

Control under Communication Constraints

Sekhar Tatikonda, *Member, IEEE*, and Sanjoy Mitter, *Fellow, IEEE*,

Abstract—There is an increasing interest in studying control systems employing multiple sensors and actuators that are geographically distributed. Communication is an important component of these distributed and networked control systems. Hence there is a need to understand the interactions between the control components and the communication components of the distributed system.

In this paper we formulate a control problem with a communication channel connecting the sensor to the controller. Our task involves designing the channel encoder and channel decoder along with the controller to achieve different control objectives. We provide upper and lower bounds on the channel rate required to achieve these different control objectives. In many cases these bounds are tight. In doing so we characterize the “information complexity” of different control objectives.

Index Terms—Linear control, Communication, Distributed systems, Networked control

I. INTRODUCTION

IN this paper we study linear, discrete time, control problems under communication constraints. Communication is an important component of distributed and networked control systems. Hence there is a need to understand the fundamental relationship between how the control parts and the communication parts of the distributed system interact. A recent report on future research directions in control listed understanding control over communication networks as a major challenge for the controls field [10].

We examine observability and stabilizability under a communication constraint. The communication constraint is modelled as a discrete time, noiseless, digital channel connecting the sensor to the controller. For each time step this channel is capable of transmitting R bits without error. Our goal is to determine the minimum rate required on the channel to achieve our control objectives. In doing so we characterize the “information complexity” of different control objectives. This paper presents a framework for treating communication issues in control problems.

This task entails specifying what the encoder, decoder, and controller, know and when they know it. This specification is called the information pattern [19]. We distinguish between two information patterns. In the first information pattern we assume that the encoder has access to the control signals being applied to the plant. This may occur, for example, when the encoder is co-located with the plant. In the second information pattern we assume that the encoder is geographically separated

from the plant and hence observes the state, or output, but not the control signals being applied. We discuss the important role different information patterns can have on the control design.

We compute a lower bound on the rate required to achieve the different control objectives. This lower bound is independent of the information pattern in place and depends only on the plant. Specifically we show that a necessary condition on the rate for asymptotic observability and asymptotic stabilizability in a linear, discrete time, system is

$$R > \sum_{\lambda(A)} \max\{0, \log |\lambda(A)|\}$$

where the sum is over the eigenvalues of the A matrix (see equation (1).) This result relates the speed of the dynamics of the plant to the information rate of the channel. In the case where the encoder observes the control signals, the co-location case, we show that for any rate larger than this lower bound there exists an encoder, decoder, and controller that achieves the control objective. Hence this bound is also sufficient. We then discuss upper bounds for systems where the encoder is geographically separated from the plant and hence does not have access to the control signals.

Typical communication channels are noisy and have delays. A complete understanding of the interaction between control and communication will need to use tools from both control theory and information theory. A necessary step in developing such a theory requires understanding the interaction between control and noiseless channels with bit rate constraints. This is the case examined in this paper. In general one is interested in the relationship between control performance and communication rate. Understanding observability and stabilizability under communication constraints is an important initial step towards that larger goal.

In our companion paper we extend these results [16] by showing the strong connection between control and the traditional information theoretic problems of source coding and channel coding. We present a general necessary condition for observability and stabilizability for a large class of noisy communication channels. Then we study sufficiency conditions for Internet-like channels that suffer erasures.

Our problem formulation was inspired by the work of Borkar and Mitter [5]. There they formulated an LGQ control problem under a rate-constrained channel. They discussed the interactions between coding, delay, and performance. The papers of Wong and Brockett were also influential [20] [21]. They provide conditions connecting the channel rate to the dynamics of the system to insure stabilizability of the system. We strengthen and generalize these results. Specifically we present necessary and sufficient conditions on the channel for the multivariate case. Furthermore we examine the role

Manuscript received March 12 2002; first revision April 15, 2003; second revision October 30, 2003. Support was provided by the Army Research Office under the MURI Grant: Data Fusion in Large Arrays of Microsensors DAAD19-00-1-0466 and the Department of Defense MURI Grant: Complex Adaptive Networks for Cooperative Control Subaward #03-132.

S. Tatikonda is with Yale University. S. Mitter is with the Massachusetts Institute of Technology.

different information patterns have in determining the channel rate.

Elia and Mitter examined the stabilizability problem in the case where the encoder is time-invariant [7]. They use a Lyapunov based synthesis scheme to design the underlying quantizers. Our work differs in that we allow for time-varying encoders. Liberzon and Brockett have also examined a Lyapunov based design with time-varying encoders [4]. They give upper bounds on the rate required to achieve stabilizability. A time-varying coding scheme is also presented in [13]. In this paper we not only present bounds but also present conditions under which these bounds are tight.

Nair and Evans give rate conditions for stabilizability of an ARMA model [11]. In [12] they provide a rate condition similar to ours for insuring convergence. In contrast, our formulation allows us to analyze systems with process disturbances and different information patterns. Baillieul presents sufficient rate conditions for the multivariate case with a single input and an A matrix with only real distinct eigenvalues [2],[3]. Here we present a general framework for analyzing the multivariate state-space case with multiple inputs and arbitrary A matrix. This framework allows us to prove necessary and sufficient rate conditions, treat different information patterns, and design encoding/decoding schemes that can treat output observations and are robust to process disturbances. Furthermore this framework lends itself to the treatment of control over more complicated communication channels as discussed in [16]. The work here can be found in the first author's Ph.D. thesis [14] and has appeared in preliminary form in [17] and [15].

We now make four observations that will motivate our analysis.

Observation 1: Why feedback?

If there is no uncertainty in the initial position, no uncertainty in the plant dynamics, and there are no process disturbances then one can achieve most control objectives using an open loop controller. A closed loop controller for the same problem is often less complex to realize. Furthermore a closed loop controller can more robustly deal with the aforementioned uncertainties in initial position, plant dynamics and process disturbances. Thus the point of feedback, if we bar complexity considerations, is to transmit from the plant to the controller information about the uncertainty in the state of the plant and the plant uncertainty itself that the controller does not know. The question then becomes what information is relevant and what communication scheme should be used to transmit that information.

Observation 2: Full observation performance

If the observation mechanism is instantaneous and the communication link is a lossless, infinite bandwidth, channel then we call the observation a full observation. We assume the control objective of interest is achievable under full observation. Clearly if an objective cannot be achieved under full observation it cannot be achieved under the rate constrained observation. Equivalently if a control objective can be achieved under a rate constrained observation then it can be achieved under full observation.

Observation 3: Number of control sequences

In a time horizon T the decoder will receive one of at most

2^{TR} channel symbol sequences. If the encoder, decoder, and controller are also deterministic then the number of different possible control sequences in this time must be smaller than or equal to 2^{TR} . Intuitively then a control objective under rate R can be achieved only if we can approximate well the control sequences for the full observation problem by one of only 2^{TR} control sequences. Thus in terms of the underlying quantization problem one may think of quantization as living in the control sequence space.

Observation 4: State estimation error

The choice of channel rate influences the level of state estimation error in our observer. If the state estimation error increases with time in an unbounded fashion there will come a point when we can no longer satisfy the control objective. In this case there is essentially no useful feedback. Thus unless the control objective can be achieved via an open loop controller, i.e. a controller without access to any feedback, we cannot hope to achieve the control objective. Hence, in a rough sense, the state estimation error should grow at a slower rate than the state dynamics. We are interested in characterizing the largest tolerable level of state estimation error that still insures that the control objective is satisfied.

We conclude this introduction with a summary of the paper. In section two we formulate the problem. In section three we provide necessary conditions on the channel rate required to achieve observability and stabilizability. These bounds are independent of the information pattern chosen. Our next objective is to provide schemes that can achieve this bound. This depends heavily on the choice of information pattern. In section four we discuss encoders with different information patterns. We conclude section four with an explicit quantizer construction and prove our main technical lemma.

In section five we treat encoder class one. This is the class where the encoder has access to the control signals. Here we show that the lower bounds provided in section three can be attained. Specifically we provide a scheme that achieves the control objectives. In section six we examine encoder class two. This encoder class has a more realistic information pattern. The rates, though, required to achieve the control objectives are larger than those in encoder class one. We conclude in section seven.

II. PROBLEM SETUP

We consider the following linear time-invariant system:

$$X_0 \in \Lambda_0, \quad X_{t+1} = AX_t + BU_t, \quad Y_t = CX_t, \quad \forall t \geq 0 \quad (1)$$

where $\{X_t\}$ is a \mathbb{R}^d -valued state process, $\{U_t\}$ is a \mathbb{R}^m -valued control process, and $\{Y_t\}$ is a \mathbb{R}^l -valued observation process. We have $A \in \mathbb{R}^{d \times d}$, $B \in \mathbb{R}^{d \times m}$, and $C \in \mathbb{R}^{l \times d}$. The initial position $X_0 \in \Lambda_0$ where $\Lambda_0 \subseteq \mathbb{R}^d$ is assumed to be an open set. If $C = I$, where I is the identity matrix, then we have full state observation at the encoder. See Figure 1.

The communication channel is modelled as a noiseless digital channel that can transmit at each time step one of 2^R symbols denoted $\sigma \in \Sigma$, $|\Sigma| = 2^R$. Specifically at each time step the channel can transmit without error R bits of information. Throughout this paper log refers to logarithm base 2.

