(\mathbb{S}_2)}\left|\Psi(\mathcal{O}\times B|s_3^{(n)})-\Psi(\mathcal{O}\times B|s_3)\right|=0 \tag{4.9}$$

if $s_3^{(n)}$ converges to s_3 in $\mathbb{S}_3$.

Theorem 4.8 shows that Assumption 4.7 is a necessary and sufficient condition for semiuniform Feller continuity.

THEOREM 4.8 (Feinberg, Kasyanov, and Zgurovsky [21, Theorem 4]). *The stochastic kernel Ψ on $\mathbb{S}_1\times\mathbb{S}_2$ given $\mathbb{S}_3$ is semiuniform Feller if and only if it satisfies Assumption* 4.7*.*

Now let $\mathbb{S}_4$ be a Borel subset of a Polish space, and let Ξ be a stochastic kernel on $\mathbb{S}_1\times\mathbb{S}_2$ given $\mathbb{S}_3\times\mathbb{S}_4$. Consider the stochastic kernel $\Xi_\int$ on $\mathbb{S}_1\times\mathbb{S}_2$ given $\mathbb{P}(\mathbb{S}_3)\times\mathbb{S}_4$ defined by

$$\begin{aligned}(4.10)\qquad &\Xi_\int(A\times B|\mu,s_4)\\ &:=\int_{\mathbb{S}_3}\Xi(A\times B|s_3,s_4)\mu(ds_3),\ A\in\mathcal{B}(\mathbb{S}_1),\ B\in\mathcal{B}(\mathbb{S}_2),\ \mu\in\mathbb{P}(\mathbb{S}_3),\ s_4\in\mathbb{S}_4.\end{aligned}$$

We observe that (4.10) becomes (3.1) with $\Xi_{\int} := R$, $\Xi := P$, $\mathbb{S}_1 := \mathbb{W}$, $\mathbb{S}_2 := \mathbb{Y}$, $\mathbb{S}_3 := \mathbb{W}$, and $\mathbb{S}_4 := \mathbb{Y}\times\mathbb{A}$. This is our main motivation for writing (4.10).

The following theorem establishes the preservation of semiuniform Fellerness of the integration operation in (4.10).

Theorem 4.9 (Feinberg, Kasyanov, and Zgurovsky [21, Theorem 5]). *The stochastic kernel $\Xi_{\int}$ on $\mathbb{S}_1\times\mathbb{S}_2$ given $\mathbb{P}(\mathbb{S}_3)\times\mathbb{S}_4$ is semiuniform Feller if and only if Ξ on $\mathbb{S}_1\times\mathbb{S}_2$ given $\mathbb{S}_3\times\mathbb{S}_4$ is semiuniform Feller.*

4.2. Continuity properties of posterior distributions. In this subsection we describe sufficient conditions for semiuniform Feller continuity of posterior distributions. The main result of this section is Theorem 4.11.

Let $\mathbb{S}_1$, $\mathbb{S}_2$, and $\mathbb{S}_3$ be Borel subsets of Polish spaces, and let Ψ on $\mathbb{S}_1\times\mathbb{S}_2$ given $\mathbb{S}_3$ be a stochastic kernel. By Bertsekas and Shreve [5, Proposition 7.27], there exists a stochastic kernel Φ on $\mathbb{S}_1$ given $\mathbb{S}_2\times\mathbb{S}_3$ such that

$$\Psi(A\times B|s_3) = \int_B \Phi(A|s_2,s_3)\Psi(\mathbb{S}_1,ds_2|s_3), \quad A\in\mathcal{B}(\mathbb{S}_1),\ B\in\mathcal{B}(\mathbb{S}_2),\ s_3\in\mathbb{S}_3. \tag{4.11}$$

The stochastic kernel $\Phi(\,\cdot\,|s_2,s_3)$ on $\mathbb{S}_1$ given $\mathbb{S}_2\times\mathbb{S}_3$ defines a measurable mapping $\Phi:\ \mathbb{S}_2\times\mathbb{S}_3\to\mathbb{P}(\mathbb{S}_1)$, where $\Phi(s_2,s_3)(\,\cdot\,)=\Phi(\,\cdot\,|s_2,s_3)$. According to Bertsekas and Shreve [5, Corollary 7.27.1], for each $s_3\in\mathbb{S}_3$ the mapping $\Phi(\,\cdot\,,s_3):\mathbb{S}_2\to\mathbb{P}(\mathbb{S}_1)$ is defined $\Psi(\mathbb{S}_1,\,\cdot\,|s_3)$-almost surely uniquely in $s_2\in\mathbb{S}_2$. Let us consider the stochastic kernel ϕ defined by

$$\begin{aligned}&\phi(D\times B|s_3)\\ &\quad := \int_B \mathbf{I}\{\Phi(s_2,s_3)\in D\}\Psi(\mathbb{S}_1,ds_2|s_3),\ D\in\mathcal{B}(\mathbb{P}(\mathbb{S}_1)),\ B\in\mathcal{B}(\mathbb{S}_2),\ s_3\in\mathbb{S}_3,\end{aligned} \tag{4.12}$$

where a particular choice of a stochastic kernel Φ satisfying (4.11) does not affect the definition of ϕ in (4.12).

In models for decision making with incomplete information, ϕ is the transition probability between belief states, which are posterior distributions of states, (3.4). Continuity properties of ϕ play the fundamental role in the studies of models with incomplete information. Theorem 4.11 characterizes such properties, and this is the reason for the title of this section. Let us consider the following assumption.

Assumption 4.10. *For a stochastic kernel Ψ on $\mathbb{S}_1\times\mathbb{S}_2$ given $\mathbb{S}_3$, there exists a stochastic kernel Φ on $\mathbb{S}_1$ given $\mathbb{S}_2\times\mathbb{S}_3$ satisfying* (4.11) *such that if a sequence $\{s_3^{(n)}\}_{n=1,2,\ldots}\subset\mathbb{S}_3$ converges to $s_3\in\mathbb{S}_3$ as $n\to\infty$, then there exist a subsequence $\{s_3^{(n_k)}\}_{k=1,2,\ldots}\subset\{s_3^{(n)}\}_{n=1,2,\ldots}$ and a measurable subset B of $\mathbb{S}_2$ such that*

$$\Psi(\mathbb{S}_1\times B|s_3)=1 \quad \textit{and} \quad \Phi(s_2,s_3^{(n_k)}) \textit{ converges weakly to } \Phi(s_2,s_3) \quad \textit{for all } s_2\in B. \tag{4.13}$$

In other words, the convergence in (4.13) *holds $\Psi(\mathbb{S}_1,\cdot|s_3)$-almost surely.*

According to Theorem 9.2.1 from [8] stating the relation between convergence in probability and almost sure convergence, Assumption 4.10 holds if and only if the following statement holds: if a sequence $\{s_3^{(n)}\}_{n=1,2,\ldots}\subset\mathbb{S}_3$ converges to $s_3\in\mathbb{S}_3$ as $n\to\infty$, then

$$\rho_{\mathbb{P}(\mathbb{S}_1)}(\Phi(s_2,s_3^{(n)}),\Phi(s_2,s_3))\to 0 \text{ in probability } \Psi(\mathbb{S}_1,ds_2|s_3), \tag{4.14}$$

where $\rho_{\mathbb{P}(\mathbb{S}_1)}$ is an arbitrary metric that induces the topology of weak convergence of probability measures on $\mathbb{S}_1$, and, in particular, $\rho_{\mathbb{P}(\mathbb{S}_1)}$ can be the Kantorovich–Rubinshtein metric defined in (2.2).

The following theorem, which is the main result of this section, provides necessary and sufficient conditions for semiuniform Fellerness of a stochastic kernel ϕ in terms of the properties of a given stochastic kernel Ψ. This theorem and the results of subsection 4.1 provide the necessary and sufficient conditions for the semiuniform Feller property of the MDPCIs in terms of the conditions on the transition kernel in the initial model for decision making with incomplete information.

THEOREM 4.11. *For a stochastic kernel* Ψ *on* $\mathbb{S}_1 \times \mathbb{S}_2$ *given* $\mathbb{S}_3$ *the following conditions are equivalent:*

(a) *the stochastic kernel* Ψ *on* $\mathbb{S}_1 \times \mathbb{S}_2$ *given* $\mathbb{S}_3$ *is semiuniform Feller;*

(b) *the marginal kernel* $\Psi(\mathbb{S}_1, \cdot\,|\,\cdot)$ *on* $\mathbb{S}_2$ *given* $\mathbb{S}_3$ *is continuous in total variation and Assumption* 4.10 *holds;*

(c) *the stochastic kernel* ϕ *on* $\mathbb{P}(\mathbb{S}_1) \times \mathbb{S}_2$ *given* $\mathbb{S}_3$ *is semiuniform Feller.*

Proof. See Appendix A for the proof. ∎

5. Markov decision processes with semiuniform Feller kernels. Let $\mathbb{X}_W$ and $\mathbb{X}_Y$ be Borel subsets of Polish spaces. In this section we consider the special class of MDPs with semiuniform Feller transition kernels, when the state space is $\mathbb{X} := \mathbb{X}_W \times \mathbb{X}_Y$. These results are important for MDPIIs with semiuniform Feller transition kernels from section 6, where $\mathbb{X}_W := \mathbb{P}(\mathbb{W})$ and $\mathbb{X}_Y = \mathbb{Y}$.

For an $\overline{\mathbb{R}}$-valued function f, defined on a nonempty subset U of a metric space $\mathbb{U}$, consider the level sets

$$\mathcal{D}_f(\lambda; U) = \{y \in U \, : \, f(y) \le \lambda\}, \qquad \lambda \in \mathbb{R}. \tag{5.1}$$

We recall that a function f is *inf-compact on* U if all the level sets $\mathcal{D}_f(\lambda; U)$ are compact.

For a metric space $\mathbb{U}$, we denote by $\mathbb{K}(\mathbb{U})$ the family of all nonempty compact subsets of $\mathbb{U}$.

DEFINITION 5.1 (Feinberg, Kasyanov, and Zadoianchuk [16, Definition 1.1]). *A function* $u : \mathbb{S}_1 \times \mathbb{S}_2 \to \overline{\mathbb{R}}$ *is called* $\mathbb{K}$*-inf-compact if this function is inf-compact on* $K \times \mathbb{S}_2$ *for each* $K \in \mathbb{K}(\mathbb{S}_1)$.

The fundamental importance of $\mathbb{K}$-inf-compactness is that Berge's theorem stating lower semicontinuity of the value function holds for possibly noncompact action sets; see Feinberg, Kasyanov, and Zadoianchuk [16, Theorem 1.2]. In particular, this fact allows us to consider the MDPII $(\mathbb{W} \times \mathbb{Y}, \mathbb{A}, P, c)$ with a possibly noncompact action space $\mathbb{A}$ and unbounded one-step cost c and examine convergence of value iterations for this model in Theorem 6.1, for Platzman's model in Corollaries 6.6 and 6.12, and for POMDPs in Corollaries 6.10 and 6.11.

