Weak Feller Property of Non-linear Filters

Ali Devran Kara, Naci Saldi and Serdar Yüksel

January 10, 2019

Abstract

Weak Feller property of controlled and control-free Markov chains lead to many desirable properties. In control-free setups this leads to the existence of invariant probability measures for compact spaces and applicability of numerical approximation methods. For controlled setups, this leads to existence and approximation results for optimal control policies. We know from stochastic control theory that partially observed systems can be converted to fully observed systems by replacing the original state space with a probability measure-valued state space, with the corresponding kernel acting on probability measures known as the non-linear filter (belief) process. Establishing sufficient conditions for the weak Feller property for such processes is a significant problem, studied under various assumptions and setups in the literature. In this paper, we prove the weak Feller property of the non-linear filter process (i) first under weak continuity of the transition probability of controlled Markov chain and total variation continuity of its observation channel, and then, (ii) under total variation continuity of the transition probability of controlled Markov chain. The former result (i) has first appeared in Feinberg et. al. [Math. Oper. Res. 41(2) (2016) 656-681]. Here, we present a concise and easy to follow alternative proof for this existing result. The latter result (ii) establishes weak Feller property of non-linear filter process under conditions, which have not been previously reported in the literature.

∗This research was supported in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada.
†Ali Devran Kara and Serdar Yüksel are with the Department of Mathematics and Statistics, Queen’s University, Kingston, ON, Canada, Emails: {16adk,yuksel}@mast.queensu.ca. Naci Saldi is with the Department of Natural and Mathematical Sciences, Ozyegin University, Cekmekoy, Istanbul, Turkey, Email: naci.saldi@ozyegin.edu.tr
1 Introduction

1.1 Preliminaries

We start with the probabilistic setup of the partially observed Markov decision processes (POMDP). Let $X \subset \mathbb{R}^n$ be a Borel set in which a controlled Markov process $\{X_t, t \in \mathbb{Z}_+\}$ takes its value. Here and throughout the paper, $\mathbb{Z}_+$ denotes the set of non-negative integers and $\mathbb{N}$ denotes the set of positive integers. Let $U$, the action space, be a Borel subset of some Euclidean space $\mathbb{R}^p$. Let $Y \subset \mathbb{R}^m$ be a Borel set, and let an observation channel $Q$ be defined as a stochastic kernel (regular conditional probability) from $X \times U$ to $Y$ such that $Q(\cdot | x, u)$ is a probability measure on the Borel $\sigma$-algebra $B(Y)$ for every $(x, u) \in X \times U$ and $Q(A \cdot) : X \times U \to [0, 1]$ is a Borel measurable function for every $A \in B(Y)$. Let a decision maker (DM) be located at the output of an observation channel $Q$, with inputs $(X_t, U_{t-1})$ and outputs $Y_t$. An admissible policy $\gamma$ is a sequence of control functions $\{\gamma_t, t \in \mathbb{Z}_+\}$ such that $\gamma_t$ is measurable with respect to the $\sigma$-algebra generated by the information variables

$$I_t = \{Y_{[0,t]}, U_{[0,t-1]}\}, \quad t \in \mathbb{N}, \quad I_0 = \{Y_0\},$$

where

$$U_t = \gamma_t(I_t), \quad t \in \mathbb{Z}_+$$

are the $U$-valued control actions and

$$Y_{[0,t]} = \{Y_s, 0 \leq s \leq t\}, \quad U_{[0,t-1]} = \{U_s, 0 \leq s \leq t - 1\}.$$

We define $\Gamma$ to be the set of all such admissible policies.

The joint distribution of the state, control, and observation processes is determined by (1) and the following system dynamics:

$$\Pr((X_0, Y_0) \in B) = \int_B Q_0(dy_0|x_0)P_0(dx_0), B \in B(\mathbb{X} \times \mathbb{Y}),$$

where $P_0$ is the prior distribution of the initial state $X_0$ and $Q_0$ is the prior control-free observation channel, and for $t \in \mathbb{N}$

$$\Pr\left((X_t, Y_t) \in B \bigg| (X, Y, U)_{[0,t-1]} = (x, y, u)_{[0,t-1]}\right)$$

$$= \int_B Q(dy_t|x_t, u_{t-1})T(dx_t|x_{t-1}, u_{t-1}), B \in B(\mathbb{X} \times \mathbb{Y}),$$

where $T$ is the transition kernel associated with the Markov process $\{X_t, t \in \mathbb{Z}_+\}$.
where $T(\cdot | x, u)$ is a stochastic kernel from $X \times U$ to $X$. This completes the probabilistic setup of the POMDPs.

Often, one is faced with an optimal control problem (or an optimal decision-making problem when control is absent in the transition dynamics). For the sake of completeness, we present typical criteria in the following. Let a one-stage cost function $c : X \times U \to [0, \infty)$, which is a Borel measurable function from $X \times U$ to $[0, \infty)$, be given. According to the Ionescu Tulcea theorem $[1]$, a policy $\gamma$ defines a unique probability measure $P^\gamma$ on $(X \times Y \times U)^\infty$. The expectation with respect to $P^\gamma$ is denoted by $E^\gamma$. Then, we denote by $W(\gamma)$ the cost function of the policy $\gamma \in \Gamma$, which can be finite-horizon total cost or infinite-horizon discounted cost or average cost criteria, which are given respectively as follows:

$$W(\gamma) := E^\gamma \left[ \sum_{t=0}^{T-1} c(X_t, U_t) \right]$$

$$W(\gamma) := E^\gamma \left[ \sum_{t=0}^{\infty} \beta^t c(X_t, U_t) \right], \beta \in (0, 1)$$

$$W(\gamma) := \lim \sup_{T \to \infty} \frac{1}{T} E^\gamma \left[ \sum_{t=0}^{T-1} c(X_t, U_t) \right].$$

With these definitions, the goal of the control problem is to find a policy $\gamma^*$ satisfying the following:

$$W(\gamma^*) = \inf_{\gamma \in \Gamma} W(\gamma).$$

### 1.2 POMDPs and Reduction to Belief-MDP

It is known that any POMDP can be reduced to a completely observable Markov decision process (MDP) $[2], [3]$, whose states are the posterior state distributions or 'beliefs' of the observer; that is, the state at time $t$ is

$$Z_t(\cdot) := \Pr\{X_t \in \cdot | Y_0, \ldots, Y_t, U_0, \ldots, U_{t-1}\} \in \mathcal{P}(X).$$

We call this equivalent MDP the belief-MDP. The belief-MDP has state space $Z = \mathcal{P}(X)$ and action space $U$. Note that $Z$ is equipped with the Borel $\sigma$-algebra generated by the topology of weak convergence $[4]$. Under this topology, $Z$ is a Borel space $[5]$. The transition probability of the belief-MDP can be constructed as follows.
The joint conditional probability on next state and observation variables given the current control action and the current state of the belief-MDP is given by

\[
R(B \times C|u_0, z_0) = \int_X \int_B Q(C|x_1, u_0)T(dx_1|x_0, u_0)z_0(dx_0), \quad (2)
\]

for all \( B \in \mathcal{B}(X) \) and \( C \in \mathcal{B}(Y) \). Then, the conditional distribution of the next observation variable given the current state of the belief-MDP and the current control action is given by

\[
P(C|u_0, z_0) = \int_X \int_X Q(C|x_1, u_0)T(dx_1|x_0, u_0)z_0(dx_0),
\]

for all \( C \in \mathcal{B}(Y) \). Using this, we can disintegrate \( R \) (see [6, Proposition 7.27]) as follows:

\[
R(B \times C|u_0, z_0) = \int_C F(B|y_1, u_0, z_0)P(dy_1|u_0, z_0) = \int_C z_1(y_1, u_0, z_0)(B)P(dy_1|u_0, z_0), \quad (3)
\]

where \( F \) is a stochastic kernel from \( Z \times Y \times U \) to \( X \) and the posterior distribution of \( x_1 \), determined by the kernel \( F \), is the state variable \( z_1 \) of the belief-MDP. Then, the transition probability \( \eta \) of the belief-MDP can be constructed as follows (see also [7]). If we define the measurable function \( F(z, u, y) := F(\cdot|y, u, z) = \Pr\{X_{t+1} \in \cdot | Z_t = z, U_t = u, Y_{t+1} = y\} \) from \( Z \times U \times Y \) to \( Z \) and use the stochastic kernel \( P(\cdot|z, u) = \Pr\{Y_{t+1} \in \cdot | Z_t = z, U_t = u\} \) from \( Z \times U \) to \( Y \), we can write \( \eta \) as

\[
\eta(\cdot|z, u) = \int_Y 1_{\{F(z,u,y)\in \cdot\}} P(dy|z, u); \quad (4)
\]

that is, \( \eta(\cdot|z, u) \) is an uncountable mixture of the probability measures \( \{F(\cdot|y, u, z)\}_{y \in Y} \).

