Asymptotic Optimality of Finite Model Approximations for Partially Observed Markov Decision Processes with Discounted Cost

Naci Saldi, Serdar Yüksel, Tamás Linder

Abstract—We consider finite model approximations of discrete-time partially observed Markov decision processes (POMDPs) under the discounted cost criterion. After converting the original partially observed stochastic control problem to a fully observed one on the belief space, the finite models are obtained through the uniform quantization of the state and action spaces of the belief space Markov decision process (MDP). Under mild assumptions on the components of the original model, it is established that the policies obtained from these finite models are nearly optimal for the belief space MDP, and so, for the original partially observed problem. The assumptions essentially require that the belief space MDP satisfies a mild weak continuity condition. We provide an example and introduce explicit approximation procedures for the quantization of the set of probability measures on the state space of POMDP (i.e., belief space).

I. INTRODUCTION

In POMDP theory, existence of optimal policies have in general been established via converting the original partially observed stochastic control problem to a fully observed one on the belief space, leading to a belief-MDP. However, computing an optimal policy for this fully observed model, and so for the original POMDP, using well known dynamic programming algorithms is challenging even if the original system has finite state and action spaces, since the state space of the fully observed model is always uncountable. One way to overcome this difficulty is to compute an approximately optimal policy instead of a true optimal policy by constructing a reduced model for the fully observed system for which one can apply well known algorithms such as policy iteration, value iteration, and Q-learning etc. to obtain the optimal policy.

In MDP theory, various methods have been developed to compute near optimal policies by reducing the original problem into a simpler one. A partial list of these techniques is as follows: approximate dynamic programming, approximate value or policy iteration, simulation-based techniques, neuro-dynamic programming (or reinforcement learning), state aggregation, etc. We refer the reader to [1]–[9] and references therein. However, existing works mostly study systems with discrete (i.e., finite or countable) state and action spaces [1], [5], [6], [10]) or those that consider general state and action

spaces (see, e.g., [7]–[9], [11], [12]) assume in general Lipschitz type continuity conditions on the transition probability and the one-stage cost function in order to provide a rate of convergence analysis for the approximation error. However, for the fully observed reduction of POMDP, a Lipschitz type regularity condition on the transition probability is in general prohibitive. Indeed, demonstrating even the arguably most relaxed regularity condition on the transition probability (i.e., weak continuity in state-action variable), is a challenging problem as was recently demonstrated in [13] for general state and action spaces (see also [14], [15] for a control-free setup). Therefore, results developed in prior literature cannot in general be applied to compute approximately optimal policies for fully observed reduction of POMDP, and so, for the original POMDP.

In [16], [17] we investigated finite action and state approximations of fully observed stochastic control problems with general state and action spaces under the discounted cost and average cost optimality criteria. For the discounted cost case, we showed that optimal policies obtained from these finite models asymptotically achieve the optimal cost for the original problem under the weak continuity assumption on the transition probability. Here, we apply and properly generalize the results in these papers to obtain approximation results for fully observed reduction of POMDPs, and so, for POMDPs. The versatility of approximation results under weak continuity conditions become particularly evident while investigating the applicability of these results to the partially observed case.

In the literature there exist various, mostly numerical and computational, results for obtaining approximately optimal policies for POMDPs. In the following, we list a number of such related results and comparisons with our paper: (i) Reference [18] develops a computational algorithm, utilizing structural convexity properties of the value function of belief-MDPs, for the solutions of POMDPs when the state space is continuous and action and measurements are discrete, and with further extensions to continuous action and measurements. Reference [19] provides an algorithm which may be regarded as a quantization of the belief space. However, no rigorous convergence results are given regarding this computational algorithm. (ii) References [20] and [21] present quantization based algorithms for the belief state, where the state, measurement, and the action sets are finite. (iii) References [22] and [23] provide an explicit quantization method for the set of probability measures containing the belief states, where the state space, unlike in many other contributions in the

literature, is continuous. The quantization is done through the approximations as measured by Kullback-Leibler divergence: Kullback-Leibler divergence (or relative entropy) is a very strong pseudo-distance measure which is even stronger than total variation (by Pinsker's inequality [24]), which in turn is stronger than weak convergence. In particular, being able to quantize the space of probability measures with finitely many balls as defined by such a distance measure requires very strict assumptions on the allowable beliefs and it in particular requires, typically equicontinuity conditions (see e.g. [25, Lemma 4.3]). (iv) In [26] the authors consider the near optimality of finite-state controllers that are finite-state probabilistic automatons taking observations as inputs and producing controls as the outputs. A special case for these type of controllers are the ones that only use finite observation history. A similar finite memory approximation is developed in [27]. (v) In [28] the authors establish finite state approximation schemes for the belief-MDPs under both discounted cost and average cost criteria using concavity properties of the corresponding value function and show that approximate costs can be used as lower bounds for the optimal cost function. A similar finite state approximation is considered in [29], [30] using concavity and convexity properties of the value function for the discounted cost criterion. We refer the reader to the survey papers [31], [32] and the book [33] for further algorithmic and computational procedures for approximating POMDPs.

Contributions of the paper. (i) We show that finite models asymptotically approximate the original partially observed Markov decision process (POMDP) in the sense that the true costs of the policies obtained from these finite models converge to the optimal cost of the original model. The finite models are constructed by discretizing both the state and action spaces of the equivalent fully-observed belief space formulation of the POMDP. We establish the result for models with general state and action spaces under mild conditions on the system components. (ii) We provide systematic procedures for the quantization of the set of probability measures on the state space of POMDPs which is the state space of belief-MDPs. By choosing an appropriate metric for the weak convergence topology, any probability measure on the state space can be approximated by the set of probability measures with fixed finite support. Then, one can use the existing methods in the literature for discretizing the probability simplex to obtain finite-state approximations of the belief space. (iii) Our rigorous results can be used to justify the novel quantization techniques presented in [19]–[21] as well as a more relaxed version for the results presented in [22] and [23]. In particular, there do not exist approximation results of the generality presented in our paper with asymptotic performance guarantees. We show that provided that the belief space is quantized according to balls generated through metrics that metrize the weak convergence topology (such as Prokhorov, bounded-Lipschitz, or stronger ones such as the Wasserstein metric), and provided that the action sets are quantized in a uniform fashion, under very weak conditions on the controlled Markov chain (namely the weak continuity of the kernel and total variation continuity of the measurement channel), asymptotic

optimality is guaranteed. (iv) Our approach also highlights the difficulties of obtaining explicit rates of convergence results for approximation methods for POMDPs. Even in fully observed models, for obtaining explicit rates of convergence, one needs strong continuity conditions of the Lipschitz type, e.g.; [17, Theorem 5.1]. As Theorem 1 shows, this is impossible under even quite strong conditions for POMDPs.

The rest of the paper is organized as follows. In Section II we introduce the partially observed stochastic control model and construct the belief space formulation. In Section III we establish the continuity properties that are satisfied by the transition probability of the belief space MDP. In Section IV we construct the finite model approximations and state approximation results. In Section V we provide explicit methods to quantize the set of probability measures on the state space. In Section VI we illustrate our results by considering a numerical example. Section VII concludes the paper.

II. PARTIALLY OBSERVED MARKOV DECISION PROCESSES

A discrete-time partially observed Markov decision process (POMDP) has the following components: (i) State space X, action space A, and observation space Y, all Borel spaces, (ii) $p(\,\cdot\,|x,a)$ is the transition probability of the next state given the current state-action pair is (x,a), (iii) $r(\,\cdot\,|x)$ is the observation channel giving the probability of the current observation given the current state variable x, and (iv) the one-stage cost function $c : \mathsf{X} \times \mathsf{A} \to [0,\infty)$.

To complete the description of the partially observed control model, we must specify how the controller designs its control law at each time step. To this end, define the history spaces $\mathsf{H}_0 = \mathsf{Y}$ and $\mathsf{H}_t = (\mathsf{Y} \times \mathsf{A})^t \times \mathsf{Y}$, $t = 1, 2, \ldots$ endowed with their product Borel σ-algebras generated by $\mathcal{B}(\mathsf{Y})$ and $\mathcal{B}(\mathsf{A})$. A *policy* $\pi = \{\pi_t\}$ is a sequence of stochastic kernels on A given H_t. We denote by Π the set of all policies.

According to the Ionescu Tulcea theorem [34], an initial distribution μ on X and a policy π define a unique probability measure P_μ^π on $\mathsf{H}_\infty \times \mathsf{X}^\infty$. The expectation with respect to P_μ^π is denoted by $\mathbb{E}_\mu^\pi$. For any initial distribution μ and policy π we can think of the POMDP as a stochastic process $\{x_t, y_t, a_t\}_{t \geq 0}$ defined on the probability space $(\Omega, \mathcal{F}, P_\mu^\pi)$, where $\Omega = \mathsf{H}_\infty \times \mathsf{X}^\infty$, the x_t are X-valued random variables, the y_t are Y-valued random variables, the a_t are A-valued random variables, and they satisfy for all $t \geq 1$

$$P_\mu^\pi(x_0 \in \,\cdot\,) = \mu(\,\cdot\,)$$
$$P_\mu^\pi(x_t \in \,\cdot\,|x_{\{0,t-1\}}, y_{\{0,t-1\}}, a_{\{0,t-1\}})$$
$$\qquad = P_\mu^\pi(x_t \in \,\cdot\,|x_{t-1}, a_{t-1}) = p(\,\cdot\,|x_{t-1}, a_{t-1})$$
$$P_\mu^\pi(y_t \in \,\cdot\,|x_{\{0,t\}}, y_{\{0,t-1\}}, a_{\{0,t-1\}})$$
$$\qquad = P_\mu^\pi(y_t \in \,\cdot\,|x_t) = r(\,\cdot\,|x_t)$$
$$P_\mu^\pi(a_t \in \,\cdot\,|x_{\{0,t\}}, y_{\{0,t\}}, a_{\{0,t-1\}}) = \pi_t(\,\cdot\,|y_{\{0,t\}}, a_{\{0,t-1\}})$$

where $x_{\{0,t\}} = (x_0, \ldots, x_t)$, $y_{\{0,t\}} = (y_0, \ldots, y_t)$, and $a_{\{0,t\}} = (a_0, \ldots, a_t)$. We denote by $J(\pi, \mu)$ the discounted cost function of the policy $\pi \in \Pi$ with initial distribution μ, which is given by

$$J(\pi, \mu) := \mathbb{E}_\mu^\pi \left[\sum_{t=0}^\infty \beta^t c(x_t, a_t) \right],$$

where $\beta \in (0,1)$ is the discount factor.

