Data Rate Theorem for Stabilization Over Time-Varying Feedback Channels

Paolo Minero, *Student Member, IEEE*, Massimo Franceschetti, *Member, IEEE*, Subhrakanti Dey, *Senior Member, IEEE*, and Girish N. Nair, *Member, IEEE*

Abstract—A data rate theorem for stabilization of a linear, discrete-time, dynamical system with arbitrarily large disturbances, over a rate-limited, time-varying communication channel is presented. Necessary and sufficient conditions for stabilization are derived, their implications and relationships with related results in the literature are discussed. The proof techniques rely on both information-theoretic and control-theoretic tools.

Index Terms—Control under communication constraints, entropy, quantized control, source coding.

I. INTRODUCTION

IN modern control theory, the *data rate theorem* refers to the smallest feedback data rate above which an unstable dynamical system can be stabilized. In its scalar form, it states that a discrete linear plant of unstable mode $|\lambda| \geq 1$ can be stabilized if and only if the data rate R over the (noise free) digital feedback link satisfies the inequality $R > \log_2 |\lambda|$ bits per sample, where $\tilde{H} = \log_2 |\lambda|$ is called the intrinsic entropy rate of the plant. From its first appearance, this result has been generalized to different notions of stability and system models, and has also been extended to multi-dimensional systems [1], [3], [5], [13], [16], [21], [24]. The survey papers [2] and [17] give an historical and technical account of the various formulations.

In many engineering applications, the aim is to control one or more dynamical systems using multiple sensors and actuators communicating over digital links. In this framework, the data rate theorem represents a point of contact where the theories of control and communication converge, as it relates the speed of the dynamics of the plant to the information rate of the communication channel. From an information-theoretic perspective, the existence of a critical positive rate below which there does not exist any quantization and control scheme able to stabilize an unstable plant is reminiscent of Shannon's source coding theorem [20]. Stated informally, this says that if one wants to communicate with a fixed-length code over a noise free channel the
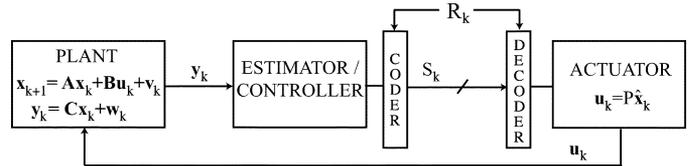

Fig. 1. Feedback loop model. The (encoded) estimated state s_k is quantized and sent to a decoder over a wireless digital link that supports error-free transmission of R_k bits per discrete unit time.

output of a finite-valued stationary ergodic source process with entropy rate $H(\mathcal{X})$, then the number of bits that must be used to represent the source sequence with arbitrarily small error probability is at least $H(\mathcal{X})$. In other words, Shannon's entropy rate, representing the amount of uncertainty of the source, poses a fundamental limit on the communication rate. Similarly, the intrinsic entropy rate $\tilde{H} = \log_2 |\lambda|$ of an unstable linear dynamical system, representing the growth of the state space spanned by the open loop system, poses a fundamental limit on the minimum data rate that must be available over the feedback loop to guarantee stability.

In this paper, we are concerned with the formulation of the data rate theorem over time-varying feedback channels. A motivating example is given by sensors and actuators communicating over a wireless channel for which the quality of the communication link varies over time because of random fading in the received signal. In the case of digital communication, this can reflect in a time variation of the rate supported by the wireless channel. However, if the channel variations are slow enough, transmitter and receiver can estimate the quality of the link by sending a training sequence, and can adapt the communication scheme to the channel's condition. We ask the following question: is it possible to design a communication scheme that changes dynamically according to the channel's condition and, at the same time, is guaranteed to stabilize the system?

To answer the above question, we assume the following model. The communication channel, at any given time k, allows transmission of R_k bits without error, where R_k fluctuates randomly over time. R_k remains constant in blocks of n consecutive channel uses and then varies according to an independent and identically distributed (i.i.d.) process across blocks. Furthermore, both encoder and decoder have causal knowledge of the rate supported by the communication link, see Fig. 1. We remark that such channel state information (CSI) can be obtained through feedback from the receiver to the transmitter if the fading variation is slow enough.

The model above includes the erasure channel as a special case, by allowing the rate process to have only a value $R > 0$, or zero if an erasure occurs. In this case, CSI at the transmitter can be simply obtained through one bit feedback that notifies the sender of erasures.

