Minimum Rate Coding for LTI Systems Over Noiseless Channels

Serdar Yüksel, *Student Member, IEEE*, and Tamer Başar, *Fellow, IEEE*

***Abstract*—This paper studies rate requirements for state estimation in linear time-invariant (LTI) systems where the controller and the plant are connected via a noiseless channel with limited capacity. Using information theoretic arguments, we obtain first for scalar systems, and subsequently for multidimensional systems, lower bounds on the data rates required for state estimation under three different stability criteria, namely monotonic boundedness of entropy, asymptotic stability of distortion, and support size stability. Further, the minimum data rate achievable by any source-encoder is computed under each of these criteria, and the best rate achievable with quantization is shown to be in agreement with the information-theoretic bounds in some specific cases (such as if the system coefficient is an integer or if the criterion is an asymptotic one). Existence of optimal variable-length and fixed-length quantizers are studied and optimal quantizers are constructed under each of these criteria. One observation is that, the uniform quantizer is, in addition to being simple, efficient in linear control systems.**

***Index Terms*—Networked control, quantization, stability.**

I. Introduction and Literature Review

We ARE seeing increasingly more the interplay between control theoretic issues in communication problems and communication theoretic issues in control problems. One class of problems in this context is control over communication channels, for which various system models and channel structures have been studied in the recent literature; see, e.g., [1]–[5], [7], and the references therein. Among these, several studies focused on noiseless systems with time-invariant encoders, where the main issue becomes one of designing a quantizer; see [7] and [8]. We note, however, that when the channels are noiseless, then structures which are not time invariant turn out to be more rate-efficient [13]–[16], because the transmitter and the receiver can make the updates in the encoding and the decoding rely on the data received, equipped with the knowledge that the data sent will make it to the other party with no ambiguity. It is this class of problems that this paper addresses.

Papers most relevant to our work here (which is an expanded version of [9] and part of [10]), are [13], [17], [18], and [14]. Among these, [18] also studies noiseless systems with noiseless channels. In [13], *directed mutual information* is utilized for systems with noiseless feedback to obtain the optimum causal encoders minimizing the rate subject to distortion criteria. In [14], using fixed-length quantization, it has been shown that for an LTI system with an uncertain initial state, where the cost is the rth moment of the state, a necessary and sufficient condition for exponential stability is that the input rate should be higher than the sum of the logarithms of the ratios of the absolute values of unstable open-loop eigenvalues and the desired exponential factor, which is naturally an asymptotic result. In contradistinction with the approaches of [13], [17] and [18], we focus here on optimal quantization problems, and show the connection between the optimal performance of causal quantizers with the lower bounds provided by information theory under three different criteria. This paper also differs from [14] in that we use information theoretic arguments to obtain bounds for any source-coder, and we study both fixed-length and variable-length quantization. We should mention that another previous study relevant to this paper is [4], where it has also been shown that capacity in the Shannon sense is not a sufficient measure for stabilizability of control systems.

We now provide a brief outline of the results of this paper. We first address scalar systems, and introduce optimal coding schemes under a zero-delay, sequential criterion, and obtain the rates achievable by any stabilizing source-coder. We then compare these rates with those corresponding to the best quantizer. We show that the information-theoretic bounds to achieve a monotonically bounded entropy sequence are operationally tight, i.e., they are achievable by a quantizer, if the system is scalar with an integer coefficient, or if the distortion criterion is an asymptotic one. If the system coefficient is not an integer (again for scalar systems), by exploiting variable-length coding we obtain the optimal quantizer meeting the stability criterion. We further show that, again for a scalar system, the probability distribution of the quantization error converges to a uniform one. We then analyze the conditions on rate required for boundedness and minimization of an asymptotic distortion, where both variable-length and fixed-length quantizers are considered. We finally consider stability in worst case estimation error for both scalar and multidimensional systems.

One of our goals in this paper is to bridge information- and quantization-theoretic results with control systems. We will see that many of the results from causal coding theory apply to control systems; however, some communication theoretic notions such as entropy and rate-distortion function have to be cautiously applied when used in a control context. Furthermore, it turns out that, the simplest and most common type of quantizer,

Manuscript received April 12, 2004; revised September 19, 2005 and April 18, 2006. Recommended by Associate Editor E. Jonckheere. This work was supported in part by the National Science Foundation under Grant CCR 00-85917 ITR. An earlier version of this paper was presented at the IEEE Conference on Control Applications (CCA), Istanbul, June 2003, and appeared in an abridged form in the CCA Proceedings.

S. Yüksel was with the Coordinated Science Laboratory, University of Illinois, Urbana, IL 61801-2307 USA. He is now with Dunham Lab, Yale University, New Haven, CT 06511 USA (e-mail: serdar.yuksel@yale.edu)

T. Başar is with the Coordinated Science Laboratory, University of Illinois, Urbana, IL 61801-2307 USA (e-mail: tbasar@control.csl.uiuc.edu).

Digital Object Identifier 10.1109/TAC.2006.886491

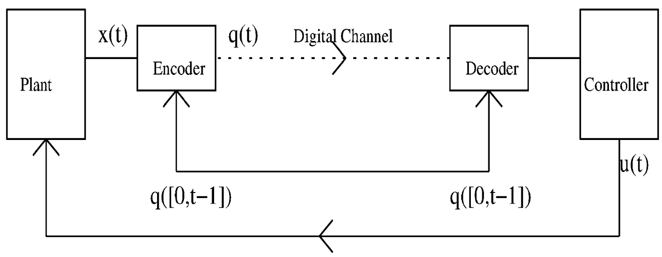


Fig. 1. Past codewords are common side information. With a slight abuse of notation here, $x(t)$ and $q(t)$ (instead of x_t and q_t) denote the state and the transmitted codewords respectively. If variable-length coding is used, the encoder consists of a quantizer followed by an entropy coder; if fixed-length coding is used the encoder is only the quantizer. The channel connecting the control to the plant is assumed to be noiseless.

the uniform quantizer, is indeed the best quantizer under some of the criteria.

This paper is organized as follows. Section II provides a precise formulation of the class of systems considered. Section III deals with quantization schemes for scalar systems, and Section IV for the multidimensional case. This paper ends with the concluding remarks of Section V.

II. Problem Formulation

The class of systems considered in this paper is described by

$$x_{t+1} = Ax_t + Bu_t, \qquad t = 0, 1, \ldots \tag{1}$$

where (A, B) is a controllable pair, with A having at least one unstable eigenvalue and no zero eigenvalues, t is the time variable, and x_t is the n-dimensional state; the initial state x_0 is a random vector with a bounded support, and u_t is the control variable of dimension m, which is allowed to depend on only a quantized version of the state x, possibly with memory. The state is quantized (source-coded) before being made available to the control u, at which site the quantized state information is recovered by a decoder; see Fig. 1.

The main goal in the design of the quantizer, the encoders, and the decoders is to achieve a certain degree of stability in state estimation error. Since (A, B) is a controllable pair, the problem of stability in state itself can be translated into an equivalent problem of stability in state estimation, as shown in [14]. Hence, without any loss of generality, we can deal with the control-free system $x_{t+1} = Ax_t$, and focus on state estimation. We now introduce some terminology and notation that will be used in the development, first for scalar (that is, one-dimensional) systems.

A *quantizer*, Q, for a scalar continuous variable is a mapping from the real line to a finite or countable set, characterized by corresponding bins $\{\mathcal{B}_i\}$ and their reconstruction levels q^i, such that $\forall i$, $Q(x) = q^i$ if and only if $x \in \mathcal{B}_i$ and $q^i \in \mathcal{B}_i$. Here (for scalar quantization), $\mathcal{B}_i$ can be taken to be non-overlapping semi-open intervals, $\mathcal{B}_i = (\delta_i, \delta_{i+1}]$, with $\delta_i < \delta_{i+1}, i = 0, \pm 1, \pm 2, \ldots$, such that δ_0 is the one closest to the origin, where $\{\delta_i\}$ are termed "bin edges." (For a comprehensive overview of quantization, see [21].)

