SIAM J. CONTROL OPTIM.
Vol. 57, No. 1, pp. 773–782


# A NOTE ON THE SEPARATION OF OPTIMAL QUANTIZATION AND CONTROL POLICIES IN NETWORKED CONTROL*

SERDAR YÜKSEL†

**Abstract.** For controlled $\mathbb{R}^n$-valued linear systems driven by Gaussian noise under quadratic cost criteria, we revisit the problem of the structure of optimal quantization and control policies. In a recent paper [*IEEE Trans. Automat. Control*, 59 (2014), pp. 1612–1617] by the author, for fully observed and partially observed systems, the global optimality of predictive encoders was established under quadratic cost criteria. Furthermore, optimal control policies were shown to be linear in the conditional estimate of the state, and a form of separation of estimation and control was established. The present note does not introduce any new results or new conditions but clarifies that the results have been mischaracterized in the recent paper [M. Rabi, C. Ramesh, and K. H. Johansson, *SIAM J. Control Optim.*, 54 (2016), pp. 662–689]. Since perhaps the arguments in [*IEEE Trans. Automat. Control*, 59 (2014), pp. 1612–1617] were concise and this led to the confusion, its key result is presented here with a more detailed proof.

**Key words.** stochastic control, networked control, source coding

**AMS subject classifications.** 93E20, 93E03, 94A29

**DOI.** 10.1137/17M1127016

## 1. Introduction.

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:

$$x_{t+1} = Ax_t + Bu_t + w_t, \qquad y_t = Cx_t + v_t. \tag{1.1}$$

Here, $\{w_t, v_t\}$ is a mutually independent, zero-mean independent and identically distributed (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 Figure 1.1, let there be an encoder which has access to the observation variable $y_t$, and which transmits its information to a receiver/controller, over a discrete noiseless channel with finite capacity.

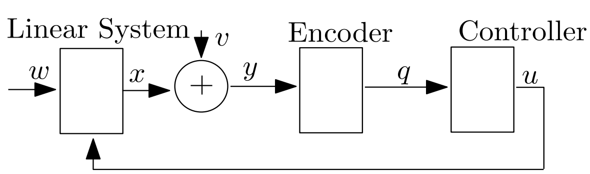


Fig. 1.1. *Joint LQG optimal design of coding and control.*

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 $\mathcal{M}$.*

*Received by the editors April 24, 2017; accepted for publication (in revised form) January 2, 2019; published electronically February 26, 2019.
http://www.siam.org/journals/sicon/57-1/M112701.html
**Funding:** This research was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).
†Department of Mathematics and Statistics, Queen's University, Kingston, Ontario, Canada, K7L 3N6 (yuksel@mast.queensu.ca).

When the spaces $\mathbb{A}$ and $\mathcal{M}$ are clear from the context, we will denote the quantizer simply by $Q$. Following [3], by *Composite Quantization (Coding) Policy* $\Pi^{comp}$, we refer to 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 $\mathcal{M} := \{1, 2, \ldots, M\}$, which is the quantization output alphabet 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 [3], 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 is used, subsequently, 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 [4] for similar reasoning).

Thus, with the information at the controller at time $t > 1$ being $\mathcal{I}_t^c = \{q_{[0,t]}\}$ and by writing $\mathcal{I}_t^e = \{y_{[0,t]}, q_{[0,t-1]}\}$, we can express the composite quantization policy as

$$Q_t^{comp}(\mathcal{I}_t^e) = (T_t(\mathcal{I}_{t-1}^c))(\mathcal{I}_t^e \setminus \mathcal{I}_{t-1}^c). \tag{1.2}$$

