

On Optimal Zero-Delay Coding of Vector Markov Sources

Tamás Linder, *Fellow, IEEE*, and Serdar Yüksel *Member, IEEE*

Abstract—Optimal zero-delay coding (quantization) of a vector-valued Markov source driven by a noise process is considered. Using a stochastic control problem formulation, the existence and structure of optimal quantization policies are studied. For a finite-horizon problem with bounded per-stage distortion measure, the existence of an optimal zero-delay quantization policy is shown provided that the quantizers allowed are ones with convex codecells. The bounded distortion assumption is relaxed to cover cases that include the linear quadratic Gaussian problem. For the infinite horizon problem and a stationary Markov source, the optimality of deterministic Markov coding policies is shown. The existence of optimal stationary Markov quantization policies is also shown provided randomization that is shared by the encoder and the decoder is allowed.

Index Terms—Real-time source coding, Markov source, quantization, stochastic control, Markov decision processes.

I. INTRODUCTION

A. Zero-Delay Coding

WE CONSIDER a zero-delay (sequential) encoding problem where a sensor encodes an observed information source without delay. It is assumed that the information source $\{x_t\}_{t \geq 0}$ is a time-homogenous \mathbb{R}^d -valued discrete-time Markov process. The initial distribution π_0 (i.e., the distribution of x_0) and the transition kernel $P(dx_{t+1}|x_t)$ uniquely determine the distribution of $\{x_t\}_{t \geq 0}$. In Assumption 1 below we will make explicit assumptions about the transition kernel.

The encoder encodes (quantizes) the source samples and transmits the encoded versions to a receiver over a discrete noiseless channel with input and output alphabet $\mathbb{M} := \{1, 2, \dots, M\}$, where M is a positive integer. Formally, the encoder is specified by a *quantization policy* Π , which is a sequence of Borel measurable functions $\{\eta_t\}_{t \geq 0}$ with $\eta_t : \mathbb{M}^t \times (\mathbb{R}^d)^{t+1} \rightarrow \mathbb{M}$. At time t , the encoder transmits the \mathbb{M} -valued message

$$q_t = \eta_t(I_t)$$

with $I_0 = x_0$, $I_t = (q_{[0,t-1]}, x_{[0,t]})$ for $t \geq 1$, where we have used the notation $q_{[0,t-1]} = (q_0, \dots, q_{t-1})$ and $x_{[0,t]} = (x_0, x_1, \dots, x_t)$. The collection of all such zero-delay policies is called the set of admissible quantization policies and is denoted by Π_A .

Manuscript received January 30, 2014; revised August 3, 2014; accepted August 5, 2014. Date of publication August 12, 2014; date of current version September 11, 2014. This work was supported by the Natural Sciences and Engineering Research Council of Canada. This paper was presented at the 51st IEEE Conference on Decision and Control in 2012 and the 2013 Workshop on Sequential and Adaptive Information Theory.

The authors are with the Department of Mathematics and Statistics, Queen’s University, Kingston, ON K7L 3N6, Canada (e-mail: linder@mast.queensu.ca; yuksel@mast.queensu.ca).

Communicated by J. Chen, Associate Editor for Shannon Theory. Digital Object Identifier 10.1109/TIT.2014.2346780

Observe that for fixed $q_{[0,t-1]}$ and $x_{[0,t-1]}$, as a function of x_t , the encoder $\eta_t(q_{[0,t-1]}, x_{[0,t-1]}, \cdot)$ is a *quantizer*, i.e., a Borel measurable mapping of \mathbb{R}^d into the finite set \mathbb{M} . Thus any quantization policy at each time $t \geq 0$ selects a quantizer $Q_t : \mathbb{R}^d \rightarrow \mathbb{M}$ based on past information $(q_{[0,t-1]}, x_{[0,t-1]})$, and then “quantizes” x_t as $q_t = Q_t(x_t)$.

Upon receiving q_t , the receiver generates its reconstruction u_t , also without delay. A zero-delay receiver policy is a sequence of measurable functions $\gamma = \{\gamma_t\}_{t \geq 0}$ of type $\gamma_t : \mathbb{M}^{t+1} \rightarrow \mathbb{U}$, where \mathbb{U} denotes the reconstruction alphabet (usually a Borel subset of \mathbb{R}^d). Thus

$$u_t = \gamma_t(q_{[0,t]}, t), \quad t \geq 0.$$

For the finite horizon setting the goal is to minimize the average cumulative cost (distortion)

$$J_{\pi_0}(\Pi, \gamma, T) := E_{\pi_0}^{\Pi, \gamma} \left[\frac{1}{T} \sum_{t=0}^{T-1} c_0(x_t, u_t) \right], \quad (1)$$

for some $T \geq 1$, where $c_0 : \mathbb{R}^d \times \mathbb{U} \rightarrow \mathbb{R}$ is a *nonnegative* Borel measurable cost (distortion) function. Here $E_{\pi_0}^{\Pi, \gamma}$ denotes expectation under initial distribution π_0 for x_0 ; the superscript signifies that the argument is a function of $\{x_t\}$ which depends on the quantization policy Π and receiver policy γ . (Later we will use the notation $E_{\pi_0}^{\Pi}$ for expectations where the argument is a function of $\{x_t\}$ that depends only on Π , and the notation E_{π_0} when the argument has no dependence on either Π or γ .) We assume that the encoder and decoder know the initial distribution π_0 .

We also consider the infinite-horizon average cost problem where the objective is to minimize

$$\begin{aligned} J_{\pi_0}(\Pi, \gamma) &:= \limsup_{T \rightarrow \infty} J_{\pi_0}(\Pi, \gamma, T) \\ &= \limsup_{T \rightarrow \infty} E_{\pi_0}^{\Pi, \gamma} \left[\frac{1}{T} \sum_{t=0}^{T-1} c_0(x_t, u_t) \right]. \end{aligned} \quad (2)$$

Our main assumption on the Markov source $\{x_t\}$ is the following.

Assumption 1: The evolution of $\{x_t\}$ is given by

$$x_{t+1} = f(x_t, w_t), \quad t = 0, 1, 2, \dots, \quad (3)$$

where $f : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ is a Borel function and $\{w_t\}$ is an independent and identically distributed (i.i.d.) vector noise sequence which is independent of x_0 . It is assumed that for each fixed $x \in \mathbb{R}^d$, the distribution of $f(x, w_t)$ admits the (conditional) density function $\phi(\cdot | x)$ (with respect to the d -dimensional Lebesgue measure) which is positive everywhere. Furthermore, $\phi(\cdot | x)$ is bounded and Lipschitz uniformly in x .

The above model includes the linear systems with Gaussian noise. Further conditions on f and the cost c_0 , and the reconstruction alphabet \mathbf{U} will be given in Sections III and IV for the finite-horizon problem (these include the case of a linear system and quadratic cost) and in Section V for the infinite-horizon problem.

Before proceeding further with formulating the results, we provide an overview of structural results for finite-horizon optimal zero-delay coding problems as well as a more general literature review.

B. Revisiting Structural Results for Finite-Horizon Problems

Structural results for the finite horizon control problem described in the previous section have been developed in a number of important papers. Among these the classic works by Witsenhausen [37] and Walrand and Varaiya [35], using two different approaches, are of particular relevance. Teneketzis [34] extended these approaches to the more general setting of non-feedback communication and [38] extended these results to more general state spaces (including \mathbb{R}^d). The following two theorems summarize, somewhat informally, these two important structural results.

Theorem 1 (Witsenhausen [37]): For the finite horizon problem, any zero-delay quantization policy $\Pi = \{\eta_t\}$ can be replaced, without any loss in performance, by a policy $\hat{\Pi} = \{\hat{\eta}_t\}$ which only uses $q_{[0,t-1]}$ and x_t to generate q_t , i.e., such that $q_t = \hat{\eta}_t(q_{[0,t-1]}, x_t)$ for all $t = 1, \dots, T-1$.

For a complete and separable (Polish) metric space \mathbf{X} and its Borel sets $\mathcal{B}(\mathbf{X})$, let $\mathcal{P}(\mathbf{X})$ denote the space of probability measures on $(\mathbf{X}, \mathcal{B}(\mathbf{X}))$, endowed with the topology of weak convergence (weak topology). This topology is metrizable with the Prokhorov metric making $\mathcal{P}(\mathbf{X})$ itself a Polish space. Given a quantization policy Π , for all $t \geq 1$ let $\pi_t \in \mathcal{P}(\mathbb{R}^d)$ be the regular conditional probability defined by

$$\pi_t(A) := P(x_t \in A | q_{[0,t-1]})$$

for any Borel set $A \subset \mathbb{R}^d$.

The following result is due to Walrand and Varaiya [35] who considered sources taking values in a finite set. For the more general case of \mathbb{R}^d -valued sources the result appeared in [38].

Theorem 2: For the finite horizon problem, any zero-delay quantization policy can be replaced, without any loss in performance, by a policy which at any time $t = 1, \dots, T-1$ only uses the conditional probability measure $\pi_t = P(dx_t | q_{[0,t-1]})$ and the state x_t to generate q_t . In other words, at time t such a policy $\hat{\eta}_t$ uses π_t to select a quantizer $Q_t = \hat{\eta}_t(\pi_t)$ (where $Q_t : \mathbb{R}^d \rightarrow \mathbf{M}$), and then q_t is generated as $q_t = Q_t(x_t)$.

A policy of the type suggested by Theorem 2 (a so-called Walrand-Varaiya-type policy) is called *stationary* if $\hat{\eta}_t = \hat{\eta}$ for all t , where $\hat{\eta}$ is a fixed policy mapping elements of $\mathcal{P}(\mathbb{R}^d)$ to the set of M -level quantizers. Stationary policies will play an important role in Section V.

As discussed in [38], the main difference between the two structural results above is the following: In the setup of Theorem 1, the encoder's memory space is not fixed and keeps expanding as the encoding block length T increases.

In the setup of Theorem 2, the memory space of an optimal encoder is fixed (note that π_t can be computed from π_{t-1} , Q_{t-1} , and q_{t-1} ; see equation (4)). Of course, in general the space of probability measures is a very large one. However, it may be the case that different quantization outputs lead to the same conditional probabilities π_t , leading to a reduction in the required memory. More importantly, the setup of Theorem 2 allows one to apply the powerful theory of Markov Decision Processes on fixed state and action spaces, thus greatly facilitating the analysis.

In this paper, we show that under quite general assumptions on the Markov process, the cost function, and the admissible quantization policies there always exists a Walrand-Varaiya-type policy that minimizes the finite horizon cost (1). For the infinite horizon problem (2), we show that there exists an optimal Walrand-Varaiya-type policy if the source is stationary. We also show that in general an optimal (possibly randomized) stationary quantization policy exists in the set of Walrand-Varaiya-type policies.

The rest of the paper is organized as follows. The next section gives a brief review of the literature. Section II contains background material on quantizers and the construction of a controlled Markov chain for our problem. Section III establishes the existence of optimal policies for the finite horizon case for bounded cost functions. Section IV considers the quadratic costs under conditions that cover linear systems. Section V considers the more involved infinite horizon case. Section VI contains concluding discussions. Most of the proofs are relegated to the Appendix.

C. Literature Review and Contributions

The existence of optimal quantizers for a one-stage ($T = 1$) cost problem has been investigated in [1], [30], and [41], among other works.

An important inspiration for our work is Borkar *et al.* [11] which studied the optimal zero-delay quantization of Markov sources. For the infinite horizon setting, this paper provided a stochastic control formulation of the optimal quantization problem with a Lagrangian cost that combined squared distortion and instantaneous entropy, and gave an elegant proof for the existence of optimal policies.

It should be noted that [11] restricted the admissible quantizers Q_t at each time stage t to so-called nearest neighbor quantizers whose reconstruction values were also suboptimally constrained to lie within a fixed compact set. Furthermore, some fairly restrictive conditions were placed on the dynamics of the system. These include requirements on the system dynamics that rule out additive noise models with unbounded support such as the Gaussian noise (see [11, p. 138]), and a uniform Lipschitz condition on the cost functions (see the condition on \hat{f} on p. 140 in [11]). These conditions made it possible to apply the discounted cost approach (see [3]) to average cost optimization problems.

Furthermore, the encoder-decoder structure in [11] has been specified *a priori*, whereas in this paper, we only relax global optimality when we restrict the quantizers to have convex codecells (to be defined later), which is a more general

condition than assuming the nearest neighbor encoding rule. On the other hand, we are unable to claim the optimality of deterministic stationary quantization policies for the infinite-horizon problem, whereas [11] establishes optimality of such policies. However, as mentioned, the conditions on the cost functions, systems dynamics, and the uniform continuity condition over all quantizers are not required in our setting.

To our knowledge, the existence of optimal quantizers for a finite horizon setting has not been considered in the literature for the setup considered in this paper.

Other relevant work include [10] which considered optimization over probability measures for causal and non-causal settings, and [23], [24], [34] and [38] which considered zero-delay coding of Markov sources in various setups. Structural theorems for zero-delay variable-rate coding of discrete Markov sources were studied in [21]. Recently [5] considered the average cost optimality equation for coding of discrete i.i.d. sources with limited lookahead and [20] studied real-time joint source-channel coding of a discrete Markov source over a discrete memoryless channel with feedback.

A different model for sequential source coding, called causal source coding, is studied in [22], [29] and [36]. In causal coding, the reconstruction depends causally on the source symbols, but in the information transmission process large delays are permitted, which makes this model less stringent (and one might argue less practical) than zero or limited-delay source coding.

For systems with control, structural results have also been investigated in the literature. In particular, for linear systems with quadratic cost criteria (known as LQG optimal control problems), it has been shown that the effect of the control policies can be decoupled from the estimation error without any loss. Under optimal control policies, [39] has shown the equivalence with the control-free setting considered in this paper (see also [6] and [28] for related results in different structural forms, where in contrast with [39] the encoders have memory). We also note that the design results developed here can be used to establish the existence of optimal quantization and control policies for LQG systems [39].

Contributions: In view of the literature review, the main contributions of the paper can be summarized as follows.

- (i) We establish a useful topology on the set of quantizers, building on [41], among other works, and show the existence of optimal coding policies for finite horizon optimization problems, under the assumption that the quantizers used have convex codecells. Notably, the set of sources considered includes LQG systems, i.e., linear systems driven by Gaussian noise under the quadratic cost criterion. The analysis requires the development of a series of technical results which facilitate establishing measurable selection criteria, reminiscent of those in [17].
- (ii) We establish, for the first time to our knowledge, the optimality of Markov (i.e., Walrand-Varaiya type) coding policies for *infinite-horizon* sequential quantization problems, using a new approach. The prior work reviewed above strictly build on dynamic programming (which is only suitable for finite-horizon problems) or does not

consider the question of global optimality of Markov policies.

- (iii) We show the existence of optimal stationary, possibly randomized, policies for a large class of sources including LQG systems. As detailed above, the assumptions are weaker than those that have appeared in prior work.

II. QUANTIZER ACTIONS AND CONTROLLED MARKOV PROCESS CONSTRUCTION

In this section, we formally define the space of quantizers considered in the paper building on the construction in [41]. Recall the notation $\mathbf{M} = \{1, \dots, M\}$.

Definition 1: An M -cell quantizer Q on \mathbb{R}^d is a (Borel) measurable mapping $Q : \mathbb{R}^d \rightarrow \mathbf{M}$. We let \mathcal{Q} denote the collection of all M -cell quantizers on \mathbb{R}^d .

Note that each $Q \in \mathcal{Q}$ is uniquely characterized by its *quantization cells* (or bins) $B_i = Q^{-1}(i) = \{x : Q(x) = i\}$, $i = 1, \dots, M$ which form a measurable partition of \mathbb{R}^d .