Among different types of quantizers, the *uniform quantizer* (where $\mathcal{B}_i$ are of equal lengths) is the most common one, because of the ease with which it can be used and of its asymptotic optimality property under the mean-square distortion measure [22]. Another type is the *logarithmic quantizer* (where the lengths of $\mathcal{B}_i$'s increase exponentially away from the origin), which has been shown to be the coarsest quantizer for stabilization of systems with quantized signals [7]. Another classification of quantizers is according to whether their rate is a constant or a variable: A *fixed-length quantizer* is one where the quantized outcomes are represented by codewords of equal length, whereas a *variable-length quantizer* can have variable lengths assigned to the quantization outcomes [24], [21]. In a dynamic, discrete-time setting, the construction of a quantizer at any time t could depend on the past quantizer values. To make this precise, let $\mathcal{X}$ be the input space, $\hat{\mathcal{X}}$ be the output space, and $\mathcal{Q}_t$ be the set of quantizer reconstruction values, q_t, at time t, $t = 0, 1, \ldots$. Then, the quantizer at time t, f_t, is a mapping from $\mathcal{X}^{t+1}$ to $\mathcal{Q}_t$, where $\mathcal{X}^{t+1}$ is the $(t+1)$-product of $\mathcal{X}$. Such a quantizer is said to be *causal*, in addition to being *dynamic*. We also introduce $g_t : \mathcal{Q}_0 \times \mathcal{Q}_1 \times \ldots \mathcal{Q}_t \rightarrow \hat{\mathcal{X}}$ as the *decoder function*, which again has causal access to the past received values.

We next introduce two standard information-theoretic notions, namely mutual information and rate distortion function. *Mutual information* between an input random variable, X, and a corresponding output, Y, is

$$I(X;Y) = H(X) - H(X|Y)$$

where $H(X)$ is the entropy of X (differential entropy if X is a continuous random variable), and $H(X|Y)$ is the conditional entropy of X given Y. The *rate distortion function*, $R_X(D)$, of a random variable, X, is the minimum (or infimum) value of the mutual information over the class of stochastic mappings from the input to the output, subject to the constraint that distortion is no higher than a given level, D, namely

$$R(D) = \inf_{p(\hat{x}|x): E[(X-\hat{X})^2] \leq D} I(X; \hat{X})$$

Finally, we introduce the notion of the *support width*, $W(X)$, of a random variable X, as the width of the connected domain over which all mass associated with the distribution of X is concentrated.

In the case of a digital noiseless channel (which is our main concern here), the quantization output values will be available error-free at the receiver, and the receiver would use the sequence of quantization outputs for estimation at the receiver. We will use the notation $\hat{x}_t$ to denote the receiver estimation output at time t, so that $\hat{x}_t = g_t(q_0^t)$, for some decoder function g_t, where $q_0^t := \{q_0, \ldots, q_t\}$ and q_t denotes output of the encoder, f_t, at time t.

Our objective here, first for scalar systems, is to design dynamic causal quantizers $Q := \{q_t = Q_t(\cdot), t = 0, 1, 2, \ldots\}$ that are optimal under one of the three criteria introduced later.

Criterion 1: Monotonic Boundedness of Differential Entropy: Find a dynamic, zero-delay encoder (quantizer) $Q = \{Q_t\}$, under which the estimation error entropy is a nonincreasing sequence (and is thus bounded), i.e., with $q_0^t := \{q_0, \ldots, q_t\}$, $\forall t = 0, 1, \ldots$

$$H(x_{t+1} - \hat{x}_{t+1}(q_0^{t+1})|q_0^{t+1}) \leq H(x_t - \hat{x}_t(q_0^t)|q_0^t)$$

and it achieves this with minimum possible rate.$\diamond$

The binary representation of the quantization outputs could be of fixed length or variable length. If fixed-length coding is used, the rate of the code is given by

$$R_f = \log_2 |K|$$

where K is the number of levels in the quantizer or, equivalently, the number of $\mathcal{B}_i$'s. If variable-length coding is used, the rate is given by

$$R_v = H(Q(x)) = -\sum_{i=1}^{K} p_i \log_2(p_i)$$

where p_i is the probability of the event $x \in \mathcal{B}_i$.

We will first obtain bounds on the performance of any type of a quantizer, and then investigate the existence of quantizers that achieve these bounds. For the cases when the bounds are not met, we will show the existence of, determine the performance of, and construct the minimum-rate quantizers that lead to monotonically bounded entropy.

Criterion 2: Asymptotic Boundedness in Distortion: Find a minimal-rate quantizer sequence under which the asymptotic terminal-time estimation error distortion, that is the quantity

$$\limsup_{t\to\infty} E\left[|x_t - \hat{x}_t(q_0^t)|^2\right]$$

does not exceed some value $D < \infty$.$\diamond$

Remark 1: A criterion similar to the previous one was considered earlier in [14], which uses fixed-length coding. The difference between the treatments there and in this paper is that we will employ information theoretic arguments to obtain the ultimate bound achievable by any source-coder, and use asymptotic quantization theory to design the optimal quantizer; furthermore we will consider here both fixed-length and variable-length encoders.$\diamond$

Criterion 3: Stability in Support Size: Find a dynamic causal quantizer Q under which $\limsup_{t\to\infty} \quad W(x_t - \hat{x}_t(q_0^t)) = 0$. $\diamond$

This third criterion can also be regarded as *stability in the almost sure sense*.

We will also consider natural extensions of these three criteria to the multidimensional case later in Section IV.

III. Scalar LTI Systems

Consider now the scalar LTI system

$$x_{t+1} = ax_t + bu_t, \qquad t = 0, 1, \ldots \tag{2}$$

where the initial state x_0 is a random variable with a given distribution with finite support, and $b \neq 0$. We assume $|a| > 1$, for otherwise there would be no need for any data exchange between the plant and the controller. At each stage t, only a particular quantized value of the state is available to the controller, possibly with memory, and the problem of interest is to design an "optimal" quantization scheme under some prespecified constraints and with respect to a particular criterion (one of the three introduced in the previous section). We will assume that the control signal is available at both the receiver and the transmitter, which is legitimate because the channel is noiseless and the control signal is transmitted error-free. As stated earlier, with this assumption, the control function will not have any effect on the evolution of the uncertainty in the initial state. Hence, in essence, we can work with the control-free system, $x_{t+1} = ax_t, \quad t = 0, 1, \ldots$, and focus on the state estimation problem. In this reformulation, the state x_t can be viewed as the "uncertainty" at time t.

A. Criterion 1: Monotonic Boundedness of Differential Entropy

The sole element of randomness is in the initial state x_0. We first show that the quantization error converges to a uniform random variable in distribution, which will be useful in the exploration of the minimum required rate.

Lemma 3.1: Let x_0 be the realization of a random variable X_0 with a continuous probability density function $(\mathrm{pdf}) f_0(\cdot)$ on a finite support set $[-\Delta_0/2, \Delta_0/2]$. Suppose that at each successive time steps, the state generated by $x_{t+1} = ax_t$ is quantized using a K level uniform quantizer with a bin size equal to $\Delta_0|a|^t/K^{t+1}$ at time t, where $K \geq |a|$. Then, the Kullback–Leibler divergence between the conditional quantization error density (the quantization error for any specific bin) and the uniform error density with support size $\Delta_t = \Delta_0|a|^t/K^{t+1}$ converges to zero as $t \to \infty$.

Proof: See the Appendix. $\diamond$

We now state a result on achievable rates under *Criterion 1*.