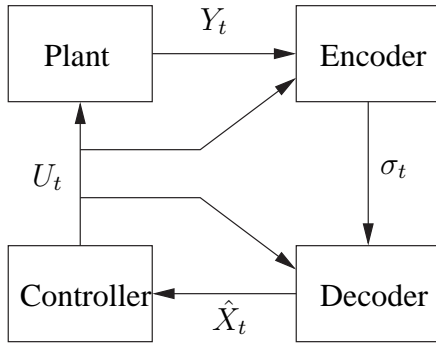


Fig. 1. System

A. Information Pattern

The control problems we look at involve the design of an encoder, decoder, and controller. Here we specify the information pattern of each component. Let $X^t \doteq (X_0, \dots, X_t)$.

a) *Encoder*:: The encoder at time t is a map

$$\mathcal{E}_t : \mathbb{R}^{l(t+1)} \times \Sigma^t \times \mathbb{R}^{mt} \rightarrow \Sigma \text{ taking } (Y^t, \sigma^{t-1}, U^{t-1}) \mapsto \sigma_t.$$

In section 4 we will discuss different restrictions on the available inputs to the encoder. The available inputs to the encoder are commonly called the *information pattern* of the encoder [19]. For example the encoder may not have access to the past controls.

b) *Decoder*:: The decoder at time t is a map

$$\mathcal{D}_t : \Sigma^{t+1} \times \mathbb{R}^{mt} \rightarrow \mathbb{R}^d \text{ taking } (\sigma^t, U^{t-1}) \mapsto \hat{X}_t.$$

The output of the decoder is an estimate of the state of the plant. We will discuss how this state estimate is computed in section 4.

c) *Controller*:: The controller at time t is a map

$$\mathcal{C}_t : \mathbb{R}^d \rightarrow \mathbb{R}^m \text{ taking } \hat{X}_t \mapsto U_t.$$

Note that we are assuming that the controller takes as input only the decoder's state estimate. Hence we are assuming a separation structure between the decoder and the controller. We will show that for encoders in encoder class one, to be defined in section 4, there is no loss of generality in making this separation assumption.

III. LOWER BOUNDS THAT ARE INDEPENDENT OF THE INFORMATION PATTERN

We now examine necessary conditions on the channel rate to insure observability and stabilizability. Note that the usual algebraic conditions for observability, e.g. certain Grammians having full rank, are still necessary but no longer sufficient. The additional necessary conditions take the form of lower bounds on the channel rate. These lower bounds will be universal in the sense that they hold independently of the actual encoder, decoder, and controller used. They hold independently of the information pattern in place.¹ We will show

¹Note the analogy with Fano's inequality used in converse theorems in information theory [6]. Fano's inequality holds independently of the actual encoder and decoder used.

in section five that there exists a particular information pattern, described in the definition of encoder class one, such that these lower bounds can be achieved. Thus these lower bounds are tight.

A. Observability

The purpose of any good observer is to distinguish points in the state space. In a time horizon of T we have at most 2^{TR} possible symbols arriving at the decoder. Thus at time T we must be able to approximate the state by one of at most 2^{TR} points.

Definition 3.1: Let the error be $e_t = X_t - \hat{X}_t$ where \hat{X}_t is the state estimate. System (1) is *asymptotically observable* if there exists an encoder and decoder such that the following hold for any control sequence $\{U_t\}$:

- (1) *Stability*: $\forall \epsilon > 0, \exists \delta(\epsilon)$ such that $\|X_0\|_2 \leq \delta(\epsilon)$ implies $\|e_t\|_2 \leq \epsilon, \forall t \geq 0$.
- (2) *Uniform attractivity*: $\forall \epsilon > 0, \forall \delta > 0 \exists T(\epsilon, \delta)$ such that $\|X_0\|_2 \leq \delta$ implies $\|e_t\|_2 \leq \epsilon, \forall t \geq T(\epsilon, \delta)$.

Point one states that the error cannot grow without bound for bounded X_0 . The second point states that the error decreases to zero uniformly in X_0 . Uniform attractivity is defined for all δ . Thus our definition of asymptotic observability is global.

Traditional definitions of observability state that given enough time one can identify the initial condition exactly. Then once you know the initial condition and the controls you can compute the state at any time in the future. In our case, at time t we can only distinguish between 2^{tR} initial positions. Hence there will be a certain amount of error in our state estimate of X_0 . This error will propagate (due to the unstable modes of the system) in our estimate of any future X_t . It is for this reason that we introduce this definition of asymptotic observability.

Now we are prepared to give a necessary condition on the rate required to achieve asymptotic observability. Let $\sum_{\lambda(A)}$ represent the sum over the eigenvalues of A . For a given set $\Omega \subseteq \mathbb{R}^d$ define the $\text{diam}(\Omega) \doteq \sup_{x,y \in \Omega} \|x - y\|_2$.

Proposition 3.1: A necessary condition for system (1) to be asymptotically observable is $R \geq \sum_{\lambda(A)} \max\{0, \log |\lambda(A)|\}$.

Proof: Assume without loss of generality that the initial uncertainty contains the bounded set $\Omega_0 = \{X : \|X\|_2 \leq L\} \subseteq \Lambda_0$.

Note, possibly after a coordinate transformation, that the matrix A can be written in the form

$$\begin{bmatrix} A_s & \\ & A_u \end{bmatrix}$$

where the A_s block corresponds to the stable subspace (that subspace corresponding to the eigenvalues of A that are strictly inside the unit circle) and the A_u block corresponds to the marginally stable and unstable subspace (that subspace corresponding to the eigenvalues of A that are either on the unit circle or outside the unit circle.) Let Π_s represent the projection onto the stable subspace.

Fix an arbitrary control sequence $\{U_t\}$. Then $X_t = A^t X_0 + \alpha_t$ where $\alpha_t = \sum_{j=0}^{t-1} A^{t-1-j} B U_j$. For any control sequence

we have

$$\lim_{t \rightarrow \infty} \Pi_s(X_t - \alpha_t) = 0.$$

Thus knowledge of the control signals alone is enough to estimate the projection of the state onto the stable subspace. Hence, without loss of generality, we can restrict our attention to A matrices that contain only marginally stable and unstable eigenvalues.

The set of points that X_t can take contains the following set

$$\Omega_t \doteq \{X : X = A^t X_0 + \alpha_t \quad \text{for some } X_0 \in \Omega_0\}.$$

If the system is asymptotically observable then it must be the case that $\forall \epsilon > 0$ and $\forall X_0 \in \Omega_0$ there is a $T(\epsilon, L)$ such that for $t \geq T(\epsilon, L)$ we have $\|e_t\|_2 \leq \epsilon$. In particular this must hold for $\epsilon < L$. A lower bound on the rate can be computed by counting the number of regions of diameter less than 2ϵ it takes to cover Ω_t for $t \geq T(\epsilon, L)$.

If the diameter of a set $\Omega \subseteq \mathbb{R}^d$ is less than 2ϵ then the volume of that set must be less than $K_d \epsilon^d$ (where K_d is the constant in the formula for the volume of a sphere in \mathbb{R}^d .) Thus to cover Ω_t by regions of diameter 2ϵ we require at least

$$\begin{aligned} R &\geq \frac{1}{t} \log \frac{\text{volume}(\Omega_t)}{K_d \epsilon^d} \\ &\stackrel{(a)}{=} \frac{1}{t} \log \frac{|\det(A^t)| \text{volume}(\Omega_0)}{K_d \epsilon^d} \\ &= \frac{1}{t} \log |\det(A^t)| + \frac{1}{t} \log \frac{K_d L^d}{K_d \epsilon^d} \\ &= \sum_{\lambda(A)} \log |\lambda(A)| + \frac{d}{t} \log \frac{L}{\epsilon} \end{aligned}$$

where (a) follows from theorem 10.38 in [1]. Note that since $\epsilon < L$ the term $\frac{d}{t} \log \frac{L}{\epsilon}$ is positive. Thus

$$R \geq \sum_{\lambda(A)} \max\{0, \log |\lambda(A)|\}$$

is a necessary condition for asymptotic observability. \square

The following corollary shows that the rate bound in proposition 3.1 continues to be necessary to insure bounded error in the case when there are additive process disturbances.

Corollary 3.1: Consider the case of bounded additive process disturbances: $X_{t+1} = AX_t + BU_t + W_t$ where $\|W_t\|_2 \leq D$. The rate $R \geq \sum_{\lambda(A)} \max\{0, \log |\lambda(A)|\}$ is necessary to insure that the state estimation error $\limsup_{t \rightarrow \infty} \|e_t\|_2$ is bounded.

Proof: As in proposition 3.1 we can assume without loss of generality that the initial uncertainty set contains the bounded set $\Omega_0 = \{X : \|X\|_2 \leq L\} \subseteq \Lambda_0$. Also we can partition the matrix A into the A_s and A_u blocks. Let Π_s represent the projection onto the stable subspace.

Fix an arbitrary control sequence $\{U_t\}$. Then $X_t = A^t X_0 + \sum_{j=0}^{t-1} A^{t-1-j} B U_j + \alpha_t$ where $\alpha_t = \sum_{j=0}^{t-1} A^{t-1-j} B U_j$. For

any control sequence we have

$$\begin{aligned} &\limsup_{t \rightarrow \infty} \|\Pi_s(X_t - \alpha_t)\|_2 \\ &= \limsup_{t \rightarrow \infty} \left(\|\Pi_s A^t\| L + \left\| \Pi_s \sum_{j=0}^{t-1} A^{t-1-j} \right\|_2 \|B\|_2 D \right) \\ &< \infty. \end{aligned}$$

Because the initial position X_0 and the disturbances W_t are bounded knowledge of the control signals alone is enough to insure that the error in the estimate of the projection of the state onto the stable subspace remains bounded. Thus, without loss of generality, we can restrict our attention to A matrices that contain only marginally stable and unstable eigenvalues.

The set of points that X_t can take contains the following set

$$\begin{aligned} \Omega_t &\doteq \{X : X = A^t X_0 + \sum_{j=0}^{t-1} A^{t-1-j} B W_j + \alpha_t, \\ &\quad \forall X_0 \in \Omega_0, \|W_t\|_2 \leq D\}. \end{aligned}$$

Let $\Omega_t^{\text{zero}} = \{X : X = A^t X_0 + \alpha_t \quad \text{for some } X_0 \in \Omega_0\}$ be the set of points $X_t \in \Omega_t$ where all the disturbances are set to zero.

If there exists an encoder and decoder such that the estimation error is bounded then there must be a $\beta < \infty$ and a $T(L)$ such that $\|e_t\|_2 \leq \beta$ for $t \geq T(L)$. A lower bound on the rate can be computed by counting the number of regions of diameter less than 2β it takes to cover Ω_t for $t \geq T(L)$. Thus we require a rate of at least

$$\begin{aligned} R &\geq \frac{1}{t} \log \frac{\text{volume}(\Omega_t)}{K_d \beta^d} \\ &\stackrel{(a)}{\geq} \frac{1}{t} \log \frac{\text{volume}(\Omega_t^{\text{zero}})}{K_d \beta^d} \\ &\stackrel{(b)}{\geq} \sum_{\lambda(A)} \log |\lambda(A)| + \frac{d}{t} \log \frac{L}{\beta} \end{aligned}$$

where (a) follows because $\Omega_t^{\text{zero}} \subset \Omega_t$ and (b) follows from proposition 3.1. Note that $\frac{d}{t} \log \frac{L}{\beta}$ is bounded and becomes negligible as $t \rightarrow \infty$. \square

Note that in the preceding proposition the necessary condition only depends on the uniform attractivity condition and not on the stability condition in the definition of asymptotic observability. Similarly, in the preceding corollary the necessary condition only depends on the boundedness condition.