Remark 1:

- (i) We allow for the possibility that some of the cells of the quantizer are empty.
- (ii) In source coding theory (see [15]), a quantizer is a mapping $Q : \mathbb{R}^d \rightarrow \mathbb{R}^d$ with a finite range. In this definition, Q is specified by a partition $\{B_1, \dots, B_M\}$ of \mathbb{R}^d and reconstruction values $\{c_1, \dots, c_M\} \subset \mathbb{R}^d$ through the mapping rule $Q(x) = c_i$ if $x \in B_i$. In our definition, we do not include the reconstruction values.

In view of Theorem 2, any admissible quantization policy can be replaced by a Walrand-Varaiya-type policy. The class of all such policies is denoted by Π_W and is formally defined as follows.

Definition 2: An (admissible) quantization policy $\Pi = \{\eta_t\}$ belongs to Π_W if there exist a sequence of mappings $\{\hat{\eta}_t\}$ of the type $\hat{\eta}_t : \mathcal{P}(\mathbb{R}^d) \rightarrow \mathcal{Q}$ such that for $Q_t = \hat{\eta}_t(\pi_t)$ we have $q_t = Q_t(x_t) = \eta_t(I_t)$.

Suppose we use a quantizer policy $\Pi = \{\hat{\eta}_t\}$ in Π_W . Let $P(dx_{t+1}|x_t)$ denote the transition kernel of the process $\{x_t\}$ determined by the system dynamics (3) and note that $P(q_t|\pi_t, x_t)$ is determined by the quantizer policy as $P(q_t|\pi_t, x_t) = 1_{\{Q_t(x_t)=q_t\}}$, where $Q_t = \hat{\eta}_t(\pi_t)$ and 1_A denotes the indicator of event A . Then standard properties of conditional probability can be used to obtain the following filtering equation for the evolution of π_t :

$$\begin{aligned} \pi_{t+1}(dx_{t+1}) &= \frac{P(dx_{t+1}, q_t|q_{[0,t-1]})}{P(q_t|q_{[0,t-1]})} \\ &= \frac{\int_{\mathbb{R}^d} \pi_t(dx_t) P(q_t|\pi_t, x_t) P(dx_{t+1}|x_t)}{\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \pi_t(dx_t) P(q_t|\pi_t, x_t) P(dx_{t+1}|x_t)} \\ &= \frac{1}{\pi_t(Q^{-1}(q_t))} \int_{Q^{-1}(q_t)} P(dx_{t+1}|x_t) \pi_t(dx_t). \quad (4) \end{aligned}$$

Hence π_{t+1} is determined by π_t , Q_t , and q_t , which implies that π_{t+1} is conditionally independent of $(\pi_{[0,t-1]}, Q_{[0,t-1]})$ given π_t and Q_t . Thus $\{\pi_t\}$ can be viewed as $\mathcal{P}(\mathbb{R}^d)$ -valued controlled Markov process [17], [18] with \mathcal{Q} -valued control $\{Q_t\}$

and average cost up to time $T - 1$ given by

$$E_{\pi_0}^{\Pi} \left[\frac{1}{T} \sum_{t=0}^{T-1} c(\pi_t, Q_t) \right] = \inf_{\gamma} J_{\pi_0}(\Pi, \gamma, T),$$

where

$$c(\pi_t, Q_t) := \sum_{i=1}^M \inf_{u \in \mathcal{U}} \int_{Q_t^{-1}(i)} \pi_t(dx) c_0(x, u). \quad (5)$$

In this context, Π_W corresponds to the class of deterministic Markov control policies. Note that this definition of average cost assumes that the decoder uses an optimal receiver policy for encoder policy Π given by (5); thus our focus is on the encoding operation.

Recall that by Assumption 1 the density $\phi(\cdot|x)$ of $f(x, w_t)$ for fixed x is bounded, positive, and Lipschitz, uniformly in x . By (3) and (4) π_t admits a density, which we also denote by π_t , given by

$$\pi_t(z) = \int_{\mathbb{R}^d} \phi(z|x_{t-1}) P(dx_{t-1}|q_{[0,t-1]}), \quad z \in \mathbb{R}^d, \quad t \geq 1.$$

Thus for any policy Π , with probability 1 we have $0 < \pi_t(z) \leq C$ for all z and $t \geq 1$, where C is an upper bound on ϕ . Also, if $\phi(z|x)$ is Lipschitz in z with constant C_1 for all x , then the bound

$$\begin{aligned} & |\pi_t(z) - \pi_t(z')| \\ & \leq \int_{\mathbb{R}^d} |\phi(z|x_{t-1}) - \phi(z'|x_{t-1})| P(dx_{t-1}|q_{[0,t-1]}), \end{aligned}$$

implies that $\{\pi_t\}_{t \geq 1}$ is uniformly Lipschitz with constant C_1 . The collection of all densities with these properties will play an important part in this paper.

Definition 3: Let \mathcal{S} denote the set of all probability measures on \mathbb{R}^d admitting densities that are bounded by C and Lipschitz with constant C_1 .

Note that viewed as a class of densities, \mathcal{S} is uniformly bounded and equicontinuous. Lemma 3 in the Appendix shows that \mathcal{S} is closed in $\mathcal{P}(\mathbb{R}^d)$. Also, the preceding argument implies the following useful lemma.

Lemma 1: For any policy $\Pi \in \Pi_W$, we have $\pi_t \in \mathcal{S}$ for all $t \geq 1$ with probability 1.

For technical reasons in most of what follows we restrict the set of quantizers by only allowing ones that have convex cells. Formally, this quantizer class \mathcal{Q}_c is defined by

$$\mathcal{Q}_c = \{Q \in \mathcal{Q} : Q^{-1}(i) \subset \mathbb{R}^d \text{ is convex for } i = 1, \dots, M\},$$

where by convention we declare the empty set convex. Note that each nonempty cell of a $Q \in \mathcal{Q}_c$ is a convex polytope in \mathbb{R}^d . The class of policies $\Pi_W^{\mathcal{Q}_c}$ is obtained by replacing \mathcal{Q} with \mathcal{Q}_c in Definition 2:

Definition 4: $\Pi_W^{\mathcal{Q}_c}$ denotes the set of all quantization policies $\Pi = \{\hat{\eta}_t\} \in \Pi_W$ such that $\hat{\eta}_t : \mathcal{P}(\mathbb{R}^d) \rightarrow \mathcal{Q}_c$, i.e., $Q_t = \hat{\eta}_t(\pi_t) \in \mathcal{Q}_c$ for all $t \geq 0$.

Remark 2:

- (i) The assumption of convex codecells is adopted for technical reasons: the structure of \mathcal{Q}_c detailed below will let us endow it with a well-behaved topology. \mathcal{Q}_c is a fairly powerful class; for example, it includes as a proper

subset the class of nearest-neighbor quantizers considered in [11]. Furthermore, it was proved in [16] that \mathcal{Q}_c contains all M -level optimal entropy-constrained quantizers when the source has a density, while the set of all M -level nearest neighbor quantizers is clearly suboptimal in this sense. On the other hand, it is likely that the convex codecell assumption results in a loss of system optimality in our case. This can be conjectured from the results of [2] where it was shown that in multiresolution quantization, requiring that quantizers have convex codecells may preclude system optimality even for continuous sources. We note that the convex codecell assumption is often made when provably optimal and fast algorithms are sought for the design of multiresolution, multiple description, and Wyner-Ziv quantizers; see [13] and [27].

- (ii) As opposed to general quantizers in \mathcal{Q} , any $Q \in \mathcal{Q}_c$ has a parametric representation. Let such a Q have cells $\{B_1, \dots, B_M\}$. As discussed in [16], by the separating hyperplane theorem, there exist pairs of complementary closed half spaces $\{(H_{i,j}, H_{j,i}) : 1 \leq i, j \leq M, i \neq j\}$ such that $B_i \subset \bigcap_{j \neq i} H_{i,j}$ for all i . Since $\bar{B}_i := \bigcap_{j \neq i} H_{i,j}$ is a closed convex polytope for each i , if $P \in \mathcal{P}(\mathbb{R}^d)$ admits a density, then $P(\bar{B}_i \setminus B_i) = 0$ for all i . We thus obtain a P -almost sure representation of Q by the $M(M-1)/2$ hyperplanes $h_{i,j} = H_{i,j} \cap H_{j,i}$. One can represent such a hyperplane h by a vector $(a_1, \dots, a_d, b) \in \mathbb{R}^{d+1}$ with $\sum_k |a_k|^2 = 1$ such that $h = \{x \in \mathbb{R}^d : \sum_k a_k x_k = b\}$, thus obtaining a parametrization over $\mathbb{R}^{(d+1)M(M-1)/2}$ of all quantizers in \mathcal{Q}_c .

In order to facilitate the stochastic control analysis of the quantization problem we need an alternative representation of quantizers. As discussed in [10] and [41], a quantizer Q with cells $\{B_1, \dots, B_M\}$ can also be identified with the stochastic kernel (regular conditional probability), also denoted by Q , from \mathbb{R}^d to \mathbf{M} defined by

$$Q(i|x) = 1_{\{x \in B_i\}}, \quad i = 1, \dots, M.$$

We will endow the set of quantizers \mathcal{Q}_c with a topology induced by the stochastic kernel interpretation. If P is a probability measure on \mathbb{R}^d and Q is a stochastic kernel from \mathbb{R}^d to \mathbf{M} , then PQ denotes the resulting joint probability measure on $\mathbb{R}^d \times \mathbf{M}$ defined through $PQ(dx dy) = P(dx)Q(dy|x)$. For some fixed $P \in \mathcal{P}(\mathbb{R}^d)$ let

$$\Gamma_P := \{PQ \in \mathcal{P}(\mathbb{R}^d \times \mathbf{M}) : Q \in \mathcal{Q}_c\}$$

It follows from [41, Theorem 5.8] that Γ_P is a compact subset of $\mathcal{P}(\mathbb{R}^d \times \mathbf{M})$ if P admits a density. If we introduce the equivalence relation $Q \equiv Q'$ if and only if $PQ = PQ'$, then the resulting set of equivalence classes, denoted by $(\mathcal{Q}_c)_P$, can be equipped with the quotient topology inherited from Γ_P . In this topology $Q_n \rightarrow Q$ if and only if (for representatives of the equivalence classes) $PQ_n \rightarrow PQ$ weakly. Also, if $PQ = PQ'$ for P admitting a positive density, then the (convex polytopal) cells of Q and Q' may differ only in their boundaries, and it follows that $(\mathcal{Q}_c)_P = (\mathcal{Q}_c)_{P'}$ for any P' also admitting a positive density. From now on we will identify \mathcal{Q}_c with $(\mathcal{Q}_c)_P$ and endow it with the resulting quotient topology,

keeping in mind that this definition does not depend on P as long as it has a positive density. Lemma 3 in the Appendix shows that \mathcal{Q}_c is compact. Note that π_t for $t \geq 1$ always has a positive density due to Assumption 1. However, in some of the results we will allow π_0 to violate this assumption (e.g., by letting π_0 be a point mass at a given $x_0 \in \mathbb{R}^d$).

For a given policy $\Pi \in \Pi_W^C$, we will consider $\{(\pi_t, Q_t)\}$ as an $\mathcal{S} \times \mathcal{Q}_c$ -valued process.

III. EXISTENCE OF OPTIMAL POLICIES: FINITE HORIZON SETTING

For any quantization policy Π in Π_W and any $T \geq 1$ we define

$$J_{\pi_0}(\Pi, T) := \inf_{\gamma} J_{\pi_0}(\Pi, \gamma, T) = E_{\pi_0}^{\Pi} \left[\frac{1}{T} \sum_{t=0}^{T-1} c(\pi_t, Q_t) \right],$$

where $c(\pi_t, Q_t)$ is defined in (5).

Assumption 2:

(i) The cost $c_0 : \mathbb{R}^d \times \mathbf{U} \rightarrow \mathbb{R}$ is nonnegative, bounded, and continuous.

(ii) \mathbf{U} is compact.

Theorem 3: Under Assumptions 1 and 2 an optimal receiver policy always exists, i.e., for any $\Pi \in \Pi_W$ there exist $\gamma = \{\gamma_t\}$ such that $J_{\pi_0}(\Pi, \gamma, T) = J_{\pi_0}(\Pi, T)$.

Proof: At any $t \geq 0$ an optimal receiver has to minimize $\int P(dx_t | q_{[0,t]}) c_0(x_t, u)$ in u . Under Assumption 2, the existence of a minimizer then follows from a standard argument, see, e.g., [41, Theorem 3.1]. \square

The following result states the existence of optimal policies in \mathcal{Q}_c for the finite horizon setting. The proof is given in Section VII-B of the Appendix.

Theorem 4: Suppose π_0 admits a density or it is a point mass $\pi_0 = \delta_{x_0}$ for some $x_0 \in \mathbb{R}^d$. For any $T \geq 1$, under Assumptions 1 and 2, there exists a policy Π in Π_W^C such that

$$J_{\pi_0}(\Pi, T) = \inf_{\Pi' \in \Pi_W^C} J_{\pi_0}(\Pi', T). \quad (6)$$

Let $J_T^T(\cdot) := 0$ and define $J_t^T(\pi)$ for $t = T-1, T-2, \dots, 1$, $\pi \in \mathcal{S}$ and $t = 0$, $\pi = \pi_0$, by the dynamic programming recursion

$$J_t^T(\pi) = \min_{Q \in \mathcal{Q}_c} \left(\frac{1}{T} c(\pi, Q) + E[J_{t+1}^T(\pi_{t+1}) | \pi_t = \pi, Q_t = Q] \right). \quad (7)$$

Then

$$J_0^T(\pi_0) = \min_{\Pi \in \Pi_W^C} J_{\pi_0}(\Pi, T).$$

IV. THE FINITE HORIZON PROBLEM FOR QUADRATIC COST

Linear systems driven by Gaussian noise are important in many applications in control, estimation, and signal processing. For such linear systems with quadratic cost (known as LQG optimal control problems), it has been shown that the effect of the control policies can be decoupled from the estimation error without any loss (see [33], [39] and for a review [40]). In this section we consider the finite

horizon problem under conditions that cover LQG systems. Let $\|x\|$ denote the Euclidean norm of $x \in \mathbb{R}^d$. We replace Assumption 2 of the preceding sections with the following.

Assumption 3:

(i) The function f in the system dynamics (3) satisfies $\|f(x, w)\| \leq K(\|x\| + \|w\|)$ for some $K > 0$ and all $x, w \in \mathbb{R}^d$.

(ii) $\mathbf{U} = \mathbb{R}^d$ and the cost is given by $c_0(x, u) = \|x - u\|^2$.

(iii) The common distribution ν_w of the w_t satisfies $\int \|z\|^2 \nu_w(dz) < \infty$.

(iv) π_0 admits a density such that $E_{\pi_0}[\|x_0\|^2] < \infty$ or it is a point mass $\pi_0 = \delta_{x_0}$.

Remark 3:

(i) The above conditions cover the case of a linear-Gaussian system

$$x_{t+1} = Ax_t + w_t, \quad t = 0, 1, 2, \dots,$$

where $\{w_t\}$ is an i.i.d. Gaussian noise sequence with zero mean, A is a square matrix, and π_0 admits a Gaussian density having zero mean.

(ii) Assumption 3(i) implies

$$\|x_t\|^2 \leq \hat{K} \left(\|x_0\|^2 + \sum_{i=0}^{t-1} \|w_i\|^2 \right)$$

for some \hat{K} that depends on t (see (34)). Together with Assumptions 3(iii) and (iv), this implies $E_{\pi_0}[\|x_t\|^2] < \infty$ for all $t \geq 0$ under any quantization policy. Therefore

$$\int_{\mathbb{R}^d} \|x_t\|^2 P(dx_t | q_{[0,t]}) dx < \infty$$

and an optimal receiver policy exists and is given by

$$\gamma_t(q_{[0,t]}) = \int_{\mathbb{R}^d} x_t P(dx_t | q_{[0,t]}). \quad (8)$$

The following is a restatement of Theorem 4 under conditions that allow unbounded cost. The proof is relegated to Section VII-C of the Appendix.

