

A Class of Solvable Markov Decision Models with Incomplete Information

September 30, 2021

Eugene A. Feinberg¹, Pavlo O. Kasyanov², and Michael Z. Zgurovsky³

Abstract

This paper investigates natural conditions for the existence of optimal policies for a Markov decision process with incomplete information (MDPII) and with expected total costs. The MDPII is the classic model of a controlled stochastic process with incomplete state observations which is more general than Partially Observable Markov Decision Processes (POMDPs). For MDPIIs we introduce the notion of a semi-uniform Feller transition probability, which is stronger than the notion of a weakly continuous transition probability. We show that an MDPII has a semi-uniform Feller transition probability if and only if the corresponding belief MDP also has a semi-uniform Feller transition probability. This fact has several corollaries. In particular, it provides new and implies all known sufficient conditions for the existence of optimal policies for POMDPs with expected total costs.

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 lucky exceptions, and Kalman's filtering is among them, problems with incomplete information are known to be difficult [23]. The general approach to solving such problems was identified long time ago in [1, 2, 7, 30], 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 an optimal policy for the original problem from this policy. This approach was introduced in [1, 2, 7, 30] for problems with finite state, observation, and action sets, and it holds for problems with Borel state, observation, and action sets [27, 32]. If there is no optimal policy for the problem with belief space, then there is no optimal policy for the original problem.

This paper deals with optimization of expected total discounted costs for models with discrete time. We describe a large class of problems, for which optimal policies exist, satisfy optimality equations, and can

¹Department of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, NY 11794-3600, USA, eugene.feinberg@sunysb.edu

²Institute for Applied System Analysis, National Technical University of Ukraine "Kyiv Polytechnic Institute", Peremogy ave., 37, build, 35, 03056, Kyiv, Ukraine, kasyanov@i.ua.

³National Technical University of Ukraine "Kyiv Polytechnic Institute", Peremogy ave., 37, build, 1, 03056, Kyiv, Ukraine, mzz@kpi.ua

be found by value iterations. For a particular model of Partially Observable Markov Decision Processes (POMDPs) the related studies are [16, 22].

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, 7, 30] 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 [26], there are two different POMDP models in the literature, which we call POMDP_1 and POMDP_2 . Platzman [26] introduced a ‘‘plant’’ model, which we call Platzman’s model, and this model is more general than POMDP_1 and POMDP_2 .

In order to explain that Platzman’s plant is a particular case of an MDPII, let us change the notation w_t to x_t in this paragraph, that is $x_t := w_t$, and this change takes place only in this paragraph. Then the transition probability for an MDP is $P(dx_{t+1}, dy_{t+1}|x_t, y_t, a_t)$, and x_t and y_t can be interpreted as a hidden state and an observation. 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 $P(dx_{t+1}, dy_{t+1}|x_t, a_t)$. POMDP_i are Platzman’s models whose transition probabilities have special structural properties. These properties are $P(dx_{t+1}, dy_{t+1}|x_t, a_t) = P_1(dx_{t+1}|x_t, a_t) Q_1(dy_{t+1}|x_t, a_t)$ for POMDP_1 and $P(dx_{t+1}, dy_{t+1}|x_t, a_t) = Q_2(dy_{t+1}|a_t, x_{t+1}) P_2(dx_{t+1}|x_t, a_t)$ for POMDP_2 .

This paper introduces the class of MDPIIs with semi-uniform Feller transition probabilities. We show that, if an MDPII has a transition probability from this class, then the transition probability of the belief-MDP also belongs to this class. Under mild conditions on cost functions, there are optimal policies for MDPs with semi-uniform Feller transition probabilities. This paper provides several sufficient conditions for the existence of optimal policies, validity of optimality equations, and convergence of value iterations. This general theory implies the following sufficient conditions for the existence of optimal policies for 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 [16]); (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 [22] for uncontrolled observation kernels, that is, $Q_2(y_{t+1}|a_t, x_{t+1}) = Q_2(y_{t+1}|x_{t+1})$.

The results are presented without proofs. The proofs can be found in [18, 19].

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 $s \in \mathbb{S}$ and $\delta > 0$ denote by $B_{\delta}(s)$ and $\bar{B}_{\delta}(s)$ respectively the open and closed balls in the metric space \mathbb{S} of radius δ with center s and by $S_{\delta}(s)$ the sphere in \mathbb{S} of radius δ with center s . Note that $S_{\delta}(s) = \bar{B}_{\delta}(s) \setminus B_{\delta}(s)$. For a subset S of \mathbb{S} let \bar{S} denote the *closure of S* and S° the *interior of S*. Then $S^{\circ} \subset S \subset \bar{S}$. S° is open and \bar{S} is closed. $\partial S := \bar{S} \setminus S^{\circ}$ denotes the *boundary of S*. In particular, $\partial B_{\delta}(s) = S_{\delta}(s)$. We denote by $\mathbb{P}(\mathbb{S})$ the *set of probability measures* on $(\mathbb{S}, \mathcal{B}(\mathbb{S}))$. A sequence of probability measures $\{\mu^{(n)}\}_{n=1,2,\dots}$ from $\mathbb{P}(\mathbb{S})$ *converges weakly* to $\mu \in \mathbb{P}(\mathbb{S})$ if for any bounded continuous function f on \mathbb{S}

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

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

$$\sup_{C \in \mathcal{B}(\mathbb{S})} |\mu^{(n)}(C) - \mu(C)| \rightarrow 0 \text{ as } n \rightarrow \infty; \tag{1}$$

see [17] 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 of probability measures, when \mathbb{S} is a separable metric space; [25, Chapter II]. Moreover, according to Bogachev [6, 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*

$$\rho_{\mathbb{P}(\mathbb{S})}(\mu, \nu) := \sup \left\{ \left| \int_{\mathbb{S}} f(s) \mu(ds) - \int_{\mathbb{S}} f(s) \nu(ds) \right| \right. \\ \left. f \in \text{Lip}_1(\mathbb{S}), \sup_{s \in \mathbb{S}} |f(s)| \leq 1 \right\}, \quad (2)$$

$\mu, \nu \in \mathbb{P}(\mathbb{S})$, where $\text{Lip}_1(\mathbb{S}) := \{f : \mathbb{S} \mapsto \mathbb{R}, |f(s_1) - f(s_2)| \leq \rho_{\mathbb{S}}(s_1, s_2), \forall s_1, s_2 \in \mathbb{S}\}$.

For a Borel subset S of a metric space $(\mathbb{S}, \rho_{\mathbb{S}})$, where $\rho_{\mathbb{S}}$ is a metric, we always consider the metric space (S, ρ_S) , where $\rho_S := \rho_{\mathbb{S}}|_{S \times S}$. A subset B of S is called open (closed) in S if B is open (closed respectively) in (S, ρ) . 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 (closed respectively). 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 $\Psi(\cdot|\cdot) : \mathcal{B}(\mathbb{S}_1) \times \mathbb{S}_2 \mapsto [0, 1]$, 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 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 [24].

