# Jointly Optimal LQG Quantization and Control Policies for Multi-Dimensional Systems

Serdar Yüksel

***Abstract*—For controlled $\mathbb{R}^n$-valued linear systems driven by Gaussian noise under quadratic cost criteria, we investigate the existence and the structure of optimal quantization and control policies. For fully observed and partially observed systems, we establish the global optimality of a class of predictive encoders and show that an optimal quantization policy exists, provided that the quantizers allowed are ones which have convex codecells. Furthermore, optimal control policies are linear in the conditional estimate of the state, and a form of separation of estimation and control holds.**

***Index Terms*—Networked control systems, stochastic systems, quantization.**

## I. Jointly Optimal Encoding and Control Policies

### A. System Model

Consider a Linear Quadratic Gaussian (LQG) setup, where a sensor encodes its noisy information to a controller. Let $x_t \in \mathbb{R}^n$ and the evolution of the system be given by the following:

$$x_{t+1} = Ax_t + Bu_t + w_t, \qquad y_t = Cx_t + v_t. \tag{1}$$

Here, $\{w_t, v_t\}$ is a mutually independent, zero-mean i.i.d. Gaussian noise sequence, $u_t$ is an $\mathbb{R}^m$-valued control action, $y_t \in \mathbb{R}^p$ is the observation variable, and $A, B, C$ are matrices of appropriate dimensions. We assume that $x_0$ is a zero-mean Gaussian random variable.

As in Fig. 1, let there be an encoder who has access to the observation variable $y_t$, and who transmits his information to a receiver/controller, over a discrete noiseless channel with finite capacity.

*Definition 1.1:* Let $\mathcal{M} = \{1, 2, \ldots, M\}$ with $M = |\mathcal{M}|$. Let $\mathbb{A}$ be a (topological) space. A quantizer $Q(\mathbb{A}; \mathcal{M})$ is a Borel measurable map from $\mathbb{A}$ to $\mathcal{M}$. $\diamond$

When the spaces $\mathbb{A}$ and $\mathcal{M}$ are clear from context, we will denote the quantizer simply by $Q$.

Following [19], we refer by a *Composite Quantization (Coding) Policy* $\Pi^{comp}$, a sequence of functions $\{Q_t^{comp}((\mathbb{R}^p)^{t+1}; \mathcal{M}), t \geq 0\}$ which are causal such that the quantization output at time $t$, $q_t$, under $\Pi^{comp}$ is generated by a function of its local information, that is, a mapping measurable on the sigma-algebra generated by $\mathcal{I}_t^e = \{y_{[0,t]}\}$ to a finite set $\mathcal{M}$, which is the quantization output alphabet given by $\mathcal{M} := \{1, 2, \ldots, M\}$, for $0 \leq t \leq T-1$. Here, we have the notation for $t \geq 1$: $y_{[0,t-1]} = \{y_s, 0 \leq s \leq t-1\}$. Let $\mathbb{I}_t = (\mathbb{R}^p)^{t+1}$, be information spaces such that for all $t \geq 0$, the realizations satisfy $\mathcal{I}_t^e \in \mathbb{I}_t$. Thus, $Q_t^{comp} : \mathbb{I}_t \to \mathcal{M}$. As elaborated on in [19], we may express the policy $\Pi^{comp}$ as a composition of a *Quantization Policy* $\Pi^i$ and a *Quantizer*. A quantization policy $\mathcal{T}$ is a sequence of functions $\{T_t\}$, such that for each $t \geq 0$, $T_t$ is a mapping from the information space $\mathbb{I}_t$ to a space of quantizers $\mathbb{Q}_t$, to be specified below. A quantizer, subsequently is used to generate the quantizer output. A quantizer will be generated based on the common information at the encoder and the controller/receiver, and the quantizer will map the relevant private information at the encoder to the quantization output (see [17] and [19] for a similar reasoning). Such a separation in the design will also allow us to use the machinery of Markov Decision Processes to obtain a structural method to design optimal quantizers, to be clarified further, without any loss in optimality. Thus, with the information at the controller at time $t$ being $\mathcal{I}_t^c = \{q_{[0,t]}\}, \quad t \geq 0$, we can express the composite quantization policy as

$$Q_t^{comp}\left(\mathcal{I}_t^e\right) = \left(T_t\left(\mathcal{I}_t^c\right)\right)\left(\mathcal{I}_t^e \setminus \mathcal{I}_t^c\right). \tag{2}$$

We note that any composite quantization policy $Q_t^{comp}$ can be expressed in the form above; that is there is no loss in the set of possible such policies, since for any $Q_t^{comp}$, one could define

$$T_t\left(\mathcal{I}_t^c\right)(\cdot) := Q_t^{comp}\left(\mathcal{I}_t^c, \cdot\right).$$

Thus, we let the encoder have policy $\mathcal{T}$ and under this policy generate quantizer actions $\{Q_t, t \geq 0\}$, $Q_t \in \mathbb{Q}_t$ (hence, $Q_t(\mathbb{I}_t \setminus \mathcal{M}^t; \mathcal{M})$ is the quantizer used at time $t$ and the realization space of $\mathcal{I}_t^e \setminus \mathcal{I}_t^c$ is quantized). Under action $Q_t$, and given the local information, the encoder generates $q_t^i$, as the *quantization output* at time $t$. The receiver/controller, upon receiving the information from the encoders, generates its decision at time $t$, also causally: An admissible causal controller policy is a sequence of functions $\gamma = \{\gamma_t\}$ such that $u_t = \gamma_t(q_{[0,t]})$, with $\gamma_t : \mathcal{M}^{t+1} \to \mathbb{R}^m, \qquad t \geq 0$. We call such encoding and control policies, *causal* or *admissible*. Now, suppose that the goal is the computation of

$$\inf_{\Pi^{comp}} \inf_{\gamma} J(\Pi^{comp}, \gamma, T) \tag{3}$$

where

$$J(\Pi^{comp}, \gamma, T) := \frac{1}{T} E_{\nu_0}^{\Pi^{comp}, \gamma}\left[\sum_{t=0}^{T-1} x_t' Q x_t + u_t' R u_t\right].$$

Here, $Q \geq 0$ a positive semi-definite matrix, and $R > 0$ a positive definite matrix.