Theorem 5: Under Assumptions 1 and 3, for any $T \geq 1$ there exists an optimal policy in Π_W^C in the sense of (6) and the dynamic programming recursion (7) for $J_t^T(\pi_t)$ also holds.

V. INFINITE HORIZON SETTING

For the infinite horizon setting, one may consider the discounted cost problem where the goal is to find policies that achieve

$$J_{\pi_0}^{\beta}(\pi_0) = \inf_{\Pi \in \Pi_W^C} J_{\pi_0}^{\beta}(\Pi) \quad (9)$$

for some $\beta \in (0, 1)$, where

$$J_{\pi_0}^{\beta}(\Pi) = \inf_{\gamma} \lim_{T \rightarrow \infty} E_{\pi_0}^{\Pi, \gamma} \left[\sum_{t=0}^{T-1} \beta^t c_0(x_t, u_t) \right].$$

The existence of optimal policies for this problem follows from the results in the previous section. In particular, it is well known that the value iteration algorithm (see [25]) will converge to an optimal solution, since the cost function is

bounded and the measurable selection hypothesis is applicable in view of Theorem 4. This leads to the fixed point equation

$$V^\beta(\pi) = \min_{Q \in \mathcal{Q}_c} \left(c(\pi, Q) + \beta \int_{\mathbb{R}^d} P(d\pi_{t+1} | \pi_t = \pi, Q_t = Q) V^\beta(\pi_{t+1}) \right).$$

The more challenging case is the average cost problem where one considers

$$J_{\pi_0}(\Pi) = \inf_{\gamma} \limsup_{T \rightarrow \infty} E_{\pi_0}^{\Pi, \gamma} \left[\frac{1}{T} \sum_{t=0}^{T-1} c_0(x_t, u_t) \right] \quad (10)$$

and the goal is to find an optimal policy attaining

$$J_{\pi_0} := \inf_{\Pi \in \Pi_A} J_{\pi_0}(\Pi). \quad (11)$$

For the infinite horizon setting the structural results in Theorems 1 and 2 are not available in the literature, due to the fact that the proofs are based on dynamic programming which starts at a finite terminal time stage and optimal policies are computed backwards. Recall that $\pi^* \in \mathcal{P}(\mathbb{R}^d)$ is called an invariant measure for $\{x_t\}$ if setting $\pi_0 = \pi^*$ results in $P(x_t \in B) = \pi^*(B)$ for every t and Borel set B (in this case $\{x_t\}$ is a strictly stationary process). The next result proves an infinite-horizon analog of Theorem 2 under the assumption that an invariant measure π^* for $\{x_t\}$ exists and $\pi_0 = \pi^*$.

Theorem 6: Assume the cost c_0 is bounded and an invariant measure π^* exists. If $\{x_t\}$ starts from π^* , then there exists an optimal policy in Π_W solving the minimization problem (11), i.e., there exists $\Pi \in \Pi_W$ such that

$$\limsup_{T \rightarrow \infty} E_{\pi^*}^{\Pi} \left[\frac{1}{T} \sum_{t=0}^{T-1} c(\pi_t, Q_t) \right] = J_{\pi^*}.$$

The proof of the theorem relies on a construction that pieces together policies from Π_W that on time segments of appropriately large lengths increasingly well approximate the minimum infinite-horizon cost achievable by policies in Π_A . Since the details are somewhat tedious, the proof is relegated to Section VII-D of the Appendix. We note that the condition that c_0 is bounded is not essential and, for example, the theorem holds for the quadratic cost if the invariant measure has a finite second moment.

Remark 4:

- (i) If the source is a positive Harris recurrent Markov chain [26], then the policy constructed in the proof of Theorem 6 achieves the optimal average cost corresponding to the stationary source even when the chain is not started from the invariant distribution π^* . This can be shown by inspecting the details of the proof and using the continuity of the *value function* as stated by Theorem 10 in the Appendix, combined with the fact that $P(x_t \in \cdot) \rightarrow \pi^*$ in total variation as $t \rightarrow \infty$ for any initial distribution π_0 .
- (ii) The proof of Theorem 6 demonstrates that for every $\epsilon > 0$ there exists a finite memory encoding policy whose performance is within ϵ of the optimum value J_{π^*} . This scheme can be computed using finite horizon dynamic programming, giving the result practical relevance.

The optimal policy constructed in the proof of Theorem 6 may not be stationary. In general, a stationary policy in a given class of policies is called optimal if it performs as well as any other policy in that class. In the next section we establish the existence of an optimal stationary policy in Π_W^C if randomization is allowed.

A. Classes of Randomized Quantization Policies

We will consider two classes of randomized policies.

1) *Randomized Walrand-Varaiya-Type (Markov) Policies:* These policies, denoted by $\bar{\Pi}_W^C$, are randomized over Π_W^C , the Walrand-Varaiya-type Markov policies with quantizers having convex cells (Definition 4). Each $\Pi \in \bar{\Pi}_W^C$ consists of a sequence of stochastic kernels $\{\bar{\eta}_t\}$ from $\mathcal{P}(\mathbb{R}^d)$ to \mathcal{Q}_c . Thus, under Π , for any $t \geq 0$,

$$\begin{aligned} P^\Pi(Q_t(x_t) = q_t | q_{[0,t-1]}, Q_{[0,t-1]}, \pi_{[0,t]}) \\ = \int_{\mathcal{Q}_c} \left(\int_{\mathbb{R}^d} 1_{\{Q(x)=q_t\}} \pi_t(dx) \right) \bar{\eta}_t(dQ | \pi_t). \end{aligned}$$

It follows from [14] or [31] that an equivalent model for randomization can be obtained by considering an i.i.d. randomization sequence $\{r_t\}$, independent of $\{x_t\}$ and uniformly distributed on $[0, 1]$, and a sequence of (measurable) randomized encoders $\{\hat{\eta}_t\}$ of the form $\hat{\eta}_t : \mathcal{P}(\mathbb{R}^d) \times [0, 1] \rightarrow \mathcal{Q}_c$ and $Q_t = \hat{\eta}_t(\pi_t, r_t)$. In this case the induced stochastic kernel encoder $\bar{\eta}_t$ is determined by

$$\bar{\eta}_t(D | \pi_t) = u \{ r : \hat{\eta}_t(\pi_t, r) \in D \}$$

for any Borel subset D of \mathcal{Q}_c , where u denotes the uniform distribution on $[0, 1]$. For randomized policies we assume that all the randomization information is shared between the encoder and the decoder, that is

$$I_t^r := (q_{[0,t-1]}, r_{[0,t-1]})$$

is known at the decoder which can therefore track π_t given by

$$\pi_t(A) := P(x_t \in A | q_{[0,t-1]}, r_{[0,t-1]})$$

for any Borel set $A \subset \mathbb{R}^d$.

We note that the cost $c(\pi_t, Q_t)$ is still defined by (5) since the decoder, having access to I_t^r can also track Q_t . Also, in computing the cost $E_{\pi_0}^{\Pi} \left[\frac{1}{T} \sum_{t=0}^{T-1} c(\pi_t, Q_t) \right]$ of policy $\Pi \in \bar{\Pi}_W^C$ after T time stages, the expectation is also taken with respect to the randomization sequence $\{r_t\}$.

2) *Randomized Stationary Walrand-Varaiya-Type (Markov) Policies:* Denoted by $\bar{\Pi}_{W,S}^C$, this class consists of all policies in $\bar{\Pi}_W^C$ that are stationary, i.e., the stochastic kernels $\bar{\eta}_t$ or the randomized encoders $\hat{\eta}_t$ do not depend on the time index t .

B. Existence of Optimal Stationary Policies

1) *The Bounded Cost Case:* In the infinite horizon setting, we add the following assumption, in addition to Assumptions 1 and 2.

Assumption 4: The chain $\{x_t\}$ is positive Harris recurrent (see [26]) with unique invariant measure π^* such that for all $x_0 \in \mathbb{R}^d$,

$$\lim_{t \rightarrow \infty} E_{\delta_{x_0}}[\|x_t\|^2] = \int_{\mathbb{R}^d} \|x\|^2 \pi^*(dx) < \infty.$$

Remark 5: A sufficient condition for Assumption 4 to hold is that $f(x, w)$ in (3) is continuous in x and satisfies $\|f(x, w)\| \leq K(\|x\| + \|w\|)$ for some $K < 1$, and w_t has zero mean and second moment $E[\|w_t\|^2] < \infty$. This follows since the upper bound on f and a straightforward calculation imply that the drift condition [26]

$$E[V(x_{t+1})|x_t = x] \leq V(x) - g(x) + b1_{\{x \in C\}}$$

holds with $V(x) = \|x\|^2$, $g(x) = (1 - (K^2 + \epsilon))\|x\|^2$, $0 < \epsilon < 1 - K^2$, and $C = \{\|x\| \leq M\}$ (a compact set), where

$$M = \frac{K^2 E[\|w\|] + \sqrt{(K^2 E[\|w\|])^2 + K^2 \epsilon E[\|w\|^2]}}{\epsilon}$$

and $b = K^2(E[\|w\|^2] + 2E[\|w\|]M)$. The continuity of $f(x, w)$ in x implies that the chain is weak Feller [26]. This and the drift condition imply through [42, Theorem 2.2] that there exists an invariant probability measure with a finite second moment. The irreducibility and aperiodicity [26] of the chain under Assumption 1 implies the uniqueness of the invariant probability measure and positive Harris recurrence, leading to Assumption 4.

To show the existence of an optimal stationary policy, we adopt the convex analytic approach of [8] (see [3] for a detailed discussion). Here we only present the essential steps.

Fix a policy $\Pi \in \bar{\Pi}_W^C$ and an initial distribution π_0 . Let $v_t \in \mathcal{P}(\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c)$ be the sequence of expected occupation measures determined by

$$v_t(D) = E_{\pi_0}^{\Pi} \left(\frac{1}{t} \sum_{i=0}^{t-1} 1_{\{(\pi_i, Q_i) \in D\}} \right)$$

for any Borel subset D of $\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c$.

Let $P(d\pi_{t+1}|\pi_t, Q_t) = P^{\Pi}(d\pi_{t+1}|\pi_t, Q_t)$ be the transition kernel determined by the filtering equation (4) and note that it does not depend on Π and t . Also note that $P(\mathcal{S}|\pi, Q) = 1$ for any π and Q , where $\mathcal{S} \subset \mathcal{P}(\mathbb{R}^d)$ is the set of probability measures, defined in Definition 3, which admit densities that satisfy the same upper bound and Lipschitz condition as the density of the additive noise w_t (\mathcal{S} contains the set of reachable states for $\{\pi_t\}$ under any quantization policy).

If \mathbf{X} is a topological space, let $\mathcal{C}_b(\mathbf{X})$ denote the set of all bounded and continuous real-valued functions on \mathbf{X} . Let \mathcal{G} be the set of so-called ergodic occupation measures on $\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c$, defined by

$$\mathcal{G} = \left\{ v \in \mathcal{P}(\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c) : \int f(\pi)v(d\pi dQ) = \int \int f(\pi')P(d\pi'|\pi, Q)v(d\pi dQ) \text{ for all } f \in \mathcal{C}_b(\mathcal{P}(\mathbb{R}^d)) \right\}.$$

Note that any $v \in \mathcal{G}$ is supported on $\mathcal{S} \times \mathcal{Q}_c$.

Any $v \in \mathcal{G}$ can be disintegrated as $v(d\pi dQ) = \hat{v}(d\pi)\bar{\eta}(dQ|\pi)$, where $\bar{\eta}$ is a stochastic kernel from $\mathcal{P}(\mathbb{R}^d)$

to \mathcal{Q}_c which corresponds to the randomized stationary policy $\Pi = \{\bar{\eta}_t\}$ in $\bar{\Pi}_{W, \mathcal{S}}^C$ such that $\bar{\eta}_t = \bar{\eta}$ for all t . Then the transition kernel of the process $\{(\pi_t, Q_t)\}$ induced by Π does not depend on t and is given by

$$P^{\Pi}(d\pi_{t+1} dQ_{t+1}|\pi_t, Q_t) = P(d\pi_{t+1}|\pi_t, Q_t)\bar{\eta}(dQ_{t+1}|\pi_t).$$

In fact, it directly follows from the definition of \mathcal{G} that

$$\begin{aligned} & \int g(\pi, Q)v(d\pi dQ) \\ &= \int \int g(\pi', Q')P^{\Pi}(d\pi' dQ'|\pi, Q)v(d\pi dQ) \end{aligned} \quad (12)$$

for all $g \in \mathcal{C}_b(\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c)$, i.e., v is an invariant measure for the transition kernel P^{Π} .

The following proposition, proved in Section VII-E, will imply the existence of optimal stationary policies.

Proposition 1: (a) For any initial distribution π_0 and policy $\Pi \in \bar{\Pi}_W^C$, if $\{v_{t_n}\}$ is a subsequence of the expected occupation measures $\{v_t\}$ such that $v_{t_n} \rightarrow \bar{v}$ weakly, then $\bar{v} \in \mathcal{G}$. Furthermore

$$\begin{aligned} & \lim_{n \rightarrow \infty} \int_{\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c} c(\pi, Q)v_{t_n}(d\pi dQ) \\ &= \int_{\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c} c(\pi, Q)\bar{v}(d\pi dQ). \end{aligned} \quad (13)$$

(b) For any $x_0 \in \mathbb{R}^d$, initial distribution $\pi_0 = \delta_{x_0}$, and policy $\Pi \in \bar{\Pi}_W^C$, $\{v_t\}$ is relatively compact.

(c) \mathcal{G} is compact.

For any initial distribution δ_{x_0} and policy $\Pi \in \bar{\Pi}_W$, we have

$$\begin{aligned} & \liminf_{T \rightarrow \infty} E_{\delta_{x_0}}^{\Pi} \left[\frac{1}{T} \sum_{t=0}^{T-1} c(\pi_t, Q_t) \right] \\ &= \liminf_{T \rightarrow \infty} \int_{\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c} c(\pi, Q)v_T(d\pi dQ). \end{aligned}$$

Let $\{v_{T_n}\}$ be a subsequence of $\{v_T\}$ such that

$$\begin{aligned} & \liminf_{T \rightarrow \infty} \int_{\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c} c(\pi, Q)v_T(d\pi dQ) \\ &= \lim_{n \rightarrow \infty} \int_{\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c} c(\pi, Q)v_{T_n}(d\pi dQ). \end{aligned}$$

By Proposition 1(b) there exists a subsequence of $\{v_{T_n}\}$, which we also denote by $\{v_{T_n}\}$, weakly converging to some \bar{v} . By Proposition 1(a) we have $\bar{v} \in \mathcal{G}$ and $\int c dv_{T_n} \rightarrow \int c d\bar{v}$. Therefore

$$\begin{aligned} & \liminf_{T \rightarrow \infty} E_{\delta_{x_0}}^{\Pi} \left[\frac{1}{T} \sum_{t=0}^{T-1} c(\pi_t, Q_t) \right] = \int_{\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c} c(\pi, Q)\bar{v}(d\pi dQ) \\ & \geq \inf_{v \in \mathcal{G}} \int_{\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c} c(\pi, Q)v(d\pi dQ). \end{aligned}$$

In addition, since c is continuous on $\mathcal{S} \times \mathcal{Q}_c$ (by Lemma 4) and each $v \in \mathcal{G}$ is supported on $\mathcal{S} \times \mathcal{Q}_c$, the mapping $v \mapsto \int c dv$ is continuous on \mathcal{G} . Since \mathcal{G} is compact by Proposition 1(c), there exists $v^* \in \mathcal{G}$ achieving the above infimum.