With this notation, the discounted value function of the control problem is defined as

$$J^*(\mu) := \inf_{\pi \in \Pi} J(\pi, \mu).$$

A policy π^* is said to be optimal if $J(\pi^*, \mu) = J^*(\mu)$.

In POMDPs, since the information available to the decision maker is a noisy version of the state, one cannot apply the dynamic programming principle directly as the one-stage cost function depends on the exact state information. A canonical way to overcome this difficulty is converting the original partially observed control problem to a fully observed one by taking the posterior state distributions $\Pr\{x_t \in \cdot | y_0, \ldots, y_t, a_0, \ldots, a_{t-1}\} \in \mathcal{P}(\mathsf{X})$ ('beliefs' of the observer) as the new state variable and the function $\tilde{c} : \mathcal{P}(\mathsf{X}) \times \mathsf{A} \to [0, \infty)$, defined by

$$\tilde{c}(z, a) := \int_\mathsf{X} c(x, a) z(dx)$$

as the new one-stage cost function. This fully-observed MDP is called the belief-MDP and 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 belief-MDP which is optimal. Therefore, finite-model approximation results developed for fully-observed MDPs can be applied to the belief-MDP and so to the POMDP.

III. CONTINUITY PROPERTIES OF BELIEF-MDPs

In this section, we first discuss the continuity properties that are satisfied by or prohibitive for the transition probability of the belief-MDP. Then, we derive the conditions satisfied by the components of the belief-MDP.

A. On the Convergence of Probability Measures

Let E be a Borel space and let $\mathcal{P}(\mathsf{E})$ denote the family of all probability measure on $(\mathsf{E}, \mathcal{B}(\mathsf{E}))$. A sequence $\{\mu_n\}$ is said to converge to $\mu \in \mathcal{P}(\mathsf{E})$ *weakly* (resp., *setwise*) if

$$\int_\mathsf{E} g(e)\mu_n(de) \to \int_\mathsf{E} g(e)\mu(de)$$

for all continuous and bounded real function g (resp., for all measurable and bounded real function g).

For any $\mu, \nu \in \mathcal{P}(\mathsf{E})$, the *total variation* norm is given by

$$
\begin{aligned}
\|\mu - \nu\|_{TV} &:= 2 \sup_{B \in \mathcal{B}(\mathsf{E})} |\mu(B) - \nu(B)| \\
&= \sup_{f : \|f\|_\infty \leq 1} \left| \int_\mathsf{E} f(e)\mu(de) - \int_\mathsf{E} f(e)\nu(de) \right|,
\end{aligned}
$$

where the supremum is over all measurable real f such that $\|f\|_\infty = \sup_{e \in \mathsf{E}} |f(e)| \leq 1$. A sequence $\{\mu_n\}$ is said to converge to $\mu \in \mathcal{P}(\mathsf{E})$ in total variation if $\|\mu_n - \mu\|_{TV} \to 0$. As it is clear from the definitions, total variation convergence implies setwise convergence, which in turn implies weak convergence.

The total variation metric leads to a stringent notion for convergence. For example a sequence of discrete probability measures on a finite-dimensional Euclidean space never converges in total variation to a probability measure which admits a density function with respect to the Lebesgue measure. Setwise convergence also induces a topology which is not easy to work with since the space under this convergence is not metrizable [35, p. 59]. However, the space of probability measures on a Borel space endowed with the topology of weak convergence is itself a Borel space [36]. The bounded-Lipschitz metric ρ_{BL} [37, p.109], for example, can be used to metrize this space:

$$\rho_{BL}(\mu, \nu) := \sup_{\|f\|_{BL} \leq 1} \left| \int_\mathsf{E} f(e)\mu(de) - \int_\mathsf{E} f(e)\nu(de) \right|, \quad (1)$$

where

$$\|f\|_{BL} := \|f\|_\infty + \sup_{e \neq e'} \frac{f(e) - f(e')}{d_\mathsf{E}(e, e')},$$

and d_E is the metric on E. Finally, the Wasserstein metric of order 1, W_1, can also be used for compact E (see [37, Theorem 6.9]):

$$W_1(\mu, \nu) = \inf_{\eta \in \mathcal{H}(\mu, \nu)} \int_{\mathsf{E} \times \mathsf{E}} d_\mathsf{E}(e, e')\eta(de, de'),$$

where $\mathcal{H}(\mu, \nu)$ denotes the set of probability measures on $\mathsf{E} \times \mathsf{E}$ with first marginal μ and second marginal ν. Indeed, W_1 can also be used as an upper bound to ρ_{BL} for non-compact E since W_1 can equivalently be written as [37, Remark 6.5]:

$$W_1(\mu, \nu) := \sup_{\|f\|_{Lip} \leq 1} \left| \int_\mathsf{E} f(e)\mu(de) - \int_\mathsf{E} f(e)\nu(de) \right|,$$

where

$$\|f\|_{Lip} := \sup_{e \neq e'} \frac{f(e) - f(e')}{d_\mathsf{E}(e, e')}.$$

Comparing this with (1), it follows that

$$\rho_{BL} \leq W_1. \quad (2)$$

This observation will be utilized for the quantization algorithms on the set of probability measures later in the paper.

B. Belief Space Formulation of POMDPs

As indicated in Section II, any POMDP can be reduced to a (completely observable) MDP [38], [39], whose states are the posterior state distributions or *beliefs* of the observer; that is, the state at time t is

$$z_t := \Pr\{x_t \in \cdot | y_0, \ldots, y_t, a_0, \ldots, a_{t-1}\} \in \mathcal{P}(\mathsf{X}).$$

In this section, we construct the components of this belief-MDP under some assumptions on the components of the POMDP. Later, we establish the conditions satisfied by the components of the belief-MDP, under which we can apply approximation results in our earlier work [16], [17] to the belief-MDP, and so, to the original POMDP.

To this end, let $v : \mathsf{X} \to [0, \infty)$ be a continuous moment function in the sense that there exists an increasing sequence of compact subsets $\{K_n\}_{n \geq 1}$ of X such that

$$\lim_{n \to \infty} \inf_{x \in \mathsf{X} \setminus K_n} v(x) = \infty.$$

The following assumptions will be imposed on the components of the POMDP.

Assumption 1.

(a) The one-stage cost function c is continuous and bounded.

(b) We have one of the following:

   (1) the stochastic kernel $p(\,\cdot\,|x,a)$ is weakly continuous in $(x,a) \in \mathsf{X} \times \mathsf{A}$, i.e., if $(x_k, a_k) \to (x,a)$, then $p(\,\cdot\,|x_k,a_k) \to p(\,\cdot\,|x,a)$ weakly. The observation channel $r(\,\cdot\,|x)$ is continuous in total variation, i.e., if $x_k \to x$, then $r(\,\cdot\,|x_k) \to r(\,\cdot\,|x)$ in total variation.

   (2) the stochastic kernel $p(\,\cdot\,|x,a)$ is total variation continuous in $(x,a) \in \mathsf{X} \times \mathsf{A}$, i.e., if $(x_k, a_k) \to (x,a)$, then $p(\,\cdot\,|x_k,a_k) \to p(\,\cdot\,|x,a)$ in total variation.

(c) A is compact.