Any composite quantization policy $Q_t^{comp}$ can be expressed in the form above; i.e., there is no loss in the set of such policies, since for any $Q_t^{comp}$, one can define

$$T_t(\mathcal{I}_{t-1}^c)(\cdot) := Q_t^{comp}(\cdot, \mathcal{I}_{t-1}^c).$$

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 $\mathcal{I}_t^e \setminus \mathcal{I}_{t-1}^c$ is quantized). Under action $Q_t$, and given the local information, the encoder generates $q_t^i$ as the *quantization output* at time $t$. An admissible controller policy is a sequence of functions $\gamma = \{\gamma_t\}$ such that $u_t = \gamma_t(q_{[0,t]})$, with $\gamma_t : \mathcal{M}^{t+1} \to \mathbb{R}^m, t \geq 0$. We call such encoding and control policies *causal* or *admissible*. The goal is the computation of

$$\inf_{\Pi^{comp}} \inf_{\gamma} J(\Pi^{comp}, \gamma, T), \tag{1.3}$$

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, $Q \geq 0$ is a positive semi-definite matrix, $R > 0$ is a positive definite matrix, and $\nu_0$ is the initial Gaussian measure on $x_0$.

Recently, [2] proposed structural results on optimal encoders for the setup provided in the previous section. The authors in [2, section 4] provide a class of encoders and establish a separation result similar to the one presented in [1]. While motivating their optimality result, the authors of [2] state that the existing results in the field are nonsatisfactory and that the arguments in [1] may not hold. In particular, they note that they *illustrate the insufficiency of the arguments offered in* 12 *papers, including* [1]*, for the optimality of separation and certainty equivalent control.*

The goal of this note is to correct the criticism claimed in [2]: There is no new result in this note, nor is there any additional new assumption; we will emphasize that the structural and separation results in [1] hold true as they were. The point of this note is to present a record with regard to the results presented in [1], but also to show that one does not need to impose any new conditions for the optimality of predictive encoders: The results in [1] on separation and optimality are general with regard to the optimality of predictive encoders without any a priori restrictions on the encoders and the controllers.

We also use this opportunity to apply some minor corrections with regard to the Riccati equation recursions in [1].

There is a large literature on jointly optimal quantization for the LQG problem dating back to the early 1960s. Since evidently this problem has caused a large amount of confusion and given the sensitivity surrounding the abundance of results in this field (some of which are unfortunately inconsistent), and to present the findings of the contribution in a proper context, we ask the reader to revisit the cautiously written literature review in [1, pages 1612–1613].

In the following, we revisit the results in [1] and present an expanded proof for the main separation result; in particular, we expand the dynamic programming argument that was crucial in the proof of [1, Lemma 3.1].

## 2. Structural results on optimal codes for controlled Markov models.

Consider the fully observed system

$$x_{t+1} = f(x_t, u_t, w_t), \qquad y_t = x_t, \quad t = 0, 1, \ldots, \tag{2.1}$$

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

$$\inf_{\Pi^{comp}} \inf_{\gamma} E_{\nu_0}^{\Pi^{comp},\gamma}\left[\sum_{t=0}^{T-1} c(x_t, u_t)\right], \tag{2.2}$$

over all quantization and control 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 $t \geq 0$. Structural results on optimal quantization policies for such controlled Markov sources have been studied in [5] in the context of finite control and action spaces and in [6] for control over noisy channels and also for finite state-actions space setting. The following theorems extend the finite state space analysis of [5] to more general spaces. The proofs of the results below essentially follow from [3, Theorems 2.4, 2.5] with additional minor modifications due to the presence of control actions. The first one can be regarded as an extension of Witsenhausen's structural theorem [7], and the second one can be regarded as an extension of the results of Walrand and Varaiya [4] (see also [8]). We note also that [3] addressed certain measurability issues which arise in the uncountable (Polish) space setting (thus including problems with real spaces as well as partially observed models) for the derivation of structural results on optimal encoders. For proofs of the results below, we refer the reader to [9, Theorem 10.3.6] and its proof.