### B. Relevant Literature and Contributions

There is a large literature on jointly optimal quantization for the LQG problem dating back to early 1960s (see for example [5] and [9]). In this literature, references [1], [3], [6], [7], [12]–[14] have considered the optimal LQG quantization and control design with various results on the optimality or the lack of optimality (with detailed comparisons reported in [7] and [13]) of the separation principle with different assumptions in the setups and various (sometimes inconsistent) conclusions on the structural properties of optimal policies.

We now highlight the differences between our technical note and the most relevant contributions in the literature. (i) A relevant paper is [7] which establishes a separation result parallel to the findings in this

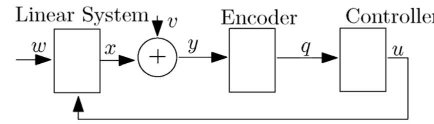


Fig. 1. Joint LQG optimal design of coding and control.

Manuscript received August 28, 2012; revised February 02, 2013, June 10, 2013, and October 27, 2013; accepted November 19, 2013. Date of publication November 28, 2013; date of current version May 20, 2014. This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC). This paper was presented in part at the Annual Allerton Conference on Communication, Control, and Computing, 2012. Recommended by Associate Editor L. Zaccarian.

The author is with the Department of Mathematics and Statistics, Queen's University, Kingston, ON K7L 3N6, Canada, (e-mail: yuksel@mast.queensu.ca).

Digital Object Identifier 10.1109/TAC.2013.2293414

technical note. Our technical note is different from [7] in that [7] assumes *a priori* a structural result on the encoders and the controllers. In particular, the encoders are restricted to use the most recent observations conditioned on the information at the controller. (ii) Our contribution is different from [13] in that, [13] establishes the optimality of a class of predictive coding schemes, however, the predictive encoders in [13] have memory and such a memory setting does not lead to a controlled Markov recursion with a fixed state-space: In this technical note, we show that restricting the memory of the encoder is without any loss, and extending the findings in [19], we obtain a dynamic program where the state space is the set of probability measures and the actions are the quantizers. In addition, we establish the existence of optimal quantization policies. (iii) Reference [1] considers the design of optimal encoders and controllers when there is a noisy channel between the encoder and the controller and the encoder has access to limited feedback information. When the information at the controller is nested in that at the encoder, the paper establishes the optimality of a class of predictive encoders (which refines the class considered in [13] further); however these coders also have memory. This paper also establishes an iterative method for the design of optimal encoders. (iv) References [3] and [14] are among the earliest contributions in the literature on jointly optimal LQG coding and control under information constraints. These papers establish separation results and optimal coding and control policies, where an innovations/predictive encoder structure is imposed *a priori*. The contribution here is different from [3] and [14] in that, this technical note establishes the optimality of such predictive coders, in addition to the existence results. (v) Reference [12] considers the design of optimal encoder and controllers for a partially observed LQG system under the assumption that the controller is memoryless. Since the memory structure of the controller is not expanding, Theorem 2.2 is not applicable (see also Section II-C of [17] and [19] for similar discussions).

Regarding structural results on optimal causal (zero-delay or real-time) coding, there have been many studies, see [19] and [20] for a review. For multi-stage settings, [10] has considered the existence problem for optimal quantization of control-free Markov sources for a class of Markov sources driven by an additive Gaussian noise under the restriction that the quantizers have convex codecells. This class of systems also includes the setup contained here except that control is not available in the systems considered in [10]. Also for optimal multi-stage vector quantizers, [4] has obtained existence results for an infinite horizon setup under a uniform boundedness assumption on the reconstruction levels. [12] established the existence of optimal coding and quantizer policies for the LQG setup under the assumption that the controller is memoryless. Chapter 10 of [20] provides a comprehensive overview of the results mentioned above and presented in this technical note.

Essentially, in the context of LQG settings, the paper unifies many of the results in the literature providing the following systematic development: (i) The paper provides a structural result for optimal encoders for controlled Markov sources taking values in a complete, separable metric space extending the findings in [16] and [19], and building on this structural result, (ii) it establishes a separation theorem between coding and control which is new to our knowledge in its generality, and (iii) it establishes an existence theorem for optimal quantizers (building on [10], [20] and [21]) subject to a restriction on the set of admissible quantizers, and also establishes the structure of optimal control policies. The rest of the technical note is structured as follows. In Section II, we establish the structure of optimal causal (zero-delay) coding policies for fully and partially observed controlled Markov sources. Section II-B introduces the set of quantizer actions. In Section III, we consider the fully observed setting in (1) (that is with $y_t = x_t$) and obtain the structure of optimal control policies. In Section IV, we establish the existence of optimal quantization policies. The partially observed setting is discussed in Section V.

## II. Structural Results for Optimal Zero-Delay Codes for Controlled Markov Sources and the Set of Quantizers

### A. Structural Results

In this section, toward obtaining a solution to (3), we develop structural results for optimal causal composite quantization policies. Consider the fully observed system

$$x_{t+1} = f(x_t, u_t, w_t), \qquad y_t = x_t, \quad t = 0, 1, \cdots \tag{4}$$

where the realizations satisfy $x_t \in \mathbb{X}, u_t \in \mathbb{U}$, with $\mathbb{X}, \mathbb{U}$ being complete, separable, metric (that is Polish) spaces (thus, including spaces such as $\mathbb{R}^n$ or a countable set). Suppose that the goal is the minimization

$$\inf_{\Pi^{comp}} \inf_{\gamma} E_{\nu_0}^{\Pi^{comp},\gamma} \left[ \sum_{t=0}^{T-1} c(x_t, u_t) \right] \tag{5}$$

over all policies $\Pi^{comp}, \gamma$ with the random initial condition $x_0$ having probability measure $\nu_0$. Here $c(\cdot,\cdot)$, is a measurable function and $u_t = \gamma_t(q_{[0,t]})$ for $t \geq 0$. Here, the information and quantization restrictions are as stated in Section I. Structural results on optimal quantization policies for such controlled Markov sources have been studied by Walrand and Varaiya [16] in the context of finite control and action spaces and by Mahajan and Teneketzis [11] for control over noisy channels, also for finite state-actions space settings (see Section V in [11] for a brief discussion on continuous state-spaces). The following extend the finite state space analysis of Walrand and Varaiya [16] to more general spaces. The proofs of the results below essentially follow from Theorems 2.4 and 2.5 in [19] with additional technical intricacies due to the presence of control actions. The first one can be regarded as an extension of Witsenhausen's structural theorem [18], and the second one can be regarded as an extension of the results of Walrand and Varaiya [17] (see also [15]). For complete proofs, see [20, Chapter 10].