(d) There exists a constant $\lambda \geq 0$ such that

$$\sup_{a \in \mathsf{A}} \int_{\mathsf{X}} v(y)p(dy|x,a) \leq \lambda v(x).$$

(e) The initial probability measure μ satisfies

$$\int_{\mathsf{X}} v(x)\mu(dx) < \infty.$$

We let

$$\mathcal{P}_v(\mathsf{X}) := \left\{ \mu \in \mathcal{P}(\mathsf{X}) : \int_{\mathsf{X}} v(x)\mu(dx) < \infty \right\}.$$

Note that since the probability law of x_t is in $\mathcal{P}_v(\mathsf{X})$, by Assumption 1-(d),(e), under any policy we have $\Pr\{x_t \in \cdot \,|y_0, \ldots, y_t, a_0, \ldots, a_{t-1}\} \in \mathcal{P}_v(\mathsf{X})$ almost everywhere. Therefore, the belief-MDP has state space $\mathsf{Z} = \mathcal{P}_v(\mathsf{X})$ instead of $\mathcal{P}(\mathsf{X})$, where Z is equipped with the Borel σ-algebra generated by the topology of weak convergence. The transition probability η of the belief-MDP can be constructed as follows (see also [40]). Let z denote the generic state variable for the belief-MDP. First consider the transition probability on $\mathsf{X} \times \mathsf{Y}$ given $\mathsf{Z} \times \mathsf{A}$

$$R(x \in A, y \in B|z,a) := \int_{\mathsf{X}} \kappa(A, B|x', a)z(dx'),$$

where $\kappa(dx, dy|x', a) := r(dy|x)p(dx|x', a)$. Let us disintegrate R as

$$R(dx, dy|z,a) = H(dy|z,a)F(dx|z,a,y).$$

Then, we define the mapping $F : \mathsf{Z} \times \mathsf{A} \times \mathsf{Y} \to \mathsf{Z}$ as

$$F(z,a,y) = F(\,\cdot\,|z,a,y). \tag{3}$$

In the literature, (3) is called the 'nonlinear filtering equation' [40]. Note that, for each $t \geq 0$, we indeed have

$$F(z,a,y)(\,\cdot\,) = \Pr\{x_{t+1} \in \cdot \,|z_t = z, a_t = a, y_{t+1} = y\}$$

and

$$H(\,\cdot\,|z,a) = \Pr\{Y_{t+1} \in \cdot \,|z_t = z, a_t = a\}.$$

Then, η can be written as

$$\eta(\,\cdot\,|z,a) = \int_{\mathsf{Y}} \delta_{F(z,a,y)}(\,\cdot\,)\, H(dy|z,a),$$

where δ_z denotes the Dirac-delta measure at point z; that is, $\delta_z(D) = 1$ if $z \in D$ and otherwise it is zero. Recall that the initial point for the belief-MDP is μ; that is, $z_0 \sim \delta_\mu$, and the one-stage cost function $\tilde{c}$ of the belief-MDP is given by

$$\tilde{c}(z,a) := \int_{\mathsf{X}} c(x,a)z(dx). \tag{4}$$

Hence, the belief-MDP is a fully-observed Markov decision process with the components $(\mathsf{Z}, \mathsf{A}, \eta, \tilde{c})$. For the belief-MDP define the history spaces $\tilde{\mathsf{H}}_0 = \mathsf{Z}$ and $\tilde{\mathsf{H}}_t = (\mathsf{Z} \times \mathsf{A})^t \times \mathsf{Z}$, $t = 1, 2, \ldots$ and let $\tilde{\Pi}$ denote the set of all policies for the belief-MDP, where the policies are defined in a usual manner. Let $\tilde{J}(\tilde{\pi}, \xi)$ denote the discounted cost function of policy $\tilde{\pi} \in \tilde{\Pi}$ for initial distribution ξ of the belief-MDP.

Notice that any history vector $\tilde{h}_t = (z_0, \ldots, z_t, a_0, \ldots, a_{t-1})$ of the belief-MDP is a function of the history vector $h_t = (y_0, \ldots, y_t, a_0, \ldots, a_{t-1})$ of the POMDP. Let us write this relation as $i(h_t) = \tilde{h}_t$. Hence, for a policy $\tilde{\pi} = \{\tilde{\pi}_t\} \in \tilde{\Pi}$, we can define $\pi^{\tilde{\pi}} = \{\pi_t^{\tilde{\pi}}\} \in \Pi$ as

$$\pi_t^{\tilde{\pi}}(\,\cdot\,|h_t) := \tilde{\pi}_t(\,\cdot\,|i(h_t)). \tag{5}$$

Let us write this as a mapping from $\tilde{\Pi}$ to Π: $\tilde{\Pi} \ni \tilde{\pi} \mapsto i(\tilde{\pi}) = \pi^{\tilde{\pi}} \in \Pi$. It is straightforward to show that the cost functions $\tilde{J}(\tilde{\pi}, \xi)$ and $J(\pi^{\tilde{\pi}}, \mu)$ are the same, where $\xi = \delta_\mu$. One can also prove that (see [38], [39])

$$\inf_{\tilde{\pi} \in \tilde{\Pi}} \tilde{J}(\tilde{\pi}, \xi) = \inf_{\pi \in \Pi} J(\pi, \mu) \tag{6}$$

and furthermore, that if $\tilde{\pi}$ is an optimal policy for the belief-MDP, then $\pi^{\tilde{\pi}}$ is optimal for the POMDP as well. Hence, the POMDP and the corresponding belief-MDP are equivalent in the sense of cost minimization. Therefore, approximation results developed for MDPs in [16], [17] can be applied to the belief-MDP and so, to the POMDP.

C. Strong and Weak Continuity Properties of the Belief MDP

The stochastic kernel η is said to be *weakly* continuous if $\eta(\,\cdot\,|z_k, a_k) \to \eta(\,\cdot\,|z,a)$ weakly, whenever $(z_k, a_k) \to (z,a)$. The kernel is said to be *strongly* continuous if, for any $z \in \mathsf{Z}$, $\eta(\,\cdot\,|z, a_k) \to \eta(\,\cdot\,|z,a)$ setwise, whenever $a_k \to a$. For the fully observed reduction of a partially observed MDP (POMDP), requiring strong continuity of the transition probability is in general too strong condition. This is illustrated through the simple example [16, Example 2.1]. The following theorem is a consequence of [41, Theorem 3.7], [53, Theorem 2] and [16, Example 2.1].

Theorem 1.

(i) Under Assumption 1-(b), the stochastic kernel η for belief-MDP is weakly continuous in (z,a).

(ii) The stochastic kernel η may not be setwise continuous in action variable a even if the observation channel is continuous in total variation.

Part (i) of Theorem 1 implies that the transition probability η of the belief-MDP is weakly continuous under Assumption 1. On the other hand, the continuity of the observation channel in total variation is not enough for the setwise continuity of η.

The above suggest that our earlier results in [16] and [17], which only require weak continuity conditions on the transition kernel of a given MDP, are particularly suitable in developing approximation methods for POMDPs (through their MDP reduction), in both quantizing the action spaces as well as state spaces.

Remark 1. We refer the reader to [41, Theorem 3.2(c)], [53, Section 4] for more general conditions implying weak continuity of the transition probability η. We also note that, in the uncontrolled setting, [42] and [14] have established similar weak continuity conditions (i.e., the weak-Feller property) of the non-linear filter process (i.e., the belief process) in continuous time and discrete time, respectively.

Remark 2. Our aim in this paper is to show the asymptotic optimality of finite model approximations under very general conditions. Establishing general regularity conditions that produce an explicit error bound on the approximation as a function of the number of bins in the quantization, and thus obtaining rates of convergence to optimality, is in general prohibitive. To obtain explicit rates of convergence even for fully observed systems, one needs strong continuity conditions of the Lipschitz type on the components of the fully observed reduction of the POMDP, see for example Theorem 5.2 in [17]. In view of Theorem 1, this does not seem possible under mild conditions since even setwise continuity does not hold. Indeed, establishing conditions that result in a Lipschitz type regularity behaviour (for the belief-MDP) on the components of the fully observed transition kernel and the measurement models is an interesting future research topic. Theorem 1 suggests that these conditions should be quite restrictive.

Example 1. In this example we consider the following partially observed model

$$
\begin{aligned}
x_{t+1} &= F(x_t, a_t, v_t), \\
y_t &= H(x_t, w_t), \ t = 0, 1, 2, \dots
\end{aligned}
\tag{7}
$$

where $\mathsf{X} = \mathbb{R}^n$, $\mathsf{A} \subset \mathbb{R}^m$, and $\mathsf{Y} \subset \mathbb{R}^d$ for some $n, m, d \geq 1$. The noise processes $\{v_t\}$ and $\{w_t\}$ are sequences of independent and identically distributed (i.i.d.) random vectors taking values in $\mathsf{V} = \mathbb{R}^p$ and $\mathsf{W} = \mathbb{R}^l$, respectively, for some $p, l \geq 1$, and they are also independent of each other. In this system, the continuity of F in (x, a) is sufficient to imply the weak continuity of the transition probability p, and no assumptions are needed on the noise process (not even the existence of a density is required). On the other hand, the continuity of the observation channel r in total variation holds, if for any $x \in \mathsf{X}$, the probability measure $r(\,\cdot\,|x)$ has a density $g(y, x)$, which is continuous in x, with respect to some reference probability measure m on Y. This follows from Scheffé's theorem (see, e.g., [43, Theorem 16.2]). For instance, this density condition holds for the following type of models:

(i) In the first model, we have $\mathsf{Y} = \mathsf{W} = \mathbb{R}^d$, $H(x, w) = H(x) + w$, H is continuous, and w has a continuous density g_w with respect to Lebesgue measure.

(ii) In the second case, Y is countable and $r(y|x)$ is continuous in x for all $y \in \mathsf{Y}$. Therefore, the transition probability η of the belief space MDP, corresponding to the model in (7), is weakly continuous.

Next, we derive conditions satisfied by the components of the belief-MDP under Assumption 1. Note first that $\mathsf{Z} = \bigcup_{m \geq 1} F_m$, where

$$
F_m := \left\{ \mu \in \mathcal{P}_v(\mathsf{X}) : \int_{\mathsf{X}} v(x) \mu(dx) \leq m \right\}.
$$

Since v is a moment function, each F_m is tight [34, Proposition E.8]. Moreover, each F_m is also closed since v is continuous. Therefore, each F_m is compact with respect to the weak topology. This implies that Z is a σ-compact Borel space. Note that by [44, Proposition 7.30], the one-stage cost function $\tilde{c}$ of the belief-MDP, which is defined in (4), is in $C_b(\mathsf{Z} \times \mathsf{A})$ under Assumption 1-(a). Therefore, the belief-MDP satisfies the following conditions under Assumption 1, which we formally state as a separate assumption.