Theorem 3.1: For the scalar LTI system (2) with a uniform, zero-mean distribution for the initial state, for any encoder to satisfy *Criterion 1*, the minimum rate is $\max(0, \log_2(|a|))$ per stage.

First, we state and prove two results (Lemmas 3.2 and 3.3) which will be needed in the proof of Theorem 3.1. Toward this end, let f_t denote the encoder function and g_t be the entropy-minimizing decoder function given f_t, at time t, $0 \leq t \leq T$, and consider the minimization, over f_t, of the differential entropy

$$H\left(x_t - g_t\left(f_t\left(x_0^t, q_0^{t-1}\right), q_0^{t-1}\right)|q_0^{t-1}, f_t\left(x_0^t, q_0^{t-1}\right)\right) \tag{3}$$

subject to the rate constraint

$$H\left(f_t\left(x_0^t, q_0^{t-1}\right)|q_0^{t-1}\right) \leq R. \tag{4}$$

Lemma 3.2: Suppose x_0 has a continuous probability density function and the conditional mean-square-error for x_t is uniformly bounded (in t). Then, there exists a solution to the constrained optimization problem of minimizing (3) subject to (4).

Proof: See the Appendix. ◇

The next result is a statement on the structure of the optimal quantizer.

Lemma 3.3: The optimal quantizer f_t, solving (3) and (4) uses only the current state and the past transmitted symbols.

Proof: See the Appendix. ◇

We are now in a position to prove Theorem 3.1.

Proof of Theorem 3.1: Past information is available at both the encoder and the decoder. We have

$$I\left(x_t; q_t | q_0^{t-1}\right) = H\left(x_t | q_0^{t-1}\right) - H\left(x_t | q_t, q_0^{t-1}\right)$$

where q_0^{t-1} can be viewed as the side information. Since $x_t = ax_{t-1}$, and $H(aX) = \log_2(|a|) + H(X)$ for any random variable X and any scalar a, we have

$$H\left(x_t | q_0^{t-1}\right) = \log_2(|a|) + H\left(x_{t-1} | q_0^{t-1}\right).$$

Thus

$$I\left(x_t; q_t | q_0^{t-1}\right) = \log_2(|a|) + H\left(x_{t-1} | q_0^{t-1}\right) - H\left(x_t | q_t, q_0^{t-1}\right).$$

Recognizing that $H(x_s - \hat{x}_s | q_0^s) = H(x_s | q_0^s)$ for both $s = t$ and $s = t-1$, we have

$$I(x_t; q_t | q_0^{t-1}) = \log_2(|a|) + H(x_{t-1} - \hat{x}_{t-1} | q_0^{t-1}) - H(x_t - \hat{x}_t | q_0^t).$$

Thus, to obtain a nonincreasing sequence in differential entropy, we need to have

$$I(x_t; q_t | q_0^{t-1}) \geq \log_2(|a|)$$

and therefore the mutual information is lower bounded by $\log_2(|a|)$. The corresponding expression that also covers the stable case is $\max(0, \log_2(|a|))$.◇

After thus presenting a bound on the rate for a general (including probabilistic) encoder, as before, we next obtain an achievable rate for a fixed-length quantizer. In this analysis, we will study the conditions for a monotonically nonincreasing sequence in distortion, $E[(x_t - \hat{x}_t)^2 | \hat{x}_t)]$, instead of in the differential entropy, because for a uniform random variable entropy and distortion are monotonic invertible functions of each other.

Theorem 3.2: For the scalar LTI system (2) with a uniform, zero-mean distribution for the initial state, for any fixed-length quantizer to satisfy Criterion 1, the required rate is $\max(0, \log_2(\lceil |a| \rceil))$ per stage, where $\lceil a \rceil$ is the smallest integer larger than or equal to a. This rate is achievable.

Proof: The estimation error is to have a nonincreasing support size, for which we need $K > |a|$, where K is an integer.◇

Thus, we observe that using fixed-length and deterministic encoding gets us close to the lower bound. Nonetheless, we may get even closer by using variable-length encoding, as we show next.

Variable-Length Encoding for Monotonic Boundedness in Distortion: Henceforth we consider the case where the encoder uses *variable-length encoding*, which is important when the parameter a is not an integer. The average code length needed to reliably express a random discrete output in a binary format is lower-bounded by the entropy rate of the process [24], and there exist coding schemes which come arbitrarily close to this bound. Our objective is to have a nonincreasing sequence

$$D_t = E[(x_t - \hat{x}_t)^2] \leq D_{t-1} = E[(x_{t-1} - \hat{x}_{t-1})^2].$$

Thus, at time t, the problem is the minimization of the entropy

$$J = -\sum_i \left[\int_{\delta_i}^{\delta_{i+1}} f(x)\, dx\right] \log g(\int_{\delta_i}^{\delta_{i+1}} f(x)\, dxg)$$

subject to the distortion and density constraints

$$\sum_i \int_{\delta_i}^{\delta_{i+1}} ((x - Q(x))^2) f(x) dx \leq D_{t-1}$$
$$\sum_i \int_{\delta_i}^{\delta_{i+1}} f(x) dx = 1. \tag{5}$$

The entropy is only a function of the quantizer threshold values, and not of the reconstruction values and, hence, given the threshold values, the partial derivative of J with respect to the quantization reconstruction values will characterize the optimality of the centroid of the cell as the reconstruction value. Now, if we restrict ourselves to a uniform input, with support set Δ, and using the centroid property for the quantization reconstruction values, J and the constraints reduce to

$$J = -\sum_i \frac{(\delta_{i+1} - \delta_i)}{\Delta} \log\left(\frac{\delta_{i+1} - \delta_i}{\Delta}\right)$$
$$\sum_i \frac{(\delta_{i+1} - \delta_i)^3}{12\Delta} \leq D_{t-1}, \frac{\delta_K - \delta_0}{\Delta} = 1. \tag{6}$$

By setting $\delta_{i+1} - \delta_i = p_i \Delta$, the cost to be minimized becomes $J = -\sum p_i \log_2(p_i)$, with the constraints

$$\sum p_i^3 \leq (1/a^2) \quad \sum p_i = 1 \quad p_i \geq 0, \forall i. \tag{7}$$

Proposition 3.1: For the linear system $x_{t+1} = ax_t$, the minimum rate required of a quantizer leading to a time-invariant distortion level is no smaller than $\max(0, \log_2(|a|))$. This rate is achievable for the class of systems where a is an integer, in which case an $|a|-$level uniform quantizer is used.

Proof: By a slight abuse of notation, let us write the entropy $H(p) = -\sum p_i \log_2(p_i)$, as $-E_p[\log_2(p)]$, where E_p denotes expectation with respect to the discrete probability mass function $(p_1, p_2, \ldots)$. Entropy is a concave function (of p) and, hence, using Jensen's inequality, $-E_p[\log_2(p)] \geq -\log_2(E_p[p]) = -\log_2(\sum_i p_i^2)$, where equality is achieved by the uniform distribution. Using Jensen's inequality again, this time for $E_p(p)$, we have $-\left(\sum p_i^2\right)^2 \geq -\sum p_i^3$ and

$$-\log_2\left(\left(\sum p_i^2\right)^2\right) \geq -\log_2\left(\sum p_i^3\right)$$

where again equality is achieved by the uniform distribution. Using the previous two results

$$\begin{aligned} H(p) = -E_p[\log_2(p)] &\geq \frac{1}{2}\log_2\left(\sum p_i^3\right) \\ &= -\frac{1}{2}\log_2(\frac{1}{a^2}) = \log_2(|a|). \end{aligned} \tag{8}$$

$\diamond$

A dual version of this problem is one of finding the minimal entropy needed to achieve a given level of distortion which is $(1/12)a^2$. Hence, one could as well analyze an entropy-constrained distortion minimizing quantizer, which has in fact been studied in the source-coding literature [25], where one can also find the following useful result.