DEFINITION 5.2. *A Borel measurable function* $u : \mathbb{S}_1 \times \mathbb{S}_2 \times \mathbb{S}_3 \to \overline{\mathbb{R}}$ *is called measurable* $\mathbb{K}$*-inf-compact on* $(\mathbb{S}_1 \times \mathbb{S}_3, \mathbb{S}_2)$ *or* $\mathbb{MK}(\mathbb{S}_1 \times \mathbb{S}_3, \mathbb{S}_2)$*-inf-compact if for each* $s_2 \in \mathbb{S}_2$ *the function* $(s_1, s_3) \mapsto u(s_1, s_2, s_3)$ *is* $\mathbb{K}$*-inf-compact on* $\mathbb{S}_1 \times \mathbb{S}_3$.

Consider a discrete-time MDP $(\mathbb{X}, \mathbb{A}, q, c)$ with a state space $\mathbb{X} = \mathbb{X}_W \times \mathbb{X}_Y$, an action space $\mathbb{A}$, one-step costs c, and transition probabilities q. Assume that $\mathbb{X}_W, \mathbb{X}_Y$, and $\mathbb{A}$ are *Borel subsets* of Polish spaces. Let $LW(\mathbb{X})$ be the class of all nonnegative Borel measurable functions $\varphi : \mathbb{X} \to \overline{\mathbb{R}}$ such that $w \mapsto \varphi(w, y)$ is l.s.c. on $\mathbb{X}_W$ for each

$y \in \mathbb{X}_Y$. For any $\alpha \geq 0$ and $u \in LW(\mathbb{X})$, we consider

$$\eta_u^\alpha(x,a) = c(x,a) + \alpha \int_{\mathbb{X}} u(\tilde{x}) q(d\tilde{x}|x,a), \quad (x,a) \in \mathbb{X} \times \mathbb{A}. \tag{5.2}$$

The following theorem is the main result of this section. It states the validity of optimality equations, convergence of value iterations, and existence of optimal policies for MDPs with semiuniform Feller transition probabilities and $\mathbb{MK}(\mathbb{W} \times \mathbb{A}, \mathbb{Y})$-inf-compact one-step cost functions, when the goal is to minimize expected total costs. For MDPs with weakly continuous transition probabilities the similar result is [15, Theorem 2], and for MDPs with setwise continuous transition probabilities the similar result is [13, Theorem 3.1]. Theorem 5.3 does not follow from these two results. In particular, the cost function is lower semicontinuous in [15, Theorem 2]. The corresponding assumption for Theorem 5.3 would be lower semicontinuity of the cost function c, but the function $c(w,y,a)$ may not be l.s.c. in y. Reference [13, Theorem 3.1] assumes setwise continuity of the transition probability q in the control parameter, which may not hold in this paper. Theorem 5.3 is applied in Theorem 6.1 to MDPCIs $(\mathbb{P}(\mathbb{W}) \times \mathbb{Y}, \mathbb{A}, q, \bar{c})$.

THEOREM 5.3 (expected total discounted costs). *Let us consider an MDP $(\mathbb{X}, \mathbb{A}, q, c)$ with $\mathbb{X} = \mathbb{X}_W \times \mathbb{X}_Y$, for each $y \in \mathbb{X}_Y$ the stochastic kernel $q(\,\cdot\,|\,\cdot\,, y, \,\cdot\,)$ on $\mathbb{X}$ given $\mathbb{X}_W \times \mathbb{A}$ being semiuniform Feller, and the nonnegative function $c : \mathbb{X} \times \mathbb{A} \to \overline{\mathbb{R}}$ being $\mathbb{MK}(\mathbb{X}_W \times \mathbb{A}, \mathbb{X}_Y)$-inf-compact. Then*

(i) *the functions $v_{t,\alpha}$, $t = 0, 1, \ldots,$ and v_α belong to $LW(\mathbb{X})$, and $v_{t,\alpha}(x) \uparrow v_\alpha(x)$ as $t \to +\infty$ for all $x \in \mathbb{X}$;*

(ii) *$v_{t+1,\alpha}(x) = \min_{a \in \mathbb{A}} \eta^\alpha_{v_{t,\alpha}}(x,a)$, $x \in \mathbb{X}$, $t = 0, 1, \ldots,$ where $v_{0,\alpha}(x) = 0$ for all $x \in \mathbb{X}$, and the nonempty sets $A_{t,\alpha}(x) := \{a \in \mathbb{A} : v_{t+1,\alpha}(x) = \eta^\alpha_{v_{t,\alpha}}(x,a)\}$, $x \in \mathbb{X}$, $t = 0, 1, \ldots,$ satisfy the following properties:* (a) *the graph* $\mathrm{Gr}(A_{t,\alpha}) = \{(x,a) : x \in \mathbb{X}, a \in A_{t,\alpha}(x)\}$, $t = 0, 1, \ldots,$ *is a Borel subset of* $\mathbb{X} \times \mathbb{A}$, *and* (b) *if $v_{t+1,\alpha}(x) = +\infty$, then $A_{t,\alpha}(x) = \mathbb{A}$ and if $v_{t+1,\alpha}(x) < +\infty$, then $A_{t,\alpha}(x)$ is compact;*

(iii) *for any $T = 1, 2, \ldots,$ there exists a Markov optimal T-horizon policy $(\phi_0, \ldots, \phi_{T-1})$, and if for a T-horizon Markov policy $(\phi_0, \ldots, \phi_{T-1})$ the inclusions $\phi_{T-1-t}(x) \in A_{t,\alpha}(x)$, $x \in \mathbb{X}$, $t = 0, \ldots, T-1$, hold, then this policy is T-horizon optimal;*

(iv) *$v_\alpha(x) = \min_{a \in \mathbb{A}} \eta^\alpha_{v_\alpha}(x,a)$, $x \in \mathbb{X}$, and the nonempty sets $A_\alpha(x) := \{a \in \mathbb{A} : v_\alpha(x) = \eta^\alpha_{v_\alpha}(x,a)\}$, $x \in \mathbb{X}$, satisfy the following properties:* (a) *the graph* $\mathrm{Gr}(A_\alpha) = \{(x,a) : x \in \mathbb{X}, a \in A_\alpha(x)\}$ *is a Borel subset of* $\mathbb{X} \times \mathbb{A}$, *and* (b) *if $v_\alpha(x) = +\infty$, then $A_\alpha(x) = \mathbb{A}$ and if $v_\alpha(x) < +\infty$, then $A_\alpha(x)$ is compact;*

(v) *for an infinite-horizon $T = \infty$ there exists a stationary discount-optimal policy ϕ_α, and a stationary policy is optimal if and only if $\phi_\alpha(x) \in A_\alpha(x)$ for all $x \in \mathbb{X}$.*

Proof. See Appendix A for the proof. ∎

Remark 5.4. Let us consider an MDP $(\mathbb{X}, \mathbb{A}, q, c)$ with $\mathbb{X} = \mathbb{X}_W \times \mathbb{X}_Y$, the stochastic kernel q on $\mathbb{X}$ given $\mathbb{X} \times \mathbb{A}$ being semiuniform Feller, and the nonnegative function $c : \mathbb{X} \times \mathbb{A} \to \overline{\mathbb{R}}$ being $\mathbb{K}$-inf-compact. Then, Lemma 4.2 implies that the stochastic kernel q on $\mathbb{X}$ given $\mathbb{X} \times \mathbb{A}$ is weakly continuous. Therefore, [15, Theorem 2] implies all assumptions and conclusions of Theorem 5.3 and, in addition, the functions $v_{t,\alpha}(\cdot)$ and $v_\alpha(\cdot)$ are l.s.c. for all $t = 0, 1, \ldots$ and $\alpha \geq 0$.

We also remark that if the cost function c is nonnegative, then optimality equations hold and stationary (Markov) optimal policies satisfy them for problems with an infinite (finite) horizons without any continuity assumptions on the transition probabilities q and cost function c; see, e.g., [5, Propositions 9.8 and 9.12 and Corollary 9.12.1] for $\alpha = 1$. This is also true in the following two cases: (a) $c, \alpha \geq 0$, and (b) $c \geq K > -\infty$ and $\alpha \in [0, 1)$. However, if transition probabilities and costs do not satisfy appropriate continuity assumptions, then min should be replaced with inf in the optimality equations stated in statements (ii) and (iv) of Theorem 5.3, the sets $A_{t,\alpha}(x)$ and $A_\alpha(x)$ can be empty, optimal policies may not exist, and, though a limit of value iterations with zero terminal costs exists, it may not be equal to the value function; see Yu [45] and references therein on value iterations for infinite-state MDPs.

6. Total-cost optimal policies for MDPII and corollaries for Platzman's model and for POMDPs. In this section we formulate Theorems 6.1 and 6.2 stating the equivalences of semiuniform Feller continuities of the transition probability P for an MDPII, stochastic kernel R defined in (3.1), and transition probability q for the MDPCI defined in (3.4). These two theorems also provide other necessary and sufficient conditions for semiuniform Feller continuity of the stochastic kernels P, R, and q. The proofs of Theorems 6.1 and 6.2 use Theorems 4.9 and 4.11, the reduction of MDPIIs to MDPCIs established in [34, 46] and described in section 3, and [19, Theorem 3.3] stating that integration of cost functions with respect to probability measures in the argument corresponding to unobservable state variables preserves $\mathbb{K}$-inf-compactness of cost functions. Then we consider Platzman's model and POMDPs and describe sufficient conditions for weak continuity of transition kernels in the reduced models, whose states are belief probabilities, and the validity of optimality equations, convergence of value iterations, and existence of optimal policies for these models.

THEOREM 6.1. *Let* $(\mathbb{W} \times \mathbb{Y}, \mathbb{A}, P, c)$ *be an MDPII,* $(\mathbb{P}(\mathbb{W}) \times \mathbb{Y}, \mathbb{A}, q, \bar{c})$ *be its MDPCI, and* $y \in \mathbb{Y}$. *Then the following conditions are equivalent:*

(a) *Assumption* 4.7 *holds with* $\mathbb{S}_1 := \mathbb{W}$, $\mathbb{S}_2 := \mathbb{Y}$, $\mathbb{S}_3 := \mathbb{W} \times \mathbb{A}$, *and* $\Psi := P(\,\cdot\,|\,\cdot\,, y, \,\cdot\,)$;

(b) *the stochastic kernel* $P(\,\cdot\,|\,\cdot\,, y, \,\cdot\,)$ *on* $\mathbb{W}\times\mathbb{Y}$ *given* $\mathbb{W}\times\mathbb{A}$ *is semiuniform Feller;*

(c) *the stochastic kernel* $R(\,\cdot\,|\,\cdot\,, y, \,\cdot\,)$ *on* $\mathbb{W} \times \mathbb{Y}$ *given* $\mathbb{P}(\mathbb{W}) \times \mathbb{A}$ *is semiuniform Feller;*

(d) *the marginal kernel* $R(\mathbb{W}, \,\cdot\,|\,\cdot\,, y, \,\cdot\,)$ *on* $\mathbb{Y}$ *given* $\mathbb{P}(\mathbb{W}) \times \mathbb{A}$ *is continuous in total variation, and the stochastic kernel* $H(\,\cdot\,|\,\cdot\,, y, \,\cdot\,, \,\cdot\,)$ *on* $\mathbb{W}$ *given* $\mathbb{P}(\mathbb{W}) \times \mathbb{A} \times \mathbb{Y}$ *defined in* (3.2) *satisfies Assumption* 4.10*;*

(e) *the stochastic kernel* $q(\,\cdot\,|\,\cdot\,, y, \,\cdot\,)$ *on* $\mathbb{P}(\mathbb{W})\times\mathbb{Y}$ *given* $\mathbb{P}(\mathbb{W})\times\mathbb{A}$ *is semiuniform Feller.*

Moreover, if the nonnegative function c *is* $\mathbb{MK}(\mathbb{W} \times \mathbb{A}, \mathbb{Y})$*-inf-compact, and for each* $y \in \mathbb{Y}$ *anyone of the above conditions* (a)–(e) *holds, then all the assumptions and conclusions of Theorem* 5.3 *hold for the MDPCI* $(\mathbb{P}(\mathbb{W}) \times \mathbb{Y}, \mathbb{A}, q, \bar{c})$.