The model, however, does not capture the possibility of having other decoding errors beside erasures. Rather than addressing general channels with noise, our aim here is to obtain crisp results in a simple setting which can be used to understand the basic trade-offs between the intrinsic entropy rate of the system, the available rate on the communication channel, and the additional randomness due to the changing conditions of the environment. In this framework, our work directly relates to the ones in [9], [13], [16], [19], [21] and we describe this relationship in more detail next, while we refer the reader interested in more general channels with noise to the work of Sahai and Mitter [18], as well as to the works in [14], [15], [22], [23].

In an influential paper, Tatikonda and Mitter [21] have studied a model similar to ours in which the rate is fixed and system disturbances are bounded. Nair and Evans [16] addressed the case in which the rate is still fixed, but disturbances can have an unbounded support (Gaussian disturbances are a special case of this). Finally, Martins, Dahleh, and Elia [13] considered the case of a scalar system with state feedback, random time-varying rate and bounded disturbances, and they provided necessary and sufficient conditions for mth moment stability. In this work, we allow both the system disturbances to have unbounded support *and* the rate to vary randomly. Furthermore, the encoder has access to output feedback rather than to state feedback and we also consider the multi-dimensional case. This formulation requires the use of an adaptive quantizer, as this must be capable of tracking the state when atypically large disturbances affect the system and must dynamically adapt to the rate that is instantaneously supported by the channel. Naturally, our results can recover the ones mentioned in the above papers.

We also want to spend a few words on a different approach that has been used in the literature to model control over time-varying channels. This has a network-theoretic flavor rather than the information-theoretic one described above. In this case, the channel uncertainty is modeled using random packet dropouts. Packets are considered as single entities, each carrying the estimated state, that can be lost independently, with some probability. Furthermore, channel state information is in this case modeled as packet acknowledgement at the transmitter. An extensive survey of different works following this approach appears in [11] and we refer the reader to this work for references. The network-theoretic equivalent of the data rate theorem is the proof of existence of a *critical dropout probability* above which the closed loop system cannot be stabilized, see for example [7], [9], [12], [19].

Our present paper reveals an important link between the network-theoretic, packet-loss model described above, and the information-theoretic approach. From an information-theoretic perspective, the packet loss model corresponds to an erasure channel in which the rate is infinity, with probability $1 - p$ and zero with probability p. This is because a single packet, representing the state of the system which is a real quantity, can

carry an infinite amount of information, as a real number can have arbitrarily many bits within its binary expansion. Now, if we apply our results to an erasure channel, where the rate is R with probability $1 - p$ and zero with probability p, in the high data rate limit ($R \to \infty$) this channel can be seen as communicating real numbers with random i.i.d. erasures, and in this case we obtain a necessary and sufficient condition for stabilization that is the same as the one in [9], obtained under the network theoretic model, with Bernoulli packet dropouts, acknowledgement of packet reception, and Gaussian system disturbances.

The rest of the paper is organized as follows. The main contributions are informally summarized and discussed next. Section III formally defines the problem. Section IV is devoted to the proof of the necessary and sufficient conditions for stabilizability in the scalar case. These are shown via the entropy-power inequality (necessary) and the construction of an adaptive, variable length encoder (sufficiency). Section V is devoted to the more complex multi-dimensional case, for which necessary and sufficient conditions are shown to be tight in some special cases.

II. OVERVIEW OF THE RESULTS

In the scalar case, we prove that a necessary and sufficient condition to stabilize a linear system of unstable mode $|\lambda| \geq 1$ in the second moment sense over a digital link of time-varying limited rate R_k as described above, is

$$\mathbb{E}\left[\left(\frac{\lambda^2}{2^{2R}}\right)^n\right] < 1 \qquad (1)$$

where n is the length of the block during which the rate on the digital link remains constant, and the rates R_k's are i.i.d. across blocks and distributed as a random variable R.

The condition above is amenable to the following intuitive interpretation. If no information is sent over the link during a transmission block, the estimation error at the decoder about the state of the system grows by λ^{2n}. The information sent by the encoder can reduce this error by at most 2^{2nR}, where nR is the total rate supported by the channel in a given block. However, if averaging over the fluctuation of the rate λ^{2n} exceeds 2^{2nR}, then the information sent over the channel cannot compensate (on average) the dynamics of the system and it is not possible to stabilize the plant. Notice that if the rate is fixed over time and equal to a constant R, then the condition in (1) reduces to the well known inequality $R > \log_2 |\lambda|$.