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 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:

$$\Psi(A, B|s_3) := \Psi(A \times B|s_3). \quad (3)$$

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)* 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 $a \in \mathbb{A}(y)$ 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) $c : \mathbb{W} \times \mathbb{Y} \times \mathbb{A} \mapsto \overline{\mathbb{R}}_+ = [0, +\infty]$ is a *one-step cost function*; Dynkin and Yushkevich [8, Chapter 8], Rhenius [27], Yushkevich [32] (see also Rieder [28] and Bäuerle and Rieder [3] for a version of this model with transition probabilities having densities)

The corresponding stochastic sequence 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 $t = 0, 1, \dots$, 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 $h_t := (y_0, a_0, y_1, a_1, \dots, y_{t-1}, a_{t-1}, y_t) \in \mathbb{H}_t$ for all $t = 1, 2, \dots$, where $\mathbb{H}_0 := \mathbb{Y}$ and $\mathbb{H}_t := \mathbb{H}_{t-1} \times \mathbb{A} \times \mathbb{Y}$ if $t = 1, 2, \dots$. Then a *policy* for the MDPII is defined as a sequence $\pi = \{\pi_t\}$ such that, for each $t = 0, 1, \dots$, π_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 [4, pp. 140-141] or Hernández-Lerma and Lasserre [21, p.178]) implies that a policy $\pi \in \Pi$, initial distribution $p \in \mathbb{P}(\mathbb{W})$ 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 of 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^π .

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

$$v_{T,\alpha}^\pi(p) := \mathbb{E}_p^\pi \sum_{t=0}^{T-1} \alpha^t c(w_t, y_t, a_t), \quad p \in \mathbb{P}(\mathbb{W}), \quad (4)$$

where $\alpha \geq 0$ is the discount factor, $v_{0,\alpha}^\pi(p) = 0$. When $T = \infty$, (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 $v_{T,\alpha}^\pi(p) = v_{T,\alpha}(p)$, and it is called *discount-optimal* if $v_\alpha^\pi(p) = v_\alpha(p)$ for all $p \in \mathbb{P}(\mathbb{W})$.

We recall that a Markov decision process (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 singleton and $\mathbb{Y} = \mathbb{X}$, where we follow the convention that $\mathbb{W} \times \mathbb{X} = \mathbb{X}$ in this case. For an MDP all states are observable. If the initial state distribution p , which is a probability on \mathbb{X} , is concentrated at $x \in \mathbb{X}$, we write $g^\pi(x)$ and $g(x)$ instead of $g^\pi(p)$ and $g(p)$. 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 an MDP called the *Markov decision process with complete information (MDPCI or belief-MDP)* ([4, 8, 21, 27, 32]). For epoch $t = 0, 1, \dots$ consider the joint conditional probability $R(B \times C | z_t, y_t, a_t)$ on next state (w_{t+1}, y_{t+1}) given the current state (z_t, y_t) and the current control action a_t :

$$R(B \times C | z, y, a) := \int_{\mathbb{W}} P(B \times C | w, y, a) z(dw), \quad (5)$$

$B \in \mathcal{B}(\mathbb{W}), C \in \mathcal{B}(\mathbb{Y}), (z, y, a) \in \mathbb{P}(\mathbb{W}) \times \mathbb{Y} \times \mathbb{A}$. According to Bertsekas and Shreve [4, 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

$$R(B \times C | z, y, a) = \int_C H(B | z, y, a, y') R(\mathbb{W}, dy' | z, y, a), \quad (6)$$

$B \in \mathcal{B}(\mathbb{W})$, $C \in \mathcal{B}(\mathbb{Y})$, $(z, y, a) \in \mathbb{P}(\mathbb{W}) \times \mathbb{Y} \times \mathbb{A}$. The stochastic kernel $H(\cdot | z, y, a, y')$ introduced in (6) defines a measurable mapping $H : \mathbb{P}(\mathbb{W}) \times \mathbb{Y} \times \mathbb{A} \times \mathbb{Y} \mapsto \mathbb{P}(\mathbb{W})$. 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 $(z, y, a) \in \mathbb{P}(\mathbb{W}) \times \mathbb{Y} \times \mathbb{A}$.

The MDPCI is defined as an MDP $(\mathbb{P}(\mathbb{W}) \times \mathbb{Y}, \mathbb{A}, q, \bar{c})$, where

- (i) $\mathbb{P}(\mathbb{W}) \times \mathbb{Y}$ is the state space;
- (ii) \mathbb{A} is the action set available at all state $(z, y) \in \mathbb{P}(\mathbb{W}) \times \mathbb{Y}$;
- (iii) 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 $C \in \mathcal{B}(\mathbb{Y})$,

$$q(D \times C | z, y, a) := \int_C \mathbf{I}\{H(z, y, a, y') \in D\} R(\mathbb{W}, dy' | z, y, a), \quad (7)$$

where $\mathbf{I}B$ denotes the *indicator of an event* B ;

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

$$\bar{c}(z, y, a) := \int_{\mathbb{W}} c(w, y, a) z(dw), \quad (8)$$

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

Additional details can be found in Yushkevich [32], Bertsekas and Shreve [4, Corollary 7.27.1, p. 139], or Dynkin and Yushkevich [8, p. 215]. Note that the measurable particular choice of stochastic kernel H from (6) does not effect the definition of q from (7).

In Section 7 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, validity of optimality equations, and convergence of value iterations for MDPCIs follow from Theorem 6.3. If ϕ is a stationary discount-optimal policy for the MDPCI, then the discount optimal information policy π for the MDPII can be defined as $\pi(y_0, \dots, y_t) = \phi(z_t)$, where z_t is the posterior distribution of the unobservable component w_t of the state x_t given the observations y_0, \dots, y_t , the initial distribution p of w_0 , and if $t > 0$, the decisions $\phi(z_s)$ for $s = 0, \dots, t - 1$. Finite-horizon discount optimal policies for MDPII can be computed similarly from Markov finite-horizon optimal policies for MDPCI.

4 Semi-Uniform Feller Kernels and Their Significance for MDMII

In this section we formulate the semi-uniform Feller property for stochastic kernels. Theorem 4.2 states that the transition kernel P of an MDMII is semi-uniform Feller if and only if the stochastic kernel q for the corresponding MDPCI satisfies this property.

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

Definition 4.1. A stochastic kernel Ψ on $\mathbb{S}_1 \times \mathbb{S}_2$ given \mathbb{S}_3 is semi-uniform 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 ,

$$\lim_{n \rightarrow \infty} \sup_{B \in \mathcal{B}(\mathbb{S}_2)} \left| \int_{\mathbb{S}_1} f(s_1) \Psi(ds_1, B | s_3^{(n)}) - \int_{\mathbb{S}_1} f(s_1) \Psi(ds_1, B | s_3) \right| = 0. \quad (9)$$

We recall that the marginal measure $\Psi(ds_1, B|s_3)$, $s_3 \in \mathbb{S}_3$, is defined in (3). The term ‘‘semi-uniform’’ is used in Definition 4.1 because the uniform property holds in (9) 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 [24]. According to Corollary 5.12, a semi-uniform Feller stochastic kernel is weakly continuous.