THEOREM 6.2. *Let* $(\mathbb{W} \times \mathbb{Y}, \mathbb{A}, P, c)$ *be an MDPII, and let* $(\mathbb{P}(\mathbb{W}) \times \mathbb{Y}, \mathbb{A}, q, \bar{c})$ *be its MDPCI. Then the following conditions are equivalent:*

(a) *Assumption* 4.7 *holds with* $\mathbb{S}_1 := \mathbb{W}$, $\mathbb{S}_2 := \mathbb{Y}$, $\mathbb{S}_3 := \mathbb{W} \times \mathbb{Y} \times \mathbb{A}$, *and* $\Psi := P$;

(b) *the stochastic kernel* P *on* $\mathbb{W} \times \mathbb{Y}$ *given* $\mathbb{W} \times \mathbb{Y} \times \mathbb{A}$ *is semiuniform Feller;*

(c) *the stochastic kernel* R *on* $\mathbb{W} \times \mathbb{Y}$ *given* $\mathbb{P}(\mathbb{W}) \times \mathbb{Y} \times \mathbb{A}$ *is semiuniform Feller;*

(d) *the marginal kernel* $R(\mathbb{W}, \,\cdot\,|\,\cdot\,)$ *on* $\mathbb{Y}$ *given* $\mathbb{P}(\mathbb{W}) \times \mathbb{Y} \times \mathbb{A}$ *is continuous in total variation, and the stochastic kernel* H *on* $\mathbb{W}$ *given* $\mathbb{P}(\mathbb{W}) \times \mathbb{Y} \times \mathbb{A} \times \mathbb{Y}$ *defined in* (3.2) *satisfies Assumption* 4.10*;*

(e) *the stochastic kernel* q *on* $\mathbb{P}(\mathbb{W})\times\mathbb{Y}$ *given* $\mathbb{P}(\mathbb{W})\times\mathbb{Y}\times\mathbb{A}$ *is semiuniform Feller.*

Moreover, if the nonnegative function c is $\mathbb{K}$-inf-compact, and any one of the above conditions (a)–(e) *holds, then all the assumptions and conclusions of Theorem* 5.3 *hold for the MDPCI* $(\mathbb{P}(\mathbb{W})\times\mathbb{Y},\mathbb{A},q,\bar{c})$*, and the functions* $v_{t,\alpha}$*,* $t=0,1,\dots,$ *and* v_α *are l.s.c. on* $\mathbb{X}$*.*

The proofs of Theorems 6.1 and 6.2 are provided in Appendix A. We recall that $c,\alpha\ge 0$ in Theorems 5.3 and 6.1. If $0\le\alpha<1$ and the function c is bounded below, then all conclusions of Theorems 5.3 and 6.1 hold with the following minor modifications: (i) the functions $v_{t,\alpha}$ and v_α are bounded below rather than nonnegative, and (ii) $v_{t,\alpha}(x)\to v_\alpha(x)$ rather than $v_{t,\alpha}(x)\uparrow v_\alpha(x)$ as $t\to\infty$. This is true for the function c bounded below by $-K>-\infty$ because such MDPII can be converted into a model with nonnegative costs by replacing costs c with $c+K$; see [19]. The suggestion to fix y in assumptions of Theorems 5.3 and 6.1 was proposed by a referee.

According to [34, 46], for each optimal policy for the MDPCI $(\mathbb{P}(\mathbb{W})\times\mathbb{Y},\mathbb{A},q,\bar{c})$ there constructively exists an optimal policy in the original MDPII $(\mathbb{W}\times\mathbb{Y},\mathbb{A},P,c)$. Reference [18, Theorem 4.4] establishes weak continuity of the transition kernel in the MDPCI under the more restrictive assumption than statement (a) of Theorem 6.1 when the countable base in Assumption 4.7 does not depend on the argument $s_3=(w,y,a)$; see also [21, Example 1]. Moreover, for any $T=1,2,\dots$ and $\alpha\ge 0$, the value functions $\tilde{V}_{T,\alpha}(z,y),\tilde{V}_\alpha(z,y)$ in the MDPCI $(\mathbb{P}(\mathbb{W})\times\mathbb{Y},\mathbb{A},q,\bar{c})$ are concave in $z\in\mathbb{P}(\mathbb{W})$. This is true because infimums of affine functions are concave functions.

The proof of Theorem 6.1 uses the following preservation property for $\mathbb{MK}(\mathbb{W}\times\mathbb{A},\mathbb{Y})$-inf-compactness.

THEOREM 6.3. *If* $c:\mathbb{W}\times\mathbb{Y}\times\mathbb{A}\to\overline{\mathbb{R}}_+$ *is an* $\mathbb{MK}(\mathbb{W}\times\mathbb{A},\mathbb{Y})$*-inf-compact function, then the function* $\bar{c}:\mathbb{P}(\mathbb{W})\times\mathbb{Y}\times\mathbb{A}\to\overline{\mathbb{R}}_+$ *defined in* (3.3) *is* $\mathbb{MK}(\mathbb{P}(\mathbb{W})\times\mathbb{A},\mathbb{Y})$*-inf-compact.*

Proof. This theorem follows from [5, Proposition 7.29] on preservation of Borel measurability and from [19, Theorem 3.3] on preservation of $\mathbb{K}$-inf-compactness. ∎

The particular case of an MDPII is a probabilistic dynamical system considered in Platzman [33].

DEFINITION 6.4. *Platzman's model is specified by an MDPII* $(\mathbb{W}\times\mathbb{Y},\mathbb{A},P,c)$*, where* P *is a stochastic kernel on* $\mathbb{W}\times\mathbb{Y}$ *given* $\mathbb{W}\times\mathbb{A}$*.*

Remark 6.5. Formally speaking, Platzman's model is an MDPII with the transition kernel $P(\,\cdot\,|w,y,\,a)$ that does not depend on y. Therefore, Theorem 6.1 implies certain corollaries for Platzman's model.

COROLLARY 6.6. *Let* $(\mathbb{W}\times\mathbb{Y},\mathbb{A},P,c)$ *be Platzman's model. Then the stochastic kernel* P *on* $\mathbb{W}\times\mathbb{Y}$ *given* $\mathbb{W}\times\mathbb{A}$ *is semiuniform Feller if and only if one of the equivalent conditions* (a)*,* (c)*,* (d)*, or* (e) *of Theorem* 6.1 *holds. Moreover, if the nonnegative function* c *is* $\mathbb{MK}(\mathbb{W}\times\mathbb{A},\mathbb{Y})$*-inf-compact and the stochastic kernel* P *on* $\mathbb{W}\times\mathbb{Y}$ *given* $\mathbb{W}\times\mathbb{A}$ *is semiuniform Feller, then all the assumptions and conclusions of Theorem* 6.1 *hold.*

Proof. According to Remark 6.5, Corollary 6.6 follows directly from Theorem 6.1. ∎

For Platzman's models we shall write $P(B\times C|w,a)$, $R(B\times C|z,a)$, $H(D|z,a,y')$, and $q(D\times C|z,a)$ instead of $P(B\times C|w,y,a)$, $R(B\times C|z,y,a)$, $H(D|z,y,a,y')$, and $q(D\times C|z,y,a)$ since these stochastic kernels do not depend on the variable y. For Platzman's models we shall also consider the marginal kernel $\hat{q}(D|z,a):=q(D,\mathbb{Y}|z,a)$

on $\mathbb{P}(\mathbb{W})$ given $\mathbb{P}(\mathbb{W})\times\mathbb{A}$. In view of (3.4), for $(z,a)\in\mathbb{P}(\mathbb{W})\times\mathbb{A}$ and for $D\in\mathcal{B}(\mathbb{P}(\mathbb{W}))$,

$$\hat{q}(D|z,a) := \int_{\mathbb{Y}} \mathbf{I}\{H(z,a,y')\in D\} R(\mathbb{W},dy'|z,a). \tag{6.1}$$

COROLLARY 6.7. *Let $(\mathbb{W}\times\mathbb{Y},\mathbb{A},P,c)$ be Platzman's model, and let the stochastic kernel P on $\mathbb{W}\times\mathbb{Y}$ given $\mathbb{W}\times\mathbb{A}$ be semiuniform Feller. Then the stochastic kernel $\hat{q}$ on $\mathbb{P}(\mathbb{W})$ given $\mathbb{P}(\mathbb{W})\times\mathbb{A}$ is weakly continuous.*

Proof. According to Corollary 6.6 and Lemma 4.2, the stochastic kernel q on $\mathbb{P}(\mathbb{W})\times\mathbb{Y}$ given $\mathbb{P}(\mathbb{W})\times\mathbb{A}$ is weakly continuous. Therefore, its marginal kernel $\hat{q}$ on $\mathbb{P}(\mathbb{W})$ given $\mathbb{P}(\mathbb{W})\times\mathbb{A}$ is also weakly continuous. ◻

As mentioned in [33], the special cases of Platzman's model include two partially observable MDPs which we denote as POMDP_1 and POMDP_2; see Definitions 6.8 and 6.9 and Figure 1.1.

Let $i=1,2$, let $\mathbb{W}$, $\mathbb{Y}$, and $\mathbb{A}$ be Borel subsets of Polish spaces, let $P_i(dw'|w,a)$ be a stochastic kernel on $\mathbb{W}$ given $\mathbb{W}\times\mathbb{A}$, let $Q_1(dy|w,a)$ be a stochastic kernel on $\mathbb{Y}$ given $\mathbb{W}\times\mathbb{A}$, let $Q_2(dy|a,w)$ be a stochastic kernel on $\mathbb{Y}$ given $\mathbb{A}\times\mathbb{W}$, let $Q_{0,i}(dy|w)$ be a stochastic kernel on $\mathbb{Y}$ given $\mathbb{W}$, and let p be a probability distribution on $\mathbb{W}$.

DEFINITION 6.8. *A* POMDP_1 $(\mathbb{W},\mathbb{Y},\mathbb{A},P_1,Q_1,c)$ *is specified by Platzman's model* $(\mathbb{W}\times\mathbb{Y},\mathbb{A},P,c)$ *with*

$$P(B\times C|w,a) := P_1(B|w,a)Q_1(C|w,a), \tag{6.2}$$

$B\in\mathcal{B}(\mathbb{W})$, $C\in\mathcal{B}(\mathbb{Y})$, $w\in\mathbb{W}$, $y\in\mathbb{Y}$, $a\in\mathbb{A}$.