Finally, it is also easy to see that when communicating over an erasure channel for which $R = \infty$ with probability $1 - p$ and $R = 0$ with probability p, then for $n = 1$ the necessary and sufficient condition for stabilization in (1) reduces to

$$p < \frac{1}{\lambda^2}$$

which is the same critical loss probability derived in [9] for systems with Gaussian (i.e. unbounded-support) disturbances under the network-theoretic model.

The proof of the result in (1) is based on an information-theoretic argument based on the entropy-power inequality (necessary condition), and on an explicit construction of an adaptive quantizer and coder-decoder pair (sufficient condition). In the

latter case, the main challenge is to design a quantizer that adapts dynamically to the exogenous rate process and can handle atypically large disturbances. The construction of the coder-decoder pair is similar to the one by Nair and Evans [16]. There are, however, some key differences vis-à-vis in the way the stabilizing scheme is constructed. In [16], time is divided into cycles of fixed duration, and system state observations are quantized using a fixed number of bits, which are transmitted over the digital link for the duration of a cycle. Thus, communication between coder and decoder occurs at a fixed transmission rate. In our case, the total number of bits available in a cycle of fixed duration is random and it is not known a priori, as the rate process is known only causally at the coder and the decoder. As a consequence, the choice of an appropriate quantization rate is not immediate. Our solution consists in dividing time into cycles of *fixed duration*, but quantizing the state observations using a random number of bits, which depends on the realization of the rate process. The fact that future realizations of the rate process are not known in advance is not a problem, since the quantizer we use is successively refinable, and can dynamically adapt to the rate that is instantaneously supported by the digital link. Hence, our scheme performs as if the future realizations of the rate process were known in advance at the coder and decoder. An alternative approach consists of quantizing the observations using a fixed number of quantization points, but allowing cycles to have *variable duration*. A scheme based on this approach is outlined in Section V-D. Finally, we remark that, as in related works in the literature [13], [16], [21], the construction provided in this paper relies on the crucial assumption that the coder and decoder can agree on the initial values of the internal states through an a priori iterative communication process.

The extension of the analysis to multi-dimensional linear systems entails the difficulty of the rate allocation to the different unstable modes. In this case, we derive necessary conditions for second moment stabilizability, which define a region with a special polymatroid structure. When the rate is fixed and equal to a constant R, the necessary conditions reduce to

$$\tilde{H} := \sum_{|\eta_i| \geq 1} \log_2 |\eta_i| < R$$

where $\eta_1, \ldots, \eta_n$ are the open loop eigenvalues (raised to their corresponding algebraic multiplicities). Again, this recovers the well known data rate theorem for vector systems with deterministic rate [16], [21]. Finally, as in the scalar case, in the high data rate limit over an erasure channel, we also recover the necessary condition on the critical dropout probability of [9].

Finally, we provide a general coder-decoder construction for vector systems and show that this is optimal in some limiting cases. For some specific rate distributions, however, it is possible to design more efficient schemes. This latter point is shown by considering stabilization over a binary erasure channel, for which a better scheme is proposed.

III. PROBLEM FORMULATION

In the sequel, the following notation is used: vectors are written in bold-faced type and sequences $\{a_i\}_{i=0}^{k}$ are denoted as a_0^k; expectation with respect to the random variable X

is written as $\mathbb{E}_X[\cdot]$, the differential entropy of a continuous random vector $\mathbf{X}$ as $h(\mathbf{X}) = -\mathbb{E}[\ln f_{\mathbf{X}}(\mathbf{x})]$ and the entropy of a discrete random vector $\mathbf{X}$ as $H(\mathbf{X}) = -\mathbb{E}[\ln P(\mathbf{X})]$; the set of non-negative integers as $\mathbb{N}$, the positive integers as $\mathbb{Z}_+$, and the rational numbers as $\mathbb{Q}$; finally, the cardinality of a finite set S is denoted as $|S|$.

Consider the partially-observed, discrete-time state-space unstable stochastic linear system

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{v}_k, \quad \mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{w}_k, \ \forall k \in \mathbb{N} \quad (2)$$

where $\mathbf{x}_k \in \mathbb{R}^d$ is the state process, $\mathbf{u}_k \in \mathbb{R}^m$ is the control input, $\mathbf{v}_k \in \mathbb{R}^d$ process disturbance, the measurement $\mathbf{y}_k$ and measurement noise $\mathbf{w}_k$ are random vectors in $\mathbb{R}^p$. Suppose $\mathbf{A}$ is uniquely composed by unstable modes (having magnitude greater or equal to unity). No Gaussian assumptions are made on the initial condition $\mathbf{x}_0$ and the disturbances, but the following is assumed to hold:

A0. $(\mathbf{A}, \mathbf{B})$ is reachable and $(\mathbf{C}, \mathbf{A})$ observable.