Assumption 2.

(i) The one-stage cost function $\tilde{c}$ is bounded and continuous.

(ii) The stochastic kernel η is weakly continuous.

(iii) A is compact and Z is σ-compact.

IV. FINITE MODEL APPROXIMATIONS

A. Finite-Action Approximation

In this section, we consider finite-action approximation of the belief-MDP and so, the POMDP. For these equivalent models, we obtain an approximate finite-action model as follows. Let d_A denote the metric on A. Since A is assumed compact and thus totally bounded, there exists a sequence of finite sets $\Lambda_n = \{a_{n,1}, \dots, a_{n,k_n}\} \subset \mathsf{A}$ such that for each n,

$$
\min_{i \in \{1, \dots, k_n\}} d_\mathsf{A}(a, a_{n,i}) < 1/n \text{ for all } a \in \mathsf{A}.
$$

In other words, Λ_n is a $1/n$-net in A. The sequence $\{\Lambda_n\}_{n \geq 1}$ is used by the finite-action model to approximate the belief-MDP and the POMDP.

In [16], for MDPs with Borel state and action spaces, we studied the problem of approximating an uncountable action set with a finite one and had established the asymptotic optimality of finite action models for such fully observed MDPs that satisfy a number of technical conditions, in particular, the weak continuity condition of the transition kernel in state and action variables. Given the belief-MDP reduction, [16, Theorem 3.2] implies the following result.

Theorem 2. Suppose Assumption 1 (and thus Assumption 2) holds for the POMDP. Then we have

$$
\lim_{n \to \infty} |\tilde{J}_n^*(z) - \tilde{J}^*(z)| = 0 \ \text{ for all } z \in \mathsf{Z},
$$

where $\tilde{J}_n^*$ is the discounted cost value function of the belief-MDP$_n$ with the components $(\mathsf{Z}, \Lambda_n, \eta, \tilde{c})$ and $\tilde{J}^*$ is the discounted cost value function of the belief-MDP with components $(\mathsf{Z}, \mathsf{A}, \eta, \tilde{c})$.

The significance of Theorem 2 is reinforced by the following observation. If we let $\Pi(\Lambda_n)$ to be the set of deterministic policies of the POMDP taking values in Λ_n, then the theorem implies that for any given $\varepsilon > 0$ there exists $n \geq 1$ and $\pi^* \in \Pi(\Lambda_n)$ such that

$$
J(\pi_\varepsilon^*, \mu) < \min_{\pi \in \Pi} J(\pi, \mu) + \varepsilon,
$$

where $\pi_\varepsilon^* = \pi^{\tilde{f}_n^*}$ (see (5)) and $\tilde{f}_n^*$ is the optimal deterministic stationary policy for the belief-MDP$_n$.

B. Finite-State Approximation

The finite-state model for the belief-MDP is obtained as in [17], by quantizing the set of probability measures $Z = \mathcal{P}_v(X)$; that is, for each m, we quantize compact set F_m similar to the quantization of A and represent the rest of the points $Z \setminus F_m$ by some pseudo-state.

If Z is compact, in the following, the index m can be fixed to $m = 1$.

We let d_Z denote a metric on Z which metrizes the weak topology. For each $m \geq 1$, since F_m is compact and thus totally bounded, there exists a sequence $\left(\{z_{n,i}^{(m)}\}_{i=1}^{k_n^{(m)}}\right)_{n \geq 1}$ of finite grids in F_m such that for all $n \geq 1$,

$$\min_{i \in \{1, \ldots, k_n^{(m)}\}} d_Z(z, z_{n,i}^{(m)}) < 1/n \text{ for all } z \in F_m. \quad (8)$$

Let $\{\mathcal{S}_{n,i}^{(m)}\}_{i=1}^{k_n^{(m)}}$ be a partition of F_m such that $z_{n,i}^{(m)} \in \mathcal{S}_{n,i}^{(m)}$ and

$$\max_{z \in \mathcal{S}_{n,i}^{(m)}} d_Z(z, z_{n,i}^{(m)}) < 1/n \quad (9)$$

for all $i = 1, \ldots, k_n^{(m)}$. Choose any $z_{n,k_n^{(m)}+1}^{(m)} \in Z \setminus F_m$ which is a so-called pseudo-state and set $\mathcal{S}_{n,k_n^{(m)}+1} = Z \setminus F_m$. Let $Z_n^{(m)} := \{z_{n,1}^{(m)}, \ldots, z_{n,k_n}^{(m)}, z_{n,k_n^{(m)}+1}^{(m)}\}$ and define function $Q_n^{(m)} : Z \to Z_n^{(m)}$ by

$$Q_n^{(m)}(z) = z_{n,i}^{(m)} \text{ when } z \in \mathcal{S}_{n,i}^{(m)}.$$

Here $Q_n^{(m)}(z)$ maps z to the representative element of the partition it belongs to.

Remark 3.

(a) Note that given $\{z_{n,i}^{(m)}\}_{i=1}^{k_n^{(m)}} \subset F_m$ that satisfies (8), one way to obtain the corresponding partition $\{\mathcal{S}_{n,i}^{(m)}\}_{i=1}^{k_n^{(m)}}$ of F_m satisfying (9) as follows. Let us define function $Q_{\text{near}} : F_m \to \{z_{n,1}^{(m)}, \ldots, z_{n,k_n^{(m)}}^{(m)}\}$ as

$$Q_{\text{near}}(z) = \arg \min_{z_{n,i}^{(m)}} d_Z(z, z_{n,i}^{(m)}),$$

where ties are broken so that Q_{near} is measurable. In the literature, Q_{near} is often called a nearest neighbor quantizer with respect to 'distortion measure' d_Z [45]. Then, Q_{near} induces a partition $\{\mathcal{S}_{n,i}^{(m)}\}_{i=1}^{k_n^{(m)}}$ of the space F_m given by

$$\mathcal{S}_{n,i}^{(m)} = \{z \in F_m : Q_n^{(m)}(z) = z_{n,i}^{(m)}\},$$

and which satisfies (9). Although one can construct, in theory, the partition using nearest neighbor sense, it is computationally difficult to find these regions when the original state space X is uncountable.

(b) The index n indicates the resolution of the quantizer that is applied to discretize the compact set F_m and index m emphasizes the size of the compact set F_m for which quantization is applied.

Let $\{\nu_n^{(m)}\}$ be a sequence of probability measures on Z satisfying

$$\nu_n^{(m)}(\mathcal{S}_{n,i}^{(m)}) > 0 \text{ for all } i, n, m. \quad (10)$$

One possible choice for $\nu_n^{(m)}$ is

$$\nu_n^{(m)}(\cdot) = \sum_{i=1}^{k_n^{(m)}+1} \delta_{z_{n,i}^{(m)}}(\cdot).$$

We let $\nu_{n,i}^{(m)}$ be the restriction of $\nu_n^{(m)}$ to $\mathcal{S}_{n,i}^{(m)}$ defined by

$$\nu_{n,i}^{(m)}(\cdot) := \frac{\nu_n^{(m)}(\cdot)}{\nu_n^{(m)}(\mathcal{S}_{n,i}^{(m)})}.$$

The measures $\nu_{n,i}^{(m)}$ will be used to define a sequence of finite-state belief MDPs, denoted as MDP$_n^{(m)}$, which approximate the belief-MDP. To this end, for each n and m define the one-stage cost function $c_n^{(m)} : Z_n^{(m)} \times A \to [0, \infty)$ and the transition probability $p_n^{(m)}$ on $Z_n^{(m)}$ given $Z_n^{(m)} \times A$ by

$$c_n^{(m)}(z_{n,i}^{(m)}, a) := \int_{\mathcal{S}_{n,i}^{(m)}} c(z, a) \nu_{n,i}^{(m)}(dz),$$

$$p_n^{(m)}(\cdot \mid z_{n,i}^{(m)}, a) := \int_{\mathcal{S}_{n,i}^{(m)}} Q_n^{(m)} * p(\cdot \mid z, a) \nu_{n,i}^{(m)}(dz), \quad (11)$$

where $Q_n^{(m)} * p(\cdot \mid z, a) \in \mathcal{P}(Z_n^{(m)})$ is the pushforward of the measure $p(\cdot \mid z, a)$ with respect to $Q_n^{(m)}$; that is,

$$Q_n^{(m)} * p(y \mid z, a) = p(\{z \in Z : Q_n^{(m)}(z) = y\} \mid z, a),$$

for all $y \in Z_n^{(m)}$. For each n and m, we define MDP$_n^{(m)}$ as a Markov decision process with the following components: $Z_n^{(m)}$ is the state space, A is the action space, $p_n^{(m)}$ is the transition probability, and $c_n^{(m)}$ is the one-stage cost function.

Given the belief-MDP, [17, Theorem 3.2][1] implies the following.

Theorem 3. Suppose Assumption 1 (and thus Assumption 2) holds for the POMDP. Then we have

$$\lim_{n,m \to \infty} |\tilde{J}(f_n^{(m)}, \mu) - \tilde{J}^*(\mu)| = 0,$$

where $f_n^{(m)}$ is obtained by extending the optimal policy of the MDP$_n^{(m)}$ to Z. Hence, by the equivalence of POMDPs and belief-MDPs, we also have

$$\lim_{n,m \to \infty} |J(\pi^{f_n^{(m)}}, \mu) - J^*(\mu)| = 0.$$

Theorem 3 implies that to find a near optimal policy for the POMDP, it is sufficient to compute an optimal policy for the finite-state belief-MDP with sufficiently many states, extend this policy to the original state space of the belief-MDP, and then construct the corresponding policy for the POMDP. Next, we discuss explicit methods to quantize the set of probability measures on X, that is, the belief-space Z.