Hence

$$\begin{aligned} c^* &:= \int_{\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c} c(\pi, Q) v^*(d\pi dQ) \\ &= \min_{v \in \mathcal{G}} \int_{\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c} c(\pi, Q) v(d\pi, dQ) \end{aligned} \quad (14)$$

provides an ultimate lower bound on the infinite-horizon cost of any policy.

The following theorem shows the existence of a stationary policy achieving this lower bound if we consider the initial distribution π_0 as a “design parameter” we can freely choose.

Theorem 7: Under Assumptions 1, 2 and 4, there exists a stationary policy Π^ in $\bar{\Pi}_{W,S}^C$ that is optimal in the sense that with an appropriately chosen initial distribution π_0^* ,*

$$\lim_{T \rightarrow \infty} E_{\pi_0^*}^{\Pi^*} \left[\frac{1}{T} \sum_{t=0}^{T-1} c(\pi_t, Q_t) \right] \leq \liminf_{T \rightarrow \infty} E_{\delta_{x_0}^\Pi} \left[\frac{1}{T} \sum_{t=0}^{T-1} c(\pi_t, Q_t) \right]$$

for any $x_0 \in \mathbb{R}^d$ and $\Pi \in \bar{\Pi}_W^C$.

Proof: We must prove the existence of $\Pi^* \in \bar{\Pi}_{W,S}^C$ which achieves infinite horizon cost c^* for some initial distribution π_0^* . Consider v^* achieving the minimum in (14), disintegrate it as $v^*(d\pi dQ) = \hat{v}^*(d\pi) \bar{\eta}^*(dQ|\pi)$, and let $\Pi^* \in \bar{\Pi}_{W,S}^C$ be the policy corresponding to $\bar{\eta}^*$. Since v^* is an invariant measure for the transition kernel P^{Π^*} (see (12)), for any $T \geq 1$,

$$c^* = E_{\hat{v}^*}^{\Pi^*} \left[\frac{1}{T} \sum_{t=0}^{T-1} c(\pi_t, Q_t) \right],$$

where the notation $E_{\hat{v}^*}^{\Pi^*}$ signifies that the initial distribution π_0 is picked randomly with distribution \hat{v}^* . Thus

$$c^* = \lim_{T \rightarrow \infty} E_{\hat{v}^*}^{\Pi^*} \left[\frac{1}{T} \sum_{t=0}^{T-1} c(\pi_t, Q_t) \right]. \quad (15)$$

From the individual ergodic theorem (see [19]) the limit

$$f(\pi_0) := \lim_{T \rightarrow \infty} E_{\pi_0}^{\Pi^*} \left[\frac{1}{T} \sum_{t=0}^{T-1} c(\pi_t, Q_t) \right]$$

exists for \hat{v}^* -a.e. π_0 and

$$\int_{\mathcal{P}(\mathbb{R}^d)} f(\pi_0) \hat{v}^*(d\pi_0) = c^*.$$

Hence for some π_0 in the support of \hat{v}^* we must have

$$\lim_{T \rightarrow \infty} E_{\pi_0}^{\Pi^*} \left[\frac{1}{T} \sum_{t=0}^{T-1} c(\pi_t, Q_t) \right] \leq c^*$$

Any such π_0 can be picked as π_0^* so that the claim of the theorem holds. \square

In the preceding theorem the initial state distribution π_0 is a design parameter which is chosen along with the quantization policy to optimize the cost. This assumption may be unrealistic. However, consider the fictitious optimal stationary policy in (15) which is allowed to pick the initial distribution π_0 according to \hat{v}^* . It follows from the analysis in the proof of Proposition 1 (see (51)) that the expectation of π_0 according to \hat{v}^* is precisely the invariant distribution π^* for $\{x_t\}$. Based on this, one can prove the following, more realistic version of

the optimality result. The proof, which is not given here, is an expanded and more refined version of the proof of Theorem 7.

Theorem 8: Under the setup of Theorem 7, assume that $\{x_t\}$ is started from the invariant distribution π^ . If the optimal stationary policy $\Pi^* \in \bar{\Pi}_{W,S}^C$ is used in such a way that the encoder and decoder’s initial belief π_0 is picked randomly according to \hat{v}^* (but independently of $\{x_t\}$), then Π^* is still optimal in the sense of Theorem 7.*

Remark 6: We have not shown that an optimal stationary policy is deterministic. In the convex analytic approach, the existence of an optimal deterministic stationary policy directly follows if one can show that the extreme points of ergodic occupation measures satisfy the following: (i) They are induced by deterministic policies; and (ii) under these policies the state invariant measures are ergodic. This property of the extreme points of the set of ergodic occupation measures has been proved by Meyn in [25, Proposition 9.2.5] for countable state spaces and by Borkar in [4] and [8] for a specific case involving \mathbb{R}^d as the state space and a non-degeneracy condition which amounts to having a density assumption on the one-stage transition kernels. Unfortunately, these approaches do not seem to apply in our setting.

2) *The Quadratic Cost Case:* In the infinite horizon setting for the important case of the (unbounded) quadratic cost function, we add the following assumption, in addition to Assumption 3.

Assumption 5: The chain $\{x_t\}$ is positive Harris recurrent with unique invariant measure π^ such that for some $\epsilon > 0$ and all $x_0 \in \mathbb{R}^d$,*

$$\lim_{t \rightarrow \infty} E_{\delta_{x_0}} [\|x_t\|^{2+\epsilon}] = \int_{\mathbb{R}^d} \|x\|^{2+\epsilon} \pi^*(dx) < \infty.$$

Remark 7: Assumption 5 holds for the LQG case $x_{t+1} = Ax_t + w_t$, with A being a $d \times d$ matrix having eigenvalues of absolute value less than 1 and w_t having a nondegenerate Gaussian distribution with zero mean.

Theorem 9: Under Assumptions 3 and 5, there exists a stationary policy Π^ in $\bar{\Pi}_{W,S}^C$ that is optimal in the sense that with an appropriately chosen initial distribution π_0^* ,*

$$\lim_{T \rightarrow \infty} E_{\pi_0^*}^{\Pi^*} \left[\frac{1}{T} \sum_{t=0}^{T-1} c(\pi_t, Q_t) \right] \leq \liminf_{T \rightarrow \infty} E_{\delta_{x_0}^\Pi} \left[\frac{1}{T} \sum_{t=0}^{T-1} c(\pi_t, Q_t) \right]$$

for any $x_0 \in \mathbb{R}^d$ and $\Pi \in \bar{\Pi}_W^C$. Furthermore, if $\{x_t\}$ is started from the invariant distribution π^* and the optimal stationary policy $\Pi^* \in \bar{\Pi}_{W,S}^C$ is used in such a way that the encoder and decoder’s initial belief π_0 is picked randomly according to \hat{v}^* (but independently of $\{x_t\}$), then Π^* is still optimal in the above sense (with π^* replacing π_0^*).

Proof: The proof is almost identical to that of Theorems 7 and 8, with the following minor adjustments, which are needed to accommodate the unboundedness of the quadratic cost function. This modification is facilitated by Assumption 5 which implies that, similar to (49) and (52) in the proof of Proposition 1, for the sequence of expected occupation measures $\{v_t\}$ corresponding to any initial distribution δ_{x_0} , we have

$$\sup_{t \geq 0} \int_{\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c} \left(\int_{\mathbb{R}^d} \|x\|^{2+\epsilon} \pi(dx) \right) v_t(d\pi dQ) < \infty,$$

as well as for all $v \in \mathcal{G}$,

$$\begin{aligned} & \int_{P(\mathbb{R}^d) \times \mathcal{Q}_c} \left(\int_{\mathbb{R}^d} \|x\|^{2+\epsilon} \pi(dx) \right) v(d\pi dQ) \\ &= \int_{\mathbb{R}^d} \|x\|^{2+\epsilon} \pi^*(dx) < \infty. \end{aligned}$$

These uniform integrability properties of $\{v_t\}$ and \mathcal{G} allow us to use the continuity result Lemma 8 for $c(\pi, Q)$. All other parts of the proof remain unchanged. \square

VI. CONCLUDING REMARKS

In this paper we established structural and existence results concerning optimal quantization policies for Markov sources. The key ingredient of our analysis was the characterization of quantizers as a subset of the space of stochastic kernels. This approach allows one to introduce a useful topology with respect to which the set of quantizers with a given number of convex codecells is compact, facilitating the proof of existence results. We note that both our assumption of convex-codecell quantizers and the more restrictive assumption of nearest neighbor-type quantizers in Borkar *et al.* [11] may preclude global optimality over all zero-delay quantization policies. The existence and finer structural characterization of such globally optimal policies are still open problems.

The existence and the structural results can be useful for the design of networked control systems where decision makers have imperfect observation of a plant to be controlled. The machinery presented here is particularly useful in the context of optimal quantized control of a linear system driven by unbounded noise: For LQG optimal control problems it has been shown that the effect of the control policies can be decoupled from the estimation error and the design results here can be used to establish existence of optimal quantization and control policies for LQG systems.

The approach developed in this paper can also be applied to the case where $\{x_t\}$ is a Markov chain with finite state space \mathbf{X} . In this case stronger results can be obtained with significantly less technical complications. In particular, when the state space is finite one does not need the convex codecell assumption since there are only a finite number of M -level quantizers on \mathbf{X} . Also, the global optimality of Walrand-Varaiya type policies for the infinite horizon discounted cost problem can be easily proved if \mathbf{X} is finite. In addition, one can prove the optimality of *deterministic* stationary policies for the average cost problem under the irreducibility condition $P(x_{t+1} = b | x_t = a) > 0$ for all $a, b \in \mathbf{X}$. Similar to [11], such an optimality result follows from a *vanishing discount* argument (see [17, Theorem 5.2.4]) using arguments similar to [11, Lemmas 4.1 and 4.2] and the fact that the filtering process forgets its initial state exponentially fast under the irreducibility condition.

A further research direction is the formulation of the communication problem over a channel with feedback. The tools and the topological analysis developed in this paper could be useful in establishing optimal coding and decoding policies and the derivation of error-exponents with feedback. Relevant efforts in the literature on this topic include [32].

VII. APPENDIX

A. Auxiliary Results

Recall that a sequence of probability measures $\{\mu_n\}$ in $\mathcal{P}(\mathbf{X})$ converges to $\mu \in \mathcal{P}(\mathbf{X})$ *weakly* if $\int_{\mathbf{X}} c(x) \mu_n(dx) \rightarrow \int_{\mathbf{X}} c(x) \mu(dx)$ for every continuous and bounded $c : \mathbf{X} \rightarrow \mathbb{R}$. For $\mu, \nu \in \mathcal{P}(\mathbf{X})$ the *total variation* metric is defined by

$$\begin{aligned} d_{TV}(\mu, \nu) &:= 2 \sup_{B \in \mathcal{B}(\mathbf{X})} |\mu(B) - \nu(B)| \\ &= \sup_{g: \|g\|_{\infty} \leq 1} \left| \int g(x) \mu(dx) - \int g(x) \nu(dx) \right|, \end{aligned} \quad (16)$$

where the second supremum is over all measurable real functions g such that $\|g\|_{\infty} := \sup_{x \in \mathbf{X}} |g(x)| \leq 1$.

Definition 5 ([41]): Let $P \in \mathcal{P}(\mathbb{R}^d)$. A *quantizer sequence* $\{Q_n\}$ *converges to* Q *weakly at* P ($Q_n \rightarrow Q$ *weakly at* P) if $PQ_n \rightarrow PQ$ *weakly*. Similarly, $\{Q_n\}$ *converges to* Q *in total variation at* P ($Q_n \rightarrow Q$ *in total variation at* P) if $PQ_n \rightarrow PQ$ *in total variation*.

The following lemma will be very useful in the upcoming optimality proofs.

Lemma 2: (a) Let $\{\mu_n\}$ be a sequence of probability density functions on \mathbb{R}^d which are uniformly equicontinuous and uniformly bounded and assume $\mu_n \rightarrow \mu$ *weakly*. Then $\mu_n \rightarrow \mu$ *in total variation*.

(b) Let $\{Q_n\}$ be a sequence in \mathcal{Q}_c such that $Q_n \rightarrow Q$ *weakly at* P for some $Q \in \mathcal{Q}_c$. If P admits a density, then $Q_n \rightarrow Q$ *in total variation at* P . If the density of P is positive, then $Q_n \rightarrow Q$ *in total variation at any* P' *admitting a density*.

(c) Let $\{Q_n\}$ be a sequence in \mathcal{Q}_c such that $Q_n \rightarrow Q$ *weakly at* P for some $Q \in \mathcal{Q}_c$ where P admits a positive density. Suppose further that $P'_n \rightarrow P'$ *in total variation* where P' admits a density. Then $P'_n Q_n \rightarrow P'Q$ *in total variation*.

Proof: (a) We will denote a density and its induced probability measure by the same symbol. By the Arzelà-Ascoli theorem the sequence of densities $\{\mu_n\}$, when restricted to a given compact subset of \mathbb{R}^d , is relatively compact with respect to the supremum norm. Considering the sequence of increasing closed balls $K_i = \{x : \|x\| \leq i\}$ of radius $i = 1, 2, \dots$, one can use Cantor's diagonal argument as in [41, Lemma 4.3] to obtain a subsequence $\{\mu_{n_k}\}$ and a non-negative continuous function $\hat{\mu}$ such that $\mu_{n_k}(x) \rightarrow \hat{\mu}(x)$ for all x , where the convergence is uniform over compact sets. Since $\int_B |\mu_{n_k}(x) - \hat{\mu}(x)| dx \rightarrow 0$ for any bounded Borel set B , and since $\{\mu_n\}$ is tight by weak convergence, it follows that $\hat{\mu}$ is a probability density. Since μ_{n_k} converges to $\hat{\mu}$ pointwise, by Scheffe's theorem [7] μ_{n_k} converges to $\hat{\mu}$ in the L_1 norm, which is equivalent to convergence in total variation. Since $\mu_n \rightarrow \mu$ *weakly*, we must have $\mu = \hat{\mu}$.

The preceding argument implies that any subsequence of $\{\mu_n\}$ has a further subsequence that converges to μ in (the metric of) total variation. This implies that $\mu_n \rightarrow \mu$ *in total variation*.

(b) It was shown in the proof of [41, Theorem 5.7] that

$$d_{TV}(PQ_n, PQ) \leq \sum_{i=1}^M P(B_i^n \Delta B_i), \quad (17)$$

where B_1^n, \dots, B_M^n and B_1, \dots, B_M are the cells of \mathcal{Q}_n and \mathcal{Q} , respectively, and $B_i^n \triangle B_i := (B_i^n \setminus B_i) \cup (B_i \setminus B_i^n)$. Since \mathcal{Q} has convex cells, the boundary ∂B_i of each cell B_i has zero Lebesgue measure, so $P(\partial B_i) = 0$ because P has a density. Since $\partial(B_i \times \{j\}) = \partial B_i \times \{j\}$, and $PQ(A \times \{j\}) = P(A \cap B_j)$, we have

$$PQ(\partial(B_i \times \{j\})) = P(\partial B_i \cap B_j) = 0,$$

for all i and j . Thus if $PQ^n \rightarrow PQ$ weakly, then $PQ^n(B_i \times \{j\}) \rightarrow PQ(B_i \times \{j\})$ by the Portmanteau theorem, which is equivalent to

$$P(B_i \cap B_j^n) \rightarrow P(B_i \cap B_j)$$

for all i and j . Since $\{B_1^n, \dots, B_M^n\}$ and $\{B_1, \dots, B_M\}$ are both partitions of \mathbb{R}^d , this implies $P(B_i^n \triangle B_i) \rightarrow 0$ for all i , which in turns proves that $PQ^n \rightarrow PQ$ in total variation via (17).