Proposition 3.2: If a is not an integer, then the optimal quantizer consists of $K = \lceil |a| \rceil$ bins where $K-1$ of them occur with equal probability, p_1, each, and the remaining one occurs with probability $p_2 \leq p_1$, such that

$$(K-1)p_1^3 + p_2^3 = \frac{1}{a^2}.$$

Furthermore, the minimum rate needed to satisfy Criterion 1, equivalently for the monotonic boundedness of the error variance, is

$$(K-1)p_1\log_2(1/p_1) + p_2\log_2(1/p_2). \tag{9}$$

B. Criterion 2: Asymptotic Boundeness in Distortion

Asymptotic stability of the state estimation error variance has been investigated earlier in [14], where asymptotic quantization analysis has been used to determine the rate requirements. Here, we take a different, information-theoretic, approach to arrive at the same rate requirements, and then employ asymptotic quantization as $T \to \infty$.

Theorem 3.3: Consider a scalar linear system $x_{t+1} = ax_t$, where the pdf of the initial state has finite support. Then, the following hold.

i) The bit rate required for boundedness of the state estimation error variance as the terminal time $T \to \infty$, is $\max(0, \log_2(|a|))$ per stage.
ii) This rate is achievable by quantization. Furthermore, uniform quantizer is the optimal variable length quantizer.

Proof: i) The entropy at time t is

$$\begin{aligned} H(x_t) &= H(a^t x_0) \\ &= H(x_0) + t\log_2(|a|) \end{aligned} \tag{10}$$

and the mutual information between x_t and $\hat{x}_t$ is

$$I(x_t;\hat{x}_t) = H(x_t) - H(x_t|\hat{x}_t) = H(x_t) - H(x_t - \hat{x}_t|\hat{x}_t). \tag{11}$$

Let $\mathcal{P}_\mathcal{D}$ be the set of all probabilistic maps that achieve a given finite distortion level D. Then, for any corresponding rate R, we have $R \geq \inf_{\mathcal{P}_\mathcal{D}} I(x_t, \hat{x}_t)$. Thus

$$\begin{aligned} R &\geq \inf_{\mathcal{P}}\{H(x_t) - H(x_t - \hat{x}_t|\hat{x}_t)\} \\ &= \inf_{\mathcal{P}}\{H(x_0) + t\log_2(|a|) - H(x_t - \hat{x}_t|\hat{x}_t)\} \\ &\geq \inf_{\mathcal{P}}\{H(x_0) + t\log_2(|a|) - H(x_t - \hat{x}_t)\} \\ &\geq H(x_0) + t\log_2(|a|) - (1/2)\log(2\pi e D) \end{aligned} \tag{12}$$

where the last two steps are due to the facts that conditioning on a random variable does not increase the entropy, and for a random variable with a finite variance D, the differential entropy is maximized by the Gaussian distribution (with that variance) [23], with the maximum value being $(1/2)\log(2\pi e D)$. Note that since we are interested in quantizers, the quantization error conditioned on the quantized values will in fact have a compact support set for its distribution. Nonetheless, using the entropy corresponding to the Gaussian density serves our purpose by providing a lower bound.

If we divide both sides of the last inequality in (12) by t, and let t approach ∞, this yields as a necessary condition on the average rate, R_{avg}, the lower bound $\log_2(|a|)$. ii) We employ asymptotic quantization theory for this part of the proof. Although the rate per stage is finite, as the terminal time goes to ∞, one can regard this problem as one where the number of quantization levels goes to ∞. Reference [22] has shown that the uniform quantizer followed by an entropy coder is at most 0.255 bits worse than the optimal quantizer. Since $0.255/T \to 0$, it readily follows that the uniform quantizer is optimal. $\diamond$

C. Criterion 3: Stability in Support Size

Theorem 3.4: Consider the scalar system (2) and suppose that fixed-length quantization is used. The rate required for stability in the support size is at least $\max(0, \log_2(|a|))$. The minimum rate achievable by a fixed-length time-invariant quantizer is $\log_2(C(|a|))$, where $C(x)$ is the smallest integer that is strictly larger than x (and is thus a modified *ceiling* function).

Proof: If a is not an integer, pick $K = \lceil(|a|\rceil)$, whereas if a is an integer, $K = |a| + 1$. Thus, the rate is: $R = \sum_i \log_2(C(|a|))$.$\diamond$

IV. Quantization and Coding for Higher-Dimensional LTI Systems

We now consider extensions of the results of the previous section to the multidimensional case—the class of systems described by (1), where (A, B) is a controllable pair, with A having at least one unstable eigenvalue and no zero eigenvalues. We can again instead consider the control-free system $x_{t+1} = Ax_t$, as the evolution of the uncertainty in the initial state does not depend on the control u. We will take the initial state vector x_0 to be the realization of a random vector X_0 with a finite-support distribution. The objective, as in the previous section, is to design optimal dynamic quantizers, $Q := \{Q_t(X), t = 0, 1, 2, \dots\}$ and corresponding encoder-decoder pairs to achieve stability. We again have three criteria, as introduced earlier, but with only the norm $|\cdot|$ appropriately changed to the n-dimensional Euclidean norm. If A is diagonalizable, even with complex eigenvalues, all the results obtained for the scalar case would be equally applicable to the decoupled components. When an eigenvalue is complex (which of course comes in conjugate pairs), then one only has to apply a rotation to the received data and, hence, only the absolute value of the eigenvalue will matter in the rate analysis.

Now, if A is not diagonalizable, it can be *block-diagonalized*, say with two blocks, where the first block has only stable eigenvalues, and the second one unstable eigenvalues. For the stable modes, one does not need to use the channel (since those modes asymptotically go to zero), and hence for the remaining discussion and analysis we can assume, without any loss of generality, that A has only unstable eigenvalues. Further, because of the reasoning in the previous paragraph, the main focus will be on the cases when A is not diagonalizable.

With this structure for A, we now study each of the criteria introduced, starting with Criterion 1.

Theorem 4.1: For an n-dimensional linear system $x_{t+1} = Ax_t$, where all eigenvalues, $\{\lambda_i\}$, of A are unstable, and the initial state x_0 is a finite-support random vector, the minimum rate required for monotonic boundedness of the state estimation error entropy is $\sum_i \log_2(|\lambda_i|)$.

Proof: The proof is almost identical to that of Theorem 3.1. Lemmas 3.2 and 3.3 apply to the multidimensional case as well, with the intervals being replaced with higher dimensional bins. We again have the mutual information as

$$I\left(x_t; q_t | q_0^{t-1}\right) = \left[H\left(x_t | q_0^{t-1}\right) - H\left(x_t | q_t, q_0^{t-1}\right)\right]$$

where q_0^{t-1} is the side information available. Since $x_t = Ax_{t-1}$, and $H(AX) = \log_2(|A|) + H(X)$ for any random variable X and any square matrix A, we have

$$H\left(x_t | q_0^{t-1}\right) = \log_2(|A|) + H\left(x_{t-1} | q_0^{t-1}\right)$$

where $|A|$ is the determinant of A. As in the Proof of Theorem 3.1, after a few steps we arrive at the inequality $I(x_t, q_t | q_0^{t-1}) \geq \log_2(|A|)$, which says that we need at least an additional $\log_2(|A|)$ bits to make the conditional differential entropy sequence $\{H(x_t - \hat{x}_t | \hat{x}_t)\}$ nonincreasing. Since $|A| = \prod_i |\lambda_i|$, the rate is $\sum_i \log_2 |\lambda_i|$.$\diamond$

