STATE ESTIMATION AND CONTROL FOR LTI SYSTEMS OVER COMMUNICATION CHANNELS

BY

SERDAR YÜKSEL

B.S., Bilkent University, 2001

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Electrical Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2003

Urbana, Illinois

ABSTRACT

In this thesis, state estimation and control problems in linear time invariant (LTI) systems where the controller and the plant are connected through digital noiseless channels and communication networks are investigated. For deterministic (noiseless) systems, uniform quantization noise is shown to be the effective one, and distortion constrained quantization is introduced as a time-invariant stable quantization method. Forward-looking schemes where the intention is the stabilization of the state estimation error, and recursive schemes where the objective is to achieve a monotonic variance in state estimation error are considered. In deterministic systems both approaches are shown to lead to the same performance in terms of the communication rate values. In stochastic (noisy) systems however, forward-looking schemes require less stringent conditions. For systems with communication networks, the rate required for the stability of the differential entropy of the state estimation error is determined as a function of the system and network parameters. The trade-off between the sampling period, packet loss probability and the rate requirements is illustrated, a Markov chain model is introduced to capture the state of reliability of the network, and conditions for stability in the mean in a control theoretic setting are studied. Conditions for the mean-square stability in the state estimation error, stability in the mean, and stability in the mean-square of an LTI system with communication networks are introduced, and in case of a packet loss, the usage of the latest control and the zero control usage are compared. The communication rate requirements for higher dimensional centralized and decentralized control systems where the controller and the plant are connected via a noiseless bandlimited channel are also studied.

Recursive time-invariant quantizers that achieve stability in the sense of worst-case state estimation error are introduced. Rate requirements for centralized schemes are shown to be lower than those for decentralized schemes. A quantification of the information sharing between the controllers, such as full, instant, and one-step delayed information sharing, is shown to be an important factor contributing to communication requirements and complexity. Slepian-Wolf coding argument is used to show that information sharing by the controllers, and not by the plants, is sufficient for optimality, and schemes confirming this efficiency are constructed. Finally, the linear quadratic regulator (LQR) problem for systems with digital channels (and hence subject to bandwidth constraints) is solved and the structure of the optimal quantizer for this class of problems is determined.

ACKNOWLEDGMENTS

I have long believed that it is more important to raise the right questions than to find the solutions to the questions, whatever the context or the science is. Now, attempting to write down some of the research results we have been working on for the last two years, I would want to acknowledge every person, every book, and every scene in my past and present that has motivated me to be a researcher, had I had enough time and space.

I am grateful to my adviser Professor Tamer Başar, for the flexibility he offered and support he provided, in addition to his scientific wisdom and expertise, and for trying make me understand how important patience is not only for the solution of a particular problem, but for the decision making process on what to devote the research career on.

I am thankful to Dr. Kivanc Mihcak for the valuable, stimulating discussions on quantization theory. I also owe a lot to our research group, in particular to Akshay, Cagri, Tansu, and Hideaki, for the stimulating and friendly environment.

A most heartful thanks is to my family, my father, mother, and brothers: Muzaffer, Melih, Nevzat, and my dear sister Serap, the support of all of whom I have always had.

I am finally thankful to my friends here and in Turkey, for their close and motivating interest during the progress of the thesis.

TABLE OF CONTENTS

LIST OF FIGURES

Figure **Page**

CHAPTER 1

INTRODUCTION

1.1 Background and Literature Overview

Classical control analysis assumes perfect information share between the system plant and the decision maker/controller, with no delay, no precision effects, and no communication constraints. However, particularly in recent applications such as remote control over wireless systems, control over the Internet, control in decentralized and distributed systems and digital control of automatic systems, a joint analysis of communications, digital signal processing and control is necessary to design the desired system with optimal performance under a specified criterion. Presence of delay in the information or feedback channel might lead to severe degradation in system stability, whereas discretization of the space might force one to target limited form of stability and controllability, and the presence of noise in the channel might lead to loss of data and necessitate the construction of robust schemes.

Thus, the general problem and the underlying issues are multifold: how to utilize available partial information about the state as received by the controller; how to generate the partial information (i.e., how to quantize, encode and decode the state information); and finally, how to use this information to generate a satisfactory control meeting a given criterion, which is also to be drawn from a finite set of decisions. The problem might be broken into two stages (the state estimation problem and the decision making problem), which are however generally coupled in the case of communication constrained control problems, in which case a joint analysis becomes inevitable to achieve the best performance. In general, the parameters in the system design might have adverse effects on each other. For instance, as the bit rate used for communication increases, the complexity in the decoder might decrease, and the optimality in the decision generated by the controller might improve, whereas the delay in the state dynamics might require the designer to be cautious, in addition to the cost due to the higher bit rate carried over the communication channel. Signal processing theory and communications theory enter the picture for the quantization, coding, and decoding aspects; however, it is control theory that motivates the underlying problem and the objectives and specifications which have not been in the repertoire of classical communications and signal processing problems so far. For instance, one cannot use block coding in a control system, unless an asymptotic analysis is being performed, as it leads to delay at unacceptable levels.

Before reviewing the literature, we would like to note that some of the work in this thesis has been partly included in [1],[2],[3].

There has been significantly increasing interest on this set of problems recently, where various different models for the systems and the channels have been taken into consideration. We now describe below some of the work that is most relevant to our work.

One of the first papers on the general topic of this thesis [4] has analyzed the effects of quantization. The author has argued that *quantization* is more than an instantaneous approximation of a continuous state, and that a white noise approximation of the error due to quantization is not generally correct. Yet another paper [5] has studied the effects of communication constraints on a continuous-time system in an information-theoretic context, where two types of recursive coder-estimator sequences have been introduced, and stabilization and convergence requirements for the sequences have been established. The companion paper to [5], [6] introduces the notion of containability as a new interpretation for limited stability for continuous-time systems, develops necessary and sufficient conditions in this context, and studies the data rate, delay and code length as factors affecting containability. In [7], an linear quadratic Gaussian (LQG) system with an analog noisy channel between the plant and controller has been considered. Under appropriate definitions for the information available at the decoder and the encoder and for a given power constraint, a bound for the required bit rate for stability is obtained. In [8], it is shown that for a system with an uncertain initial state, where the cost is the mth moment of the state, the necessary and sufficient condition for exponential stability is that the rate should be higher than the sum of the logarithms of the ratios between the unstable open-loop eigenvalues and the desired exponential factor. In [9], the quantization problem in a different model, where the sensor measurements and the plant are connected via a channel, has been analyzed. In [10], conditions on the encoder and the norm of the transition matrix are obtained to achieve global asymptotic stabilization. There exist information theoretic approaches to the problem as well, which is more in line with the spirit of this thesis. Sahai [11] shows that the classical Shannon capacity is not the best measure to analyze capacity in systems with feedback. In the paper, Sahai introduces a different notion of capacity where the probability of error decays exponentially, where the code length is a variable parameter that can be used to decrease the probability of error.

This thesis develops information theoretic, quantization theoretic, and coding and control theoretic approaches for different classes of systems with a variety of channels. As the title suggests, the main intention is to achieve a degree of stability either in state estimation or at the actual state at the plant.

We will consider the following system in its most general form:

$$x_{t+1} = Ax_t + Bu_t + Dw_t, \tag{1.1}$$

where x and u are the state and control variables of dimensions n and k, respectively; A is an $n \times n$ matrix; B is an $n \times k$ matrix; D is an $n \times m$ matrix; and w_t is an $m-$dimensional noise.

Throughout the first few chapters of the thesis, we will consider systems of the form as shown in Figure 1.1. The state information will be sent over the channel

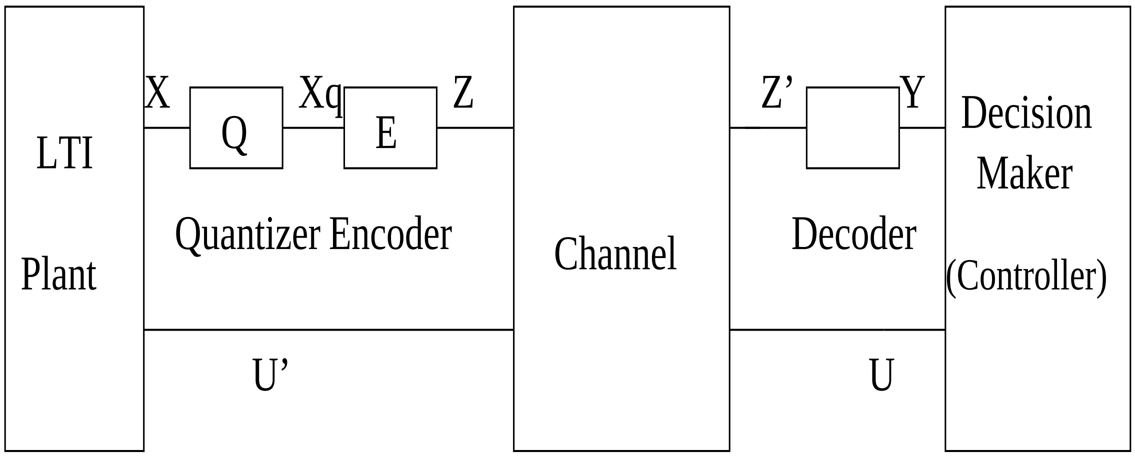

Figure 1.1 Control and state estimation over a channel.

(after being quantized and encoded) from the plant to the controller. The controller, after having decoded the message and estimated x will determine the control input and send it back to the plant, again over the channel, thus closing the loop.

We will start off with the analysis of scalar systems and then generalize the results to higher dimensional systems. In Chapter 2 we analyze the state estimation problem for a digital noiseless channel. In Chapter 3 we investigate stochastic linear time-invariant (LTI) systems over digital noiseless channels. In Chapter 4 we consider the scenario where the channel between the controller and the plant is a communication network, such as the Internet. Subsequently, in Chapter 5 we extend the analysis to higher dimensional systems, before considering decentralized LTI systems in Chapter 6. In Chapter 7 the linear quadratic regulator (LQR) problem for LTI systems with digital noiseless channels is introduced, and the optimal quantizer is presented along with a discussion on dynamic system analysis. Finally, in Chapter 8 we summarize the main results of the thesis and identify directions for future research.

CHAPTER 2

QUANTIZATION AND CODING FOR DETERMINISTIC LTI SYSTEMS WITH NOISELESS CHANNELS

2.1 Introduction

In this chapter we analyze and construct quantization schemes under various cost criteria and carry out performance analyses of recursive and forward-looking schemes for deterministic (noiseless) systems, where the plant and the controller are connected via a digital noiseless channel. The criteria include having a nonincreasing sequence of estimation error variances, mean square stability in the Lyapunov sense, and exponential stability of the estimation error variance. Asymptotic mean square stability has been studied before in [8] using asymptotic quantization theory. In the approach here, however, information and rate-distortion theoretic arguments are used to determine the requirements for mean square stability in the sense of Lyapunov and for exponential stability.

The class of systems considered in this chapter is described by

$$x_{t+1} = Ax_t + Bu_t + Dw_t\,,\; t = 0, 1, \ldots, \tag{2.1}$$

where x_0 is random with a known distribution; $w_t\,,\; t \geq 0$, is a stage-independent zero-mean driving noise with a known distribution; and the control has access to only a quantized version of the state x. The state is quantized (source-coded) and then channel-coded before being made available to the control u, at which site the quantized state information is recovered by a decoder.

The chapter is organized as follows: First a brief review of the topic of quantization is provided to the extent of its relevance to the development in this chapter. Then quantization schemes for a deterministic scalar system (i.e., (2.1) with $D = 0$ and $\dim(x) = 1$) are investigated, and subsequently they are extended to higher dimensions. Finally, the chapter ends with a few concluding remarks.

2.2 Quantization

A quantizer $Q(x)$ is a function that maps a large, possibly infinite, set (where a variable takes values) into a smaller finite set (where the quantized values lie), and is characterized by a set of thresholds or bin edges that partition the input space into quantization bins and a set of reconstruction values as the representative values of the corresponding bins (see, for example, [12], [13], and [14]). We briefly discuss below three main types of quantizers, which will be used in the development of this chapter.

2.2.1 Different types of quantization for random variables

2.2.1.1 Uniform quantizer

Among the different types of quantizers, the *uniform quantizer* is the most popular one, because of the ease with which it can be designed and implemented, and because of its asymptotic optimality under the mean square distortion measure [12]. In uniform quantization, the support of the probability density function is partitioned evenly (see Figure 2.1), and the reconstruction values are taken to be the mid-points between consecutive bin edges. Although uniform quantizers are mostly applied to sources with densities of finite support, they are also used for infinite sources, with the least likely outermost bins in this case regarded as overload regions and analyzed separately.

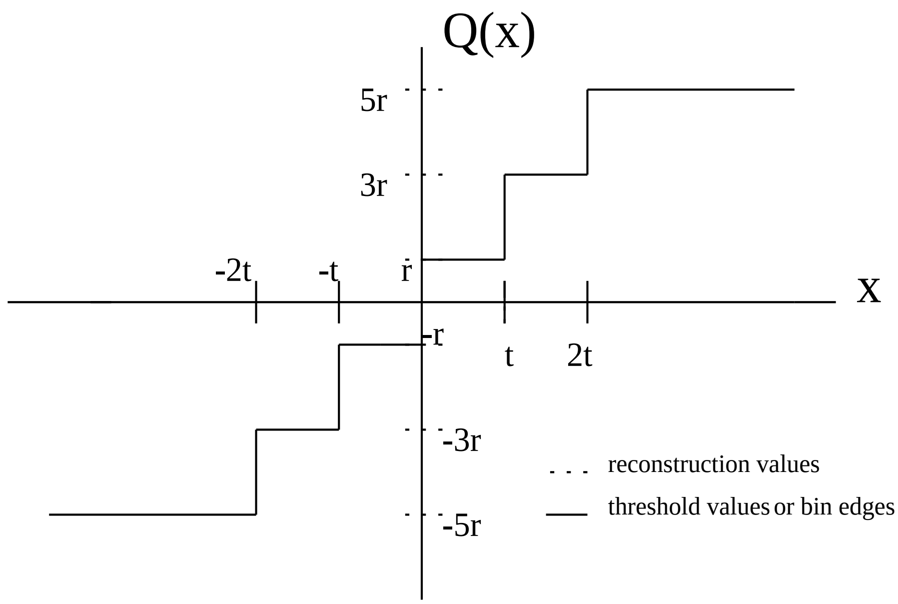

Figure 2.1 Uniform quantizer.

2.2.1.2 Logarithmic quantizer

In the *logarithmic quantizer*, the support set of the density of a random variable is partitioned logarithmically; in other words, lengths of the partitioning intervals increase exponentially with the quantization parameter. This way, the ratio between the quantized input and the quantization bin length remains smooth throughout the input space. The logarithmic quantizer has been proven to be optimal for stabilization of systems with quantized signals [15].

2.2.1.3 Lloyd-Max quantizer

The third class of quantizers we will discuss is the *Lloyd-Max quantizer*, which features an important optimality property. Let x be a random variable, $q = Q(x)$ be a particular quantization, and $\hat{x}(q)$ be the least squares estimate for x using the information $Q(x)$, that is an $\hat{x}$ that minimizes the quadratic distortion measure $E[(x-\hat{x})^2]$. Let $\mathcal{Q}_{\mathcal{K}}$ be the class of all quantizers with a fixed finite number K of quantization levels. Then, a problem of interest is the minimization of $E[\big(x - \hat{x}(Q(k))\big)^2]$ over all $Q \in \mathcal{Q}_{\mathcal{K}}$. An algorithm to solve this optimization problem was proposed independently by Lloyd [16] and Max [17], and is based on a Gauss-Seidel-type recursive optimization of the values of the quantizer thresholds and quantizer reconstruction values, which is the following: Given the reconstruction values, the thresholds are determined as the midpoints between the reconstruction values, and given the threshold values, the reconstruction values are obtained as the centroids of the intervals between the corresponding reconstruction levels. Hence, at each step of the iteration, either the threshold values or the reconstruction values are updated according to these rules. The algorithm converges at least to a local minimum, and if the logarithm of the probability density function (pdf) of the random variable x is concave, then the algorithm is known to be globally convergent. If the pdf is not concave, then one can still achieve global convergence if the initial conditions are chosen carefully.

The Lloyd-Max quantizer has the following appealing properties:

1. The error is orthogonal to the quantizer output, so the input variance can be expressed as the sum of variances of the quantization error and the quantizer output.

2. The quantization error is zero mean.

3. The variance contributions of each bin to the error variance are equal.

A quantizer is said to be a *high-rate quantizer* if the number of levels is very high, so that the quantization error at each bin is approximately uniform [12]. The minimum quantization error variance,

$$\sigma_K^2 := \min_{Q \in \mathcal{Q}_K} E\left[(x - \hat{x}(q))^2\right],$$

is given, for high rates, by the Panter-Dite, or the Aigrain formula [12]:

$$\sigma_K^2 \approx \frac{1}{12}\frac{1}{K^2}\Big(\int (f_x(x))^{\frac{1}{3}} dx\Big)^3, \tag{2.2}$$

where K is the number of quantization levels. If the input to the quantizer has a uniform distribution, then the Lloyd-Max quantizer becomes the uniform quantizer. In this case, the Panter-Dite formula is exact, and it becomes:

$$\sigma_K^2 = \frac{1}{12}\frac{\Delta^2}{K^2}, \tag{2.3}$$

where Δ is the support of the uniform quantization input.

2.2.2 Quantization for random sequences: causality

When the quantity to be quantized is a sequence of (correlated) random variables rather than a single random variable, the time variable also enters the picture, and the dynamic information upon which the quantization is allowed to depend at each point of time becomes an important factor in the performance of the quantizer. Let $X := \{x_t,\ t = 0, \pm 1, \pm 2, \ldots\}$ be a doubly infinite sequence of random variables, with $X_{-\infty}^t$ denoting its truncation to the semi-infinite interval $(-\infty, t]$. Then we say that a quantizer $Q := \{q_t = Q_t(X),\ t = 0, \pm 1, \pm 2, \ldots\}$ is *causal*, if it depends only on the past and present values of the random sequence, i.e., $q_t = Q_t(X_{-\infty}^t)\ \ \forall t$.

When different elements of the sequence are independent, it has been shown that particularly for a stationary, independent, and identically distributed random sequence, optimal causal coding is identical to memoryless coding where only the current source outcome is used for coding [18]. For sources with memory, however, feedback from the receiver is useful [19], and if the source is kth order Markov, then the optimal causal coding uses only the last k source symbols and the current state at the receiver's memory [20]. In [21], it is shown that the results of [18] are applicable to schemes with side information for the case where the side information is available to both the receiver and the transmitter, and the causal side information, that is only available to the encoder is useless. We will find an occasion to use these results in our analysis in this chapter.

2.2.3 Time-invariance of the quantizer

We introduce here two notions of time invariance, which will be used in the chapter.

Definition 2.1 *A time-invariant quantizer is one that can change only up to a scaling constant. A strictly time-invariant quantizer is one that does not change over time at all.*

2.3 Optimal Quantization

We now start off with the scalar deterministic LTI system:

$$x_{t+1} = ax_t + bu_t \,, \quad t = 0, 1, \ldots$$

where the initial state x_0 is a random variable with a given pdf. 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 a quantization scheme under some prespecified constraints and according to a particular optimality criterion. We will assume that the control function is available at both the receiver and the transmitter, which is a legitimate assumption since the channel is noiseless and the control signal is transmitted error-free; thus, the control function will not have any effect on the evolution of the uncertainty in the initial state. Hence, in essence, we can work without any loss of generality with the control-free system:

$$x_{t+1} = ax_t \,, \quad t = 0, 1, \ldots \tag{2.4}$$

Note however that the system is still a control system, and, for instance, one cannot use block quantization where outcomes belonging to different time stages are quantized jointly, since there is a recursive evolution at the receiver and the transmitter.

One can regard the state in the above equation as the uncertainty at a particular time.

We will assume, unless otherwise stated, that the initial state x_0 is uniformly distributed, with finite support. Note that under this assumption x_t is also uniformly distributed for each t. Our objective is to design optimal dynamic quantizers, $Q := \{q_t = Q_t(X),\ t = 0, 1, 2, \ldots\}$ under one of the following two criteria, where in both cases $\hat{x}$ denotes the least squares estimate (equivalently, the conditional mean) of x, and we use throughout the notation already introduced for quantizers. The class of admissible quantizers, $\mathcal{Q}$, to be used in each case will be specified later in individual subsections.

Criterion 1: *Monotonic boundedness of variance with recursive quantization.*

Find a recursive Q under which the estimation error variance is a nonincreasing sequence, and thus will be bounded, that is, $\forall\, t = 0, 1, \ldots,$

$$E\big[\big(x_{t+1} - \hat{x}_{t+1}(q_{t+1})\big)^2\big] \leq E\big[\big(x_t - \hat{x}_t(q_t)\big)^2\big] .$$

For this criterion, we will investigate the existence of strictly time-invariant quantizers that achieve a bounded variance and time-invariant quantizers that will achieve exponential stability.

Criterion 2: *Forward-looking stabilizing quantization.*

Find a forward-looking quantizer Q_t that minimizes for each given t the estimation error variance

$$E\big[\,\big(x_t - \hat{x}_t(q)\big)^2\,\big]\,,$$

and keeps it bounded as t goes to ∞.

We use here the terminology, *forward-looking*, since the objective is to achieve a minimum distortion at a particular future time t.

Remarks. A problem similar to the second one above has been considered previously in [8], where asymptotic quantization theory has been used to find the requirements for asymptotic stability. Here we use a different, information-theoretic approach to find the requirements for stability in the sense of Lyapunov, which leads to a different, less stringent condition.