If P has a positive density and P' admits a density, then P' is absolutely continuous with respect to P and so $P(B_i^n \triangle B_i) \rightarrow 0$ implies $P'(B_i^n \triangle B_i) \rightarrow 0$. Combined with the preceding argument this proves the second statement in part (b).

(c) For any $A \in \mathcal{B}(\mathbf{X} \times \mathbf{M})$ let $A(x) := \{y : (x, y) \in A\}$. Then

$$\begin{aligned} & |P'_n Q_n(A) - P' Q_n(A)| \\ &= \left| \int_{\mathbb{R}^d} Q_n(A(x)|x) P'_n(dx) - \int_{\mathbb{R}^d} Q_n(A(x)|x) P'(dx) \right| \\ &\leq d_{TV}(P'_n, P'), \end{aligned}$$

where the inequality is due to (16). Taking the supremum over all A yields

$$d_{TV}(P'_n Q_n, P' Q_n) \leq d_{TV}(P'_n, P').$$

Hence

$$\begin{aligned} d_{TV}(P'_n Q_n, P' Q) &\leq d_{TV}(P'_n Q_n, P' Q_n) + d_{TV}(P' Q_n, P' Q) \\ &\leq d_{TV}(P'_n, P) + d_{TV}(P' Q_n, P' Q). \end{aligned}$$

From part (b) we know that $Q_n \rightarrow Q$ in total variation at P' . Since $P'_n \rightarrow P$ in total variation, we obtain $d_{TV}(P'_n Q_n, P' Q) \rightarrow 0$. \square

Recall from Definition 3 in Section II the set $\mathcal{S} \subset \mathcal{P}(\mathbb{R}^d)$ of probability measures admitting densities that are uniformly bounded and uniformly Lipschitz (with constants determined by the conditional density $\phi(\cdot|x)$ of $x_{t+1} = f(x_t, w_t)$ given $x_t = x$). In Lemma 1 we showed that \mathcal{S} contains all reachable states, i.e., $\pi_t \in \mathcal{S}$ for all $t \geq 1$ with probability 1 under any policy $\Pi \in \Pi_W$.

Lemma 2(a) immediately implies that for any sequence $\{\mu_n\}$ in \mathcal{S} and $\mu \in \mathcal{S}$, $\mu_n \rightarrow \mu$ weakly if and only if $\mu_n \rightarrow \mu$ in total variation. In this case we simply say that $\{\mu_n\}$ converges to μ in \mathcal{S} .

As discussed in Section II, we can define the (quotient) topology on \mathcal{Q}_c induced by weak convergence of sequences at a given P admitting a positive density. Lemma 2(b) implies that any sequence in \mathcal{Q}_c converging in this topology will converge both weakly and in total variation at any P' admitting

a density. In the rest of this section, to say that $\{Q_n\}$ converges in \mathcal{Q}_c will mean *convergence in this topology*. We equip $\mathcal{S} \times \mathcal{Q}_c$ with the corresponding product topology, and continuity of any $F : \mathcal{S} \times \mathcal{Q}_c \rightarrow \mathbb{R}$ will be meant in this sense, unless specifically stated otherwise.

Lemma 3: (a) \mathcal{S} is closed in $\mathcal{P}(\mathbb{R}^d)$.

(b) \mathcal{Q}_c is compact.

(c) If $\{(\mu_n, Q_n)\}$ converges in $\mathcal{S} \times \mathcal{Q}_c$ to $(\mu, Q) \in \mathcal{S} \times \mathcal{Q}_c$ then $\mu_n Q_n \rightarrow \mu Q$ in total variation. Thus any $F : \mathcal{S} \times \mathcal{Q}_c \rightarrow \mathbb{R}$ is continuous if $F(\mu_n, Q_n) \rightarrow F(\mu, Q)$ whenever $\mu_n Q_n \rightarrow \mu Q$ in total variation.

Proof: (a) Recall that \mathcal{S} is a uniformly bounded and uniformly equicontinuous family of densities. Lemma 2(a) shows that if $\{\mu_n\}$ is a sequence in \mathcal{S} and $\mu_n \rightarrow \mu$ weakly, then μ has a density. The proof also shows that some subsequence of (the densities of) $\{\mu_n\}$ converges to (the density of) μ pointwise. Thus μ must admit the same uniform upper bound and Lipschitz constant as all densities in \mathcal{S} , proving that $\mu \in \mathcal{S}$.

(b) The compactness of \mathcal{Q}_c was shown in [41, Theorem 5.8].

(c) If $\{(\mu_n, Q_n)\}$ converges in $\mathcal{S} \times \mathcal{Q}_c$ to $(\mu, Q) \in \mathcal{S} \times \mathcal{Q}_c$ then $\mu_n \rightarrow \mu$ in total variation. Since μ has a density, $Q_n \rightarrow Q$ in \mathcal{Q}_c implies that $Q_n \rightarrow Q$ in total variation at μ . Thus $\mu_n Q_n \rightarrow \mu Q$ in total variation by Lemma 2(c). \square

B. Proof of Theorem 4

The first statement of the following theorem immediately implies Theorem 4.

Theorem 10: For $t = T-1, \dots, 0$ define the value function J_t^T at time t recursively by

$$J_t^T(\pi) = \inf_{Q \in \mathcal{Q}_c} \left(\frac{1}{T} c(\pi, Q) + E[J_{t+1}^T(\pi_{t+1}) | \pi_t = \pi, Q_t = Q] \right)$$

with $J_T^T := 0$ and $c(\pi, Q)$ defined in (5). Then for any $t \geq 1$ and $\pi \in \mathcal{S}$ or $t = 0$ and $\pi \in \mathcal{S} \cup \{\pi_0\}$, the infimum is achieved by some Q in \mathcal{Q}_c . Moreover, $J_t^T(\pi)$ is continuous on \mathcal{S} .

The rest of this section is devoted to proving Theorem 10. The proof is through backward induction in t combined with a series of lemmas that show the continuity of both $c(\pi, Q)$ and $E[J_{t+1}^T(\pi_{t+1}) | \pi_t = \pi, Q_t = Q]$ in (π, Q) .

Lemma 4: $c(\pi, Q)$ is continuous on $\mathcal{S} \times \mathcal{Q}_c$.

Proof: If $\{(\pi_n, Q_n)\}$ converges in $\mathcal{S} \times \mathcal{Q}_c$ then $\pi_n Q_n \rightarrow \pi Q$ in total variation by Lemma 3(c). We have to show that in this case

$$\begin{aligned} c(\pi_n, Q_n) &= \inf_{\gamma} \int_{\mathbb{R}^d} \pi_n(dx) \sum_{i=1}^M Q_n(i|x) c_0(x, \gamma(i)) \\ &\rightarrow \inf_{\gamma} \int_{\mathbb{R}^d} \pi(dx) \sum_{i=1}^M Q(i|x) c_0(x, \gamma(i)) \\ &= c(\pi, Q). \end{aligned}$$

This follows verbatim from the proof of [41, Theorem 3.4] where for any bounded c_0 the convergence for a fixed π and $Q_n \rightarrow Q$ was shown. \square

We now start proving Theorem 10. At $t = T-1$ we have

$$J_{T-1}^T(\pi) = \inf_{Q \in \mathcal{Q}_c} c(\pi, Q).$$

By Lemma 4 and the compactness of the set of quantizers \mathcal{Q}_c (Lemma 3(b)) there exists an optimal quantizer that achieves the infimum. The following lemma will be useful.

Lemma 5: If $F : \mathcal{S} \times \mathcal{Q}_c \rightarrow \mathbb{R}$ is continuous then $\inf_{Q \in \mathcal{Q}_c} F(\pi, Q)$ is achieved by some Q in \mathcal{Q}_c and $\min_{Q \in \mathcal{Q}_c} F(\pi, Q)$ is continuous in π on \mathcal{S} .

Proof: The existence of an optimal Q in \mathcal{Q}_c achieving $\inf_{Q \in \mathcal{Q}_c} F(\pi, Q)$ is a consequence of the continuity of F and the compactness of \mathcal{Q}_c . Assume $\pi_n \rightarrow \pi$ in \mathcal{S} and let Q_n be optimal for π_n and Q optimal for π . Then

$$\begin{aligned} & \left| \min_{Q'} F(\pi_n, Q') - \min_{Q'} F(\pi, Q') \right| \\ & \leq \max \left(F(\pi_n, Q) - F(\pi, Q), F(\pi, Q_n) - F(\pi_n, Q_n) \right). \end{aligned}$$

The first term in the maximum converges to zero since F is continuous. To show that the second converges to zero, suppose to the contrary that for some $\epsilon > 0$ and for a subsequence $\{(\pi_{n_k}, Q_{n_k})\}$,

$$|F(\pi, Q_{n_k}) - F(\pi_{n_k}, Q_{n_k})| \geq \epsilon. \quad (18)$$

By Lemma 3(a), there is a further subsequence $\{n'_k\}$ of $\{n_k\}$ such that $\{Q_{n'_k}\}$ converges to some Q' in \mathcal{Q}_c . Then $\{(\pi, Q_{n'_k})\}$ and $\{(\pi_{n'_k}, Q_{n'_k})\}$ both converge to (π, Q') , which contradicts (18) since F is continuous. \square

As a consequence of Lemmas 4 and 5, $J_{T-1}^T(\pi)$ is continuous on \mathcal{S} , proving Theorem 10 for $t = T - 1$. To prove the theorem for all $t = T - 2, \dots, 0$, we apply backward induction. Assume that the both statements of the theorem hold for $t' = T - 1, \dots, t + 1$. We want to show that the minimization problem

$$J_t^T(\pi) = \min_{Q \in \mathcal{Q}_c} \left(\frac{1}{T} c(\pi, Q) + E[J_{t+1}^T(\pi_{t+1}) | \pi_t = \pi, Q_t = Q] \right) \quad (19)$$

has a solution and $J_t^T(\pi)$ is continuous on \mathcal{S} .

Consider the conditional probability distributions given by

$$\begin{aligned} \hat{\pi}(m, \pi, Q)(C) & := P(x_{t+1} \in C | \pi_t = \pi, Q_t = Q, q_t = m) \\ & = \frac{1}{\pi(Q^{-1}(m))} \int_C \left(\int_{\mathbb{R}^d} \pi(dx) 1_{\{x \in B_m\}} \phi(z|x) \right) dz \quad (20) \end{aligned}$$

(if $\pi(Q^{-1}(m)) = 0$, then $\hat{\pi}(m, \pi, Q)$ is set arbitrarily). Note that

$$\begin{aligned} E[J_{t+1}^T(\pi_{t+1}) | \pi_t = \pi, Q_t = Q] \\ & = \sum_{m=1}^M J_{t+1}^T(\hat{\pi}(m, \pi, Q)) \pi(Q^{-1}(m)), \quad (21) \end{aligned}$$

where

$$\pi(Q^{-1}(m)) = P(q_t = m | \pi_t = \pi, Q_t = Q).$$

The following lemma will imply that if $(\pi_n, Q_n) \rightarrow (\pi, Q)$ in $\mathcal{S} \times \mathcal{Q}_c$, then

$$\begin{aligned} J_{t+1}^T(\hat{\pi}(m, \pi_n, Q_n)) \pi_n(Q_n^{-1}(m)) \\ \rightarrow J_{t+1}^T(\hat{\pi}(m, \pi, Q)) \pi(Q^{-1}(m)) \quad (22) \end{aligned}$$

for all m .

Lemma 6: If $\pi_n Q_n \rightarrow \pi Q$ in total variation, then $\hat{\pi}(m, \pi_n, Q_n) \rightarrow \hat{\pi}(m, \pi, Q)$ in total variation for every $m = 1, \dots, M$ with $\pi(Q^{-1}(m)) > 0$.

Proof: Let B_1, \dots, B_M and B_1^n, \dots, B_M^n denote the cells of Q and Q_n , respectively. Since for any Borel set A , $\pi_n Q_n(A \times \{j\}) = \pi_n(A \cap B_j^n)$, the convergence of $\pi_n Q_n$ to πQ implies that $\pi_n(A \cap B_m^n) \rightarrow \pi(A \cap B_m)$. This implies $\pi_n(B_i \cap B_j^n) \rightarrow \pi(B_i \cap B_j)$ for all i and j , from which we obtain for all $m = 1, \dots, M$,

$$\pi_n(B_m^n) \rightarrow \pi(B_m), \quad \pi_n(B_m^n \Delta B_m) \rightarrow 0, \quad (23)$$

where $B_m^n \Delta B_m = (B_m^n \setminus B_m) \cup (B_m \setminus B_m^n)$.

If $\pi(B_m) > 0$, the probability distribution $\hat{\pi}(m, \pi, Q)$ has density

$$\hat{\pi}(m, \pi, Q)(z) = \frac{1}{\pi(B_m)} \int_{B_m} \pi(dx) \phi(z|x)$$

so by Scheffe's theorem [7] it suffices to show that $\hat{\pi}(m, \pi_n, Q_n)(z) \rightarrow \hat{\pi}(m, \pi, Q)(z)$ for all z . As $\pi(B_m) > 0$ by assumption and $\pi_n(B_m^n) \rightarrow \pi(B_m)$, it is enough to establish the convergence of $v_n^m(z) := \int_{B_m^n} \pi_n(dx) \phi(z|x)$ to $v^m(z) := \int_{B_m} \pi(dx) \phi(z|x)$.

For any $z \in \mathbb{R}^d$ we have

$$\begin{aligned} |v_n^m(z) - v^m(z)| \\ & \leq \left| \int_{\mathbb{R}^d} \pi_n(dx) (1_{\{x \in B_m^n\}} - 1_{\{x \in B_m\}}) \phi(z|x) \right| \\ & \quad + \left| \int_{\mathbb{R}^d} 1_{\{x \in B_m\}} (\pi_n(x) - \pi(x)) \phi(z|x) dx \right| \\ & \leq \int_{B_m^n \Delta B_m} \pi_n(dx) \phi(z|x) + \int_{\mathbb{R}^d} |\pi_n(x) - \pi(x)| \phi(z|x) dx \\ & \leq C \left[\pi_n(B_m^n \Delta B_m) + d_{TV}(\pi_n, \pi) \right], \quad (24) \end{aligned}$$

where C is a uniform upper bound on ϕ . Since both terms in the brackets converge to zero as $n \rightarrow \infty$, the proof is complete. \square

Now if $(\pi_n, Q_n) \rightarrow (\pi, Q)$ in $\mathcal{S} \times \mathcal{Q}_c$, then by Lemma 3 and (23) we have $\pi_n(Q_n^{-1}(m)) \rightarrow \pi(Q^{-1}(m))$ for all m . If $\hat{\pi}(m, \pi, Q) > 0$ for some m , then Lemma 6 implies that $\hat{\pi}(m, \pi_n, Q_n) \rightarrow \hat{\pi}(m, \pi, Q)$ in total variation; hence (22) holds in this case by the continuity of J_{t+1}^T . If $\pi(Q^{-1}(m)) = 0$, then by (23) and the boundedness of the cost (22) holds again. In view of (21), we obtain that $E[J_{t+1}^T(\pi_{t+1}) | \pi_t = \pi, Q_t = Q]$ is continuous on $\mathcal{S} \times \mathcal{Q}_c$.