Remark 2: In the above, if we had not taken out the stable modes of A, then the stable eigenvalues would also have entered the computation of the rate required to keep the differential entropy bounded, as the summation in $\max(0, \sum_i \log_2 |\lambda_i|)$ would be over all eigenvalues of A, whereas as we have argued above, stable eigenvalues can be safely left out. This shows that entropy is not a good measure to use in the multidimensional case provided that one carefully separates out the stable part of the system, since otherwise the analysis leads to a very loose bound. This all stems from the fact that the volume of any cube of dimension 2 or higher (corresponding in this case to a bin) can be made arbitrarily small without making the lengths of all edges arbitrarily small. As a parallel argument, asymptotic quantization theory is not directly applicable here either. In high-rate quantization, the quantization error is taken to be uniformly distributed over cells of arbitrarily small volume, and that $f(x) \approx f(q_i)\ \forall x \in \mathcal{B}_i$, where $\mathcal{B}_i$ is one such bin, and f is a pdf. However, as we argued previously, $V(\mathcal{B}_i) \to 0$, where V denotes the volume, does not imply that $\max_i |x_i - Q(x)_i| \to 0$; hence, the property $f(x) \approx f(Q(x))\ \forall x \in \mathcal{B}_i$, which follows from the continuity of the pdf, does not apply. Thus, to employ asymptotic quantization theory in a meaningful way, one should only consider the unstable eigenvalues of the system.$\diamond$

We now move on to Criterion 2, and have the following result.

Theorem 4.2: Consider the n-dimensional system $x_{t+1} = Ax_t$, where A has only unstable eigenvalues, and x_0 is a random vector. Then, the following hold.

i) Under Criterion 2, a lower bound on the required rate is $\sum_i \log_2 |\lambda_i|$.

ii) The bound in i) is achievable with both fixed-length or variable-length quantizers.

Proof: i) The entropy at time step t is

$$H(x_t) = H(A^t x_0) = H(x_0) + t \log_2(|A|). \tag{13}$$

Let $\mathcal{P}_D$ be the set of all probabilistic maps achieving a bounded final distortion level of D. Then, $R \geq \inf_{\mathcal{P}_D} I(x_t, \hat{x}_t)$. For any corresponding rate R, we have

$$\begin{aligned} R &\geq \inf_{\mathcal{P}_D} \{H(x_t) - H(x_t - \hat{x}_t | \hat{x}_t)\} \\ &= \inf_{\mathcal{P}_D} \{H(x_0) + t \log_2(|A|) - H(x_t - \hat{x}_t | \hat{x}_t)\} \\ &\geq \inf_{\mathcal{P}_D} \{H(x_0) + t \log_2(|A|) - H(x_t - \hat{x}_t) \end{aligned} \tag{14}$$

where the inequality uses the fact that conditioning does not increase the entropy. Let the Euclidean distortion for the vector $x_t - \hat{x}_t$ be denoted by D_t, and the covariance matrix of the componentwise errors be denoted by C_t. Clearly, $D_t = \mathrm{Trace}(C_t)$. Let D_t be finite, which also makes C_t a matrix with finite entries. Now, among random vectors with a fixed covariance matrix, the differential entropy is maximized by a jointly Gaussian distribution, which in turn has a finite entropy. Hence, the minimum rate, $R^*(D_t)$, achieving a distortion level of D_t, satisfies the following inequality for any nonnegative–definite matrix C_t whose trace equals D_t:

$$\begin{aligned} R^*(D_t) &\geq H(x_0) \\ &\quad + t \left(\sum_i \log_2(|\lambda_i|)\right) (1/2) \log((2\pi e)^n |C_t|) \end{aligned} \tag{15}$$

Dividing both sides by t, and letting $t \to \infty$, the desired result follows.

ii) In this part of the proof, we employ results and bounds from asymptotic quantization theory [26]. Suppose that K is the number of quantization levels, $D_{2,n}(K,f)$ is the second moment of the distance between the quantizer's input and output, and f is the signal pdf. Then, for large K, the distortion of the quantization satisfies $D_{2,n}(K,f) \leq K^{-\frac{2}{n}} B$ where

$$B = b\left[\int f^p(x)dx\right]^{1/p}, \qquad p = n/(n+2)$$

and b is a constant. Without any loss of generality, let the initial state be confined to the n-dimensional unit cube. Suppose that at time t, x_t has pdf f_t. The resulting distortion then satisfies

$$\lim_{K\to\infty} D_{2,n}(K,f)K^{2t/n} \leq b(\int (f_t(x_t))^{\frac{n}{n+2}} dx)^{(n+2)/n}. \quad (16)$$

Now, given a real number p, $0 < p < 1$, and a support volume V, the quantity $\int_V f(x)^p dx$ is maximized over all pdf's $f(x)$ by the uniform distribution. Since $x_t = A^t x_0$, and $\mathrm{vol}(f_t(x_t)) = |A|^t \mathrm{vol}(f_0(x_0))$, we have

$$\lim_{K\to\infty} D_{2,n}(K,f)(K/|A|)^{2t/n} \leq b(\mathrm{vol}(f_0(x_0)))^{2/n}.$$

Hence, distortion remains bounded in t, if $K \geq |A|$. For variable length encoding, the distortion satisfies [26], [27]

$$D_n(R) \leq c_n 2^{(2/n)H(A^t x)} 2^{-(2/n)R} \quad (17)$$

where c_n is a term independent of the source distribution.

Using the facts that $H(A^t x) = t\log_2(|A|) + H(x)$ and $\log_2(|A|) = \sum_i \log_2(|\lambda_i|)$ and noting that by our assumption, all eigenvalues of A are unstable, the rate required for boundedness of the terminal-time distortion is $\sum_i \log_2 |\lambda_i|$. $\diamond$

Remark 3: In the result of Proposition 4.2, if A had both stable and unstable eigenvalues, then the expression for the minimum rate would be $\sum_i \max(0, \log_2(|\lambda_i|))$. $\diamond$

Finally, we consider stability in support size.

Proposition 4.1: Consider the n-dimensional system $x_{t+1} = Ax_t$. Under time-invariant quantization, the minimum achievable fixed-length quantizer rate for stability in support size is $R = \sum_i \log_2(C(|\lambda_i|))$, where $C(x)$ is the smallest integer that is strictly larger than x (and is thus a modified *ceiling* function).

Proof: For diagonalizable systems, the condition for stability follows from that of scalar systems.

For the nondiagonalizable case, suppose that the system is in Jordan form, with J_m be the Jordan block corresponding to an eigenvalue λ_m. Then, a bound on the uncertainty region, after quantization, satisfies

$$W(e_{t+1}^m) \leq (1/K_m)[(|\lambda_m|W(e_t^m) + W(e_t^l + \cdots + W(e_t^s)]$$

for the mode λ_m. Letting K be a diagonal matrix, with the reciprocal of the number of levels on its diagonals, we can rewrite the above as a linear system: $W_{t+1} \leq KJ_m W_t$, with

$$KJ = \begin{bmatrix} \lambda_m/K_m & 1 & 0 \\ 0 & \lambda_m/K_m & 1 \\ 0 & 0 & \lambda_m/K_m \end{bmatrix}$$

and the condition for stability becomes $(1/K_m)|\lambda_m| < 1$ for each m. Thus, the result follows. $\diamond$

V. Concluding Remarks

In this paper, we have studied the rate requirements for stability of the state estimation error in LTI systems under three types of criteria. Information theory and source-coding theory have been used to derive achievable bounds, and optimal quantizers have been designed as implementable codes. The optimal performance under each criterion has been quantified, with lower bounds which are tight in some cases. We have observed that uniform quantization is indeed very efficient. We also demonstrated that entropy is not necessarily an appropriate measure for rate analysis in multidimensional control systems, unless one carefully separates out the stable and unstable modes of the system.