B. Stabilizability

In this section we discuss stabilizability under a rate constraint. The lower bound uses a counting argument similar to that given in proposition 3.1.

Definition 3.2: System (1) is *asymptotically stabilizable* if there exists an encoder, decoder, and controller such that

- (1) **Stability:** $\forall \epsilon > 0, \exists \delta(\epsilon)$ such that $\|X_0\|_2 \leq \delta(\epsilon)$ implies $\|X_t\|_2 \leq \epsilon, \forall t \geq 0$.
- (2) **Uniform attractivity:** $\forall \epsilon > 0, \delta > 0 \exists T(\epsilon, \delta)$ such that $\|X_0\|_2 \leq \delta$ implies $\|X_t\|_2 \leq \epsilon, \forall t \geq T$.

Point one states that the state cannot grow unbounded for bounded X_0 . The second point states that the state decreases to zero uniformly in X_0 .

Proposition 3.2: Assume (A, B) is a stabilizable pair. A necessary condition for system (1) to be asymptotically stabilizable is $R \geq \sum_{\lambda(A)} \max\{0, \log |\lambda(A)|\}$.

Proof: As in proposition 3.1 we can assume without loss of generality that the initial uncertainty set contains the bounded set $\Omega_0 = \{X : \|X\|_2 \leq L\} \subseteq \Lambda_0$ and that the matrix A has only unstable and marginally stable eigenvalues.

For a given control sequence U_0, U_1, \dots, U_{t-1} we have

$$X_t = A^t X_0 + \sum_{i=0}^{t-1} A^{t-1-i} B U_i.$$

If the system is asymptotically stabilizable then it must be the case that $\forall \epsilon > 0$ and $\forall X_0 \in \Omega_0$ there exists a $T(\epsilon, L)$ such that $\forall t \geq T(\epsilon, L)$ we have $\|X_t\|_2 \leq \epsilon$. In particular this must hold for $\epsilon < L$. For this value of ϵ define the sets Γ , parameterized by the control sequences U_0, \dots, U_{t-1} , to be

$$\Gamma_{U_0^{t-1}} = \{X_0 : \|X_t\|_2 \leq \epsilon\}.$$

Note that X_t depends linearly on U_0, \dots, U_{t-1} hence all the Γ sets are linear translations of each other. Thus each Γ set has the same volume: $\text{volume}(\Gamma) = |\det A^{-t}| K_d \epsilon^d$. A lower bound on the rate can be computed by counting how many Γ sets it takes to cover Ω_0 . Now

$$\begin{aligned} R &\geq \frac{1}{t} \log \frac{\text{volume}(\Omega_0)}{\text{volume}(\Gamma)} \\ &= \frac{1}{t} \log \frac{K_d L^d}{|\det A^{-t}| K_d \epsilon^d} \\ &= \sum_{\lambda(A)} \log(|\lambda(A)|) + \frac{d}{t} \log\left(\frac{L}{\epsilon}\right) \end{aligned}$$

Since $\epsilon < L$ the term $\frac{d}{t} \log \frac{L}{\epsilon}$ is positive. We may conclude that

$$R \geq \sum_{\lambda(A)} \max\{0, \log |\lambda(A)|\}$$

is a necessary condition for asymptotic stabilizability. \square

IV. ENCODER CLASSES, PRIMITIVE QUANTIZERS, AND THE KEY TECHNICAL LEMMA

In this section we define two different encoder classes of interest, then we present the primitive quantizer, and we end with a statement and proof of the key technical lemma.

A. Encoder Classes

Recall that the encoder at time t is a map \mathcal{E}_t that takes $(Y^t, \sigma^{t-1}, U^{t-1}) \mapsto \sigma_t$. In this case the encoder knows the past outputs, past channel symbols, and past controls. In some cases it is unreasonable to allow the encoder access to the past controls. For example the encoder may be geographically separated from the plant.

At one extreme one can consider an encoder with access to all the past information $(Y^t, \sigma^{t-1}, U^{t-1})$. At the other extreme one can consider an encoder with access to only Y_t . There are

many cases in between. We will provide an encoder, decoder, and controller for both of these two extreme cases. It will turn out that the encoder with access to the control signals has a lower rate requirement than the encoder without access to the control signals. These two cases shed light on the importance of the information pattern at the encoder for determining the channel rates required to achieve different control objectives.

Encoder Class One

In this class the encoder at time t is a map, \mathcal{E}_t , that takes $(Y^t, \sigma^{t-1}, U^{t-1}) \mapsto \sigma_t$. The decoder at time t is a map, \mathcal{D}_t , that takes $(\sigma^t, U^{t-1}) \mapsto \hat{X}_t$. We assume that both the encoder and decoder have knowledge of the dynamics of the plant. Furthermore we assume that the encoder knows \mathcal{D}_t and the decoder knows \mathcal{E}_t . We do not assume that the encoder or decoder knows the controller maps \mathcal{C}_t .

Encoder Class Two

In this class the encoder at time t is a map, \mathcal{E}_t , taking $Y_t \mapsto \sigma_t$. The encoder does not know the values of the control signals. We will assume, though, that it knows the control policy \mathcal{C}_t . The decoder at time t is a map that takes $(\sigma^t, U^{t-1}) \mapsto \hat{X}_t$. Finally we assume that the encoder knows \mathcal{D}_t and the decoder knows \mathcal{E}_t .

B. Equi-memory

The encoder and decoder need to work together. The job of the decoder upon receiving the channel symbol σ_t is to recover the quantization region the observation Y_t fell into. To achieve this the decoder needs to know what the encoder operation is. Knowledge of the map \mathcal{E}_t is not enough to insure this. To remedy this we now introduce the notion of equi-memory [8].

We define a state for the encoder and decoder. This requires specifying the information pattern [19] of each encoder and decoder. The maximal amount of information the encoder can observe at time t is $(Y^{t-1}, \sigma^{t-1}, U^{t-1}) \in \mathbb{R}^{lt} \times \Sigma^t \times \mathbb{R}^{mt}$. Let the information pattern of the encoder be $I_t \subset \mathbb{R}^{lt} \times \Sigma^t \times \mathbb{R}^{mt}$. For the decoder let $J_t = (\sigma^{t-1}, U^{t-1}) \in \Sigma^t \times \mathbb{R}^{mt}$. Finally let $\Omega_{I_t, t}(\sigma) \doteq \{Y_t : \mathcal{E}_t(Y_t, I_t) = \sigma\}$.

Definition 4.1: An encoder/decoder pair are said to be *equi-memory* if the information (J_t, σ_t) is sufficient to determine the set $\Omega_{I_t, t}(\sigma_t)$. Specifically there exists a map taking $(\sigma^t, U^{t-1}) \mapsto \Omega_{I_t, t}(\sigma_t)$.

In words this definition states that the information in J_t along with the quantization symbol σ_t is sufficient to determine which particular quantization region the observation Y_t fell into. We will assume throughout this paper that the encoder/decoder pairs we use are equi-memory. For encoder class one this means that $I_t = (Y^{t-1}, \sigma^{t-1}, U^{t-1})$ and hence the primitive quantizer should be chosen on the basis of the information (σ^{t-1}, U^{t-1}) . For encoder class two this means that $I_t = \emptyset$ and hence the primitive quantizer should not depend on any of the control or channel signals.

C. Primitive Quantizer

The encoders in both encoder classes will be restricted to apply the following primitive quantizer at each time step.

Definition 4.2: A *primitive quantizer* is a four-tuple $(c, \underline{R}, \underline{L}, \Phi)$ where $c \in \mathbb{R}^d$ represents the centroid, $\underline{R} =$

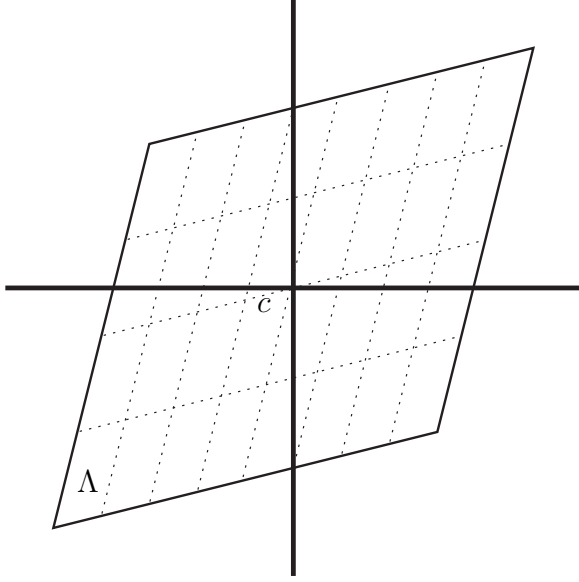


Fig. 2. Primitive Quantizer

$(R_1, \dots, R_d)' \in \mathbb{R}^{d,+}$ represents the rate vector, $\underline{L} = (L_1, \dots, L_d)' \in \mathbb{R}^{d,+}$ represents the dynamic range, and Φ is an invertible matrix that represents a coordinate transformation. This quantizer partitions the region

$$\Lambda = \{X \in \mathbb{R}^d : \Phi(X - c) \in \{[-L_1, L_1] \times \dots \times [-L_d, L_d]\}\}$$

into boxes with side lengths $\frac{2L_i}{2^{R_i}}$. Let $R = \sum_{i=1}^d R_i$. Each of the 2^R boxes is represented by an element $\sigma \in \Sigma$. Upon observing X the $(c, \underline{R}, \underline{L}, \Phi)$ -quantizer

- (1) subtracts off c ,
- (2) applies the coordinate transformation Φ ,
- (3) determines which box it falls into (points that land on the boundary of more than one box are given the label of one of those boxes according to some fixed priority rule),
- (4) and then transmits the σ representing that box.

If X falls outside the region Λ then the quantizer transmits a special symbol representing an overflow. Thus we have $2^R + 1$ symbols. The set Λ is called the *support* of the quantizer.