Let $(\mathbb{W},\mathbb{Y},\mathbb{A},P_1,Q_1,c)$ be a POMDP_1. Then, the stochastic kernel R on $\mathbb{W}\times\mathbb{Y}$ given $\mathbb{P}(\mathbb{W})\times\mathbb{A}$, which is defined for MDPIIs in (3.1), takes the following form:

$$R(B\times C|z,a) := \int_{\mathbb{W}} Q_1(C|w,a)P_1(B|w,a)z(dw), \tag{6.3}$$

$B\in\mathcal{B}(\mathbb{W})$, $C\in\mathcal{B}(\mathbb{Y})$, $z\in\mathbb{P}(\mathbb{W})$, $a\in\mathbb{A}$.

DEFINITION 6.9. *A* POMDP_2 $(\mathbb{W},\mathbb{Y},\mathbb{A},P_2,Q_2,c)$ *is specified by Platzman's model* $(\mathbb{W}\times\mathbb{Y},\mathbb{A},P,c)$ *with*

$$P(B\times C|w,a) := \int_B Q_2(C|a,w')P_2(dw'|w,a), \tag{6.4}$$

$B\in\mathcal{B}(\mathbb{W})$, $C\in\mathcal{B}(\mathbb{Y})$, $w\in\mathbb{W}$, $y\in\mathbb{Y}$, $a\in\mathbb{A}$.

We recall that Figure 1.1 describes the relations between an MDPII, Platzman's model, POMDP_1, and POMDP_2 based on the generality of transition probabilities P. In addition, POMDP_1 and POMDP_2 are two different models. For example, for a POMDP_1 the random variables w_{t+1} and y_{t+1} are conditionally independent given the values w_t and a_t. This is not true for POMDP_2.

Other relations between these models also take place. In particular, a reduction of an MDPII to a POMDP_2 is described in [18, section 6] and in [19, section 8.3]. Therefore, in some sense an MDPII, Platzman's model, and a POMDP_2 can be viewed as equivalent models. This reduction was used in [19] to prove Theorem 8.1 there stating sufficient conditions for weak continuity of transition probabilities for MDPCIs. This reduction transforms an MDPII with a weakly continuous transition probability into a POMDP_2 with weakly continuous transition and observation probabilities.

Since weak continuity of transition and observation probabilities for POMDP$_2$ are not sufficient for continuity of transition probabilities for the corresponding belief-MDP (see [19, Example 4.1]), [19, Theorem 8.1] contains an additional assumption on the transition probability P of the MDPII. This assumption is relaxed in [18, Theorem 6.2]. As shown in [21, Example 1], semiuniform Feller continuity of the transition probability P assumed in this paper is a more general property than the assumption on P in [18, Theorem 6.2].

For a POMDP$_2$ $(\mathbb{W},\mathbb{Y},\mathbb{A},P_2,Q_2,c)$ the stochastic kernel R on $\mathbb{W}\times\mathbb{Y}$ given $\mathbb{P}(\mathbb{W})\times\mathbb{A}$, which is defined for MDPIIs in (3.1), takes the following form:

$$R(B\times C|z,a):=\int_{\mathbb{W}}\int_B Q_2(C|a,w')P_2(dw'|w,a)z(dw), \tag{6.5}$$

$B\in\mathcal{B}(\mathbb{W})$, $C\in\mathcal{B}(\mathbb{Y})$, $z\in\mathbb{P}(\mathbb{W})$, $a\in\mathbb{A}$. A POMDP$_1$ is Platzman's model with observations y_{t+1} being "random functions" of w_t and a_t, and a POMDP$_2$ is Platzman's model with observations y_{t+1} being "random functions" of a_t and w_{t+1}. Let us apply Theorem 6.1 to a POMDP$_1$ and POMDP$_2$.

Corollary 6.6 establishes necessary and sufficient conditions for semiuniform Feller continuity of the transition probabilities P for Platzman's model $(\mathbb{W}\times\mathbb{Y},\mathbb{A},P,c)$ in terms of the same property for the transition probabilities q of the respective belief-MDP $(\mathbb{P}(\mathbb{W})\times\mathbb{Y},\mathbb{A},q,\bar{c})$. Since a POMDP$_i$, $i=1,2$, is a particular case of Platzman's model, Corollary 6.6 implies the necessary and sufficient conditions for semiuniform Feller continuity of the stochastic kernel q on $\mathbb{P}(\mathbb{W})\times\mathbb{Y}$ given $\mathbb{W}\times\mathbb{A}$ in terms of the same property for the transition probability P defined in (6.2) for a POMDP$_1$ and in (6.4) for a POMDP$_2$, respectively.

Corollary 6.10. *For a* POMDP$_1$ $(\mathbb{W},\mathbb{Y},\mathbb{A},P_1,Q_1,c)$*, the following two conditions hold together:*

(a) *the stochastic kernel P_1 on $\mathbb{W}$ given $\mathbb{W}\times\mathbb{A}$ is weakly continuous;*

(b) *the stochastic kernel Q_1 on $\mathbb{Y}$ given $\mathbb{W}\times\mathbb{A}$ is continuous in total variation;*

is equivalent to semiuniform Feller continuity of the stochastic kernel P on $\mathbb{W}\times\mathbb{Y}$ given $\mathbb{W}\times\mathbb{A}$. Moreover, if these two conditions hold, then the following statements are true:

(i) *statements* (a) *and* (c)–(e) *of Theorem* 6.1 *hold;*

(ii) *if the nonnegative function $c:\mathbb{W}\times\mathbb{Y}\times\mathbb{A}\to\overline{\mathbb{R}}$ is $\mathbb{MK}(\mathbb{W}\times\mathbb{A},\mathbb{Y})$-inf-compact, then all the conclusions of Theorem* 6.1 *hold;*

(iii) *the stochastic kernel $\hat{q}$ on $\mathbb{P}(\mathbb{W})$ given $\mathbb{P}(\mathbb{W})\times\mathbb{A}$ defined in* (6.1) *is weakly continuous.*

Proof. See Appendix A for the proof. ∎

Corollary 6.11. *For a* POMDP$_2$ $(\mathbb{W},\mathbb{Y},\mathbb{A},P_2,Q_2,c)$ *each of the conditions:*

(a) *the stochastic kernel P_2 on $\mathbb{W}$ given $\mathbb{W}\times\mathbb{A}$ is weakly continuous, and the stochastic kernel Q_2 on $\mathbb{Y}$ given $\mathbb{A}\times\mathbb{W}$ is continuous in total variation;*

(b) *the stochastic kernel P_2 on $\mathbb{W}$ given $\mathbb{W}\times\mathbb{A}$ is continuous in total variation, and the observation kernel Q_2 on $\mathbb{Y}$ given $\mathbb{A}\times\mathbb{W}$ is continuous in a in total variation;*

implies semiuniform Feller continuity of the stochastic kernel P on $\mathbb{W}\times\mathbb{Y}$ given $\mathbb{W}\times\mathbb{A}$. Moreover, each of the conditions (a) *and* (b) *implies the validity of conclusions* (i)–(iii) *of Corollary* 6.10 *for the* POMDP$_2$.

Proof. See Appendix A for the proof. ∎

Regarding Corollary 6.11, weak continuity of the stochastic kernel $\hat{q}$ on $\mathbb{P}(\mathbb{W})\times\mathbb{A}$

for a POMDP$_2$ under condition (a) from Corollary 6.11 is stated in [19, Theorem 3.6], and another proof of this statement is provided in [28, Theorem 1]. Weak continuity of the stochastic kernel $\hat{q}$ on $\mathbb{P}(\mathbb{W}) \times \mathbb{A}$ for a POMDP$_2$ under condition (b) from Corollary 6.11 is an extension of [28, Theorem 2], where this weak continuity is proved under the assumption that the stochastic kernel P_2 on $\mathbb{W}$ given $\mathbb{W} \times \mathbb{A}$ is continuous in total variation and the observation kernel Q_2 does not depend on actions.

Different sufficient conditions for weak continuity of the kernel $\hat{q}$ for a POMDP$_2$ are formulated in monographs [24, 37]. In both cases these conditions are stronger than condition (a) from Corollary 6.11. In terms of the current paper, weak continuity of the stochastic kernel $\hat{q}$ on $\mathbb{P}(\mathbb{W})$ given $\mathbb{P}(\mathbb{W}) \times \mathbb{A}$ is stated in [24, p. 92] under condition (a) from Corollary 6.11 and under the assumption that the observation space $\mathbb{Y}$ is denumerable. The proof in [24, p. 93] is based on the existence of a transition kernel $H(z, a, y')$, which is weakly continuous in (z, a, y') and satisfies (6.1). However, [19, Example 4] shows that such kernel may not exist even for a POMDP$_2$ with finite sets $\mathbb{X}$, $\mathbb{Y}$ and continuous in a functions $P_2(x'|x, a)$ and $Q_2(y|a, x)$. A POMDP$_2$ is considered in [37, Chapter 2] under additional assumptions that the state space $\mathbb{X}$ is locally compact, observations y_t belong to a Euclidean space, and the observation kernel does not depend on actions and has a density, that is, $Q(dy|x) = r(x, y)dy$. Weak continuity of the kernel $\hat{q}$ is stated in [37, Corollary 1.5] under four assumptions, which taken together are stronger than condition (a) in Corollary 6.11.

Let us consider Platzman's model $(\mathbb{W} \times \mathbb{Y}, \mathbb{A}, P, c)$ with the cost function c that does not depend on observations y, that is, $c(w, y, a) = c(w, a)$. In this case the MDPCI $(\mathbb{P}(\mathbb{W}) \times \mathbb{Y}, \mathbb{A}, q, \bar{c})$ can be reduced to a smaller MDP $(\mathbb{P}(\mathbb{W}), \mathbb{A}, \hat{q}, \hat{c})$ with the state space $\mathbb{P}(\mathbb{W})$, action space $\mathbb{A}$, transition probability $\hat{q}$ defined in (6.1), and one-step cost function $\hat{c} : \mathbb{P}(\mathbb{W}) \times \mathbb{A} \to \overline{\mathbb{R}}$, defined for $z \in \mathbb{P}(\mathbb{W})$ and $a \in \mathbb{A}$ as

$$\hat{c}(z, a) := \int_{\mathbb{W}} c(w, a) z(dw). \tag{6.6}$$

The reduction of an MDPCI $(\mathbb{P}(\mathbb{W}) \times \mathbb{Y}, \mathbb{A}, q, \bar{c})$ to the belief-MDP $(\mathbb{P}(\mathbb{W}), \mathbb{A}, \hat{q}, \hat{c})$ holds in view of [11, Theorem 2] because in the MDPCI transition probabilities from states $(z_t, y_t) \in \mathbb{P}(\mathbb{W}) \times \mathbb{Y}$ to states $z_{t+1} \in \mathbb{P}(\mathbb{W})$ and costs $c(z_t, a_t)$ do not depend on y_t. If a Markov or stationary optimal policy is found for the belief-MDP $(\mathbb{P}(\mathbb{W}), \mathbb{A}, \hat{q}, \hat{c})$, it is possible, as described at the end of section 3, to construct an optimal policy for Platzman's models following the same procedures as constructing an optimal policy for an MDPII given a Markov or stationary optimal policy for the corresponding MDPCI.