Multisensor and multicontroller systems can also be studies in an analogous fashion. In a multisensor problem, without system and channel noise, one can use the techniques here to show that the information theoretic lower bounds are tight [19], [11]. Likewise, in the multicontroller case, one can, under the strong connectivity assumption, obtain encoding schemes based on quantization in the absence of a centralized plant decoder, where the plant acts as a relay [12].

The techniques used here can also be applied to systems driven by noise. However, the construction for coder and controllers requires a study that involves stochastic analysis, since almost sure stabilizability is not possible in that case [6]. The techniques used here can be used to obtain necessary and sufficient conditions on the channels, coder and controllers for the existence of an invariant distribution [28], [29].

Appendix

A. Proof of Lemma 3.1

The Kullback–Leibler divergence between two probability density functions (pdfs) $g(x)$ and $h(x)$ is defined as

$$D(g,h) = \int g(x)\log(\frac{g(x)}{h(x)})\,dx.$$

Since $X_{t+1} = aX_t$, the pdf $f_{t+1}(y)$ of $Y = X_{t+1}$, in terms of the pdf $f_t(x)$ of $X = X_t$ is $(1/|a|)f_t(y/a)$. Suppose that at each time $t > 0$, a uniform quantizer with a spacing $\Delta_t = \Delta_{t-1}|a|/K$ is used. Then, at time t, the centralized support interval of the density will be $[-(|a|^t/2K^{t+1})\Delta_0, (|a|^t/2K^{t+1})\Delta_0]$. Hence, after t steps, the conditional quantization error probability density, given that $x \in \mathcal{B}_k$ where $\mathcal{B}_k$ is the kth quantization bin, becomes

$$f_Q(q)(t) = \frac{(1/|a^t|) f_0(y/a^t)}{\int_{k\Delta_t}^{(k+1)\Delta_t} (1/|a^t|) f_0(\xi/a^t) d\xi}.$$

Thus, the Kullback–Leibler divergence between the uniform pdf with a support size Δ_t and the quantization error pdf is

$$D = -\int_{k\Delta_t}^{(k+1)\Delta_t} \frac{1}{\Delta_t} \log\left(\frac{\frac{\Delta_t}{|a^t|} f_0(y/a^t)}{\int_{k\Delta_t}^{(k+1)\Delta_t} \frac{1}{|a^t|} f_0(\xi/a^t) d\xi}\right) dy$$

and by a change of variables, this becomes

$$D = -\int_{k\frac{\Delta_t}{|a^t|}}^{(k+1)\frac{\Delta_t}{|a^t|}} \frac{|a^t|}{\Delta} \log\left(\frac{\frac{\Delta_t}{|a^t|} f_0(y)}{\int_{k\frac{\Delta_t}{|a^t|}}^{(k+1)\frac{\Delta_t}{|a^t|}} f_0(z) dz}\right) dy.$$

By the continuity of f_0, and using the mean-value theorem, together with the fact that $\Delta_t/|a^t| = \Delta_0 K^t$, the distance expression is equivalent to

$$D = -\int_{k\frac{\Delta_0}{K^t}}^{(k+1)\frac{\Delta_0}{K^t}} \frac{K^t}{\Delta_0} \log(\frac{f_0(y)}{f_0(\bar{z}_t)}) dy$$

for some $\bar{z}_t \in ((k\Delta_0/K^t), ((k+1)\Delta_0/K^t))$. Again by the continuity of f_0, given any $\epsilon > 0, \exists T$ such that $\forall t > T, 1-\epsilon < (f_0(y)/f_0(\bar{z}_t)) < 1+\epsilon$. Therefore, for $t > T, D$ is bounded from above by

$$-\int_{k\frac{\Delta_0}{K^t}}^{(k+1)\frac{\Delta_0}{K^t}} \frac{K^t}{\Delta_0} \log(1-\epsilon) dy = -\log(1-\epsilon)$$

and bounded from below by $-\log(1+\epsilon)$. Hence, the Kullback–Leibler divergence asymptotically converges to $-\log(1) = 0$. $\diamond$

B. *Proof of Lemma 3.2*

A weak* continuous functional on a weak* compact set admits a minimum [30]. Hence, the lemma follows by showing that, for each t, the set of quantization output distributions satisfying the conditional entropy bound is weak* compact, and the differential entropy to be minimized is weak* continuous over the set of quantizers that lead to satisfaction of the conditional entropy bound.

For the former, we use the properties of entropy for discrete random variables. Fix t, and let Y_t be a random variable taking values on the set of integers, corresponding to the bin index of the quantizer. Since the quantizer bins are nonoverlapping, there exists a $\psi > 0$, such that $\delta_i > \delta_{i-1} + \psi$ for all i. Hence, by Markov's inequality, for each positive integer n

$$\begin{aligned} P(|Y_t| > n) &\leq P(|x_t| > (n-1)\psi | q_0^{t-1}) \\ &\leq E[\frac{x_t^2}{(n-1)^2\psi^2} | q_0^{t-1}] \leq \frac{a^2\sigma^2}{(n-1)^2\psi^2} \end{aligned}$$

This shows that the tails of the distribution of Y_t are arbitrarily small and, hence, by Prohorov's theorem [31], the set of such quantizer output distributions is tight. Tightness is equivalent to relative weak* compactness (that is, weak* compactness of the closure of the set), but since the set of output distributions is closed (because of the inequality constraint on the entropy), it is weak* compact.

We next show the second desired property, that is the differential entropy (3) is weak* continuous over the set of admissible quantizers.

Let e_t denote the conditional estimation error. We first argue that the density $p(e_t)$ is continuous in f_t, where the metric is the total variation. This result follows from the definition of a quantizer and the fact that x_t has a continuous probability density function (pdf) for all t. Hence, it will suffice to show that the differential entropy is weak* continuous in the pdf of the estimation error, and this will be true (as shown below) if the error entropy is uniformly integrable.

By assumption, using Markov's inequality, the following holds for all t:

$$\lim_{K\to\infty} P(e_t > K) \leq \lim_{K\to\infty} \sigma^2/K^2 = 0.$$

Again, following Prohorov's theorem, at time t, the set of input distributions is compact in weak* topology [31]. Let $\mathcal{E}$ be the set of estimation error densities. We suppress the time index, and let $p(e)$ denote the pdf for the error.

The Gaussian density maximizes the entropy among distributions with the same variance, and hence, via Jensen's inequality

$$P(-\log_2(p(e)) > K) \leq (1/2)\log_2(2\pi e\sigma^2)/K^2$$

for all $p(e) \in \mathcal{E}$. This implies that

$$\lim_{N\to\infty} \sup_{p(e)\in\mathcal{E}} \int -p(e)|\log_2(p(e)) 1_{\{\log_2(p(e))|>N\}} de = 0$$

where $1_{\{.\}}$ is the indicator function. Thus, $p(e)\log_2(p(e))$ is uniformly integrable and, hence, the entropy contributions from distant sets are negligible. Then, in view of tightness, we can essentially work with values of $p(e)$ from a compact set. Now, for N a sufficiently large positive integer, let $S_N := \{e : |\log_2(p(e))| \leq N\}$. Let $p^1(e), p^2(e)$ be two pdf's in $\mathcal{E}$. By a slight abuse of notation, let $H(p^i(e))$ denote the entropy of e with pdf's $p^i(e)$, $i = 1, 2$.

We then have

$$\begin{aligned} &|H(p^1(e)) - H(p^2(e))| \\ &\leq \int_{S_N} |p^1(e)\log_2(p^1(e)) - p^2(e)\log_2(p^2(e))| de \\ &+ \int_{S_N^C} |p^1(e)\log_2(p^1(e)) - p^2(e)\log_2(p^2(e))| de \end{aligned}$$

where S_N^C denotes the complement of S_N. We have a uniform bound on the second term for each $N$$, which we denote by η_N for all $p^i(e) \in \mathcal{E}$.