Figure 2 shows a two-dimensional primitive quantizer with $R_1 = 3$ and $R_2 = 2$. Here the dynamic range in both directions are equal: $L_1 = L_2 = 1$ and c is the origin. The rate of this primitive quantizer is $R = 5$.

For both encoder class one and encoder class two the encoding map \mathcal{E}_t based on its information I_t^* selects a $(c, \underline{R}, \underline{L}, \Phi)$ -quantizer. Upon observing Y_t it computes the appropriate σ_t and transmits it across the channel. The decoder needs to know which quantizer was selected so that it may decode the received symbol σ_t appropriately. This is assured by equi-memory. Specifically the equi-memory condition forces the encoder and decoder to make decisions, in this case choose a primitive quantizer, based on the same information.

One may ask why we have chosen boxes instead of more general polytopes to partition Λ . Using general polytopes may lead to a lower rate than the rate one gets when restricting oneself to boxes. However the analysis for the boxes case is

much easier and they are simpler to implement in practice. Moreover we will show that for certain cases encoders using primitive quantizers are sufficient to achieve the information theoretic lower bounds proved in section three.

D. Key Technical Lemma

The growth of the uncertainty in the state is determined by the A matrix. Let us consider the uncontrolled system: $X_{t+1} = AX_t$. Assume $X_t \in \{[-L, L] \times \dots \times [-L, L]\}$. We would like a way to compute a box that bounds the set that X_{t+1} lives in given the box that X_t lives in. In particular we will characterize the growth of the box that upper bounds the state in terms of the dynamics of the state.

We first treat the simple case where the A matrix has only real eigenvalues each with geometric multiplicity one. Let Φ diagonalize A : $\Phi A \Phi^{-1} = \Upsilon = \text{diag}[\lambda_1, \dots, \lambda_d]$. Let $Z_t = \Phi X_t$. Then

$$Z_{t+1} = \Phi X_{t+1} = \Phi A X_t = \Phi A \Phi^{-1} Z_t = \Upsilon Z_t.$$

If $Z_t \in \{[-L, L] \times \dots \times [-L, L]\}$ then $Z_{t+1} \in \{[-|\lambda_1|L, |\lambda_1|L] \times \dots \times [-|\lambda_d|L, |\lambda_d|L]\}$.

We now generalize this idea to general A matrices. The construction we present requires the real Jordan canonical form. We quickly review this now:

Theorem 4.1: For $A \in \mathbb{R}^{d \times d}$ there exists a real valued nonsingular matrix Φ and a real valued matrix Υ such that $\Phi A \Phi^{-1} = \Upsilon = \text{diag}[J_1, \dots, J_m]$. Where each J_j , $j = 1, \dots, m$, is a Jordan block of dimension (geometric multiplicity) d_j . Clearly $d_1 + \dots + d_m = d$. The Jordan block associated with a real eigenvalue λ takes the form

$$\begin{bmatrix} \lambda & 1 & & \\ & \lambda & 1 & \\ & & \ddots & \ddots \\ & & & \lambda \end{bmatrix}.$$

The Jordan block associated with the complex conjugate pair of eigenvalues $\lambda = \rho(\cos\theta \pm i \sin\theta)$ takes the form

$$\begin{bmatrix} D & I & & \\ & D & I & \\ & & \ddots & \ddots \\ & & & D \end{bmatrix}$$

where $D = \rho r(\theta)$. Where $r(\theta) = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$.

Proof: See theorem 2 of section 6.4 in [9]. (Note $\rho \geq 0$). \square

Define $H = \text{diag}[H_1, \dots, H_m]$. Where each H_j is associated with one of the Jordan blocks J_j . Specifically $H_j = I$ if J_j is the Jordan block with real eigenvalue λ_j . And $H_j = \text{diag}[r(\theta)^{-1}, \dots, r(\theta)^{-1}]$ if J_j is the Jordan block associated with the complex conjugate eigenvalues $\rho(\cos\theta \pm i \sin\theta)$. If A has all real eigenvalues then $H = I$. The following lemma, proved in the appendix, shows that Υ and any power of the matrix H commute.

Lemma 4.1: $H^t \Upsilon H^{-t} = \Upsilon$

As before let $X_{t+1} = AX_t$ and let $Z_t = H^t \Phi X_t$. Note that if A has real eigenvalues then $Z_t = \Phi X_t$. The H is

needed to undue the effects caused by the dynamics of the complex conjugate eigenvalue pairs. In general Υ will not be an upper triangular matrix but the matrix $H\Upsilon$ will be an upper triangular matrix with real-valued eigenvalues. The magnitude of these eigenvalues are the same as the magnitude of the corresponding eigenvalues of Υ . Now

$$\begin{aligned} Z_{t+1} &= H^{t+1}\Phi X_{t+1} \\ &= H^{t+1}\Phi A X_t \\ &= H^{t+1}\Phi A \Phi^{-1} H^{-t} Z_t \\ &= H^{t+1}\Upsilon H^{-t} Z_t \\ &= H\Upsilon Z_t \quad \text{by lemma 4.1} \end{aligned}$$

We need a way to bound the growth of the operator $H\Upsilon$. This is an upper triangular matrix which can be written in the following block diagonal form $H\Upsilon = \text{diag}[K_1, \dots, K_m]$ where $K_j = J_j$ if J_j is the Jordan block associated with a real eigenvalue and

$$K_j = H_j J_j = \begin{bmatrix} \rho I & r(\theta)^{-1} & & \\ & \rho I & r(\theta)^{-1} & \\ & & \ddots & \\ & & & \rho I \end{bmatrix}$$

if J_j is the Jordan block associated with a complex conjugate eigenvalue pair. The eigenvalues of the upper triangular matrix K_j are all equal to ρ .

For each block K_j associated with a real eigenvalue λ_j define

$$\bar{K}_j \doteq \begin{bmatrix} |\lambda| & 1 & & \\ & |\lambda| & 1 & \\ & & \ddots & \\ & & & |\lambda| \end{bmatrix}.$$

For each block K_j associated with a complex eigenvalue $\rho(\cos \theta + i \sin \theta)$ define

$$\bar{K}_j \doteq \begin{bmatrix} \rho I & O & & \\ & \rho I & O & \\ & & \ddots & \\ & & & \rho I \end{bmatrix}$$

where $O = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.

Finally define $\bar{\Upsilon} \doteq \text{diag}[\bar{K}_1, \dots, \bar{K}_m]$.

Lemma 4.2: If Z_t is in the box determined by \underline{L} (i.e. $Z_t \in \{[-L_1, L_1] \times \dots \times [-L_d, L_d]\}$) then Z_{t+1} is in the box determined by $\bar{\Upsilon}\underline{L}$.

Proof: We know $Z_{t+1} = H\Upsilon Z_t$. By construction $\bar{\Upsilon}$ is a matrix whose ij th entry is greater than or equal to the absolute value of the ij th entry of $H\Upsilon$. Thus if Z_t is in the box determined by \underline{L} then Z_{t+1} will be in the box determined by $\bar{\Upsilon}\underline{L}$. \square

We are now in a position to state the key technical lemma. This lemma shows how the growth in the box representing the uncertainty in the location of the state is determined by both the dynamics of the plant as well as the information

transmitted across the channel. For a given rate vector $\underline{R} = (R_1, \dots, R_d)'$ define:

$$F_{\underline{R}} \doteq \begin{bmatrix} 2^{-R_1} & & & \\ & 2^{-R_2} & & \\ & & \ddots & \\ & & & 2^{-R_d} \end{bmatrix}.$$

Lemma 4.3: If for all i we have $R_i > \max\{0, \log |\lambda_i|\}$ then $\bar{\Upsilon}F_{\underline{R}}$ is stable. If there exists at least one i such that $R_i < \max\{0, \log |\lambda_i|\}$ then $\bar{\Upsilon}F_{\underline{R}}$ is unstable.

Proof: The matrix $\bar{\Upsilon}$ is an upper triangular matrix whose eigenvalues have the same magnitude as the corresponding eigenvalues of the A matrix. The matrix $\bar{\Upsilon}F_{\underline{R}}$ is an upper triangular matrix with the values $|\lambda_i|2^{-R_i}$ along its diagonal.

If for all i the rate $R_i > \max\{0, \log |\lambda_i|\}$ then $0 \leq |\lambda_i|2^{-R_i} < 1$ and hence all of the eigenvalues of $\bar{\Upsilon}F_{\underline{R}}$ will be stable. If $R_i < \max\{0, \log |\lambda_i|\}$ for at least one i then $\bar{\Upsilon}F_{\underline{R}}$ will have at least one unstable eigenvalue. \square

Example: Let $A = \begin{bmatrix} -3 & 1 \\ 0 & -3 \end{bmatrix}$, $\underline{R} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, and $\underline{L}(t) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. The size of the box determined by $\underline{L}(t)$ will expand due to the dynamics of the system and then shrink due to the rate of information we transmit. Specifically

$$\begin{aligned} \underline{L}(t+1) &= \bar{\Upsilon}F_{\underline{R}}\underline{L}(t) \\ &= \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 2^{-2} & 0 \\ 0 & 2^{-3} \end{bmatrix} \underline{L}(t) \\ &= \begin{bmatrix} \frac{7}{8} \\ \frac{3}{8} \end{bmatrix}. \end{aligned}$$

V. ENCODER CLASS ONE

In this section we provide sufficient conditions on the rate for observability and stabilizability for systems with encoders restricted to encoder class one. In particular we show that we can achieve the lower bounds proved in section three. Furthermore we show that the rate of convergence depends on the difference between the channel rate used and the lower bound.²

A. Observability

Our first proposition treats asymptotic observability when the encoder observes the state ($Y_t = X_t$.) We follow this with two corollaries. The first provides a rate of convergence and the second shows that we can treat the case where the bound on the initial condition is unknown. We then generalize our basic observability result to systems with additive process disturbances and systems with output observations.