COROLLARY 6.12. *Let us consider Platzman's model* $(\mathbb{W} \times \mathbb{Y}, \mathbb{A}, P, c)$ *with the one-step cost function* $c : \mathbb{W} \times \mathbb{A} \to \overline{\mathbb{R}}_+$. *If the stochastic kernel* P *on* $\mathbb{W} \times \mathbb{Y}$ *given* $\mathbb{W} \times \mathbb{A}$ *is semiuniform Feller, and the one-step cost function* c *is* $\mathbb{K}$*-inf-compact on* $\mathbb{W} \times \mathbb{A}$, *then the transition kernel* $\hat{q}$ *on* $\mathbb{P}(\mathbb{W})$ *given* $\mathbb{P}(\mathbb{W}) \times \mathbb{A}$ *is weakly continuous, the one-step cost function* $\hat{c}$ *is* $\mathbb{K}$*-inf-compact on* $\mathbb{P}(\mathbb{W}) \times \mathbb{A}$, *and all the conclusions of* [19, Theorem 2.1] *hold for the belief-MDP* $(\mathbb{P}(\mathbb{W}), \mathbb{A}, \hat{q}, \hat{c})$, *that is,*

(i) *optimality equations hold, and they define optimal policies;*
(ii) *value iterations converge to optimal values if zero terminal costs are chosen;*
(iii) *Markov optimal policies exist for finite-horizon problems;*
(iv) *stationary optimal policies exist for infinite-horizon problems.*

Moreover, all these conclusions hold for a POMDP$_1$ $(\mathbb{W}, \mathbb{Y}, \mathbb{A}, P_1, Q_1, c)$ *with the transition and observation kernels* P_1 *and* Q_1 *satisfying conditions* (a) *and* (b) *from Corollary* 6.10 *and for* a POMDP$_2$ $(\mathbb{W}, \mathbb{Y}, \mathbb{A}, P_2, Q_2, c)$ *with the transition and observation* *kernels* P_2 *and* Q_2 *satisfying either condition* (a) *or condition* (b) *from Corollary* 6.11.

Proof. Weak continuity of the stochastic kernel $\hat{q}$ on $\mathbb{P}(\mathbb{W})$ given $\mathbb{P}(\mathbb{W})\times\mathbb{A}$ is stated in Corollary 6.7. $\mathbb{K}$-inf-compactness of the function $\hat{c}$ on $\mathbb{P}(\mathbb{W})\times\mathbb{A}$ follows from [19, Theorem 3.3]. The remaining statements of the corollary follow from [19, Theorem 2.1]. The transition probability P for POMDP$_1$ $(\mathbb{W},\mathbb{Y},\mathbb{A},P_1,Q_1,c)$ defined in (6.2) is semiuniform Feller according to Corollary 6.10, and the transition probability P for POMDP$_2$ $(\mathbb{W},\mathbb{Y},\mathbb{A},P_2,Q_2,c)$ defined in (6.4) is semiuniform Feller due to Corollary 6.11. ∎

Appendix A. Proofs of Theorems 4.11, 5.3, and 6.1, and Corollaries 6.10 and 6.11. We use the following fact in the proofs of equalities (A.1) and (A.2) below: if $\{G^{(n)},G\}_{n=1,2,\dots}$ is a sequence of finite measures on a metric space $\mathcal{S}$ and $\{g^{(n)},g\}_{n=1,2,\dots}$ is a uniformly bounded sequence of Borel measurable functions on $\mathcal{S}$ such that

$$\lim_{n\to\infty}\sup_{B\in\mathcal{B}(\mathcal{S})}\left|\int_B g^{(n)}(s)G^{(n)}(ds)-\int_B g^{(n)}(s)G(ds)\right|=0,$$

then

$$\lim_{n\to\infty}\sup_{B\in\mathcal{B}(\mathcal{S})}\left|\int_B g^{(n)}(s)G^{(n)}(ds)-\int_B g(s)G(ds)\right|=0$$

holds if and only if

$$\lim_{n\to\infty}\sup_{B\in\mathcal{B}(\mathcal{S})}\left|\int_B g^{(n)}(s)G(ds)-\int_B g(s)G(ds)\right|=0.$$

Proof of Theorem 4.11. (a) $\Rightarrow$ (b). Since the stochastic kernel Ψ on $\mathbb{S}_1\times\mathbb{S}_2$ given $\mathbb{S}_3$ is semiuniform Feller, the marginal kernel $\Psi(\mathbb{S}_1,\cdot|\cdot)$ is continuous in total variation. Moreover, for each bounded continuous function f on $\mathbb{S}_1$, we have from (4.1) and (4.11) that

$$\begin{aligned}&\lim_{n\to\infty}\sup_{B\in\mathcal{B}(\mathbb{S}_2)}\left|\int_B\int_{\mathbb{S}_1} f(s_1)\Phi(ds_1|s_2,s_3^{(n)})\Psi(\mathbb{S}_1,ds_2|s_3)\right.\\ &\left.\quad-\int_B\int_{\mathbb{S}_1} f(s_1)\Phi(ds_1|s_2,s_3)\Psi(\mathbb{S}_1,ds_2|s_3)\right|=0\end{aligned}\tag{A.1}$$

because the family of Borel measurable functions $\{s_2\mapsto\int_{\mathbb{S}_1}f(s_1)\Phi(ds_1|s_2,s_3^{(n)}) : n=1,2,\dots\}$ is uniformly bounded on $\mathbb{S}_2$ by the same constant as f on $\mathbb{S}_1$. This is equivalent to $\int_{\mathbb{S}_1}f(s_1)\Phi(ds_1|\cdot,s_3^{(n)})\to\int_{\mathbb{S}_1}f(s_1)\Phi(ds_1|\cdot,s_3)$ in $L_1(\mathbb{S}_2,\mathcal{B}(\mathbb{S}_2),\nu)$ with $\nu(\cdot):=\Psi(\mathbb{S}_1,\cdot|s_3)$. Therefore,

$$\int_{\mathbb{S}_1} f(s_1)\Phi(ds_1|\cdot,s_3^{(n_k)})\to\int_{\mathbb{S}_1} f(s_1)\Phi(ds_1|\cdot,s_3)\quad\nu\text{-almost surely, as }k\to\infty,$$

for some sequence $\{n_k\}_{k=1,2,\dots}$ ($n_k\uparrow\infty$ as $k\to\infty$). We apply the diagonalization procedure to extract a subsequence $\{\tilde{n}_k\}_{k=1,2,\dots}$ ($\tilde{n}_k\uparrow\infty$ as $k\to\infty$) such that

$$\int_{\mathbb{S}_1} g(s_1)\Phi(ds_1|\cdot,s_3^{(\tilde{n}_k)})\to\int_{\mathbb{S}_1} g(s_1)\Phi(ds_1|\cdot,s_3)\quad\nu\text{-almost surely, as }k\to\infty,$$

for each $g\in\mathcal{G}$, where $\mathcal{G}$ is a countable uniformly bounded family of continuous functions on $\mathbb{S}_2$ that determines weak convergence of probability measures on $\mathbb{S}_2$

according to Parthasarathy [32, Theorem 6.6, p. 47]. Thus, $\Phi(\,\cdot\,, s_3^{(\tilde{n}_k)})$ converges weakly to $\Phi(\,\cdot\,, s_3)$ ν-almost surely, and Assumption 4.10 holds.

(b) $\Rightarrow$ (c). Let f be a bounded continuous function on $\mathbb{P}(\mathbb{S}_1)$. Since $\Psi(\mathbb{S}_1, \cdot|\cdot)$ is continuous in total variation, to prove that (4.1) holds for the stochastic kernel ϕ, it is sufficient to show that

(A.2)
$$\lim_{n\to\infty} \sup_{B\in\mathcal{B}(\mathbb{S}_2)} \left| \int_B f(\Phi(s_2, s_3^{(n)}))\Psi(\mathbb{S}_1, ds_2|s_3) - \int_B f(\Phi(s_2, s_3))\Psi(\mathbb{S}_1, ds_2|s_3) \right| = 0.$$

For the probability space $\Sigma := (\mathbb{S}_2, \mathcal{B}(\mathbb{S}_2), \mu)$ with $\mu(\,\cdot\,) := \Psi(\mathbb{S}_1, \cdot\,|s_3)$, the $\mathbb{P}(\mathbb{S}_1)$-valued random variables $\Phi(\,\cdot\,, s_3^{(n)}) \to^{\mu} \Phi(\,\cdot\,, s_3)$ as $n \to \infty$, according to Assumption 4.10 and (4.14), where $\nu^{(n)} \to^{\mu} \nu$ denotes the convergence in probability μ, that is, $\rho_{\mathbb{P}(\mathbb{S}_1)}(\nu^{(n)}, \nu) \to 0$ in probability μ. Then $f(\Phi(\,\cdot\,, s_3^{(n)})) \to^{\mu} f(\Phi(\,\cdot\,, s_3))$ because f is continuous on $\mathbb{P}(\mathbb{S}_1)$. In turn, since f is bounded on $\mathbb{P}(\mathbb{S}_1)$, this implies that $f(\Phi(\,\cdot\,, s_3^{(n)})) \to f(\Phi(\,\cdot\,, s_3))$ in $L_1(\Sigma)$, from which the desired relation (A.2) follows.

(c) $\Rightarrow$ (a). Let a sequence $\{s_3^{(n)}\}_{n=1,2,\dots} \subset \mathbb{S}_3$ converge to $s_3 \in \mathbb{S}_3$ as $n \to \infty$. Since the stochastic kernel ϕ on $\mathbb{P}(\mathbb{S}_1) \times \mathbb{S}_2$ given $\mathbb{S}_3$ is semiuniform Feller, for every nonnegative bounded lower semicontinuous function f on $\mathbb{P}(\mathbb{S}_1)$, according to Theorem 4.6(a,e),

$$\liminf_{n\to\infty} \inf_{B\in\mathcal{B}(\mathbb{S}_2)} \left(\int_{\mathbb{P}(\mathbb{S}_1)} f(\mu)\phi(d\mu, B|s_3^{(n)}) - \int_{\mathbb{P}(\mathbb{S}_1)} f(\mu)\phi(d\mu, B|s_3) \right) = 0. \tag{A.3}$$

For each $B \in \mathcal{B}(\mathbb{S}_2)$, formula (4.12) establishes the equality of two measures on $(\mathbb{P}(\mathbb{S}_1), \mathcal{B}(\mathbb{P}(\mathbb{S}_1)))$. Therefore, for every Borel measurable nonnegative functions f on $\mathbb{P}(\mathbb{S}_1)$,

$$\int_B f(\Phi(s_2, \tilde{s}_3))\Psi(\mathbb{S}_1, ds_2|\tilde{s}_3) = \int_{\mathbb{P}(\mathbb{S}_1)} f(\mu)\phi(d\mu, B|\tilde{s}_3), \qquad \tilde{s}_3 \in \mathbb{S}_3. \tag{A.4}$$

Let us fix an arbitrary open set $\mathcal{O} \subset \mathbb{S}_1$ and consider nonnegative bounded l.s.c. function $f(\mu) := \mu(\mathcal{O})$, $\mu \in \mathbb{P}(\mathbb{S}_1)$. Then

$$\begin{aligned}
&\liminf_{n\to\infty} \inf_{B\in\mathcal{B}(\mathbb{S}_2)} \left(\int_B \Phi(\mathcal{O}|s_2, s_3^{(n)})\Psi(\mathbb{S}_1, ds_2|s_3^{(n)}) - \int_B \Phi(\mathcal{O}|s_2, s_3)\Psi(\mathbb{S}_1, ds_2|s_3) \right) \\
&\quad = \liminf_{n\to\infty} \inf_{B\in\mathcal{B}(\mathbb{S}_2)} \left(\int_B f(\Phi(s_2, s_3^{(n)}))\Psi(\mathbb{S}_1, ds_2|s_3^{(n)}) \right. \\
&\quad \left. - \int_B f(\Phi(s_2, s_3))\Psi(\mathbb{S}_1, ds_2|s_3) \right) = 0,
\end{aligned}$$

where the first equality follows from the definition of f, and the second equality follows from (A.4) and from (A.3). Thus, the stochastic kernel Ψ on $\mathbb{S}_1 \times \mathbb{S}_2$ given $\mathbb{S}_3$ is WTV-continuous, and therefore it is semiuniform Feller. $\square$

Remark A.1. Theorem 4.11 can be proved in multiple ways using equivalent characterizations of semiuniform Feller kernels. The original proofs [22, Proof of Theorem 5.10, pp. 16–20] were based on some of these characterizations, while the current proofs of (a) $\Rightarrow$ (b) $\Rightarrow$ (c) were suggested by a referee.