We have

$$\int_{S_N} p^1(e)|\log_2(p^1(e)) - \log_2(p^2(e))|de$$
$$+ \int_{S_N} |(p^1(e) - p^2(e))\log_2(p^2(e))|de \le c_N \int_{S_N} |p^1(e) - p^2(e)|de$$
$$+ d_N \int_{S_N} |p^1(e) - p^2(e)|de = (c_N + d_N)\int_{S_N} |p^1(e) - p^2(e)|de$$

where c_N and d_N are finite positive constants. The finiteness of c_N (and validity of the first bound) follows from the facts that $\log_2(.)$ is a continuous function $\log_2(1+x) \le [\log_2(e)]x$ for $x \ge 0$, and $p^i(e) > 2^{-N}$ for $i = 1, 2$. Finiteness of d_N (and validity of the second bound) follows from $p^2(e) > 2^{-N}$. Let $\eta'_N := \int_{S_N^C} |p^1(e) - p^2(e)|de$.

Now, consider two quantizers $Q^1(x) \in \mathcal{Q}$ and $Q^2(x) \in \mathcal{Q}$. Let $\mathcal{B}_i^j$ denote the bins in the quantizer $Q^j(x), j \in \{1, 2\}$. Define

$$p_i^j := P(x \in \mathcal{B}_i^j), \qquad j = 1, 2, i \in \mathcal{Z}.$$

Given $\epsilon > 0$, we can pick N large enough so that $\eta_N < \epsilon$. Then, there exist an $\alpha > 0$, and

$$\delta := \frac{1}{\alpha}[(\epsilon - \eta_N)/(c_N + d_N)] + 2\eta'_N$$

such that for every uniformly integrable p^1, p^2, the bound $\|p_i^1 - p_i^2\|_{TV} \le \delta$ implies $\|p^1(e) - p^2(e)\|_{TV} \le \alpha\delta$ which in turn implies $|H(p^1) - H(p^2)| < \epsilon$, where $\|.\|_{TV}$ denotes the total variation norm. This then completes the proof of weak* continuity. $\diamond$

C. Proof of Lemma 3.3

We first show that both the constraint functional and the objective functional are Fréchet differentiable, and admit Fréchet derivatives, with the Fréchet derivative for the former being onto. This will then imply the existence of a Lagrange multiplier. Let $\mathbf{h} = \{h_i, i = \pm 0, 1, 2, \ldots\}$ be a probability mass function. The Gateaux differential of $H(\mathbf{p}) = -\sum_i p_i \log_2(p_i)$ at $\mathbf{p} := \{p_i, i = \pm 0, 1, 2, \ldots\}$, with increment $\mathbf{h}$, is

$$\delta H(\mathbf{p}, \mathbf{h}) = \lim_{\alpha \to 0} \left\{ 1/\alpha \left(\sum_i (p_i + \alpha h_i)\log_2(p_i + \alpha h_i) - p_i \log_2(p_i) \right) \right\}$$

which can be written as:

$$\begin{aligned}\delta H(\mathbf{p}, \mathbf{h}) &= -\lim_{\alpha \to 0} \{ (\sum_i (p_i)(\log_2(p_i + \alpha h_i) - \log_2(p_i)))/\alpha \\ &\quad - \sum_i h_i \log_2(p_i + \alpha h_i)\} \\ &= -\sum_i h_i(\log_2(e) + \log_2(p_i)) \\ &=: (\mathbf{g}, \mathbf{h})\end{aligned}$$

where $\mathbf{g} = \{g_i, i = \pm 0, 1, 2, \ldots\}$ with $g_i = -(\log_2(e) + \log_2(p_i))$, and $(\cdot, \cdot)$ denotes the ℓ_2 inner product.

Thus, the Gateaux differential exists, and so do the Gateaux and Fréchet derivatives, which are both equal to $\mathbf{g}$, and are *onto*. The objective functional, the entropy, is continuously differentiable in $g_t(\cdot, \cdot)$. We show that $g_t(.,.)$ is continuously differentiable in f_t. Here, g_t is the decoder function minimizing the entropy of the estimation error. Given f_t, we rewrite $H(x_t - g_t(f_t, q_0^{t-1})|q_0^{t-1}, f_t)$ as $H(x_t - g_{t,q_0^{t-1}}(f_t)|q_0^{t-1}, f_t)$ where the decoder function $g_{t,q_0^{t-1}}$ now takes only f_t as its argument. We have

$$\begin{aligned}&H(x_t - g_{t,q_0^{t-1}}(f_t)|q_0^{t-1}, f_t) \\ &\quad \ge H(x_t - g_{t,q_0^{t-1}}(f_t)|g_{t,q_0^{t-1}}(f_t), q_0^{t-1}, f_t) \\ &\quad = H(x_t|g_{t,q_0^{t-1}}(f_t), q_0^{t-1}, f_t) \\ &\quad \ge H(x_t|q_0^{t-1}, f_t).\end{aligned}$$

Since the estimation error is independent of the reconstruction given the reconstruction, the first inequality becomes an equality. For any deterministic, one-to-one decoder function, the second inequality also becomes an equality since

$$\begin{aligned}&H\left(x_t|g_{t,q_0^{t-1}}(f_t), q_0^{t-1}\right) - H\left(x_t|f_t, q_0^{t-1}\right) \\ &\quad = H\left(x_t|g_{t,q_0^{t-1}}(f_t), q_0^{t-1}\right) - H\left(x_t|f_t, q_0^{t-1}, g_{t,q_0^{t-1}}(f_t)\right) \\ &\quad = I\left(x_t; f_t|q_0^{t-1}, g_{t,q_0^{t-1}}(f_t)\right) \\ &\quad = H\left(f_t|q_0^{t-1}, g_{t,q_0^{t-1}}(f_t)\right) - H\left(f_t|x_t, q_0^{t-1}, g_{t,q_0^{t-1}}(f_t)\right) \\ &\quad = 0.\end{aligned}$$

However, the conditional entropy $H(x_t|f_t, q_0^{t-1})$ is Fréchet differentiable in f_t, as was shown previously in the study of the constraint functional. These then imply the existence of a Lagrange multiplier [30], which is further positive (because of the nature of the constraint).

Now, for $\lambda > 0$, introduce the Lagrangian

$$\begin{aligned}J_\lambda := H(x_t - g_t\left(f_t(x_0^t, q_0^{t-1}), q_0^{t-1}\right)|q_0^{t-1}) \\ + \lambda H(f_t\left(x_0^t, q_0^{t-1}\right)|q_0^{t-1}).\end{aligned}$$

Since $q_t = f_t(x_0^t, q_0^{t-1})$, we have

$$H\left(x_t - g_t(q_t, q_0^{t-1})|q_0^{t-1}\right) + \lambda H(q_t) =: k\left(q_t, q_0^{t-1}, x_t\right)$$

for some functional $k(\cdot, \cdot, \cdot)$. However, since there is perfect access to x_t and q_0^{t-1}, the optimal choice for q_t will be a function of only (x_t, q_0^{t-1}). $\diamond$

Acknowledgment

The authors acknowledge the useful discussions with Dr. H. Ishii on this general topic and on an earlier version of this paper. They also acknowledge the constructive comments and suggestions of three reviewers and, particularly, the Associate Editor.

References

[1] D. F. Delchamps, "Stabilizing a linear system with quantized state feedback," *IEEE Trans. Autom. Control*, vol. 35, no. 8, pp. 916–924, Aug. 1990.

[2] G. N. Nair and R. J. Evans, "Stabilizability of stochastic linear systems with finite feedback data rates," *SIAM J. Control Optim.*, vol. 43, pp. 413–436, Jul. 2004.