The one-stage cost function \( \tilde{c} : Z \times U \to [0, \infty) \) of the belief-MDP is given by

\[
\tilde{c}(z, u) := \int_X c(x,u)z(dx),
\]

which is a Borel measurable function. Hence, the belief-MDP is a completely observable Markov decision process with the components \((Z, U, \tilde{c}, \eta)\)
For the belief-MDP, the information variables is defined as

\[ \tilde{I}_t = \{Z_{[0,t]}, U_{[0,t-1]}\}, \quad t \in \mathbb{N}, \quad \tilde{I}_0 = \{Z_0\}. \]

Let \( \tilde{\Gamma} \) denote the set of all policies for the belief-MDP, where the policies are defined in an usual manner. It is a standard result that an optimal control policy of the original POMDP will use the belief \( Z_t \) as a sufficient statistic for optimal policies (see [2], [3]). More precisely, the belief-MDP is equivalent to the original POMDP in the sense that for any optimal policy for the belief-MDP, one can construct a policy for the original POMDP which is optimal.

The following observations are crucial in obtaining the equivalence of these two models: (i) any information vector of the belief-MDP is a function of the information vector of the original POMDP and (ii) information vectors of the belief-MDP is a sufficient statistic for the original POMDP.

### 1.3 Problem Formulation

We first give definitions of two convergence notions for probability measures, and also, we define the weak Feller property of Markov decision processes. Let \((S,d)\) be a separable metric space. A sequence \( \{\mu_n, n \in \mathbb{N}\} \) in the set of probability measures \( \mathcal{P}(S) \) is said to converge to \( \mu \in \mathcal{P}(S) \) weakly if

\[ \int_S f(x)\mu_n(dx) \to \int_S f(x)\mu(dx) \]

for every continuous and bounded \( f : S \to \mathbb{R} \). One thing to note here is that the topology of weak convergence on the set of probability measures on a separable metric space is metrizable. One such metric is the bounded-Lipschitz metric. For any two probability measures \( \mu \) and \( \nu \), the bounded-Lipschitz metric is defined as:

\[ \rho_{BL}(\mu, \nu) = \sup_{\|f\|_{BL} \leq 1} \left| \int_S f(x)\mu(dx) - \int_S f(x)\nu(dx) \right| \]

where \( \|f\|_{BL} = \|f\|_\infty + \sup_{x \neq y} \frac{|f(x) - f(y)|}{d(x,y)} \) and \( \|f\|_\infty = \sup_{x \in S} |f(x)| \). Another metric that metrizes the weak topology on \( \mathcal{P}(S) \) is the following:

\[ \rho(\mu, \nu) = \sum_{m=1}^{\infty} 2^{-m+1} \left| \int_S f_m(x)\mu(dx) - \int_S f_m(x)\nu(dx) \right|, \]

where \( \{f_m\}_{m \geq 1} \) is an appropriate sequence of continuous and bounded functions such that \( \|f_m\|_\infty \leq 1 \) for all \( m \geq 1 \) (see [3] Theorem 6.6, p. 47)).
For probability measures $\mu, \nu \in \mathcal{P}(\mathcal{S})$, the total variation norm is given by
\[
\|\mu - \nu\|_{TV} = 2 \sup_{B \in \mathcal{B}(\mathcal{S})} |\mu(B) - \nu(B)|
\]
where the supremum is taken over all Borel measurable real $f$ such that $\|f\|_\infty \leq 1$. A sequence $\{\mu_n, n \in \mathbb{N}\}$ in $\mathcal{P}(\mathcal{S})$ is said to converge in total variation to $\mu \in \mathcal{P}(\mathcal{S})$ if $\|\mu_n - \mu\|_{TV} \to 0$.

We say that a Markov decision process with transition kernel $T(\cdot | x, u)$ has weak Feller property if $T$ is weakly continuous in $x$ and $u$; that is, if $(x_n, u_n) \to (x, u)$, then $T(\cdot | x_n, u_n) \to T(\cdot | x, u)$ weakly.

In this paper, we study the following problem.

(P) Under what conditions on the transition probability and the observation channel of the POMDP, the transition probability of the belief-MDP is weakly continuous; that is, for any $f \in C_b(\mathcal{P}(X))$ and for any $(z_n, u_n) \to (z, u)$, we have
\[
\int f(z_1) \eta(dz_1 | z_n, u_n) \to \int f(z_1) \eta(dz_1 | z, u).
\]

1.4 Practical and Mathematical Significance of the Problem

In this section, we motivate the operational (in view of engineering applications) and the mathematical importance of the problem in view of existence and invariance properties, approximations and computational results involving non-linear filters and stochastic control, and related applications involving particle filtering.

For finite-horizon problems and a large class of infinite-horizon discounted cost problems, it is a standard result that an optimal control policy will use the belief-process as a sufficient statistic for optimal policies (see [2] [3] [8]). Hence, the POMDP and the corresponding belief-MDP are equivalent in the sense of cost minimization. Therefore, results developed for the standard MDP problems (e.g., measurable selection criteria as summarized in [1] Chapter 3) can be applied to the belief-MDPs, and so, to the POMDPs. In MDP theory, weak continuity of the transition kernel is an important condition leading to both the existence of optimal control policies for finite-horizon and infinite-horizon discounted cost problems as well as the continuity properties of the value function (see, e.g., [9] Section 8.5).
For partially observed stochastic control problems with the average cost criterion, the conditions of existence of optimal policies stated in the literature are somewhat restrictive, with the most relaxed conditions to date being reported in [10] [11], to our knowledge. For such average cost stochastic control problems, the weak Feller property can be utilized to obtain a direct method to establish the existence of optimal stationary (and possibly randomized) control policies. Indeed, for such problems, the convex analytic method ([12] and [13]) is a powerful approach to establish the existence of optimal policies. If one can establish the weak Feller property of the belief-MDP, then the continuity and compactness conditions utilized in the convex program of [13] would lead to the existence of optimal average cost policies.

In addition to existence of optimal policies, the weak Feller property has also recently been shown to lead to (asymptotic) consistency in approximation methods for MDPs with uncountable state and action spaces. In [14] [15], authors showed that optimal policies obtained from finite-model approximations to infinite-horizon discounted cost MDPs with Borel state and action spaces asymptotically achieve the optimal cost for the original problem under weak Feller property. Hence, the weak Feller property of belief process suggests that approximation results in [14] [15], which only require weak continuity conditions on the transition probability of a given MDP, are particularly suitable in developing approximation methods for POMDPs (through their MDP reduction).

For control-free setups, the weak Feller property of the belief-processes leads to the existence of invariant probability measures when the hidden Markov processes take their values in compact spaces or more general spaces under appropriate tightness conditions [16] [17].