Theorem 4.2. *Let us consider an MDMII $(\mathbb{W} \times \mathbb{Y}, \mathbb{A}, P, c)$ and its corresponding MDPCI $(\mathbb{P}(\mathbb{W}) \times \mathbb{Y}, \mathbb{A}, q, \bar{c})$. The transition kernel q on $\mathbb{P}(\mathbb{W}) \times \mathbb{Y}$ given $\mathbb{P}(\mathbb{W}) \times \mathbb{Y} \times \mathbb{A}$ for the MDPCI $(\mathbb{P}(\mathbb{W}) \times \mathbb{Y}, \mathbb{A}, q, \bar{c})$ is semi-uniform Feller if and only if the transition kernel P on $\mathbb{W} \times \mathbb{Y}$ given $\mathbb{W} \times \mathbb{Y} \times \mathbb{A}$ for the MDMII $(\mathbb{W} \times \mathbb{Y}, \mathbb{A}, P, c)$ is semi-uniform Feller.*

5 Properties of Semi-Uniform Feller Stochastic Kernels

Let us consider some basic definitions.

Definition 5.1. *Let \mathbb{S} be a metric space. A function $f : \mathbb{S} \mapsto \overline{\mathbb{R}}$ is called*

- (i) *lower semi-continuous (l.s.c.) at a point $s \in \mathbb{S}$ if $\liminf_{s' \rightarrow s} f(s') \geq f(s)$;*
- (ii) *upper semi-continuous at $s \in \mathbb{S}$ if $-f$ is lower semi-continuous at s ;*
- (iii) *continuous at $s \in \mathbb{S}$ if f is both lower and upper semi-continuous at s ;*
- (iv) *lower / upper semi-continuous (continuous respectively) (on \mathbb{S}) if f is lower / upper semi-continuous (continuous respectively) 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 lower semi-continuous functions, and all real-valued continuous functions respectively defined on the metric space \mathbb{S} . The following definitions are taken from [11].

Definition 5.2. *A set $F \subset \mathbb{F}(\mathbb{S})$ of real-valued functions on a metric space \mathbb{S} is called (i) lower semi-equicontinuous at a point $s \in \mathbb{S}$ if $\liminf_{s' \rightarrow s} \inf_{f \in F} (f(s') - f(s)) \geq 0$; (ii) upper semi-equicontinuous at a point $s \in \mathbb{S}$ if the set $\{-f : f \in F\}$ is lower semi-equicontinuous at $s \in \mathbb{S}$; (iii) equicontinuous at a point $s \in \mathbb{S}$, if F is both lower and upper semi-equicontinuous at $s \in \mathbb{S}$, that is, $\limsup_{s' \rightarrow s} \sup_{f \in F} |f(s') - f(s)| = 0$;*

(iv) *lower / upper semi-equicontinuous (equicontinuous respectively) (on \mathbb{S}) if it is lower / upper semi-equicontinuous (equicontinuous respectively) at all $s \in \mathbb{S}$; (v) uniformly bounded (on \mathbb{S}), if there exists a constant $M < +\infty$ such that $|f(s)| \leq M$ for all $s \in \mathbb{S}$ and for all $f \in F$.*

Obviously, if a set $F \subset \mathbb{F}(\mathbb{S})$ is lower semi-equicontinuous, then $F \subset \mathbb{L}(\mathbb{S})$. Moreover, if a set $F \subset \mathbb{F}(\mathbb{S})$ is equicontinuous, then $F \subset \mathbb{C}(\mathbb{S})$.

5.1 WTV-continuous Stochastic Kernels

In this subsection we introduce WTV-continuous stochastic kernels, which are more general than semi-uniform Feller, and investigate their basic properties. In particular, Theorem 5.4 provides the equivalent definition of a WTV-continuous stochastic kernel under the additional condition of continuity in total variation of its marginal kernel. Theorem 5.8 establishes a necessary and sufficient condition for a stochastic kernel to be WTV-continuous and to have its marginal kernel to be continuous in total variation. This condition is Assumption 5.6, whose stronger version was introduced in [15, Theorem 4.4]. Theorem 5.9 describes the preservation of WTV-continuity 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. For each set $A \in \mathcal{B}(\mathbb{S}_1)$ consider the set of functions

$$F_A^\Psi = \{s_3 \mapsto \Psi(A \times B|s_3) : B \in \mathcal{B}(\mathbb{S}_2)\} \quad (10)$$

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 5.3. 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 set of functions $F_{\mathcal{O}}^{\Psi}$ is lower semi-equicontinuous on \mathbb{S}_3 .

Similarly to Parthasarathy [25, 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 WTV-continuous stochastic kernels.

Theorem 5.4. For a stochastic kernel Ψ on $\mathbb{S}_1 \times \mathbb{S}_2$ given \mathbb{S}_3 , such that the marginal kernel $\Psi(\mathbb{S}_1, \cdot | \cdot)$ on \mathbb{S}_2 given \mathbb{S}_3 is continuous in total variation, the following conditions are equivalent: (a) the stochastic kernel Ψ on $\mathbb{S}_1 \times \mathbb{S}_2$ given \mathbb{S}_3 is semi-uniform 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

$$\lim_{n \rightarrow \infty} \sup_{B \in \mathcal{B}(\mathbb{S}_2)} \left(\Psi(C \times B | s_3^{(n)}) - \Psi(C \times B | s_3) \right) = 0; \quad (11)$$

(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 $\Psi(\mathbb{S}_1, \partial A | s_3) = 0$,

$$\lim_{n \rightarrow \infty} \sup_{B \in \mathcal{B}(\mathbb{S}_2)} |\Psi(A \times B | s_3^{(n)}) - \Psi(A \times B | s_3)| = 0; \quad (12)$$

(e) if $s_3^{(n)}$ converges to s_3 in \mathbb{S}_3 , then, for each nonnegative bounded lower semi-continuous function f on \mathbb{S}_1 ,

$$\liminf_{n \rightarrow \infty} \inf_{B \in \mathcal{B}(\mathbb{S}_2)} \left(\int_{\mathbb{S}_1} f(s_1) \Psi(ds_1, B | s_3^{(n)}) - \int_{\mathbb{S}_1} f(s_1) \Psi(ds_1, B | s_3) \right) = 0. \quad (13)$$

Corollary 5.5. A stochastic kernel Ψ on $\mathbb{S}_1 \times \mathbb{S}_2$ given \mathbb{S}_3 is semi-uniform Feller if and only if the marginal kernel $\Psi(\mathbb{S}_1, \cdot | \cdot)$ on \mathbb{S}_2 given \mathbb{S}_3 is continuous in total variation and at least one of conditions (b)–(e) of Theorem 5.4 holds.

Let us consider the following assumption. According to Example 5.7, Assumption 5.6 is weaker than combined assumptions (i) and (ii) in [15, 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 5.6. 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 set of functions $F_{\mathcal{O}}^{\Psi}$, defined in (10), is equicontinuous at s_3 .

Note that Assumption 5.6(ii) holds if and only if for each finite intersection $\mathcal{O} = \cap_{i=1}^k \mathcal{O}_i$ of sets $\mathcal{O}_i \in \tau_b^{s_3}(\mathbb{S}_1)$, $i = 1, 2, \dots, k$,

$$\lim_{n \rightarrow \infty} \sup_{B \in \mathcal{B}(\mathbb{S}_2)} \left| \Psi(\mathcal{O} \times B | s_3^{(n)}) - \Psi(\mathcal{O} \times B | s_3) \right| = 0 \quad (14)$$

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

The following example demonstrates that the version of Assumption 5.6 with the same base $\tau_b(\mathbb{S}_1)$ for all $s_3 \in \mathbb{S}_3$ is essentially stronger than Assumption 5.6.