Definition 2.2 *Communication rate is the amount of data, in term of bits, that is transferred from one party to another over a communication channel.*

We assume that the pdf of the initial state x_0 has finite support, is uniform, and is zero mean. Under this assumption, we first state a basic fact (after introducing some notions from information theory, which will be used throughout the chapter), and then obtain a result on the minimum rate required among all quantization schemes that lead to satisfaction of the given criterion.

Mutual information is a quantity that captures the amount of information one random variable, the output of the channel, carries about a second random variable, the input to the channel, and is expressed as the entropy of the input conditioned on the output subtracted from the entropy of the input:

$$I(X,Y) = H(X) - H(X|Y). \tag{2.5}$$

The minimum achievable rate of data as a function of the average distortion D is called the *source-distortion function* or the *rate-distortion function*, and is denoted by $B(D)$ or $R(D)$ [22]. We will use the latter notation throughout the chapter. The rate distortion function is the minimum mutual information satisfying a given distortion constraint between the input and the output.

Quantization and channel estimation problems are intrinsically different problems and generally entail different optimality conditions. However, in the case of a digital error-free, or noiseless, channel, the quantization output is identical to the value that the receiver would estimate at the receiver. Therefore $\hat{x}$ will be used as both the quantization output and the receiver estimation output, and we hope that this common use of notation wil not lead to any confusion.

2.3.1 Criterion 1

Fact 2.1 *For any random variable X, and scalar a,*

$$H(aX) = \log_2(|a|) + H(X). \tag{2.6}$$

Theorem 2.1 *For the scalar LTI system (2.4) with a uniform distribution for the initial state, for any recursive strictly-time-invariant quantizer to satisfy criterion 1, the required rate is* $\max\big(0, \log_2(|a|)\big)$ *per stage.*

Proof. **Necessity.** Regarding the past information as the side information, since the actual past outcomes are not available at the receiver, the only useful information to enhance the system performance is the past quantized states since they are available at both the encoder and the decoder.

For any stage i, the rate is lower bounded by the minimum mutual information between the state and its quantized value:

$$\begin{aligned} R &\geq \min_{Q\in\mathcal{Q}} I(\hat{x}_i|\hat{x}_0^{i-1}; x_i|\hat{x}_0^{i-1}) \\ &= H(x_i|\hat{x}_0^{i-1}) - H(x_i|\hat{x}_i, \hat{x}_0^{i-1}), \end{aligned} \tag{2.7}$$

where $\hat{x}_0^{i-1}$ denotes the past information of the quantizer outputs, which are available both at the transmitter and the receiver. Since the system under consideration is Markov, the entire information that can be inferred from the past is captured by the last outcome. Thus, the above becomes

$$\begin{aligned} R &\geq \min_{Q\in\mathcal{Q}} I(\hat{x}_i|\hat{x}_{i-1}; x_i|\hat{x}_{i-1}) \\ &= H(x_i|\hat{x}_{i-1}) - H(x_i|\hat{x}_i, \hat{x}_{i-1}). \end{aligned} \tag{2.8}$$

Hence,

$$H(x_i - \hat{x}_i|\hat{x}_i) \geq H(x_i|\hat{x}_{i-1}) - R,$$

and since $x_i = ax_{i-1}$,

$$H(x_i - \hat{x}_i|\hat{x}_i) \geq \log_2(|a|) + H(x_{i-1} - \hat{x}_{i-1}|\hat{x}_{i-1}) - R\,.$$

Therefore, we need to have at least an additional $\log_2(|a|)$ bits to make the entropy nonincreasing. Since the entropy changes monotonically with the variance for a given distribution (uniform in this case) the variance is nonincreasing.

Sufficiency and construction. The error in the estimation of the state will be quantized and fed into the channel, as illustrated in Figure 2.2. This is done to minimize the range of the quantizer around the origin. The scalar system equation is

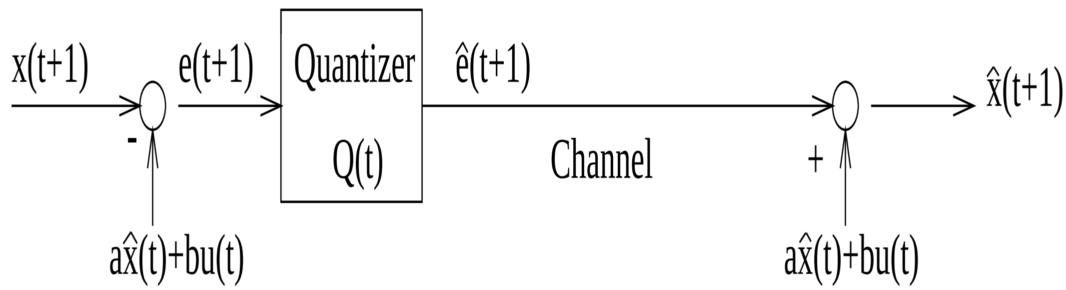

Figure 2.2 Coding and quantization scheme.

$$x_{t+1} = ax_t + bu_t. \tag{2.9}$$

Define the error e as

$$e_{t+1} := x_{t+1} - a\hat{x}_t - bu_t\,, \tag{2.10}$$

and at the receiver we have

$$\hat{x}_{t+1} = a\hat{x}_t + bu_t + \hat{e}_{t+1}. \tag{2.11}$$

From (2.9) and (2.10), we get

$$e_{t+1} = a(x_t - \hat{x}_t). \tag{2.12}$$

From (2.10) and (2.11), with t shifted,

$$x_t - \hat{x}_t = e_t - \hat{e}_t. \tag{2.13}$$

By virtue of the last equation, we reduce the error on the state to the error in the estimation. The difference between the actual state and the quantized state is equal to the difference between the actual error and the quantizer error. From (2.12) and (2.13) the following recursion follows:

$$e_{t+1} = a(e_t - \hat{e}_t). \tag{2.14}$$

Thus, the error introduced by the quantization becomes the signal to be quantized at the next stage. Hence it is the quantizer that will determine the power of the uncertainty in the system, which will propagate recursively. To ensure that the estimation error is a nonincreasing sequence, it suffices to enforce $E[e_t^2] \leq E[e_{t-1}^2]$, which corresponds to

$$\frac{E[e_t^2]}{E[(e_t - \hat{e}_t)^2]} \geq a^2. \tag{2.15}$$

The proof of the theorem is now completed in view of the following result.

Lemma 2.1 *In an LTI system with a uniformly distributed initial state, for uniform quantization to ensure that the estimation error variance is a nonincreasing sequence with respect to time, the rate should be at least* $\max(0, \log_2(|a|))$ *bits per sample.*

Proof. Assume that the current support of the uniform pdf is $[-\Delta, \Delta]$. If we use a K-level quantizer, the variance in the quantization error will be $(1/3)(\Delta/K)^2$, and the variance of the error at the next stage will be $(a^2/3)(\Delta/K)^2$, and from (2.14) we want this to be smaller than $\Delta^2/3$, which means that we should have $a \leq K$. We therefore need a quantizer with the number of levels at least as large as $|a|$, and to achieve a given K number of levels, we know that we need $R = \log_2(K)$ (since the symbols will be uniformly distributed, the bit rate is identical to the entropy of the random variable). Hence, $R \geq \max(0, \log_2(|a|))$. ◇

Corollary 2.1 *Given a uniformly distributed input, for any recursive time-invariant quantizer to achieve exponential stability of the state estimation error variance, the rate has to be strictly greater than* $\log_2 |a|$. *This is achievable by any* K *level uniform quantizer with* $K > |a|$.

Proof. The uncertainty and hence the support of the error decreases over time. At each stage, the support of the quantizer will be decreased by

$$\Delta_{t+1} = \frac{|a|}{K}\Delta_t \,, \tag{2.16}$$

which will converge to zero exponentially, with $|a|/K$, in case $K > |a|$. In this scheme, the quantizer values will also be scaled by the ratio $|a|/K$. ◇

2.3.1.1 Simulation results

Consider the system

$$x_{t+1} = 3x_t, \tag{2.17}$$

where the initial state is a uniformly distributed random variable with support $[-2, 2]$. In case the data rate is 2 bits per symbol, from Figure 2.3 it can be seen that the estimation error remains finite. However, the case where $R = 1$ bit per stage causes

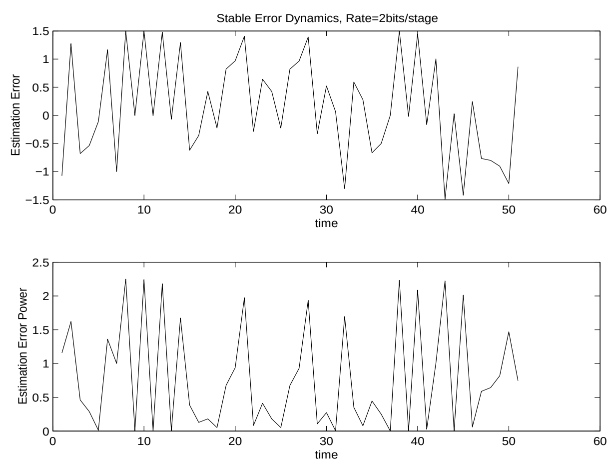

Figure 2.3 Stable error dynamics.

unstable error dynamics as shown in Figure 2.4.

2.3.2 Distortion constrained entropy minimization

In this section we present a time-invariant quantization scheme for a uniformly distributed source input that will satisfy a given constraint on the quantization error

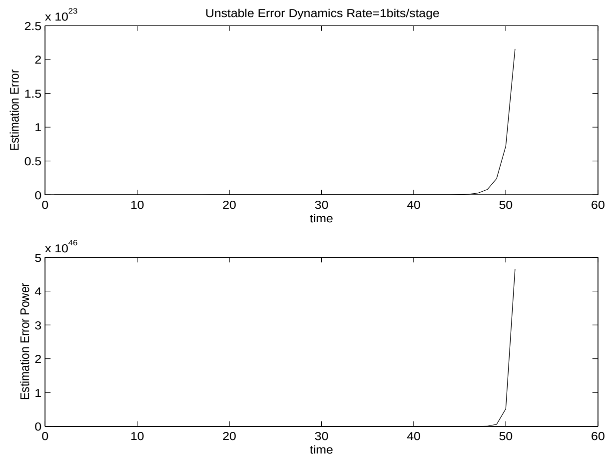

Figure 2.4 Unstable error dynamics.

while minimizing the output entropy. Although the analysis is coupled with the analysis in the previous subsection, here we consider specifically the case where the parameter a is not an integer, and the intention is to satisfy a minimal bit rate with a constraint on the quantization error.

It is well known that the data rate required, or the average code length needed to send a random output over a channel is lower-bounded by the entropy rate of the input random variable [22]. In fact there are effective coding schemes, like Huffman and arithmetic coding, which generate codewords whose average code length gets very close to the entropy rate. Shannon has shown that the bit rate required is lower-bounded by the entropy of the quantized signal and can be upper-bounded by the entropy plus one. Huffman coding [22] is an efficient algorithm where the average code length is very close to the entropy rate. Therefore entropy is indeed a very efficient measure for the average code length, and for optimization one can use the entropy instead of the bit rate. Generally one is interested in minimizing both the entropy and the distortion; however, since entropy and distortion are inversely proportional, one could minimize one under a constraint on the other. Since we assume the quantization error to be uniform, one meaningful problem here is to find the minimum achievable entropy of the quantization output under the distortion constraint $E[(x-\hat{x})^2] \leq D$.

This constraint is a convex function of the quantization outputs, and the entropy $H(\hat{x}_i|\hat{x}_0^{i-1})$ is concave; hence, the constraint inequality should end up with an equality for the optimal case. To solve this optimization problem, we introduce the Lagrangian:

$$\begin{aligned} J &= -\sum_i \int_{\delta_i}^{\delta_{i+1}} f(x)\,dx \,\log\left(\int_{\delta_i}^{\delta_{i+1}} f(x)\,dx\right) \\ &\quad + \sum_i \int_{\delta_i}^{\delta_{i+1}} \left(\lambda_1 (x-q_i)^2 + \lambda_2\right) f(x)dx\,, \end{aligned} \tag{2.18}$$

where λ_1 and λ_2 are the Lagrange multipliers associated with the distortion constraint and the fact that f is a legitimate pdf. The entropy is just a function of the quantizer threshold values, and not a function of the quantizer, 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 property. Given the quantizer values, the threshold values will be different from the ones found by the Lloyd-Max analysis. If we restrict ourselves to a uniform input, with support set Δ, and using the centroid property for the quantization reconstruction values, J reduces to

$$\begin{aligned} J &= -\sum_i \frac{(\delta_{i+1}-\delta_i)}{\Delta} \log\left(\frac{\delta_{i+1}-\delta_i}{\Delta}\right) \\ &\quad + \lambda_1 \sum_i \frac{(\delta_{i+1}-\delta_i)^3}{12\Delta} + \lambda_2 \frac{\delta_K-\delta_0}{\Delta}. \end{aligned} \tag{2.19}$$

The distortion constraint D will be the distortion that will guarantee stability in the sense of boundedness of the estimation error variance. Referring to the analysis in the previous subsection, the target distortion will be $\Delta^2/12a^2$. Taking the derivative of J with respect to δ_i, except at $i = K$ and $i = 0$, we get

$$\frac{1}{\Delta}\log[\frac{(\delta_i-\delta_{i-1}}{(\delta_{i+1}-\delta_i)}] + \lambda_1 \frac{(\delta_i-\delta_{i-1})^2-(\delta_{i+1}-\delta_i)^2}{4},$$

which we set equal to zero. We know that $\delta_0 = 0$ and $\delta_K = \Delta$. Hence, we can propose a scheme for encoding, assuming all the random variables to be quantized at each stage are uniform.

1. Find the optimal distortion constrained quantizer Q^* satisfying the distortion constraint.

2. For the next stages, just scale the quantizer with the support of the corresponding quantizer bin, based on the data sent or the data received.

3. The quantizer is time-invariant.

This will ensure stability with minimum bit rate, provided Huffman coding or some other type of entropy coding is employed.

As the number of quantization levels increases, one might expect to obtain more improved results, since the class of quantizers with fewer number of levels might be regarded as a subset of the class of quantizers with a higher number of levels. This, however, turns out not to be the case, since the entropy is a concave function, having significant contributions from low probability components. By setting $\delta_{i+1} - \delta_i = p_i \Delta$, and redefining the Lagrangian multipliers as λ_1' and λ_2' to absorb scaling factors, J becomes

$$J = -\sum p_i \log(p_i) + \lambda_1' \sum p_i^3 + \lambda_2' \sum p_i. \tag{2.20}$$

The constraints associated with this Lagrangian are

$$\begin{aligned} \sum p_i^3 &\leq (1/a^2), \\ \sum p_i &= 1, \\ p_i &\geq 0, \forall i. \end{aligned} \tag{2.21}$$

The first constraint is the cost in distortion (expressed in terms of probabilities) and the remaining two are the requirements for being a legitimate probability mass function. The constraints in (2.21) are independent of the support, which is why the quantizer is time-invariant. Taking the derivative of J above with respect to p_i, we obtain

$$-\log_2(p_i) - \log_2(e) + 3\lambda_1' p_i^2 + \lambda_2' = 0, \tag{2.22}$$

which is a quadratic equation and hence admits at most two solutions for each λ_1' and λ_2'.

Proposition 2.1 *For the linear system $x_{t+1} = ax_t$, the minimum rate required of a quantizer satisfying a time-invariant distortion level is greater than or equal to $max(0, \log 2(|a|))$. This rate is achievable if a is an integer and a $K = |a|$ level uniform quantizer is used.*

Proof. The entropy, $-E_p[\log_2(p)] = -\sum p_i \log_2(p_i)$, is a concave function, and hence, using Jensen's inequality,

$$-E_p[\log_2(p)] \geq -\log_2(E_p[p]) = -\log_2\big(\sum_i p_i^2\big), \tag{2.23}$$

where equality is achieved by a uniform distribution. Using Jensen's inequality again, this time for $E_p(p)$, we have

$$-(\sum p_i^2)^2 \geq -\sum p_i^3\,, \tag{2.24}$$

which is equivalent to

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

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

$$\begin{aligned} H(p) &= -E_p[\log_2(p)] \geq \frac{1}{2}\log_2(\sum p_i^3) \\ &= -\frac{1}{2}\log_2(\frac{1}{a^2}) = \log_2(|a|). \end{aligned} \tag{2.25}$$

◇

Remarks. The inequality in (2.25) becomes an equality only when the probability distribution is uniform, and is equal to $(1/|a|)$. Hence, if $K = |a|$, the uniform quantizer will minimize the entropy. Surprisingly, an increase in K will not decrease the entropy, since the inequality will not turn into an equality.

Proposition 2.2 *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 of p_1 each, and the remaining one occurs with probability $p_2 \leq p_1$, such that*

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

Proof. Since the distortion constraint is convex, and the entropy is concave, the solution satisfies the constraint with equality. We also know from (2.22) that the number of different probabilities cannot exceed two. In [23] an entropy-constrained quantizer was designed, and an analysis was performed for a similar set of functions, where the optimal numbers of allocations for the two different probabilities were obtained through a rigorous analysis of the convexity properties of the functions and their derivatives with respect to the numbers of allocations. As a result, the allocations were found to be $K-1$ times the larger bin, and a single usage of the smaller bin. Since the entropy-constrained distortion minimizing quantization and distortion-constrained entropy minimizing quantization are duals of each other, a similar result holds in this case. $\diamond$

Proposition 2.3 *The quantizer designed above achieves boundedness of the mean-square error variance.*

Proof. Referring to (2.15), the constraint guarantees that the mean-square error variance is nonincreasing at each stage, which readily leads to boundedness. $\diamond$

Remarks. For some relevant results, we refer the reader to [24], which conducts a joint analysis of entropy and distortion for some non-Gaussian sources, including uniform sources. $\diamond$

The quantizer of the two propositions above will not generally be unique, since the bin probabilities cannot characterize the quantizer values uniquely. However, up to a permutation term of $K = \lceil |a| \rceil$ there will be a unique probability distribution. The rate required will be less than $\log_2(\lceil |a| \rceil)$.

Another approach depending on our flexibility in the design would be the following: For the case where a is not an integer, at time n, the uncertainty in the state will be $a^n x_0$. If we sample the system such that the system coefficient becomes closer to an integer, then uniform quantization could be applied without much loss of optimality. For instance if $a = 3.17$, instead of taking samples with duration T, taking them with a period of $3T$, since $(3.17)^3 = 31.86$, a uniform $\log_2(32) = 5$-level quantizer might be used as a more practical one.

2.3.3 On the uniformness of the initial state

In the above analysis, the sole element of randomness was the initial state x_0, and the distribution was assumed to be uniform. Now we will justify this assumption. We state the following lemma.

Lemma 2.2 *Let x_0 be the realization of a random variable X_0 with a continuous pdf $f_0(.)$ and with finite support $[-|a|\Delta, |a|\Delta]$, and suppose that at each time step, $x_{t+1} = ax_t$ is quantized successively using an $|a|$ level uniform quantizer, where $|a|$ is a nonzero integer. Then, the Kullback-Leibler distance between the uniform error density and conditional quantization error density (the quantization error for a specific bin) converges to zero as $t \to \infty$.*

Proof. The Kullback-Leibler distance between two pdf's $g(x)$ and $h(x)$ is defined as

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

Since $X_{t+1} = aX_t$, the pdf $f_{t+1}(y)$ of X_{t+1}, in terms of $f_t(x)$ of X_t is

$$\frac{1}{|a|} f_t\left(\frac{y}{a}\right).$$

Assume that a uniform quantizer with a spacing Δ is used at each stage. After n steps, the conditional quantization error probability density in the kth bin becomes

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

Thus, the Kullback-Leibler distance between the uniform pdf and the quantization error pdf is

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

and by a change of variables, this becomes

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

By the continuity of f_0, and using the mean-value theorem, this expression is equivalent to

$$-\int_{k\frac{\Delta}{|a^n|}}^{(k+1)\frac{\Delta}{|a^n|}} \frac{|a^n|}{\Delta} \log\left(\frac{f_0(y)}{f_0(\bar{z}_n)}\right) dy$$

for some $\bar{z}_n \in \left(\frac{k\Delta}{|a^n|}, \frac{(k+1)\Delta}{|a^n|}\right)$.

Now, again by the continuity of f_0, given any $\epsilon > 0$, $\exists\, N$ such that $\forall\, n > N$,

$$1 - \epsilon < \frac{f_0(y)}{f_0(\bar{z}_n)} < 1 + \epsilon\,.$$

Therefore, for $n > N$, D is bounded from above by

$$-\int_{k\frac{\Delta}{|a^n|}}^{(k+1)\frac{\Delta}{|a^n|}} \frac{|a^n|}{\Delta} \log(1-\epsilon) = -\log(1-\epsilon)\,,$$

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

It thus follows from the proof above that the probability density can be regarded as effectively being uniform after a sufficient number of time steps (and this number could in fact be only moderately large in practice).

2.3.4 Forward-looking quantization for stability

Asymptotic stability of state estimation error variance has been investigated earlier in [8]. Here, however, we investigate stability in the Lyapunov sense, using a rate-distortion theoretic argument. This leads to minimum requirements, which are different from those for asymptotic stability.