A1. $\mathbf{x}_0$, $\mathbf{v}_k$ and $\mathbf{w}_j$ are mutually independent for all $k, j \in \mathbb{N}$.

A2. $\exists \epsilon > 0$ such that $\mathbf{x}_0$, $\mathbf{v}_k$ and $\mathbf{w}_j$ have uniformly bounded $(2 + \epsilon)$th absolute moments over $k \in \mathbb{N}$.

A3. $\inf_{k \in \mathbb{N}} h(\mathbf{v}_k) > -\infty$. Thus, $\exists \beta > 0$ such that $e^{(2/d)h(\mathbf{v}_k)} > \beta$ for all $k \in \mathbb{N}$ and $\mathbf{v}_k \in \mathbb{R}^d$.

Suppose that coder and decoder are connected by a *time-varying* digital link, see Fig. 1. The transmission rate supported by the digital link is assumed constant over blocks of $n \in \mathbb{N}$ channel uses but changes independently from block to block according to a given probability distribution. Formally, at time $k \in \mathbb{N}$ the digital link is an identity map on an alphabet $\{1, \cdots, |S_k|\}$; $\log_2 |S_k|$ denotes the transmission rate supported by the digital link, and coincides with R_j if and only if $k \in \{jn, jn + 1, \cdots, (j + 1)n - 1\}$. At time $k \in \{jn, jn + 1, \cdots, (j + 1)n - 1\}$, coder and decoder know R_j, while the realization of the rate process in future blocks, $\{R_i\}_{i=j+1}^{\infty}$, is unknown to them. The $\{R_j\}_{j \in \mathbb{N}}$ are i.i.d. random variables distributed as R, where R is an integer-valued random variable taking values on $\mathcal{R} \subseteq \mathbb{N}$. We denote by r_{min} the minimum value in the set $\mathcal{R}$.

This definition of the rate process is motivated by communication over wireless channels. In fact, the rate supported by a block fading wireless channel can be modeled as a random variable, since this is a function of the (random) channel gain that attenuates the transmitted signal. The block fading model captures a fading scenario where the fading channel state remains invariant over a block of time but changes from block to block. If the fading variation is slow enough, feedback from the receiver to the transmitter can be used to acquire channel state information. If the channel state information is known, then the rate supported by the channel is also known at both transmitter and receiver. Finally, the rate can be modeled as an i.i.d. random process across the channel blocks if we assume that block lengths are similar to coherence time intervals (length of time over which the channel's statistical properties do not change) of the channel. For example, the i.i.d. assumption is valid for a slow frequency hopped time division multiple access channel.

Example 3.1: We call *erasure channel* a digital link where $R = r$ with probability $1 - p$ and $R = 0$ with probability p, for some nonnegative integer r and $p \in (0, 1)$. If $r = 1$, we call the channel *binary erasure channel*.

Each transmitted symbol can depend on all past and present measurements, the present channel state and the past symbols

$$S_k = g_k\left(\mathbf{y}_0^k, S_0^{k-1}, R_j\right), \quad k \in \{jn, \dots, (j+1)n-1\}, \quad \forall j \in \mathbb{N}$$

where $g_k(\cdot)$ is the coder mapping at time k. The control sequence, on the other hand, can depend on all past and present channel symbols

$$\mathbf{u}_k = \delta_k\left(S_0^k\right) \quad \forall k \in \mathbb{N}$$

where $\delta_k(\cdot)$ is the controller mapping at time k.

We want to construct a coder-decoder pair which stabilizes the plant in the mean square sense

$$\sup_{k \in \mathbb{N}} \mathbb{E}\left[\|\mathbf{x}_k\|^2\right] < \infty \tag{3}$$

using the finite data rate provided by the time-varying digital feedback link.

IV. Scalar Systems

In this section it is assumed that the plant in (2) is scalar and has a representation of the following type:

$$x_{k+1} = \lambda x_k + u_k + v_k, \quad y_k = x_k + w_k, \quad \forall k \in \mathbb{N} \tag{4}$$

where $|\lambda| \geq 1$, so that the system is unstable. The result for the scalar case is now stated:

Theorem 4.1: Under assumptions **A0.–A3.** above, necessary and sufficient condition for stabilizing the plant (4) in the mean square sense (3) is that

$$\mathbb{E}\left[\left(\frac{|\lambda|^2}{2^{2R}}\right)^n\right] < 1 \tag{5}$$

where n is the length of the channel block with the same rate.