The following lemma, Lemma A.2, is useful for establishing continuity properties of the value functions $v_{n,\alpha}(x)$ and $v_\alpha(x)$ in $x \in \mathbb{X}$ stated in Theorem 5.3.

LEMMA A.2. *Let the MDP* $(\mathbb{X},\mathbb{A},q,c)$ *satisfy the assumptions of Theorem* 5.3, *and let* $\alpha \geq 0$. *Then the function* $u^*(x) := \inf_{a\in\mathbb{A}} \eta_u^\alpha(x,a)$, $x \in \mathbb{X}$, *where the function* η_u^α *is defined in* (5.2), *belongs to* $LW(\mathbb{X})$, *and there exists a stationary policy* $f : \mathbb{X}\to\mathbb{A}$ *such that* $u^*(x) := \eta_u^\alpha(x,f(x))$, $x \in \mathbb{X}$. *Moreover, the sets* $A_*(x) = \{a \in \mathbb{A} : u^*(x) = \eta_u^\alpha(x,a)\}$, $x \in \mathbb{X}$, *which are nonempty, satisfy the following properties:* (a) *the graph* $\mathrm{Gr}(A_*) = \{(x,a) : x \in \mathbb{X},\ a \in A_*(x)\}$ *is a Borel subset of* $\mathbb{X}\times\mathbb{A}$; (b) *if* $u^*(x) = +\infty$, *then* $A_*(x) = \mathbb{A}$, *and if* $u^*(x) < +\infty$, *then* $A_*(x)$ *is compact.*

Proof. The function $(x,a) \mapsto \eta_u^\alpha(x,a)$ is *nonnegative* because c, u, and α are nonnegative. Therefore, since u is a Borel measurable function, and q is a stochastic kernel, [5, Proposition 7.29] implies that the function $(x,a) \mapsto \int_{\mathbb{X}} u(\tilde{x})q(d\tilde{x}|x,a)$ is Borel measurable on $\mathbb{X}\times\mathbb{A}$, which implies that the function $(x,a) \mapsto \eta_u^\alpha(x,a)$ is *Borel measurable* on $\mathbb{X}\times\mathbb{A}$ because c is Borel measurable.

Let us prove that the function $(w,a) \mapsto \int_{\mathbb{X}} u(\tilde{x})q(d\tilde{x}|w,y,a)$ is l.s.c. on $\mathbb{X}_W\times\mathbb{A}$ for each $y \in \mathbb{X}_Y$. On the contrary, if this function is not l.s.c., then there exist a sequence $\{(w^{(n)},a^{(n)})\}_{n=1,2,\dots} \subset \mathbb{X}_W\times\mathbb{A}$ converging to some $(w,a) \in \mathbb{X}_W\times\mathbb{A}$ and a constant λ such that for each $n = 1,2,\dots$

$$\int_{\mathbb{X}_W\times\mathbb{X}_Y} u(\tilde{w},\tilde{y})q(d\tilde{w}\times d\tilde{y}|w^{(n)},y,a^{(n)}) \leq \lambda < \int_{\mathbb{X}} u(\tilde{x})q(d\tilde{x}|w,y,a). \tag{A.5}$$

According to Theorem 4.11(a,b) applied to $\Psi := q$, $\mathbb{S}_1 := \mathbb{X}_W$, $\mathbb{S}_2 := \mathbb{X}_Y$, and $\mathbb{S}_3 := \mathbb{X}_W\times\{y\}\times\mathbb{A}$, there exists a stochastic kernel Φ on $\mathbb{X}_W$ given $\mathbb{X}_Y\times\mathbb{X}_W\times\{y\}\times\mathbb{A}$ such that (4.11) and Assumption 4.10 hold. In particular, (A.5) implies that for each $n = 1,2,\dots$

$$\int_{\mathbb{X}_Y}\left[\int_{\mathbb{X}_W} u(\tilde{w},\tilde{y})\Phi(d\tilde{w}|\tilde{y},w^{(n)},y,a^{(n)})\right] q(\mathbb{X}_W,d\tilde{y}|w^{(n)},y,a^{(n)}) \leq \lambda,$$

and there exist a subsequence $\{(w^{(n_k)},a^{(n_k)})\}_{k=1,2,\dots} \subset \{(w^{(n)},a^{(n)})\}_{n=1,2,\dots}$ and a Borel set $Y \in \mathcal{B}(\mathbb{X}_Y)$ such that $q(\mathbb{X}_W\times Y|w,y,a) = 1$ and $\Phi(\tilde{y},w^{(n)},y,a^{(n)})$ converges weakly to $\Phi(\tilde{y},w,y,a)$ in $\mathbb{P}(\mathbb{X}_W)$ as $k\to\infty$ for all $\tilde{y} \in Y$. Therefore, since the function $\tilde{w} \mapsto u(\tilde{w},\tilde{y})$ is nonnegative and l.s.c. for each $\tilde{y} \in Y$, Fatou's lemma for weakly converging probabilities [17, Theorem 1.1] implies that for each $\tilde{y} \in Y$,

$$\int_{\mathbb{X}_W} u(\tilde{w},\tilde{y})\Phi(d\tilde{w}|\tilde{y},w,y,a) \leq \liminf_{k\to\infty}\int_{\mathbb{X}_W} u(\tilde{w},\tilde{y})\Phi(d\tilde{w}|\tilde{y},w^{(n_k)},y,a^{(n_k)}). \tag{A.6}$$

For a fixed $N = 1,2,\dots$, we set $\varphi_k^N(\tilde{y}) := \min\{\int_{\mathbb{X}_W} u(\tilde{w},\tilde{y})\Phi(d\tilde{w}|\tilde{y},w^{(n_k)},y,a^{(n_k)}),N\}$ and $\varphi^N(\tilde{y}) := \min\{\int_{\mathbb{X}_W} u(\tilde{w},\tilde{y})\Phi(d\tilde{w}|\tilde{y},w,y,a),N\}$, where $\tilde{y} \in Y$, $k = 1,2,\dots$. Note that $\varphi^N(\tilde{y}) \leq \liminf_{k\to\infty}\varphi_k^N(\tilde{y})$, $\tilde{y} \in Y$, in view of (A.6). Therefore, uniform Fatou's lemma [20, Corollary 2.3] implies that for each $N = 1,2,\dots$

$$\begin{aligned}\int_{\mathbb{X}_Y}\varphi^N(\tilde{y})q(\mathbb{X}_W,d\tilde{y}|w,y,a) &\leq \liminf_{k\to\infty}\int_{\mathbb{X}_Y}\varphi_k^N(\tilde{y})q(\mathbb{X}_W,d\tilde{y}|w^{(n_k)},y,a^{(n_k)})\\ &\leq \liminf_{k\to\infty}\int_{\mathbb{X}_Y}\left[\int_{\mathbb{X}_W} u(\tilde{w},\tilde{y})\Phi(d\tilde{w}|\tilde{y},w^{(n_k)},y,a^{(n_k)})\right] q(\mathbb{X}_W,d\tilde{y}|w^{(n_k)},y,a^{(n_k)}) \leq \lambda.\end{aligned}$$

Thus, the monotone convergence theorem implies

$$\int_{\mathbb{X}} u(\tilde{x})q(d\tilde{x}|w,y,a) = \lim_{N\to\infty}\int_{\mathbb{X}_Y}\varphi^N(\tilde{y})q(\mathbb{X}_W,d\tilde{y}|w,y,a) \leq \lambda.$$

This is a contradiction with (A.5). Therefore, the function $(w,a)\mapsto \int_{\mathbb{X}} u(\tilde{x})q(d\tilde{x}|w,y,a)$ is l.s.c. on $\mathbb{X}_W\times\mathbb{A}$ for each $y\in\mathbb{X}_Y$.

For an arbitrary fixed $y\in\mathbb{X}_Y$ the function $(w,a)\mapsto \eta_u^\alpha(w,y,a)$ is $\mathbb{K}$-*inf-compact on* $\mathbb{X}_W\times\mathbb{A}$ as a sum of a $\mathbb{K}$-inf-compact function $(w,a)\mapsto c(w,y,a)$ and a nonnegative l.s.c. function $(w,a)\mapsto \alpha\int_{\mathbb{X}} u(\tilde{x})q(d\tilde{x}|w,y,a)$ on $\mathbb{X}_W\times\mathbb{A}$. Moreover, Berge's theorem for noncompact image sets [16, Theorem 1.2] implies that for each $(y,a)\in\mathbb{X}_Y\times\mathbb{A}$ the function $w\mapsto u^*(w,y):=\inf_{a\in\mathbb{A}}\eta_u^\alpha(w,y;a)$ is l.s.c. on $\mathbb{X}_W$. The Borel measurability of the function u^* on $\mathbb{X}$ and the existence of a stationary policy $f:\mathbb{X}\to\mathbb{A}$ such that $u^*(x):=\eta_u^\alpha(x,f(x))$, $x\in\mathbb{X}$, follow from [13, Theorem 2.2 and Corollary 2.3(i)] because the function $(x,a)\mapsto\eta_u^\alpha(x,a)$ is Borel measurable on $\mathbb{X}\times\mathbb{A}$ and it is inf-compact in a on $\mathbb{A}$. Property (a) for nonempty sets $\{A_*(x)\}_{x\in\mathbb{X}}$ follows from Borel measurability of $(x,a)\mapsto\eta_u^\alpha(x,a)$ on $\mathbb{X}\times\mathbb{A}$ and $x\mapsto u^*(x)$ on $\mathbb{X}$. Property (b) for $\{A_*(x)\}_{x\in\mathbb{X}}$ follows from inf-compactness of $a\mapsto\eta_u^\alpha(x,a)$ on $\mathbb{A}$ for each $x\in\mathbb{X}$. ∎