THEOREM 2.1 (see [1], [9, Theorem 10.3.6]). *For system* (2.1)*, under the information structure described in the previous section and the objective given in* (2.2)*, 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, and define $\pi_t \in \mathcal{P}(\mathbb{X})$ to be the regular conditional probability measure given by $\pi_t(\cdot) = P(x_t \in \cdot | q_{[0,t-1]}, u_{0,t-1})$.

THEOREM 2.2 (see [1], [9, Theorem 10.3.6]). *For system* (2.1)*, under the information structure described in the previous section and the objective given in* (2.2)*, any composite quantization policy can be replaced, without any loss in performance, by one which only uses the conditional probability measure $\pi_t$, 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 $t$.*

One can also consider the partially observed setting; see [1], [9, section 10.3].

We next revisit the following construction in [10] on the set of quantizers.

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 $\mathbb{R}^n$.*

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 [10], we allow for the possibility that some of the cells of the quantizer are empty.

As discussed in [10], a quantizer $Q$ with cells $\{B_1, \ldots, B_M\}$ can be characterized as a stochastic kernel $Q$ from $\mathbb{R}^n$ to $\{1, \ldots, M\}$ defined by

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

Reference [10] endows 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 $\mathbb{R}^n \times \mathcal{M}$. Consider the set of probability measures

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

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 [11]). 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, [10] defines

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

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

## 3. Fully observed LQG: Separation of estimation error and control.

Consider the LQG problem for the system given in (1.1) with the cost function given in (1.3), 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 $\Pi_W$. In the following, we adopt a dynamic programming approach and establish that the optimal controller is linear in its estimate. This fact applies naturally to the terminal time stage control. That this also applies to 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 $u_{T-1} = 0$ almost surely. To obtain a solution for $t = T-2$, we look for a solution to

$$\min_{\gamma_t} E\left[x_t'Qx_t + u_t'Ru_t + E[(Ax_t + Bu_t + w_t)'Q(Ax_t + Bu_t + w_t)|\mathcal{I}_t^c, u_t]\middle|\mathcal{I}_t^c\right].$$

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 $L_{T-2} = -(R + B'QB)^{-1}B'QA$. 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

$$\begin{aligned}\tilde{m}_{t+1} &= E[x_{t+1}|\mathcal{I}_{t+1}^c] = E\left[x_{t+1} - E[x_{t+1}|\mathcal{I}_t^c] + E[x_{t+1}|\mathcal{I}_t^c]\middle|\mathcal{I}_{t+1}^c\right] \\ &= E\left[E[Ax_t + Bu_t + w_t|\mathcal{I}_t^c] + \left(x_{t+1} - E[x_{t+1}|\mathcal{I}_t^c]\right)\middle|\mathcal{I}_{t+1}^c\right] \\ &= A\tilde{m}_t + Bu_t + \left(E[x_{t+1}|\mathcal{I}_{t+1}^c] - E[x_{t+1}|\mathcal{I}_t^c]\right) = A\tilde{m}_t + Bu_t + \bar{w}_t,\end{aligned} \tag{3.1}$$

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 $\mathcal{I}_t^c$ (in the sense that for any appropriate measurable bounded $g$, $E[\bar{w}_t g(\mathcal{I}_t^c)] = 0$).

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 by an uncorrelated noise (though not necessarily an independent) process $\bar{w}_t = E[x_{t+1}|\mathcal{I}_{t+1}^c] - E[x_{t+1}|\mathcal{I}_t^c]$. However, this lack of independence may be important, as elaborated on in [12]. 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 $\bar{w}_t'Q\bar{w}_t$ is not affected by the control policies $\{\gamma_k, k \leq t-1\}$ (that is, the changes in the control actions $\{u_k, k \leq t-1\}$ do not affect $\bar{w}_t'Q\bar{w}_t$) under an optimal coding policy for all time stages $t$. A sufficient condition for this is that the encoder is a predictive one (see [13], [12], and [14] for related discussions), as is derived in the following analysis.