Fact 2.2 *For a given finite variance random variable, the differential entropy is maximized by the Gaussian distribution, and is consequently finite.*

Theorem 2.2 *For a scalar linear system $x_{t+1} = ax_t$, where the initial state uncertainty has a uniform distribution and the type of quantization is uniform, the bit rate required for boundedness of the mean-square state estimation error variance is* $\max(0, \log_2(|a|))$. *This rate is achievable.*

Proof. **Necessity.** If $|a|$ is less than 1, since the system is stable, it is not necessary to send any data. Assume therefore, for the remainder of the proof, that $|a| \geq 1$. Consider now the following information theoretic analysis: The entropy at time step n is

$$H(a^n x_0) = H(x_0) + n \log_2(|a|) \tag{2.26}$$

which can be evaluated using the definition of the differential entropy. Rate distortion function is the minimum mutual information under a given level of distortion, that is

$$R(D) = \min_{p(x):E[(X-\hat{X})^2] \leq D} I(X, \hat{X}), \tag{2.27}$$

where I denotes the mutual information [22]. Now,

$$\begin{aligned} I(x_n, \hat{x}_n) &= H(x_n) - H(x_n | \hat{x}_n) \\ &= H(x_n) - H(x_n - \hat{x}_n | \hat{x}), \end{aligned} \tag{2.28}$$

and since $H(x_n - \hat{x}_n | \hat{x}) \leq H(x_n - \hat{x}_n)$,

$$\begin{aligned} I(x_n, \hat{x}_n) &\geq H(x_n) - H(x_n - \hat{x}_n) \\ &= H(x_0) + n \log_2(|a|) - f(D), \end{aligned} \tag{2.29}$$

where $f(D)$ is a finite number from Fact (2.2), since the distortion is intended to be finite. If we divide the expression by n, and let n approach ∞, this yields the rate required, which is $\log_2(|a|)$.

Sufficiency. The quantization gain for a quantizer is defined to be the ratio between the variances of the input and the noise. To achieve finiteness as n approaches ∞, a quantization gain of a^{2n} is sufficient, which means that the minimum number of levels should be a^n. Therefore $n \log_2(|a|)$ bits are needed for the stability of the state estimation error, which corresponds to a bit rate requirement of $\log_2(|a|)$ for each time stage. Thus, a uniform quantizer with a^n levels can be used, and coded successively. ◇

Remarks. Note that the quantizer in the scheme above is forward-looking. The codes sent at any stage i, $i < n$, depend on the source value at stage n. ◇

2.4 Summary

This chapter has considered the state estimation problem in a control system with random initial state, where the plant and the controller are connected via a digital noiseless channel. It has been shown that quantization theoretic results achieve the ultimate lower bound requirements dictated by information theory, and that the system parameters determine the required rate for stability in state estimation. A time-invariant, minimum entropy quantizer has been introduced, which further improves the performance for the cases where the system parameter a is not an integer.

CHAPTER 3

QUANTIZATION AND CODING FOR STOCHASTIC LTI SYSTEMS WITH NOISELESS CHANNELS

3.1 Introduction

In this chapter we consider a stochastic system described by

$$x_{t+1} = Ax_t + Bu_t + Dw_t, \tag{3.1}$$

where w_t is a noise process. Since we consider the scalar case, we take $D = B = 1$ without loss of generality and let $A = a$. As in the last chapter, we assume that the control function is available at both the transmitter and the receiver. The discussion here parallels that of the deterministic case, and begins with recursive quantization.

For stochastic systems, entropy power is a more appropriate measure of the uncertainty. Entropy power of a random variable is the mean-square power of the Gaussian random variable which has the same entropy as the random variable itself [22].

3.2 Recursive Quantization and a Quantizer-Filter

Lemma 3.1 *For two independent random variables X and Y, $H(X+Y) > H(X)$.*

Proof. Conditioning cannot increase entropy:

$$H(X+Y) > H(X+Y|Y).$$

The inequality above is strict since the random variables $X+Y$ and Y are not independent. Now, since X and Y are independent, $H(X+Y|Y) = H(X|Y) = H(X)$, and hence the result follows. $\diamond$

Thus the entropy of the source at time $t+1$ conditioned on the past quantized values increases by more than $\log_2(|a|)$ when compared to the conditional entropy at time t. Since $H(x_{t+1}|\hat{x}_0^t) = H(ax_t + w_t|\hat{x}_0^t)$ and

$$H(ax_t|\hat{x}_0^t) = \log_2(|a|) + H(x_t|\hat{x}_0^t),$$

we have

$$H(x_{t+1}|\hat{x}_0^t) \geq \log_2(|a|) + H(x_t|\hat{x}_0^t).$$

In view of this, we now have the following result.

Theorem 3.1 *In an unstable linear system, $x_{t+1} = ax_t + w_t$, $|a| > 1$, for any recursive scheme, where there is access only to the past information, to achieve a nonincreasing sequence for the entropy power of the state estimation, the minimal bit rate should be strictly larger than* $\log_2(|a|)$.

Proof. Since the source is Markov, first-order memory for the coded outputs is the only tool needed for optimal quantization and the actual state outputs are useless for performance improvement, since they are not available at the receiver. Using the first-order memory, the conditional entropy of the source given the last coded outcome is

$$H(x_{t+1}|\hat{x}_t). \tag{3.2}$$

Conditioning makes the term effectively memoryless, and thus, we can apply the memoryless quantization scheme without any loss of optimality.

Minimum rate is equal to the mutual information between the process and its quantized version, and thus any rate will be greater than or equal to this mutual information:

$$R \geq H(x_t|\hat{x}_{t-1}) - H(x_t|\hat{x}_t, \hat{x}_{t-1}). \tag{3.3}$$

Since

$$H(x_t|\hat{x}_{t-1}) = H(a(x_{t-1}|\hat{x}_{t-1}) + w_{t-1}),$$

we have

$$H(x_t|\hat{x}_t) \geq H(x_t|\hat{x}_{t-1}) - R > \log_2(|a|) + H(x_{t-1}|\hat{x}_{t-1}) - R, \tag{3.4}$$

which uses (3.1). Thus, the rate has to be greater than $\log_2(|a|)$ to have a nonincreasing state estimation entropy sequence, which leads to the desired conclusion. $\diamond$

3.2.1 A recursive quantizer-filter

Here we present a recursive filter that combines the quantizer constructed for the deterministic case with the classical Kalman filter. In the system under consideration, the system state x is the variable to be observed (denoted also by y, the observation variable). The quantized observations are received at the receiver, thus by regarding the quantization error as the noise, quantized observations are interpreted as the noisy observations in a standard Kalman filtering setup. Since a Kalman filter is the best linear filter, even for non-Gaussian and nonstationary processes for minimizing the signal to noise error ratio, it is used here as a practical filter, although it is evident that it is optimal only for the cases where the noise is Gaussian, which is not the case for quantization error noise. Thus, in the system, y_t is identical to x_t, and $\hat{y}_t$ is the quantized outcome of y_t (see Figure 3.1). The state equation is

$$x_{t+1} = ax_t + w_t, \tag{3.5}$$

and the Kalman estimator equation is

$$\hat{x}_{t+1} = a\hat{x}_t + l(\hat{y}_{t+1} - a\hat{x}_t), \tag{3.6}$$

for some constant gain l. Defining ξ_t as $x_t - \hat{x}_t$,

$$\xi_{t+1} = a\xi_t + w_t - l(\hat{y}_{t+1} - a\hat{x}_t). \tag{3.7}$$

And the term $\hat{y}_{t+1} - a\hat{x}_t$ can be expressed as

$$\hat{y}_{t+1} - a\hat{x}_t = y_{t+1} - ax_t + a(x_t - \hat{x}_t) + \hat{y}_{t+1} - y_{t+1}. \tag{3.8}$$

Defining the quantizer error as $q_{t+1} = y_{t+1} - \hat{y}_{t+1}$,

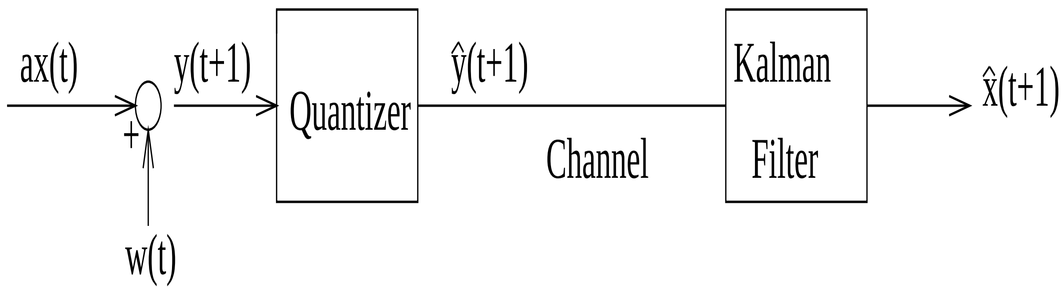

Figure 3.1 Coding and quantization with a Kalman filter.

$$\begin{aligned}\xi_{t+1} &= a\xi_t + w_t - l(w_t + a\xi_t - q_{t+1}) \\ &= a(1-l)\xi_t + (1-l)w_t + lq_{t+1}.\end{aligned} \tag{3.9}$$

We now consider the recursion for the quantization. As in the case of the deterministic system, a dithering function which is $a\hat{x}_t$ in this case, is subtracted from the actual state, and the resulting signal is quantized and sent through the channel. Thus,

$$y_{t+1} - \hat{y}_{t+1} = a\xi_t + w_t - [\widehat{a\xi_t + w_t}] = q_{t+1}, \tag{3.10}$$

where $[\widehat{a\xi_t + w_t}]$ is the quantization output of $a\xi_t + w_t$. The recursion for ξ becomes

$$\begin{aligned}\xi_{t+1} = a(1-l)\xi_t + w_t(1-l) + lq_{t+1} &= (1-l)(a\xi_t + w_t) \\ &+ l(a\xi_t + w_t - [\widehat{a\xi_t + w_t}]).\end{aligned} \tag{3.11}$$

We further have

$$E[\xi_{t+1}^2] = E[(1-l)^2(a\xi_t + w_t)^2 + l^2(q_{t+1}^2) + 2(1-l)l((a\xi_t + w_t)q_{t+1})]. \tag{3.12}$$

We know that, for a given input statistics, the distortion minimizing quantizer is the Lloyd-Max type quantizer, and this minimizes the correlation between the input and the quantization error

$$E[\xi_{t+1}^2] = (1-l)^2 E[(a\xi_t + w_t)^2] + (2l - l^2)E[(q_{t+1}^2)]. \tag{3.13}$$

The Kalman filter is related to the quantization error only through the variance of the error:

$$E[q_{t+1}^2] = \frac{a^2 E[\xi_t^2] + \sigma_w^2}{K^2} C_t^3, \tag{3.14}$$

where C_t is the quantizer gain coefficient, given by

$$C_t := \int (f_x(x))^{\frac{1}{3}} dx.$$

Equation (3.13) shows that Lloyd-Max is optimal for the one-stage estimation problem. However, most of the reasonably spaced quantizers, including the uniform quantizer, achieve a distortion that is inversely proportional to the square of the number of levels and can be used. To find the optimal Kalman gain coefficient, taking the derivative of (3.13) with respect to l and setting it equal to zero gives the minimizing l, since the function is convex. The unique value is

$$l^* = 1, \tag{3.15}$$

and using this value in (3.13) yields

$$E[q_{t+1}^2] = \frac{a^2 E[\xi_t^2] + \sigma_w^2}{K^2} C_t^3. \tag{3.16}$$

If the number of levels is moderate, then the Panter-Dite formula will not hold with high precision, and since the conditioning of the past data will possibly cause fluctuations in the C_t values, the sequence of C_t values will not be time-invariant. However a fixed value on the C_t can be used to guarantee monotonic boundedness. Note that if the noise is not bounded, conditioning on the past values might lead to schemes that are very time dependent, since there will be quantization bins that have relatively large support sets. However if the noise is assumed to be bounded, then this scheme might lead to more feasible implementations.

3.2.2 Forward-looking quantization

Now, rewriting the system equation as

$$x_n = a^n x_0 + a^{n-1} w_0 + \ldots + w_{n-1}, \tag{3.17}$$

since all the terms are independent, the variance of the expression will be the sum of the individual variances. Our objective is to make the quantization error of this term finite.

Proposition 3.1 *In an unstable linear system, $x_{t+1} = ax_t + w_t$, $|a| > 1$, with any distribution for w_t, to achieve a finite estimation error variance (distortion), the bit rate required is at least* $\log_2(|a|)$ *per sample.*

Proof. Rate required is bounded by the minimum value of the mutual information:

$$I(X, \hat{X}) = H(X) - H(X|\hat{X}) \tag{3.18}$$

$$H(X) = H(a^n x_0 + a^{n-1} w_0 + \ldots + w_{n-1}) = H(a^n [\tilde{x}_0]),$$

where

$$\tilde{x}_0 = x_0 + a^{-1} w_0 + \ldots + a^{-n} w_{n-1}.$$

Let $g(D) = H(X, \hat{X}) = H(X - \hat{X}|\hat{X})$. This is finite and bounded by the entropy of a Gaussian distribution, where D is the finite expected distortion. Thus,

$$R \geq n \log_2(|a|) + H(\tilde{x}_0) - g(D). \tag{3.19}$$

Dividing by n, and taking n to ∞, the average rate required becomes $\log_2(|a|)$. ◇

Proposition 3.2 *For any noise statistics, under fixed code length requirements, the rate* $\log_2(|a|)$ *is achievable.*

Proof. The above rate is actually achievable, and Lloyd-Max quantizer can be employed as the optimal quantizer given the fixed-length constraint for minimizing the distortion. Since the distortion in the quantizer is inversely proportional to the square of the number of levels, to achieve stability, the number of levels will be a^n. This will correspond to a rate of $n \log_2(|a|)$, which yields $\log_2(|a|)$ per unit time. $\diamond$

Not only Lloyd-Max, but most of the other quantization schemes also achieve the distortion. Asymptotic quantization theory [25], [26] shows that the distortion is inversely related to the square of the number of levels, and the constant depends on the type of the quantizer, the minimizing one being the Lloyd-Max quantizer. Such a quantizer is forward-looking. The codes sent at a stage i, $i < n$, depend on the source value at stage n. The higher information set ends up with a lower data rate requirement, when compared with the recursive scheme considered in the previous subsection.

3.3 Summary

In this chapter we have considered a stochastic scalar LTI system where the plant and the controller are connected through a digital noiseless channel. Unlike the results of the Chapter 2, the recursive scheme and the forward-looking scheme are shown to have different rate requirements for stability.

CHAPTER 4

STATE ESTIMATION AND CONTROL FOR LTI SYSTEMS OVER COMMUNICATION NETWORKS

4.1 Introduction and Problem Definition

As regards communication constraints, there are various models that capture the effects of communication in the problem of state estimation and control for systems with channels. In the previous chapters, we studied the optimal coding for state estimation in linear systems when the digital channel is noiseless. In this chapter, we consider the effects of a digital erasure channel, such as the Internet, on remote control performance, and present a new approach to this class of problems under some specific objectives.

Consider a standard LTI control system described by

$$x_{t+1} = Ax_t + Bu_t. \tag{4.1}$$

In the case of a noise-free digital channel, where also the propagation delay is not a factor, the degradation in performance of a controller is solely due to quantization effects, with the controller being a function of the quantized state, $u_t = u_t(\hat{x}_t)$. However, if the controller and the plant are connected over a communication network,

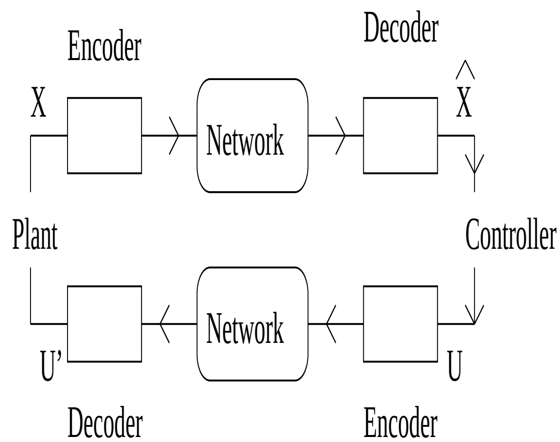

Figure 4.1 Control over networks.

as in Figure 4.1, then because of the uncertainties due to propagation delay and the packet losses, there will be stochastic factors impacting the system.

4.1.1 System model

In a typical system the random delays and occasional packet losses might be from the controller to the plant as well as from the plant to the controller, and in case of a packet loss, it is critical that the sender of the packet be informed of this event. To achieve this, an acknowledgment bit, which will be 1 in case of a successful arrival and 0 otherwise, will be embedded in the packet. In Figure 4.2 we depict all four possibilities of packet losses. In all cases the sender becomes informed of what has happened with regard to his last message, and if there are no acknowledgments, as in the second case, the controller assumes that its last message was lost and by letting this be known to the receiver, the receiver will update its state accordingly.

In all the cases above, the state of the control is made available to both the controller and the plant. Thus, the system estimation error is not affected by the control signals, and it is only the state uncertainty that causes the discrepancy between the state estimated at the controller and the actual state at the plant.

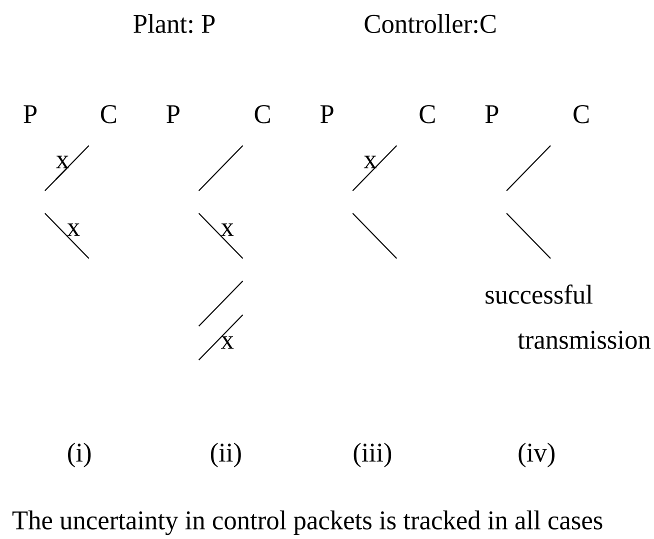

Figure 4.2 Different possibilities for packet losses.

Letting p' denote the probability of a packet being lost in the network, and assuming that the probability of packet losses in a pair of transmissions from the controller to the plant and from the plant to the controller are independent and identical, the effective probability of a packet loss p in the above scheme is

$$p = 1 - (1 - p')(1 - p') = 2p' - p'^2. \tag{4.2}$$

However, we will argue that in a typical network, packet losses occur in bursts, and the joint probability of successive packets to be lost is not very different from the probability of a single packet to be lost, due to the high level of correlation. This point will be further elaborated upon later. We will assume throughout that the initial state is the outcome of a zero-mean random variable with a known distribution. In Chapter 2, it was shown that, regardless of the input distribution, the error in quantization may be assumed to be uniformly distributed without much loss of optimality.

Time-stamping is another important issue to avoid ambiguity in the received packets, which might arise in case a packet belonging to a previous stage arrives. To enable the receiver to distinguish between the ones that are expected to arrive at the current stage, time labels will be carried by the header information in the packets, as will be discussed in later sections.

We consider scalar systems, for ease in presentation, but without much loss of conceptual generality since scalar systems in this context capture all the essential characteristics of higher dimensional systems. For future reference, we rewrite the state equation for scalar LTI systems as

$$x_{t+1} = ax_t + bu_t\,,\ t = 0, 1, \dots \tag{4.3}$$

4.2 Stability in Differential Entropy

Definition 4.1 *The communication rate is the amount of data, measured in bits, that is transferred from one party to another (from a sender to a receiver) over a communication channel.*

The problem of state estimation for LTI systems of the type above was investigated in the Chapter 2, and the required communication rates under various criteria have been obtained. Let $q_t := Q_t(x_t)$ be a quantized value of state x_t under a quantization scheme $Q_t(\cdot)$, and let $\hat{x}_k$ denote the least-squares estimate for x_k using the noise-free measurements of the present and past values of q_t, that is $\hat{x}_k = f_k(q_k, \dots, q_0)$. Then, it was shown that for the variance $E[(x_k - \hat{x}_k)^2]$ to be finite as k approaches ∞, the quantization rate for Q_t should be at least $\log_2(|a|)$. The following theorem now generalizes this result to the case where the channel is a communication network (and hence is noiseless). First two more definitions.

Definition 4.2 *Information bits in a packet are all the bits except the header and stamping bits. Average information rate is the average number of information bits, per time stage, sent over the network.*

Definition 4.3 *A system is stable in the differential entropy of the estimation error if the expected value of the differential entropy, $E(H(e_t))$, is less than a positive finite* *value M. The system is asymptotically stable in the differential entropy if the steady state expected value of the differential entropy is $-\infty$.*

Theorem 4.1 *For stability in the differential entropy of the state estimation error in an LTI system connected through a communication network with a rational, effective packet loss probability of p under the model above, the average information rate required is at least $max(0, \frac{\log_2(|a|)}{1-p})$. Furthermore, this rate is sufficient for the stability in estimation error differential entropy. If the differential entropy of the error is to be asymptotically stabilized, the rate should be strictly greater than $\frac{\log_2(|a|)}{1-p}$, if $|a| > 1$.*

Proof. If the system is stable (that is $|a| < 1$), then there is no need for communication.