Proof of Theorem 5.3. According to [5, Proposition 8.2], the functions $v_{t,\alpha}(x)$, $t=0,1,\dots$, recursively satisfy the optimality equations with $v_{0,\alpha}(x)=0$ and $v_{t+1,\alpha}(x)=\inf_{a\in A(x)}\eta^\alpha_{v_{t,\alpha}}(x,a)$ for all $x\in\mathbb{X}$. So, Lemma A.2 sequentially applied to the functions $v_{0,\alpha}(x)$, $v_{1,\alpha}(x),\dots$ implies statement (i) for them. According to [5, Proposition 9.17], $v_{t,\alpha}(x)\uparrow v_\alpha(x)$ as $t\to+\infty$ for each $x\in\mathbb{X}$. Therefore, $v_\alpha\in LW(\mathbb{X})$. Thus, statement (i) is proved. In addition, [5, Lemma 8.7] implies that a Markov policy defined at the first T steps by the mappings $\phi_0^\alpha,\dots,\phi_{T-1}^\alpha$, which satisfy for all $t=1,\dots,T$ the equations $v_{t,\alpha}(x)=\eta^\alpha_{v_{t-1,\alpha}}(x,\phi^\alpha_{T-t}(x))$ for each $x\in\mathbb{X}$, is optimal for the horizon T. According to [5, Propositions 9.8 and 9.12], v_α satisfies the discounted cost optimality equation $v_\alpha(x)=\inf_{a\in A(x)}\eta^\alpha_{v_\alpha}(x,a)$ for each $x\in\mathbb{X}$; and a stationary policy ϕ_α is discount-optimal if and only if $v_\alpha(x)=\eta^\alpha_{v_\alpha}(x,\phi_\alpha(x))$ for each $x\in\mathbb{X}$. Statements (ii)–(v) follow from these facts and Lemma A.2. ∎

Proof of Theorem 6.1. The equivalence of statements (a) and (b) follows directly from Theorem 4.8 applied to $\mathbb{S}_1:=\mathbb{W}$, $\mathbb{S}_2:=\mathbb{Y}$, $\mathbb{S}_3:=\mathbb{W}\times\mathbb{A}$, and $\Psi:=P(\,\cdot\,|\,\cdot\,,y,\,\cdot\,)$. According to (3.1), Theorem 4.9 applied to $\mathbb{S}_1:=\mathbb{W}$, $\mathbb{S}_2:=\mathbb{Y}$, $\mathbb{S}_3=\mathbb{W}$, $\mathbb{S}_4:=\mathbb{A}$, and $\Xi:=P(\,\cdot\,|\,\cdot\,,y,\,\cdot\,)$ implies that the stochastic kernel $P(\,\cdot\,|\,\cdot\,,y,\,\cdot\,)$ on $\mathbb{W}\times\mathbb{Y}$ given $\mathbb{W}\times\mathbb{A}$ is semiuniform Feller if and only if the stochastic kernel $R(\,\cdot\,|\,\cdot\,,y,\,\cdot\,)$ on $\mathbb{W}\times\mathbb{Y}$ given $\mathbb{P}(\mathbb{W})\times\mathbb{A}$ is semiuniform Feller. Therefore, statement (b) holds if and only if the stochastic kernel $R(\,\cdot\,|\,\cdot\,,y,\,\cdot\,)$ on $\mathbb{W}\times\mathbb{Y}$ given $\mathbb{P}(\mathbb{W})\times\mathbb{A}$ is semiuniform Feller, that is, statement (c) holds. Thus, the equivalence of statements (c)–(e) follows directly from Theorem 4.11 applied to $\mathbb{S}_1:=\mathbb{W}$, $\mathbb{S}_2:=\mathbb{Y}$, $\mathbb{S}_3:=\mathbb{P}(\mathbb{W})\times\mathbb{A}$, $\Psi:=R(\,\cdot\,|\,\cdot\,,y,\,\cdot\,)$, $\Phi:=H(\,\cdot\,|\,\cdot\,,y,\,\cdot\,,\,\cdot\,)$, and $\phi:=q(\,\cdot\,|\,\cdot\,,y,\,\cdot\,)$.

Moreover, let the nonnegative function c be $\mathbb{MK}(\mathbb{W}\times\mathbb{A},\mathbb{Y})$-inf-compact, and for each $y\in\mathbb{Y}$ let one of the equivalent conditions (a)–(d) hold. Then, in view of (3.3) and Theorem 6.3, $\bar{c}$ is nonnegative and $\mathbb{MK}(\mathbb{P}(\mathbb{W})\times\mathbb{A},\mathbb{Y})$-inf-compact. Thus, the assumptions and conclusions of Theorem 5.3 hold for the MDPCI $(\mathbb{P}(\mathbb{W})\times\mathbb{Y},\mathbb{A},q,\bar{c})$. ∎

Proof of Theorem 6.2. The equivalence of statements (a) and (b) follows directly from Theorem 4.8 applied to $\mathbb{S}_1:=\mathbb{W}$, $\mathbb{S}_2:=\mathbb{Y}$, $\mathbb{S}_3:=\mathbb{W}\times\mathbb{Y}\times\mathbb{A}$, and $\Psi:=P$. According to (3.1), Theorem 4.9 applied to $\mathbb{S}_1:=\mathbb{W}$, $\mathbb{S}_2:=\mathbb{Y}$, $\mathbb{S}_3=\mathbb{W}$, $\mathbb{S}_4:=\mathbb{Y}\times\mathbb{A}$, and $\Xi:=P$ implies that the stochastic kernel P on $\mathbb{W}\times\mathbb{Y}$ given $\mathbb{W}\times\mathbb{Y}\times\mathbb{A}$ is semiuniform Feller if and only if the stochastic kernel R on $\mathbb{W}\times\mathbb{Y}$ given $\mathbb{P}(\mathbb{W})\times\mathbb{Y}\times\mathbb{A}$ is semiuniform Feller. Therefore, statement (b) holds if and only if the stochastic kernel R on $\mathbb{W}\times\mathbb{Y}$ given $\mathbb{P}(\mathbb{W})\times\mathbb{Y}\times\mathbb{A}$ is semiuniform Feller, that is, statement (c) holds. Thus, the equivalence of statements (c)–(e) follows directly from Theorem 4.11 applied to $\mathbb{S}_1:=\mathbb{W}$, $\mathbb{S}_2:=\mathbb{Y}$, $\mathbb{S}_3:=\mathbb{P}(\mathbb{W})\times\mathbb{Y}\times\mathbb{A}$, $\Psi:=R$, $\Phi:=H$, and $\phi:=q$.

Moreover, let the nonnegative function c be $\mathbb{K}$-inf-compact, and let one of the equivalent conditions (a)–(d) hold. Then, in view of (3.3) and [19, Theorem 3.3] on preservation of $\mathbb{K}$-inf-compactness, $\bar{c}$ is nonnegative and $\mathbb{K}$-inf-compact. Thus, according to Remark 5.4, the assumptions and conclusions of Theorem 5.3 hold for the MDPCI $(\mathbb{P}(\mathbb{W}) \times \mathbb{Y}, \mathbb{A}, q, \bar{c})$, and the functions $v_{t,\alpha}$, $t = 0, 1, \ldots,$ and v_α are l.s.c. ◻

Proof of Corollary 6.10. Let us prove that semiuniform Feller continuity of the stochastic kernel P on $\mathbb{W} \times \mathbb{Y}$ given $\mathbb{W} \times \mathbb{A}$ implies conditions (a) and (b). Indeed, Definition 4.1 implies weak continuity of the stochastic kernel P_1 on $\mathbb{W}$ given $\mathbb{W}\times\mathbb{A}$ and continuity in the total variation of the stochastic kernel Q_1 on $\mathbb{Y}$ given $\mathbb{W}\times\mathbb{A}$ because $P_1(\,\cdot\,|\,\cdot\,) = P(\,\cdot\,, \mathbb{Y}|\,\cdot\,)$ is weakly continuous and $Q_1(\,\cdot\,|\,\cdot\,) = P(\mathbb{W}, \,\cdot\,|\,\cdot\,)$ is continuous in total variation. Conversely, let us prove that conditions (a) and (b) imply semiuniform Feller continuity of the stochastic kernel P on $\mathbb{W}\times\mathbb{Y}$ given $\mathbb{W}\times\mathbb{A}$. Indeed, P on $\mathbb{W}\times\mathbb{Y}$ given $\mathbb{W}\times\mathbb{A}$ is WTV-continuous since

$$
\begin{aligned}
\liminf_{(w'a')\to(w,a)} \inf_{C\in\mathcal{B}(\mathbb{Y})} & (Q_1(C|w',a')P_1(\mathcal{O}|w',a') - Q_1(C|w,a)P_1(\mathcal{O}|w,a)) \\
&\geq \liminf_{(w'a')\to(w,a)} (P_1(\mathcal{O}|w',a') - P_1(\mathcal{O}|w,a))^- \\
&\quad - \lim_{(w'a')\to(w,a)} \sup_{C\in\mathcal{B}(\mathbb{Y})} |Q_1(C|w',a') - Q_1(C|w,a)| = 0
\end{aligned}
$$

for each $\mathcal{O} \in \tau(\mathbb{W})$, where $a^- := \min\{a, 0\}$ for each $a \in \mathbb{R}$, the equality follows from weak continuity of P_1 on $\mathbb{W}$ given $\mathbb{W}\times\mathbb{A}$ and continuity in the total variation of Q_1 on $\mathbb{Y}$ given $\mathbb{A}\times\mathbb{W}$. Therefore, according to Theorem 4.6(a,b), conditions (a) and (b) from Corollary 6.10 taken together are equivalent to semiuniform Feller continuity of the stochastic kernel P on $\mathbb{W}\times\mathbb{Y}$ given $\mathbb{W}\times\mathbb{A}$. Thus, Theorem 6.1 implies all statements of Corollary 6.10. ◻

Proof of Corollary 6.11. For each $B \in \mathcal{B}(\mathbb{W})$ consider the family of functions

$$
\mathcal{G}(B) := \Big\{(w,a) \mapsto \int_B Q_2(C|a,w')P_2(dw'|w,a) \,:\, C \in \mathcal{B}(\mathbb{Y})\Big\}.
$$

Let condition (a) hold. Fix an arbitrary open set $\mathcal{O} \in \tau(\mathbb{W})$. Feinberg, Kasyanov, and Zgurovsky [21, Theorem 1], applied to the lower semiequicontinuous and uniformly bounded family of functions $\{(w',a) \mapsto \mathbf{I}\{w' \in \mathcal{O}\}Q_2(C|a,w') \,:\, C \in \mathcal{B}(\mathbb{Y})\}$ and weakly continuous stochastic kernel $P_2(dw'|w,a)$ on $\mathbb{W}$ given $\mathbb{W}\times\mathbb{A}$, implies that the family of functions $\mathcal{G}(\mathcal{O})$ is lower semiequicontinuous at all the points $(w,a) \in \mathbb{W}\times\mathbb{A}$, that is, the stochastic kernel P on $\mathbb{W}\times\mathbb{Y}$ given $\mathbb{W}\times\mathbb{Y}\times\mathbb{A}$ defined in (6.4) is WTV-continuous. Therefore, Theorem 4.6(a,b) applied to the stochastic kernel P on $\mathbb{W}\times\mathbb{Y}$ given $\mathbb{W}\times\mathbb{Y}\times\mathbb{A}$ implies that this kernel is semiuniform Feller. Thus, assumption (a) of Theorem 6.1 holds, and this conclusion and Theorem 6.1 imply all statements of Corollary 6.11 under condition (a).