For the empirical consistency and convergence results involving the very popular particle filtering algorithms, weak Feller property of the belief-process is a commonly imposed, implicit, assumption (see e.g. [18] [19]). Finally, for the study of ergodicity and asymptotic stability of nonlinear filters, weak Feller property also plays an important role (see [20] [21] [17]).

2 Main Results and Connections with the Literature

2.1 Statement of Main Results

In this paper, we show the weak Feller property of the belief-MDP under two different set of assumptions.
**Assumption 1.**

(i) The transition probability $T(\cdot|\cdot, u)$ is weakly continuous in $(x, u)$, i.e., for any $(x_n, u_n) \to (x, u)$, $T(\cdot|x_n, u_n) \to T(\cdot|x, u)$ weakly.

(ii) The observation channel $Q(\cdot|\cdot, u)$ is continuous in total variation, i.e., for any $(x_n, u_n) \to (x, u)$, $Q(\cdot|x_n, u_n) \to Q(\cdot|x, u)$ in total variation.

**Assumption 2.**

(i) The transition probability $T(\cdot|x, u)$ is continuous in total variation in $(x, u)$, i.e., for any $(x_n, u_n) \to (x, u)$, $T(\cdot|x_n, u_n) \to T(\cdot|x, u)$ in total variation.

(ii) The observation channel $Q(\cdot|x)$ is independent of the control variable.

We now formally state the main results of our paper.

**Theorem 1** (Feinberg et. al. [26]). Under Assumption 1, the transition probability $\eta(\cdot|z, u)$ of the belief-MDP is weakly continuous in $(z, u)$.

**Theorem 2.** Under Assumption 2, the transition probability $\eta(\cdot|z, u)$ of the belief-MDP is weakly continuous in $(z, u)$.

Theorem 1 is originally due to Feinberg et. al. [26]. The contribution of our paper is that the proof presented here is more direct and significantly more concise. Theorem 2 establishes that under Assumption 2 (with no assumptions on the measurement model), the belief-MDP is weakly continuous. This result has not been previously reported in the literature.

The proofs of these results are presented in Section 3.

### 2.2 Prior Results on Weak Feller Property of the Belief-Process

Weak Feller property of the control-free transition probability of the belief-MDPs has been established using different approaches and different conditions. In [21] it has been shown that, for continuous-time belief-processes, if the signal process (state process of the POMDP) is weak Feller and the measurement channel is an additive channel in the form $Y_t = \int_0^t h(X_u)du + V_t$, where $h$ is assumed to be continuous and possibly unbounded and $V_t$ is a standard Wiener process, then the belief-process itself is also weak Feller. In [17], the authors study the discrete-time belief-processes, where the state process noise may not be independent of the observation process noise; it has been shown that if the observation model is additive in the form $Y_t = h(X_t) + V_t$, where $h$ is assumed to be continuous and $V_t$ is an i.i.d. noise process which
admits a continuous and bounded density function, then the observation
and belief processes \((Y_t, Z_t)\) are jointly weak Feller. In [20], the weak Feller
property of the belief-process has been shown for both discrete and continuous
time setups when the channel is additive, \(Y_t = h(X_t) + V_t\), where \(h\) is
bounded and continuous and \(V_t\) is an i.i.d. noise process with a continuous,
bounded and positive density function.

Weak Feller property of the control-free belief-processes has also been
proven in some particle filtering consistency and convergence studies. [18,
19] have studied the consistency of the particle filter methods where the weak
Feller property of the belief-process has been used to establish the convergence results. In [18], it has been shown that the belief-process is weak
Feller under the assumption that the transition probability of the partially
observed system is weak Feller and the measurement channel is an additive
channel in the form \(Y_t = h(X_t) + V_t\), where \(h\) is assumed to be continuous and
\(V_t\) is an i.i.d. noise process, which admits a continuous and bounded density
function with respect to the Lebesgue measure. In [19], the weak Feller
property of the belief-process has been established under the assumption
that the measurement channel admits a continuous and bounded density
function with respect to the Lebesgue measure; i.e., the channel can be
written in the following form: \(Q(y \in A|x) = \int_A g(x, y)\mu(dy)\) for any \(A \in \mathcal{B}(\mathbb{Y})\)
and for any \(x \in \mathbb{X}\), where \(g\) is a continuous and bounded function.

Weak Feller property of the controlled transition probability of the belief-
MDPs has been established, in the most general case to date, by Feinberg
et.al. [26]. Under the assumption that the measurement channel is continuous in total variation and the transition kernel of the POMDP is weak Feller, the authors have established the weak Feller property of the transition probability of the belief-process. In Section 4 we will give a detailed discussion
on the methods used by Feinberg et.al. [26], and also, we will compare their
approach with ours.

2.3 A Comparative Discussion

As reviewed above, the prior literature had imposed quite strong structural
conditions on the measurement models, while imposing weak continuity of
the transition probability of the hidden Markov process. Note that when the
observation channel is additive \(Y_t = h(X_t, U_{t-1}) + V_t\), where \(h\) is continuous
and \(V_t\) admits a continuous density function with respect to some measure \(\mu\),
one can show that the channel also admits a continuous density function, i.e.,
\(Q(y \in A|x, u) = \int_A g(x, u, y)\mu(dy)\) for any \(A \in \mathcal{B}(\mathbb{Y})\) and for any \((x, u) \in \mathbb{X} \times \mathbb{U}\). When the observation channel has a continuous density function,
the pointwise convergence of the density functions implies the total variation convergence by Scheffé’s Lemma [22, Theorem 16.12]. Thus, \( g(x_k, u_k, y) \to g(x, u, k) \) for some \((x_k, u_k) \to (x, u)\) implies that \( Q(\cdot|x_k, u_k) \to Q(\cdot|x, u) \) in total variation, i.e., the observation channel is continuous in total variation. As noted in [26], the structural condition \( Y_t = h(X_t, U_{t-1}) + V_t \) can be relaxed; it suffices that the measurement channel is continuous under total variation.

In the following, we develop a relationship between the total variation continuity of the channel (as required by [26] and in our Theorem 1) and the more restrictive density conditions on the measurement channels presented in the prior works [21, 17, 19, 18].

In the proposition below, we show that having a continuous density is almost equivalent to the condition that the observation channel is continuous in total variation. Although, we will not use this proposition in the proofs of the main results, it is important on its own as it states that if the observation channel is continuous with respect to the total variation norm, then it admits a density with respect to some reference measure, which satisfies some regularity conditions, thereby placing the results in the literature in a unified context. The proof can be found in Appendix A.

**Proposition 1.** Suppose that the observation channel \( Q(dy|x, u) \) is continuous in total variation. Then, for any \((z, u) \in Z\), we have, \( T(\cdot|z, u)\)-a.s., that

\[
Q(dy|x, u) \ll P(dy|z, u)
\]

and

\[
Q(dy|x, u) = g(x, u, y)P(dy|z, u)
\]

for a measurable function \( g \), which satisfies for any \( A \in B(Y) \) and for any \( x_k \to x \)

\[
\int_A |g(x_k, u, y) - g(x, u, y)|P(dy|z, u) \to 0.
\]

We again emphasize that weak Feller property of the belief-process under Assumption 1 has been first established by [26] using different method compared to ours. Our method is significantly more concise and direct. It is also important to note that Assumption 2 completely eliminates any restriction on the observation channel to establish the weak Feller property of belief-MDP. This relaxation is quite important in practice since modelling the noise on the observation channel in control problems is quite cumbersome, and in general, infeasible. But, in many problems that arise in practice, the transition probability has a continuous density with respect to some reference measure, which directly implies, via Scheffé’s Lemma, the total variation...
continuity of the transition probability. We also note that the weak Feller property under only Assumption 2-(i) cannot be established. Indeed, Example 4.1 of [26] shows that the total variation continuity assumption on the observation channel cannot be relaxed even when the transition kernel is continuous in total variation to prove weak Feller property of the belief-process under controlled observation channel model. A careful look at the counterexample shows that it indeed uses the discontinuity of the observation channel in the control action to prove that the belief-process cannot be weak Feller when the observation channel is not continuous in total variation and the transition kernel is continuous in total variation.