Example 5.7. Let $\mathbb{S}_1 = \mathbb{S}_3 := \mathbb{R}$, $\text{Card } \mathbb{S}_2 = 1$, and $\Psi(S_1|s_3) := \mathbf{I}\{s_3 \in S_1\} \forall S_1 \in \mathcal{B}(\mathbb{S}_1)$ and $\forall s_3 \in \mathbb{S}_3$.

Let us prove that Assumption 5.6 holds. Indeed, for a fixed $s_3 \in \mathbb{R}$ let us consider the countable base $\tau_b^{s_3}(\mathbb{R}) = \{\mathbb{R}\} \cup \{(a + \sqrt{2}, b + \sqrt{2}) : a, b \in \mathbb{Q}, a < b\}$ for $s_3 \in \mathbb{Q}$, and $\tau_b^{s_3}(\mathbb{R}) = \{\mathbb{R}\} \cup \{(a, b) : a, b \in \mathbb{Q}, a < b\}$ for $s_3 \notin \mathbb{Q}$, where \mathbb{Q} is the set of rational numbers. Note that this base satisfies the following properties : (a) $\mathbb{R} \in \tau_b^{s_3}(\mathbb{R})$, (b) $\mathcal{O} = \cap_{i=1}^k \mathcal{O}_i \in \tau_b^{s_3}(\mathbb{R})$ for any $k = 1, 2, \dots$ and $\{\mathcal{O}_i\}_{i=1}^k \subset \tau_b^{s_3}(\mathbb{R})$, and (c) $s_3 \notin \partial \mathcal{O}$ for all $\mathcal{O} \in \tau_b^{s_3}(\mathbb{R})$. Statement (a) implies that Assumption 5.6(i) holds. Assumption 5.6(ii) holds because, according to (b) each finite intersection $\mathcal{O} = \cap_{i=1}^k \mathcal{O}_i$ of sets $\mathcal{O}_i \in \tau_b^{s_3}(\mathbb{R})$, $i = 1, 2, \dots, k$, belongs to $\tau_b^{s_3}(\mathbb{R})$, and according to (c) the function $s \mapsto \mathbf{I}\{s \in \mathcal{O}\}$ is continuous at s_3 . Thus, Assumption 5.6 holds.

Assumption 5.6 does not hold with the same base $\tau_b(\mathbb{S}_1)$ for all $s_3 \in \mathbb{S}_3$ because for any nonempty open set $\mathcal{O} \in \tau(\mathbb{S}_1) \setminus \{\mathbb{S}_1\}$ there exist $s_3^* \in \partial \mathcal{O}$ and a sequence $\{s_3^{(n)}\}_{n=1,2,\dots} \subset \mathcal{O}$ such that $s_3^{(n)} \rightarrow s_3^*$ in \mathbb{S}_3 as $n \rightarrow \infty$, and, therefore, $\Psi(\mathcal{O}|s_3^{(n)}) = \mathbf{I}\{s_3^{(n)} \in \mathcal{O}\} = 1 \not\rightarrow 0 = \mathbf{I}\{s_3^* \in \mathcal{O}\} = \Psi(\mathcal{O}|s_3^*)$ as $n \rightarrow \infty$, that is, the set of functions $F_{\mathcal{O}}^{\Psi}$ is not equicontinuous at s_3^* . \square

Theorem 5.8 shows that Assumptions 5.6 is a necessary and sufficient condition for semi-uniform Feller continuity.

Theorem 5.8. *The stochastic kernel Ψ on $\mathbb{S}_1 \times \mathbb{S}_2$ given \mathbb{S}_3 is semi-uniform Feller if and only if Assumption 5.6 holds for this kernel.*

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 Ξ_f on $\mathbb{S}_1 \times \mathbb{S}_2$ given $\mathbb{P}(\mathbb{S}_3) \times \mathbb{S}_4$ defined by

$$\Xi_f(A \times B|\mu, s_4) := \int_{\mathbb{S}_3} \Xi(A \times B|s_3, s_4)\mu(ds_3), \quad (15)$$

$A \in \mathcal{B}(\mathbb{S}_1)$, $B \in \mathcal{B}(\mathbb{S}_2)$, $\mu \in \mathbb{P}(\mathbb{S}_3)$, $s_4 \in \mathbb{S}_4$.

The following theorem establishes the preservation of WTV-continuity of the integration operation in (15).

Theorem 5.9. *The stochastic kernel Ξ_f on $\mathbb{S}_1 \times \mathbb{S}_2$ given $\mathbb{P}(\mathbb{S}_3) \times \mathbb{S}_4$ is WTV-continuous if and only if Ξ on $\mathbb{S}_1 \times \mathbb{S}_2$ given $\mathbb{S}_3 \times \mathbb{S}_4$ is WTV-continuous.*

5.2 Continuity Properties of Posterior Distributions

In this section we establish sufficient conditions for semi-uniform Feller continuity of posterior distributions. The main result of this section is Theorem 5.11.

Let $\mathbb{S}_1, \mathbb{S}_2$, and \mathbb{S}_3 be Borel subsets of Polish spaces, and Ψ on $\mathbb{S}_1 \times \mathbb{S}_2$ given \mathbb{S}_3 be a stochastic kernel. By Bertsekas and Shreve [4, Proposition 7.27], there exists a stochastic kernel Φ on \mathbb{S}_1 given $\mathbb{S}_2 \times \mathbb{S}_3$ such that

$$\Psi(A \times B|s_3) = \int_B \Phi(A|s_2, s_3)\Psi(\mathbb{S}_1, ds_2|s_3), \quad (16)$$

$A \in \mathcal{B}(\mathbb{S}_1)$, $B \in \mathcal{B}(\mathbb{S}_2)$, $s_3 \in \mathbb{S}_3$.

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 $\Phi : \mathbb{S}_2 \times \mathbb{S}_3 \rightarrow \mathbb{P}(\mathbb{S}_1)$, where $\Phi(s_2, s_3)(\cdot) = \Phi(\cdot|s_2, s_3)$. According to Bertsekas and Shreve [4, Corollary 7.27.1], for each $s_3 \in \mathbb{S}_3$ the mapping $\Phi(\cdot, s_3) : \mathbb{S}_2 \rightarrow \mathbb{P}(\mathbb{S}_1)$ is defined $\Psi(\mathbb{S}_1, \cdot|s_3)$ -almost surely uniquely in $s_2 \in \mathbb{S}_2$. Consider the stochastic kernel

$$\phi(D \times B|s_3) := \int_B \mathbf{I}\{\Phi(s_2, s_3) \in D\}\Psi(\mathbb{S}_1, ds_2|s_3), \quad (17)$$

$D \in \mathcal{B}(\mathbb{P}(\mathbb{S}_1))$, $B \in \mathcal{B}(\mathbb{S}_2)$, $s_3 \in \mathbb{S}_3$. In models for decision making with incomplete information, ϕ is the transition probability between belief states, which are posterior distributions of states; (7). Continuity properties of ϕ play the fundamental role in the studies of models with incomplete information. Theorem 5.11 characterizes such properties, and this is the reason for the title of this section.