---

[1]Although Theorem 3.2 in [17] is proved under the assumption that the state space is locally compact, a careful examination of the proof reveals that σ-compactness of the state space is also sufficient to establish the result.

V. QUANTIZATION OF THE BELIEF-SPACE

An explicit construction for an application requires a properly defined metric on Z. As stated in Section III-A, one can metrize the set of probability measures defined on a Borel space under the weak topology using various distance measures. Building on this fact, in the following we present explicit methods for the quantization of Z for the cases where X is finite, a compact subset of a finite dimensional Euclidean space, or the finite-dimensional Euclidean space itself, and $p(\cdot|a)$ is independent of the state variable x.

A. Construction with Finite X

If the state space is finite with $|X| = m$, then $Z = \mathcal{P}_v(X) = \mathcal{P}(X)$, and Z is a simplex in $\mathbb{R}^m$. In this case, Euclidean distance can be used to metrize Z. Indeed, one can make use of the algorithm in [46] (see also [47]) to quantize Z in a nearest neighbor manner. To this end, for each $n \geq 1$, define

$$Z_n := \left\{ (p_1, \ldots, p_m) \in \mathbb{Q}^m : p_i = \frac{k_i}{n}, \sum_{i=1}^{m} k_i = n \right\}, \quad (12)$$

where $\mathbb{Q}$ is the set of rational numbers and $n, k_1, \ldots, k_m \in \mathbb{Z}_+$. The set Z_n is called *type lattice* by analogy with the concept of *types* in information theory [48, Chp. 12]. Then, the algorithm that computes the nearest neighbor levels can be described as follows:

Algorithm. Given $z \in Z$, find nearest $y \in Z_n$:
(1) Compute values ($i = 1, \ldots, m$)

$$k_i' = \left\lfloor nz_i + \frac{1}{2} \right\rfloor \text{and } n' = \sum_{i=1}^{m} k_i'.$$

(2) If $n' = n$ the nearest y is given by $(\frac{k_1'}{n}, \ldots, \frac{k_m'}{n})$. Otherwise, compute the errors

$$\delta_i = k_i' - nz_i,$$

and sort them

$$\frac{-1}{2} \leq \delta_{i_1} \leq \delta_{i_2} \leq \ldots \leq \delta_{i_m} \leq \frac{1}{2}.$$

(3) Let $\Delta = n' - n$. If $\Delta > 0$, set

$$k_{i_j} = \begin{cases} k_{i_j}' & \text{if } j = 1, \ldots, m - \Delta - 1 \\ k_{i_j}' - 1 & \text{if } j = m - \Delta, \ldots, m. \end{cases}$$

If $\Delta < 0$, set

$$k_{i_j} = \begin{cases} k_{i_j}' + 1 & \text{if } j = 1, \ldots, |\Delta| \\ k_{i_j}' & \text{if } j = |\Delta| + 1, \ldots, m. \end{cases}$$

Then, the nearest y is given by $(\frac{k_1}{n}, \ldots, \frac{k_m}{n})$.

One can also compute the maximum radius of the quantization regions for this algorithm. To this end, let d_∞ and d_p denote respectively the metrics induced by L_∞ and L_p ($p \geq 1$) norms on $\mathbb{R}^m$, which metrizes the weak topology on Z. Then, we have [46, Proposition 2]

$$b_\infty := \max_{z \in Z} \min_{y \in Z_n} d_\infty(z, y) = \frac{1}{n} \left(1 - \frac{1}{m} \right),$$

$$b_2 := \max_{z \in Z} \min_{y \in Z_n} d_2(z, y) = \frac{1}{n} \sqrt{\frac{a(m - a)}{m}},$$

$$b_1 := \max_{z \in Z} \min_{y \in Z_n} d_1(z, y) = \frac{1}{n} \frac{2a(m - a)}{m},$$

where $a = \lfloor m/2 \rfloor$. Hence, for each $n \geq 1$, the set Z_n is an b_j-net in Z with respect to d_j metric, where $j \in \{\infty, 2, 1\}$.

B. Construction with Compact X

The analysis in the previous subsection shows that a finitely supported measure can be approximated through *type lattice*s. Thus, if compactly supported probability measures can be approximated with those having finite support, the analysis in Section V-A yields approximately optimal policies. In the following, we assume that X is a compact subset of $\mathbb{R}^d$ for some $d \geq 1$. Then $Z := \mathcal{P}_v(X) = \mathcal{P}(X)$ is also compact (under the weak convergence topology) and can be metrized using the Wasserstein metric W_1 (here, in defining W_1, we use the metric on X induced by the Euclidean norm $\|\cdot\|$) as discussed in Section III-A.

For each $n \geq 1$, let Q_n be some lattice quantizer [45] on X such that $\|x - Q_n(x)\| < 1/n$ for all $x \in X$. Set $X_n = Q_n(X)$, i.e., the output levels of Q_n (note that X_n is finite since X is compact). Then, one can approximate any probability measure in Z with probability measures in

$$\mathcal{P}(X_n) := \left\{ \mu \in \mathcal{P}(X) : \mu(X_n) = 1 \right\}.$$

Indeed, for any $\mu \in Z$, we have [49, Theorem 2.6]

$$\inf_{\mu' \in \mathcal{P}(X_n)} W_1(\mu, \mu') \leq \inf_{Q:X \to X_n} \int_X \|x - Q(x)\| \mu(dx)$$

$$\leq \int_X \|x - Q_n(x)\| \mu(dx) \leq \frac{1}{n}.$$

Once this is obtained, we can further approximate the probability measure induced by Q_n via the algorithm introduced in Section V-A with asymptotic performance guarantees. Thus, through a sequence of type lattices Z_{m_n} as given in (12) with a successively refined support set so that $X_n \subset X_{n+1}$ for $n \in \mathbb{N}$ with $m_n = |X_n|$, one can quantize Z to obtain a sequence of finite state-action MDPs through (11) leading to Theorem 3.

For some related properties of approximations of probability measures with those with finite support, and the relation to optimal quantization, we refer the reader to [49].

C. Construction with Non-compact X

Here we assume that $X = \mathbb{R}^d$ for some $d \geq 1$ and that Assumption 1 holds for $v(x) = \|x\|^2$. In this case, $Z := \mathcal{P}_v(X)$ becomes the set of probability measures with finite second moment and F_m is the set of probability measures with finite second moments bounded by m. We endow here Z with the bounded-Lipschitz metric ρ_{BL}, which metrizes weak convergence (see Section III-A).

We first describe the discretization procedure for F_m. For each $n \geq 1$, set $K_n := [-n, n]^d$ and let q_n denote a lattice quantizer on K_n satisfying

$$\sup_{x \in K} \|x - q_n(x)\| < 1/n.$$

Let X_n denote the set of output levels of q_n; that is, $\mathsf{X}_n = q_n(K_n)$. Define

$$Q_n(x) = \begin{cases} q_n(x) & \text{if } x \in K \\ 0 & \text{if } x \in K^c, \end{cases}$$

Then, any measure in F_m can be approximated by probability measures in

$$\mathcal{P}(\mathsf{X}_n) := \left\{ \mu \in \mathcal{P}(\mathsf{X}) : \mu(\mathsf{X}_n) = 1 \right\}.$$

Indeed, for any $\mu \in F_m$, we have

$$\inf_{\mu' \in \mathcal{P}(\mathsf{X}_n)} \rho_{BL}(\mu, \mu') \leq \inf_{\mu' \in \mathcal{P}(\mathsf{X}_n)} W_1(\mu, \mu') \tag{13}$$

$$\leq \inf_{Q:\mathsf{X} \to \mathsf{X}_n} \int_\mathsf{X} \|x - Q(x)\| \mu(dx)$$

$$\leq \int_\mathsf{X} \|x - Q_n(x)\| \mu(dx)$$

$$= \int_K \|x - Q_n(x)\| \mu(dx) + \int_{K^c} \|x\| \mu(dx)$$

$$\leq \frac{1}{n} + \int_{\{\|x\|>n\}} \|x\|^2 \mu(dx) \frac{1}{n}$$

$$\leq \frac{(1+m)}{n} \tag{14}$$

In the derivation above, (13) follows from (2). Thus, μ in F_m can be approximated by the $\mu_n \in \mathcal{P}(\mathsf{X}_n)$, which is induced by the quantizer Q_n, with a bound $\rho_{BL}(\mu, \mu_n) \leq (1+m)/n$. Then, similar to Section V-B, we can further approximate probability measure μ_n via the algorithm introduced in Section V-A with again asymptotic performance guarantees by Theorem 3. Thus, analogous to compact case, using a sequence of type lattices Z_{m_n} as given in (12) with a successively refined support set $\mathsf{X}_n \subset \mathsf{X}_{n+1}$ for $n \in \mathbb{N}$ with $m_n = |\mathsf{X}_n|$, one can quantize Z to obtain a sequence of finite state-action MDPs through (11).

Remark 4. Note that the complexity of the quantization algorithm scales with the dimension of the space on which it is applied. Therefore, quantizing a belief space when the underlying state space is continuous (compact or noncompact) is in general computationally very demanding since, in this case, the quantized space is infinite dimensional. One way to overcome this issue is to (approximately or exactly) parameterize the reachable beliefs with a low dimensional parameter as will be discussed in detail in the next section.