A. Necessity

In order to prove the statement, we find a lower bound for the second moment of the state, and show that (5) is a necessary condition for this lower bound to be finite. We focus on the times $k = jn$ with $j \in \mathbb{N}$, i.e. on the beginning of each channel block. Let $\bar{S}_j = \{S_0, \dots, S_{(j+1)n-1}\}$, denote the symbols sent over the noiseless channel until the end of the jth channel block. By iteration of (4), we have

$$x_{(j+1)n} = \lambda^n x_{jn} + \sum_{k=jn}^{(j+1)n-1} \lambda^{(j+1)n-1-k}\left[\delta_k\left(S_0^k\right) + v_k\right].$$

Let $n_j = (1/2\pi e)\mathbb{E}_{\bar{S}_{j-1}}[e^{2h(x_{jn}|\bar{S}_{j-1}=\bar{s}_{j-1})}]$ be the conditional entropy power of x_{jn} conditioned on the event $\{\bar{S}_{j-1} = \bar{s}_{j-1}\}$,

averaged over all possible $\bar{s}_{j-1}$. The second moment of x_{jn} is lower bounded by n_j

$$
\begin{aligned}
\mathbb{E}\left[x_{jn}^2\right] &= \mathbb{E}_{\bar{S}_{j-1}}\left[\mathbb{E}\left[x_{jn}^2|\bar{S}_{j-1}=\bar{s}_{j-1}\right]\right] \\
&= \frac{1}{2\pi e}\mathbb{E}_{\bar{S}_{j-1}}\left[e^{\ln(2\pi e \mathbb{E}[x_{jn}^2|\bar{S}_{j-1}=\bar{s}_{j-1}])}\right] \\
&\geq \frac{1}{2\pi e}\mathbb{E}_{\bar{S}_{j-1}}\left[e^{2h(x_{jn}|\bar{S}_{j-1}=\bar{s}_{j-1})}\right] \\
&= n_j
\end{aligned}
$$

where the inequality follows from the maximum entropy theorem [4, Theorem 9.4.1]. It follows that a necessary condition for (3) to hold is that $\sup_{j \in \mathbb{N}} n_j < \infty$. We now complete the proof by showing that this necessary condition is violated whenever (5) does not hold. We make use of the following technical lemma (proved in the Appendix A),

Lemma 4.2: For all non-negative random variables R_j, the following inequality holds:

$$\mathbb{E}_{\bar{S}_j|\bar{S}_{j-1}, R_j}\left[e^{2h(x_{jn}|\bar{S}_j=\bar{s}_j)}\right] \geq \frac{1}{2^{2nR_j}}e^{2h(x_{jn}|\bar{S}_{j-1}=\bar{s}_{j-1})}.$$

First, we show that n_j evolves according to a recursive equation. Using standard properties of entropy [4] (translation invariance, conditional version of entropy power inequality), and assumptions **A1.** and **A3.**, it follows that:

$$
\begin{aligned}
&\mathbb{E}_{\bar{S}_j}\left[e^{2h\left(x_{(j+1)n}|\bar{S}_j=\bar{s}_j\right)}\right] \\
&= \mathbb{E}_{\bar{S}_j}\left[e^{2h\left(\lambda^n x_{jn} + \sum_{k=jn}^{(j+1)n-1}\lambda^{(j+1)n-1-k}v_k|\bar{S}_j=\bar{s}_j\right)}\right] \\
&\geq \lambda^{2n}\mathbb{E}_{\bar{S}_j}\left[e^{2h(x_{jn}|\bar{S}_j=\bar{s}_j)}\right] + \gamma \\
&= \lambda^{2n}\mathbb{E}_{\bar{S}_{j-1}, R_j}\left[\mathbb{E}_{\bar{S}_j|\bar{S}_{j-1}, R_j}\left[e^{2h(x_{jn}|\bar{S}_j=\bar{s}_j)}\right]\right] + \gamma \\
&\geq \lambda^{2n}\mathbb{E}_{\bar{S}_{j-1}, R_j}\left[\frac{1}{2^{2nR_j}}e^{2h(x_{jn}|\bar{S}_{j-1}=\bar{s}_{j-1})}\right] + \gamma \\
&= \lambda^{2n}\mathbb{E}_{R_j}\left[\frac{1}{2^{2nR_j}}\right]\mathbb{E}_{\bar{S}_{j-1}}\left[e^{2h(x_{jn}|\bar{S}_{j-1}=\bar{s}_{j-1})}\right] + \gamma
\end{aligned}
$$