According to Bertsekas and Shreve [4, Corollary 7.27.1], the particular choice of a stochastic kernel Φ satisfying (16) does not effect the definition of ϕ in (17) because for each $s_3 \in \mathbb{S}_3$ the mapping $\Phi(\cdot, s_3) : \mathbb{S}_2 \rightarrow \mathbb{P}(\mathbb{S}_1)$ is defined $\Psi(\mathbb{S}_1, \cdot | s_3)$ -almost surely uniquely in $s_2 \in \mathbb{S}_2$.

Consider the following assumption.

Assumption 5.10. *There exists a stochastic kernel Φ on \mathbb{S}_1 given $\mathbb{S}_2 \times \mathbb{S}_3$ satisfying (16) such that, if a sequence $\{s_3^{(n)}\}_{n=1,2,\dots} \subset \mathbb{S}_3$ converges to $s_3 \in \mathbb{S}_3$ as $n \rightarrow \infty$, then there exists a subsequence $\{s_3^{(n_k)}\}_{k=1,2,\dots} \subset \{s_3^{(n)}\}_{n=1,2,\dots}$ and a measurable subset B of \mathbb{S}_2 such that*

$$\begin{aligned} \Psi(\mathbb{S}_1 \times B | s_3) = 1 \quad \text{and} \quad \Phi(s_2, s_3^{(n_k)}) \\ \text{converges weakly to } \Phi(s_2, s_3), \quad \text{for all } s_2 \in B. \end{aligned} \tag{18}$$

In other words, the convergence in (18) holds $\Psi(\mathbb{S}_1, ds_2 | s_3)$ -almost surely.

The following theorem, which is the main result of this section, provides necessary and sufficient conditions for WTV-continuity of a stochastic kernel ϕ in terms of the properties of a given stochastic kernel Ψ . This theorem and the results of Subsection 5.1 provide the necessary and sufficient conditions for the semi-uniform 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 5.11. *For a given stochastic kernel Ψ on $\mathbb{S}_1 \times \mathbb{S}_2$ given \mathbb{S}_3 , let the marginal kernel $\Psi(\mathbb{S}_1, \cdot | \cdot)$ on \mathbb{S}_2 given \mathbb{S}_3 is continuous in total variation. Then the following conditions are equivalent: (a) the stochastic kernel Ψ on $\mathbb{S}_1 \times \mathbb{S}_2$ given \mathbb{S}_3 is semi-uniform Feller; (b) Assumption 5.10 holds; (c) if a sequence $\{s_3^{(n)}\}_{n=1,2,\dots} \subset \mathbb{S}_3$ converges to $s_3 \in \mathbb{S}_3$ as $n \rightarrow \infty$, then $\rho_{\mathbb{P}(\mathbb{S}_1)}(\Phi(s_2, s_3^{(n)}), \Phi(s_2, s_3)) \rightarrow 0$ in probability $\Psi(\mathbb{S}_1, ds_2 | s_3)$, 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 , in particular, Kantorovich-Rubinshtein metric defined in (2); (d) the stochastic kernel ϕ on $\mathbb{P}(\mathbb{S}_1) \times \mathbb{S}_2$ given \mathbb{S}_3 is semi-uniform Feller; and each of these statements implies that the stochastic kernels Ψ on $\mathbb{S}_1 \times \mathbb{S}_2$ given \mathbb{S}_3 and ϕ on $\mathbb{P}(\mathbb{S}_1) \times \mathbb{S}_2$ given \mathbb{S}_3 are weakly continuous.*

Corollary 5.12. *A semi-uniform Feller stochastic kernel Ψ on $\mathbb{S}_1 \times \mathbb{S}_2$ given \mathbb{S}_3 is weakly continuous.*

6 Markov Decision Processes with Semi-Uniform Feller Kernels

Let \mathbb{X}_W and \mathbb{X}_Y are Borel subsets of Polish spaces. In this section we consider the special class of MDPs with semi-uniform Feller transition kernels, when $\mathbb{X} := \mathbb{X}_W \times \mathbb{X}_Y$. These results are important for MDPIIs with semi-uniform Feller transition kernels from Section 7, 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

$$\mathcal{D}_f(\lambda; U) = \{y \in U : f(y) \leq \lambda\}, \quad \lambda \in \mathbb{R}. \tag{19}$$

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 6.1. ([13, Definition 1.1]) *A function $u : \mathbb{S}_1 \times \mathbb{S}_2 \mapsto \overline{\mathbb{R}}$ 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 on lower semicontinuity of the minimum holds with possibly noncompact action sets; [13, Theorem 1.2]. In particular, it allows us to consider the MDPII $(\mathbb{W} \times \mathbb{Y}, \mathbb{A}, P, c)$ with possibly noncompact action space \mathbb{A} and unbounded one-step cost c and examine the convergence of value and policy iterations for this model in Theorem 7.1, in particular, for Platzman's models and partially observable MDPs in Theorem 7.7 and its corollaries.

Definition 6.2. A Borel-measurable function $u : \mathbb{S}_1 \times \mathbb{S}_2 \times \mathbb{S}_3 \mapsto \overline{\mathbb{R}}$ is called measurable \mathbb{K} -inf-compact (MKK-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 $\mathbb{S}_1 \times \mathbb{S}_3$.

As follows from Theorem 7.2 below, the function $\bar{c} : \mathbb{P}(\mathbb{S}_1) \times \mathbb{S}_2 \times \mathbb{S}_3 \mapsto \overline{\mathbb{R}}$ defined in (8) is MKK-inf-compact if $c : \mathbb{S}_1 \times \mathbb{S}_2 \times \mathbb{S}_3 \mapsto \overline{\mathbb{R}}_+$ is an MKK-inf-compact function.

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 $\varphi : \mathbb{X} \rightarrow \overline{\mathbb{R}}$ such that $w \mapsto \varphi(w, y)$ is lower semi-continuous on \mathbb{X}_W for each $y \in \mathbb{X}_Y$. For any $\alpha \geq 0$ and $u \in LW(\mathbb{X})$, we consider

$$\eta_u^\alpha(x, a) = c(x, a) + \alpha \int_{\mathbb{X}} u(\tilde{x})q(d\tilde{x}|x, a), \quad (20)$$

$(x, a) \in \mathbb{X} \times \mathbb{A}$.

The following theorem is the main result of this section. It is applied in Theorem 7.1 to MDPCIs $(\mathbb{P}(\mathbb{W}) \times \mathbb{Y}, \mathbb{A}, q, \bar{c})$.