If the system is unstable, the uncertainty in the state, which can be measured by the differential entropy [22], increases by $\log_2(|a|)$ at each stage. Thus, after n stages the total amount of information bits the receiver should receive to keep the uncertainty the same as the original uncertainty is $n \log_2(|a|)$, and this should be achieved even in the case of packet loss. Thus a Markov process approach can be used for this system, the states of which represent the uncertainty in the system in terms of the entropy. There will be an increase in the uncertainty at each stage by an amount $\log_2(|a|)$ with probability p, and there will be a decrement in the uncertainty by the difference between the length of the successfully transmitted packet, which we will call R, and $\log_2(|a|)$with probability $1-p$. Therefore we can interpret this as a Markov system which jumps to a higher state with probability p, where the increment is proportional to $\log_2(|a|)$, and which jumps to a lower state with probability $1-p$, where the decrement is $R - \log_2(|a|)$.

Since the probability of packet loss is assumed to be a rational number, the process can be discretized such that the successful arrivals occur with probability p where entropy increases by k units, and a decrement occurs with probability $1-p$, where l units of information are provided to decrease the entropy. This will form an irreducible ergodic Markov chain, with a stationary distribution. Since the process can be discretized, let i denote the index of the state which ranges from $-\infty$ to $+\infty$, and π be the steady state probability distribution. This Markov process has its mean $\sum_i i\pi_i$ equal to zero if $pk = (1-p)l$ or

$$l = \frac{pk}{1-p}. \tag{4.4}$$

The probability of the uncertainty being unbounded is zero, since under the equality in (4.4) above, the probability of each state will be identical and the probability of being infinite is zero. The above Markov process has its mean as $-\infty$ if $pk < (1-p)l$, and in that case the uncertainty of the state has the potential to become zero and thus, asymptotic stability can be achieved with an appropriate coding. In this case, the probability of the state of the uncertainty being positive is zero. Thus, if k is some multiple of $\log_2(|a|)$, then l should be $\log_2(|a|)\frac{p}{1-p}$. Adding $\log_2(|a|)$ to this term, we get the net information rate as

$$R = \frac{\log_2(|a|)}{1-p}. \tag{4.5}$$

For asymptotic stability, the equality above should be a strict inequality. ◇

Figure 4.3 depicts the state evolution that could be viewed as an interpretation of Little's theorem [27]. The steps show successful transmissions. At times 2, 9, and 11 there are packet losses.

Note that the arrivals do not have to be uniform; they are displayed here to be uniform, but the controller can be designed to be event-driven, where the control signal is generated as soon as the state information arrives and is sent back to the plant. Our assumption throughout is that the controller waits until the end of the sampling time and then sends the control signal.

In case of packet loss, it might not be possible to find out which packets were lost; therefore, the redundancy that we will have in channel coding should be introduced carefully to prevent data loss and to enable perfect recovery of the data. We propose to use finer quantization and send a higher amount of information per packet, and therefore each of the packets that will be received at the receiver will provide

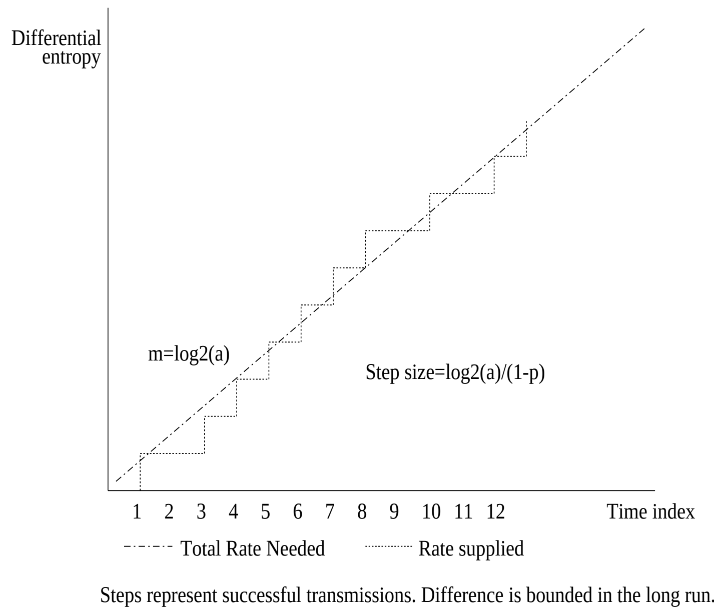

Figure 4.3 Arrival of packets.

useful information. Other coding schemes are also possible, such as embedding the information in the adjacent packets in a neighboring packet. However, in the case of multiple packet losses, this approach becomes inferior. Therefore the number of quantization levels will be $2^{\frac{\log_2(|a|)}{1-p}} = |a|^{\frac{1}{1-p}}$. One other restriction for packet networks is that the codewords should be of fixed length, because variable lengths would make the analysis for the packet protocols more difficult.

4.3 Delay and Quantization

4.3.1 Data packets in networks

In the Internet, the data are sent through packets which are prepared based on various protocols [28], such as the User Datagram Protocol (UDP) or the Transmission Control Protocol (TCP), among many others. In all these protocols there exist header bits and information data, where header bits are used for addressing, error correction, and data characterization.

In control applications, the number of packets for a state or a control signal may vary depending on the system. For example, the quantized state information for the control of a high-precision image might take a large number of packets to be transmitted. In this case, the randomness in the network might change the order of packet transmission. However in this chapter our running assumption is that only one packet is sent per stage, since as was argued in the last section the bit rate required is related to the logarithm of the eigenvalues, which is usually small. In general the length of the header data required for the protocols is much bigger than that of the information bits for a control system.

Thus, in a packet carrying the state and control information, the information bits will be related to the packet length in an affine way. Using a K-level uniform quantization, the packet length will be

$$l = d_0 + \log_2(K), \tag{4.6}$$

where d_0 consists of the time-stamp and the acknowledgment bits, and the fixed header lengths of the protocols. For instance, the fixed header is 8 bytes long for UDP and 32 bytes long for TCP.

4.3.2 Delay in communication networks

The total delay t_d in a communication network can be considered as the sum of processing, queuing, transmission, and propagation delays:

$$t_d = d_{proc} + d_{que} + d_{trans} + d_{prop}. \tag{4.7}$$

The transmission delay is related to the network capacity and the length of the packet. The propagation delay is the time it takes for the packet to reach the next router on its path, and is generally negligible if the physical distance between the transmitter and the receiver is not very long. Processing delay is related to the quantization, encoding, and decoding complexity, and is generally negligible. Queuing delay is the most difficult one to analyze and is related to the rate of congestion at the routers. In case the network becomes congested the finite buffer queues will drop packets, causing losses.

4.3.3 Optimal sampling rate for a simple network delay model

In this subsection we present a simple network model as a prelude to a more realistic one introduced in the next subsection. In the model, the delay in the network is assumed to be inversely proportional to the bandwidth of the network and proportional to the packet length, unless there is no congestion. The congestion takes place with probability p, in which case the packets will be lost. In case there is no congestion, the delay is equal to the delay in the link, where the bandwidth is C bits per second.

Assuming fixed length coding for quantization, the bit rate information will be $\log_2(K)$, where K is the number of levels. Thus, the delay will be

$$d = d_{trans} = \frac{d_0 + (\frac{\log_2(K)}{1-p})}{C}. \tag{4.8}$$

The term above provides a bound on the sampling period. The network provides a lower bound on the sampling period, which should be twice the delay. Hence

$$T_s \geq 2\frac{d_0 + (\frac{\log_2(K)}{1-p})}{C}. \tag{4.9}$$

We should note that the multiplying factor of 2 may not be the optimal selection. The quantization in state information and the quantization in control data might require different lengths for the packets. They are assumed to be equal in this case.

However, note that the rate required is also a function of the sampling period. There is a linear relationship between the rate required and the sampling interval, for there will be a linear increase in uncertainty. Thus, expanding the term $K(T_s) = k_0^{T_s}$, where k_0 is the rate required for a unit sampling period, the effect of the sampling period on delay becomes more transparent,

$$d = \frac{d_0 + (\log_2(k_0)T_s/(1-p))}{C}. \qquad (4.10)$$

To satisfy $T_s \geq 2d$, we need to have:

$$T_s \geq 2d_0(1-p)/(C(1-p) - 2log(k_0)).$$

4.3.4 Random delays and optimal design

In this subsection the queuing delay will be modeled by a random variable. A commonly accepted density function for network queuing delay is the Pareto distribution:

$$F_D(d) = 1 - [\beta/d]^\alpha \quad \text{for } d \geq \beta \qquad (4.11)$$

where F_D is the cumulative distribution. This has been argued to be a good approximation [29] regardless of the time of the day, where the statistics change depending on time-varying conditions and times.

In case the packets are of zero length, as a degenerate case, the probability of packet loss will be $1 - F(\frac{T_s}{2})$, where T_s is the sampling period. Thus, the delay we have is

$$d = d_{que} + d_{trans} = D + \left[\frac{d_0 + \log_2(K)/(1-p)}{C}\right], \qquad (4.12)$$

where C is the effective channel capacity, and D is the random delay whose distribution is given by (4.11). Here $K = k_0 T_s$ is the number of levels we would need in case of perfect transmission, i.e., without packet loss, and can be assumed to be $\lceil aT_s \rceil$.

The choice of the sampling time will affect the probability of packet loss. Thus, there is a trade-off between the precision in quantization and delay, and between the sampling rate and packet loss.

Theorem 4.2 *The following relation holds between the sampling rate, probability of packet loss and the information bit rate:*

$$p = \left(\beta/\left[\frac{T_s}{2} - (\frac{d_0 + \frac{\log_2(k_0)T_s}{1-p})}{C})\right]\right)^\alpha. \qquad (4.13)$$

Proof. Using (4.12) to find d, and then computing the cumulative density function in (4.11), one can find the probability of packet loss as a function of the sampling rate and the number of levels as in (4.13). ◇

4.3.5 Markov model for packet loss statistics

The network delay and packet losses are correlated random processes, and once a packet is lost, it is likely that the adjacent packets will also be lost. The main reason for this is the slow change in the state of networks, i.e., once the network is congested, this will continue for some time and packets will continue to be lost.

Here we introduce a Markov model for the channel's (e.g., Internet's) packet loss characteristics, as can be seen in Figure 4.4. The number of states in the Markov chain could be increased to have finer levels, but having two states, one of which corresponds to the reliable mode (of the channel) and the other one to the unreliable mode, captures the underlying philosophy in a simple model.

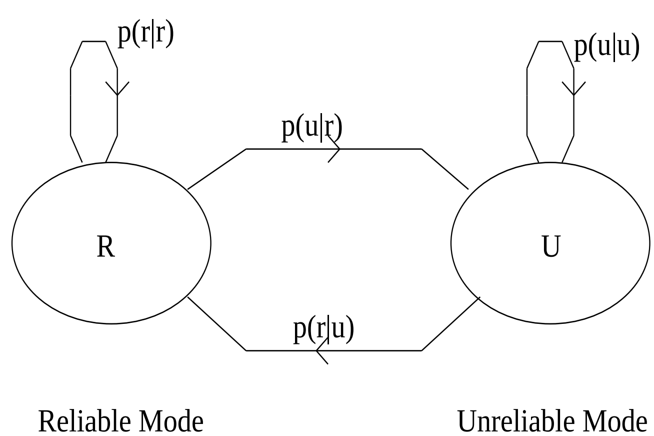

Figure 4.4 Markov model for packet loss statistics.

For the coding of a Markov modeled channel, the quantizers and the decoders can be designed so as to track the mode of the channel. Since we know that in the unreliable mode, the data sent is likely to be lost, we might choose not to send any data, and in the case of the reliable mode, we might send as much data as we can to compensate for the idle information transmission during the unreliable mode.

In this case, the effective probability of packet loss will be $p(u|r)$, and the probability of an excursion from the reliable mode to the unreliable mode. This approach could decrease the data rate required significantly.

The extra bit can be used to inform the controller and the plant regarding the state of the channel. The channel is assumed to be in the reliable mode if during a sampling period, the message from the controller to the plant and from the plant to the controller are transmitted with no loss. In case of a packet loss from the controller to the plant, or from the plant to the controller, the system will be in the unreliable mode.

Theorem 4.3 *Using the Markov model introduced above, let p_r and p_u be the steady state probabilities of the reliable mode and the unreliable mode, respectively. Likewise, let R_r and R_u denote the rates corresponding to each of these modes. Then, the rate required for stability in differential entropy of the state estimation error is*

$$R = \frac{\log_2(K)}{(1 - p(u|r))} + (d_0 + 1 + d_s), \tag{4.14}$$

where d_s is the length of the time-stamp.

Proof. Since we will send no information during the unreliable mode, except the few bits to track the channel mode, the rate required per symbol will be

$$R = R_r p_r + R_u p_u. \tag{4.15}$$

Thus, the amount of information that has to be sent by the plant during the period when packets are transmitted will be $\big(\log_2(K)/(1 - p_u)\big)$. To compensate for the packet losses in the reliable mode, we need an extra factor of $\frac{1}{1-p(u|r)}$ for the data to be sent. During the unreliable mode, the only data that will be sent will be the ones required to track the mode of the channel, which is the sum of the header bits and a single information bit. Thus, the rate will be

$$R = [\log_2(K)/(1-(p_u|r))(1-p_u) + (d_0+1+d_s)](1-p_u) + (d_0+1+d_s)p_u,$$

which leads to (4.14). $\diamond$

Assuming that the noninformation bits are negligible, the data rate will be approximately $\frac{\log_2(K)}{1-p(r|u)}$, which is much smaller than the rate (4.5), since typically the probability of excursion from the reliable state to the unreliable mode is much lower than the steady state probability of the unreliable mode.

4.4 Mean-Square Stability in State Estimation

In this section, we investigate the conditions on the communication rate to ensure stability in mean square state estimation error.

Proposition 4.1 *Consider a scalar LTI system $x_{t+1} = ax_t + bu_t$. Assume time stamping and packet headers are all available and that a uniform quantizer is used to discretize the state. Then the following has to be satisfied to achieve mean-square stability in state estimation:*

$$K > \sqrt{(1-p)/(\frac{1}{a^2}-p)}.$$

Proof. Let e_t denote the state estimation error at time t.

In case there is packet loss, the expected mean square state estimation error will be scaled by a^2:

$$E[e_{t+1}^2] = a^2 E[e_t^2].$$

On the other hand, if there is a successful transmission, the expected mean square state estimation error will be scaled by $E[e_{t+1}^2] = (a/K)^2 E[e_t^2]$, where K is the number of levels in the quantizer which is assumed to be uniform.

Thus the expected value of the second moment at time n, $E[e_n^2]$ will be:

$$E[e_n^2] = pa^2E[e_{n-1}^2] + (1-p)(a/K)^2E[e_{n-1}^2] = a^2(p+(1-p)/K^2)E[e_{n-1}^2]. \quad (4.16)$$

To achieve stability, $a^2(p+(1-p)/K^2)$ has to be less than unity in magnitude. Thus, K has to satisfy

$$K > \sqrt{(1-p)/(\frac{1}{a^2} - p)}.$$

◇

4.5 Stability of an LTI System State in the Mean

In this section control theoretic arguments will be combined with the preceding discussion to investigate the effects of the network as a channel on the stability of the original system.

For control applications, the decision on what the controller should do in case of a packet loss is important. In this section we have two approaches to realize in case of a packet loss. One is to use zero control and the other one is to have the control function of the previous sampling instant to be effective. In the latter case as we could regard the system as event-driven, as soon as the control signal arrives, the new control is used, otherwise the last control will continue to be effective.

In the case of a networked system, let us assume that the packets are timed out, i.e., they are lost with probability p. Given the system (4.3), with probability p the control function u_t will not be available. In that case, if we are to use the last control available, u_t can be regarded as a random variable that is $u_{t-1}(\hat{x}_{t-1})$ with probability p and $u_t(\hat{x}_t)$ with probability $1-p$; thus it is the outcome of a binomial random process. If the plant is to use zero control, in case of a packet loss from the controller, the control function can be modeled as $u_t(\hat{x}_t)$ with probability $1-p$, and 0 with probability $1-p$. The only assumption here is that $E[\hat{x}_t] = E[x_t]$, which holds provided that the quantization error has zero-mean, which is the case for Lloyd-Max quantizer [12],[13] or any type of quantizer with centroid property for the reconstruction values.

In case of packet loss, if the last available control is continued to be used in the next stage,

$$\begin{aligned} x_{t+1} &= ax_t + bu_t \\ u_t &= f(\hat{x}_t) \quad w.p. \ (1-p) \\ &= u_{t-1} \quad w.p. \ p, \end{aligned} \tag{4.17}$$

where f is the controller function, which does not have to be linear in its argument.

Now, the following theorem provides a condition for stability, where we take the controller to be linear in its argument.

Theorem 4.4 *For a control system connected by a network, if the plant uses the most recent control value, which is a linear feedback control, $u_t = k\hat{x}_t$, in case of a packet loss from the controller, to ensure stability in the mean, the following condition has to be satisfied:*

$$\zeta := \left| \frac{(p + a + bk(1-p) \pm \eta)}{2} \right| < 1, \tag{4.18}$$

where

$$\eta := \sqrt{(p + a + bk(1-p))^2 - 4ap}.$$

If zero control is used, the requirement becomes:

$$|a + bk(1-p)| < 1 \,.$$

Proof. If the last control value is used in case of a packet loss, since $E[\hat{x}_t] = E[x_t]$, the recursion takes the form:

$$\begin{aligned} E[x_{t+1}] &= E[ax_t + bu_t] \\ &= E[(a + bk(1-p))x_t + bpu_{t-1}] \\ E[u_t] &= (1-p)kE[x_t] + pE[u_{t-1}] \end{aligned}$$

This two-dimensional system is stable if and only if the eigenvalues of the following matrix are in the unit circle:

$$\mathbf{A} = \begin{bmatrix} (a + bk(1-p)) & bp \\ (1-p)k & p \end{bmatrix}.$$

The condition for this stability is precisely that of (4.18). On the other hand, if zero control is used in case of a packet loss, the underlying system will be (in the mean)

$$E[x_t] = \big[a + bk(1-p)\big]E[x_{t-1}],$$

from which the requirement for stability (in the mean) is as follows:

$$|a + bk(1-p)| < 1. \tag{4.19}$$

◇

Note that, without any packet loss, the stability condition is $|a + bk| < 1$. Clearly this does not imply (4.18). Hence, in a network with delay and packet loss, an otherwise stable system could lose stability. Note that the only assumption made in deriving the result above is that the quantization error should have zero mean.

4.5.1 Simulations

Figures 4.5 and 4.6 depict two scenarios, in both cases with latest control in case of packet loss. With $p = 0.15$, $a = 1.5$, $bk = -1.7$, and rate $= 1$ bit per stage, we have stability in mean as shown in Figure 4.5. The computed largest (in magnitude) eigenvalue has the absolute value of 0.474, which satisfies (4.18).

However, if $p = 0.75$, with everything else as before, the resulting dynamics are unstable, as (4.18) fails, which is also corroborated by Figure 4.6. The maximum (in magnitude) eigenvalue has a magnitude of 1.067. Thus, the outcomes are corroborated by Theorem 4.4.

If zero control is used in case of a packet loss, for the same system as above, we obtain the following: With probability of packet loss as 0.15, Figure 4.7 depicts stability. In this case $|a + bk(1-p)|$ is 0.05.

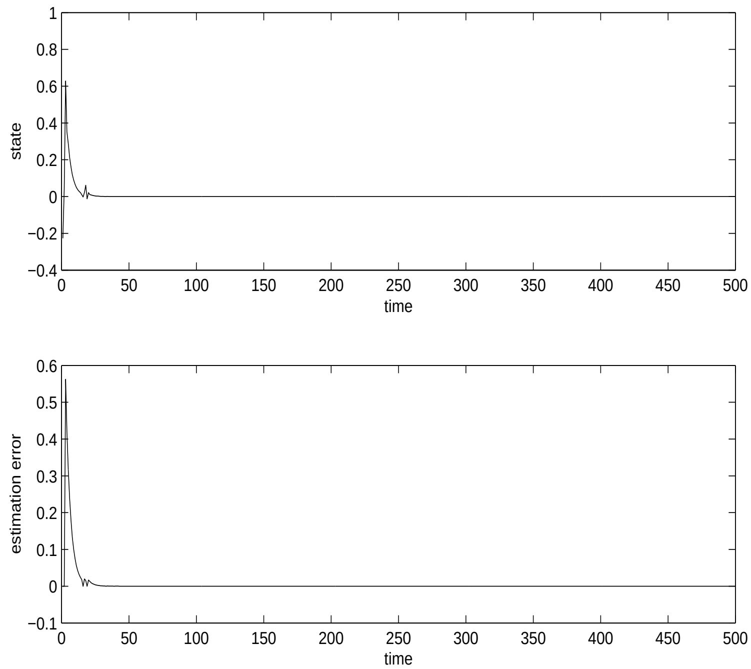

Figure 4.5 Evolution of state and state estimation error, if latest available control is used in case of a packet loss. Packet loss probability is 0.15.

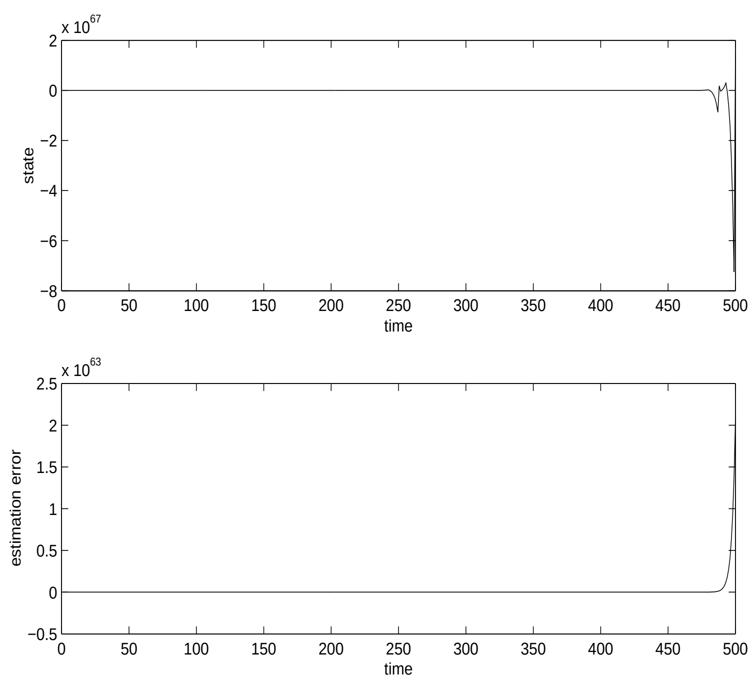

Figure 4.6 Evolution of state and state estimation error, if latest available control is used in case of a packet loss. Packet loss probability is 0.75.