*Theorem 2.1:* For system (4), under the information structure described in the previous section and the objective given in (5), any composite quantization policy (with a given control policy) can be replaced, without any loss in performance, by one which only uses $x_t$ and $q_{[0,t-1]}$ at time $t \geq 1$ while keeping the control policy unaltered. This can be expressed as a quantization policy which only uses $q_{[0,t-1]}$ to generate a quantizer, where the quantizer uses $x_t$ to generate the quantization output at time $t$. ◇

Let $\mathcal{P}(\mathbb{X})$ denote the set of probability measures on $\mathcal{B}(\mathbb{X})$ (where $\mathcal{B}(\mathbb{X})$ denotes the Borel $\sigma$-field on $\mathbb{X}$) under the topology of weak convergence (see Section II-B in [19] for a discussion on the use of such a topology) and define $\pi_t \in \mathcal{P}(\mathbb{X})$ to be the regular conditional probability measure given by $\pi_t(\cdot) = P(x_t \in \cdot | q_{[0,t-1]}, u_{0,t-1})$ or since the control actions are determined by quantizer outputs given a deterministic control policy (we note that such policies are optimal without any loss), $\pi_t(\cdot) = P(x_t \in \cdot | q_{[0,t-1]})$, that is $\pi_t(A) = P(x_t \in A | q_{[0,t-1]})$, $A \in \mathcal{B}(\mathbb{X})$.

*Theorem 2.2:* For system (4), under the information structure described in the previous section and the objective given in (5), any composite quantization policy can be replaced, without any loss in performance, by one which only uses the conditional probability measure $\pi_t(\cdot) = P(x_t \in \cdot | q_{[0,t-1]})$, the state $x_t$, and the time information $t$, at time $t$. This can be expressed as a quantization policy which only uses $\{\pi_t, t\}$ to generate a quantizer, where the quantizer uses $x_t$ to generate the quantization output at time $t$. ◇

We can also consider the partially observed setting. Instead of (4), the system considered is a discrete-time system described by

$$x_{t+1} = f(x_t, u_t, w_t), \qquad y_t = g(x_t, v_t), t = 0, 1, \cdots \tag{6}$$

for (Borel) measurable functions $f, g$, with $\{w_t, v_t, i = 1, 2\}$ i.i.d., mutually independent noise processes and $x_0$ a random variable with probability measure $\nu_0$. Here, we let $x_t \in \mathbb{X}, u_t \in \mathbb{U}$, and $y_t \in \mathbb{Y}$, where $\mathbb{X}, \mathbb{U}, \mathbb{Y}$ are Polish spaces.

For such a setup, results similar to Theorems 2.1 and 2.2 hold: (i) We first note that the space $\mathcal{P}(\mathbb{X})$ under the topology of weak convergence is itself a Polish space. (ii) We then recognize the fact that $\{\tilde{\pi}_t\}$ forms a controlled Markov source, and that (iii) the cost function can be expressed as $\tilde{c}(\tilde{\pi}, u) = \int_{\mathbb{X}} c(x, u)\tilde{\pi}(dx)$, where $\tilde{c} : \mathcal{P}(\mathbb{X}) \times \mathbb{U} \to \mathbb{R}$. Thus, one could directly apply Theorems 2.1 and 2.2 to obtain structural results, see [19] for further discussions. For a further case where the decoder's memory is limited or imperfect, the results may apply by replacing the full information considered so far at the receiver with the limited one with additional assumptions on the decoder's update of its memory.

### *B. The Space of Quantizers*

In this section, we construct a topology on the set of quantizers which will be used in the subsequent analysis.

*Definition 2.1:* An $M$-cell quantizer $Q$ on $\mathbb{R}^n$ is a (Borel) measurable mapping $Q : \mathbb{R}^n \to \mathcal{M}$, and $\mathcal{Q}$ denotes the collection of all $M$-cell quantizers on $\mathbb{R}^n$. ⋄

Note that each $Q \in \mathcal{Q}$ is uniquely characterized by its *quantization cells* (or bins) $B_i = \{x : Q(x) = i\}, i = 1, \ldots, M$ which form a measurable partition of $\mathbb{R}^n$. As in [21], we allow for the possibility that some of the cells of the quantizer are empty.

As discussed in [21], a quantizer $Q$ with cells $\{B_1, \ldots, B_M\}$ can also be characterized as a stochastic kernel $Q$ from $\mathbb{R}^n$ to $\{1, \ldots, M\}$ defined by

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

We will endow the quantizers with a topology induced by such a stochastic kernel interpretation. If $P$ is a probability measure on $\mathbb{R}^n$ and $Q$ is a stochastic kernel from $\mathbb{R}^n$ to $\mathcal{M}$, then $PQ$ denotes the resulting joint probability measure on $\mathbb{R}^n \times \mathcal{M}$. Let $\mathcal{P}(\mathbb{R}^{\mathbb{N}})$ denote the family of all probability measures on $(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$ for some $N \in \mathbb{N}$. Let $\{\mu_n, n \in \mathbb{N}\}$ be a sequence in $\mathcal{P}(\mathbb{R}^N)$. It is said to converge to $\mu \in \mathcal{P}(\mathbb{R}^N)$ *weakly* if $\int_{\mathbb{R}^N} c(x)\mu_n(dx) \to \int_{\mathbb{R}^N} c(x)\mu(dx)$ for every continuous and bounded $c : \mathbb{R}^N \to \mathbb{R}$. The following sequential convergence notion is considered.