We have shown that both expressions on the right side of (19) are continuous on $\mathcal{S} \times \mathcal{Q}_c$. By Lemma 5 the minimization problem (19) has a solution and $J_t^T(\pi)$ is continuous on \mathcal{S} , proving the induction hypothesis for $t' = t$.

To finish the proof we have to consider the last step $t = 0$ separately. For $t = 0$ we have that if π admits a positive density, then there exists a minimizing Q for

$$J_0^T(\pi) = \inf_{Q \in \mathcal{Q}_c} \left(\frac{1}{T} c(\pi_0, Q) + E[J_1^T(\pi_1) | \pi_0 = \pi, Q_t = Q] \right)$$

by Lemma 5 since the preceding proofs readily imply that both $c(\pi, Q)$ and $E[J_1^T(\pi_1) | \pi_0 = \pi, Q_t = Q]$ are continuous in Q as long as π admits a positive density. If π is a point mass on x_0 , then any Q is optimal. This establishes Theorem 10. \square

C. Proof of Theorem 5

The first statement of the following counterpart of Theorem 10 immediately implies Theorem 5.

Theorem 11: Consider Assumption 3. For $t = T - 1, \dots, 0$ define the value function J_t^T at time t recursively by

$$J_t^T(\pi) = \inf_{Q \in \mathcal{Q}_c} \left(\frac{1}{T} c(\pi, Q) + E[J_{t+1}^T(\pi_{t+1}) | \pi_t = \pi, Q_t = Q] \right)$$

with $J_T^T := 0$ and $c(\pi, Q)$ defined in (5). Then for any $t \geq 1$ and $\pi \in \mathcal{S}$ or $t = 0$ and $\pi \in \mathcal{S} \cup \{\pi_0\}$ the infimum is achieved by some Q in \mathcal{Q}_c .

Moreover, $J_t^T(\pi)$ is continuous on \mathcal{S} in the sense that if $\pi_n \rightarrow \pi$ and $\{\pi_n\}$ satisfies the uniform integrability condition

$$\limsup_{L \rightarrow \infty} \sup_{n \geq 1} \int_{\{\|x\|^2 \geq L\}} \|x\|^2 \pi_n(dx) = 0, \quad (25)$$

then $J_t^T(\pi_n) \rightarrow J_t^T(\pi)$.

To prove Theorem 11 we need to modify the proof of Theorem 10 only in view of the unboundedness of the cost, which affects the proof of the continuity of $c(\pi, Q)$ and $E[J_{t+1}^T(\pi_{t+1}) | \pi_t = \pi, Q_t = Q]$.

We first establish the continuity of $c(\pi, Q)$ in a more restricted sense than in Lemma 4. We know from (8) that given $\pi_t = \pi$ and $Q_t = Q$ with cells B_1, \dots, B_M , the unique optimal receiver policy is given, for any m such that $\pi(B_m) > 0$, by

$$\gamma(m) = \int_{B_m} x \pi(dx).$$

If $\pi(B_m) = 0$, then $\gamma(m)$ is arbitrary. Using this optimal receiver policy, define $\bar{Q} : \mathbb{R}^d \rightarrow \mathbb{R}^d$ by

$$\bar{Q}(x) = \gamma(Q(x)).$$

Note that $c(\pi, Q) = \int \|x - \bar{Q}(x)\|^2 \pi(dx)$ and that for all m ,

$$\begin{aligned} \int \|x - \bar{Q}(x)\|^2 1_{\{x \in B_m\}} \pi(dx) &= \int_{B_m} \|x - \gamma(m)\|^2 \pi(dx) \\ &\leq \int_{B_m} \|x\|^2 \pi(dx) \end{aligned} \quad (26)$$

which implies

$$c(\pi, Q) \leq \int \|x\|^2 \pi(dx). \quad (27)$$

Lemma 7: Assume $(\pi_n, Q_n) \rightarrow (\pi, Q)$ in $\mathcal{S} \times \mathcal{Q}_c$ and $\{\pi_n\}$ satisfies the uniform integrability condition (25). Then $c(\pi_n, Q_n) \rightarrow c(\pi, Q)$.

Proof: If B_1^n, \dots, B_M^n denote the cells of Q_n and let B_1, \dots, B_M be the cells of Q . By (23) we have $\pi_n(B_m^n) \rightarrow \pi(B_m)$ and $\pi_n(B_m^n \triangle B_m) \rightarrow 0$. Let $I = \{m \in \{1, \dots, M\} : \pi(B_m) > 0\}$. We have for any $L > 0$ and $m \in I$,

$$\int_{B_m^n} x 1_{\{\|x\|^2 < L\}} \pi_n(dx) \rightarrow \int_{B_m} x 1_{\{\|x\|^2 < L\}} \pi(dx).$$

This and a standard truncation argument that makes use of (25) imply

$$\int_{B_m^n} x \pi_n(dx) \rightarrow \int_{B_m} x \pi(dx), \quad m \in I$$

so the optimal receiver policy γ_n for Q_n satisfies $\gamma_n(m) \rightarrow \gamma(m)$ for all $m \in I$. In particular, this implies that for all $m \in I$,

$$D_m := \sup_{n \geq 1} \sup_{x \in B_m^n} \|\bar{Q}_n(x)\|^2 = \sup_{n \geq 1} \gamma_n(m) < \infty.$$

In turn, the parallelogram law gives for $m \in I$ and $x \in B_m^n$,

$$\|x - \bar{Q}_n(x)\|^2 \leq 2\|x\|^2 + 2\|\bar{Q}_n(x)\|^2 \leq 2\|x\|^2 + 2D_m \quad (28)$$

so for $m \in I$ we obtain

$$\begin{aligned} \lim_{L \rightarrow \infty} \sup_{n \geq 1} \int_{B_m^n} \pi_n(dx) \|x - \bar{Q}_n(x)\|^2 1_{\{\|x\|^2 \geq L\}} \\ \leq \lim_{L \rightarrow \infty} \sup_{n \geq 1} \int_{\mathbb{R}^d} \pi_n(dx) (2\|x\|^2 + 2D_m) 1_{\{\|x\|^2 \geq L\}} \\ = 0, \end{aligned} \quad (29)$$

where the second limit is zero due to (25).

Since $\pi_n \rightarrow \pi$ in total variation, $\pi_n(B_m^n) \rightarrow \pi(B_m)$ and $\pi_n(B_m^n \triangle B_m) \rightarrow 0$, and since $\|x - \bar{Q}_n(x)\|^2$ is uniformly bounded if $\|x\|^2 < L$ by (28), we have

$$\begin{aligned} \int_{B_m^n} \pi_n(dx) \|x - \bar{Q}_n(x)\|^2 1_{\{\|x\|^2 < L\}} \\ \rightarrow \int_{B_m} \pi(dx) \|x - \bar{Q}(x)\|^2 1_{\{\|x\|^2 < L\}}. \end{aligned}$$

Then uniform integrability (29) and a standard truncation argument yield for $m \in I$

$$\int_{B_m^n} \pi_n(dx) \|x - \bar{Q}_n(x)\|^2 \rightarrow \int_{B_m} \pi(dx) \|x - \bar{Q}(x)\|^2. \quad (30)$$

Assume $m \notin I$. Then we have

$$\int_{B_m^n} \|x\|^2 \pi_n(dx) \leq \int_{\{\|x\|^2 \geq L\}} \|x\|^2 \pi_n(dx) + L \pi_n(B_m^n) \rightarrow 0$$

from (25) and since $\pi_n(B_m^n) \rightarrow 0$. In view of (26) we obtain

$$\int_{B_m^n} \pi_n(dx) \|x - \bar{Q}_n(x)\|^2 \rightarrow 0.$$

This and (30) give

$$\begin{aligned} c(\pi_n, Q_n) &= \int_{\mathbb{R}^d} \pi_n(dx) \|x - \bar{Q}_n(x)\|^2 \\ &\rightarrow \int_{\mathbb{R}^d} \pi(dx) \|x - \bar{Q}(x)\|^2 = c(\pi, Q) \end{aligned}$$

which proves the lemma. \square

The following variant of Lemma 5 lemma will be useful.

Lemma 8: Assume $F : \mathcal{S} \times \mathcal{Q}_c \rightarrow \mathbb{R}$ is continuous in the sense that if $(\pi_n, Q_n) \rightarrow (\pi, Q)$ in $\mathcal{S} \times \mathcal{Q}_c$ and $\{\pi_n\}$ satisfies the uniform integrability condition (25), then $F(\pi_n, Q_n) \rightarrow F(\pi, Q)$. Then $\inf_{Q \in \mathcal{Q}_c} F(\pi, Q)$ is achieved by some Q in \mathcal{Q}_c and $\min_Q F(\pi, Q)$ is continuous in π in the sense that if $\pi_n \rightarrow \pi$ in \mathcal{S} and $\{\pi_n\}$ is uniformly integrable, then $\min_Q F(\pi_n, Q) \rightarrow \min_Q F(\pi, Q)$.

Proof: The existence of an optimal Q for any $\pi \in \mathcal{S}$ is a consequence of the compactness of \mathcal{Q}_c . The rest of the proof follows verbatim the proof of Lemma 5 with the convergence sequence $\{\pi_n\}$ also assumed to be uniformly integrable. \square

Lemmas 7 and 8 prove Theorem 11 for $t = T - 1$. To prove the theorem for all t , we apply backward induction. Assume that the both statements of the theorem hold for $t' = T - 1, \dots, t + 1$.

Recall the conditional distribution $\hat{\pi}(m, \pi, Q)$ defined in (20). The following lemma shows that the uniform integrability condition is inherited in the induction step.

Lemma 9: Assume $(\pi_n, Q_n) \rightarrow (\pi, Q)$ in $\mathcal{S} \times \mathcal{Q}_c$ and $\{\pi_n\}$ satisfies the uniform integrability condition (25). If cell B_m of Q satisfies $\pi(B_m) > 0$, then $\{\hat{\pi}(m, \pi_n, Q_n)\}$ is uniformly integrable in the sense of (25).

Proof: Let B_m^n denote the m th cell of Q_n . Since $\pi(B_m) > 0$, we have $\pi_n(B_m^n) > 0$ for n large enough, so

$$\int_{\mathbb{R}^d} \|z\|^2 \hat{\pi}(m, \pi_n, Q_n)(z) 1_{\{\|z\|^2 \geq L\}} dz = \frac{1}{\pi(B_m^n)} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \|f(x, w)\|^2 1_{\{\|f(x, w)\|^2 \geq L\}} 1_{\{x \in B_m^n\}} \pi_n(dx) \nu_w(dw).$$

Since $\|f(x, w)\|^2 \leq 2K^2(\|x\|^2 + \|w\|^2)$ by Assumption 3(i), we have

$$1_{\{\|f(x, w)\|^2 \geq L\}} \leq 1_{\{\|x\|^2 \geq L/(4K^2)\}} + 1_{\{\|w\|^2 \geq L/(4K^2)\}}$$

and so

$$\begin{aligned} & \int_{\mathbb{R}^d} \|z\|^2 \hat{\pi}(m, \pi_n, Q_n)(z) 1_{\{\|z\|^2 \geq L\}} dz \\ & \leq \frac{1}{\pi(B_m^n)} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} 2K^2(\|x\|^2 + \|w\|^2) \\ & \quad \times (1_{\{\|x\|^2 \geq L/(4K^2)\}} + 1_{\{\|w\|^2 \geq L/(4K^2)\}}) \pi_n(dx) \nu_w(dw) \\ & = \frac{2K^2}{\pi(B_m^n)} \int_{\mathbb{R}^d} \|x\|^2 1_{\{\|x\|^2 \geq L/(4K^2)\}} \pi_n(dx) \\ & \quad + \frac{2K^2}{\pi(B_m^n)} \left(\int_{\mathbb{R}^d} \|x\|^2 \pi_n(dx) \right) \left(\int_{\mathbb{R}^d} 1_{\{\|w\|^2 \geq L/(4K^2)\}} \nu_w(dw) \right) \\ & \quad + \frac{2K^2}{\pi(B_m^n)} \left(\int_{\mathbb{R}^d} \|w\|^2 \nu_w(dw) \right) \left(\int_{\mathbb{R}^d} 1_{\{\|x\|^2 \geq L/(4K^2)\}} \pi_n(dx) \right) \\ & \quad + \frac{2K^2}{\pi(B_m^n)} \int_{\mathbb{R}^d} \|w\|^2 1_{\{\|w\|^2 \geq L/(4K^2)\}} \nu_w(dw). \end{aligned}$$

Recall that $\pi_n(B_m^n) \rightarrow \pi(B_m)$. Thus the first term in sum above converges to zero as $L \rightarrow \infty$ uniformly in n by (25). The uniform convergence to zero of the other three terms in the sum follows since $\int \|w\|^2 \nu_w(dw) < \infty$ and $\sup_{n \geq 1} \int \|x\|^2 \pi_n(x) dx < \infty$ by (25). This proves

$$\limsup_{L \rightarrow \infty} \sup_{n \geq 1} \int_{\mathbb{R}^d} \|z\|^2 \hat{\pi}(m, \pi_n, Q_n)(z) 1_{\{\|z\|^2 \geq L\}} dz = 0$$

as claimed. \square

The next lemma shows the continuity of $E[J_{t+1}^T(\pi_{t+1}) | \pi_t = \pi, Q_t = Q]$.

Lemma 10: $E[J_{t+1}^T(\pi_{t+1}) | \pi_t = \pi, Q_t = Q]$ is continuous on $\mathcal{S} \times \mathcal{Q}_c$ in the sense of Lemma 8.

Proof: Assume $(\pi_n, Q_n) \rightarrow (\pi, Q)$ in $\mathcal{S} \times \mathcal{Q}_c$ and $\{\pi_n\}$ satisfies the uniform integrability condition (25). Let B_1, \dots, B_M and B_1^n, \dots, B_M^n denote the cells of Q and Q_n ,

respectively. In view of (21) and the fact that $\pi_n(B_m^n) \rightarrow \pi(B_m)$, we need to prove that for all m with $\pi(B_m) > 0$,

$$J_{t+1}^T(\hat{\pi}(m, \pi_n, Q_n)) \rightarrow J_{t+1}^T(\hat{\pi}(m, \pi, Q)) \quad (31)$$

and for m with $\pi(B_m) = 0$,

$$J_{t+1}^T(\hat{\pi}(m, \pi_n, Q_n)) \pi_n(B_m^n) \rightarrow 0. \quad (32)$$

The convergence in (31) follows from Lemmas 6 and 9, and the induction hypothesis that $J_{t+1}^T(\cdot)$ is continuous along convergent and uniformly integrable sequences in \mathcal{S} .

To prove (32) first note that from (27) we have

$$J_{t+1}^T(\pi_{t+1}) \leq E \left[\frac{1}{T} \sum_{i=t+1}^{T-1} \|x_i\|^2 \right],$$

where x_{t+1} has distribution π_{t+1} and $x_i = f(x_{i-1}, w_{i-1})$, where w_{t+1}, \dots, w_{T-1} are independent of x_{t+1} . Accordingly,

$$\begin{aligned} & J_{t+1}^T(\hat{\pi}(m, \pi_n, Q_n)) \pi_n(B_m^n) \\ & \leq E \left[\frac{1}{T} \left(\sum_{i=t+1}^{T-1} \|x_{i,n}\|^2 \right) 1_{\{x_{t,n} \in B_m^n\}} \right], \end{aligned} \quad (33)$$

where $x_{t,n}$ has distribution π_n .