2.4 Examples

In this section, we give concrete examples for the system and observation channel models which satisfy Assumption 1 or Assumption 2. Suppose that the system dynamics and the observation channel are represented as follows:

\[
X_{t+1} = H(X_t, U_t, W_t)
\]
\[
Y_t = G(X_t, U_{t-1}, V_t),
\]

where \(W_t\) and \(V_t\) are i.i.d. noise processes. This is, without loss of generality, always the case; that is, you can transform the dynamics of any POMDP into this form.

- Suppose that \(H(x, u, w)\) is a continuous function in \(x\) and \(u\). Then, the corresponding transition kernel is weakly continuous. To see this, observe that, for any \(c \in C_b(\mathbb{X})\), we have

\[
\int c(x_1)T(dx_1|x_0^n, u_0^n) = \int c(H(x_0^n, u_0^n, w_0))\mu(dw_0)
\]
\[
\rightarrow \int c(H(x_0, u_0, w_0))\mu(dw_0) = \int c(x_1)T(dx_1|x_0, u_0),
\]

where we use \(\mu\) to denote the probability model of the noise.

- Suppose that \(G(x, u, v) = g(x, u) + v\), where \(g\) is a continuous function and \(V_t\) admits a continuous density function \(\varphi\) with respect to some reference measure \(\nu\). Then, the channel is continuous in total variation. Notice that under this setup, we can write \(Q(dy|x, u) = \varphi(y - h(x, u))\nu(dy)\). Hence, the density of \(Q(dy|x_n, u_n)\) converges to the density of \(Q(dy|x, u)\) pointwise, and so, \(Q(dy|x_n, u_n)\) converges to \(Q(dy|x, u)\) in total variation by Scheffé’s Lemma [22, Theorem 16.12].
Hence, $Q(dy|x,u)$ is continuous in total variation under these conditions.

- Suppose that we have $H(x,u,w) = h(x,u) + w$, where $f$ is continuous and $W_t$ admits a continuous density function $\varphi$ with respect to some reference measure $\nu$. Then, the transition probability is continuous in total variation. Again, notice that with this setup we have $T(dx_1|x_0,u_0) = \varphi(x_1 - h(x_0,u_0))\nu(dx_1)$. Thus, continuity of $\varphi$ and $h$ guarantees the pointwise convergence of the densities, so we can conclude that the transition probability is continuous in total variation by again Scheffé’s Lemma.

The analysis in the paper will provide weak Feller results for a large class of partially observed control systems as reviewed in the aforementioned examples. In particular, if the state dynamics are affected by an additive noise process which admits a continuous density, we can guarantee weak Feller property of the belief-process by means of Theorem 2 without referring to the noise model of the observation channel.

3 Proofs

The following result will play a key role for the proofs of main results. The result mainly builds on ([23, Lemma A.2]), however, we give a separate proof in the appendix for completeness.

**Lemma 1.** Let $X$ be a Borel space. Suppose that we have a family of uniformly bounded real Borel measurable functions $\{f_{n,\lambda}\}_{n \geq 1, \lambda \in \Lambda}$ and $\{f_{\lambda}\}_{\lambda \in \Lambda}$, for some set $\Lambda$. If, for any $x_n \to x$ in $X$, we have

$$\lim_{n \to \infty} \sup_{\lambda \in \Lambda} |f_{n,\lambda}(x_n) - f_{\lambda}(x)| = 0$$

and

$$\lim_{n \to \infty} \sup_{\lambda \in \Lambda} |f_{\lambda}(x_n) - f_{\lambda}(x)| = 0,$$

then, for any $\mu_n \to \mu$ weakly in $\mathcal{P}(X)$, we have

$$\lim_{n \to \infty} \sup_{\lambda \in \Lambda} \left| \int_X f_{n,\lambda}(x)\mu_n(dx) - \int_X f_{\lambda}(x)\mu(dx) \right| = 0.$$
3.1 Proof of Theorem 1

We show that, for every \((z^n_0, u_n) \to (z_0, u)\) in \(Z \times U\), we have

\[
\sup_{\|f\|_{BL} \leq 1} \left| \int_Z f(z_1) \eta(dz_1 | z^n_0, u_n) - \int_Z f(z_1) \eta(dz_1 | z_0, u) \right| \to 0,
\]

where we equip \(Z\) with the metric \(\rho\) to define bounded-Lipschitz norm \(\|f\|_{BL}\) of any Borel measurable function \(f : Z \to \mathbb{R}\). We can equivalently write this as

\[
\sup_{\|f\|_{BL} \leq 1} \left| \int_Y f(z_1(z^n_0, u_n, y_1)) P(dy_1 | z^n_0, u_n) \right.
\]

\[
- \int_Y f(z_1(z_0, u, y_1)) P(dy_1 | z_0, u) \right| \to 0. \tag{7}
\]

Our first claim is that \(P(dy_1 | z_0, u)\) is continuous in total variation. To see this, let \((z^n_0, u_n) \to (z_0, u)\). Then, we write

\[
\sup_{A \in B(Y)} \left| P(A | z^n_0, u_n) - P(A | z_0, u) \right|
\]

\[
= \sup_{A \in B(Y)} \left| \int_X Q(A | x_1, u_n) \mathcal{T}(dx_1 | z^n_0, u_n) \right.
\]

\[
- \int_X Q(A | x_1, u) \mathcal{T}(dx_1 | z_0, u) \right|.
\]

where \(\mathcal{T}(dx_1 | z^n_0, u_n) := \int_X \mathcal{T}(dx_1 | x_0, u_n) z^n_0(dx_0)\). Note that, by Lemma 1 we can show that \(\mathcal{T}(dx_1 | z^n_0, u_n) \to \mathcal{T}(dx_1 | z_0, u)\) weakly. Indeed, if \(g \in C_b(X)\), then we define \(r_n(x_0) = \int_X g(x_1) \mathcal{T}(dx_1 | x_0, u_n)\) and \(r(x_0) = \int_X g(x_1) \mathcal{T}(dx_1 | x_0, u)\). Since \(\mathcal{T}(dx_1 | x_0, u)\) is weakly continuous, we have \(r_n(x^n_0) \to r(x_0)\) when \(x^n_0 \to x_0\). Hence, by Lemma 1 we have

\[
\lim_{n \to \infty} \left| \int_X r_n(x_0) z^n_0(dx_0) - \int_X r(x_0) z_0(dx_0) \right| = 0.
\]

Hence, \(\mathcal{T}(dx_1 | z^n_0, u_n) \to \mathcal{T}(dx_1 | z_0, u)\) weakly. Moreover, the families of functions \(\{Q(A | \cdot, u_n)\}_{n \geq 1, A \in B(Y)}\) and \(\{Q(A | \cdot, u)\}_{A \in B(Y)}\) satisfy the conditions of Lemma 1 as \(Q\) is continuous in total variation distance. Therefore, Lemma 1 yields that

\[
\lim_{n \to \infty} \sup_{A \in B(Y)} \left| \int_X Q(A | x_1, u_n) \mathcal{T}(dx_1 | z^n_0, u_n) \right|
\]

\[
= \lim_{n \to \infty} \sup_{A \in B(Y)} \left| \int_X Q(A | x_1, u) \mathcal{T}(dx_1 | z_0, u) \right| = 0.
\]

Therefore, \(\mathcal{T}(dx_1 | z^n_0, u_n) \to \mathcal{T}(dx_1 | z_0, u)\) weakly.
Thus, $P(dy_1|z_0, u)$ is continuous in total variation.