Now let condition (b) hold. Let us prove that for each $B \in \mathcal{B}(\mathbb{W})$ the family of functions $\mathcal{G}(B)$ is equicontinuous at all the points $(w,a) \in \mathbb{W}\times\mathbb{A}$, which implies condition (a) of Theorem 6.1. Indeed, for $n = 1, 2, \ldots,$

$$
\begin{aligned}
\text{(A.7)} \quad \sup_{C\in\mathcal{B}(\mathbb{Y})} \Big| \int_B Q_2(C|a^{(n)},w')P_2(dw'|w^{(n)},a^{(n)}) - \int_B Q_2(C|a,w')P_2(dw'|w,a)\Big| \\
\leq I_1^{(n)} + I_2^{(n)},
\end{aligned}
$$

where $(w^{(n)}, a^{(n)}) \to (w, a)$ as $n \to \infty$,

$$I_1^{(n)} := \sup_{C \in \mathcal{B}(\mathbb{Y})} \left| \int_B Q_2(C|a^{(n)}, w')P_2(dw'|w^{(n)}, a^{(n)}) - \int_B Q_2(C|a^{(n)}, w')P_2(dw'|w, a) \right|$$

$$I_2^{(n)} := \sup_{C \in \mathcal{B}(\mathbb{Y})} \int_B |Q_2(C|a^{(n)}, w') - Q_2(C|a, w')|P_2(dw'|w, a).$$

Let $C^{(n)} \in \mathcal{B}(\mathbb{Y})$, $n = 1, 2, \dots$, be chosen to satisfy the inequality

$$I_2^{(n)} \le \int_B |Q_2(C^{(n)}|a^{(n)}, w') - Q_2(C^{(n)}|a, w')|P_2(dw'|w, a) + \frac{1}{n}. \tag{A.8}$$

Note that $I_1^{(n)} \to 0$ as $n \to \infty$ because the family of measurable functions $\{w' \mapsto Q_2(C \,|\, a^{(n)}, w') \,:\, n = 1, 2, \dots \text{and } C \in \mathcal{B}(\mathbb{Y})\}$ is uniformly bounded by 1, and the stochastic kernel P_2 on $\mathbb{W}$ given $\mathbb{W} \times \mathbb{A}$ is continuous in total variation. Moreover, the convergence $I_2^{(n)} \to 0$ as $n \to \infty$ follows from (A.8) and Lebesgue's dominated convergence theorem because the family of functions $\{w' \mapsto |Q_2(C^{(n)}|a^{(n)}, w') - Q_2(C^{(n)}|a, w')| \,:\, n = 1, 2, \dots\}$ is uniformly bounded by 1 and pointwise convergent to 0, according to (2.1). Therefore, the family of functions $\mathcal{G}(B)$ is equicontinuous at all the points $(w, a) \in \mathbb{W} \times \mathbb{A}$. Thus, assumption (a) of Theorem 6.1 holds, and this conclusion and Theorem 6.1 imply all statements of Corollary 6.11 under condition (b). ∎

Acknowledgments. We thank Janey (Huizhen) Yu for valuable remarks. We thank the referees for insightful remarks. In particular, one of the referees suggested a short proof of weak continuity of semiuniform Feller kernels, observed the equivalence of WTV-continuity and semiuniform Feller continuity, suggested strengthening Theorems 5.3 and 6.1 to their current formulations, proposed the provided proof of Theorem 4.11, and made other valuable comments.

REFERENCES

[1] M. Aoki, *Optimal control of partially observable Markovian systems*, J. Franklin Inst., 280 (1965), pp. 367–386.
[2] K. J. Åström, *Optimal control of Markov processes with incomplete state information*, J. Math. Anal. Appl., 10 (1965), pp. 174–205.
[3] E. J. Balder, *On compactness of the space of policies in stochastic dynamic programming*, Stoch. Proc. Appl., 32 (1989), pp. 141–150.
[4] N. Bäuerle and U. Rieder, *Markov Decision Processes with Applications to Finance*, Springer-Verlag, Berlin, 2011.
[5] D. P. Bertsekas and S. E. Shreve, *Stochastic Optimal Control: The Discrete-Time Case*, Academic Press, New York, 1978.
[6] P. Billingsley, *Convergence of Probability Measures*, John Wiley, New York, 1968.
[7] V. I. Bogachev, *Measure Theory, Volume* II, Springer-Verlag, Berlin, 2007.
[8] R. M. Dudley, *Real Analysis and Probability*, Cambridge University Press, Cambridge, 2002.
[9] E. B. Dynkin, *Controlled random sequences*, Theory Probab. Appl., 10 (1965), pp. 1–14.
[10] E. B. Dynkin and A. A. Yushkevich, *Controlled Markov Processes*, Springer-Verlag, New York, 1979.
[11] E. A. Feinberg, *On essential information in sequential decision processes*, Math. Meth. Oper. Res., 62 (2005), pp. 399–410.
[12] E. A. Feinberg, *Optimality conditions for inventory control*, in Tutorials in Operations Research, Optimization Challenges in Complex, Networked, and Risky Systems, A. Gupta and A. Capponi, eds., INFORMS, Cantonsville, MD, 2016, pp. 14–44.
[13] E. A. Feinberg and P. O. Kasyanov, *MDPs with setwise continuous transition probabilities*, Oper. Res. Lett., 49 (2021), pp. 734–740.

[14] E. A. Feinberg, P. O. Kasyanov, and Y. Liang, *Fatou's lemma in its classical form and Lebesgue's convergence theorems for varying measures with applications to Markov decision processes*, Theory Probab. Appl., 65 (2020), pp. 270–291.
[15] E. A. Feinberg, P. O. Kasyanov, and N. V. Zadoianchuk, *Average-cost Markov decision processes with weakly continuous transition probabilities*, Math. Oper. Res., 37 (2012), pp. 591–607.
[16] E. A. Feinberg, P. O. Kasyanov, and N. V. Zadoianchuk, *Berge's theorem for noncompact image sets*, J. Math. Anal. Appl., 397 (2013), pp. 255–259.
[17] E. A. Feinberg, P. O. Kasyanov, and N. V. Zadoianchuk, *Fatou's lemma for weakly converging probabilities*, Theory Probab. Appl., 58 (2014), pp. 683–689.
[18] E. A. Feinberg, P. O. Kasyanov, and M. Z. Zgurovsky, *Convergence of probability measures and Markov decision models with incomplete information*, Proc. Steklov Inst. Math., 287 (2014), pp. 96–117.
[19] E. A. Feinberg, P. O. Kasyanov, and M. Z. Zgurovsky, *Partially observable total-cost Markov decision processes with weakly continuous transition probabilities*, Math. Oper. Res., 41 (2016), pp. 656–681.
[20] E. A. Feinberg, P. O. Kasyanov, and M. Z. Zgurovsky, *Uniform Fatou's lemma*, J. Math. Anal. Appl., 444 (2016), pp. 550–567.
[21] E. A. Feinberg, P. O. Kasyanov, and M. Z. Zgurovsky, *Semi-uniform Feller Stochastic Kernels*, preprint, https://arxiv.org/abs/2107.02207, 2021.
[22] E. A. Feinberg, P. O. Kasyanov, and M. Z. Zgurovsky, *Markov Decision Processes with Incomplete Information and Semi-uniform Feller Transition Probabilities*, preprint, https://arxiv.org/abs/2108.09232v1, 2021.
[23] E. A. Feinberg, P. O. Kasyanov, and M. Z. Zgurovsky, *A class of solvable Markov decision models with incomplete information*, in Proceedings of the 60th IEEE Conference on Decision and Control (CDC), 2021, pp. 1615–1620.
[24] O. Hernández-Lerma, *Adaptive Markov Control Processes*, Springer-Verlag, New York, 1989.
[25] O. Hernández-Lerma, *Average optimality in dynamic programming on Borel spaces—unbounded costs and controls*, Systems and Control Lett., 17 (1991), pp. 237–242.
[26] O. Hernández-Lerma and J. B. Lasserre, *Discrete-Time Markov Control Processes: Basic Optimality Criteria*, Springer, New York, 1996.
[27] K. Hinderer and K.-H. Waldmann, *The critical discount factor for finite Markovian decision processes with an absorbing set*, Math. Meth. Oper. Res., 57 (2003), pp. 1–19.
[28] A. D. Kara, N. Saldi, and S. Yüksel, *Weak Feller property of non-linear filters*, Systems Control Lett., 134 (2019), 104512.
[29] G. E. Monahan, *State of the art–a survey of partially observable Markov decision processes: Theory, models, and algorithms*, Management Sci. 28 (1982), pp. 1–16.
[30] C. H. Papadimitriou and J. N. Tsitsiklis, *The complexity of Markov decision processes*, Math. Oper. Res., 12 (1987), pp. 441–450.
[31] G. C. Papanicolaou, *Asymptotic analysis of stochastic equations*, in Studies in Probability Theory, M. Rosenblatt, ed., MAA Stud. Math. 18, Math. Assoc. America, Washington, DC, 1978, pp. 111–179.
[32] K. R. Parthasarathy, *Probability Measures on Metric Spaces*, Academic Press, New York, 1967.
[33] L. K. Platzman, *Optimal infinite-horizon undiscounted control of finite probabilistic systems*, SIAM J. Control Optim., 18 (1980), pp. 362–380, https://doi.org/10.1137/0318028.
[34] D. Rhenius, *Incomplete information in Markovian decision models*, Ann. Statist., 2 (1974), pp. 1327–1334.
[35] U. Rieder, *Bayesian dynamic programming*, Adv. Appl. Probab., 7 (1975), pp. 330–348.
[36] W. Rudin, *Principles of Mathematical Analysis*, 2nd ed., McGraw-Hill, New York, 1964.
[37] W. J. Runggaldier and L. Stettner, *Approximations of Discrete Time Partially Observed Control Problems*, Applied Mathematics Monographs CNR, Giardini Editori, Pisa, 1994.
[38] M. Schäl, *On dynamic programming: Compactness of the space of policies*, Stochastic Process. Appl., 3 (1975), pp. 345–364.
[39] M. Schäl, *Conditions for optimality in dynamic programming and for the limit of n-stage optimal policies to be optimal*, Z. Wahrsch. verw. Gebiete, 32 (1975), pp. 179–196.
[40] M. Schäl, *Average optimality in dynamic programming with general state space*, Math. Oper. Res., 18 (1993), pp. 163–172.
[41] A. N. Shiryaev, *Some new results in the theory of controlled random processes*, in Transactions of the Fourth Prague Conference on Information Theory, Statistical Decision Functions, Random Processes, (Prague, 1965), 1967, pp. 131–201 (in Russian); Engl. transl. in Select. Transl. Math. Statist. Probab., 8 (1969), pp. 49–130.

[42] A. N. SHIRYAEV, *Probability*, 2nd ed., Springer-Verlag, New York, 1996.
[43] R. D. SMALLWOOD AND E. J. SONDIK, *The optimal control of partially observable Markov processes over a finite horizon*, Oper. Res., 21 (1973), pp. 1071–1088.
[44] E. J. SONDIK, *The optimal control of partially observable Markov processes over the infinite horizon: Discounted costs*, Oper. Res., 26 (1978), pp. 282–304.
[45] H. YU, *On convergence of value iteration for a class of total cost Markov decision processses*, SIAM J. Control Optim., 53 (2015), pp. 1982–2016, https://doi.org/10.1137/141000294.
[46] A. A. YUSHKEVICH, *Reduction of a controlled Markov model with incomplete data to a problem with complete information in the case of Borel state and control spaces*, Theory Probab. Appl., 21 (1976), pp. 153–158.