The results in this section are based on the following idea. At time t the encoder first computes the decoder's estimate of the state. It can compute this estimate because it knows both the policy of the decoder and the control and channel

²We can make an analogy with the rate distortion theorem in information theory. If the channel rate C is less than the rate distortion rate R then one cannot transmit information reliably (i.e. we cannot insure that the probability of decoding error converges to zero with time.) If $R < C$ then one can transmit reliably [6], [16].

signals available to the decoder. The encoder then computes the difference, the *innovation*, between the true value of the state and the decoder's estimate of the state. Next the encoder chooses a primitive quantizer and quantizes this innovation. It then transmits to the decoder the appropriate channel symbol corresponding to this quantization value. The decoder, due to equi-memory, knows which primitive quantizer the encoder used and hence can decode this channel symbol and update its state estimate. In the following we show that such a scheme leads to asymptotic observability for all rates satisfying the given rate bound.

There are many situations where the rate bounds presented in section 3 will be non-integer. Clearly we cannot send a non-integer number of bits at each time step. But if we use a time-sharing scheme we can, on average, send a non-integer number of bits. We define the *average rate* of an encoding scheme to be $\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} R(t)$ where $R(t)$ is the number of bits transmitted at time t . We treat both the fixed rate and average rate cases in the following proposition. Let $\lceil x \rceil$ represent the smallest integer larger than or equal to x .

Proposition 5.1: For system (1), encoders restricted to encoder class one, $C = I$, and bounded initial set Λ_0

(a) a sufficient condition on the *fixed* rate for asymptotic observability is

$$R > \sum_{\lambda(A)} \max\{0, \lceil \log |\lambda(A)| \rceil\}.$$

(b) a sufficient condition on the *average* rate for asymptotic observability is

$$R > \sum_{\lambda(A)} \max\{0, \log |\lambda(A)|\}.$$

Proof: We will restrict our encoder and decoder to be equi-memory. We prove part (a) first. Because Λ_0 is bounded we know there exists a constant L such that $\Lambda_0 \subset \{X : \|X\|_2 \leq L\}$. Let Φ diagonalize A into real Jordan canonical form: $\Phi A \Phi^{-1} = \Upsilon$. For $X \in \Lambda_0$ we have $\|\Phi X\|_2 \leq \|\Phi\| \|X\|_2 \leq \|\Phi\| L$. At time zero choose a $(c(0), \underline{R}, \underline{L}(0), \Phi(0))$ -quantizer where $c(0)$ is the origin and $\Phi(0) = \Phi$. Let $L_i(0) = \|\Phi\| L \forall i$ and choose the rate vector \underline{R} such that each component R_i is an integer and $R_i > \max\{0, \lceil \log |\lambda_i| \rceil\}$, $i = 1, \dots, d$. Apply this quantizer to X_0 and transmit σ_0 . Because the box determined by $\underline{L}(0)$ contains Λ_0 the channel symbol σ_0 will not be the overflow symbol.

At time t let the state estimate, \hat{X}_t , be the centroid of the region defined by σ_t . Due to equi-memory both the encoder and decoder can determine this centroid. We update the quantizer parameters as follows:

(1) The centroid of the $t+1$ th quantizer is just the one-step ahead state prediction (the encoder observes the controls):

$$c(t+1) = A\hat{X}_t + BU_t.$$

(2) The coordinate transformation evolves as:

$$\Phi(t+1) = H^{t+1}\Phi = H\Phi(t).$$

As in section 4 let $Z_t = H^t \Phi X_t$ represent the state in the new coordinates. Then

$$Z_{t+1} = H\Upsilon Z_t + H^{t+1}\Phi BU_t.$$

(3) The size of the dynamic range of the $t+1$ th quantizer will evolve according to:

$$\underline{L}(t+1) = \bar{\Upsilon} F_{\underline{R}} \underline{L}(t).$$

The $F_{\underline{R}}$ term captures the decrease in state estimation error at time t due to applying the primitive quantizer while $\bar{\Upsilon}$ term captures the growth in the one-step ahead error due to the dynamics of the plant. By lemma 4.3 $\bar{\Upsilon} F_{\underline{R}}$ is a stable matrix. Thus the dimensions of the dynamic range are decreasing in time.

This completes our description of the encoder and decoder.

By construction the state X_t never leaves the range of the t -th quantizer with support Λ_t . Where

$$\Lambda_t = \{X \in \mathbb{R}^d : \Phi(t)(X - c(t)) \in \{[-L_1(t), L_1(t)] \times \dots \times [-L_d(t), L_d(t)]\}\}.$$

Let

$$\Omega_t = \{X \in \mathbb{R}^d : \Phi(t)(X - \hat{X}_t) \in \left\{ \left[-\frac{L_1(t)}{2R_1}, \frac{L_1(t)}{2R_1} \right] \times \dots \times \left[-\frac{L_d(t)}{2R_d}, \frac{L_d(t)}{2R_d} \right] \right\}.$$

One can think of Λ_t as containing the one-step ahead state prediction error for X_t and Ω_t as containing the state estimation error for X_t based on Λ_t and the new information provided by the channel symbol σ_t . Now

$$\begin{aligned} \|e_t\|_2 &\leq \sup_{X \in \Omega_t} \|X - \hat{X}_t\|_2 \\ &= \sup_{X \in \Omega_t} \|\Phi(t)^{-1} \Phi(t)(X - \hat{X}_t)\|_2 \\ &\leq \sup_{X \in \Omega_t} \|\Phi(t)^{-1}\| \|\Phi(t)(X - \hat{X}_t)\|_2 \\ &\leq \|\Phi(t)^{-1}\| \|F_{\underline{R}} \underline{L}(t)\|_2 \\ &\leq \|H^{-t}\| \|\Phi^{-1}\| \|F_{\underline{R}}\| \|(\bar{\Upsilon} F_{\underline{R}})^t\| \|\underline{L}(0)\|_2 \\ &\leq \sqrt{d} L \|\Phi\| \|\Phi^{-1}\| \|F_{\underline{R}}\| \|(\bar{\Upsilon} F_{\underline{R}})^t\| \end{aligned}$$

Since $\|\Phi\|$, $\|\Phi^{-1}\|$, and $\|F_{\underline{R}}\|$ are all bounded we see that there exists a constant γ such that $\|e_t\|_2 \leq \gamma L \|(\bar{\Upsilon} F_{\underline{R}})^t\|$. Since $\bar{\Upsilon} F_{\underline{R}}$ is a stable matrix the error goes to zero as $t \rightarrow \infty$. Thus we have shown the uniform attractivity condition in the definition of asymptotic observability holds. Furthermore, for any $\epsilon > 0$ we can find an L small enough so that for all $t \geq 0$ we have $\|e_t\|_2 \leq \epsilon$. Thus the stability condition in the definition of asymptotic observability holds as well.

We will now prove part (b) by specifying a time-sharing encoder and decoder scheme. In this case the channel rate used at each time step can vary. Pick any $R_i > \max\{0, \log |\lambda_i|\}$, $i = 1, \dots, d$. There exist nonnegative integers M , α_i , β_i , $i = 1, \dots, d$ such that

$$R_i > \alpha_i + \frac{\beta_i}{M} > \max\{0, \log |\lambda_i|\}, \quad i = 1, \dots, d.$$

We will show that there exists an encoder and decoder scheme with average rate less than or equal to $\sum_{i=1}^d R_i$.

First we need to specify a periodic schedule based on epochs of length M . At time t apply a $(c(t), \underline{R}(t), \underline{L}(t), \Phi(t))$ -quantizer where $c(t)$ and $\Phi(t)$ evolve as before. The sides of the dynamic range now evolve as

$$\underline{L}(t+1) = \bar{\Upsilon} F_{\underline{R}(t)} \underline{L}(t)$$

where $\underline{R}(t) = [R_1(t), \dots, R_d(t)]'$ is defined as follows $\forall i = 1, \dots, d$:

$$R_i(t) = \begin{cases} \alpha_i + 1 & \text{if } t \bmod M \in \{0, 1, \dots, \beta_i - 1\} \\ \alpha_i & \text{if } t \bmod M \in \{\beta_i, \dots, M - 1\} \end{cases}$$

For each $i = 1, \dots, d$ we have

$$\begin{aligned} & \frac{1}{T} \sum_{t=0}^{T-1} R_i(t) \\ &= \frac{1}{T} \left((T - T \bmod M) \left(\alpha_i + \frac{\beta_i}{M} \right) + \sum_{\tau=1}^{T \bmod M} R_i(\tau) \right) \\ &= \left(\alpha_i + \frac{\beta_i}{M} \right) + \\ & \quad \frac{1}{T} \left((-T \bmod M) \left(\alpha_i + \frac{\beta_i}{M} \right) + \sum_{\tau=1}^{T \bmod M} R_i(\tau) \right) \end{aligned}$$

The second addend goes to zero as $T \rightarrow \infty$. Thus for T large enough we have:

$$\max\{0, \log |\lambda_i|\} < \frac{1}{T} \sum_{t=0}^{T-1} R_i(t) < R_i.$$

Hence for T large enough we have: $\sum_{i=1}^d \max\{0, \log |\lambda_i|\} < \frac{1}{T} \sum_{i=1}^d \sum_{t=0}^{T-1} R_i(t) < \sum_{i=1}^d R_i$.

Now we need to show that under this time-sharing scheme the estimation error goes to zero as $t \rightarrow \infty$. As before

$$\|e_t\|_2 \leq \|\Phi(t)^{-1}\| \|F_{\underline{R}(t)} \underline{L}(t)\|_2$$

where $\underline{L}(t) = \left(\prod_{\tau=1}^t \bar{\Upsilon} F_{\underline{R}(\tau)} \right) \underline{L}(0)$. The matrix $\prod_{\tau=1}^M \bar{\Upsilon} F_{\underline{R}(\tau)}$ is an upper triangular matrix. Using an argument similar to lemma 4.3 we see that it is stable. Thus

$$\begin{aligned} \|e_t\|_2 &\leq \|\Phi(t)^{-1}\| \|F_{\underline{R}(t)} \underline{L}(t)\|_2 \\ &\leq \|\Phi(t)^{-1}\| \|F_{\underline{R}(t)}\| \left\| \left(\prod_{\tau=1}^M \bar{\Upsilon} F_{\underline{R}(\tau)} \right)^{\frac{t - (t \bmod M)}{M}} \right\| \\ &\quad \times \left\| \prod_{\tau=1}^{(t \bmod M)} \bar{\Upsilon} F_{\underline{R}(\tau)} \right\| \|\underline{L}(0)\|_2 \end{aligned}$$

which goes to zero as $t \rightarrow \infty$. \square

For the rest of the paper the statement ‘‘a sufficient condition for property X to hold is $R > \sum_{\lambda(A)} \max\{0, \log |\lambda(A)|\}$ ’’ will be taken to mean we are using the fixed rate scheme for integer rates and the average rate scheme for non-integer rates.