*Definition 2.2:* [21] A quantizer sequence $Q_n$ converges to $Q$ weakly at $P$ ($Q_n \to Q$ weakly at $P$) if $PQ^n \to PQ$ weakly. ⋄

Consider the set of probability measures

$$\Theta := \{\zeta \in P(\mathbb{R}^n \times \mathcal{M}) : \zeta = PQ, Q \in \mathcal{Q}\}$$

on $\mathbb{R}^n \times \mathcal{M}$ having fixed input marginal $P$, equipped with weak topology. This is the Borel measurable set of the extreme points of the set of probability measures on $\mathbb{R}^n \times \mathcal{M}$ with a fixed input marginal $P$ (see [2]). In view of this observation, and that the class of quantization policies which admit the structure suggested in Theorem 2.2 is an important one, we define

$$\Pi_W := \{\Pi^{comp} = \{Q_t^{comp}, t \geq 0\} : \quad \exists \Upsilon_t : \mathcal{P}(\mathbb{X}) \to \mathcal{Q} \\ Q_t^{comp}(I_t) = (\Upsilon_t(\pi_t))(x_t), \forall I_t\} \tag{7}$$

to represent this class of policies. Here, the input measure is time varying and is given by $\pi_t$.

## III. Fully Observed LQG: Separation of Estimation Error and Control

We now consider the original LQG problem given in (1) with the cost function given in (3), but with a fully observed setup where $y_t = x_t$. By Theorem 2.2, an optimal composite quantization policy will be within the class $\Pi_W$. Let us fix such a composite quantization policy. In the following, we adopt a dynamic programming approach and establish that the optimal controller is linear in its estimate. This fact applies naturally for the terminal time stage control. That this also applies for the previous time stages follows from dynamic programming as we observe in the following.

First consider the terminal time $t = T-1$. For this time stage, to minimize $E[x_t'Qx_t + u_t'Ru_t]$, the optimal control is $u_{T-1} = 0$ a.s. To obtain a solution for $t = T-2$, we look for a solution to: $\min_{\gamma_t} E[(x_t'Qx_t + u_t'Ru_t + E[(Ax_t + Bu_t + w_t)'Q(Ax_t+Bu_t + w_t)|\mathcal{I}_t^c, u_t])|\mathcal{I}_t^c]$. By completing the squares, and using the *Orthogonality Principle*, we obtain that the optimal control is linear and is given by $u_{T-2} = L_{T-2}E[x_{T-2}|q_{[0,T-2]}]$, with $L_{T-2} = -R^{-1}B'QA$.

For $t < T-2$, to obtain the solutions, we will first establish that the estimation errors are uncorrelated. Towards this end, define for $1 \leq t \leq T-1$ (recall that the control actions are determined by the quantizer outputs): $\mathcal{I}_t^c = \{q_{[0,t]}, u_{[0,t-1]}\}$ and note that

$$\tilde{m}_{t+1} := E[x_{t+1}|\mathcal{I}_{t+1}^c] = E[Ax_t + Bu_t + w_t|\mathcal{I}_{t+1}^c].$$

It then follows that

$$\begin{aligned} &\tilde{m}_{t+1} \\ &= E\left[x_{t+1}|\mathcal{I}_{t+1}^c\right] = E[x_{t+1} - E\left[x_{t+1}|\mathcal{I}_t^c\right] + E\left[x_{t+1}|\mathcal{I}_t^c\right]|\mathcal{I}_{t+1}^c] \\ &= A\tilde{m}_t + Bu_t + \left(E\left[x_{t+1}|\mathcal{I}_{t+1}^c\right] - E[x_{t+1}|\mathcal{I}_t^c]\right) \\ &= A\tilde{m}_t + Bu_t + \bar{w}_t \end{aligned} \tag{8}$$

with $\bar{w}_t = (E[x_{t+1}|\mathcal{I}_{t+1}^c] - E[x_{t+1}|\mathcal{I}_t^c])$. The variable $\bar{w}_t$ is orthogonal to the control action variable $u_t$, as control actions are determined by the past quantizer outputs and iterated expectation leads to the result that conditioned on $\mathcal{I}_t^c$, $\bar{w}_t$ is zero mean, and is orthogonal to $\mathcal{I}_t^c$.

Now, for going into earlier time stages, the dynamic programming recursion for linear systems driven by an uncorrelated noise process would normally apply, since the estimate process $\tilde{m}_t$ is driven an uncorrelated noise (though, not necessarily an independent) process $E[x_{t+1}|\mathcal{I}_{t+1}^c] - E[x_{t+1}|\mathcal{I}_t^c]$. However, this lack of independence may be important, as elaborated on in [13]. Using the completion of the squares method, we can establish that the optimal controller at any time will be linear in its estimate, provided that the random variable $\bar{w}_t'Q\bar{w}_t$ does not depend on $u_k, k \leq t$ under an optimal coding policy for all time stages. A sufficient condition for this is that the encoder is a predictive one (see [3], [13] and [14] for related discussions).

*Definition 3.1:* A predictive quantizer policy is one where for each time stage $t$, the quantization has the form that the quantizer at all time stages subtracts the effect of the past control terms, that is, at time $t$ it has the form $Q_t(x_t - \sum_{k=0}^{t-1} A^{t-k-1}Bu_k)$, and the past control terms are added at the receiver. Hence, the encoder quantizes a control free process, defined by:

$$\bar{x}_{t+1} = A\bar{x}_t + w_t, \tag{9}$$

the receiver generates the quantized estimate and adds $\sum_{k=0}^{t-1} A^{t-k-1}Bu_k$ to compute the estimate of the state at time $t$. ⋄

A predictive quantizer is depicted in Fig. 2. One question, which has not been addressed in [1], [3], [13], [14], and [7], is *whether restriction to this class of quantization policies (given in Definition 3.1) is without loss*. We have the following key lemma.

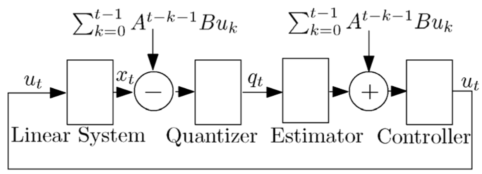


Fig. 2. For the LQG problem, a predictive encoder is optimal.