Now, we go back to (7). We have

$$-\int X Q(A|x_1, u)T(dx_1|z_0, u) = 0.$$
where we have used Fubini’s theorem with the fact that \( \sup_{m} \|f_m\|_{\infty} \leq 1 \). For each \( m \), let us define

\[
I_{+}^{(n)} := \left\{ y_1 \in \mathbb{Y} : \int_{X} f_m(x_1)z_1(z_0^n, u_n, y_1)(dx_1) > \int_{X} f_m(x_1)z_1(z_0, u, y_1)(dx_1) \right\}
\]

\[
I_{-}^{(n)} := \left\{ y_1 \in \mathbb{Y} : \int_{X} f_m(x_1)z_1(z_0^n, u_n, y_1)(dx_1) \leq \int_{X} f_m(x_1)z_1(z_0, u, y_1)(dx_1) \right\}.
\]

(9)

Then, we can write

\[
\int_{\mathbb{Y}} \left| \int_{X} f_m(x_1)z_1(z_0^n, u_n, y_1)(dx_1) \right. \left. - \int_{X} f_m(x_1)z_1(z_0, u, y_1)(dx_1) \right| P(dy_1|z_0, u)
\]

\[
= \int_{I_{+}^{(n)}} \left( \int_{X} f_m(x_1)z_1(z_0^n, u_n, y_1)(dx_1) \right. \left. - \int_{X} f_m(x_1)z_1(z_0, u, y_1)(dx_1) \right) P(dy_1|z_0, u)
\]

\[
+ \int_{I_{-}^{(n)}} \left( \int_{X} f_m(x_1)z_1(z_0^n, u_n, y_1)(dx_1) \right. \left. - \int_{X} f_m(x_1)z_1(z_0^n, u_n, y_1)(dx_1) \right) P(dy_1|z_0, u).
\]

In the sequel, we only consider the term with the set \( I_{+}^{(n)} \). The analysis for the other one follows from the same steps. We have

\[
\int_{I_{+}^{(n)}} \left( \int_{X} f_m(x_1)z_1(z_0^n, u_n, y_1)(dx_1) \right. \left. - \int_{X} f_m(x_1)z_1(z_0^n, u_n, y_1)(dx_1) \right) P(dy_1|z_0, u)
\]

\[
\leq \int_{I_{+}^{(n)}} \int_{X} f_m(x_1)z_1(z_0^n, u_n, y_1)(dx_1) P(dy_1|z_0, u)
\]

\[
- \int_{I_{+}^{(n)}} \int_{X} f_m(x_1)z_1(z_0^n, u_n, y_1)(dx_1) P(dy_1|z_0^n, u_n)
\]
\begin{align*}
+ \int_{I_+^{(n)}} \int_X f_m(x_1)z_1(z_0^n, u, y_1)(dx_1)P(dy_1|z_0^n, u, y_1) \\
- \int_{I_+^{(n)}} \int_X f_m(x_1)z_1(z_0, u, y_1)(dx_1)P(dy_1|z_0, u) \\
\leq \|P(dy_1|z_0, u) - P(dy_1|z_0^n, u, y_1)\|_{TV} \\
+ \int_X \int_{I_+^{(n)}} f_m(x_1)Q(dy_1|x_1, u)\mathcal{T}(dx_1|z_0^n, u) \\
- \int_X \int_{I_+^{(n)}} f_m(x_1)Q(dy_1|x_1, u)\mathcal{T}(dx_1|z_0, u),
\end{align*}

where we have used \(\|f_m\|_{\infty} \leq 1\) in the last inequality. The first term above goes to 0 since \(P(dy_1|z_0, u)\) is continuous in total variation. For the second term, we use Lemma 1. Indeed, families of functions \(\{f_m(\cdot)Q(A|\cdot, u_n) : n \geq 1, A \in \mathcal{B}(Y)\}\) and \(\{f_m(\cdot)Q(A|\cdot, u) : A \in \mathcal{B}(Y)\}\) satisfy the conditions in Lemma 1 as \(Q\) is continuous in total variation. Hence, the second term converges to 0 by Lemma 1 since \(\mathcal{T}(dx_1|z_0^n, u) \to \mathcal{T}(dx_1|z_0, u)\) weakly. Hence, for each \(m\), we have

\[
\lim_{n \to \infty} \int_Y \left| \int_X f_m(x_1)z_1(z_0^n, u, y_1)(dx_1) - \int_X f_m(x_1)z_1(z_0, u, y_1)(dx_1) \right| P(dy_1|z_0, u) = 0.
\]

By dominated convergence theorem, we then have

\[
\lim_{n \to \infty} \sup_{\|f\|_{BL} \leq 1} \int_Y |f(z_1(z_0^n, u, y_1)) - f(z_1(z_0, u, y_1))| P(dy_1|z_0, u) \\
\leq \sum_{m=1}^{\infty} 2^{-m+1} \lim_{n \to \infty} \int_Y \left| f_m(x_1)z_1(z_0^n, u, y_1)(dx_1) - f_m(x_1)z_1(z_0, u, y_1)(dx_1) \right| P(dy_1|z_0, u) = 0.
\]

This implies that (8) converges to 0 as \(n \to 0\). This completes the proof.

### 3.2 Proof of Theorem 2

A careful look at the proof of Theorem 1 reveals that the weak Feller property of the belief-process follows from the following facts:
(i) \( P(dy_1|z_0, u_0) \) is continuous in total variation,

(ii) \( \lim_{n \to \infty} \int_Y \rho(z_1(z^n_0, u_n, y_1), z_1(z_0, u, y_1)) P(dy_1|z_0, u) = 0 \) as \((z^n_0, u_n) \to (z_0, u)\).

We first show that \( P(dy_1|z_0, u_0) \) is continuous total variation. Let \((z^n_0, u_n) \to (z_0, u)\). Then, we have

\[
\sup_{A \in \mathcal{B}(Y)} |P(A|z^n_0, u_n) - P(A|z_0, u)| = \sup_{A \in \mathcal{B}(Y)} \left| \int_X \int_X Q(A|x_1) \mathcal{T}(dx_1|x_0, u_n) \rho_0(dz_1) - \int_X \int_X Q(A|x_1) \mathcal{T}(dx_1|x_0, u) \rho_0(dz_1) \right|.
\]

For each \( A \in \mathcal{B}(Y) \) and \( n \geq 1 \), we define

\[
f_{n,A}(x_0) = \int_X Q(A|x_1) \mathcal{T}(dx_1|x_0, u_n)
\]

and

\[
f_A(x_0) = \int_X Q(A|x_1) \mathcal{T}(dx_1|x_0, u).
\]

Then, for all \( x^n_0 \to x_0 \), we have

\[
\lim_{n \to \infty} \sup_{A \in \mathcal{B}(Y)} |f_{n,A}(x^n_0) - f_A(x_0)|
\]

\[
= \lim_{n \to \infty} \sup_{A \in \mathcal{B}(Y)} \left| \int_X Q(A|x_1) \mathcal{T}(dx_1|x^n_0, u_n) - \int_X Q(A|x_1) \mathcal{T}(dx_1|x_0, u) \right|
\]

\[
\leq \lim_{n \to \infty} \| \mathcal{T}(dx_1|x^n_0, u_n) - \mathcal{T}(dx_1|x_0, u) \|_{TV} = 0
\]

and

\[
\lim_{n \to \infty} \sup_{A \in \mathcal{B}(Y)} |f_A(x^n_0) - f_A(x_0)|
\]

\[
= \lim_{n \to \infty} \sup_{A \in \mathcal{B}(Y)} \left| \int_X Q(A|x_1) \mathcal{T}(dx_1|x^n_0, u) - \int_X Q(A|x_1) \mathcal{T}(dx_1|x_0, u) \right|
\]

17
Indeed, we have

\[ \lim_{n \to \infty} \| T(dx_1|x_0^n, u) - T(dx_1|x_0, u) \|_{TV} = 0. \]

Then, by Lemma 1 we have

\[
\begin{align*}
\lim_{n \to \infty} \sup_{A \in \mathcal{B}(Y)} \left| \int_X f_n,A(x_0) z_0^n(d x_0) & - \int_X f_A(x_0) z_0(d x_0) \right| \\
= \lim_{n \to \infty} \sup_{A \in \mathcal{B}(Y)} \left| \int_X \int_X Q(A|x_1) T(dx_1|x_0,u_n) z_0^n(d x_0) \\
- \int_X \int_X Q(A|x_1) T(dx_1|x_0,u) z_0(d x_0) \right| \\
= 0.
\end{align*}
\]

Hence, \( P(dy_1|z_0,u_0) \) is continuous in total variation.