Corollary 5.1: For both the fixed rate and average rate encoding schemes described in proposition 5.1 the rate of convergence is $\|e_t\|_2 \leq \kappa t^{(d_{\max}-1)} 2^{-t(\min_i(R_i - \log |\lambda_i(A)|))}$ where κ is a constant independent of t .

Proof: We treat the fixed rate case. A similar argument can be used for the average rate case. The decay of $\|(\bar{\Upsilon} F_{\underline{R}})^t\|$ is determined by the largest eigenvalue of $\bar{\Upsilon} F_{\underline{R}}$. The largest eigenvalue is given by $\max_i 2^{(-R_i + \log |\lambda_i(A)|)}$. There exists a constant η such that $\|(\bar{\Upsilon} F_{\underline{R}})^t\| \leq$

$\eta t^{(d_{\max}-1)} 2^{-t(\min_i(R_i - \log |\lambda_i(A)|))}$ where d_{\max} is the multiplicity of the largest eigenvalue. Thus

$$\begin{aligned} \|e_t\|_2 &\leq \sqrt{d} L \|\Phi\| \|\Phi(t)^{-1}\| \|F_{\underline{R}}\| \|(\bar{\Upsilon} F_{\underline{R}})^t\| \\ &\leq \kappa t^{(d_{\max}-1)} 2^{-t(\min_i(R_i - \log |\lambda_i(A)|))} \end{aligned}$$

for a constant κ independent of t (Recall $\|\Phi(t)^{-1}\| \leq \|H^{-t}\| \|\Phi^{-1}\| \leq \|\Phi^{-1}\|$.) \square

Next we treat the case when Λ_0 is bounded but the encoder does not *a priori* know that bound.

Corollary 5.2: For system (1), encoders restricted to encoder class one, and $C = I$ a sufficient condition on the rate for asymptotic observability is $R > \sum_{\lambda(A)} \max\{0, \log |\lambda(A)|\}$.

Proof: For the case when a bound on the initial uncertainty, Λ_0 , is unknown one must first ‘‘capture’’ the state in the quantizer domain. Let Φ be as in proposition 5.1. At time zero apply the primitive quantizer $(c(0), \underline{R}(0), \underline{L}(0), \Phi)$ where $c(0)$ is the origin, $\Phi(0) = \Phi$, and $L_i(0) = L \forall i$ for some fixed L . If upon observing X_t the quantizer transmits an overflow symbol then update the quantizer as follows: $c(t+1) = Ac(t) + BU_t$, $\Phi(t) = H^t \Phi$, and $\underline{L}(t+1) = 2\bar{\Upsilon} \underline{L}(t)$. Since $\underline{L}(t)$ is growing at a rate faster than the dynamics of the state process eventually the quantizer will capture the state. At this point proceed as we did in proposition 5.1. (Strictly speaking we are using one extra symbol for the overflow message. But as $t \rightarrow \infty$ the overhead due to these overflow messages is amortized out.) \square

We now consider the case of bounded additive process disturbances.

$$X_{t+1} = AX_t + BU_t + W_t, \quad Y_t = X_t, \quad t \geq 0 \quad (2)$$

where $\|W_t\|_2 \leq D$. In corollary 3.1 we provided a necessary condition to insure bounded error. The following proposition provides a related sufficient condition.

Proposition 5.2: For system (2), encoders restricted to encoder class one, and bounded initial set Λ_0 the rate $R > \sum_{\lambda(A)} \max\{0, \log |\lambda(A)|\}$ is sufficient to insure $\limsup_{t \rightarrow \infty} \|e_t\|_2$ is bounded.

Proof: First we prove the fixed rate case. We follow the proof in proposition 5.1. In this case the state dynamics in the new coordinates are:

$$Z_{t+1} = H\Upsilon Z_t + \Phi(t+1)BU_t + \Phi(t+1)W_t.$$

We update the $\underline{L}(t)$ as follows:

$$\underline{L}(t+1) = \bar{\Upsilon} F_{\underline{R}} \underline{L}(t) + D \|\Phi(t+1)\| \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

Here we expand the size of the dynamic range of the quantizer to take into account the dynamics of the plant, the information transmitted across the channel, and the bounded process disturbances. For shorthand denote the second addend by $h(t)$. Then we can write

$$\underline{L}(t) = (\bar{\Upsilon} F_{\underline{R}})^t \underline{L}(0) + \sum_{j=0}^{t-1} (\bar{\Upsilon} F_{\underline{R}})^{t-1-j} h(j).$$

As in proposition 5.1 we have

$$\begin{aligned}
 & \|e_t\|_2 \\
 & \leq \|\Phi(t)^{-1}\| \|F_{\underline{R}}\underline{L}(t)\|_2 \\
 & \leq \|\Phi(t)^{-1}\| \|F_{\underline{R}}\| \\
 & \quad \times \left(\|(\bar{\Upsilon}F_{\underline{R}})^t \underline{L}(0)\|_2 + \left\| \sum_{j=0}^{t-1} (\bar{\Upsilon}F_{\underline{R}})^{t-1-j} h(j) \right\|_2 \right) \\
 & \leq \sqrt{d} \|\Phi(t)^{-1}\| \|F_{\underline{R}}\| \times \\
 & \quad \left(\|(\bar{\Upsilon}F_{\underline{R}})^t \underline{L}(0)\|_2 + \sum_{j=0}^{t-1} D \|(\bar{\Upsilon}F_{\underline{R}})^{t-1-j}\| \|H^{j+1}\| \|\Phi\| \right) \\
 & \leq \sqrt{d} \|\Phi(t)^{-1}\| \|F_{\underline{R}}\| \\
 & \quad \times \left(\|(\bar{\Upsilon}F_{\underline{R}})^t \underline{L}(0)\|_2 + D \|\Phi\| \sum_{j=0}^{t-1} \|(\bar{\Upsilon}F_{\underline{R}})^{t-1-j}\| \right)
 \end{aligned}$$

The matrix $(\bar{\Upsilon}F_{\underline{R}})$ is stable. Hence the first term goes to zero as $t \rightarrow \infty$. For the second term note that there exists a constant κ such that $\|(\bar{\Upsilon}F_{\underline{R}})^t\| \leq \kappa t^{(d_{\max}-1)} 2^{-t(\min_i(R_i - \log |\lambda_i(A)|))}$. Thus the series is summable. Hence there exists an η such that $\lim_{t \rightarrow \infty} \sum_{j=0}^{t-1} \|(\bar{\Upsilon}F_{\underline{R}})^{t-1-j}\| \leq \frac{\eta}{1-2^{-\min_i(R_i - \log |\lambda_i(A)|)}}$. Therefore $\lim_{t \rightarrow \infty} \|e_t\| \leq \frac{\eta \sqrt{d} D}{1-2^{-\min_i(R_i - \log |\lambda_i(A)|)}} \|\Phi(t)^{-1}\| \|\Phi\| \|F_{\underline{R}}\|$. A similar argument can be used for the average rate case. \square

Due to the noise term W_t there will always be a nonzero state estimation error. But if the channel rate goes to infinity then we see that $\|F_{\underline{R}}\| = \max_i 2^{-R_i} \rightarrow 0$ and, as we expect, the bound on the estimation error in the previous proposition goes to zero.

Example: Consider the scalar system: $X_{t+1} = aX_t + bU_t + W_t$. In this case the limit of the error is bounded as $\lim_{t \rightarrow \infty} |e_t| \leq \frac{D}{2^R - |a|}$.

Now we consider system (1) with general observation equation $Y_t = CX_t$. Assume that the pair (A, C) is observable. We will need the following technical lemma which is proved in the appendix.

Lemma 5.1: Let A be a stable matrix. Let B_t be a set of matrices such that $\|B_t\| \leq K < \infty$ and the $\lim_{t \rightarrow \infty} \|B_t\| = 0$. Let $S_t = \sum_{i=0}^{t-1} A^{t-1-i} B_i$. Then $\lim_{t \rightarrow \infty} \|S_t\| = 0$.

Proposition 5.3: For system (1), encoders restricted to encoder class one, and bounded initial set Λ_0 a sufficient condition for asymptotic observability is $R > \sum_{\lambda(A)} \max\{0, \log |\lambda(A)|\}$.

Proof: At time t the encoder has received Y^t, U^{t-1} . We will assume that the encoder has access to an observer that produces a state estimate \bar{X}_t . Specifically the encoder applies a Luenberger observer:

$$\bar{X}_t = A\bar{X}_{t-1} + BU_{t-1} + M(Y_{t-1} - C\bar{X}_{t-1})$$

where M is chosen so that $A - MC$ is stable. The estimation error of this observer is $\bar{e}_t = X_t - \bar{X}_t = (A - MC)\bar{e}_{t-1}$. Thus $\bar{e}_t = (A - MC)^t \bar{e}_0$ and $\|\bar{e}_t\| \leq \|(A - MC)^t\| \|\bar{e}_0\| \leq \eta \alpha^t \|\bar{e}_0\|$ for some constant η and $0 \leq \alpha < 1$. The observer's initial state error $\|\bar{e}_0\|$ is bounded because Λ_0 is bounded. Hence as $t \rightarrow \infty$ the observer's state error \bar{e}_t goes to zero.

We will construct an encoder and decoder that asymptotically observes the observer state \bar{X}_t . The observer's estimation error goes to zero hence the encoder and decoder will asymptotically observe X_t . The observer's state estimate evolves as

$$\bar{X}_t = A\bar{X}_{t-1} + BU_{t-1} + MC\bar{e}_{t-1}.$$

We will treat the $MC\bar{e}_{t-1}$ term as a process disturbance.