*Lemma 3.1:* For problem (3), for any quantizer policy in class $\Pi_W$ (which is without any loss as a result of Theorem 2.2), there exists a predictive quantizer in the sense of Definition 3.1 which attains the same performance under an optimal control policy for problem (3). $\diamond$

*Proof:* We apply dynamic programming. Let for the time-stage, $t = T-2$, $f_t(q_{[0,t-1]}) := \sum_{k=0}^{t-1} A^{t-k-1}Bu_k$. If the policy considered is in $\Pi_W$, the quantization policy is of the form $Q_t(\bar{x}_t + \sum_{k=0}^{t-1} A^{t-k-1}Bu_k, P(\bar{x}_t + \sum_{k=0}^{t-1} A^{t-k-1}Bu_k \ \in \cdot|q_{[0,t-1]}))$. For this time-stage, there exists an optimal decoder and controller for which a sufficient statistic for the optimal control policy is $E[x_t|q_{[0,t]}]$. Observe that

$$\begin{aligned}E\left[\bar{x}_t + f_t\left(q_{[0,t-1]}\right)|q_{[0,t]}\right] &= E\left[\bar{x}_t|q_{[0,t]}\right] + f_t\left(q_{[0,t-1]}\right)\\ &= E\big[\bar{x}_t|q_{[0,t-1]},q_t\big]+f_t\big(q_{[0,t-1]}\big).\end{aligned}$$

The quantization output $q_t$ represents the bin information for $x_t$. By shifting each of the finitely many quantizer bins by $f_t(q_{[0,t-1]})$, a new quantizer which quantizes $\bar{x}_t$ (see (9)), can generate the same bin information on $\bar{x}_t$ through $q_t$, that is, can encode the event $1_{\{\bar{x}_t \in B_i\}}$ for some bin $B_i$ almost surely. Hence, there is no information loss due to the elimination of the past control actions. This new quantizer, by adding $f_t(q_{[0,t-1]})$ to the receiver output, generates the same conditional estimate of the state as the original quantizer. Thus, corresponding to a quantizer policy in $\Pi_W$ at time $t$, there exists a quantizer of the form $\tilde{Q}_t(\bar{x}_t, P(\bar{x}_t \in \cdot|q_{[0,t-1]}))$ with the following property: The estimation error realization and hence the estimation is the same almost surely. Furthermore, under such a predictive scheme (with $\tilde{Q}_t(\bar{x}_t, P(\bar{x}_t \in \cdot|q_{[0,t-1]}))$ fixed), $\bar{w}_{T-2}$ does not depend on the control actions applied earlier; for a predictive quantizer, the error only depends on the control-free process. By the analysis following (8), an optimal controller at time $t = T-3$ will then use $E[x_{T-3}|q_{[0,T-3]}]$ as a sufficient statistic (note that the optimal controls for $t = T-1$ and $t = T-2$ have been derived earlier). To design the quantizer at $T-3$, by a similar reasoning as above for $t = T-1$ and $T-2$, a predictive quantizer can be used so that $\bar{w}_k, k \geq T-3$ is independent of the control actions applied earlier, inductively leading to the optimality of linear policies, for all $t \geq 0$. $\diamond$

*Remark 3.1:* We note that the structure in Definition 3.1 separates the estimation from the control process in the sense that the estimation errors do not depend on the control actions or policies. Hence, there is no *dual effect* of the control actions, in that the estimation error at any given time does not depend on the past applied control actions. $\diamond$

We have thus established above that the optimal control is linear for all time stages, by the proof of Lemma 3.1. We have the following result (see also [13] which assumes a different version of the structure given in Definition 3.1 for an essentially identical result on optimal control policies):

*Theorem 3.1:* For the minimization problem (3), with the new effective state dynamics in (8), an optimal control policy is given by $u_t = L_t E[x_t|q_{[0,t]}]$, where $L_t = -(R + B'P_{t+1}B)^{-1}B'K_{t+1}A$, $P_t = A_t'K_{t+1}B(R + B'K_{t+1}B)^{-1}B'K_{t+1}A$, and $K_t = A_t'K_{t+1}A_t - P_t + Q$, with $K_T = P_{T-1} = 0$. $\diamond$

Therefore, we obtain for $t \geq 0$, the unnormalized value function for any time stage $t$ as

$$\begin{aligned}J_t\left(\mathcal{I}_t^c\right) = E&\left[x_t'K_tx_t|\mathcal{I}_t^c\right]\\ &+ \sum_{k=t}^{T-1}\big(E\left[(x_k - E\left[x_k|\mathcal{I}_k^c\right])'Q\left(x_k - E\left[x_k|\mathcal{I}_k^c\right]\right)\right]\\ &\qquad + E\left[\bar{w}_k'K_{k+1}\bar{w}_k\right]\big)\end{aligned}$$

with $J(\Pi^{comp},\gamma,T) = (1/T)J_0(\mathcal{I}_0^c)$. To obtain a more explicit expression for the value function $J_t$, we have the following analysis. Given a positive definite matrix $\Lambda$ define an inner-product as $\langle z_1, z_2\rangle_\Lambda = z_1'\Lambda z_2$, and the norm generated by this inner-product as $\|z\|_\Lambda = \sqrt{z'\Lambda z}$. We now note the following:

$$\begin{aligned}&E\left[\|E\left[x_{t+1}|\mathcal{I}_{t+1}^c\right] - E[x_{t+1}|\mathcal{I}_t^c]\|_\Lambda^2\right]\\ &=E\Big[\|((E\left[x_{t+1}|\mathcal{I}_{t+1}^c\right]-x_{t+1})+(x_{t+1}-E\left[x_{t+1}|\mathcal{I}_t^c\right]))\|_\Lambda^2\Big]\\ &=E\Big[\|(E[x_{t+1}|\mathcal{I}_{t+1}^c]-x_{t+1})\|_\Lambda^2\Big]+E\Big[\|(x_{t+1}-E[x_{t+1}|\mathcal{I}_t^c])\|_\Lambda^2\Big]\\ &\quad+2E\left[\langle(E\left[x_{t+1}|\mathcal{I}_{t+1}^c\right]-x_{t+1}),(x_{t+1}-E\left[x_{t+1}|\mathcal{I}_t^c\right])\rangle_\Lambda\right].\end{aligned}$$

Note that

$$\begin{aligned}&E\left[\langle(E\left[x_{t+1}|\mathcal{I}_{t+1}^c\right] - x_{t+1}),(x_{t+1} - E\left[x_{t+1}|\mathcal{I}_t^c\right])\rangle_\Lambda\right]\\ &\quad= E\left[-\langle(E\left[x_{t+1}|\mathcal{I}_{t+1}^c\right] - x_{t+1}),(E\left[x_{t+1}|\mathcal{I}_t^c\right])\rangle_\Lambda\right.\\ &\qquad\left.+\langle(E\left[x_{t+1}|\mathcal{I}_{t+1}^c\right] - x_{t+1}),(x_{t+1})\rangle_\Lambda\right]\\ &\quad= E\left[\langle(E\left[x_{t+1}|\mathcal{I}_{t+1}^c\right] - x_{t+1}),(x_{t+1})\rangle_\Lambda\right]\\ &\quad= -E\left[\|(E\left[x_{t+1}|\mathcal{I}_{t+1}^c\right] - x_{t+1})\|_\Lambda^2\right]\end{aligned}\tag{10}$$

where (10) follows from the orthogonality property of minimum mean-square estimation and that $E[x_{t+1}|\mathcal{I}_t^c]$ is measurable on $\sigma(\mathcal{I}_{t+1}^c)$, the sigma-field generated by $\mathcal{I}_{t+1}^c$.