Now note that the assumption $\|f(x, w)\| \leq K(\|x\| + \|w\|)$ and the inequality $\|x + y\|^2 \leq 2\|x\|^2 + 2\|y\|^2$ imply the upper bound

$$\|x_{t+j,n}\|^2 \leq (2K^2)^j \|x_{t,n}\|^2 + \sum_{i=0}^{j-1} (2K^2)^{j-i} \|w_{t+i}\|^2. \quad (34)$$

Thus for any $j = 1, \dots, T - t - 1$ we have

$$\begin{aligned} & E[\|x_{t+j,n}\|^2 1_{\{x_{t,n} \in B_m^n\}}] \\ & \leq (2K^2)^j E[\|x_{t,n}\|^2 1_{\{x_{t,n} \in B_m^n\}}] \\ & \quad + \sum_{i=1}^j (2K^2)^{j-i} E[\|w_{t+i-1}\|^2 1_{\{x_{t,n} \in B_m^n\}}] \\ & = (2K^2)^j \left(E[\|x_{t,n}\|^2 1_{\{x_{t,n} \in B_m^n\}}] \right. \\ & \quad \left. + \sum_{i=1}^j (2K^2)^{j-i} E[\|w_{t+i-1}\|^2] \pi_n(B_m^n) \right), \end{aligned}$$

where we used the independence of w_t, \dots, w_{T-1} and $x_{t,n}$. The first expectation in the last equation converges to zero as $n \rightarrow \infty$ since $\{\pi_n\}$ is uniformly integrable and $\pi_n(B_m^n) \rightarrow \pi(B_m) = 0$, while the second one converges to zero since $\pi_n(B_m^n) \rightarrow 0$. This proves that the right side of (33) converges to zero, finishing the proof of the lemma. \square

Lemmas 7 and 10 show that

$$F_t(\pi, Q) := \frac{1}{T} c(\pi, Q) + E[J_{t+1}^T(\pi_{t+1}) | \pi_t = \pi, Q_t = Q]$$

satisfies the conditions of Lemma 8, which in turn proves the induction hypothesis for $t' = t$. For the last step $t = 0$ a similar argument as in the proof of Theorem 10 applies (but here we also need the condition $E_{\pi_0}[\|x_0\|^2] < \infty$). This finishes the proof of Theorem 11. \square

D. Proof of Theorem 6

Define

$$J_{\pi^*}(T) := \inf_{\Pi \in \Pi_A} \inf_{\gamma} E_{\pi^*}^{\Pi, \gamma} \left[\frac{1}{T} \sum_{t=0}^{T-1} c_0(x_t, u_t) \right]$$

and note that $\limsup_{T \rightarrow \infty} J_{\pi^*}(T) \leq J_{\pi^*}$. Thus there exists an increasing sequence of time indices $\{T_k\}$ such that for all $k = 1, 2, \dots$,

$$J_{\pi^*}(T_k) \leq J_{\pi^*} + \frac{1}{k}. \quad (35)$$

A key observation is that by Theorem 2 for all k there exists $\Pi_k = \{\hat{\eta}_t^{(k)}\} \in \Pi_W$ (a Markov policy) such that

$$J_{\pi^*}(\Pi_k, T_k) := E_{\pi^*}^{\Pi_k} \left[\frac{1}{T_k} \sum_{t=0}^{T_k-1} c(\pi_t, Q_t) \right] \leq J_{\pi^*}(T_k) + \frac{1}{k}. \quad (36)$$

Now let $n_1 = 1$ and for $k = 2, 3, \dots$, choose the positive integers n_k inductively as

$$n_k = \left\lceil k \cdot \max \left(\frac{T_{k+1}}{T_k}, \frac{n_{k-1} T_{k-1}}{T_k} \right) \right\rceil, \quad (37)$$

where $\lceil x \rceil$ denotes the smallest integer greater than equal to x . Note that the definition of n_k implies $n_k T_k \geq k n_{k-1} T_{k-1}$. Thus letting $T'_k = n_k T_k$ for all k we have

$$T'_k \geq k T'_{k-1}, \quad (38)$$

and hence

$$\lim_{k \rightarrow \infty} \frac{\sum_{l=1}^k T'_l}{T'_k} = 1. \quad (39)$$

Now let $N_0 = 0$, $N_k = \sum_{i=1}^k T'_i$ for $k \geq 1$, and define the policy $\Pi = \{\hat{\eta}_t\} \in \Pi_W$ by piecing together, in a periodic fashion, the initial segments of Π_k as follows:

- (1) For $t = N_{k-1} + jT_k$, where $k \geq 1$ and $0 \leq j < n_k$, let $\hat{\eta}_t(\cdot) \equiv \hat{\eta}_0^{(k)}(\pi^*)$;
- (2) For $t = N_{k-1} + jT_k + i$, where $k \geq 1$, $0 \leq j < n_k$, and $1 \leq i < T_k$, let $\hat{\eta}_t = \hat{\eta}_i^{(k)}$.

In the rest of the proof we show that Π is optimal. First note that by the stationarity of $\{x_t\}$ we have, for all $k \geq 1$ and $j = 0, \dots, n_k - 1$,

$$E_{\pi^*}^{\Pi} \left[\sum_{t=N_{k-1}+jT_k}^{N_{k-1}+(j+1)T_k-1} c(\pi_t, Q_t) \right] = T_k J_{\pi^*}(\Pi_k, T_k).$$

Hence, for $T = N_{k-1} + jT_k + i$, where $k \geq 3$, $0 \leq j < n_k$, and $0 \leq i < T_k$, we have

$$\begin{aligned} & E_{\pi^*}^{\Pi} \left[\frac{1}{T} \sum_{t=0}^{T-1} c(\pi_t, Q_t) \right] \\ &= E_{\pi^*}^{\Pi} \left[\frac{1}{T} \sum_{t=0}^{N_{k-2}-1} c(\pi_t, Q_t) \right] + E_{\pi^*}^{\Pi} \left[\frac{1}{T} \sum_{t=N_{k-2}}^{T-1} c(\pi_t, Q_t) \right] \\ &= \frac{1}{T} \sum_{l=1}^{k-2} T'_l J_{\pi^*}(\Pi_l, T_l) \end{aligned} \quad (40)$$

$$+ \frac{1}{T} \left(T'_{k-1} J_{\pi^*}(\Pi_{k-1}, T_{k-1}) + j T_k J_{\pi^*}(\Pi_k, T_k) \right) \quad (41)$$

$$+ E_{\pi^*}^{\Pi} \left[\frac{1}{T} \sum_{t=N_{k-1}+jT_k}^{T-1} c(\pi_t, Q_t) \right] \quad (42)$$

(the last sum is empty if $i = 0$).

Let \hat{C} be a uniform upper bound on the cost c_0 . Since $T \geq N_{k-1}$, (40) can be bounded as

$$\begin{aligned} \frac{1}{T} \sum_{l=1}^{k-2} T'_l J_{\pi^*}(\Pi_l, T_l) &\leq \hat{C} \frac{1}{N_{k-1}} \sum_{l=1}^{k-2} T'_l = \hat{C} \frac{N_{k-2}}{N_{k-1}} \\ &= \hat{C} \frac{\frac{N_{k-2}}{T'_{k-2}}}{\frac{N_{k-2}}{T'_{k-2}} + \frac{T'_{k-1}}{T'_{k-2}}} \rightarrow 0 \end{aligned} \quad (43)$$

as $k \rightarrow \infty$ since $\frac{N_{k-2}}{T'_{k-2}} \rightarrow 1$ from (39) and $\frac{T'_{k-1}}{T'_{k-2}} \geq k - 1$ from (38).

Since $T'_{k-1} + jT_k \leq T$, (41) can be upper bounded as

$$\begin{aligned} & \frac{1}{T} \left(T'_{k-1} J_{\pi^*}(\Pi_{k-1}, T_{k-1}) + j T_k J_{\pi^*}(\Pi_k, T_k) \right) \\ & \leq \max \left(J_{\pi^*}(\Pi_{k-1}, T_{k-1}), J_{\pi^*}(\Pi_k, T_k) \right). \end{aligned} \quad (44)$$

Finally, the expectation in (42) is upper bounded as

$$\begin{aligned} E_{\pi^*}^{\Pi} \left[\frac{1}{T} \sum_{t=N_{k-1}+jT_k}^{T-1} c(\pi_t, Q_t) \right] &\leq \hat{C} \frac{T_k}{T} \leq \hat{C} \frac{T_k}{T'_{k-1}} \\ &\leq \frac{\hat{C}}{k-1} \rightarrow 0 \end{aligned} \quad (45)$$

as $k \rightarrow \infty$, where the last inequality holds since by (37) we have $T'_k = n_k T_k \geq k T_{k+1}$ for all k .

Combining (40)–(45) we obtain

$$\limsup_{T \rightarrow \infty} E_{\pi^*}^{\Pi} \left[\frac{1}{T} \sum_{t=0}^{T-1} c(\pi_t, Q_t) \right] \leq \limsup_{k \rightarrow \infty} J_{\pi^*}(\Pi_k, T_k) \leq J_{\pi^*}$$

which proves the optimality of Π .

E. Proof of Proposition 1

Proof of (a): Here we show that any weak limit of $\{v_t\}$ must belong to \mathcal{G} . For $v \in \mathcal{P}(\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c)$ and $g \in \mathcal{C}_b(\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c)$ or $f \in \mathcal{C}_b(\mathcal{P}(\mathbb{R}^d))$ define

$$\langle v, g \rangle := \int g(\pi, Q) v(d\pi dQ), \quad \langle v, f \rangle := \int f(\pi) v(d\pi dQ).$$

Also define $vP \in \mathcal{P}(\mathcal{P}(\mathbb{R}^d))$ by

$$vP(A) := \int P(\pi_{t+1} \in A | \pi_t = \pi, Q_t = Q) v(d\pi dQ)$$

for any measurable $A \subset \mathcal{P}(\mathbb{R}^d)$. Note that $v \in \mathcal{G}$ is equivalent to

$$\langle vP, f \rangle = \langle v, f \rangle \quad \text{for all } f \in \mathcal{C}_b(\mathcal{P}(\mathbb{R}^d)). \quad (46)$$

From the definition of $v_t P$, we have for any $f \in \mathcal{C}_b(\mathcal{P}(\mathbb{R}^d))$,

$$\langle v_t, f \rangle - \langle v_t P, f \rangle = \frac{1}{t} E_{\pi_0} \left[\sum_{i=0}^{t-1} f(\pi_i) - \sum_{i=1}^t f(\pi_i) \right]$$

$$= \frac{1}{t} E_{\pi_0} [f(\pi_0) - f(\pi_t)] \rightarrow 0 \quad (47)$$

as $t \rightarrow \infty$. Now suppose that $v_{t_k} \rightarrow \bar{v}$ weakly along a subsequence of $\{v_t\}$. Then $\langle v_{t_k}, f \rangle \rightarrow \langle \bar{v}, f \rangle$ for all $f \in \mathcal{C}_b(\mathcal{P}(\mathbb{R}^d))$, and (47) implies

$$\langle v_{t_k} P, f \rangle \rightarrow \langle \bar{v}, f \rangle. \quad (48)$$

The following lemma is proved at the end of this section.

Lemma 11: The transition kernel $P(d\pi_{t+1}|\pi_t, Q_t)$ is continuous in the weak-Feller sense, i.e., for any $f \in \mathcal{C}_b(\mathcal{P}(\mathbb{R}^d))$,

$$Pf(\pi, Q) := \int_{\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c} f(\pi') P(d\pi'|\pi, Q)$$

is continuous on $\mathcal{S} \times \mathcal{Q}_c$,

The lemma implies that $Pf \in \mathcal{C}_b(\mathcal{S} \times \mathcal{Q}_c)$, so $\langle v_{t_k}, Pf \rangle \rightarrow \langle \bar{v}, Pf \rangle$. However, since for all v ,

$$\langle vP, f \rangle = \int_{\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c} f(\pi) P(d\pi'|\pi, Q) v(d\pi dQ) = \langle v, Pf \rangle,$$

this is equivalent to $\langle v_{t_k} P, f \rangle \rightarrow \langle \bar{v} P, f \rangle$. Combining this with (48) yields $\langle \bar{v} P, f \rangle = \langle \bar{v}, f \rangle$ which finishes the proof that $\bar{v} \in \mathcal{G}$.

Although $c(\pi, Q)$ is continuous on $\mathcal{S} \times \mathcal{Q}_c$ by Lemma 4, the limit relation (13) does not follow immediately since π_0 may not be in \mathcal{S} and thus v_t may not be supported on $\mathcal{S} \times \mathcal{Q}_c$. However, since $\pi_t \in \mathcal{S}$ for all $t \geq 1$ with probability 1, we have $v_t(\mathcal{S} \times \mathcal{Q}_c) \geq 1 - 1/t$, and we can proceed as follows: Recall that $\mathcal{S} \times \mathcal{Q}_c$ is a closed subset of $\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c$ by Lemma 3 and the topology on $\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c$ is metrizable. Thus by the Tietze-Urysohn extension theorem [12] there exists $\tilde{c} \in \mathcal{C}_b(\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c)$ which coincides with c on $\mathcal{S} \times \mathcal{Q}_c$. Then since $v_{t_n}(\mathcal{S} \times \mathcal{Q}_c) \geq 1 - 1/t_n$ and both c and \tilde{c} are bounded,

$$\lim_{n \rightarrow \infty} \int_{\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c} |\tilde{c}(\pi, Q) - c(\pi, Q)| v_{t_n}(d\pi dQ) = 0.$$

On the other hand, $v_{t_n} \rightarrow \bar{v}$ implies

$$\begin{aligned} \lim_{n \rightarrow \infty} \int_{\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c} \tilde{c}(\pi, Q) v_{t_n}(d\pi dQ) \\ &= \int_{\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c} \tilde{c}(\pi, Q) \bar{v}(d\pi dQ) \\ &= \int_{\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c} c(\pi, Q) \bar{v}(d\pi dQ), \end{aligned}$$

where the last equality holds since $\bar{v} \in \mathcal{G}$ is supported on $\mathcal{S} \times \mathcal{Q}_c$. This proves (13).

Proof of (b): We need the following simple lemma.

Lemma 12: Let H be a collection of probability measures on $\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c$ such that

$$R := \sup_{v \in H} \int_{\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c} \left(\int_{\mathbb{R}^d} \|x\|^2 \pi(dx) \right) v(d\pi dQ) < \infty$$

Then H is tight and is thus relatively compact.

Proof: For any $\alpha > 0$ let

$$K_\alpha := \left\{ \pi \in \mathcal{P}(\mathbb{R}^d) : \int_{\mathbb{R}^d} \|x\|^2 \pi(dx) \leq \alpha \right\}.$$

Then $\pi(\{x : \|x\|^2 > L\}) \leq \alpha/L$ for all $\pi \in K_\alpha$ by Markov's inequality. Hence K_α is tight and thus relatively compact.

A standard truncation argument shows that if $\pi_k \rightarrow \pi$ (weakly) for a sequence $\{\pi_k\}$ in K_α , then

$$\alpha \geq \limsup_{k \rightarrow \infty} \int_{\mathbb{R}^d} \|x\|^2 \pi_k(dx) \geq \int_{\mathbb{R}^d} \|x\|^2 \pi(dx)$$

so K_α is also closed. Thus K_α is compact.