D. Construction for Special Models leading to Quantized Beliefs with Continuous Support

So far, we have obtained quantized beliefs where each such quantized belief measure was supported on a finite set. For some applications, this may not be efficient and it may be more desirable to quantize the measurement space appropriately. For some further applications, a parametric representation of the set of reachable beliefs may be present and the construction of bins may be more immediate through quantizing the parameters in a parametric class. What is essential in such models is that the bins designed to construct the finite belief-MDP correspond to balls which are *small* under the

metrics that metrize the weak convergence as discussed in Section III-A. Therefore, although the belief space is originally infinite dimensional in such models, because of the special parametrization of the reachable beliefs, the complexity of the finite model approximation algorithm can be reduced to the complexity of the quantization of the parameter space. This drastically decreases the complexity of the algorithm as can be seen in the numerical example in Section VI-B. For instance, if we obtain the discretized states by quantizing the observation space (which can be done if the transition probability only depends on the actions as discussed in the next section), then the complexity of the finite model approximation algorithm scales with the dimension of the observation space. Hence, exact or approximate parametrization of reachable beliefs can drastically reduce the complexity of the algorithm. This is a very efficient way to overcome the *curse-of-dimensionality* in approximation of POMDPs.

1) Quantized Measures Through Quantized Measurements: For this section, we assume that transition probability $p(\cdot|a)$ is independent of the state variable x, $\mathsf{Y} \subset \mathbb{R}^p$ for some $p \geq 1$, and Assumption 1 holds for some v. In the view of Theorem 2, as a pre-processing set-up, we quantize the action space A, where the finite set A_q represents the output levels of this quantizer. Hence, in the sequel, we assume that the action space is A_q.

Since $\kappa(dx, dy|a) := r(dy|x) \otimes p(dx|a)$, we have

$$R(x \in A, y \in B|z, a) = \int_\mathsf{X} \kappa(A, B|a) z(dx')$$
$$= \kappa(A, B|a),$$

and so, the disintegration of R becomes

$$R(dx, dy|a) = H(dy|a) \otimes F(dx|a, y).$$

Then, η is given by

$$\eta(\cdot|a) = \int_\mathsf{Y} \delta_{F(\cdot|a,y)}(\cdot) \, H(dy|a).$$

This implies that we can take the following set as the state space Z of the fully-observed model instead of $\mathcal{P}_v(\mathsf{X})$:

$$\mathsf{Z} = \left\{ F(\cdot|a, y) : (a, y) \in \mathsf{A}_q \times \mathsf{Y} \right\}.$$

We endow Z with the bounded-Lipschitz metric ρ_{BL}. For each $n \geq 1$, set $L_n := [-n, n]^p$ and let l_n denote a lattice quantizer on L_n satisfying

$$\sup_{y \in L_n} \|y - l_n(y)\| < 1/n.$$

Let Y_n denote the set of output levels of l_n; that is, $\mathsf{Y}_n = l_n(L_n)$. Define

$$q_n(y) = \begin{cases} l_n(y) & \text{if } y \in L_n \\ 0 & \text{if } y \in L_n^c. \end{cases}$$

Then, finite set $\mathsf{Z}_n \subset \mathsf{Z}$, which is used to quantize Z, is given by

$$\mathsf{Z}_n = \left\{ F(\cdot|a, y) : (a, y) \in \mathsf{A}_q \times \mathsf{Y}_n \right\},$$

and the corresponding quantizer $Q_n : \mathsf{Z} \to \mathsf{Z}_n$ is defined as follows: given $z = F(\cdot|a, y)$, we define

$$Q_n(z) = F(\cdot|a, q_n(y)).$$

Note that to use Q_n for constructing finite models, we have to obtain an upper bound on the ρ_{BL}-distance between z and $Q_n(z)$. This can be achieved under various assumptions on the system components. One such assumption is the following: (i) $\mathsf{X} = \mathbb{R}^d$ for some $d \geq 1$, (ii) Y is compact, (ii) $p(dx|a) = g_p(x|a)m(dx)$ and $r(dy|x) = g_r(y|x)m(dy)$, (iv) g_r is Lipschitz continuous with Lipschitz constant K_r, $g_r > \theta$ for some $\theta > 0$, and $\sup_{\{(y,x)\in\mathsf{Y}\times\mathsf{X}\}} |g_r(y, x)| =: \|g_r\| < \infty$. Since Y is compact, there exists $\epsilon(n)$ for each n such that $\epsilon(n) \to 0$ as $n \to \infty$ and $\|y - q_n(y)\| \leq \epsilon(n)$ for all $y \in \mathsf{Y}$. Under the above assumptions, we have

$$F(dx|a, y) = f(x|a, y)m(dx),$$

where

$$f(x|a, y) = \frac{g_r(y|x)g_p(x|a)}{\int_\mathsf{X} g_r(y|x)g_p(x|a)m(dx)}.$$

Since the bounded-Lipschitz metric ρ_{BL} is upper bounded by the total variation distance, we obtain

$$\rho_{BL}(z, Q_n(z)) \leq \|z - Q_n(z)\|_{TV}$$
$$= \int_\mathsf{X} \big|f(x|a, y) - f(x|a, q_n(y))\big|m(dx)$$
$$\leq \frac{2\|g_r\|K_r}{\theta^2}\|y - q_n(y)\| \leq \frac{2\|g_r\|K_r}{\theta^2}\epsilon(n).$$

Hence, Q_n is a legitimate quantizer for constructing the finite models. Section VI-B exhibits another example where we have such an upper bound.

2) Construction from a Parametrically Represented Class: For some applications, the set of belief measures can be first approximated by some parametric class of measures, where parameters belong to some low-dimensional space [23], [50], [51]. For instance, in [23], densities of belief measures are projected onto exponential family of densities using the *Kullback-Leibler (KL) divergence*, where it was assumed that projected beliefs are close enough to true beliefs in terms of cost functions. In [51], densities are parameterized by unimodal Gaussian distributions and parameterized MDP are solved through Monte Carlo simulation based method. In [50], densities are represented by sufficient statistics, and in particular represented by Gaussian distributions, and the parameterized MDP is solved through fitted value iteration algorithm. However, among these works, only the [23] develop rigorous error bounds for their algorithms using the KL divergence and the other works do not specify distance measures to quantify parametric representation approximations.

In these methods, if the parameterized beliefs are good enough to represent true beliefs as it was shown in [23], then the method presented in the earlier sections (of first quantizing the state space, and then quantizing the beliefs on the state space) may not be necessary and one can, by quantizing the parameters for the class of beliefs considered, directly construct the finite belief-MDP. As noted earlier, what is essential in such methods is that the bins designed to construct the finite belief-MDP correspond to balls which are *small* under the metrics that metrize the weak convergence as discussed in Section III-A. This possible if the projected beliefs are proved to be close to the true beliefs with respect to some metric that generates the weak topology or with respect to some (pseudo) distance which is stronger than weak topology. For instance, since convergence in KL-divergence is stronger than weak convergence, the projected beliefs constructed in [23] indeed satisfies this requirement. Hence, one can apply our results to conclude the convergence of the reduced model to the original model in [23]. As noted earlier, relative entropy is a very strong pseudo-distance measure which is even stronger than total variation (by Pinsker's inequality [24]) and for being able to quantize a set of probability measures with finitely many balls as defined by such a distance measure requires very strict assumptions on the allowable beliefs and it in particular requires, typically equicontinuity conditions (see e.g. [25, Lemma 4.3]). In turn, it is in general necessary to assume that transition probability and observation channel have very strong regularity conditions.

VI. NUMERICAL EXAMPLES

A. Example with Finite X

We consider a machine repair problem in order to illustrate our results numerically for finite state POMDPs. In this model, we have $\mathsf{X} = \mathsf{A} = \mathsf{Y} = \{0, 1\}$ with the following interpretation:

$$x_t = \begin{cases} 1 & \text{machine is working at time } t \\ 0 & \text{machine is not working at time } t, \end{cases}$$

$$a_t = \begin{cases} 1 & \text{machine is being repaired at time } t \\ 0 & \text{machine is not being repaired at time } t, \end{cases}$$

Below κ is the probability that the machine repair was successful given an initial 'not working' state:

$$\Pr\{x_{t+1} = 1|x_t = 0, a_t = 1\} = \kappa.$$

Finally, the probability that the machine does not break down in one time step is denoted by α:

$$\Pr\{x_{t+1} = 0|x_t = 1, a_t = 0\} = \alpha.$$

The probability that the measured state is not the true state is given by ε; that is,

$$\Pr\{y_t = 0|x_t = 1\} = \Pr\{y_t = 1|x_t = 0\} = \varepsilon.$$

The one-stage cost function for this model is given by:

$$c(x, a) = \begin{cases} R + E & x = 0 \text{ and } a = 1 \\ E & x = 0 \text{ and } a = 0 \\ 0 & x = 1 \text{ and } a = 0 \\ R & x = 1 \text{ and } a = 1, \end{cases}$$

where R is defined to be the cost of repair and E is the cost incurred by a broken machine. The cost function to be minimized is the discounted cost function with a discount factor β. In order to find the approximately optimal policies, we first construct the belief space formulation of the above

model. Note that the state space of the belief space model is the interval $[0, 1]$. Hence, we can use uniform quantization on $[0, 1]$ to obtain the finite model. For the numerical results, we use the following parameters: $\varepsilon = 0.17$, $\kappa = 0.9$, $\alpha = 0.9545$, and $\beta = 0.3$. We selected 20 different values for the number n of grid points to discretize $[0, 1]$: $n = 10, 20, 30, \ldots, 200$. The grid points are chosen uniformly. For each n, the finite state models are constructed as in [17, Section 2].