Theorem 6.3. (Expected Total Discounted Costs) *Let us consider an MDP $(\mathbb{X}, \mathbb{A}, q, c)$ with $\mathbb{X} = \mathbb{X}_W \times \mathbb{X}_Y$, q on \mathbb{X} given $\mathbb{X} \times \mathbb{A}$ being semi-uniform Feller, and the nonnegative function $c : \mathbb{X} \times \mathbb{A} \mapsto \overline{\mathbb{R}}$ being MKK-inf-compact. Then*

- (i) *the functions $v_{t,\alpha}$, $t = 0, 1, \dots$, and v_α belongs to $LW(\mathbb{X})$, and $v_{t,\alpha}(x) \uparrow v_\alpha(x)$ as $t \rightarrow +\infty$ for all $x \in \mathbb{X}$;*
- (ii) *$v_{t+1,\alpha}(x) = \min_{a \in \mathbb{A}} \eta_{v_{t,\alpha}}^\alpha(x, a)$, $x \in \mathbb{X}$, $t = 0, 1, \dots$, where $v_{0,\alpha}(x) = 0$ for all $x \in \mathbb{X}$, and the nonempty sets $A_{t,\alpha}(x) := \{a \in \mathbb{A} : v_{t+1,\alpha}(x) = \eta_{v_{t,\alpha}}^\alpha(x, a)\}$, $x \in \mathbb{X}$, $t = 0, 1, \dots$, satisfy the following properties: (a) the graph $\text{Gr}(A_{t,\alpha}) = \{(x, a) : x \in \mathbb{X}, a \in A_{t,\alpha}(x)\}$, $t = 0, 1, \dots$, 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 $T = 1, 2, \dots$, there exists a Markov optimal T -horizon policy $(\phi_0, \dots, \phi_{T-1})$ and if, for an T -horizon Markov policy $(\phi_0, \dots, \phi_{T-1})$ the inclusions $\phi_{T-1-t}(x) \in A_{t,\alpha}(x)$, $x \in \mathbb{X}$, $t = 0, \dots, T-1$, hold then this policy is T -horizon optimal;*
- (iv) *for $\alpha \in [0, 1)$ $v_\alpha(x) = \min_{a \in \mathbb{A}} \eta_{v_\alpha}^\alpha(x, a)$, $x \in \mathbb{X}$, and the nonempty sets $A_\alpha(x) := \{a \in \mathbb{A} : v_\alpha(x) = \eta_{v_\alpha}^\alpha(x, a)\}$, $x \in \mathbb{X}$, satisfy the following properties: (a) the graph $\text{Gr}(A_\alpha) = \{(x, a) : x \in \mathbb{X}, a \in A_\alpha(x)\}$ 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 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 $x \in \mathbb{X}$.*

7 Total-cost Optimal Policies for MDPII and Corollaries for Platzman's Model and Partially Observable MDPs

Let us state Theorem 7.1 which is the corollary to Theorems 4.2, 5.11, and to the reduction of MDPIIs to MDPCIs. Then we consider Platzman's model and POMDPs.

Theorem 7.1. *Let $(\mathbb{W} \times \mathbb{Y}, \mathbb{A}, P, c)$ be an MDPII and $(\mathbb{P}(\mathbb{W}) \times \mathbb{Y}, \mathbb{A}, q, \bar{c})$ be its MDPCI. Then the following conditions are equivalent :*

- (a) *Assumption 5.6 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 semi-uniform Feller;*
- (c) *the stochastic 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 at least one of equivalent statements (a)–(d) of Theorem 5.11 with $\mathbb{S}_1 := \mathbb{W}$, $\mathbb{S}_2 := \mathbb{Y}$, $\mathbb{S}_3 := \mathbb{P}(\mathbb{W}) \times \mathbb{Y} \times \mathbb{A}$, $\Psi := R$, $\Phi := H$, and $\phi := q$ holds;*
- (d) *the stochastic kernel q on $\mathbb{P}(\mathbb{W}) \times \mathbb{Y}$ given $\mathbb{P}(\mathbb{W}) \times \mathbb{Y} \times \mathbb{A}$ is semi-uniform Feller;*

and each of them implies that the stochastic kernel q on $\mathbb{P}(\mathbb{W}) \times \mathbb{Y}$ given $\mathbb{P}(\mathbb{W}) \times \mathbb{Y} \times \mathbb{A}$ is weakly continuous. Moreover, if nonnegative function c is $\mathbb{M}\mathbb{K}$ -inf-compact and one of the equivalent conditions (a)–(d) holds, then assumptions and all the conclusions of Theorem 6.3 for the MDPCI $(\mathbb{P}(\mathbb{W}) \times \mathbb{Y}, \mathbb{A}, q, \bar{c})$ hold.

According to [27, 32], 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 $(\mathbb{W} \times \mathbb{Y}, \mathbb{A}, P, c)$. [15, Theorem 4.4] establishes weak continuity of the transition kernel for the MDPCI under the more restrictive assumption than statement (a) of Theorem 7.1 when the countable base in this condition does not depend on the argument $s_3 = (w, y, a)$; see also Example 5.7. Moreover, for any $T = 1, 2, \dots$ and $\alpha \geq 0$, the value functions $\tilde{V}_{T,\alpha}(z, y), \tilde{V}_\alpha(z, y)$ in the MDPCI $(\mathbb{P}(\mathbb{W}) \times \mathbb{Y}, \mathbb{A}, q, \bar{c})$, under the conditions of Theorem 7.1, are concave by $z \in \mathbb{P}(\mathbb{W})$ for any fixed $y \in \mathbb{Y}$,

The proof of Theorem 7.1 uses the following preservation property for $\mathbb{M}\mathbb{K}$ -inf-compactness.

Theorem 7.2. *If $c : \mathbb{W} \times \mathbb{Y} \times \mathbb{A} \mapsto \overline{\mathbb{R}}_+$ is an $\mathbb{M}\mathbb{K}$ -inf-compact function, then the function $\bar{c} : \mathbb{P}(\mathbb{W}) \times \mathbb{Y} \times \mathbb{A} \mapsto \overline{\mathbb{R}}$ defined in (8) is $\mathbb{M}\mathbb{K}$ -inf-compact too.*

A particular case of an MDPII is Platzman's model introduced in [26].

Definition 7.3. *Platzman's model is an MDPII $(\mathbb{W} \times \mathbb{Y}, \mathbb{A}, P, c)$, where P is a stochastic kernel on $\mathbb{W} \times \mathbb{Y}$ given $\mathbb{W} \times \mathbb{A}$.*

Platzman's model is an MDPII with the transition kernel $P(\cdot | w, y, a)$, which does not depend on y . Theorem 7.1 implies the following corollary.

Corollary 7.4. *Let $(\mathbb{W} \times \mathbb{Y}, \mathbb{A}, P, c)$ be a Platzman's model. Then the stochastic kernel P on $\mathbb{W} \times \mathbb{Y}$ given $\mathbb{W} \times \mathbb{A}$ is semi-uniform Feller if and only if at least one of the equivalent conditions (a), (c), or (d) of Theorem 7.1 holds. Moreover, if the nonnegative function c is $\mathbb{M}\mathbb{K}$ -inf-compact and P on $\mathbb{W} \times \mathbb{Y}$ given $\mathbb{W} \times \mathbb{A}$ is semi-uniform Feller, then all assumptions and statements of Theorem 7.1 hold.*

As noticed in [26], the special cases of Platzman's model are two MDPs defined below, which we denote as POMDP₁ and POMDP₂; see Definitions 7.5, 7.6.

Let $i = 1, 2$, \mathbb{X} , \mathbb{Y} , and \mathbb{A} be Borel subsets of Polish spaces, $P_i(dx'|x, a)$ be a stochastic kernel on \mathbb{X} given $\mathbb{X} \times \mathbb{A}$, $Q_i(dy|a, x)$ be a stochastic kernel on \mathbb{Y} given $\mathbb{A} \times \mathbb{X}$, $Q_{0,i}(dy|x)$ be a stochastic kernel on \mathbb{Y} given \mathbb{X} , p be a probability distribution on \mathbb{X} .