Definition 3.1 (see [1, 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*

$$\bar{x}_{t+1} = A\bar{x}_t + w_t, \tag{3.2}$$

*and 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 Figure 3.1. One question which had not been addressed in [15], [13], [12], [14], or [16] is *whether restriction to this class of quantization policies (given in Definition* 3.1*) is without loss.* We have the following lemma, *which was the key result in* [1] *on the structure of optimal encoders, the optimality of predictive quantizers, and the associated separation result.*

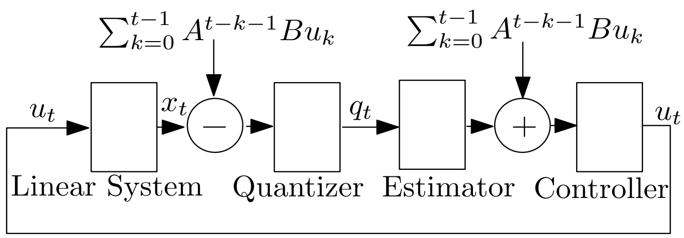


FIG. 3.1. *For the LQG problem, a predictive encoder is optimal.*

LEMMA 3.1 (see [1, Lemma 3.1]). *For problem* (1.3)*, 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 given an optimal control policy for problem* (1.3)*.*

*Proof.* We apply backwards induction and dynamic programming. For $t = T-1$, the optimal control is zero; therefore, the quantizer's design does not affect the expected cost. We therefore may use a predictive quantizer for $t = T-1$ without any loss. Now, for the time stage, $t = T-2$, let $f_t(q_{[0,t-1]}) := \sum_{k=0}^{t-1} A^{t-k-1}Bu_k$. If the policy considered is in $\Pi_W$, the quantization policy is of the form $Q_t(\bar{x}_t + f_t(q_{[0,t-1]}), P(\bar{x}_t + f_t(q_{[0,t-1]}) \in \cdot|q_{[0,t-1]}))$. For this time stage, the optimal decoder and controller uses a sufficient statistic to generate the optimal control policy, which is $E[x_t|q_{[0,t]}]$. Observe that

$$E[\bar{x}_t + f_t(q_{[0,t-1]})|q_{[0,t]}] = E[\bar{x}_t|q_{[0,t]}] + f_t(q_{[0,t-1]}) = E[\bar{x}_t|q_{[0,t-1]}, q_t] + f_t(q_{[0,t-1]}).$$

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 (3.2)) 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.

Here, we note that $\bar{w}_{T-2}$ does not (functionally) depend on the control actions in that if one changes the control policies $\{\gamma_s, 0 \le s \le T-3\}$, $\bar{w}_{T-2}$ is not affected. This does not imply that $\bar{w}_{T-2}$ is statistically independent from the past control actions; however, this is not relevant for the analysis, as we demonstrate in the following.

First, observe that through the law of iterated expectations, and the orthogonality principle, we have

$$\frac{1}{T}E_{\nu_0}^{\Pi^{comp},\gamma}\left[\sum_{t=0}^{T-1} x_t'Qx_t + u_t'Ru_t\right]$$
$$= \frac{1}{T}E_{\nu_0}^{\Pi^{comp},\gamma}\left[\sum_{t=0}^{T-1}(x_t - \tilde{m}_t)'Q(x_t - \tilde{m}_t) + \tilde{m}_t'Q\tilde{m}_t + u_t'Ru_t\right]. \tag{3.3}$$