[3] R. Bansal and T. Başar, "Simultaneous design of measurement and control strategies for stochastic systems with feedback," *Automatica*, vol. 25, no. 5, pp. 679–694, 1989.

[4] A. Sahai and S. Mitter, "The necessity and sufficiency of anytime capacity for stabilization of a linear system over a noisy communication link Part I: Scalar systems," *IEEE Trans. Inform. Theory*, vol. 52, no. 8, pp. 3369–3395, Aug. 2006.

[5] A. V. Savkin and I. R. Petersen, "Set-valued state estimation via a limited capacity communication channel," *IEEE Trans. Autom. Control*, vol. 48, no. 4, pp. 676–680, Apr. 2003.

[6] A. S. Matveev and A. V. Savkin, "Shannon zero error capacity and the problem of almost sure observability over noisy communication channels," in *Proc. IEEE Conf. Decision Control and Eur. Control Conf.*, Sevilla, Spain, 2005, pp. 3219–3224.

[7] N. Elia and S. Mitter, "Stabilization of linear systems with limited information," *IEEE Trans. Autom. Control*, vol. 46, no. 9, pp. 384–1400, Sep. 2001.

[8] F. Fagnani and S. Zampieri, "Stability analysis and synthesis for scalar linear systems with a quantized feedback," *IEEE Trans. Autom. Control*, vol. 48, no. 9, pp. 1569–1584, Sep. 2003.

[9] S. Yüksel and T. Başar, "Quantization for LTI systems with noiseless channels," in *Proc. IEEE Conf. Control Appl.*, Istanbul, Turkey, Jun. 2003, pp. 423–429.

[10] S. Yüksel, "State estimation and control for LTI systems over communication channels," M.S. thesis, Univ. Illinois, Urbana-Champaign, IL, Oct. 2003.

[11] S. Yüksel and T. Başar, "On the absence of rate loss in decentralized sensor and controller structure for asymptotic stability," in *Proc. IEEE Amer. Control Conf.*, Minneapolis, MN, Jun. 2006, pp. 5562–5567.

[12] S. Yüksel and T. Başar, Communications for decentralized stability under multi-controller structure UIUC-CSL Tech. Rep., May 2005.

[13] S. Tatikonda, "The sequential rate-distortion function and joint source-channel coding with feedback," in *Proc. Allerton Conf.*, Oct. 2003, pp. 191–200.

[14] G. N. Nair and R. J. Evans, "Exponential stabilisability of multidimensional linear systems with finite data rates," *Automatica*, vol. 39, no. 4, pp. 585–593, Apr. 2003.

[15] H. Ishii and B. A. Francis*, Limited Data Rate in Control Systems with Networks*, ser. Lecture Notes in Control and Information Sciences. Berlin, Germany: Springer-Verlag, 2002, vol. 275.

[16] S. V. Sarma, M. A. Dahleh, and S. Salapaka, "Synthesis of efficient time-varying bit-allocation strategies maintaining input-ouput stability," in *Proc. Allerton Conf.*, Oct. 2004, pp. 220–229.

[17] S. Tatikonda and S. Mitter, "Control under communication constraints," *IEEE Trans. Autom. Control*, vol. 49, no. 7, pp. 1056–1068, Jul. 2004.

[18] S. Tatikonda and S. Mitter, "Control over noisy channels," *IEEE Trans. Autom. Control*, vol. 49, no. 7, pp. 1196–1201, Jul. 2004.

[19] S. Tatikonda, "Control under communication constraints," Ph.D. dissertation, Mass. Inst. Technol., Cambridge, MA, 2000.

[20] 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.

[21] R. M. Gray and D. L. Neuhoff, "Quantization," *IEEE Trans. Inform. Theory*, vol. 44, no. 10, pp. 2325–2384, Oct. 1998.

[22] H. Gish and J. N. Pierce, "Asymptotically efficient quantization," *IEEE Trans. Inform. Theory*, vol. IT-14, no. 9, pp. 676–683, Sep. 1968.

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

[24] R. E. Blahut*, Principles of Information Theory*, Aug. 2002, Course Notes at UIUC.

[25] A. György and T. Linder, "Optimal entropy-constrained scalar quantization of a uniform source," *IEEE Trans. Inform. Theory*, vol. 46, no. 11, pp. 2704–2711, Nov. 2000.

[26] P. Zador, "Asymptotic quantization error of continuous signals and the quantization dimension," *IEEE Trans. Inform. Theory*, vol. IT-28, no. 3, pp. 139–148, Mar. 1982.

[27] M. P. Schutzenberger, "On the quantization of finite-dimensional messages," *Inform. Control*, vol. 1, pp. 153–158, 1958.

[28] S. Yüksel and T. Başar, "Coding and control over discrete noisy forward and feedback channels," in *Proc. IEEE Conf. Decision and Control and European Control Conf.*, Seville, Spain, 2005, pp. 2517–2522.

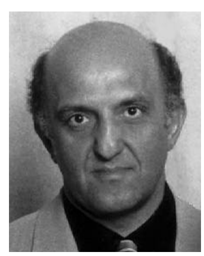
[29] S. Yüksel and T. Başar, "Control of multi-dimensional linear systems on noisy channels," in *Proc. IEEE Conf. Decision and Control*, San Diego, CA, Dec. 2006.

[30] D. G. Luenberger*, Optimization by Vector Space Methods*. New York: Wiley, 1969.

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

Serdar Yüksel (S'02) received the B.S. degree in electrical engineering from Bilkent University, Ankara, Turkey, in 2001, and the M.S. and Ph.D. degrees in electrical engineering from the University of Illinois at Urbana-Champaign (UIUC), in 2003 and 2006, respectively.

He is currently a Postdoctoral Researcher with the Department of Electrical Engineering, Yale University, New Haven, CT. His research interests include stochastic control theory, information theory, control over heterogeneous networks and channels, decentralized control, and estimation and control.



Tamer Başar (S'71–M'73–SM'79–F'83) received the B.S.E.E. degree from Robert College, Istanbul, Turkey, in 1969, and the M.S., M.Phil, and Ph.D. degrees from Yale University, New Haven, CT, in 1970–1972.

He has been with the University of Illinois at Urbana-Champaign (UIUC) since 1981, where he currently holds the positions of the Fredric G. and Elizabeth H. Nearing Endowed Professor of Electrical and Computer Engineering, Center for Advanced Study Professor, and Research Professor at the Coordinated Science Laboratory. He joined UIUC in 1981 after holding positions at Harvard University, Cambridge, MA, and Marmara Research Institute, Turkey. He has published extensively in systems, control, communications, and dynamic games, and has current research interests in modeling and control of communication networks; control over heterogeneous networks; resource allocation, management and pricing in networks; mobile computing; security issues in computer networks; and robust identification, estimation and control.

Dr. Başar is currently the Editor-in-Chief of *Automatica*, Editor of the Birkhäuser Series on *Systems and Control*, Managing Editor of the *Annals of the International Society of Dynamic Games (ISDG)*, and a member of editorial and advisory boards of several international journals in control, wireless networks, and applied mathematics. He has received several awards and recognitions over the years, among which are the Medal of Science of Turkey (1993); Distinguished Member Award (1993), Axelby Outstanding Paper Award (1995), and Bode Lecture Prize (2004) of the IEEE Control Systems Society (CSS); Millennium Medal of IEEE (2000); Tau Beta Pi Drucker Eminent Faculty Award of UIUC (2004); the Outstanding Service Award (2005) and the Giorgio Quazza Medal (2005) of the International Federation of Automatic Control (IFAC); and the Richard Bellman Heritage Award of the American Automatic Control Council (2006). He is a member of the National Academy of Engineering (of USA), a member of the European Academy of Sciences, a Fellow of IFAC, a past President of CSS, and the founding President of ISDG.