Definition 7.5. POMDP₁ $(\mathbb{X}, \mathbb{Y}, \mathbb{A}, P_1, Q_1, c)$ is specified by Platzmann's model $(\mathbb{X} \times \mathbb{Y}, \mathbb{A}, P, c)$ with $P(B \times C|x, y, a) := P_1(B|x, a)Q_1(C|x, a)$, $B \in \mathcal{B}(\mathbb{X})$, $C \in \mathcal{B}(\mathbb{Y})$, $x \in \mathbb{X}$, $y \in \mathbb{Y}$, $a \in \mathbb{A}$.

Let $(\mathbb{X}, \mathbb{Y}, \mathbb{A}, P_1, Q_1, c)$ be POMDP₁. Then, the stochastic kernel R on $\mathbb{X} \times \mathbb{Y}$ given $\mathbb{P}(\mathbb{X}) \times \mathbb{Y} \times \mathbb{A}$ specified in (5) takes the following form,

$$R(B \times C|z, y, a) := \int_{\mathbb{X}} Q_1(C|x, a)P_1(B|x, a)z(dx), \quad (21)$$

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

Definition 7.6. POMDP₂ $(\mathbb{X}, \mathbb{Y}, \mathbb{A}, P_2, Q_2, c)$ is specified by Platzmann's model $(\mathbb{X} \times \mathbb{Y}, \mathbb{A}, P, c)$ with

$$P(B \times C|x, y, a) := \int_B Q_2(C|a, x')P_2(dx'|x, a),$$

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

For a POMDP₂ $(\mathbb{X}, \mathbb{Y}, \mathbb{A}, P_2, Q_2, c)$ the stochastic kernel R on $\mathbb{X} \times \mathbb{Y}$ given $\mathbb{P}(\mathbb{X}) \times \mathbb{Y} \times \mathbb{A}$ specified in (5) takes the following form,

$$R(B \times C|z, y, a) := \int_{\mathbb{X}} \int_B Q_2(C|a, x')P_2(dx'|x, a)z(dx), \quad (22)$$

$B \in \mathcal{B}(\mathbb{X})$, $C \in \mathcal{B}(\mathbb{Y})$, $z \in \mathbb{P}(\mathbb{X})$, $a \in \mathbb{A}$. POMDP₁ is Platzman's model with observations y_{t+1} being "random functions" of w_t and a_t , and POMDP₂ is Platzmann's model, when observations y_{t+1} being "random functions" of a_t and w_{t+1} . Let us apply Theorem 7.1 to POMDP₁ and POMDP₂.

In the following theorem and its corollaries we establish the semi-uniform Feller property for the stochastic kernel q .

Theorem 7.7. Let $i = 1, 2$. The following conditions are equivalent for a POMDP _{i} :

- (a) the stochastic kernel P on $\mathbb{X} \times \mathbb{Y}$ given $\mathbb{X} \times \mathbb{A}$ is semi-uniform Feller;
- (b) Assumption 5.6 holds with $\mathbb{S}_1 := \mathbb{X}$, $\mathbb{S}_2 := \mathbb{Y}$, $\mathbb{S}_3 := \mathbb{X} \times \mathbb{A}$, and $\Psi := P$;
- (c) the stochastic kernel $R(\mathbb{X}, \cdot | \cdot)$ on \mathbb{Y} given $\mathbb{P}(\mathbb{X}) \times \mathbb{Y} \times \mathbb{A}$ is continuous in total variation, and one of equivalent statements (a)–(d) of Theorem 5.11 holds with $\mathbb{S}_1 := \mathbb{X}$, $\mathbb{S}_2 := \mathbb{Y}$, $\mathbb{S}_3 := \mathbb{P}(\mathbb{X}) \times \mathbb{Y} \times \mathbb{A}$, $\Psi := R$, $\Phi := H$, and $\phi := q$;
- (d) the stochastic kernel q on $\mathbb{Y} \times \mathbb{P}(\mathbb{X})$ given $\mathbb{X} \times \mathbb{A}$ is semi-uniform Feller;

and each of them implies that the stochastic kernel q on $\mathbb{P}(\mathbb{X}) \times \mathbb{Y}$ given $\mathbb{P}(\mathbb{X}) \times \mathbb{Y} \times \mathbb{A}$ is weakly continuous. Moreover, if the nonnegative function $c : \mathbb{X} \times \mathbb{Y} \times \mathbb{A} \mapsto \overline{\mathbb{R}}$ is MIIK-inf-compact, and one of the equivalent conditions (a)–(d) holds, then assumptions and all the conclusions of Theorem 7.1 hold for the MDPII $(\mathbb{P}(\mathbb{X}) \times \mathbb{Y}, \mathbb{A}, q, \bar{c})$.

Corollary 7.8. Let us consider POMDP₁. The stochastic kernel P_1 on \mathbb{X} given $\mathbb{X} \times \mathbb{A}$ is weakly continuous and the stochastic kernel Q_1 on \mathbb{Y} given $\mathbb{X} \times \mathbb{A}$ is continuous in total variation if and only if condition (a) of Theorem 7.7 hold and, then, all the conclusions of Theorem 7.7 hold.

Corollary 7.9. Let us consider POMDP₂. Each of the following assumptions :

- (a) the stochastic kernel $P_2(\cdot|x, a)$ on \mathbb{X} given $\mathbb{X} \times \mathbb{A}$ is weakly continuous, and the stochastic kernel $Q_2(\cdot|a, x')$ on \mathbb{Y} given $\mathbb{A} \times \mathbb{X}$ is continuous in total variation; cp. [16, Theorem 3.6];

- (b) *the stochastic kernel $P_2(\cdot|x, a)$ on \mathbb{X} given $\mathbb{X} \times \mathbb{A}$ is continuous in total variation, and the observation kernel $Q_2(\cdot|a, x')$ on \mathbb{Y} given $\mathbb{A} \times \mathbb{X}$ is continuous in a in total variation; cp. [22, Theorem 2];*

implies condition (a) of Theorem 7.7 and all conclusions of Theorem 7.7.

Corollary 7.9 is a more general statement than [16, Theorem 3.6] and [22, Theorem 2]. The first difference is that the cost function in [16] does not depend on observations, and the corresponding cost function $c(x, a)$ is \mathbb{K} -inf-compact. In this paper the cost function $c(x, y, a)$ may depend on the observation y . If $c(x, y, a)$ is $\mathbb{M}\mathbb{K}$ -compact and y is a dummy variable, then the function $c(x, a) = c(x, y, a)$ is \mathbb{K} -inf-compact. The second difference is that in [22, Theorem 2] the observation kernel is $Q_2(\cdot|x)$, and it does not depend on actions. The only assumption in [22, Theorem 2] is continuity of $P_2(\cdot|x, a)$ in total variation. The conclusion of [22, Theorem 2] is weak continuity of the filter transition probability, which follows from Theorem 7.7 and Corollary 5.12.