Now, once we have that the quantizer at time $t = T-2$ is a predictive one, we can write the cost for $t = T-3$ as follows: Through (3.1), with $u_{T-2} = L_{T-2}\tilde{m}_{T-2} = -(R + B'QB)^{-1}B'QA\tilde{m}_{T-2}$, the cost to go for a policy $\gamma_{T-3}$ would write as

$$E\left[E\left[\tilde{m}_{N-2}'\left(Q + A'QA - A'QB(R+B'QB)^{-1}B'QA\right)\tilde{m}_{N-2}\middle|\mathcal{I}_{N-3}\right]\right. \tag{3.4}$$
$$+E\left[\bar{w}_{N-2}'Q\bar{w}_{N-2}\middle|\mathcal{I}_{N-3}\right] \tag{3.5}$$
$$+E\left[(x_{N-2} - \tilde{m}_{N-2})'Q(x_{N-2} - \tilde{m}_{N-2})\middle|\mathcal{I}_{N-3}\right] \tag{3.6}$$
$$+E\left[(x_{N-1} - \tilde{m}_{N-1})'Q(x_{N-1} - \tilde{m}_{N-1})\middle|\mathcal{I}_{N-3}\right] \tag{3.7}$$
$$\left.+2E\left[(x_{N-2} - \tilde{m}_{N-2})'A'QBu_{N-2}\middle|\mathcal{I}_{N-3}\right]\right]. \tag{3.8}$$

The last term (3.8) is zero since $(x_{T-2} - \tilde{m}_{T-2})$ is orthogonal to $u_{T-2}$. By the use of the predictive quantizer, the terms (3.6) and (3.7) do not depend on the control policy at time $T-3$. Note that

$$\begin{aligned}
&\bar{w}_t = E[x_{t+1}|\mathcal{I}_{t+1}] - E[x_{t+1}|\mathcal{I}_t] = (x_{t+1} - E[x_{t+1}|\mathcal{I}_{t+1}]) - (x_{t+1} - E[x_{t+1}|\mathcal{I}_t])\\
&= (x_{t+1} - E[x_{t+1}|\mathcal{I}_{t+1}]) - (Ax_t + Bu_t + w_t - E[Ax_t + Bu_t + w_t|\mathcal{I}_t])\\
&= (x_{t+1} - E[x_{t+1}|\mathcal{I}_{t+1}]) - (Ax_t + w_t - E[Ax_t + w_t|\mathcal{I}_t])\\
&= (x_{t+1} - E[x_{t+1}|\mathcal{I}_{t+1}]) - (A(x_t - E[Ax_t|\mathcal{I}_t]) + w_t),
\end{aligned}$$

and this term does not depend on the past applied control policies. Thus, (3.5) is not affected by the control policy at time $T-3$. Hence, these terms can be taken out from the optimization so that the cost to go that is relevant for optimization is

$$E\left[E\left[\tilde{m}_{T-2}'\left(Q + A'QA - A'QB(R+B'QB)^{-1}B'QA\right)\tilde{m}_{T-2}\middle|\mathcal{I}_{T-3}^c\right]\right].$$

To obtain a solution for $t = T-3$, we then look for a solution to

$$\begin{aligned}
&\min_{\gamma_{T-3}} E\left[(x_{T-3}'Qx_{T-3} + u_{T-3}'Ru_{T-3})\right.\\
&\left.+E\left[\tilde{m}_{T-2}'\left(Q + A'QA - A'QB(R+B'QB)^{-1}B'QA\right)\tilde{m}_{T-2}\middle|\mathcal{I}_{T-3}^c, u_{T-3}\right]\middle|\mathcal{I}_{T-3}^c\right].
\end{aligned}$$

As in (3.3), writing $E[x_{T-3}'Qx_{T-3}]$ as

$$E\left[(x_{T-3} - \tilde{m}_{T-3} + \tilde{m}_{T-3})'Q(x_{T-3} - \tilde{m}_{T-3} + \tilde{m}_{T-3})\right],$$

and noting the orthogonality of $x_{T-3}-\tilde{m}_{T-3}$ and $\tilde{m}_{T-3}$, the estimation cost $E[(x_{T-3}-\tilde{m}_{T-3})'Q(x_{T-3}-\tilde{m}_{T-3})]$ can be left out from the optimization, and the cost relevant for the control policy at time $T-3$ is

$$\min_{\gamma_{T-3}} E\Bigg[E\Bigg[(\tilde{m}'_{T-3}Q\tilde{m}_{T-3}+u'_{T-3}Ru_{T-3}) \\ +E\Bigg[\tilde{m}'_{T-2}\Bigg(Q+A'QA-A'QB(R+B'QB)^{-1}B'QA\Bigg)\tilde{m}_{T-2}\Bigg|\mathcal{I}^c_{T-3},u_{T-3}\Bigg]\Bigg|\mathcal{I}^c_{T-3}\Bigg]\Bigg].$$