Figure 1 shows the graph of the value functions of the finite models corresponding to the different values of n (number of grid points), when the initial state is $x = 1$. The resulting graph suggest convergence of the value functions (to the value function of the original model). Note that the rigorous justification of eventual convergence is given in the paper; see Theorems 2 and 3.
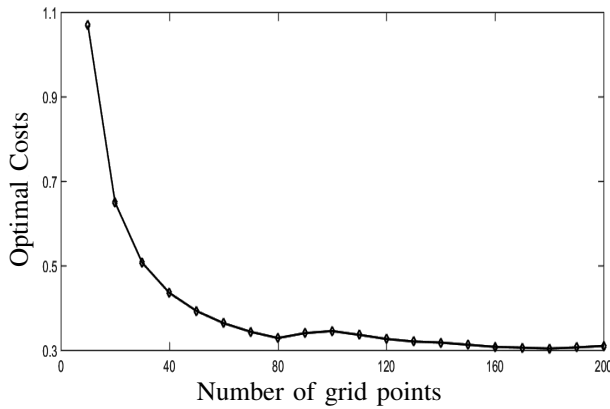

Figure 1. Optimal costs of the finite models when the initial state is $x = 1$

B. Example with Compact X

In this example we consider the following model:

$$x_{t+1} = \exp\{-\theta_1 a_t + v_t\}, \ t = 0, 1, 2, \ldots \quad (15)$$
$$y_t = x_t + \xi_t, \ t = 0, 1, 2, \ldots \quad (16)$$

where $\theta_1 \in \mathbb{R}_+$, x_t is the state at t, and a_t is the action at t. The one-stage 'reward' function is $u(x_t - a_t)$, where u is some utility function. In this model, the goal is to maximize the discounted reward. This model is the modified and partially observed version of the population growth model in [34, Section 1.3].

The state and action spaces are $X = A = [0, L]$, for some $L \in \mathbb{R}_+$, and the observation space is $Y = [0, K]$ for some $K \in \mathbb{R}_+$. Since θ_1 is merely a constant, by taking $[0, \frac{L}{\theta_1}]$ as our new action space, instead of dynamics in equation (15) we can write the dynamics of the state as

$$x_{t+1} = \exp\{-a_t + v_t\}, \ t = 0, 1, 2 \ldots.$$

The noise processes $\{v_t\}$ and $\{\xi_t\}$ are sequences of independent and identically distributed (i.i.d.) random variables which have common densities g_v supported on $[0, \lambda]$ and g_ξ supported on $[0, \tau]$, respectively. Therefore, the transition probability $p(\cdot | x, a)$ is given by

$$p(D|x, a) = \int_D g_v(\log(v) + a) \frac{1}{v} m(dv),$$

for all $D \in \mathcal{B}(\mathbb{R})$ and the observation kernel $r(\cdot | x)$ is given by

$$r(B|x, a) = \int_B g_\xi(\xi - x) m(d\xi),$$

for all $B \in \mathcal{B}(\mathbb{R})$. To make the model consistent, we must have $\exp\{-a + v\} \in [0, L]$ for all $(a, v) \in [0, L] \times [0, \lambda]$. We assume that g_v and g_ξ are uniform probability density functions; that is, $g_v = \frac{1}{\lambda}$ on $[0, \lambda]$ and $g_\xi = \frac{1}{\tau}$ on $[0, \tau]$. Hence, Assumption 1 holds for this model with $v(x) = 1$.

In the view of Theorem 2, as a pre-processing set-up, we quantize the action space A, where the finite set $A_q = \{a_1, a_2, \ldots, a_q\}$ represents the output levels of this quantizer with $0 < a_1 < a_2 < \ldots < a_q$. In the remainder of this example we assume that the action space is A_q.

We now obtain the stochastic kernels $H(\cdot | z, a)$ and $F(\cdot | z, a, y)$ that describe the transition probability η of the reduced MDP. Indeed, we have $H(dy|z, a) = h(y|a) m(dy)$ where $h(y|a)$ is given by

$$h(y|a) = \int_X g_\xi(y - x) g_v(\log(x) + a) \frac{1}{x} m(dx).$$

Similarly, we have $F(dx|z, a, y) = f(dx|a, y) m(dx)$ where $f(x|z, a, y)$ is given by

$$f(x|a, y) = \frac{g_\xi(y - x) g_v(\log(x) + a) \frac{1}{x}}{\int_X g_\xi(y - x) g_v(\log(x) + a) \frac{1}{x} m(dx)} \quad (17)$$

Hence, for any (z, a), the transition probability $\eta(\cdot | z, a)$ has a support on the set of probability measures on X having densities given by (17). This implies that we can take the following set as the state space Z of the fully-observed model instead of $\mathcal{P}(X)$:

$$Z = \left\{ f(x|a, y) m(dx) : (a, y) \in A \times Y \text{ and } f \text{ as in (17)} \right\}.$$

For each n, let q_n denote the uniform quantizer on Y having n output levels; that is, $q_n : Y \to \{y_1, \ldots, y_n\} =: Y_n \subset Y$ where $y_j = (j - \frac{1}{2}) \Delta_n$, $j = 1, \ldots, n$, and

$$q_n^{-1}(y_j) = \left[y_j - \frac{\Delta_n}{2}, y_j + \frac{\Delta_n}{2} \right),$$

where $\Delta_n = \frac{K}{n}$. We define

$$Z_n := \left\{ f(x|a, y) m(dx) \in Z : (a, y) \in A_q \times Y_n \right\}.$$

Then, the quantizer $Q_n : Z \to Z_n$, which is used to construct the finite model, is defined as follows: given $z = f(x|a, y) m(dx)$, we define $Q_n(z) = f(x|a, q_n(y)) m(dx)$.

To be able to use Q_n for constructing finite models, we need to obtain an upper bound on the ρ_{BL}-distance between z and $Q_n(z)$. Indeed, one can prove that

$$\rho_{BL}(z, Q_n(z)) \le \|z - Q_n(z)\|_{TV} \le M \Delta_n, \quad (18)$$

for some constant M. The utility function u is taken to be quadratic function; i.e., $u(t) = t^2$

For the numerical results, we use the parameters as: $\lambda = 1$, $\tau = 0.5$, $\beta = 0.2$. As a pre-processing set-up, we uniformly discretize the action space A by using 20 grid points.

Then, we select 99 different values for the number n of grid points to discretize the state space Z using the quantizer Q_n, where n varies from 29 to 1436.

We use the value iteration algorithm to compute the value functions of the finite models. Since the components of the transition matrices and the cost function matrix can be computed analytically in this example, complexity of constructing these matrices is polynomial in the size of the action space (i.e. $|A_q|$) and state space (i.e. $|Z_n|$). In addition, it is known that the computation of an optimal policy using value iteration algorithm is also polynomial in $|A_q|$, $|Z_n|$, $\frac{1}{1-\beta}$, and the maximum number of bits B needed to represent any numerator or denominator of β or one of the components of the transition matrices or cost function matrix [52]. Therefore, the computation of an optimal policy for any finite model with n grid points is polynomial in $|A_q|$, $|Z_n|$, $\frac{1}{1-\beta}$, and B.

The simulation was implemented by using MATLAB and it took 411.75 seconds using an HP EliteDesk 800G2 SFF desktop computer with CPU Intel Core i7-6700 3.4 GHz 4-core. Figure 2 displays the graph of the value functions corresponding to the different values for the number of grid points when the initial state is $x = 2$. The resulting graph suggest convergence of the value functions; the rigorous justification of convergence is given in Theorems 2 and 3.
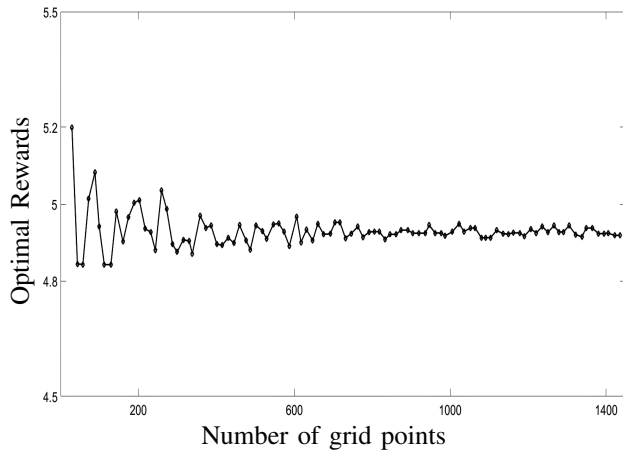

Figure 2. Optimal rewards of the finite models when the initial state is $x = 2$

VII. Concluding Remarks

We studied the approximation of discrete-time partially observed Markov decision processes under the discounted cost criterion. An essential observation was that establishing strong continuity properties for the reduced (belief) model is quite difficult for general state and action models, whereas weak continuity can be established under fairly mild conditions on the transition kernel of the original model and the measurement equations. This allowed us to apply our prior approximation results [16], [17], developed under weak continuity conditions, to partially observed models. Extending the analysis to average cost problems is a future task.

VIII. Acknowledgements

References

[1] B. Fox, "Finite-state approximations to denumerable state dynamic programs," *J. Math. Anal. Appl.*, vol. 34, pp. 665–670, 1971.

[2] W. Whitt, "Approximations of dynamic programs I," *Math. Oper. Res.*, vol. 3, no. 3, pp. 231–243, 1978.

[3] H. Langen, "Convergence of dynamic programming models," *Math. Oper. Res.*, vol. 6, no. 4, pp. 493–512, Nov. 1981.