Let us consider Platzman's plant $(\mathbb{W} \times \mathbb{Y}, \mathbb{A}, P, c)$ 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} ; the one-step cost function $\hat{c} : \mathbb{P}(\mathbb{W}) \times \mathbb{A} \mapsto \overline{\mathbb{R}}$, defined as

$$\hat{c}(z, a) := \int_{\mathbb{W}} c(w, a) z(dw), \quad (23)$$

$z \in \mathbb{P}(\mathbb{W})$, $a \in \mathbb{A}$, and the transition probability

$$\hat{q}(D|z, a) := \int_{\mathbb{Y}} \mathbf{I}\{H(z, a, y') \in D\} R(\mathbb{W}, dy'|z, a), \quad (24)$$

for $(z, a) \in \mathbb{P}(\mathbb{W}) \times \mathbb{A}$ and for $D \in \mathcal{B}(\mathbb{P}(\mathbb{W}))$, where the function H is defined in (6), and it does not depend on y because the function R defined in (5) does not depend on y for Platzman's model; Formulae (23) and (24) follow from (8) and (7) respectively. The described reduction of an MDPCI $(\mathbb{P}(\mathbb{W}) \times \mathbb{Y}, \mathbb{A}, q, \bar{c})$ to the MDP $(\mathbb{P}(\mathbb{W}), \mathbb{A}, \hat{q}, \hat{c})$ holds [9, 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 $c(z_t, a)$ do not depend on y_t . For POMDPs this smaller MDP is well-known; see, e.g., [16, 20].

8 CONCLUSIONS AND FUTURE WORKS

8.1 Conclusions

The paper introduces the property of semi-uniform Feller continuity for transition probabilities and applies it to discrete-time stochastic control problems with incomplete information. The analysis of this property provides new insights and results on the structure of optimal policies for POMDPs, and gives sufficient conditions for the existence of optimal policies and convergence of value iterations.

8.2 Future Works

The future work will investigate whether the developed procedures for statistical learning can be combined with deep learning for solving POMDPs. Another research direction is to try to relax the semi-uniform Feller condition. As the results of this paper show, this condition holds in the known situations when an optimal policy exists, even if this is not stated explicitly in the model, as this takes place with POMDPs.

References

- [1] Aoki, M. (1965) Optimal control of partially observable Markovian systems. *J. Franklin Inst.* 280(5): 367–386.
- [2] Åström, K.J. (1965). Optimal control of Markov processes with incomplete state information. *Journal of Mathematical Analysis and Applications* 10: 174–205.
- [3] Bäuerle, N., Rieder, U. (2011) *Markov Decision Processes with Applications to Finance*, Springer-Verlag, Berlin.
- [4] Bertsekas, D.P., Shreve S.E. (1978) *Stochastic Optimal Control: The Discrete-Time Case*, Academic Press, New York
- [5] Billingsley, P. (1968) *Convergence of Probability Measures*, John Wiley, New York.
- [6] Bogachev, V.I. (2007) *Measure Theory, Volume II*, Springer-Verlag, Berlin.
- [7] Dynkin, E.B. (1965) Controlled random sequences. *Theory Probab. Appl.* 10(1): 1–14.
- [8] Dynkin, E.B., Yushkevich AA (1979) *Controlled Markov Processes*, Springer-Verlag, New York.
- [9] Feinberg, E.A. (2005) On essential information in sequential decision processes. *Math. Meth. Oper. Res.* 62, 399–410.
- [10] Feinberg, E.A., Kasyanov, P.O. (2020) MDPs with Setwise Continuous Transition Probabilities, arXiv:2011.01325
- [11] Feinberg, E.A., Kasyanov, P.O., Liang, Y. (2020) 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(2): 270–291.
- [12] Feinberg, E.A., Kasyanov, P.O., Zadoianchuk, N.V. (2012) Average-cost Markov decision processes with weakly continuous transition probabilities. *Math. Oper. Res.* 37(4): 591–607.
- [13] Feinberg, E.A., Kasyanov, P.O., Zadoianchuk, N.V. (2013) Berge’s theorem for noncompact image sets, *J. Math. Anal. Appl.* 397(1): 255–259.
- [14] Feinberg, E.A., P.O. Kasyanov, P.O., Zadoianchuk, N.V. (2014) *Fatou’s lemma for weakly converging probabilities*, *Theory Probab. Appl.*, 58(4), 683–689.
- [15] Feinberg, E.A., Kasyanov, P.O., Zgurovsky, M.Z. (2014) Convergence of probability measures and Markov decision models with incomplete information, *Proceedings of the Steklov Institute of Mathematics*, 287 (1), 96–117.
- [16] Feinberg, E.A., Kasyanov, P.O., Zgurovsky, M.Z. (2016) Partially observable total-cost Markov decision processes with weakly continuous transition probabilities, *Math. Oper. Res.*, 41(2), 656–681.
- [17] Feinberg, E.A., Kasyanov, P.O., Zgurovsky, M.Z. (2016) Uniform Fatou’s lemma, *Journal of Mathematical Analysis and Applications*, 444(1), 550–567.
- [18] Feinberg, E.A., Kasyanov, P.O., Zgurovsky, M.Z. (2021) Semi-Uniform Feller Stochastic Kernels. arXiv:2107.02207

- [19] Feinberg, E.A., Kasyanov, P.O., Zgurovsky, M.Z. (2021) Markov decision processes with incomplete information and semi-uniform Feller transition probabilities. arXiv:2108.09232
- [20] Hernández-Lerma, O. (1989) *Adaptive Markov Control Processes*, Springer-Verlag, New York.
- [21] Hernández-Lerma, O., Lasserre, J.B. (1996) *Discrete-Time Markov Control Processes: Basic Optimality Criteria*, Springer, New York.
- [22] Kara, A.D., Saldi, N., Yüksel, S. (2019) Weak Feller property of non-linear filters. *Systems & Control Letters* 134: 104512.
- [23] Papadimitriou, C.H., Tsitsiklis, J.N. (1987) The complexity of Markov decision processes. *Math. Oper. Res.*, 12(3), 441–681.
- [24] Papanicolaou, G.C. (1978) Asymptotic analysis of stochastic equations. Rosenblatt M, ed. *Studies in Probability Theory* Mathematical Association of America, Washington DC, 111–179.
- [25] Parthasarathy, K.R. (1967) *Probability Measures on Metric Spaces*, Academic Press, New York.
- [26] Platzman, L.K. (1980) Optimal infinite-horizon undiscounted control of finite probabilistic systems. *SIAM Journal on Control and Optimization* 18(4): 362-380.
- [27] Rhenius, D. (1974) Incomplete information in Markovian decision models. *Ann. Statist.* 2(6): 1327-1334.
- [28] Rieder, U. (1975) Bayesian dynamic programming. *Adv. Appl. Probab.* 7(2): 330-348.
- [29] Rudin, W. (1964) *Principles of Mathematical Analysis*, Second edition, McGraw-Hill Inc,
- [30] Shiryaev, A.N. (1967) Some new results in the theory of controlled random processes. *Transactions of the Fourth Prague Conference on Information Theory, Statistical Decision Functions, Random Processes* (Prague, 1965), pp. 131-201 (in Russian); Engl. transl. in *Select. Transl. Math. Statist. Probab.* 8(1969), 49-130.
- [31] Shiryaev AN (1996) *Probability*, Second edition, Springer-Verlag, New York.
- [32] Yushkevich AA (1976) 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(1): 153-158.