Now, using (3.1), by completing the squares, and using the orthogonality principle with $u_{T-3}$ being orthogonal to $\bar{w}_{T-3}$, we obtain that the optimal control is linear and is given by $u_{T-3}=L_{T-3}\tilde{m}_{T-3}$ with $L_{T-3}=-(R+B'K_{T-2}B)^{-1}B'K_{T-2}A$, where $K_{T-2}$ satisfies the recursion $K_t=A'_tK_{t+1}A_t-P_t+Q$, with $P_t=A'_tK_{t+1}B(R+B'K_{t+1}B)^{-1}B'K_{t+1}A$ and $K_T=P_{T-1}=0$.

An optimal controller at time $t=T-3$ will then use $\tilde{m}_{T-3}$ 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 reasoning similar to that 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 policies $\{\gamma_s, s<T-3\}$ (and thus does not functionally depend on the control actions) applied earlier. This inductively leads to the optimality of linear policies and the optimality of predictive quantizers for all $t\geq 0$. ◻

We have also thus established above that the optimal control is linear for all time stages, by the proof of Lemma 3.1.

*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 policies. Hence, *there is no dual effect of the control policies in the sense that the estimation error at any given time does not depend on past-applied control policies (or is not affected by past-applied actions).*

*Remark* 3.2. For the proof presented, it was essential to show first that the coding policies adopted can be taken to be in class $\Pi_W$. Indeed, in the absence of such a restriction (which we showed to be without any loss), a counterexample presented in [2, Example 3] utilizing a coding policy which does not belong to $\Pi_W$ reveals that the aforementioned separation result does not hold.

We have the following (see also [12], which establishes a more restrictive structure than that given in Definition 3.1 for a similar result).

THEOREM 3.1. *For the minimization problem* (1.3)*, with the new effective state dynamics in* (3.1)*, an optimal control policy is given by* $u_t=L_tE[x_t|q_{[0,t]}]$*, where* $L_t=-(R+B'K_{t+1}B)^{-1}B'K_{t+1}A$*, where* $K_t$ *satisfies the recursions* $K_t=A'_tK_{t+1}A_t-P_t+Q$ *with* $P_t=A'_tK_{t+1}B(R+B'K_{t+1}B)^{-1}B'K_{t+1}A$ *and* $K_T=P_{T-1}=0$.

With the cost written as (3.3) and with the preceding analysis, we obtain for $t\geq 0$ the unnormalized value function for any time stage $t$ as

$$J_t(\mathcal{I}^c_t)=E[\tilde{m}'_tK_t\tilde{m}_t|\mathcal{I}^c_t]+\sum_{k=t}^{T-1}\Bigg(E[(x_k-E[x_k|\mathcal{I}^c_k])'Q(x_k-E[x_k|\mathcal{I}^c_k])]+E[\bar{w}'_kK_{k+1}\bar{w}_k]\Bigg),$$

with $J(\Pi^{comp},\gamma,T)=\frac{1}{T}J_0(\mathcal{I}^c_0)$. 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:

$$E\bigg[||E[x_{t+1}|\mathcal{I}^c_{t+1}] - E[x_{t+1}|\mathcal{I}^c_t]||^2_\Lambda\bigg] = E\bigg[||(E[x_{t+1}|\mathcal{I}^c_{t+1}] - x_{t+1}) + (x_{t+1} - E[x_{t+1}|\mathcal{I}^c_t])||^2_\Lambda\bigg]$$
$$= E\bigg[||(E[x_{t+1}|\mathcal{I}^c_{t+1}] - x_{t+1})||^2_\Lambda] + E[||(x_{t+1} - E[x_{t+1}|\mathcal{I}^c_t])||^2_\Lambda\bigg]$$
$$+2E[\langle(E[x_{t+1}|\mathcal{I}^c_{t+1}] - x_{t+1}), (x_{t+1} - E[x_{t+1}|\mathcal{I}^c_t])\rangle_\Lambda].$$