wherein the second inequality follows from assumption **A3.** above, i.e. $e^{2h(v_k)} > \beta$. The constant γ is defined as $\gamma := \sum_{k=jn}^{(j+1)n-1}\lambda^{2[(j+1)n-1-k]}\beta$. Finally, the last inequality follows from Lemma 4.2 and the fact that R_j is independent of x_{jn} and $\bar{S}_{j-1}$. Thus, using the fact that the rate process is i.i.d., we have

$$n_{j+1} \geq \mathbb{E}\left[\left(\frac{\lambda^2}{2^{2R}}\right)^n\right]n_j + \frac{\gamma}{2\pi e}.$$

Therefore, $\mathbb{E}[(\lambda^2/2^{2R})^n] \geq 1$ implies that $\sup_{j \in \mathbb{N}} n_j = \infty$.

B. Sufficiency

We first describe the adaptive quantizer that is at the base of the constructive scheme. A fundamental property of this quantizer is then stated as a lemma, whose proof appears in [16].

1) Quantizer: The quantizer partitions the real line into non-uniform regions, and a parameter $\rho > 1$ determines the speed

at which the quantizer range increases. The quantizer generates 2^ν, $\nu \geq 0$, quantization intervals labeled from left to right by $I_\nu(0), \ldots, I_\nu(2^\nu - 1)$. Let $I_0(0) := (-\infty, \infty)$, $I_1(0) := (-\infty, 0]$ and $I_1(1) := (0, \infty)$. If $\nu \geq 2$ the quantization intervals are generated by

- partitioning the set $[-1, 1]$ into $2^{\nu-1}$ intervals of equal length,
- partitioning the sets $(\rho^{i-2}, \rho^{i-1}]$, $[-\rho^{i-2}, -\rho^{i-1})$ into $2^{\nu-1-i}$ intervals of equal length, $i \in \{2, \ldots, \nu - 1\}$.

The two open sets $(-\infty, -\rho^{\nu-2}]$ and $(\rho^{\nu-2}, \infty)$ are respectively the leftmost and rightmost intervals of the quantizer. Let

- $\kappa_\nu(\omega)$ be half-length of interval $I_\nu(\omega)$ for $\omega \in \{1, \ldots, 2^\nu - 2\}$, be equal to $\rho^\nu - \rho^{\nu-1}$ when $\omega = 2^\nu - 1$ and equal to $-(\rho^\nu - \rho^{\nu-1})$ when $\omega = 0$.
- $q_\nu(x) := \bar{\omega}_\nu(\omega)$ be midpoint of interval $x \in I_\nu(\omega)$ for $\omega \in \{1, \ldots, 2^\nu - 2\}$, be equal to ρ^ν when $\omega = 2^\nu - 1$ and equal to $-\rho^\nu$ when $\omega = 0$.

A property of this construction is that for $\nu \geq 2$ the quantization intervals $I_\nu(\cdot)$ can be generated recursively starting from $q_2(\cdot)$. In fact, for any integer $i \geq 2$ the quantizer intervals for $q_{i+1}(\cdot)$ are formed by partitioning each bounded interval $I_i(\omega), \omega \in \{1, \ldots, 2^i - 2\}$, into two uniform subintervals, and partitioning the semi-infinite interval $I_i(0) = (-\infty, -\rho^{i-2}]$ into two intervals $I_{i+1}(0) = (-\infty, -\rho^{(i+1)-2}]$ and $I_{i+1}(1) = (-\rho^{(i+1)-2}, -\rho^{i-2}]$ and, similarly, partitioning the semi-infinite interval $I_i(2^i - 1) = (\rho^{i-2}, \infty)$ into two intervals $I_{i+1}(2^{i+1} - 2) = (\rho^{i-2}, \rho^{(i+1)-2}]$ and $I_{i+1}(2^{i+1} - 1) = (\rho^{(i+1)-2}, +\infty)$.