Therefore, we have

$$\begin{aligned}&E\left[\|(E\left[x_{t+1}|\mathcal{I}_{t+1}^c\right] - E\left[x_{t+1}|\mathcal{I}_t^c\right])\|_{K_{t+1}}^2\right]\\ &\quad= -E\left[(x_{t+1} - E\left[x_{t+1}|\mathcal{I}_{t+1}^c\right])'\left(K_{t+1}\right)\right.\\ &\qquad\times\left(x_{t+1} - E\left[x_{t+1}|\mathcal{I}_{t+1}^c\right]\right)\\ &\qquad+E\big[(x_t - E\left[x_t|\mathcal{I}_t^c\right])'\left(A'K_{t+1}A\right)\left(x_t - E\left[x_t|\mathcal{I}_t^c\right]\right)\big]\\ &\qquad\left.+E[w'K_{t+1}w]\right].\end{aligned}$$

After some algebra, for $t < T-1$, the optimal cost can be written as

$$\begin{aligned}&J_t\left(\mathcal{I}_t^c\right)\\ &\quad= E\left[x_t'K_tx_t|\mathcal{I}_t^c\right]\\ &\qquad+ E\left[(x_t - E\left[x_t|\mathcal{I}_t^c\right])'\left(Q + A'K_{t+1}A\right)\left(x_t - E\left[x_t|\mathcal{I}_t^c\right]\right)\right]\\ &\qquad+ \sum_{k=t+1}^{T-1} E\left[(x_k - E\left[x_k|\mathcal{I}_k^c\right])'\left(Q + A'K_{k+1}A - K_k\right)\right.\\ &\qquad\qquad\left.\times\left(x_k - E\left[x_k|\mathcal{I}_k^c\right]\right)\right]\\ &\qquad+ \sum_{k=t}^{T-1} E\left[w_k'K_{k+1}w_k\right].\end{aligned}\tag{11}$$

Given the optimal control solution, we address the optimal quantization problem.

## IV. Existence of Optimal Quantization Policies

In (11) above, we have separated the costs due to control and quantization. Since under the optimal structure considered, control actions do not affect the estimation performance, for the optimal quantization policy we can effectively consider the setting where in (1), $u_t = 0$ and the quantizer is designed for this system. Hence, we consider below the system $x_{t+1} = Ax_t + w_t$.

We first note that, for $K > 0$

$$\begin{aligned} E&\left[(x_t - E[x_t|\mathcal{I}_t^c])' K (x_t - E[x_t|\mathcal{I}_t^c]) |\mathcal{I}_{t-1}^c\right] \\ &= \sum_{i\in\mathcal{M}} P\left(q_t = i|q_{[0,t-1]}\right) \\ &\quad\times \inf_{u\in\mathbb{R}^n}\left(\int_{\mathbb{R}^n} P\left(dx_t|q_{[0,t-1]}, q_t = i\right)(x_t - u)'K(x_t - u)\right) \\ &= \sum_{i\in\mathcal{M}} \inf_{\gamma_t(i)} \int_{\mathbb{R}^n} 1_{\{q_t=i\}}\pi_t(dx)\,(x_t - \gamma_t(i))' K (x_t - \gamma_t(i)). \end{aligned} \tag{12}$$

Thus, from (11), for $T \in \mathbb{N}$, we can define a cost to be minimized, $J(\Pi^{comp}, T)$, as

$$E_{\nu_0}^{\Pi^{comp}}\left[\frac{1}{T}\left(x_0'K_0x_0 + \sum_{t=0}^{T-1} c_t(\pi_t, Q_t) + E\left[w_t'K_{t+1}w_t\right]\right)\right]$$

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

$$c_t(\pi_t, Q_t) = \sum_{i\in\mathcal{M}} \inf_{\gamma_t(i)} \int_{\mathbb{R}^n} 1_{\{q_t=i\}}\pi_t(dx_t) \times (x_t - \gamma_t(i))' P_t (x_t - \gamma_t(i))$$

with $\tilde{\gamma} = \{\gamma_t, t \geq 0\}$ now denoting the receiver policy and $P_t = (Q + A'K_{t+1}A - K_t)$, and $P_0 = Q + A'K_1A$. Note that

$$E\left[1_{\{q_t=i\}}(\bar{x}_t - \gamma_t(i))' P_t (\bar{x}_t - \gamma_t(i))\right]$$

is minimized by the conditional expectation given the bin information. As a consequence, an optimal receiver and hence control policy always exists. In the analysis for an optimal quantization policy, as was also motivated in [21], we will restrict the quantizers to have convex codecells.

*Assumption 4.1:* The quantizers have convex codecells with at most a given number, $M$, of cells. The set of such quantizers is denoted by $\mathcal{Q}_c(M)$. ◇

We note that the assumption on convex codecells is adopted for technical reasons, and it may lead to a loss in optimality. However, such quantizers are very desirable in practice due to their parametric representability: As discussed in [8], 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 for all $i = 1, \ldots, M$

$$B_i \subset \bigcap_{j\neq i} H_{i,j}.$$

Since $\bar{B}_i := \bigcap_{j\neq i} H_{i,j}$ is a closed convex polytope for each $i$, if the input probability measure $P$ admits a density function, then one has $P(\bar{B}_i \setminus B_i) = 0$ for all $i = 1, \ldots, M$. One can thus obtain a ($P$-a.s) representation of a quantizer $Q$ by the $M(M-1)/2$ hyperplanes $h_{i,j} = H_{i,j} \cap H_{j,i}$. One can represent a hyperplane in $\mathbb{R}^n$ by a vector of $n+1$ components $a_1, a_2, \ldots, b$ with $\sum_k |a_k|^2 = 1$, and $h = \{x \in \mathbb{R}^n : \sum a_i x_i = b\}$. See [8] and [21] for further discussions on such quantizers.