Now, we show that, for any \( (z_0^n,u_n) \to (z_0,u) \), we have

\[ \lim_{n \to \infty} \int_Y \rho(z_1(z_0^n,u_n,y_1),z_1(z_0,u,y_1)) P(dy_1|z_0,u) = 0. \]

From the proof of Theorem 2 it suffices to show that

\[
\begin{align*}
\lim_{n \to \infty} \int_{I_{+}^{(n)}} f_m(x_1) Q(dy_1|x_1) T(dx_1|z_0^n,u_n) \\
- \int_{I_{+}^{(n)}} f_m(x_1) Q(dy_1|x_1) T(dx_1|z_0,u) = 0. \quad (10)
\end{align*}
\]

Indeed, we have

\[
\begin{align*}
\left| \int_X \int_{I_{+}^{(n)}} f_m(x_1) Q(dy_1|x_1) T(dx_1|z_0^n,u_n) \\
- \int_X \int_{I_{+}^{(n)}} f_m(x_1) Q(dy_1|x_1) T(dx_1|z_0,u) \right| \\
\leq \left| \int_{X^2} f_m(x_1) Q(I_{+}^{(n)}|x_1) T(dx_1|x_0,u_n) z_0^n(d x_0) \\
- \int_{X^2} f_m(x_1) Q(I_{+}^{(n)}|x_1) T(dx_1|x_0,u) z_0^n(d x_0) \right| \\
+ \left| \int_{X^2} f_m(x_1) Q(I_{+}^{(n)}|x_1) T(dx_1|x_0,u) z_0^n(d x_0) \\
- \int_{X^2} f_m(x_1) Q(I_{+}^{(n)}|x_1) T(dx_1|x_0,u) z_0(d x_0) \right|
\end{align*}
\]
\[
\begin{align*}
&\leq \int_X \|T(dx_1|x_0, u_n) - T(dx_1|x_0, u)\|_{TV} z_0^n(dx_0) \\
&+ \left| \int_{X^2} f_m(x_1)Q(I_{+}^{(n)}|x_1)T(dx_1|x_0, u)z_0^n(dx_0) \\
&\quad - \int_{X^2} f_m(x_1)Q(I_{+}^{(n)}|x_1)T(dx_1|x_0, u)z_0(dx_0) \right|,
\end{align*}
\]

where we have used \(\sup_{n \geq 1} \sup_{x_1 \in X} |f_m(x_1)Q(I_{+}^{(n)}|x_1)| \leq 1\) in the last inequality. If we define \(r_n(x_0) = \|T(dx_1|x_0, u_n) - T(dx_1|x_0, u)\|_{TV}\), then \(r_n(x_0^n) \to 0\) whenever \(x_0^n \to x_0\). Then, the first term converges to 0 by Lemma 1 as \(z_0^n \to z_0\) weakly. The second term also converges to 0 by Lemma 1 since \(\{\int_X f(x_1)Q(I_{+}^{(n)}|x_1)T(dx_1|\cdot, u) : n \geq 1\}\) is a family of uniformly bounded and equicontinuous functions by total variation continuity of \(T(dx_1|x_0, u)\). This completes the proof.

4 An Extension and a Discussion

In this section, we prove the weak Feller property of the belief-process under more general condition than those in Theorem 1 and Theorem 2. But, we note that it is indeed generally infeasible to establish this condition without imposing assumptions similar to the assumptions in Theorem 1 and Theorem 2. Therefore, although this condition is more general than those in Theorem 1 and Theorem 2, this generalization is not excessively important in practice.

We first note that our proof technique brings to light the main ingredients that is necessary to prove the weak Feller property of the belief-process via the item (i) and eq. (10), in the proof of Theorem 2 that is,

- \(P(dy_1|z_0, u_0)\) is continuous in total variation,
- \(\lim_{n \to \infty} \int_X f_m(x_1)Q(I_{+}^{(n)}|x_1)T(dx_1|z_0^n, u_n) - \int_X f_m(x_1)Q(I_{+}^{(n)}|x_1)T(dx_1|z_0, u) = 0\).

This observation suggests the following condition that generalize the conditions in our previously stated main results. Let \(F = \{f_m\}_{m \geq 1} \subset C_0(X)\) be a countable set of continuous and bounded functions such that \(\|f_m\|_{\infty} \leq 1\) for all \(m \geq 1\), \(1_X \in F\), and \(F\) metrizes the weak topology on \(P(X)\) via the metric \(\rho\) introduced in Section 1.3. Then, we state the following assumption:
For each $f \in F$, the family of functions

$$
(z_0, u_0) \mapsto \int_X f(x)Q(A|x, u_0)\mathcal{T}(dx|z_0, u_0)
$$

is equicontinuous when indexed by $A \in B(Y)$.

Using Lemma 1, it is fairly straightforward to prove that conditions in Theorem 1 and Theorem 2 both imply the assumption (M). Hence, assumption (M) is more general than those in Theorem 1 and Theorem 2.

**Theorem 3.** Under assumption (M), the transition probability $\eta(\cdot|z, u)$ of the belief-MDP is weakly continuous in $(z, u)$.

**Proof.** By proof of Theorem 1, it is sufficient to prove the following:

(i) $P(dy_1|z_0, u_0)$ is continuous in total variation,

(ii) $\lim_{n \to \infty} \int_X \rho(z_1(z_0^n, u_n, y_1), z_1(z_0, u, y_1))P(dy_1|z_0, u) = 0$ as $(z_0^n, u_n) \to (z_0, u)$.

Firstly, (i) is true since $1_X \in F$. For (ii), it suffices to show that

$$
\lim_{n \to \infty} \int_X f_m(x)Q(I_x^n|z_0^n, u_n)\mathcal{T}(dx|z_0^n, u_n) - \int_X f_m(x)Q(I_x^n|z_0^n)\mathcal{T}(dx|z_0, u) = 0.
$$

But this immediately follows from assumption (M). \qed

A careful look at the proof of the weak Feller property of the belief-process in Feinberg et. al. reveals that they have first established the weak Feller property under a condition somewhat similar to the assumption (M), and then, establish Theorem 1 by proving that assumptions in Theorem 1 imply this more general condition. Indeed, let $\tau_b = \{O_j\} \subset X$ be a countable base for the topology on $X$ such that $X \in \tau_b$. Then, under the following assumption:

(F) For each finite intersection $O = \bigcap_{n=1}^N O_{j_n}$, where $O_{j_n} \in \tau_b$, the family of functions

$$
(z_0, u_0) \mapsto \int_X \int_X 1_O(x)Q(A|x, u_0)\mathcal{T}(dx|x_0, u_0)z_0(dx_0)
$$

is equicontinuous when indexed by $A \in B(Y)$,
they have proved that the weak Feller property of the belief-process holds (see [26, Lemma 5.3] and [25, Theorem 5.5]). We observe that (F) is very similar to (M) except that, in (F), Feinberg et. al. use open sets instead of continuous and bounded functions. However, proving that conditions in Theorem 1 imply the assumption (F) as in [26] requires quite tedious mathematical methods. By using open sets instead of continuous and bounded functions, one needs to work with inequalities and limit infimum operation as a result of Portmanteau theorem [4, Theorem 2.1] (and the associated proof program involving generalized Fatou’s lemma [24, 27]), in place of equalities and limit operation, which are significantly easier to analyze than the former leading to a much more concise analysis that we have presented in this paper. For instance, [26, Theorem 5.1] is the key result to prove that the weak Feller condition of the transition probability and the total variation continuity of the observation channel imply the assumption (F). We note that if one states this result using continuous and bounded functions in place of open sets, then this version of [26, Theorem 5.1] becomes a corollary of Lemma 1 which has a concise and easy to follow proof. But, the proof of [26, Theorem 5.1] with open sets requires quite tedious mathematical concepts from topology and weak convergence of probability measures. In view of this discussion, we also note that Theorem 2 can also be proved using the condition (F) rather than our approach building on (M) through some additional argumentation.