[4] D. Bertsekas and J. Tsitsiklis, *Neuro-Dynammic Programming*. Athena Scientific, 1996.

[5] R. Ortner, "Pseudometrics for state aggregation in average reward Markov decision processes," in *Algorithmic Learning Theory*. Springer-Verlag, 2007.

[6] D. White, "Finite-state approximations for denumerable state infinite horizon discounted Markov decision processes," *J. Math. Anal. Appl.*, vol. 74, pp. 292–295, 1980.

[7] D. Bertsekas, "Convergence of discretization procedures in dynamic programming," *IEEE Trans. Autom. Control*, vol. 20, no. 3, pp. 415–419, Jun. 1975.

[8] F. Dufour and T. Prieto-Rumeau, "Finite linear programming approximations of constrained discounted Markov decision processes," *SIAM J. Control Optim.*, vol. 51, no. 2, pp. 1298–1324, 2013.

[9] ——, "Approximation of average cost Markov decision processes using empirical distributions and concentration inequalities," *Stochastics*, pp. 1–35, 2014.

[10] B. Roy, "Performance loss bounds for approximate value iteration with state aggregation," *Math. Oper. Res.*, vol. 31, no. 2, pp. 234–244, May 2006.

[11] F. Dufour and T. Prieto-Rumeau, "Approximation of Markov decision processes with general state space," *J. Math. Anal. Appl.*, vol. 388, pp. 1254–1267, 2012.

[12] C.-S. Chow and J. N. Tsitsiklis, "An optimal one-way multigrid algorithm for discrete-time stochastic control," *IEEE Transactions on Automatic Control*, vol. 36, no. 8, pp. 898–914, 1991.

[13] E. Feinberg, P. Kasyanov, and N. Zadioanchuk, "Average cost Markov decision processes with weakly continuous transition probabilities," *Math. Oper. Res.*, vol. 37, no. 4, pp. 591–607, Nov. 2012.

[14] 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.

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

[16] 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.

[17] N. Saldi, S. Yüksel, and T. Linder, "Asymptotic optimality of finite approximations to Markov decision processes with Borel spaces," *Math. Oper. Res.*, pp. 1–34, March 2017.

[18] J. M. Porta, N. Vlassis, M. T. J. Spaan, and P. Poupart, "Point-based value iteration for continuous POMDPs," *Journal of Machine Learning Research*, vol. 7, no. Nov, pp. 2329–2367, 2006.

[19] N. Vlassis and M. T. J. Spaan, "Perseus: Randomized point-based value iteration for POMDPs," *Journal of artificial intelligence research*, vol. 24, pp. 195–220, 2005.

[20] T. Smith and R. Simmons, "Point-based POMDP algorithms: Improved analysis and implementation," *arXiv preprint arXiv:1207.1412*, 2012.

[21] J. Pineau, G. Gordon, and S. Thrun, "Anytime point-based approximations for large POMDPs," *Journal of Artificial Intelligence Research*, vol. 27, pp. 335–380, 2006.

[22] E. Zhou, M. C. Fu, and S. I. Marcus, "A density projection approach to dimension reduction for continuous-state POMDPs," in *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*. IEEE, 2008, pp. 5576–5581.

[23] ——, "Solving continuous-state POMDPs via density projection," *IEEE Transactions on Automatic Control*, vol. 55, no. 5, pp. 1101–1116, 2010.

[24] R. M. Gray, *Entropy and Information Theory*. New York: Springer-Verlag, 1990.

[25] S. Yüksel and T. Linder, "Optimization and convergence of observation channels in stochastic control," *SIAM J. on Control and Optimization*, vol. 50, pp. 864–887, 2012.

[26] H. Yu and D. Bertsekas, "On near optimality of the set of finite-state controllers for average cost POMDP," *Math. Oper. Res.*, vol. 33, no. 1, pp. 1–11, Feb. 2008.

[27] C. White and W. Scherer, "Finite-memory suboptimal design for partially observed Markov decision processes," *Operations Research*, vol. 42, no. 3, pp. 439–455, 1994.

[28] H. Yu and D. Bertsekas, "Discretized approximations for POMDP with average cost," in *Conference on Uncertainity in Artifical Intelligence*, July 2004, pp. 619–627.

[29] W. Lovejoy, "Computationally feasible bounds for partially observed Markov decision processes," *Operations Research*, vol. 39, no. 1, pp. 162–175, 1991.

[30] R. Zhou and E. Hansen, "An improved grid-based approximation algorithm for POMDPs," in *Int. J. Conf. Artificial Intelligence*, Aug. 2001, pp. 707–714.

[31] W. Lovejoy, "A survey of algorithmic methods for partially observed Markov decision processes," *Annals of Operations Research*, vol. 28, pp. 47–66, 1991.

[32] C. White, "A survey of solution techniques for the partially observed Markov decision process," *Annals of Operations Research*, vol. 32, pp. 215–230, 1991.

[33] V. Krishnamurthy, *Partially observed Markov decision processes: from filtering to controlled sensing*. Cambridge University Press, 2016.

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

[35] J. K. Ghosh and R. V. Ramamoorthi, *Bayesian Nonparametrics*. New York: Springer, 2003.

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

[37] C. Villani, *Optimal transport: old and new*. Springer, 2009.

[38] 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.

[39] D. Rhenius, "Incomplete information in Markovian decision models," *Ann. Statist.*, vol. 2, pp. 1327–1334, 1974.

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

[41] E. Feinberg, P. Kasyanov, and M. Zgurovsky, "Partially observable total-cost Markov decision process with weakly continuous transition probabilities," arXiv:1401.2168, 2014.

[42] 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.

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

[44] D. P. Bertsekas and S. E. Shreve, *Stochastic optimal control: The discrete time case*. Academic Press New York, 1978.

[45] G. Gray and D. Neuhoff, "Quantization," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2325–2383, Oct. 1998.

[46] Y. A. Reznik, "An algorithm for quantization of discrete probability distributions," in *DCC 2011*, March 2011, pp. 333–342.

[47] G. Böcherer and B. C. Geiger, "Optimal quantization for distribution synthesis," *IEEE Transactions on Information Theory*, vol. 62, no. 11, pp. 6162–6172, 2016.

[48] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[49] W. Kreitmeier, "Optimal vector quantization in terms of Wasserstein distance," *Journal of Multivariate Analysis*, vol. 102, no. 8, pp. 1225–1239, 2011.

[50] A. Brooks, A. Makarenko, S. Williams, and H. Durrant-Whyte, "Parametric POMDPs for planning in continuous state spaces," *Robot. Auton. Syst.*, vol. 54, no. 11, pp. 887–897, 2006.

[51] A. Brooks and S. Williams, "A Monte Carlo update for parametric POMDPs," in *Proc. Int. Symp. Robot. Res.*, 2007.

[52] M. Littman, T. Dean, and L. Kaelbling, "On the complexity of solving markov decision problems," in *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*, ser. UAI'95, 1995, pp. 394–402.

[53] A.D. Kara and N. Saldi and S. Yüksel, "Weak Feller Property of Non-linear Filters," *arXiv preprint arXiv:1812.05509*, 2019.

PLACE PHOTO HERE

Naci Saldi (S'14, M'16) received the B.Sc. and M.S. degrees in Electrical and Electronics Engineering from Bilkent University in 2008 and 2010, respectively and the Ph.D. degree in Department of Mathematics and Statistics from Queen's University in 2015. He was a postdoctoral researcher at the University of Illinois at Urbana-Champaign before joining the Department of Natural and Mathematical Sciences at Ozyegin University as an Assistant Professor. His research interests include stochastic and decentralized control, source coding, mean-field games, and applied probability.

PLACE PHOTO HERE

Serdar Yüksel (S'02, M'11) received his B.Sc. degree in Electrical and Electronics Engineering from Bilkent University in 2001; M.S. and Ph.D. degrees in Electrical and Computer Engineering from the University of Illinois at Urbana- Champaign in 2003 and 2006, respectively. He was a post-doctoral researcher at Yale University for a year before joining Queen's University as an Assistant Professor of Mathematics and Engineering in the Department of Mathematics and Statistics, where he is now an Associate Professor. He has been awarded the 2013 CAIMS/PIMS Early Career Award in Applied Mathematics. His research interests are on stochastic control, decentralized control, information theory and applied probability. He is an Associate Editor for the IEEE TRANSACTIONS ON AUTOMATIC CONTROL and Automatica.

PLACE PHOTO HERE

Prof. Tamás Linder (S'92, M'93, SM'00, F'13) received the M.S. degree in electrical engineering from the Technical University of Budapest, Hungary, in 1988, and the Ph.D degree in electrical engineering from the Hungarian Academy of Sciences in 1992. He was a post-doctoral researcher at the University of Hawaii in 1992 and a Visiting Fulbright Scholar at the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign during 1993-1994. From 1994 to 1998 he was a faculty member in the Department of Computer Science and Information Theory at the Technical University of Budapest. From 1996 to 1998 he was also a visiting research scholar in the Department of Electrical and Computer Engineering, University of California, San Diego. In 1998 he joined Queen's University where he is now a Professor of Mathematics and Engineering in the Department of Mathematics and Statistics. His research interests include communications and information theory, source coding and vector quantization, machine learning, and statistical pattern recognition. Dr. Linder received the Premier's Research Excellence Award of the Province of Ontario in 2002 and the Chancellor's Research Award of Queen's University in 2003. He was an Associate Editor for Source Coding of the IEEE TRANSACTIONS ON INFORMATION THEORY in 2003-2004.