Let $\Pi_W^C$ denote the set of all policies in $\Pi_W$ (defined in (7)) which in addition satisfy Assumption 4.1 (i.e., $Q_t \in \mathcal{Q}_c(M)$ for all $t \geq 0$). The properties of conditional probability lead to the filtering expression:

$$\pi_t(dx_t) = \frac{\int_{x_{t-1}} \pi_{t-1}(dx_{t-1})P(q_{t-1}|\pi_{t-1}, x_{t-1})P(dx_t|x_{t-1})}{\int_{x_{t-1}}\int_{x_t} \pi_{t-1}(dx_{t-1})P(q_{t-1}|\pi_{t-1}, x_{t-1})P(dx_t|x_{t-1})}.$$

Here, the term $P(q_{t-1}|\pi_{t-1}, x_{t-1})$ is determined by the quantizer action $Q_{t-1}$. With $\mathcal{P}(\mathbb{R}^n)$ denoting the set of probability measures on $\mathcal{B}(\mathbb{R}^n)$ under weak convergence topology, the conditional probability measure process and the quantization process $(\pi_t, Q_t)$ form a controlled Markov process in $\mathcal{P}(\mathbb{R}^n) \times \mathcal{Q}_c(M)$ [19].

*Theorem 4.1:* Under Assumption 4.1, there exists an optimal composite coding policy in $\Pi_W^C$ such that

$$\inf_{\Pi^{comp}\in\Pi_W^C} J(\Pi^{comp}, T)$$

is achieved. With, $J_0'(\pi) := \min_{\Pi^{comp}\in\Pi_W^C} J(\Pi^{comp}, T)$, the following dynamic programming recursion holds for $0 \leq t \leq T-1$:

$$TJ_t'(\pi_t) = \min_{Q\in\mathcal{Q}_c(M)}\left(c_t(\pi_t, Q_t) + TE\left[J_{t+1}'(\pi_{t+1})|\pi_t, Q_t\right]\right)$$

with $J_T'(\cdot) = 0$. Furthermore, the optimal control policy is linear in the conditional estimate and is given in Theorem 3.1. ◇

The proof of Theorem 4.1 follows from the separation argument considered since one can consider a control-free Markov source which is to be quantized. Therefore, the existence result follows from [10] (see also [20]) which considers a control-free setting. We also note that, even though [10] assumes a time-invariant $P_t$, the analysis is identical.

## V. Partially Observed Case

In this section, we consider the partially observed model (1) with $W = E[w_tw_t'], V = E[v_tv_t']$.

To obtain a solution, we again first separate the estimation and control terms, as in the fully observed case. The solution to the control terms then will follow from classical results in LQG theory. The solution for the quantization component will follow from the results earlier and Theorem 4.1 in [19]. Define $\bar{m}_t := E[x_t|y_{[0,t]}]$, which is computed through a Kalman Filter. Recall that by the Kalman Filter with $\Sigma_{0|-1} = E[x_0x_0']$ and for $t \geq 0$,

$$\begin{aligned}\Sigma_{t+1|t} = A\Sigma_{t|t-1}A' + W - (A\Sigma_{t|t-1}C') \\ \times(C\Sigma_{t|t-1}C' + V)^{-1}(C\Sigma_{t|t-1}A')\end{aligned}$$

the following recursion holds for $t \geq 0$ and with $\tilde{m}_{-1} = 0$:

$$\begin{aligned}\bar{m}_t = A\bar{m}_{t-1} + Bu_{t-1} + \Sigma_{t|t-1}C' \\ \times(C\Sigma_{t|t-1}C' + V)^{-1}(CA(x_{t-1} - \tilde{m}_{t-1}) + v_t).\end{aligned}$$

Note that the cost

$$\inf_{\Pi^{comp}} \inf_{\gamma} J(\Pi^{comp}, \gamma, T) \tag{13}$$

with $J(\Pi^{comp}, \gamma, T) = (1/T)E_{\nu_0}^{\Pi^{comp},\gamma}[\sum_{t=0}^{T-1} x_t'Qx_t + u_t'Ru_t]$, can be written equivalently as

$$\frac{1}{T}E_{\nu_0}^{\Pi^{comp},\gamma}\left[\sum_{t=0}^{T-1} \bar{m}_t'Q\bar{m}_t + u_t'Ru_t\right] + \frac{1}{T}E_{\nu_0}\left[\sum_{t=0}^{T-1}(x_t - \bar{m}_t)'Q(x_t - \bar{m}_t)\right]$$

since the quadratic error $(x_t - \bar{m}_t)'Q(x_t - \bar{m}_t)$ is independent of the coding or the control policy (and only depends on the estimation performance at the encoder). Thus, the process $(\bar{m}_t, \Sigma_{t+1|t})$ and $u_t$ form a controlled Markov chain and we can invoke Theorem 2.2: Any causal quantizer, can be, without any loss replaced with one in $\Pi_W$ (where the state is now $(\bar{m}_t, \Sigma_{t+1|t})$ instead of $x_t$) as a consequence of Theorem 2.2. Furthermore, any quantizer in $\Pi_W$ can be replaced without any loss with a predictive quantizer with the new state $\bar{m}_t$, as

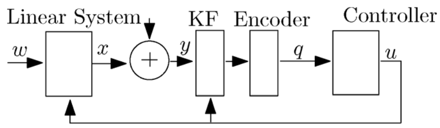


Fig. 3. Separated structure of the encoder above involving a Kalman Filter (KF) is optimal. Here, the encoder is a predictive encoder without any loss.

a consequence of Lemma 3.1 applied to the new state with identical arguments: Observe that the past control actions do not affect the evolution of $\Sigma_{t+1|t}$.