First we treat the fixed rate scheme. As in proposition 5.2 we update $\underline{L}(t)$ as follows:

$$\underline{L}(t+1) = \bar{\Upsilon}F_{\underline{R}}\underline{L}(t) + c\alpha^t \|MC\| \|\bar{e}_0\| \|\Phi(t+1)\| \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

where $c = \frac{\eta}{\alpha}$. Denote the second addend by $\alpha^t h(t)$. Then

$$\underline{L}(t) = (\bar{\Upsilon}F_{\underline{R}})^t \underline{L}(0) + \sum_{j=0}^{t-1} (\bar{\Upsilon}F_{\underline{R}})^{t-1-j} \alpha^j h(j)$$

Since $\bar{\Upsilon}F_{\underline{R}}$ is stable the first term goes to zero. Because $\lim_{t \rightarrow \infty} \alpha^t h(t) = 0$ ($h(t)$ is uniformly bounded in t) lemma 5.1 implies that the second term goes to zero. Thus as $t \rightarrow \infty$ the error in the decoder's estimate of the encoder's observer state converges to zero. Thus the rate condition is sufficient for asymptotic observability. A time-sharing argument can be used for the average rate case. \square

The general prescription for observability in encoder class one is to transmit a finer and finer description of the zero control input response state trajectory. If there are no disturbances this is equivalent to successively refining the initial position. Specifically if we allow the encoder "infinite" memory then it need only transmit a finer and finer description of X_0 . But assuming that the encoder can have in memory a perfect description of X_0 for all time t is unrealistic. Furthermore it is not robust to disturbances. For this reason we have proposed a recursive structure for the encoders in encoder class one. As we have shown this recursive structure can be used for systems with process disturbances and output observations.

B. Stabilizability

For encoder class one we can combine the properties of asymptotic observability and full state feedback stability to get output feedback stability. Assume that the pair (A, B) is stabilizable and the pair (A, C) is observable.

Proposition 5.4: For system (1), encoder restricted to encoder class one, and the initial set Λ_0 bounded a sufficient condition for asymptotic stabilizability is $R > \sum_{\lambda(A)} \max\{0, \log |\lambda(A)|\}$.

Proof: Let K be a stabilizing controller, i.e. $A + BK$ is stable. Apply the certainty equivalent controller $U_t = K\hat{X}_t$ where \hat{X}_t is the decoder's state estimate. We need to show that under this control law the system is stable. As before let $e_t = X_t - \hat{X}_t$. Then

$$X_t = (A + BK)^t X_0 - \sum_{j=0}^{t-1} (A + BK)^{t-1-j} BK e_j$$

By proposition 5.3 we can asymptotically observe the state. Thus the $\lim_{t \rightarrow \infty} \|e_t\|_2 = 0$. Since $A + BK$ is stable and Λ_0 is bounded the first addend in the above equation goes to zero. By lemma 5.1 so does the second. Hence $\lim_{t \rightarrow \infty} \|X_t\|_2 = 0$. Thus we have shown the uniform attractivity condition in the definition of asymptotic stabilizability holds. Furthermore, for any $\epsilon > 0$ we can find an L small enough so that for all $t \geq 0$ we have $\|X_t\|_2 \leq \epsilon$. Thus the stability condition in the definition of asymptotic stabilizability holds as well. \square

This result is related to a general result of Vidyasagar that states if a system is state feedback stabilizable and output observable then it is output feedback stabilizable [18]. Furthermore the certainty equivalent controller applied to the state estimate is a stabilizing controller. The difference here is that due to the encoder and decoder the ‘‘observation equation’’ will depend on the past states and controls.

We can treat the case when Λ_0 is bounded but the encoder does not *a priori* know that bound.

Corollary 5.3: For system (1), encoder restricted to encoder class one, and $C = I$ a sufficient condition for asymptotic stabilizability is $R > \sum_{\lambda(A)} \max\{0, \log |\lambda(A)|\}$.

Proof: Apply the zero control until the encoder, using the technique in corollary 5.2, ‘‘captures’’ the state. Then proceed as in proposition 5.4. \square

Example: Consider the scalar system $X_{t+1} = aX_t + bU_t$, $a > 1$, $|X_0| \leq L$. Choose controller k such that $|a + bk| < 1$. Then under full state feedback the magnitude of the state is strictly decreasing to the origin $|X_t| = |a + bk|^t |X_0|$. Under a rate $R > \log a$ and the scheme proposed in the last proposition we see that

$$|X_t| \leq |a + bk|^t |X_0| + \sum_{j=0}^{t-1} |a + bk|^{t-1-j} |bk| \frac{1}{2R^j} L.$$

There can exist trajectories that initially are not strictly decreasing to the origin. One can consider this the price of learning the state under a rate constraint. See [7].

VI. ENCODER CLASS TWO

In this section we examine observability and stabilizability for encoders restricted to encoder class two. For this encoder class we assume that the encoder knows the control law though not the actual control signals.

A. Observability

For encoder class one we showed there exist encoders that can asymptotically observe the state. Furthermore, because the encoder observed the controls, asymptotic observability held independently of the control signals chosen. Here we will show that the condition of observability for encoders in encoder class two depends on the control signals chosen.

Proposition 6.1: For system (1) with unstable A , encoder restricted to encoder class two, initial set Λ_0 , $C = I$, and controls set to zero there is no finite rate encoder in encoder class two such that the estimation error is bounded.

Proof: As in proposition 3.1 we can assume without loss of generality that the initial uncertainty set contains the bounded

set $\Omega_0 = \{X : \|X\|_2 \leq L\} \subseteq \Lambda_0$ and that the A matrix contains only marginally stable and unstable eigenvalues. The set of points that X_t can take contains the following set $\Omega_t \doteq \{X : X = A^t X_0, X_0 \in \Omega_0\}$.

If there exists an encoder and decoder in encoder class two such that the estimation error is bounded then there must be a $\beta < \infty$ and a $T(L)$ such that $\|e_t\|_2 \leq \beta$ for $t \geq T(L)$. A lower bound on the rate can be computed by counting the number of regions of diameter less than 2β it takes to cover Ω_t for $t \geq T(L)$.

At any time $t_0 > T(L)$ we know that any $X_{t_0} \in \Omega_{t_0}$ is a realizable state. To insure an error of size β at time t_0 we would need to transmit a total of at least

$$\log \frac{\text{volume}(\Omega_{t_0})}{K_d \beta^d} = t_0 \sum_{\lambda(A)} \log |\lambda(A)| + d \log \frac{L}{\beta}$$

bits over the t_0 time steps. Thus for the first t_0 steps we would need to transmit bits at a rate of at least $R \geq \sum_{\lambda(A)} \log |\lambda(A)| + \frac{d}{t_0} \log \frac{L}{\beta}$. By the same argument, at time $t_1 = t_0 + N$ we would need to transmit a total of at least $t_1 \sum_{\lambda(A)} \log |\lambda(A)| + d \log \frac{L}{\beta}$ bits over t_1 time steps. Encoders in encoder class two are memoryless hence the encoder at any given time cannot depend on any of the previous encodings of the state. Thus we cannot use any of the bits transmitted during the first t_0 time steps to help aid the reconstruction of X_{t_1} . Therefore over the time interval from t_0 to t_1 we would need to transmit bits at a rate of at least $R \geq \frac{1}{N} \left(t_1 \sum_{\lambda(A)} \log |\lambda(A)| + d \log \frac{L}{\beta} \right)$. For fixed N we see that the right hand side grows to infinity as t_0 , and hence t_1 , goes to infinity. Thus there does not exist a finite rate scheme that will insure bounded error. \square

In summary, the encoder at time t cannot depend on the previous encodings and hence cannot compute the innovation with respect to the previous encodings. Thus there can be no finite rate encoder in encoder class two such that the estimation error is bounded.

This result may seem like bad news. However in most cases we are interested in observability when the controls being applied are in the closed loop. We will now show that we can achieve asymptotic observability and as a by-product of asymptotic stabilizability.

B. Stabilizability

Assume (A, B) is a stabilizable pair. Because the encoders in encoder class two do not observe the control signals they need to operate with the closed loop dynamics of $A + BK$ and not the open loop dynamics of A . We have assumed that the encoder knows the control law K . Computing the minimal rate is difficult for this class of encoders because in general the rate will depend on the particular controller policy K .

Proposition 6.2: For system (1), encoder restricted to encoder class two, $C = I$, and bounded initial set Λ_0 there exists a finite rate such that the system is asymptotically stabilizable.

Proof: Choose K such that $A + BK$ is stable. Let Φ diagonalize $(A + BK)$ into real Jordan canonical form: $\Phi(A + BK)\Phi^{-1} = \Upsilon$. Assume $\Lambda_0 \subset \{X : \|X\|_2 \leq L\}$. For $X \in \Lambda_0$ we have $\|\Phi X\|_2 \leq \|\Phi\| \|X\|_2 \leq \|\Phi\| L$. At time zero choose

a $(c(0), \underline{R}, \underline{L}(0), \Phi(0))$ -quantizer where $c(0)$ is the origin and $\Phi(0) = \bar{\Phi}$. Let $L_i(0) = \|\bar{\Phi}\|L \forall i$. The rate vector \underline{R} will be determined shortly. Apply this quantizer to X_0 and transmit σ_0 . Note that σ_0 will not be the overflow symbol.

The encoder does not have access to the controls or the past channel symbols. Thus it can only evolve according to a schedule based on the control law. The state evolves as:

$$X_{t+1} = (A + BK)X_t - BK e_t.$$

Update the quantizer parameters as follows: $c(t) = 0$ for all time and $\Phi(t+1) = H\Phi(t)$. Where H is defined as in section 4 but in this case it is with respect to the matrix $A + BK$. Then the state in the new coordinates evolves as:

$$Z_{t+1} = H\Upsilon Z_t - \Phi(t+1)BK e_t$$

where $e_t = X_t - \hat{X}_t$. Let $f_t = Z_t - \hat{Z}_t = \Phi(t)e_t$. Then

$$Z_{t+1} = H\Upsilon Z_t - \Phi(t+1)BK\Phi^{-1}(t)f_t.$$

Observe that Z_t is in the box determined by $\underline{L}(t)$ and the error f_t is in the box determined by $F_{\underline{R}}\underline{L}(t)$. Let the boundary lengths of the dynamic range evolve as:

$$\underline{L}(t+1) = \{\bar{\Upsilon} + \|\Phi(t+1)BK\Phi^{-1}(t)\|F_{\underline{R}}\} \underline{L}(t).$$

The operator $\bar{\Upsilon}$ is stable because Υ is stable. We can then find a rate vector \underline{R} with components large enough so that $\bar{\Upsilon} + \|\Phi(t+1)BK\Phi^{-1}(t)\|F_{\underline{R}}$ is uniformly stable in t . Thus $\lim_{t \rightarrow \infty} \underline{L}(t) = 0$ and the system is asymptotically stabilizable. \square

Example: Consider the scalar system $X_{t+1} = aX_t + bU_t$. Let k be such that $|a + bk| < 1$. Then

$$L(t+1) = (|a + bk| + \frac{|bk|}{2R})L(t)$$

Letting $R > \max\{0, \log \frac{|bk|}{1-|a+bk|}\}$ is sufficient to ensure asymptotic stabilizability. If $a + bk = 0$ then the rate bound becomes $R > \max\{0, \log |a|\}$.