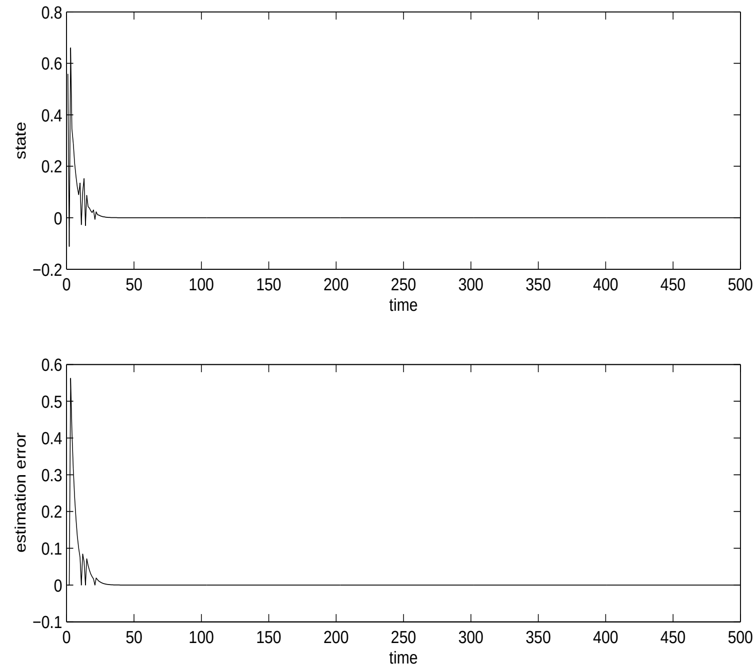

Figure 4.7 Evolution of state and state estimation error, if zero control is used in case of a packet loss. Packet loss probability is 0.15.

However, when the probability of packet loss is increased to 0.75, the system becomes unstable as is illustrated in Figure 4.8. In this case $|a + bk(1-p)|$ is 1.075, and in both cases the outcomes confirm Theorem 4.4.

4.6 Mean-Square Stability of State

In the previous section, we analyzed the stability in the mean for an LTI system. Although stability in the mean is a useful analytical tool, it might not say much about a behavior of a random system. In this section, we analyze the conditions on the rate and packet loss probability for a system to remain stable in the mean square sense. Two different scenarios are covered: the case where zero control is used in case of a packet loss and the case where the latest available control is used.

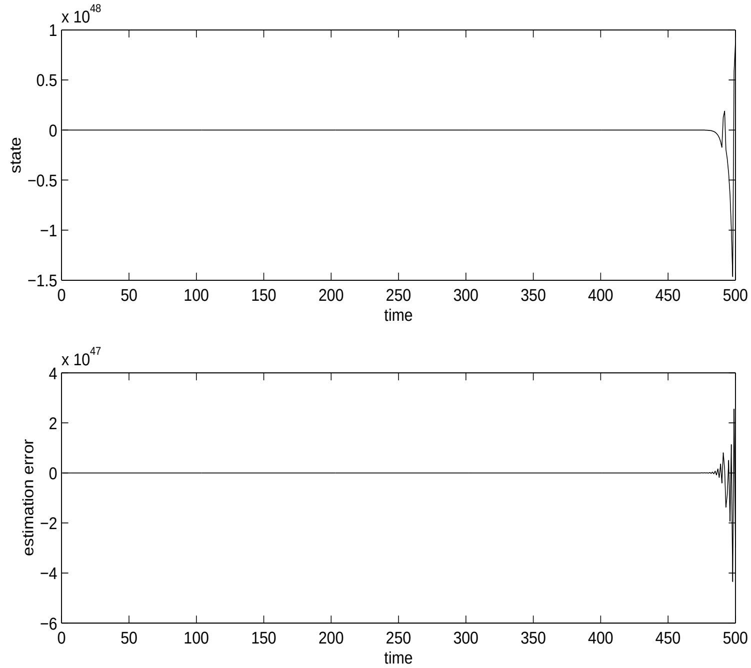

Figure 4.8 Evolution of state and state estimation error, if zero control is used in case of a packet loss. Packet loss probability is 0.75.

4.6.1 Case where zero control is used in case of a packet loss

In this subsection we consider the case where zero control is used in case of a packet loss, which occurs with probability p. Thus, with probability p the system will have a different evolution, than it would have in case of a successful transmission.

In case of a successful transmission, the evolution of the state will be as

$$x_{t+1} = (a + bk)x_t - bk(x_t - \hat{x}_t),$$

but since

$$e_t = x_t - \hat{x}_t,$$

we have,

$$E[x_{t+1}^2] = (a + bk)^2 E[x_t^2] + (bk)^2 E[e_t^2] - 2(a + bk)(bk)E[x_t e_t].$$

Since, by the orthogonality of the quantization error and the quantizer outputs,

$$E[x_t e_t] = E[(\hat{x}_t + e_t)e_t] = E[e_t^2],$$

we have

$$E[x_{t+1}^2] = (a+bk)^2 E[x_t^2] - (bk)(2a+bk)E[e_t^2]. \tag{4.20}$$

And since in case of successful transmission $E[e_{t+1}^2] = (a/K)^2 E[e_t^2]$, defining a two-dimensional state $Y_t = \begin{bmatrix} E[x_t^2] \\ E[e_t^2] \end{bmatrix}$, with probability $1-p$ the mean square evolution will be $Y_{t+1} = BY_t$, where B is

$$B = \begin{bmatrix} (a+bk)^2 & -bk(bk+2a) \\ 0 & a^2/K^2 \end{bmatrix}.$$

In case of a packet loss,

$$E[x_{t+1}] = a^2 E[x_t^2]$$

$$E[e_{t+1}^2] = a^2 E[e_t^2],$$

and hence the system evolution will be $Y_{t+1} = AY_t$, where A is

$$A = \begin{bmatrix} a^2 & 0 \\ 0 & a^2 \end{bmatrix}.$$

These together lead to a Bernoulli random evolution of the state:

$$E[Y_{t+1}] = pAY_t + (1-p)BY_t,$$

which leads to

$$E[Y_{t+1}] = (pA + (1-p)B)E[Y_t].$$

Defining $M := (pA + (1-p)B)$,

$$M = \begin{bmatrix} (a+bk)^2(1-p) + pa^2 & -(1-p)(2a+bk)bk \\ 0 & (1-p)(a^2/K^2) + pa^2 \end{bmatrix},$$

we have

$$E[Y_{t+1}] = ME[Y_t].$$

Thus, the system will be stable if the eigenvalues, λ_i, of M are in the unit circle: Since $\lambda_1 = (a+bk)^2(1-p) + pa^2$ and $\lambda_2 = (1-p)(a^2/K^2) + pa^2$, the condition for stability is:

$$ap^2 + (1-p)\max((a+bk)^2, a^2/K^2) < 1. \tag{4.21}$$

Note that, to achieve stability ap^2 should always be less than 1, as a necessary condition.

4.6.2 Case where latest available control is used

In case there is no loss, we have

$$x_{t+1} = (a+bk)x_t - bk(x_t - \hat{x}_t) = (a+bk)x_t - bke_t,$$

and

$$E[e_{t+1}^2] = (a/K)^2 E[e_t^2].$$

For the state,

$$E[x_{t+1}^2] = (a+bk)^2 E[x_t^2] + (bk)^2 E[e_t^2] - 2bk(a+bk)E[x_t e_t],$$

which, using the orthogonality of the quantization error and the quantization outcomes, becomes

$$E[x_{t+1}^2] = (a+bk)^2 E[x_t^2] - bk(2a+bk)E[e_t^2].$$

We now add two other equations involving the control term. Since

$$u_t = k\hat{x}_t = kx_t - k(x_t - \hat{x}_t),$$

again using the orthogonality principle,

$$E[u_t^2] = k^2 E[x_t^2] - k^2 E[e_t^2].$$

Finally

$$E[x_{t+1}u_t] = E[(ax_t + bu_t)u_t] = E[((a+bk)x_t - bke_t)(kx_t - ke_t)],$$

which becomes

$$E[x_{t+1}u_t] = (a+bk)kE[x_t^2] - (a+bk)kE[e_t^2].$$

Thus, defining a four-dimensional state

$$Y_t = \begin{bmatrix} E[x_t^2] \\ E[u_{t-1}^2] \\ E[x_t u_{t-1}] \\ E[e_t^2] \end{bmatrix},$$

with probability $1-p$ the system evolution will be $Y_{t+1} = BY_t$, where B is

$$B = \begin{bmatrix} (a+bk)^2 & 0 & 0 & -bk(2a+bk) \\ k^2 & 0 & 0 & -k^2 \\ (a+bk)k & 0 & 0 & -(a+bk)k \\ 0 & 0 & 0 & a^2/K^2 \end{bmatrix}.$$

On the other hand, in case of a packet loss, the system equation will be

$$x_{t+1} = ax_t + bu_t,$$

and

$$u_t = u_{t-1}.$$

Thus, we have

$$E[x_{t+1}^2] = a^2E[x_t^2] + b^2E[u_{t-1}^2] + 2abE[x_t u_{t-1}].$$

For the cross-term, using the orthogonality principle

$$E[x_{t+1}u_t] = E[(ax_t + bu_{t-1})(u_{t-1})] = aE[x_t u_{t-1}] + bE[u_{t-1}^2].$$

For the control, we have:

$$E[u_{t+1}^2] = E[u_t^2],$$

and for the state estimation error the following holds

$$E[e_{t+1}^2] = a^2 E[e_t].$$

Thus, in case of a packet loss the system evolution will be $Y_{t+1} = AY_t$, where A is

$$A = \begin{bmatrix} a^2 & b^2 & 2ab & 0 \\ 0 & 1 & 0 & 0 \\ 0 & b & a & 0 \\ 0 & 0 & 0 & a^2 \end{bmatrix}.$$

Again, there will be a Bernoulli random evolution of the state:

$$E[Y_{t+1}] = pAY_t + (1-p)BY_t,$$

$$E[Y_{t+1}] = (pA + (1-p)B)Y_t.$$

This implies

$$E[Y_{t+1}] = (pA + (1-p)B)E[Y_t].$$

Defining $N := pA + (1-p)B$,

$$N = \begin{bmatrix} (a+bk)^2(1-p) + pa^2 & pb^2 & p2ab & -(1-p)bk(2a+bk) \\ (1-p)k^2 & p & 0 & (1-p)k^2 \\ (a+bk)(1-p) & bp & ap & -(1-p)(a+bk)k \\ 0 & 0 & 0 & (1-p)a^2/K^2 + pa^2 \end{bmatrix},$$

we have

$$E[Y_{t+1}] = NE[Y_t], \tag{4.22}$$

which is stable if all the eigenvalues of N, λ_i, are within the unit circle:

$$|\lambda_i| \leq 1, i = 1, 2, 3, 4. \tag{4.23}$$

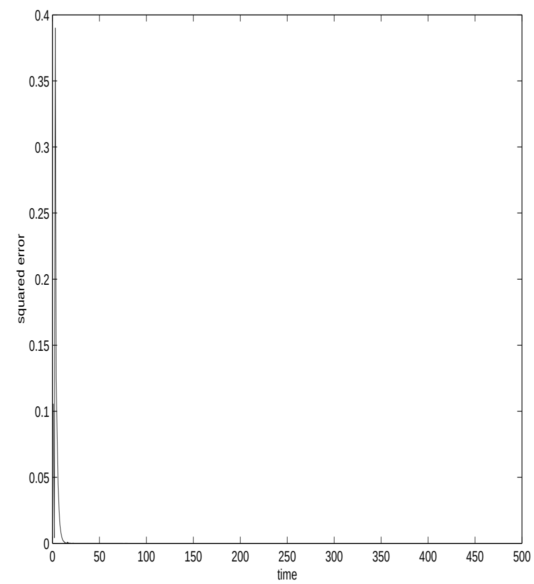

Figure 4.9 Latest available control in case of a packet loss, when packet loss probability is 0.1.

4.6.3 Simulations

Consider the same system considered in Section 4.5.1 connected through a communication network:

$$x_{t+1} = ax_t + bu_t.$$

System parameters are: $a = 1.5$, $bk = -1.7$, and rate $= 1$ bit per stage.

In case latest available control is used in case of a packet loss, with packet loss probability p as 0.1, the system remains stable as is seen in Figure 4.9. In this case the computed matrix N has its largest eigenvalue as 0.86, satisfying (4.23).

However, if $p = 0.5$, with everything else as before, the resulting dynamics are unstable in mean square (Figure 4.10). In this case the largest (in magnitude) computed eigenvalue is 1.41. Thus, both simulations confirm the condition for stability in (4.21).

If zero control is used in case of a packet loss, using the same system as above, we obtain the following:

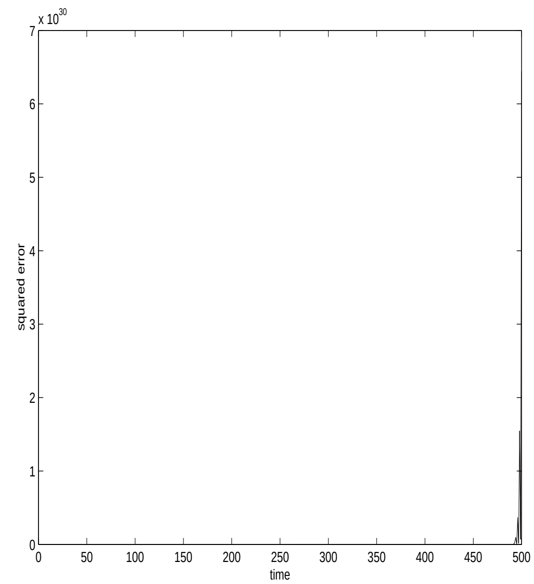

Figure 4.10 Latest available control in case of a packet loss; unstable error dynamics when packet loss probability is 0.5.

With probability of packet loss as 0.1, Figure 4.11 depicts stability, with a largest computed eigenvalue for M as 0.73, which satisfies 4.21. However, when the probability of packet loss is increased to 0.7, the system becomes unstable, as is illustrated in Figure (4.12). The maximum magnitude of the eigenvalues is 1.40. Thus, both simulations confirm the conditions for stability in (4.21).

Finally, in Figure 4.13, the maximum tolerable probability of packet loss for a given data rate is plotted for both latest control usage and zero control usage in case of a packet loss for the same scalar LTI system. In the figure, one can see that the system which uses zero control in case of a packet loss has a wider range of tolerance for packet loss probability to achieve stability in the mean-square sense. Thus, using zero control in case of a pakcet loss is more robust than using the latest available control.

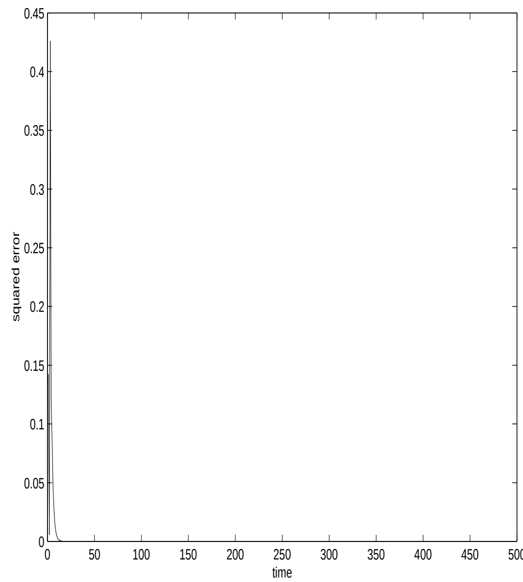

Figure 4.11 Zero control applied in case of packet loss. Loss probability is 0.1.

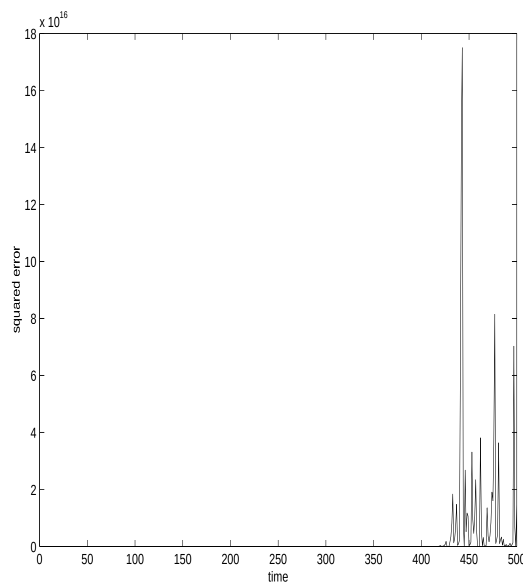

Figure 4.12 Zero control applied in case of packet loss. Loss probability is 0.5.

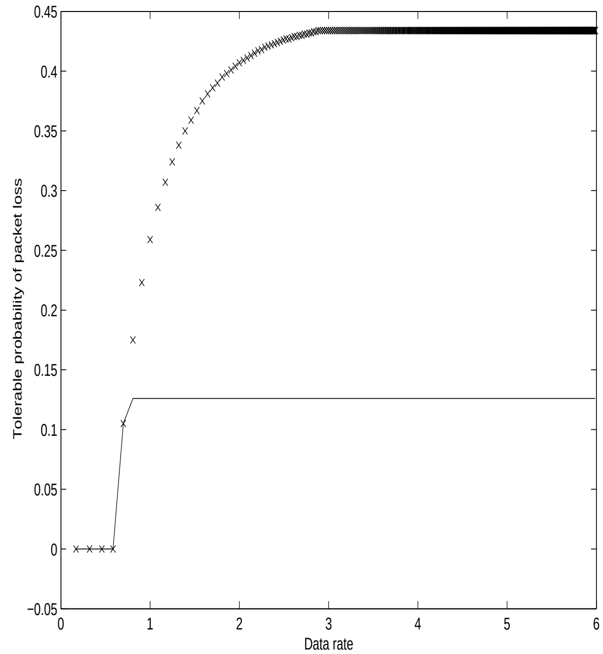

Figure 4.13 Packlet loss tolerance of the system for zero control 'x' and latest control '-'. Zero control has a wider range.

4.7 Summary

In this chapter, the state estimation problem in a communication network has been studied. The effects and the interaction of packet loss probabilities, quantization rate, sampling period, system stability and random delays have been introduced and various network models have been considered in the analysis. A Markov model has been introduced for network reliability and packet losses. Stability in the mean and stability in the mean square have been analyzed, and sufficient conditions (for stability) have been obtained. Extensive simulations included illustrate and corroborate these results.

CHAPTER 5

HIGHER DIMENSIONS AND CENTRALIZED LTI SYSTEMS

5.1 Introduction

In this chapter we analyze the effects of communication constraints on a multi-variable *centralized* system. We will assume, unless otherwise stated, that the initial state vector x_0 is the realization of a random vector X_0 with a finite support set. We consider systems of the form

$$x_{t+1} = Ax_t + Bu_t \, , \; t \geq 0,$$

where A is an $n \times n$ matrix, B is an $n \times k$ matrix, and u is an $k-$dimensional control vector. For ease of presentation, we assume throughout that the matrix A has only real eigenvalues. Our objective is to design optimal dynamic quantizers, $Q := \{Q_t(X), \; t = 0, 1, 2, \ldots\}$ under the two separate criteria, where in all cases $\hat{x}$ denotes the least-squares estimate for x at the receiver (controller) based on all the quantization outcomes available at the receiver. The class of admissible quantizers $\mathcal{Q}$ to be used in each case will be specified later in individual subsections.

We first state the following definition:

Definition 5.1 *For a scalar (one-dimensional) random variable, the support width, W, is the width of the domain over which the probability density function is nonzero:*

$$W(X) = \int 1_{(f(x)\neq 0)} dx.$$

For a multidimensional random variable, the support width is the sum of the support widths corresponding to each of the marginal densities.

Based on the definition above, the two criteria are now given as follows:

Criterion 1: *Monotonic boundedness in uncertainty.*

Find a recursive time-invariant quantizer Q such that the worst-case estimation error is a monotonic bounded sequence, which means that the width of the support set is nonincreasing for each state dimension. Then the goal is to find a quantizer Q that will lead to, $\forall\, t = 0, 1, \ldots,$

$$W_{t+1}(e) \leq W_t(e),$$

where W_t denotes the width W of the estimation error e at the tth step of the iteration.

Criterion 2: *Exponential stability in uncertainty.*

Find a recursive time-invariant quantizer Q such that the width of the support set of the estimation error converges exponentially to zero, i.e., $\forall\, t = 0, 1, \ldots,$

$$W_{t+1}(e) \leq \alpha W_t(e),$$

for some $\alpha < 1$.

If the error happens to be uniform—an issue that will be addressed in detail—the above will imply monotonic boundedness and stability, respectively, in the entropy of the state estimation error.