In summary, our approach allows for a more direct and concise approach which also makes the proof of Theorem 1 more accessible. Once again, we note that Theorem 2 has not been reported in the literature.

5 Conclusion

In this paper, there are two main contributions: (i) the weak Feller property of the belief-process is established under a new condition, which assumes that the state transition probability is continuous under the total variation convergence with no assumptions on the measurement model, and (ii) a concise and easy to follow proof of the same result under the weak Feller condition of the transition probability and the total variation continuity of the observation channel, which was first established in [26], is also given. Implications of these results have also been presented.
6 Acknowledgements

The authors are very grateful to Prof. Eugene Feinberg for generously providing technical feedback and pointing out to related results.

Appendix

A Proof of Proposition

Fix any \((z, u)\). We first show that \(Q(dy|x, u) \ll P(dy|z, u), T(\cdot|z, u)\)-a.s.

Note that \(Q(dy|x, u) \ll P(dy|z, u)\) if and only if, for all \(\varepsilon > 0\), there exists \(\delta > 0\) such that \(Q(A|x, u) < \varepsilon\) whenever \(P(A|z, u) < \delta\). For each \(n \geq 1\), let \(K_n \subset X\) be compact such that \(T(K_n|z, u) > 1 - 1/n\). As \(Q(dy|x, u)\) is continuous in total variation norm, the image of \(K_n \times \{u\}\) under \(Q(dy|x, u)\) is compact in \(\mathcal{P}(\mathcal{Y})\). Hence, there exist \(\{\nu_1, \ldots, \nu_l\} \subset \mathcal{P}(\mathcal{Y})\) such that

\[
\max_{x \in K_n} \min_{i=1, \ldots, l} \|Q(\cdot|x, u) - \nu_i\|_{TV} < 1/3n.
\]

Define the following stochastic kernel \(\nu_n(\cdot|x, u) = \arg\min_{\nu_i} \|Q(\cdot|x, u) - \nu_i\|_{TV}\). Then, we define \(P_n(\cdot|z, u) = \int_X \nu_n(\cdot|x, u)T(dx|z, u)\). One can prove that \(\|P(\cdot|z, u) - P_n(\cdot|z, u)\|_{TV} < 1/n\). Moreover, since \(P_n(\cdot|z, u)\) is a mixture of finite probability measures \(\{\nu_1, \ldots, \nu_l\}\), we have that \(\nu_n(\cdot|x, u) \ll P_n(\cdot|z, u)\) for all \(x \in C_n\), where \(T(C_n|z, u) = 1\). Let \(C = \bigcap_n C_n\), and so, \(T(C|z, u) = 1\). We claim that if \(x \in C\), then \(Q(dy|x, u) \ll P(dy|z, u)\), which completes the proof of the first statement. To prove the claim, fix any \(\varepsilon > 0\) and choose \(n \geq 1\) such that \(\varepsilon > 2/3n\). Then, there exists \(\delta > 0\) such that \(\nu_n(A|x, u) < \varepsilon/2\) whenever \(P_n(A|z, u) < \delta\). This implies that \(Q(A|x, u) < \varepsilon\) whenever \(P(A|z, u) < \delta + 1/n\). Hence, \(Q(dy|x, u) \ll P(dy|z, u)\).

To show the second claim, for any \(A \in \mathcal{B}(\mathcal{Y})\) and for any \(x_k \to x\), we define

\[
A_+^{(k)} := \{y \in A : g(x_k, u, y) > g(x, u, y)\}
\]

\[
A_-^{(k)} := \{y \in A : g(x_k, u, y) < g(x, u, y)\}.
\]

With these sets, we have

\[
\int_A |g(x_k, u, y) - g(x, u, y)|P(dy|z, u)
\]

\[
= \int_{A_+^{(k)}} g(x_k, u, y)P(dy|z, u) - \int_{A_-^{(k)}} g(x, u, y)P(dy|z, u)
\]
\[ + \int_{A^{(k)}} g(x, u, y) P(dy|z, u) - \int_{A^{(k)}} g(x_k, u, y) P(dy|z, u) \]
\[ \leq |Q(A^{(k)}|x_k, u) - Q(A^{(k)}|x, u)| \]
\[ + |Q(A^{(k)}|x_k, u) - Q(A^{(k)}|x, u)| \]
\[ \leq 2\|Q(\cdot|x_k, u) - Q(\cdot|x, u)\|_{TV} \to 0. \]

### B Proof of Lemma 1

Note that since, for any \( x_n \to x \) in \( X \), we have

\[ \lim_{n \to \infty} \sup_{\lambda \in \Lambda} |f_\lambda(x_n) - f_\lambda(x)| = 0, \]  
we see that \( \{f_\lambda\}_{\lambda \in \Lambda} \) is an equicontinuous family of functions. Thus, by the Arzela-Ascoli Theorem [28], for any given compact set \( K \subset X \) and \( \epsilon > 0 \), there is a finite set of continuous and bounded functions \( F := \{f_1, \ldots, f_N\} \), so that, for any \( \lambda \in \Lambda \), there is \( f_i \in F \) with

\[ \sup_{x \in K} |f_\lambda(x) - f_i(x)| \leq \epsilon. \]

Now, we claim that, for the same \( f_i \in F \), we have \( \sup_{x \in K} |f_n,\lambda(x) - f_i(x)| \leq 3\epsilon/2 \) for large enough \( n \), which is independent of \( \lambda \). To see this, observe the following:

\[ \sup_{x \in K} |f_{n,\lambda}(x) - f_i(x)| \leq \sup_{x \in K} |f_{n,\lambda}(x) - f_\lambda(x)| \]
\[ + \sup_{x \in K} |f_\lambda(x) - f_i(x)|. \]

Note that the second term is less than \( \epsilon \) and the first term can be made arbitrarily small as \( f_{n,\lambda} \to f_\lambda \) uniformly on compact sets and on \( \Lambda \), which can be easily proved using the assumptions in the lemma.