Given a real-valued random variable X, if its realization x is in $I(\omega)$ for some $\omega \in \{0, \ldots, 2^\nu - 1\}$, then the quantizer approximates x with $\bar{\omega}_\nu(\omega)$. The quantization error is not uniform over $x \in \mathbb{R}$, but is bounded by $\kappa_\nu(\omega)$ for all $\omega \in \{1, \ldots, 2^\nu - 2\}$. A fundamental property of the quantizer is that the average quantization error diminishes like the inverse square of the number of levels, $2^{-2\nu}$. More precisely, if the $(2 + \epsilon)$th moment of X is bounded for some $\epsilon > 0$, then an upper bound of the second moment of the estimation error decays as $2^{-2\nu}$. The higher moment of X is useful to bound the estimation error (using Chebyshev's inequality) when X lies in one of the two open intervals $(\rho^{\nu-1}, \infty)$ and $(-\infty, -\rho^{\nu-1}]$.

Let $L \in \mathbb{R}_+$ be any random variable, define the functional

$$M_\epsilon[X, L] \equiv \mathbb{E}[L^2 + X^{2+\epsilon}L^{-\epsilon}]. \quad (6)$$

The functional $M_\epsilon[X, L]$ is an upper bound to the second moment of X

$$\mathbb{E}[X^2] = \mathbb{E}\left[X^2\left(1_{|X|\leq L} + 1_{|X|>L}\right)\right] \leq M_\epsilon[X, L]. \quad (7)$$

Define the conditional version of $M_\epsilon[X, L]$ given a random variable R as $M_\epsilon[X, L|R] \equiv \mathbb{E}[L^2 + X^{2+\epsilon}L^{-\epsilon}|R]$. The fundamental property of the quantizer described above is given by the following result:

Lemma 4.3: [16, Lemma 5.2] Let X and $L > 0$ and be random variables with $\mathbb{E}[X^{2+\epsilon}] < \infty$ for some $\epsilon > 0$, and $n \in \mathbb{N}$. If $\rho > 2^{2/\epsilon}$, then for any $R \in \mathbb{N}$ the quantization errors $X - Lq_{nR}(X/L)$ satisfy

$$M_\epsilon\left[X - Lq_{nR}(X/L), L\kappa_{nR}(\omega)\right] \leq \frac{\zeta}{2^{2nR}}M_\epsilon[X, L]$$

where $\omega \in \{0, \ldots, 2^{nR} - 1\}$ is the index of the quantizer level $q_{nR}(X/L)$, and ζ is a constant greater than 2 determined only by ϵ and ρ.

Next, the coder and decoder are described.

2) Coder: The first stage of the encoding process consists of computing the linear minimum variance estimator of the plant state based on the previous measurements and control sequences. The filter process satisfies a recursive equation of the same form as (4), namely

$$\bar{x}_{k+1} = \lambda \bar{x}_k + u_k + z_k, \ \forall k \in \mathbb{N} \quad (8)$$

where $z_k := (y_k - \bar{x}_k)l_k$ is the product of the innovation $y_k - \bar{x}_k$ and the appropriate optimal gain l_k. The $(2 + \epsilon)$th moment of z_k can be shown to be bounded, under assumption A2., for any finite k. From the orthogonality principle the stability of $\bar{x}$ is equivalent to that of x. The output $\bar{x}_k$ of the filter (or a function of it) must be transmitted using the finite number of bits supported on the digital channel. Coder and decoder share a state estimator $\hat{x}_k$ based uniquely on the symbols sent over the digital link. Since $\hat{x}_k$ is available both at the coder and decoder, while the minimum variance estimator is available at the coder only, the encoder uses the quantizer described in the previous section to encode the error between $\bar{x}_k$ and $\hat{x}_k$. The error is scaled by an appropriate coefficient and then recursively encoded using the quantizer in Section IV-B-1. An accurate approximation of the error is obtained by transmitting the quantization index across many channel blocks. The fact that the random rate available at future times is not known in advance is not a problem, as the quantizer is successively refinable and can dynamically adapt to the rate that is instantaneously supported by the channel. By transmitting for a large enough number of blocks, the error between the two estimators can be kept bounded.

Define the coder error at time $k \in \mathbb{N}$ as $f_k = \bar{x}_k - \hat{x}_k$. Times $k \in \mathbb{N}$ are divided into *cycles* $\{jn\tau, \ldots, (j+1)n\tau\}, j \in \mathbb{N}$, of integer duration $n\tau, \tau \in \mathbb{Z}_+$. Notice that each cycle consists of τ channel blocks.