Note that

$$E\bigg[\langle(E[x_{t+1}|\mathcal{I}^c_{t+1}] - x_{t+1}), (x_{t+1} - E[x_{t+1}|\mathcal{I}^c_t])\rangle_\Lambda\bigg]$$
$$= E\bigg[-\langle(E[x_{t+1}|\mathcal{I}^c_{t+1}] - x_{t+1}), (E[x_{t+1}|\mathcal{I}^c_t])\rangle_\Lambda + \langle(E[x_{t+1}|\mathcal{I}^c_{t+1}] - x_{t+1}), (x_{t+1})\rangle_\Lambda\bigg]$$
$$= E[\langle(E[x_{t+1}|\mathcal{I}^c_{t+1}] - x_{t+1}), (x_{t+1})\rangle_\Lambda] = -E[||(E[x_{t+1}|\mathcal{I}^c_{t+1}] - x_{t+1})||^2_\Lambda], \tag{3.9}$$

where (3.9) follows from the orthogonality property of minimum mean-square estimation and that $E[x_{t+1}|\mathcal{I}^c_t]$ is measurable on $\sigma(\mathcal{I}^c_{t+1})$, the sigma-field generated by $\mathcal{I}^c_{t+1}$. Therefore, we have

$$E\bigg[||(E[x_{t+1}|\mathcal{I}^c_{t+1}] - E[x_{t+1}|\mathcal{I}^c_t])||^2_{K_{t+1}}\bigg]$$
$$= -E\bigg[(x_{t+1} - E[x_{t+1}|\mathcal{I}^c_{t+1}])'(K_{t+1})(x_{t+1} - E[x_{t+1}|\mathcal{I}^c_{t+1}])$$
$$+E[(x_t - E[x_t|\mathcal{I}^c_t])'(A'K_{t+1}A)(x_t - E[x_t|\mathcal{I}^c_t])] + E[w'K_{t+1}w]\bigg].$$

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

$$J_t(\mathcal{I}^c_t) = E[\tilde{m}_t'K_t\tilde{m}_t|\mathcal{I}^c_t] + E[(x_t - \tilde{m}_t)'(Q + A'K_{t+1}A)(x_t - \tilde{m}_t)]$$
$$+\sum_{k=t+1}^{T-1} E[(x_k - \tilde{m}_k)'(Q + A'K_{k+1}A - K_k)(x_k - \tilde{m}_k)] + \sum_{k=t}^{T-1} E[w_k'K_{k+1}w_k]. \tag{3.10}$$

In particular, *it follows that the quantization problem can be separated from the control problem*. Once this separation result is established, [1] then studies the existence problem for the optimal quantization policies (building on [17]) and the extensions to the partially observed setups.

**4. Conclusion.** In this brief note, a clarification on some technical concerns presented in [2] questioning the separation results in [1] is presented. We have expanded the original proof of [1, Lemma 3.1] with no new assumptions. Thus, the joint optimization problem of encoding and control policies for networked Linear Quadratic Gaussian systems with a discrete noiseless channel is studied, the global optimality of predictive encoders established in [1] is revisited, and it is shown that a form of separation of estimation and control applies. These results further refine the existing structural and separation results in [18, 19, 20, 15, 21, 14, 12, 13, 16].