In Chapter 2, an analysis for forward-looking schemes was also performed, and an improvement in system performance in terms of data rate requirements was shown to be possible for noisy systems. In this chapter we do not consider forward-looking schemes.

5.2 Centralized Scheme

In a classical centralized scheme, there is a single multidimensional plant with a single multidimensional controller, and each subcomponent acts in full cooperation with the others. Thus, in essence, the system is a single multidimensional system, as depicted in Figure 5.1.

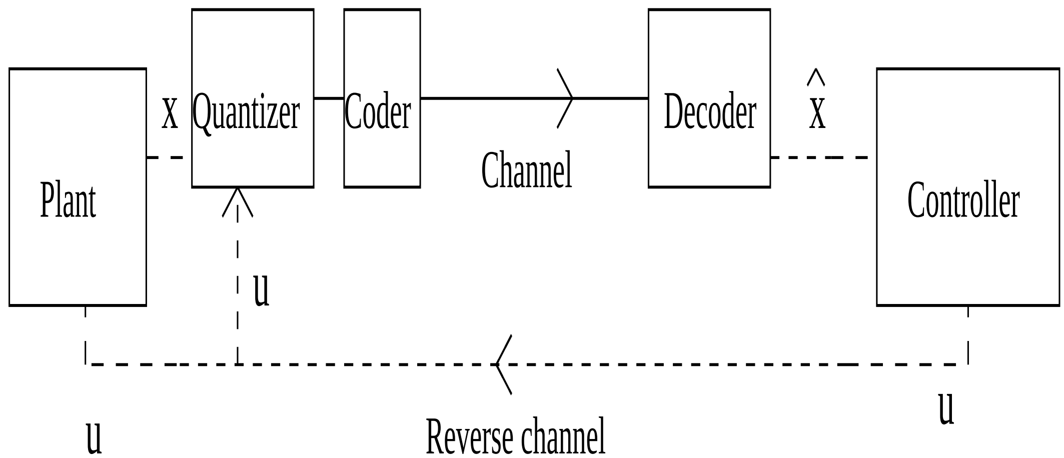

Figure 5.1 Centralized control with communication channel.

We start with the sequential quantization and coding of one-dimensional components.

5.2.1 Sequential scalar quantization of vector components

Vector quantization is a slight generalization of scalar quantization, where the quantization bins and reconstruction values belong to higher dimensions (see [25]). Vector quantization has a higher degree of freedom over scalar quantization of vector components; since for the quantization, the components of the vector can have a joint density that does not have to be of product form, whereas scalar components can only generate joint density functions that are all in product form. However, vector quantization is rather difficult to construct in practice. Implementation of optimal vector quantization under mean square distortion measure as a generalization of the

Lloyd-Max algorithm has been addressed by several authors in the data compression community [13]. In these results, since the analytic formulas and formal derivations for the optimal quantizer are difficult to obtain, the optimization has been performed over a training sequence that is constituted of a set of samples generated according to the underlying stochastic distribution. In such an analysis, the ergodic nature of the processes is essential, and the law of large numbers, consistency and learning theory form the basis of the approach [30].

For the two-dimensional space, for instance, again assuming a uniform distribution for the input vector, if we use squares as the quantizer bins, what we obtain would be identical to sequential scalar quantization of vector elements. However, we might as well use hexagons, or structures with higher numbers of vertices, which will further decrease the expected distortion with the same number of quantization bins in the quantizer. Since scalar quantizers are better understood and are easily implementable, we consider sequential quantization of vector components in this subsection.

In general, the advantage of vector quantization over scalar quantization of vector components diminishes as the statistical dependence between the vector components decreases. One way to increase the efficiency is to apply a linear transformation to the input vector and diagonalize the covariance matrix, and thus decorrelate the vector components:

$$e_{t+1} = (AC^{-1})(Ce_t - C\hat{\mathbf{e}}_{\mathbf{t}}) \tag{5.1}$$

where C is the whitening matrix. Since decorrelation does not imply independence, we found this approach to be less appropriate. A scheme to achieve independence would be to apply a mapping by conditioning, so as to have statistically independent components. This method is appropriate for control systems, which can be converted to a diagonal or a lower triangular form where there is a causality in the evolution of the vector state components, and the conditioning on the previous components will yield a scalar independent distribution for each component. For instance for a two-dimensional system we can use

$$f(x,y) = f(x)f(y|x), \tag{5.2}$$

which will not only decorrelate the inputs, but will also make them independent, and there will be independent evolution for each component if the system matrix is diagonalized:

$$x_{t+1} = (SAS^{-1})(e_t - \hat{e}_t) \tag{5.3}$$

We now consider the general n-dimensional system, the first result of which is the following.

Proposition 5.1 *Consider the n-dimensional system $x_{t+1} = Ax_t + Bu_t$, with A being diagonalizable, with real eigenvalues λ_i, $i = 1, \ldots, n$. If one applies sequential time-invariant scalar quantization, the rate required for the boundedness of the worst-case estimation error is at least $\sum_i \max(0, \log_2(|\lambda_i|))$. For exponential stability, the rate required is strictly larger than this amount.*

Proof. First we will use an information theoretic approach to obtain the ultimate lower bound and then use a quantization viewpoint to show that this bound is achievable. A can be written in the form $A = \Pi\Lambda\Pi^{-1}$, where $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n)$. If the components are made statistically independent by conditioning along with the coordinate transformation the rate required will be the sum of the requirements for each scalar subsystem. Using Lemma 2.2, the distribution for the error will be uniform after a sufficiently large number of steps. In this analysis, since the control functions are transferred with no loss, they do not contribute to the evolution of the uncertainty. However, their existence should always be considered since the reason for quantization to be used instead of, for instance, long block codes is the recursion in the evolution of the system and control functions.

For any scalar subsystem, for any stage i, the rate is lower bounded by the minimum mutual information between the state and its quantized value:

$$\begin{aligned} R &\geq \min_{Q\in\mathcal{Q}} I(\hat{x}_i|\hat{x}_0^{i-1}; x_i|\hat{x}_0^{i-1}) \\ &= H(x_i|\hat{x}_0^{i-1}) - H(x_i|\hat{x}_i, \hat{x}_0^{i-1}), \end{aligned} \tag{5.4}$$

where $\hat{x}_0^{i-1}$ denotes the past information of the quantizer outputs, which are available both at the transmitter and the receiver. Since the system is Markov, the entire information that can be inferred from the past is captured by the latest quantization outcome [20], and is useful in the construction of the encoder and decoder since it is available at both sites. Thus (5.4) becomes

$$\begin{aligned} R &\geq \min_{Q\in\mathcal{Q}} I(\hat{x}_i|\hat{x}_{i-1}; x_i|\hat{x}_{i-1}) \\ &= H(x_i|\hat{x}_{i-1}) - H(x_i|\hat{x}_i, \hat{x}_{i-1}) \,. \end{aligned} \tag{5.5}$$

We now have

$$H(x_i - \hat{x}_i|\hat{x}_i) \geq H(x_i|\hat{x}_{i-1}) - R,$$

and since $x_i = \lambda_i x_{i-1}$,

$$H(x_i - \hat{x}_i|\hat{x}_i) \geq \log_2(|\lambda_i|) + H(x_{i-1} - \hat{x}_{i-1}|\hat{x}_{i-1}) - R\,.$$

Therefore, we need to have at least an additional $\log_2(|\lambda_i|)$ bits per stage to make the entropy nonincreasing. Since the quantizer is intended to be a time-invariant quantizer, to achieve stability the rate has to be strictly larger than $\log_2(|\lambda_i|)$.

Using a quantization theoretic argument, let the ith eigenvalue be λ_i, and let the associated estimation error have a support width of Δ_i, which can be interpreted as the worst-case state estimation error. After one time step, the support width will be $|\lambda_i|\Delta_i$. If a time-invariant K_i-level uniform quantizer is used at each stage, the support width in the next stage becomes $\frac{|\lambda_i|}{K_i}\Delta$, which will be stable only when $K_i \geq |\lambda_i|$. If one uses fixed code length for each of the quantization bins, with a length of $\log_2(K_i)$, the minimum rate required for boundedness is $\log_2(|\lambda_i|)$. Thus, the information theoretic bound is indeed achievable by a single step quantization approach, if the eigenvalue is an integer. If the eigenvalue is not an integer, a rate better than $\log_2(\lceil|\lambda_i|\rceil)$ can be achieved using distortion constrained entropy minimizing quantizer introduced earlier in Chapter 2.

Sufficiency and construction. The error in the estimation of the state will be quantized and fed into the channel, as depicted in Figure 5.2. This is done to minimize the range of the quantizer around the origin. Let λ_i be the eigenvalue of the ith subsystem after diagonalization, whose state is denoted by x^i.

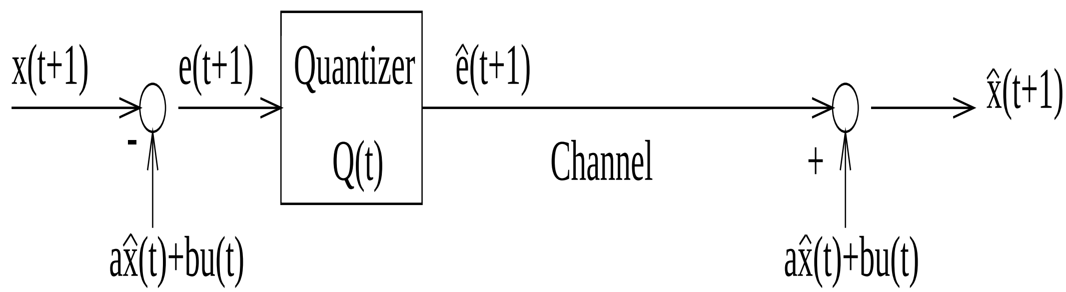

Figure 5.2 The quantization scheme; a represents λ_i.

Then, the scalar system equation for the ith subsystem is

$$x^i_{t+1} = \lambda_i x^i_t + b^i u_t, \tag{5.6}$$

where b^i is a vector. Let $\hat{x}^i_t$ denote the estimate for x^i_t using quantized information. Define the error e as

$$e^i_{t+1} := x^i_{t+1} - \lambda_i \hat{x}^i_t - b^i u_t. \tag{5.7}$$

Let $\hat{e}^i$ be the quantization outcome of the error, e^i. Then, at the receiver we have

$$\hat{x}^i_{t+1} = \lambda_i \hat{x}^i_t + b^i u_t + \hat{e}^i_{t+1}. \tag{5.8}$$

From (5.6) and (5.7), we get

$$e^i_{t+1} = \lambda_i (x^i_t - \hat{x}^i_t). \tag{5.9}$$

From (5.7) and (5.8), with t shifted,

$$x^i_t - \hat{x}^i_t = e^i_t - \hat{e}^i_t. \tag{5.10}$$

By virtue of the last equation, we reduce the error on the state to the error in the estimation. The difference between the actual state and the quantized state is equal to the difference between the actual error and the quantizer error. From (5.9) and (5.10) the following recursion follows:

$$e^i_{t+1} = \lambda_i(e^i_t - \hat{e}^i_t). \tag{5.11}$$

Thus, the error introduced by the quantization becomes the signal to be quantized at the next stage. Hence, it is the quantizer which will determine the power of the uncertainty in the system, which will propagate recursively. To ensure the estimation error to be a nonincreasing sequence, it suffices to enforce $E[(e^i_t)^2] \leq E[(e^i_{t-1})^2]$, which corresponds to

$$\frac{E[(e^i_t)^2]}{E[(e^i_t - \hat{e}^i_t)^2]} \geq \lambda_i^2. \tag{5.12}$$

The proof of the proposition is now completed in view of the following result:

Lemma 5.1 *In an LTI system with a uniformly distributed initial state, for uniform quantization to ensure that the estimation error variance is a nonincreasing sequence with respect to time, the rate should be at least* $\sum_i \max(0, \log_2(|\lambda_i|))$ *bits per sample.*

Proof. Assume that the current support of the uniform pdf is $[-\Delta, \Delta]$. If we use a K-level quantizer, the variance in the quantization error will be $(1/3)(\Delta/K)^2$, and the variance of the error at the next stage will be $(\lambda_i^2/3)(\Delta/K)^2$ and from (5.11) we want this to be smaller than $\Delta^2/3$, which means that we should have $|\lambda_i| \leq K$. We therefore need a quantizer with the number of levels at least as large as $|\lambda_i|$, and to achieve a given K number of levels, we know that we need $R = \log_2(K)$ (since the symbols will be uniformly distributed the bit rate is identical to the entropy of the random variable). Hence,

$$R \geq \sum_i \max(0, \log_2(|\lambda_i|)). \tag{5.13}$$

◇

Corollary 5.1 *Given a uniformly distributed input, for any recursive time-invariant quantizer to achieve exponential stability of the state estimation error variance, the rate has to be strictly greater than* $\sum_i \max(0, \log_2 |\lambda_i|)$*. This is achievable by any* K_i *level uniform quantizer with* $K_i > |\lambda_i|$*.*

Proof. By an information-theoretic argument, the uncertainty and hence the support width of the error decreases over time. At each stage, the support width of the quantizer will be decreased by

$$\Delta_{t+1} = \frac{|\lambda_i|}{K_i} \Delta_t \,, \tag{5.14}$$

which will converge to zero exponentially, with $|\lambda_i|/K_i$, in case $K_i > |\lambda_i|$. In this scheme, the quantizer values will also be scaled by the ratio $|\lambda_i|/K_i$. ◇

5.2.2 Rate-distortion theoretic approach

Rate-distortion theory provides the ultimate bound for the rate required to achieve a given level of distortion [22]. In most cases, the rate distortion bound is not achievable by a single step quantization and requires long block codes, which is not appropriate for a control system because of the delay needed for block coding. Nonetheless, we find it useful to consider the bound obtained by such an analysis.

We state the following proposition.

Proposition 5.2 *For a linear system* $x_{t+1} = Ax_t$*, where the initial state* x_0 *is uniformly distributed and the type of quantization is uniform, the bit rate required for boundedness of the state estimation error variance is* $\max(0, \sum_i \log_2(|\lambda_i|))$*, where* λ_i*'s are the eigenvalues of* A*.*

Proof.

$$R(D) = \min_{p(x):E[|X-\hat{X}|^2 \leq D} I(X, \hat{X}), \tag{5.15}$$

where I denotes the mutual information [22]. Now,

$$\begin{aligned} I(x_n, \hat{x}_n) &= H(x_n) - H(x_n|\hat{x}_n) \\ &= H(x_n) - H(x_n - \hat{x}_n|\hat{x}), \end{aligned} \tag{5.16}$$

and since $H(x_n - \hat{x}_n|\hat{x}) \leq H(x_n - \hat{x}_n)$,

$$I(x_n, \hat{x}_n) \geq H(x_n) - H(x_n - \hat{x}_n)\,.$$

Also,

$$H(A^n x_0) = \log_2(|\det(A)|^n) + H(x_0) \tag{5.17}$$

where $|\det(A)|$ is the absolute value of the determinant of A. Since $\det(A) = \prod \lambda_i$,

$$\begin{aligned} I(x_n, \hat{x}_n) &\geq H(x_n) - H(x_n - \hat{x}_n) \\ &= n \log_2(|\prod \lambda_i|) + H(x_0) - f(D), \end{aligned} \tag{5.18}$$

where $f(D)$ is a finite number. If we divide this expression by n, and let n approach ∞, this leads to the rate required: $\max(0, \sum_i \log_2(|\lambda_i|))$. ◇

5.3 Simulations

In this section, a numerical example is presented to confirm and illustrate the theoretical results obtained in the previous sections.

We consider a centralized system with a random initial state vector, where each component is independently distributed according to a uniform distribution on the interval $[-1, 1]$. The state evolution is

$$x_{t+1} = Ax_t + Bu_t,$$

where

$$A = \begin{bmatrix} -4 & 1 & 2 \\ 1 & 3 & 2 \\ 2 & 2 & 5 \end{bmatrix}$$

The control terms are arbitrary, and in the framework of this chapter, since controls have no effect on the evolution of uncertainty, they are assumed to be zero in the simulations without loss of generality.

In Figure 5.3, the stability analysis of the state estimation error is illustrated. The rate used is $\sum_i \log_2(\lceil|\lambda_i|\rceil) = 6.49$ bits/sample, λ_i being the eigenvalues, and stability in estimation error is achieved. The results confirm the analytical results, since the rate used is greater than the minimum required rate in (5.13).

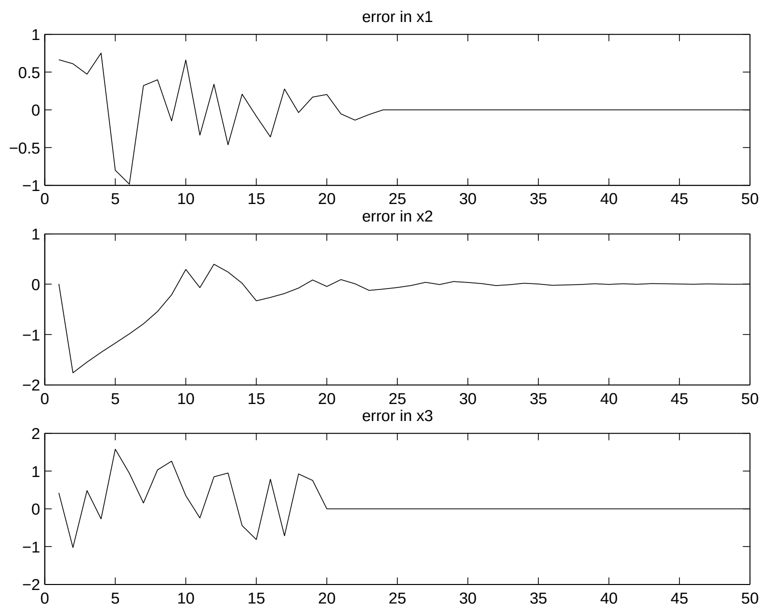

Figure 5.3 Centralized scheme.

5.4 Summary

In this chapter we have generalized the results of Chapter 2 to higher dimensions and showed that the eigenvalues of the system matrix determine the rate requirements in the system. After a discussion on vector quantization, we introduced sequential quantization for each component of the state vector, which are made independent by conditioning on the other components in a causal way. Finally rate distortion theoretic analysis is performed to establish an ultimate lower bound for the required data rate.

CHAPTER 6

DECENTRALIZED LTI SYSTEMS WITH DIGITAL NOISELESS CHANNELS

6.1 Introduction

In this chapter, we investigate the communication problem in decentralized systems and consider coding and decoding schemes for LTI systems with different structures and information sharing topologies involving the plants and the controllers, which are connected via a band-limited channel. Simulations and case studies are presented in Section 6.3, and the chapter ends with the concluding remarks.

6.2 Decentralized Scheme

We consider an LTI system of the following form:

$$x_{t+1} = Ax_t + Bu_t, \tag{6.1}$$

where $B = \text{diag}(b_1, \ldots, b_n)$ is a diagonal $n \times n$ matrix, A is an $n \times n$ diagonalizable matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, x_t is an n-dimensional state vector, and u is an n-dimensional control vector. Since the control matrix B is diagonal, all the control terms have direct impact on the associated state with the same index i, $i \leq n$. Thus, the system is decentralized in the sense that each controller state pair, a subsystem,

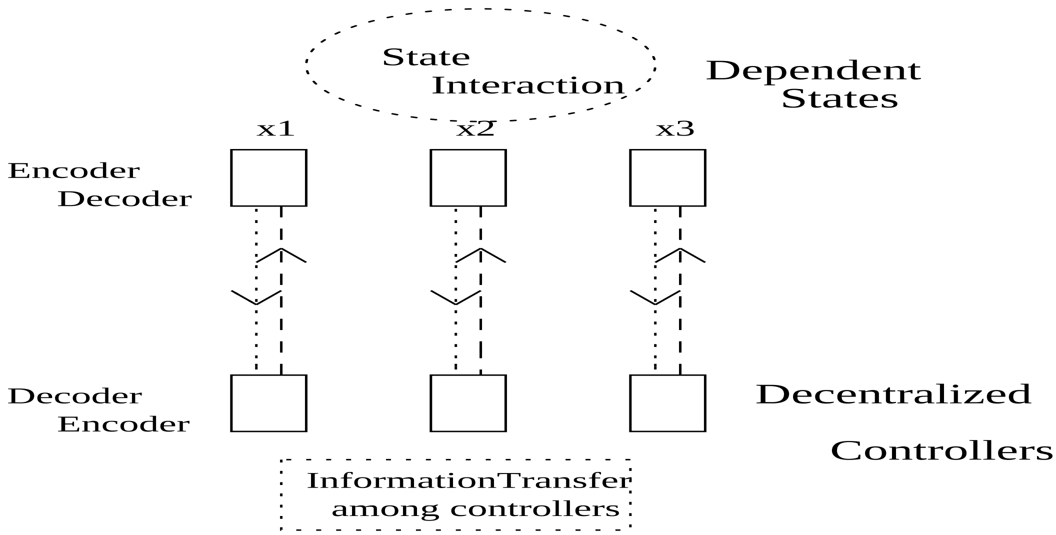

Figure 6.1 Decentralized structure.

is decoupled from the other ones as is depicted in Figure 6.1. Thus, for instance, for the ith component (x^i) of the state, the evolution in time will be:

$$x_{t+1}^i = a_{i1}x_t^1 + a_{i2}x_t^2 + a_{i3}x_t^3 + \ldots + a_{in}x_t^n + b_i u_t^i. \tag{6.2}$$

In case there are no communication and quantization effects, u_t^i will be a function of the perfect state information available to it. If the controller has access to all the observations corresponding to other state components and the control functions, then the system will effectively be a centralized one. In contrast, we consider in this section the situation where each controller does not have access to the control functions used by the remaining controllers, and we investigate different cases where various different types of information is available to the controllers, which is related to the amount of communication among the controllers. Here, due to channel effects, the controller will not be able to follow the state evolution perfectly, and the error in state tracking will depend on the communication scheme and the rate used for the transmission of information. The primary interest is to introduce a rate analysis for reliable state estimation.