For both encoder class one and two our strategy has been to keep track of a region in which the state lies. The size and location of this region are recursively updated. One might wonder how these regions, determined by the dynamic range and support of the quantizer, are related to the Lyapunov level sets used in Lyapunov theory. In Lyapunov theory one finds a suitable Lyapunov function V such that the value of V decreases along trajectories [7] [4]. For a finite channel rate it is impossible to insure that V decreases at every time step. This is because there is always a region around the origin that is not under the influence of any control.

VII. CONCLUSION

In this paper we formulated a discrete time, linear systems control problem with a noiseless digital communication link. We discussed the role of information patterns and control policy knowledge in this context.

We first provided lower bounds on the rates required to achieve asymptotic observability and asymptotic stabilizability. These bounds hold independently of the information pattern chosen. To compute upper bounds we explicitly described

the encoder, decoder, and controller schemes. We characterized two different encoder structures based on whether the encoder observed the control signals or not. Under the added structural assumptions of equi-memory and use of a primitive quantizer we showed that encoders in encoder class one can achieve these lower bounds. For encoders in encoder class a weaker result was provided.

APPENDIX

Let $A \in \mathbb{R}^{d \times d}$. Then by theorem 4.1 A has a real Jordan canonical form.

Lemma 4.1 $H^t \Upsilon H^{-t} = \Upsilon$.

Proof: $H^t \Upsilon H^{-t}$ is the product of three block diagonal matrices. Thus we need only check that the result holds for each of the blocks. The blocks come in two types: those associated with real eigenvalues and those associated with complex eigenvalues. For the real eigenvalue case H_j is identity. Thus clearly $I^t J_j I^{-t} = J_j$. Let us examine the complex conjugate eigenvalue case:

$$\begin{aligned} & H_j^t J_j H_j^{-t} \\ &= \text{diag}[r(\theta)^{-t}] \begin{bmatrix} \rho r(\theta) & I & & \\ & \rho r(\theta) & I & \\ & & \ddots & \\ & & & \rho r(\theta) \end{bmatrix} \text{diag}[r(\theta)^t] \\ &= \begin{bmatrix} \rho r(\theta) & I & & \\ & \rho r(\theta) & I & \\ & & \ddots & \\ & & & \rho r(\theta) \end{bmatrix} = J_j. \end{aligned}$$

\square

Lemma 5.1 Let A be a stable matrix. Let B_t be a set of matrices such that $\|B_t\| \leq K$ and the limit $\lim_{t \rightarrow \infty} \|B_t\| = 0$. Let $S_t = \sum_{i=0}^{t-1} A^{t-1-i} B_i$. Then $\lim_{t \rightarrow \infty} \|S_t\| = 0$.

Proof: Since A is stable there exists $c \geq 0$ and $0 \leq \lambda < 1$ such that $\|A^t\| \leq c\lambda^t$. For all $\epsilon > 0$ there exists a $T(\epsilon)$ such that $\|B_t\| \leq \epsilon \forall t \geq T(\epsilon)$. Let $t > T(\epsilon)$. Then

$$\begin{aligned} & \left\| \sum_{j=0}^{t-1} A^{t-j-1} B_j \right\| \\ & \leq \sum_{j=0}^{t-1} \|A^{t-j-1}\| \|B_j\| \\ & \leq c \sum_{j=0}^{t-1} \lambda^{t-j-1} \|B_j\| \\ & \leq c \left\{ \lambda^{t-T(\epsilon)-1} \sum_{j=0}^{T(\epsilon)} \lambda^{T(\epsilon)-j} K + \sum_{j=T(\epsilon)+1}^{t-1} \lambda^{t-j-1} \epsilon \right\} \\ & \leq \frac{c}{1-\lambda} \left\{ \lambda^{t-T(\epsilon)-1} K + \epsilon \right\} \end{aligned}$$

Now $\epsilon > 0$ can be chosen arbitrarily small and $t > T(\epsilon)$ can be chosen arbitrarily large hence the bound can be made arbitrarily small. \square

ACKNOWLEDGMENT

The authors would like to thank Nicola Elia for many helpful discussions and the reviewers for their detailed comments.

REFERENCES

- [1] S. Axler, *Linear Algebra Done Right*, Springer-Verlag, New York, 1997.
- [2] J. Baillieul, "Feedback Designs for Controlling Device Arrays with Communication Channel Bandwidth Constraints." In *Lecture Notes of the Fourth ARO Workshop on Smart Structures*, Penn State University, August 1999.
- [3] J. Baillieul, "Feedback Designs in Information-Based Control." Proceedings of the Workshop on Stochastic Theory and Control, University of Kansas, October 2001.
- [4] R. Brockett and D. Liberzon, "Quantized Feedback Stabilization of Linear Systems." *IEEE Trans. on Automatic Control*, Vol. 45, No. 7, pp. 1279-1289, July 2000.
- [5] V. Borkar and S. Mitter, "LQG Control with Communication Constraints." In *Communications, Computation, Control, and Signal Processing: a Tribute to Thomas Kailath*. Norwell, MA, Kluwer Academic Publishers: 1997.
- [6] T. Cover and J. Thomas, *Elements of Information Theory*, John Wiley, New York, 1991.
- [7] N. Elia and S. Mitter, "Stabilization of Linear Systems with Limited Information." *IEEE Trans. on Automatic Control*, Vol. 46, No. 9, pp. 1384-1400, September 2001.
- [8] G. Gabor and Z. Györfi, *Recursive Source Coding*. New York, NY: Springer-Verlag, 1986.
- [9] M. Hirsch and S. Smale, *Differential Equations, Dynamical Systems, and Linear Algebra*. San Diego, CA: Academic Press, Inc., 1974.
- [10] R. Murray, editor., "Control in an Information Rich World: Report of the Panel on Future Directions in Control, Dynamics, and Systems." SIAM, June 30, 2002. Also available at: <http://www.cds.caltech.edu/%7Emurray/cdspanel/>
- [11] G. Nair and R. Evans, "Stabilization with Data-Rate-Limited Feedback: Tightest Attainable Bounds." *Systems and Control Letters*, Vol. 41, pp. 49-76, 2000.
- [12] G. Nair and R. Evans, "Exponential Stabilisability of Finite-Dimensional Linear Systems with Limited Data Rates." *Automatica*, Vol. 39, pp. 585-593, 2003.
- [13] I. Petersen and A. Savkin, "Multi-rate Stabilization of Multivariable Discrete-Time Linear Systems via a Limited Capacity Communication Channel." Proceedings of the 40th IEEE Conference on Decision and Control, December 2001.
- [14] S. Tatikonda, "Control Under Communication Constraints." MIT Ph.D. thesis, August 2000. Available at *****
- [15] S. Tatikonda and S. Mitter, "Control Under Communication Constraints." Proceedings of the 38th Annual Allerton Conference on Communication, Control, and Computing, October 2000.
- [16] S. Tatikonda and S. Mitter, "Control Over Noisy Channels." Submitted to the *IEEE Trans. on Automatic Control*, February 2002.
- [17] S. Tatikonda, A. Sahai, and S. Mitter, "Control of LQG Systems Under Communication Constraints." Proceedings of the 37th IEEE Conference on Decision and Control, Tampa, Florida, December 1998.
- [18] M. Vidyasagar, "On the Stabilization of Nonlinear Systems Using State Detection." *IEEE Trans. on Automatic Control*, Vol. 25, No. 3, pp. 504-509, June 1980.
- [19] H. Witsenhausen, "Separation of Estimation and Control for Discrete Time Systems." Proceedings of the IEEE, Volume 59, No. 11, November 1971.
- [20] W. Wong and R. Brockett, "Systems with Finite Communication Bandwidth Constraints I: State Estimation Problems," *IEEE Trans. on Automatic Control*, Vol. 42(9), pp. 1294-1299, Sept. 1997.
- [21] W. Wong and R. Brockett, "Systems with Finite Communication Bandwidth Constraints II: Stabilization with Limited Information Feedback," *IEEE Trans. on Automatic Control*, Vol. 44 (5), pp. 1049 -1053, May 1999.
- [22] H. Kopka and P. W. Daly, *A Guide to L^AT_EX*, 3rd ed. Harlow, England: Addison-Wesley, 1999.

Sekhar Tatikonda received his Ph.D. in Electrical Engineering and Computer Science from MIT in 2000. From 2000-02 he was a postdoctoral fellow in the Computer Science Department at the University of California, Berkeley. He is currently an Assistant Professor of Electrical Engineering at Yale University. His research interests include communication theory, information theory, stochastic control, distributed estimation and control, statistical machine learning and inference.

Sanjoy Mitter received his Ph.D. degree from the Imperial College of Science and Technology in 1965. He taught at Case Western Reserve University from 1965 to 1969. He joined MIT in 1969 where he has been a Professor of Electrical Engineering since 1973. He was the Director of the MIT Laboratory for Information and Decision Systems from 1981 to 1999. He has also been a Professor of Mathematics at the Scuola Normale, Pisa, Italy from 1986 to 1996. He has held visiting positions at Imperial College, London; University of Groningen, Holland; INRIA, France; Tata Institute of Fundamental Research, India and ETH, Zurich, Switzerland. He was the McKay Professor at the University of California, Berkeley in March 2000, and held the Russell-Severance-Springer Chair in Fall 2003. He has held visiting positions in several American universities. He is a Fellow of the IEEE and a Member of the National Academy of Engineering. He is the winner of the 2000 IEEE Control Systems Award. His current research interests are Communication and Control in a Networked Environment, the relationship of Statistical and Quantum Physics to Information Theory and Control and Autonomy and Adaptiveness for Integrative Organization.