Note that \( \mu_n \to \mu \) weakly. Hence, \( \{\mu_n\} \) is a tight family of probability measures by Prokhorov theorem [29, Theorem 5.2]. Therefore, for any \( \epsilon > 0 \), there exists a compact subset \( K_\epsilon \) of \( X \) such that, for all \( n \),

\[ \mu_n(K_\epsilon) \geq 1 - \epsilon. \]

Now, we fix any \( \epsilon > 0 \) and choose a compact set \( K_\epsilon \) such that, for all \( n \), \( \mu_n(K_\epsilon) \geq 1 - \epsilon \). We also fix a finite family of continuous and bounded functions \( F := \{f_1, \ldots, f_N\} \) such that, for any \( \lambda \), we can find \( f_i \in F \) with
\[
\sup_{x \in K_\varepsilon} |f_\lambda(x) - f_i(x)| \leq \varepsilon. \text{ Moreover, we choose a large } N \text{ such that }
\sup_{x \in K_\varepsilon} |f_{n, \gamma}(x) - f_i(x)| \leq 3\varepsilon/2 \text{ for all } n \geq N.
\]

With this setup, we go back to the main statement:

\[
\sup_{\lambda \in \Lambda} \left| \int_{X \setminus K_\varepsilon} f_{n, \lambda}(x) \mu_n(dx) - \int_{X \setminus K_\varepsilon} f_\lambda(x) \mu(dx) \right|
\leq \sup_{\lambda \in \Lambda} \left| \int_{K_\varepsilon} f_{n, \lambda}(x) \mu_n(dx) - \int_{K_\varepsilon} f_\lambda(x) \mu(dx) \right|
\]

\[
\quad + \sup_{\lambda \in \Lambda} \left| \int_{K_\varepsilon} f_{n, \lambda}(x) \mu_n(dx) - \int_{K_\varepsilon} f_i(x) \mu(dx) \right|
\]

\[
\leq 2\varepsilon C + \sup_{\lambda \in \Lambda} \left| \int_{K_\varepsilon} (f_{n, \lambda}(x) - f_i(x)) \mu_n(dx) \right|
\]

\[
\quad + \int_{K_\varepsilon} f_i(x) \mu_n(dx) - \int_{K_\varepsilon} f_i(x) \mu(dx)
\]

\[
\quad + \int_{K_\varepsilon} (f_i(x) - f_\lambda(x)) \mu(dx)
\]

\[
\leq 2\varepsilon C + \left| \int_{K_\varepsilon} f_i(x) \mu_n(dx) - \int_{K_\varepsilon} f_i(x) \mu(dx) \right| + 5\varepsilon/2
\]

\[
\leq 2\varepsilon C + \left| \int_{X \setminus K_\varepsilon} f_i(x) \mu_n(dx) - \int_{K_\varepsilon} f_i(x) \mu(dx) \right|
\]

\[
\quad + \left| \int_{K_\varepsilon} f_i(x) \mu(dx) - \int_{K_\varepsilon} f_i(x) \mu_n(dx) \right| + 5\varepsilon/2
\]

\[
\leq 4\varepsilon C + 5\varepsilon/2 + \left| \int_{X \setminus K_\varepsilon} f_i(x) \mu_n(dx) - \int_{X \setminus K_\varepsilon} f_i(x) \mu(dx) \right|
\]

where \(C\) is the uniform bound on \(\{f_{n, \lambda}\}\) and \(\{f_\lambda\}\). Since \(\varepsilon\) is arbitrary and \(\mu_n\) converges weakly to \(\mu\), the result follows.

**References**

[1] O. Hernández-Lerma and J. Lasserre, *Discrete-Time Markov Control Processes: Basic Optimality Criteria*. Springer, 1996.

[2] 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 Prob. Appl.*, vol. 21, pp. 153–158, 1976.
[3] D. Rhenius, “Incomplete information in Markovian decision models,” *Ann. Statist.*, vol. 2, pp. 1327–1334, 1974.

[4] P. Billingsley, *Convergence of probability measures*, 2nd ed. New York: Wiley, 1999.

[5] K. Parthasarathy, *Probability Measures on Metric Spaces*. AMS Bookstore, 1967.

[6] D. P. Bertsekas and S. Shreve, *Stochastic Optimal Control: The Discrete Time Case*. New York: Academic Press, 1978.

[7] O. Hernández-Lerma, *Adaptive Markov Control Processes*. Springer-Verlag, 1989.

[8] D. Blackwell, “Memoryless strategies in finite-stage dynamic programming,” *Annals of Mathematical Statistics*, vol. 35, pp. 863–865, 1964.

[9] O. Hernández-Lerma and J. Lasserre, *Further Topics on Discrete-Time Markov Control Processes*. Springer, 1999.

[10] V. S. Borkar, “Average cost dynamic programming equations for controlled Markov chains with partial observations,” *SIAM Journal on Control and Optimization*, vol. 39, no. 3, pp. 673–681, 2000.

[11] V. S. Borkar and A. Budhiraja, “A further remark on dynamic programming for partially observed Markov processes,” *Stochastic processes and their applications*, vol. 112, no. 1, pp. 79–93, 2004.

[12] A. S. Manne, “Linear programming and sequential decision,” *Management Science*, vol. 6, pp. 259–267, April 1960.

[13] V. S. Borkar, “Convex analytic methods in Markov decision processes,” in *Handbook of Markov Decision Processes*, E. A. Feinberg, A. Shwartz (Eds.). Kluwer, Boston, MA, 2001, pp. 347–375.

[14] N. Saldi, S. Yüksel, and T. Linder, “Near optimality of quantized policies in stochastic control under weak continuity conditions,” *J. Math. Anal. Appl.*, vol. 435, pp. 321–337, 2016.

[15] N. Saldi, S. Yüksel, and T. Linder, “On the asymptotic optimality of finite approximations to Markov decision processes with Borel spaces,” *Mathematics of Operations Research*, vol. 42, no. 4, pp. 945–978, 2017.
[16] J. B. Lasserre, “Invariant probabilities for Markov chains on a metric space,” *Statistics and Probability Letters*, vol. 34, pp. 259–265, 1997.

[17] A. Budhiraja, “On invariant measures of discrete time filters in the correlated signal-noise case,” *The Annals of Applied Probability*, vol. 12, no. 3, pp. 1096–1113, 2002.

[18] P. Del Moral, “Measure-valued processes and interacting particle systems. Application to nonlinear filtering problems,” *Ann. Appl. Probab.*, vol. 8, no. 2, pp. 438–495, 05 1998. [Online]. Available: https://doi.org/10.1214/aoap/1028903535

[19] D. Crisan and A. Doucet, “A survey of convergence results on particle filtering methods for practitioners,” *IEEE Transactions on Signal Processing*, vol. 50, no. 3, pp. 736–746, March 2002.

[20] L. Stettner, “On invariant measures of filtering processes,” in *Stochastic Differential Systems*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1989, pp. 279–292.

[21] A. G. Bhatt, A. Budhiraja, and R. L. Karandikar, “Markov property and ergodicity of the nonlinear filter,” *SIAM Journal on Control and Optimization*, vol. 39, no. 3, pp. 928–949, 2000.

[22] P. Billingsley, *Probability and Measure*, 3rd ed. Wiley, 1995.

[23] A. D. Kara and S. Yüksel, “Robustness to incorrect system models in stochastic control,” *arXiv preprint arXiv:1803.06046*, 2018.

[24] E. Feinberg, P. Kasyanov, and Y. Liang, "Fatou’s Lemma for Weakly Converging Measures under the Uniform Integrability Condition," *arXiv preprint arXiv:1807.07931*.

[25] E. Feinberg, P. Kasyanov, and M. Zgurovsky, “Convergence of probability measures and Markov decision models with incomplete information,” *Proceedings of the Steklov Institute of Mathematics*, vol. 287, pp. 96–117, 2014.

[26] E. Feinberg, P. Kasyanov, and M. Zgurovsky, “Partially observable total-cost Markov decision process with weakly continuous transition probabilities,” *Mathematics of Operations Research*, vol. 41, no. 2, pp. 656–681, 2016.
[27] E. Feinberg, P. Kasyanov, and M. Zgurovsky, "Uniform Fatou's lemma,” Journal of Mathematical Analysis and Applications, vol. 444, issue 1 pp. 550–567, 2016.

[28] R. M. Dudley, Real Analysis and Probability. Cambridge University Press, 2004.

[29] P. Billingsley, Convergence of Probability Measures. New York: Wiley, 1968.