Various schemes are possible for information structures at the controllers. In spite of being impractical, the case where state information is made available to each subsystem is worth analyzing since it provides a lower bound on the system requirements.

6.2.1 Case where encoders and decoders share the estimates

In this scheme, the estimated states and not the control signals are shared between the controllers. As a further assumption, the encoder is assumed to have access to the corresponding receiver's information. In this case if we condition on the immediate past, due to the Markov property of the state, the further past data will be of no use [20]. Let us consider the first component: If we subtract the predicted value

$$a_{11}\hat{x}^1 + a_{12}\hat{x}^2 + \ldots + a_{1n}\hat{x}^n + b_1 u^1,$$

from the actual state x_{t+1}, the innovation to be sent will be the sum of the individual quantization errors, which are independent given the past state estimates.

Here an asymptotic analysis is possible, as well as a causality-based one. Even if the initial distributions are uniform for each component, the quantization error at a later time will not be uniform due to the additive nature in the noise terms, whose probability density computation requires a series of convolution operations. Note that the main difference between the centralized scheme and this scheme lies in the unavailability of the control signals. If those were also known, then the system could be transformed into a diagonal form, which would lead to the analysis done in the Chapter 5.

Proposition 6.1 *Assume that the initial state of each subsystem has a finite support pdf. If fixed length coding is used, the rate required for exponential stability of the worst-case error is strictly larger than*

$$\sum_i \max(0, \log_2(|\lambda_i|)).$$

To achieve boundedness, the rate required is lower bounded by

$$\sum_i \max(0, \log_2(|\lambda_i|)).$$

Proof. Since the information structure above is inferior to the sequential scalar quantization case discussed in Chapter 5 (where the control functions are shared as well), the rate required for that case can be regarded as a lower bound for the scheme under consideration for this case. Note that in the decentralized scheme, sequential analysis is inevitable, since a joint analysis would require information sharing between the subsystems. ◇

The above analysis is based on an information theoretic argument. In the following proposition, we restrict the analysis to time-invariant quantizers with fixed length encoders. In this case, the signal to be coded will be the state conditioned on the information set.

Proposition 6.2 *To achieve boundedness of the quantization error with a time-invariant quantizer, with fixed length codewords, the minimum rate required is given by the solution of the following optimization problem: Let A^+ be obtained by replacing all entries of A with their absolute values, and let $\mathcal{K}$ denote the class of diagonal $n \times n$ matrices K (with the ith diagonal element denoted by K_{ii}), where $\frac{1}{K_{ii}}$ are integers, and further KA^+ is stable. Then, the optimization problem is, to find the minimum rate, R:*

$$R = \min_{\mathcal{K}} \sum_i \log_2(1/|K_{ii}|). \tag{6.3}$$

Proof. Due to the Markovness of the system, for the ith component, the signal to be quantized is the conditioned state on the quantized data of the previous state:

$$e_{t+1} = (x^i_{t+1} | \hat{x}^1_t + \hat{x}^2_t + \ldots + \hat{x}^n + bu^1_t),$$

which is identical to

$$a_{i1}(x^1_t - \hat{x}^1_t) + a_{i2}(x^2_t - \hat{x}^2_t) + \ldots + a_{in}(x^n_t - \hat{x}^n_t). \tag{6.4}$$

Defining the quantization error for each component

$$q_t^i := (x_t^i - \hat{x}_t^i),$$

we have

$$e_{t+1}^i = a_{i1} q_t^1 + a_{i2} q_t^2 + \ldots + a_{in} q_t^n.$$

In matrix form, this is equivalent to

$$e_{t+1} = A q_t.$$

Assuming the initial condition for each component x_0^i to be bounded, not necessarily being uniform, at time t before quantization, the support width W of the random variable to be quantized e^i will be

$$|a_{i1}|\Delta_t^1 + |a_{i2}|\Delta_t^2 + \ldots + |a_{in}|\Delta_t^n.$$

If one uses a K_{ii} level uniform quantization, the support width of each bin will be

$$\frac{|a_{i1}|\Delta_t^1 + |a_{i2}|\Delta_t^2 + \ldots + |a_{in}|\Delta_t^n}{K_{ii}}, \tag{6.5}$$

and using this approach for each dimension,

$$\Delta_{t+1} = K A^+ \Delta_t, \tag{6.6}$$

where K is a diagonal matrix with the reciprocal of the number of levels forming the nonzero diagonal terms.

Thus, the problem is reduced to a linear system problem, and the stability of the estimation error can be investigated by finding the eigenvalues of the matrix KA^+, and requiring them all to be in the unit circle. The rate required is the sum of the logarithms of the numbers of levels for each subcomponent in the system, and is thus equal to

$$\sum_i \log_2(\max(1, \frac{1}{K_{ii}})).$$

◇

In case no information is transmitted for a particular subcomponent of the control system, the number of levels required can be interpreted to be 1, since the rate required will be zero. Therefore the matrix K has elements at most equal to 1 on the diagonal. Thus, an optimization problem can be posed as follows:

$$\min_{\mathcal{K}} \sum_i \log_2(1/|K_{ii}|), \tag{6.7}$$

which is the minimum achievable rate. But since $\sum_i \log_2(\frac{1}{|K_{ii}|}) = \log_2 \prod \frac{1}{|K_{ii}|}$, logarithm is a monotonic function, and K is diagonal, the problem reduces to one of finding the optimal K_0 matrix solving

$$\det(K_0) = \max_{\mathcal{K}} \det(K) \tag{6.8}$$

Proposition 6.3 *When A is a diagonal matrix, the minimum fixed code length data rate as the solution of (6.7) 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. When A is diagonal, the condition for stability becomes:

$$|K_{ii}\lambda_i| < 1.$$

Thus,

$$\frac{1}{|K_{ii}|} > |\lambda_i|.$$

To achieve strict inequality using the constraint that $1/K_{ii}$ is an integer: If λ_i is not an integer,

$$\frac{1}{K_{ii}} = \lceil(|\lambda_i|\rceil)\,,$$

whereas if λ is an integer,

$$\frac{1}{K_{ii}} = |\lambda_i| + 1\,.$$

Thus, the rate is:

$$R = \sum_i \log_2(C(|\lambda_i|)).$$

◇

Remarks If the probability density is assumed to be uniform, which is a reasonable assumption for a diagonal system since there is no state interaction, a lower rate can be achieved using the distortion-constrained entropy minimizing quantizer studied in Chapter 2, which exploits the advantage of variable length coding. ◇

Proposition 6.4 *The optimization problem (6.3) admits a solution for any diagonalizable matrix A.*

Proof. If one uses

$$\frac{1}{K_{ii}} = C(\max_i |\lambda_i|) =: |\lambda_m|,$$

the matrix KA^+ can be expressed as:

$$KA^+ = KS^{-1}\Lambda S = \frac{1}{|\lambda_m|}S^{-1}\Lambda S = S^{-1}\frac{\Lambda}{|\lambda_m|}S,$$

which has all its eigenvalues in the unit circle. The rate in this case becomes:

$$R = n\log_2(|\lambda_m|)$$

There is only a finite number of choices for the matrix K which are better than the one proposed. Thus, there should exist at least one optimal solution. ◇

Remarks In this problem, and in this chapter in general, the intention is to bound the maximum error, i.e., to have a bounded error. For the rate analysis, the information theoretic analysis is not used, and quantization theoretic approach is used for the sake of practicality, since the probability density function at each stage changes dynamically and having nonuniform quantization schemes, combined with the convolution analysis would then make the problem intractable. Furthermore, having uniform quantization will be important in the construction of different information share structures among the controllers. The construction for asymptotic stability will be time-invariant in the sense that the quantizer coefficients change by just a scaling at each time. $\diamond$

6.2.2 Case where controllers fully share their estimates

We now consider a structure where the controllers share their data regarding estimation, and since each is able to compute the necessary control function after receiving the estimates from the other controllers, the controllers do not have to share the control functions. We now state the following proposition.

Proposition 6.5 *To achieve boundedness of the quantization error with a time-invariant quantizer, with fixed-length codewords, the minimum rate required is given by the solution of the optimization problem introduced in Proposition 6.2, and thus there is no loss of performance when compared with the case where the subplants have access to the estimates at the controllers.*

Proof. The conceptual proof is based on the Slepian-Wolf coding theorem [31]. Slepian and Wolf showed that if some random data are not available at the encoder, but are available at the decoder, then the data rate can still be decreased by conditioning on the information available at the decoder. This is at first a surprising result because the encoder does not have the side information and without it how it would achieve optimal quantization is not immediately clear.

Consider a three-dimensional system. For this case, the Slepian-Wolf theorem says that (also using the Markov property) the rate required for each channel is bounded from below by:

$$R > H(x_{t+1}^1 | \hat{x}_t^1, \hat{x}_t^2, \hat{x}_t^3)\,. \tag{6.9}$$

Thus, in theory, there is no loss in optimality, and the communication between the controller and the plant is unnecessary. Below we show that this is indeed the case in practice as well.

Construction. Coming up with constructions that achieve Slepian-Wolf coding efficiency remains a difficult and open problem. Schemes that use long block codes and codes with memory are available, but these are mostly impractical for control systems because of the significant delay they entail. There are approaches based on binning and coset formation ([32], [33]) that are more practical but mostly optimal only in the asymptotic sense. However, for the problem we consider, Slepian-Wolf gain is indeed achievable with a single step quantization using a binning argument. The approach we take is based on a uniform quantization interpretation of binning and exploits the optimality based on the assumption that the quantization and estimation errors are bounded.

The data to be sent, x^i_{t+1}, can be written as

$$x^i_{t+1} = a_i \hat{x}_t + a_i q_t, \tag{6.10}$$

where $\hat{x}_t$ is the vector of the state estimates available at each subcontroller, q_t is the vector of quantization error at time t, and a_i is the ith row of matrix A^+. In this problem the side information available is the term $\hat{x}_t$.

The goal again is to have a nonincreasing sequence for the support width of the estimation error density. Assume that d^i_t is the support width of the uncertainty of x^i_t, and suppose that d^i_{qt} is the support width of the random variable $a_i q_t$. Since the support width d^i_{qt} is to be finite, the side information and the state differ at most by the width of the support, and this will be the idea behind the construction scheme.

As seen from Figure 6.2, there are n cosets, where n is 4 in the figure. The quantizer therefore needs only inform the receiver which coset the data belongs to. The receiver will use the coset and decode the signal using the available side information. The rate required will then be $\log_2(d^i_q/d^i)$.

Let us assume that the encoder sent $C1$. In this case, the receiver will just need to choose between the possible bins in the coset. Since it is assumed that there is always a nearest one without ambiguity, the evolution of the vector d_q will be

$$(d_q)_{t+1} = A^+ d_t. \tag{6.11}$$

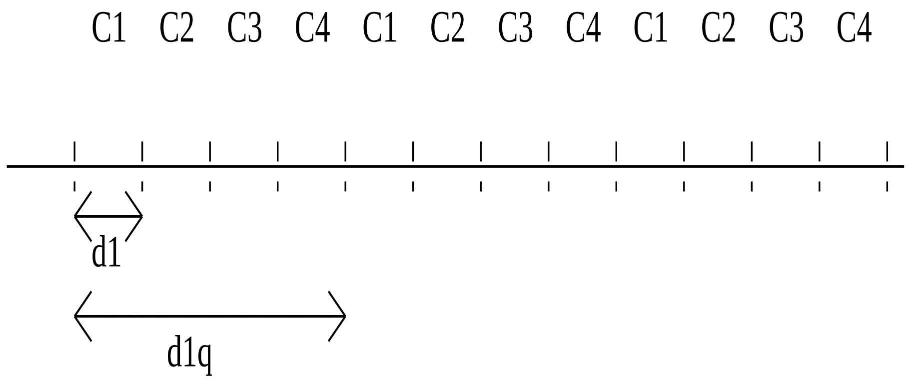

Figure 6.2 The decoder can decode the signal by just knowing the coset.

After quantization,

$$d_{t+1} = (\frac{1}{K_i})d_{qt},$$

which is equivalent to

$$d_{t+1} = (\frac{1}{K_i})A^+d_t.$$

Now, define K as the diagonal matrix consisting of the reciprocals of the number of quantization levels on its diagonals. Recognizing this expression from the previous section, an optimization problem can be proposed as follows:

$$\min_{\mathcal{K}} \sum_i \log_2(\frac{1}{|K_{ii}|}), \tag{6.12}$$

where $\mathcal{K}$ is as defined in the previous section. Note that this is equivalent to the optimization problem introduced in (6.7). ◇

6.2.3 Information sharing is delayed

We now consider a different scheme where information on the stage before the last one, including the control functions, is available.

Consider the first subsytem. Before quantization,

$$q^1_{t+1} = a_{11}q^1_t + a_{12}\sum_i a_{2i}q^i_{t-1} + a_{13}\sum_i a_{3i}q^i_{t-1} + \ldots$$

Thus, in vector form, we have:

$$q_{t+1} = K((A^+ - \mathrm{diag}(A^+))A^+)q_{t-1} + K\mathrm{diag}(A^+)q_t. \tag{6.13}$$

In (6.13), since the control functions corresponding to different controllers are not available and cannot be estimated at time t, the only information to be used regarding those subsytems is the estimation error at the stage for which the estimation is shared.

Let $\mathcal{K}_n$ denote the set of diagonal positive and real matrices K satisfying stability condition in system (6.13). The second-order system in (6.13) can be reduced to a first-order system in the usual way: Define

$$q_{t-1} = r_t,\ q_t = p_t.$$

Then (6.13) becomes

$$\begin{bmatrix} r_{t+1} \\ p_{t+1} \end{bmatrix} = B \begin{bmatrix} r_t \\ p_t \end{bmatrix},$$

where

$$B = \begin{bmatrix} 0 & I \\ K[(A^+ - \mathrm{diag}(A^+_{ii}))A^+ & K(\mathrm{diag}(A^+_{ii})) \end{bmatrix}.$$

The optimization problem is to find a K_0 such that

$$\det(K_0) = \max_{\mathcal{K}} \det(K). \tag{6.14}$$

Proposition 6.6 *When A is diagonal, the maximum value in (6.14) is*

$$R = \sum_i \log_2(C(|\lambda_i|)),$$

where $C(x)$ is the smallest integer strictly larger than x (a modified ceiling function).

Proof. The analysis is identical to the full information sharing case. ◇

Proposition 6.7 *There exists a solution to the optimization problem above.*

Proof. The cardinality of the solution set is countably infinite and unbounded from above (in terms of the numbers of levels in the quantizers), and thus, if a conservative solution is provided, then the set of better possible solutions will be finite since the solution set is bounded from below. A very conservative solution is provided below.

The matrices involved in (6.13) all have positive entries. If x is a positive vector with bounded entries (i.e., $x_i \leq D$ for some positive D), then there exists a scalar m such that

$$((A^+ - \text{diag}(A^+))A^+)x \leq mx.$$

Likewise, there exists a scalar n such that

$$\text{diag}(A^+)x \leq nx,$$

where the relation between m and n can be, for instance,

$$m = n(n-1)\max_{i,j}(A^+(i,j))^2 D\,,$$

and n can be

$$n = \max_i A^+(i,i).$$

If the above inequalities are used and are multiplied by a diagonal $K = kI$ matrix from the left, we get the following scalar relation for each dimension:

$$q^i_{t+1} \leq kmq^i_t + knq^i_{t-1}.$$

Since the entries are all positive in the above expression and the system is scalar, equality is a worst-case. Finding the roots of the characteristic polynomial, $s^2 - kms - kn = 0$, suffices to show that the value of

$$k = (1/2)\max(m, 4n)$$

will lead to stability. Thus, a diagonal matrix with $K_{ii} = 1/k$ will achieve stability.

Since this conservative solution provides an upper bound, and there exists a finite number of candidate quantizers outperforming the conservative scheme, there indeed exists an optimal scheme. ◇

There is a trade-off between the communication in the plant-controller pairs and the communication among the controllers. There might be different schemes, such as different delay structures for each controller, depending on the physical distance between the controllers, or a variety of communication rates that might be used among the controllers, such as the controllers not sharing their estimates with precision in order to save from bandwidth usage. In such cases as well, the Slepian-Wolf argument can be used to find the underlying structure for the optimization of this trade-off. We will not address these extensions in this thesis, and leave them as future research directions.

6.3 Simulations

In this section numerical examples are presented to illustrate and corroborate the theoretical results obtained in the this chapter.

We consider a system with a random initial state vector, where each of the three dimensions has a continuous uniform distribution in the interval $[-1, 1]$ and with the following evolution relation:

$$x_{t+1} = Ax_t + Bu_t,$$

where A is (as was the case for the simulation in Chapter 5)

$$A = \begin{bmatrix} -4 & 1 & 2 \\ 1 & 3 & 2 \\ 2 & 2 & 5 \end{bmatrix}$$

and B and u are arbitrary at this point. In fact, in the framework of this chapter, since controls have no effect on the evolution of uncertainty, they will be taken to be zero in the simulations without loss of generality.

6.3.1 Case where side information is available

The system is used for the simulation in a decentralized setting. The rate $R = 6.49$ bits/sample, which was a stabilizing rate for the centralized scheme in Section 5.3, is used, and an unstable output is obtained (Figure 6.3), showing the inferiority of the decentralized scheme.

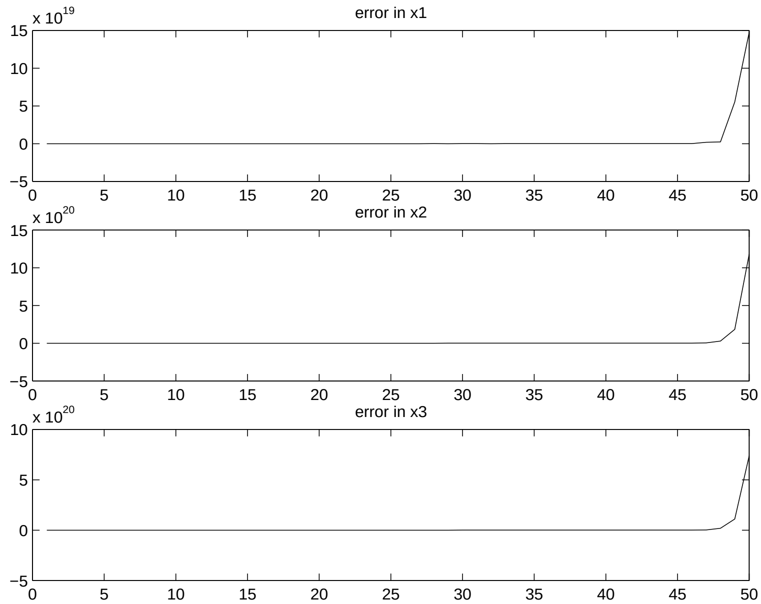

Figure 6.3 Decentralized unstable system.

6.3.2 No side information at the plant

However, if R is increased to, say $R = 8.4221$ bits/s, stability is achieved as can be seen in Figure 6.4. In this case the number of levels is taken to be the ceiling function of the eigenvalue of A^+ with the maximum absolute value, which is 7. Thus this rate is higher than the lower bound introduced in (6.7).

We consider a scalar system, where the plant does not have access to the side information at the receiver (which corresponds to the case where the encoder is memoryless). The system we consider is of the form

$$x_{t+1} = 1.5x_t + bu_t.$$

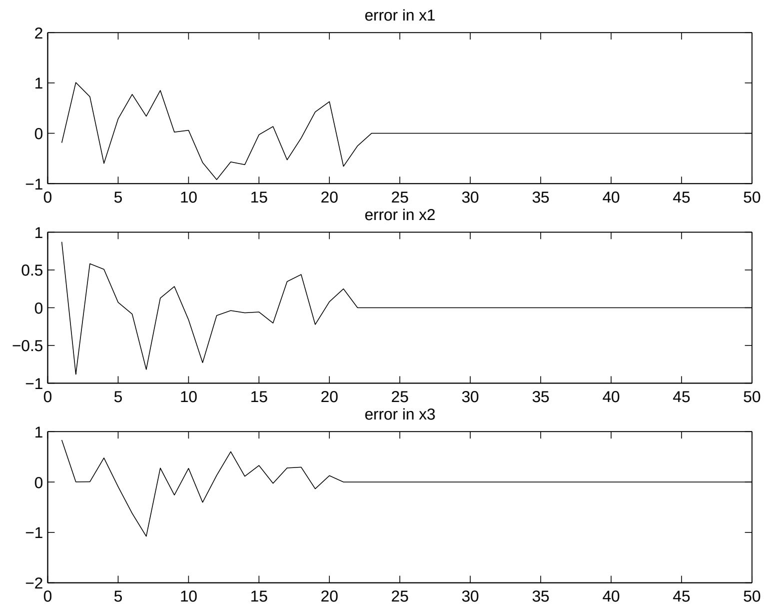

Figure 6.4 Decentralized stable system.

The rate used in the simulation is 2 bits/s which should be enough to stabilize the estimation error by (6.12). A time-invariant quantizer is used, and Slepian-Wolf efficiency and stability is achieved (Figure 6.5).