*Theorem 5.1:* For the minimization problem (13), the optimal control policy is given by $u_t = L_t E[x_t|q_{[0,t]}]$, where $L_t = -(R + B'P_{t+1}B)^{-1}B'K_{t+1}A$, $P_t = A'_t K_{t+1}B(R + B'K_{t+1}B)^{-1}B'K_{t+1}A$, and $K_t = A'_t K_{t+1}A_t - P_t + Q$, with $K_T = P_{T-1} = 0$. The optimal cost is given by $(1/T)J_0(\Pi^{comp}, T)$, where

$$\begin{aligned} &J_0(\Pi^{comp}, T) \\ &\quad = E\left[x'_0 K_0 x_0\right] \\ &\qquad + E\left[(x_0 - E\left[x_0|\mathcal{I}^c_0\right])'\,(Q + A'K_1A)\,(x_0 - E\left[x_0|\mathcal{I}^c_0\right])\right] \\ &\qquad + \sum_{t=1}^{T-1} E\left[(x_t - E\left[x_t|\mathcal{I}^c_t\right])'\,(Q + A'K_{t+1}A - K_t)\right. \\ &\qquad\qquad \left.\times\,(x_t - E\left[x_t|\mathcal{I}^c_t\right])\right] \\ &\qquad + \sum_{t=0}^{T-1} E\left[(x_t - \bar{m}_t)'Q(x_t - \bar{m}_t) + w'_t K_{t+1} w_t\right]. \end{aligned}$$

Now that we have separated the cost terms, and given that we can use a predictive encoder without any loss, we have the following.

*Theoram 5.2:* For the minimization of the cost in (3), any composite quantization policy can be replaced, without any loss in performance, by an encoder which only uses the output of the Kalman Filter and the information available at the receiver. Furthermore, any causal coder can be replaced with one which only uses the conditional probability on $\bar{m}_t$, $P(d\bar{m}_t|q_{[0,t-1]})$, and the realization $(\bar{m}_t, \Sigma_{t|t-1}, t)$ at time $t$ (see Fig. 3). An optimal quantization policy exists in $\Pi^C_W$. $\diamond$

## VI. Conclusion

In this technical note, joint optimization of encoding and control policies is investigated for the LQG problem. Global optimality of predictive encoders is established and it is shown that separation of estimation and control applies. Furthermore, an optimal quantizer is shown to exist under mild technical assumptions on the space of policies considered and an optimal control policy is linear in its conditional estimate. Results have been extended to the partially observed case, where the structure of optimal coding and control policies is presented. As a side result, towards obtaining the main results of the technical note, structural results in the literature for optimal causal (zero-delay) quantization of Markov sources is extended to systems driven by control.

Future work will focus on the relaxation of the convex-codecell assumption (Assumption 4.1) in Theorem 4.1. Another direction is the extension of the results here to discrete noisy channels with noiseless feedback, for which the separation results apply identically; see [20] for some related discussion.

## References

[1] 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.
[2] V. S. Borkar, “White-noise representations in stochastic realization theory,” *SIAM J. Control and Optimiz.*, vol. 31, pp. 1093–1102, 1993.
[3] V. S. Borkar and S. K. Mitter, “LQG control with communication constraints,” in *Kailath Festschrift*. Boston, MA: Kluwer, 1997.
[4] V. S. Borkar, S. K. Mitter, and S. Tatikonda, “Optimal sequential vector quantization of Markov sources,” *SIAM J. Control and Optimiz.*, vol. 40, pp. 135–148, 2001.
[5] R. E. Curry, *Estimation and Control With Quantized Measurements*. Cambridge, MA: MIT Press, 1969.
[6] T. R. Fischer, “Optimal quantized control,” *IEEE Trans. Autom. Control*, vol. AC-27, no. 8, pp. 996–998, Aug. 1982.
[7] M. Fu, “Lack of separation principle for quantized linear quadratic Gaussian control,” *IEEE Trans. Autom. Control*, vol. 57, no. 9, pp. 2385–2390, Sep. 2012.
[8] A. György and T. Linder, “Codecell convexity in optimal entropy-constrained vector quantization,” *IEEE Trans. Inform. Theory*, vol. 49, no. 7, pp. 1821–1828, Jul. 2003.
[9] J. B. Lewis and J. T. Tou, “Optimum sampled-data systems with quantized control signals,” *IEEE Trans. Appl. Ind.*, vol. 82, pp. 229–233, Jul. 1965.
[10] T. Linder and S. Yüksel, On Optimal Zero-Delay Quantization of Vector Markov Sources Preprint, available on arXiv, 2013.
[11] A. Mahajan and D. Teneketzis, “Optimal performance of networked control systems with non-classical information structures,” *SIAM J. Control and Optimiz.*, vol. 48, pp. 1377–1404, May 2009.
[12] A. S. Matveev and A. V. Savkin, “The problem of LQG optimal control via a limited capacity communication channel,” *Syst. and Control Lett.*, vol. 53, pp. 51–64, Feb. 2004.
[13] G. N. Nair, F. Fagnani, S. Zampieri, and J. R. Evans, “Feedback control under data constraints: An overview,” *Proc. IEEE*, vol. 95, no. 1, pp. 108–137, Jan. 2007.
[14] S. Tatikonda, A. Sahai, and S. Mitter, “Stochastic linear control over a communication channels,” *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1549–1561, Sep. 2004.
[15] D. Teneketzis, “On the structure of optimal real-time encoders and decoders in noisy communication,” *IEEE Trans. Inform. Theory*, vol. 52, no. 9, pp. 4017–4035, Sep. 2006.
[16] J. C. Walrand and P. Varaiya, “Causal coding and control of markov chains,” *Syst. & Control Lett.*, vol. 3, pp. 189–192, 1983.
[17] J. C. Walrand and P. Varaiya, “Optimal causal coding-decoding problems,” *IEEE Trans. Inform. Theory*, vol. 19, no. 11, pp. 814–820, Nov. 1983.
[18] H. S. Witsenhausen, “On the structure of real-time source coders,” *Bell Syst. Tech. J*, vol. 58, pp. 1437–1451, Jul./Aug. 1979.
[19] S. Yüksel, “On optimal causal coding of partially observed Markov sources in single and multi-terminal settings,” *IEEE Trans. Inform. Theory*, vol. 59, no. 1, pp. 424–437, Jan. 2013.
[20] S. Yüksel and T. Başar, *Stochastic Networked Control Systems: Stabilization and Optimization under Information Constraints*. Boston, MA: Birkhäuser, 2013.
[21] S. Yüksel and T. Linder, “Optimization and convergence of observation channels in stochastic control,” *SIAM J. Control and Optimiz.*, vol. 50, pp. 864–887, 2012.