Let $f(\pi) := \int_{\mathbb{R}^d} \|x\|^2 \pi(dx)$. Then

$$\int_{\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c} f(\pi) v(d\pi dQ) \leq R \quad \text{for all } v \in H$$

Again by Markov's inequality,

$$\int_{\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c} f(\pi) v(d\pi dQ) \geq \alpha v((K_\alpha)^c \times \mathcal{Q}_c)$$

implying, for all $v \in H$,

$$v(K_\alpha \times \mathcal{Q}_c) \geq 1 - \frac{R}{\alpha}.$$

Since \mathcal{Q}_c is compact and K_α is compact for all $\alpha > 0$, we obtain that H is tight. \square

Let Π be an arbitrary fixed policy in $\bar{\Pi}_W^C$, fix the initial distribution δ_{x_0} , and consider the corresponding sequence of expected occupation measures $\{v_t\}$. Then

$$\begin{aligned} \int_{\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c} \left(\int_{\mathbb{R}^d} \|x\|^2 \pi(dx) \right) v_t(d\pi dQ) \\ = E_{\delta_{x_0}} \left[\frac{1}{t} \sum_{k=0}^{t-1} \|x_k\|^2 \right] \rightarrow \int_{\mathbb{R}^d} \|x\|^2 \pi^*(dx) < \infty \end{aligned}$$

by Assumption 4. Hence

$$\sup_{t \geq 0} \int_{\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c} \left(\int_{\mathbb{R}^d} \|x\|^2 \pi(dx) \right) v_t(d\pi dQ) < \infty. \quad (49)$$

Thus $\{v_t\}$ is relatively compact by Lemma 12, proving part (b) of the proposition.

Proof of (c): We will show that \mathcal{G} is closed and relatively compact. To show closedness, let $\{v_n\}$ be a sequence in \mathcal{G} such that $v_n \rightarrow \bar{v}$. Using the notation introduced in the proof of part (a), we have for any $f \in \mathcal{C}_b(\mathcal{P}(\mathbb{R}^d))$ by (46),

$$\langle v_n P, f \rangle = \langle v_n, f \rangle \rightarrow \langle \bar{v}, f \rangle.$$

But we also have

$$\langle v_n P, f \rangle = \langle v_n, Pf \rangle \rightarrow \langle \bar{v}, Pf \rangle = \langle \bar{v} P, f \rangle,$$

where the limit holds by the weak-Feller property of P (Lemma 11). Thus $\langle \bar{v} P, f \rangle = \langle \bar{v}, f \rangle$, showing that $\bar{v} \in \mathcal{G}$. Hence \mathcal{G} is closed.

To show relative compactness, recall from (20) the conditional distributions

$$\hat{\pi}(m, \pi, Q)(dx_{t+1}) = P(dx_{t+1}|\pi_t = \pi, Q_t = Q, q_t = m)$$

for $m = 1, \dots, M$. For any (π, Q) and Borel set $A \subset \mathbb{R}^d$,

$$\begin{aligned} \int_{\mathcal{P}(\mathbb{R}^d)} \pi'(A) P(d\pi'|\pi, Q) \\ = \sum_{m=1}^M \hat{\pi}(m, \pi, Q)(A) P(\hat{\pi}(m, \pi, Q)|\pi, Q) \end{aligned}$$

$$\begin{aligned}
&= \sum_{m=1}^M \left(\frac{1}{\pi(Q^{-1}(m))} \int_{Q^{-1}(m)} P(x_{t+1} \in A | x_t) \pi(dx_t) \right) \pi(Q^{-1}(m)) \\
&= \int_{\mathbb{R}^d} P(x_{t+1} \in A | x_t) \pi(dx_t). \tag{50}
\end{aligned}$$

Now let $v \in \mathcal{G}$ and consider the ‘‘average’’ π_v under v determined by

$$\pi_v(A) = \int_{\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c} \pi(A) v(d\pi dQ) = \int_{\mathcal{P}(\mathbb{R}^d)} \pi(A) \hat{v}(d\pi),$$

where \hat{v} is obtained from $v(d\pi dQ) = \bar{\eta}(dQ|\pi) \hat{v}(d\pi)$. Recall that v is supported on $\mathcal{S} \times \mathcal{Q}_c$. If A has boundary of zero Lebesgue measure, the mapping $\pi \mapsto \pi(A)$ is continuous on \mathcal{S} and the definition of \mathcal{G} implies

$$\begin{aligned}
\pi_v(A) &= \int_{\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c} \pi(A) v(d\pi dQ) \\
&= \int_{\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c} \int_{\mathcal{P}(\mathbb{R}^d)} \pi'(A) P(d\pi'|\pi, Q) v(d\pi dQ) \\
&= \int_{\mathcal{P}(\mathbb{R}^d)} \int_{\mathcal{Q}_c} \int_{\mathcal{P}(\mathbb{R}^d)} \pi'(A) P(d\pi'|\pi, Q) \bar{\eta}(dQ|\pi) \hat{v}(d\pi). \tag{51}
\end{aligned}$$

Substituting (50) into the last integral, we obtain

$$\begin{aligned}
\pi_v(A) &= \int_{\mathcal{P}(\mathbb{R}^d)} \int_{\mathcal{Q}_c} \int_{\mathbb{R}^d} P(x_{t+1} \in A | x_t) \pi(dx_t) \bar{\eta}(dQ|\pi) \hat{v}(d\pi) \\
&= \int_{\mathbb{R}^d} P(x_{t+1} \in A | x_t) \pi_v(dx_t).
\end{aligned}$$

Since the Borel sets in \mathbb{R}^d having boundaries of zero Lebesgue measure form a separating class for $\mathcal{P}(\mathbb{R}^d)$, the above holds for all Borel sets A , implying that $\pi_v = \pi^*$, the unique invariant measure for $\{x_t\}$. Thus

$$\begin{aligned}
&\int_{\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c} \left(\int_{\mathbb{R}^d} \|x\|^2 \pi(dx) \right) v(d\pi dQ) \\
&= \int_{\mathbb{R}^d} \|x\|^2 \pi_v(dx) = \int_{\mathbb{R}^d} \|x\|^2 \pi^*(dx) \tag{52}
\end{aligned}$$

for all $v \in \mathcal{G}$. Since the last integral is finite by Assumption 4, Lemma 12 implies that \mathcal{G} is relatively compact. \square

Proof of Lemma 11: Consider a sequence $\{(\pi_n, Q_n)\}$ converging to some (π, Q) in $\mathcal{S} \times \mathcal{Q}_c$. Then for any $f \in C_b(\mathcal{P}(\mathbb{R}^d))$,

$$\begin{aligned}
&\int_{\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c} f(\pi') P(d\pi'|\pi_n, Q_n) - \int_{\mathcal{P}(\mathbb{R}^d) \times \mathcal{Q}_c} f(\pi') P(d\pi'|\pi, Q) \\
&= \sum_{m=1}^M \left(f(\hat{\pi}(m, \pi_n, Q_n)) P(\hat{\pi}(m, \pi_n, Q_n) | \pi_n, Q_n) \right. \\
&\quad \left. - f(\hat{\pi}(m, \pi, Q)) P(\hat{\pi}(m, \pi, Q) | \pi, Q) \right) \\
&= \sum_{m=1}^M \left(f(\hat{\pi}(m, \pi_n, Q_n)) \pi_n(Q_n^{-1}(m)) \right. \\
&\quad \left. - f(\hat{\pi}(m, \pi, Q)) \pi(Q^{-1}(m)) \right).
\end{aligned}$$

From Lemma 3 we have that $\pi_n Q_n \rightarrow \pi Q$ in total variation which implies via Lemma 6 that $\hat{\pi}(m, \pi_n, Q_n) \rightarrow \hat{\pi}(m, \pi, Q)$ in total variation and thus weakly for all m with $\pi(Q^{-1}(m)) > 0$. The proof of the same lemma shows that $\pi_n(Q_n^{-1}(m)) \rightarrow \pi(Q^{-1}(m))$ for all $m = 1, \dots, M$. Since f is continuous and bounded, the last sum converges to zero as $n \rightarrow \infty$, proving the claim of the lemma. \square

ACKNOWLEDGEMENTS

We are grateful to Vivek S. Borkar and Naci Saldi for technical discussions related to the paper. We also thank two anonymous reviewers for constructive comments.

REFERENCES

- [1] E. F. Abaya and G. L. Wise, ‘‘Convergence of vector quantizers with applications to optimal quantization,’’ *SIAM J. Appl. Math.*, vol. 44, no. 1, pp. 183–189, Feb. 1984.
- [2] A. Antos, ‘‘On codecell convexity of optimal multiresolution scalar quantizers for continuous sources,’’ *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 1147–1157, Feb. 2012.
- [3] A. Arapostathis, V. S. Borkar, E. Fernandez-Gaucherand, M. K. Ghosh, and S. I. Marcus, ‘‘Discrete-time controlled Markov processes with average cost criterion: A survey,’’ *SIAM J. Control Optim.*, vol. 31, no. 2, pp. 282–344, 1993.
- [4] A. Arapostathis, V. S. Borkar, and M. K. Ghosh, *Ergodic Control of Diffusion Processes*. Cambridge, U.K.: Cambridge Univ. Press, 2012.
- [5] H. Asnani and T. Weissman, ‘‘On real time coding with limited lookahead,’’ *IEEE Trans. Inf. Theory*, vol. 59, no. 6, pp. 3582–3606, Jun. 2013.
- [6] L. Bao, M. Skoglund, and K. H. Johansson, ‘‘Iterative encoder-controller design for feedback control over noisy channels,’’ *IEEE Trans. Autom. Control*, vol. 57, no. 2, pp. 265–278, Feb. 2011.
- [7] P. Billingsley, *Probability and Measure*, 2nd ed. New York, NY, USA: Wiley, 1986.
- [8] V. S. Borkar, ‘‘Convex analytic methods in Markov decision processes,’’ in *Handbook of Markov Decision Processes: Methods and Applications*. Boston, MA, USA: Kluwer, 2002.
- [9] V. S. Borkar and S. K. Mitter, ‘‘LQG control with communication constraints,’’ in *Communications, Computation, Control, and Signal Processing*. Boston, MA, USA: Kluwer, 1997.
- [10] V. S. Borkar, S. K. Mitter, A. Sahai, and S. Tatikonda, ‘‘Sequential source coding: An optimization viewpoint,’’ in *Proc. 44th IEEE Conf. Decision Control*, Seville, Spain, Dec. 2005, pp. 1035–1042.
- [11] V. S. Borkar, S. K. Mitter, and S. Tatikonda, ‘‘Optimal sequential vector quantization of Markov sources,’’ *SIAM J. Control Optim.*, vol. 40, no. 1, pp. 135–148, 2001.
- [12] R. M. Dudley, *Real Analysis and Probability*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2002.
- [13] S. Dumitrescu and X. Wu, ‘‘On properties of locally optimal multiple description scalar quantizers with convex cells,’’ *IEEE Trans. Inf. Theory*, vol. 55, no. 12, pp. 5591–5606, Dec. 2009.
- [14] I. I. Gikhman and A. V. Skorokhod, *Controlled Stochastic Processes*. New York, NY, USA: Springer-Verlag, 1979.
- [15] R. M. Gray and D. L. Neuhoff, ‘‘Quantization,’’ *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2325–2383, Oct. 1998.
- [16] A. Gyorgy and T. Linder, ‘‘Codecell convexity in optimal entropy-constrained vector quantization,’’ *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1821–1828, Jul. 2003.
- [17] O. Hernandez-Lerma and J. B. Lasserre, *Discrete-Time Markov Control Processes, Basic Optimality Criteria*. New York, NY, USA: Springer-Verlag, 1996.
- [18] O. Hernandez-Lerma and J. B. Lasserre, *Further Topics on Discrete-Time Markov Control Processes*. New York, NY, USA: Springer-Verlag, 1999.
- [19] O. Hernandez-Lerma and J. B. Lasserre, *Markov Chains and Invariant Probabilities*. Cambridge, MA, USA: Birkhuser, 2003.
- [20] T. Javidi and A. Goldsmith, ‘‘Dynamic joint source-channel coding with feedback,’’ in *Proc. IEEE Int. Symp. Inf. Theory*, Istanbul, Turkey, Jul. 2013, pp. 16–20.
- [21] Y. Kaspi and N. Merhav, ‘‘Structure theorems for real-time variable rate coding with and without side information,’’ *IEEE Trans. Inf. Theory*, vol. 58, no. 12, pp. 7135–7153, Dec. 2012.

- [22] T. Linder and R. Zamir, "Causal coding of stationary sources and individual sequences with high resolution," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 662–680, Feb. 2006.
- [23] A. Mahajan and D. Teneketzis, "On the design of globally optimal communication strategies for real-time noisy communication systems with noisy feedback," *IEEE J. Special Areas Commun.*, vol. 26, no. 4, pp. 580–595, May 2008.
- [24] A. Mahajan and D. Teneketzis, "Optimal design of sequential real-time communication systems," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 5317–5338, Nov. 2009.
- [25] S. P. Meyn, *Control Techniques for Complex Networks*. Cambridge, U.K.: Cambridge Univ. Press, 2008.
- [26] S. P. Meyn and R. Tweedie, *Markov Chains and Stochastic Stability*. London, U.K.: Springer-Verlag, 1993.
- [27] D. Muresan and M. Effros, "Quantization as histogram segmentation: Optimal scalar quantizer design in network systems," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 344–366, Jan. 2008.
- [28] G. N. Nair, F. Fagnani, S. Zampieri, and J. R. Evans, "Feedback control under data rate constraints: An overview," *Proc. IEEE*, vol. 95, no. 1, pp. 108–137, Jan. 2007.
- [29] D. L. Neuhoff and R. Gilbert, "Causal source codes," *IEEE Trans. Inf. Theory*, vol. 28, no. 5, pp. 701–713, Sep. 1982.
- [30] D. Pollard, "Quantization and the method of k -means," *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 199–205, Mar. 1982.
- [31] N. Saldi, T. Linder, and S. Yüksel, "Randomized quantization and optimal design with a marginal constraint," in *Proc. IEEE Int. Symp. Inf. Theory*, Istanbul, Turkey, Jul. 2013, pp. 2349–2353.
- [32] S. Tatikonda and S. Mitter, "The capacity of channels with feedback," *IEEE Trans. Inf. Theory*, vol. 55, no. 1, pp. 323–349, Jan. 2009.
- [33] S. Tatikonda, A. Sahai, and S. Mitter, "Stochastic linear control over a communication channel," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1549–1561, Sep. 2004.
- [34] D. Teneketzis, "On the structure of optimal real-time encoders and decoders in noisy communication," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 4017–4035, Sep. 2006.
- [35] J. Walrand and P. Varaiya, "Optimal causal coding—Decoding problems," *IEEE Trans. Inf. Theory*, vol. 19, no. 6, pp. 814–820, Nov. 1983.
- [36] T. Weissman and N. Merhav, "On causal source codes with side information," *IEEE Trans. Inf. Theory*, vol. 51, no. 11, pp. 4003–4013, Nov. 2005.
- [37] H. S. Witsenhausen, "On the structure of real-time source coders," *Bell Syst. Tech. J.*, vol. 58, no. 6, pp. 1437–1451, Jul./Aug. 1979.
- [38] S. Yüksel, "On optimal causal coding of partially observed Markov sources in single and multi-terminal settings," *IEEE Trans. Inf. Theory*, vol. 59, no. 1, pp. 424–437, Jan. 2013.
- [39] S. Yüksel, "Jointly optimal LQG quantization and control policies for multi-dimensional systems," *IEEE Trans. Autom. Control*, vol. 59, no. 6, pp. 1612–1617, Jun. 2014.
- [40] S. Yüksel and T. Başar, *Stochastic Networked Control Systems: Stabilization and Optimization Under Information Constraints*. Boston, MA, USA: Birkhäuser, 2013.
- [41] S. Yüksel and T. Linder, "Optimization and convergence of observation channels in stochastic control," *SIAM J. Control Optim.*, vol. 50, no. 2, pp. 864–887, 2012.
- [42] S. Yüksel and S. P. Meyn, "Random-time, state-dependent stochastic drift for Markov chains and application to stochastic stabilization over erasure channels," *IEEE Trans. Autom. Control*, vol. 58, no. 1, pp. 47–59, Jan. 2013.

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.

Serdar Yüksel (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 and decentralized control, information theory and applied probability.