6.4 Summary

In this chapter, communication rate requirements for noiseless decentralized LTI control systems have been investigated, and recursive quantizers leading to stability have been introduced. Rates required for centralized schemes have been shown to be lower than those for decentralized schemes. Different structures for sharing information are shown to be important factors for communication requirements and complexity. Slepian-Wolf coding argument is used to show that the information sharing by the controllers, and not by the plants, is sufficient for optimality, and schemes confirming this efficiency are constructed.

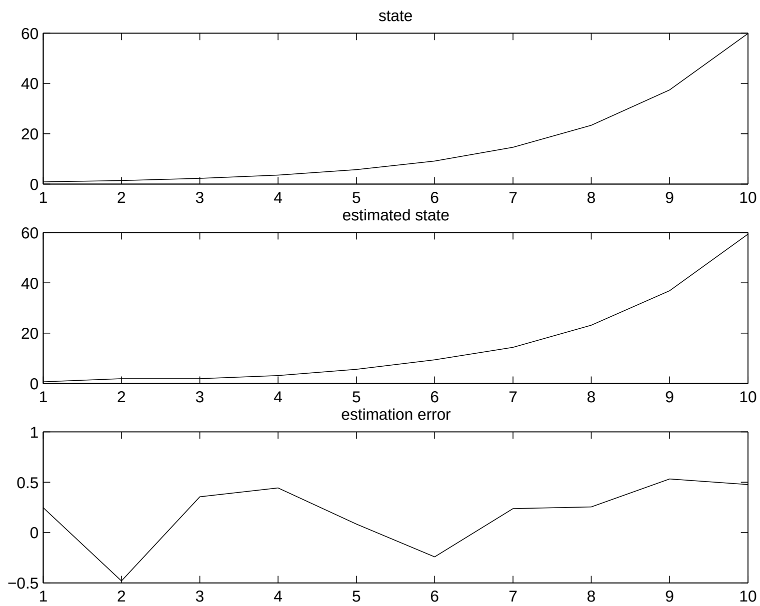

Figure 6.5 Stability in estimation with no estimator information at the plant.

CHAPTER 7

LINEAR QUADRATIC REGULATOR (LQR) MINIMIZATION WITH COMMUNICATION CONSTRAINTS

7.1 Introduction

In this chapter we consider a linear quadratic regulator (LQR) problem with communication constraints, and analyze the effects of the constraints on the overall system performance. The data available to the controller will be constrained by the rate restrictions, which is to be less than the bandwidth of the channel. Thus the problem can be regarded as an application of control with communication constraints to optimal control.

We consider the following scalar LTI system

$$x_{t+1} = ax_t + bu_t, \tag{7.1}$$

where x, u are the state and control variables. The control term here is a function of the estimate of the state. The structure of the system is depicted in Figure 7.1.

The goal here is to solve an LQR problem in a remote LTI system connected via a digital noiseless channel with a bandwidth B. The LQR cost is

$$J = E[\sum_{k=t_0}^{\infty}(|x_k|_Q^2 + |u(\hat{x}_k)|_R^2)], \tag{7.2}$$

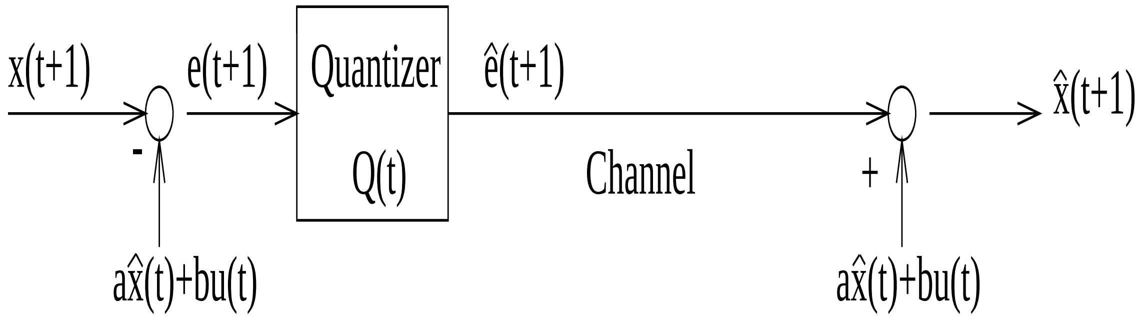

Figure 7.1 The quantization scheme, where control is a function of the estimate.

where $Q > 0, R > 0$, and the control u is only a function of the quantized state $\hat{x}$ available at the receiver at the control side. Furthermore, the rate to be used is upper bounded by the available channel bandwidth. We further assume that the quantization error is uniform, as is motivated by Lemma 2.2.

We start off the analysis with a dynamic quantization problem.

7.2 A Dynamic Quantization Problem

Let us consider the following problem, with a dynamic aggregate quantization cost:

$$\sum_k a^{2k} E[(x_k - \hat{x}_k)^2], \tag{7.3}$$

where a is the LTI system coefficient and $\hat{x}$ is the quantized state. Since the quantization at any particular time instant affects the distortion in the next stages, the analysis of optimal quantization leads to a dynamic analysis. From Lemma 2.2, the quantization error is taken to be uniform. In case the initial error is not uniform one can use a high-rate quantizer initially to force the quantization error in the subsequent stages to be uniform.

Proposition 7.1 *If one uses high-rate quantization at the first stage, to minimize (7.3), Lloyd-Max quantization applied at the first stage followed by uniform quantization applied at the remaining stages, is the optimal quantization scheme, among the*

quantizers with a given number of levels K at each stage (equivalently, given the bit rate, under fixed length codeword assumption).
In case the initial input is uniform, uniform quantizer applied at every stage is the optimal quantization scheme.

Proof. Since high-rate is assumed, the quantization error conditioned on a specific bin will be uniform.

Given a fixed rate for quantization at each stage, let $J_{f_t(i)}$ denote the cost of the ith bin at time t, $p_i(t)$ denote the probability of occurrence of values in bin i at time t, and q_i^e as the quantization error for bin i. The cost will be

$$J_{f_t}(.) = \min_{Quan(.)} \sum_i p_i(t)[E[(q_i^e(t))^2] + a^2 \min_{Quan(t+1),Quan(t+2)..} J_{f_{t+1}(i)}].$$

Thus, there is a dynamic evolution, and the quantization values for a particular time will affect the quantization values at the future times.

If we expand the cost in terms of the bin edge levels $x_i(t)$, and reconstruction values $\hat{x}_i(t)$,

$$J_{f_t} = \min_{Quan(t)} \sum_i p_i(t) \frac{1}{p_i(t)} \int_{x_i(t)}^{x_{i+1}(t)} (x - \hat{x}_i(t))^2 f_t(x) dx$$
$$+a^2 p_i(t) \frac{(x_i(t) - x_{i-1}(t))^2}{12K^2} + a^4 p_i(t) \frac{(x_i(t) - x_{i-1}(t))^2}{12K^4} + \dots , \tag{7.4}$$

which finally yields

$$J_{f_t} = \min_{Quan(.)} \sum_i p_i(t) (\frac{1}{12}(x_i(t) - x_{i-1}(t))^2 (1 + \sum_{l=1}^{\infty} (\frac{a^2}{K^2})^l). \tag{7.5}$$

Since the quantizer which minimizes the term

$$\sum_i \int_{x_i(t)}^{x_{i+1}(t)} (x - q_i(t))^2 f_x(x) dx = \sum_i \frac{(x_i(t) - x_{i-1}(t))^2}{12}$$

is the Lloyd-Max quantizer, and since the cost is just a scalar multiple of the distortion in the first stage, (7.5) is minimized by the Lloyd-Max quantizer applied at the first stage, and by the uniform quantizer applied in the subsequent stages.

If we assume the quantization error has a uniform error distribution, since the Lloyd-Max quantizer for a uniformly distributed input is the uniform quantizer, the long-run distortion will be minimized by a uniform quantizer. Being analogous with (7.5), the proof follows from the fact that, for uniform distribution, the cost in the next stages can be expressed in terms of the cost in the first stage, and the cost in the next stage is a linear function of the cost of the first term. To see that uniform quantization is optimal for a uniform source, let $p_i = \frac{\Delta_i}{\Delta}$. Then the distortion will be

$$D = \frac{1}{12} \sum_i p_i \Delta_i^2 = \frac{\Delta^2}{12} \sum p_i^3. \tag{7.6}$$

Using a Lagrangian multiplier for the constraint $\sum p_i = 1$, and embedding this into the cost function, the optimality conditions for each p_i are identical. Thus, uniform quantizer is the the optimal quantizer. $\diamond$

Note that, in case the error does not become uniform after the first stage (which is the case if the initial error is not uniform and the quantization in the first stage is not a high-rate one), the Lloyd-Max quantizer is not the optimal quantizer, since the optimality conditions for the bin edges and the reconstruction values of the Lloyd-Max quantizer [16] would not be the same as the quantizer minimizing (7.3).

For instance, even for the two-step cost $E[(x(0) - \hat{x(0)})^2 + a^2(x(1) - \hat{x(1)})^2]$, the optimality condition for the lth threshold value, which can be obtained by equating the derivative of the expression with respect to the quantizer parameters to zero, will be

$$f_x(x_l)[(x_l - \hat{x}_l(0))^2 - (x_l - \hat{x}_{l-1}(0))^2 + a^2 p_l (x_l - \hat{x}_l(1))^2 - a^2 p_{l-1}(x_l - \hat{x}_{l-1}(1))^2$$
$$+ a^2(-J(f_l(1) + J(f_{l-1}(1)) = 0, \tag{7.7}$$

which is not the same equation as the Lloyd-Max optimality condition [16], since the optimal x_l is not the mid-point between $\hat{x}_{l-1}$ and $\hat{x}_l$.

We now return to the original problem of the chapter.

7.3 Solution to the LQR Problem

We state the following theorem for the solution of the LQR problem with communication constraints.

Theorem 7.1 *The quantizer, minimizing the cost function (7.2), under the constraint that it has a given fixed number of levels, is the uniform quantizer.*

Proof. Let us first consider the finite horizon problem:

$$J = E[|x_k|_{S_N}^2 + \sum_{k=t_0}^{N_1}(|x_k|_Q^2 + |u(\hat{x}_k)|_R^2)], \tag{7.8}$$

where $S_N \geq 0$. The relation between the state estimation and the actual state is

$$x_t = \hat{x}_t + q_t^e,$$

where q_t^e is the quantization error. In chapter 2, the recursion for the estimated state was found to be

$$\hat{x}_{t+1} = a\hat{x}_t + bu_t + \hat{e}_t,$$

where $\hat{e}_t$ is the quantized innovation term.

If the quantization error is assumed to be orthogonal to the quantizer outcome, the problem above can be formulated as a stochastic control problem, which is very well studied [34],[35].

The structure of the cost function will be quadratic at each time instant [34] with an additional noise term. At the last stage, N,

$$J_N = P_N x_N^2 = S_N \hat{x}_N^2 + w_N,$$

where $w(N) = S_N E[{q_N^e}^2]$ and $P_N = S_N$. The recursion for P_k is an Riccati equation:

$$P_k = Q + a^2 R P_{k+1}/(R + b^2 P_{k+1}). \tag{7.9}$$

And the optimal control is

$$u_k^* = -baP_{k+1}/(R + b^2 P_{k+1})\hat{x}_k.$$

The overall cost will have a structure as

$$J = P_0 \hat{x}_0^2 + E[\sum_{k \leq N} (P_k a^2 + Q) {q_{k-1}^e}^2]. \tag{7.10}$$

In the infinite horizon case, by setting $P_k = P_{k+1} = P$, we get the quadratic equation for P

$$P^2 + (R(1-a^2) - Qb^2)/b^2 P - QR/b^2 = 0. \tag{7.11}$$

Thus, the infinite horizon cost will become

$$J = P \hat{x}_0^2 + E[\sum_{k}^{\infty} (P a^2 + Q) {q_{k-1}^e}^2]. \tag{7.12}$$

We note that (7.16) has a structure similar to the problem introduced in Proposition 7.1. Thus the optimal quantizer is the uniform quantizer.

◇

One can use variable length encoding for the quantizer outputs, in which case, the data rate might be lower than $\log_2(K)$ where K is the number of levels used in the quantizer. As discussed in Chapter 2, the entropy of the variable can be used as a measure of the data rate required. Thus, the constraint on (7.2) will be

$$H(\hat{e}) \leq B. \tag{7.13}$$

Now we state the following theorem.

Theorem 7.2 *The time-invariant quantizer minimizing the cost problem (7.2), with the constraint that the quantizer satisfies (7.13) has the same structure as the distortion-constrained entropy minimizing quantizer introduced in Section 2.3.2.*

Proof. Since the constraint is a concave function of the quantizer parameters and (7.2) is a convex function, the optimal rate would be equal to the bandwidth, and a Lagrangian multiplier λ can adjoin the constraint, which is now an equality, into the cost to be minimized. Thus, we have

$$J = E[\sum_{k=t_0}^{\infty}(|x_k|_Q^2 + |u(\hat{x}_k)|_R^2) + \lambda H(\hat{e}_k)]. \tag{7.14}$$

Thus, the problem is to find the quantizer among the set of quantizers $Quan(.)$ satisfying the constraint (7.13) and minimizing the cost:

$$\min_{Quan(.)} \min_{u} E[\sum_{k=t_0}^{\infty}(|x_k|_Q^2 + |u(\hat{x}_k)|_R^2) + \lambda H(\hat{e}_k)]. \tag{7.15}$$

The cost, paralleling Theorem 7.1, becomes

$$J = P\hat{x}_0^2 + E[\sum_{k}^{\infty}(Pa^2 + Q){q_{k-1}^e}^2 + \lambda H(\hat{e}_k)]. \tag{7.16}$$

The problem here is an entropy-constrained distortion minimization. However, as was argued in Section 2.3.2, the distortion-constrained entropy minimization and the entropy-constrained distortion minimization problems are duals of each other. Thus, the cost function to be minimized has a structure similar to the distortion-constrained entropy minimizing quantizer proposed in Section 2.3.2 .

$\diamond$

Remarks. The above analysis can be generalized to multidimensional systems, since the orthogonality property holds for higher dimensions as well. However, note that the vector quantization implementations have a high degree of complexity, especially for the cases where the quantizer is not uniform. $\diamond$

7.4 Summary

In this chapter we considered an LQR problem with communication constraints. A dynamic quantization problem has been introduced, solved, and is used together with the principle of separation and estimation to provide a solution to the LQR problem.

CHAPTER 8

CONCLUSION

In this thesis, we have considered the general problem of state estimation and control for LTI systems where the controller/decision-maker and the plant are connected via communication channels with different characteristics.

For the case where the channel is a digital noiseless channel, the problem has been formulated as a quantization problem. A forward-looking scheme and a monotonic variance achieving scheme have been constructed, and their optimality have been established using information theoretic arguments. It has been shown that the system has a higher performance in terms of the bit rate with forward-looking criterion for stochastic systems, whereas for noisy systems the rate requirements are identical under both criteria. A time-invariant distortion-constrained entropy minimizing quantizer has been introduced as a stabilizing optimal quantizer.

The analysis was then extended to channels which have packet drop characteristics, such as the Internet. Here the main difference is the nondeterministic nature of the signal sent; thus, the problem is not merely a quantization problem. The relation between the sampling rate, quantization rate and the packet loss probability has been investigated. Rates required to achieve stability in state estimation error and the state itself under different measures have been obtained.

Subsequently, the problem of state estimation and control in centralized higher dimensional systems and decentralized systems has been considered. Centralized and decentralized systems have been shown to entail different optimality conditions, and it has been shown using information and quantization theoretic arguments that the degree of the information sharing between decentralized controllers is crucially important.

Finally in Chapter 7 the dynamic nature of quantization and control problems in LTI systems have been exposed, and an LQR problem for an LTI system with digital noiseless channels has been investigated. This might also be regarded as an application of the results in the earlier chapters to a control problem.

Extensions of the work presented in this thesis to binary noisy channels and additive white Gaussian channels are currently under study and are not included in this thesis. Quantization of control signals is a problem in its own, and requires a separate treatment. Another topic of interest and relevance is game theoretic analysis for optimal quantization and decentralized detection and control.

REFERENCES

[1] S. Yüksel and T. Başar, "Quantization for LTI systems with noiseless channels," in *IEEE Conference on Control Applications*, Istanbul, June 2003, pp. 423-429.

[2] S. Yüksel and T. Başar, "State estimation and control for LTI systems over communication networks," in *IEEE Conference on Control Applications*, Istanbul, June 2003, pp. 409-414.

[3] S. Yüksel and T. Başar, "Quantization and coding for decentralized LTI systems," to appear in *IEEE Conference on Decision and Control*, Hawai, December 2003.

[4] D. F. Delchamps, "Stabilizing a linear system with quantized state feedback," *IEEE Trans. Automatic Control*, vol.35, pp. 916-924, Aug 1990.

[5] W. S. Wong and R. W. Brockett, "Systems with finite communication bandwidth constraints - part I: State estimation problems," *IEEE Trans. Automatic Control*, vol. 42, no. 9, pp. 1294-1299, Sept. 1997.

[6] W. S. Wong and R. W. Brockett, "Systems with finite communication bandwidth constraints - part II: Stabilization with limited information feedback, *IEEE Trans. Automatic Control*, vol. 44, pp. 1049-1052, 1999.

[7] S. Tatikonda, A. Sahai, and S. K. Mitter, "Control of LQG systems under communication constraints," in *Proc. 37th IEEE Conf. Decision and Control*, Tampa, Florida, Dec. 1998, pp. 1165-1170.

[8] G. N. Nair and R. J. Evans, "Exponential stabilisability of multidimensional linear systems with finite data rates," in *Proc. IFAC World Congress*, Barcelona, Spain, July 2002.

[9] H. Ishii and T. Başar, "Remote control of LTI systems over networks with state quantization," in *Proc. 41st IEEE Conf. Decision and Control*, Las Vegas, Nevada, Dec. 2002 pp. 830-835.

[10] D. Liberzon, "On stabilization of linear systems with limited information," *IEEE Trans. Automatic Control*, vol. 48, pp. 304 -307, Feb 2003.

[11] A. Sahai, "Evaluating channels for control: capacity reconsidered," in *Proc. American Control Conference*, Chicago, Illinois, June 2000, pp. 2358-2362.

[12] N. S. Jayant and P. Noll, *Digital Coding of Waveforms.* Englewood Cliffs, NJ: Prentice Hall, 1984.

[13] R. M. Gray and D. L. Neuhoff, "Quantization," *IEEE Trans. Information Theory*, vol. 44, pp. 2325-2384, October 1998.

[14] R. M. Gray, *Source Coding Theory.* Amsterdam:Kluwer, 1990.

[15] N. Elia and S. Mitter. "Quantization of linear systems," in *Proc. of the 38th IEEE Conference on Decision and Control*, Phoenix, AZ, December 1999.

[16] S. P. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Information Theory*, vol. 28, pp. 129-137, 1982.

[17] J. Max, "Quantizing for minimum distortion," *IEEE Trans. Information Theory*, vol. 8, pp. 7-12, March 1960.

[18] D. L. Neuhoff and R. K. Gilbert, "Causal source codes," *IEEE Trans. Information Theory*, vol. 28, pp. 701-713, Sept. 1982.

[19] J. C. Walrand and P. Varaiya, "Optimal causal coding-decoding problems," *IEEE Trans. Information Theory*, vol. 19, pp. 814-820, Nov. 1983.

[20] H. S. Witsenhausen, "On the structure of real-time source coders," *Bell Syst. Tech. J.*, vol. 58, pp. 1437-1451, Jul./Aug. 1979.

[21] T. Weissman and N. Merhav, "On causal source codes with side information," in *40th Annual Allerton Conference on Communication, Control, and Computing*, University of Illinois, October 2002.

[22] "Principles of information theory," class notes for ECE 463, Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Fall 2002.

[23] A. György and T. Linder, "Optimal entropy-constrained scalar quantization of a uniform source," *IEEE Trans. Information Theory*, vol. 46, pp. 2704-2711, Nov 2000.

[24] N. Farvardin and J. W. Modestino, "Optimum quantizer performance for a class of non-Gaussian memoryless sources," *IEEE Trans. Information Theory*, vol. 30, pp. 485-497, May 1984.

[25] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression.* Boston:Kluwer, 1992.

[26] H. Gish and J. N. Pierce, "Asymptotically efficient quantization," *IEEE Trans. Information Theory*, vol. 14, pp. 676-683, Sept. 1968.

[27] D. Bertsekas and R. Gallager, *Data Networks.* New Jersey:Prentice-Hall, Inc., 1992.

[28] J. Kurose and K. Ross, *Computer Networking: A Top-Down Approach Featuring the Internet.* Boston:Addison-Wesley 2000.

[29] K. Fujimoto, S. Ata, and M. Murata, "Statistical analysis of packet delays in the Internet and its application to playout control for streaming applications," *IEICE Transactions on Communications*, vol. E84-B, no.6, pp. 1504-1512, June 2001.

[30] T. Linder, "Learning-theoretic methods in vector quantization," lecture notes for the Advanced School on the Principles of Nonparametric Learning, Udine, Italy, July 9-13, 2001.

[31] J. D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 471-480, Jul. 1973.

[32] A. D. Wyner, "On source coding with side information at the decoder," *IEEE Trans. Information Theory*, vol. 21, pp. 244-300, May 1975.

[33] J. Kusuma, L. Doherty, and K. Ramchandran, "Distributed compression for sensor networks," in *IEEE Information Conference on Image Processing*, Thessaloniki, Greece, October 2001.

[34] T. Başar and P. Bernhard, *H-infinity Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach.* Boston:Birkhäuser, August 1995.

[35] A. P. Sage and C. C. White, *Optimum Systems Control.* New Jersey:Prentice Hall, 1977.