## REFERENCES


[1] S. Yüksel, *Jointly optimal LQG quantization and control policies for multi-dimensional linear Gaussian sources*, IEEE Trans. Automat. Control, 59 (2014), pp. 1612–1617.

[2] M. Rabi, C. Ramesh, and K. H. Johansson, *Separated design of encoder and controller for networked linear quadratic optimal control*, SIAM J. Control Optim., 54 (2016), pp. 662–689, https://doi.org/10.1137/14M0970987.

[3] S. Yüksel, *On optimal causal coding of partially observed Markov sources in single and multi-terminal settings*, IEEE Trans. Inform. Theory, 59 (2013), pp. 424–437.

[4] J. C. Walrand and P. Varaiya, *Optimal causal coding-decoding problems*, IEEE Trans. Inform. Theory, 19 (1983), pp. 814–820.

[5] J. C. Walrand and P. Varaiya, *Causal coding and control of Markov chains*, Systems Control Lett., 3 (1983), pp. 189–192.

[6] A. Mahajan and D. Teneketzis, *Optimal performance of networked control systems with nonclassical information structures*, SIAM J. Control Optim., 48 (2009), pp. 1377–1404, https://doi.org/10.1137/060678130.

[7] H. S. Witsenhausen, *On the structure of real-time source coders*, Bell Syst. Tech. J., 58 (1979), pp. 1437–1451.

[8] D. Teneketzis, *On the structure of optimal real-time encoders and decoders in noisy communication*, IEEE Trans. Inform. Theory, 52 (2006), pp. 4017–4035.

[9] S. Yüksel and T. Başar, *Stochastic Networked Control Systems: Stabilization and Optimization under Information Constraints*, Birkhäuser/Springer, New York, 2013.

[10] S. Yüksel and T. Linder, *Optimization and convergence of observation channels in stochastic control*, SIAM J. Control Optim., 50 (2012), pp. 864–887, https://doi.org/10.1137/100808976.

[11] V. S. Borkar, *White-noise representations in stochastic realization theory*, SIAM J. Control Optim., 31 (1993), pp. 1093–1102, https://doi.org/10.1137/0331050.

[12] G. N. Nair, F. Fagnani, S. Zampieri, and J. R. Evans, *Feedback control under data constraints: An overview*, Proc. IEEE, 95 (2007), pp. 108–137.

[13] V. S. Borkar and S. K. Mitter, *LQG control with communication constraints*, in Communications, Computation, Control, and Signal Processing: A Tribute to Thomas Kailath, Kluwer Academic Publishers, Boston, MA, 1997, pp. 365–373.

[14] S. Tatikonda, A. Sahai, and S. Mitter, *Stochastic linear control over a communication channels*, IEEE Trans. Automat. Control, 49 (2004), pp. 1549–1561.

[15] L. Bao, M. Skoglund, and K. H. Johansson, *Iterative encoder-controller design for feedback control over noisy channels*, IEEE Trans. Automat. Control, 57 (2011), pp. 265–278.

[16] M. Fu, *Lack of separation principle for quantized linear quadratic Gaussian control*, IEEE Trans. Automat. Control, 57 (2012), pp. 2385–2390.

[17] T. Linder and S. Yüksel, *On optimal zero-delay quantization of vector Markov sources*, IEEE Trans. Inform. Theory, 60 (2014), pp. 2975–5991.

[18] J. B. Lewis and J. T. Tou, *Optimum sampled-data systems with quantized control signals*, IEEE Trans. Appl. Industry, 82 (1965), pp. 229–233.

[19] R. E. Curry, *Estimation and Control with Quantized Measurements*, MIT Press, Cambridge, MA, 1969.

[20] T. R. Fischer, *Optimal quantized control*, IEEE Trans. Automat. Control, 27 (1982), pp. 996–998.

[21] A. S. Matveev and A. V. Savkin, *The problem of LQG optimal control via a limited capacity communication channel*, Systems Control Lett., 53 (2004), pp. 51–64.