At time $k = jn\tau$, just before the start of the jth cycle, the coder sets the quantization rate equal to $nR_{j\tau}$, i.e. the rate in the first channel block in the jth cycle, and computes

$$\bar{\omega}_{nR_{j\tau}}(\omega_{jn\tau}) = q_{nR_{j\tau}}(f_{jn\tau}/l_j)$$

where l_j is a scaling factor updated at the beginning of each cycle. This factor is used to scale $f_{jn\tau}$ close to the origin, where the quantizer provides better estimates. The index $\omega_{jn\tau}$ of the quantization level is converted into a string of $nR_{j\tau}$ bits and transmitted using the n channel uses of the $j\tau$th channel block. Denote by $I_{nR_{j\tau}}(\omega_{jn\tau})$ the quantization interval labeled by $\omega_{jn\tau}$. After the first n transmissions in the cycle, coder and decoder agree on the fact that $f_{jn\tau}/l_j \in I_{nR_{jn}}(\omega_{jn\tau})$. The remaining $(n-1)\tau$ transmissions in the cycle are devoted to reducing the size of the uncertainty interval $I_{nR_{j\tau}}(\omega_{jn\tau})$.

At time $k = jn\tau + n$, the rate $R_{j\tau+1}$ supported during the next channel block becomes known at both coder and decoder. Thus, coder and decoder divide up $I_{nR_{j\tau}}(\omega_{jn\tau})$ into $2^{nR_{j\tau+1}}$ sub-intervals in the manner described above (uniform partitions of bounded intervals and exponential partition of semi-infinite intervals), sequentially generating the

partitions $I_{nR_{j\tau}+nR_{j\tau+1}}(\cdot) \subseteq I_{nR_{j\tau}}(\omega_{jn\tau})$ of the quantizer $q_{nR_{j\tau}+nR_{j\tau+1}}(f_{jn\tau}/l_j)$.

Then, the coder sends to the decoder the index of the sub-interval containing $f_{jn\tau}/l_j$. At the end of the second channel block in the cycle, coder and decoder agree on the fact that $f_{jn\tau}/l_j \in I_{nR_{j\tau}+nR_{j\tau+1}}(\omega_{jn\tau+n})$.

Continue this process until the end of the τth channel block. After receiving the last sequence of bits, the decoder computes the final uncertainty interval $I_{\nu_j}(\omega_{(j+1)n\tau-n})$, corresponding to the uncertainty set formed by the quantizer $q_{\nu_j}(f_{jn\tau}/l_j)$, where the random variable

$$\nu_j := nR_{j\tau} + nR_{j\tau+1} + \ldots + nR_{(j+1)\tau-1}$$

indicates the cumulative number of bits sent in the jth cycle.

Before the beginning of the $(j+1)$th cycle, the coder updates the state estimator as follows:

$$\hat{x}_{(j+1)n\tau} = \lambda^{n\tau} \left[\hat{x}_{jn\tau} + l_j q_{\nu_j}(f_{jn\tau}/l_j) \right] \\ + \sum_{k=jn\tau}^{(j+1)n\tau-1} \lambda^{(j+1)n\tau-1-k} P\hat{x}_k \quad (9)$$

where

$$\hat{x}_{k+1} = (\lambda + P)\hat{x}_k \quad \forall k \in \{jn\tau, \ldots, j(n+1)\tau-2\} \quad (10)$$

and $\hat{x}_0 = 0$. P is the certainty-equivalent control coefficient such that $|\lambda + P| < 1$. Finally, the scaling coefficient l_j is updated as follows:

$$l_{j+1} = \max \left\{ \sigma, l_j |\lambda|^{n\tau} \kappa_{\nu_j}\left(\omega_{(j+1)n\tau-n} \right) \right\} \quad (11)$$

with $l_0 = \sigma$ and where $\sigma^{2+\epsilon}$ is a uniform bound for the $(2+\epsilon)$-moment of

$$g_j := \sum_{i=1}^{n\tau-1} \lambda^{n\tau-i} z_{jn\tau+i}, \quad j \in \mathbb{N}. \quad (12)$$

3) Decoder: At time $k = jn\tau$ coder and decoder are synchronized and have common knowledge of the state estimator $\hat{x}_{jn\tau}$. During times $jn\tau, \ldots, (j+1)n\tau - 1$, the decoder sends to the plant a certainty-equivalent control signal

$$u_k = P\hat{x}_k \quad \forall k \in \{jn\tau, \ldots, (j+1)n\tau-1\} \quad (13)$$

where $\hat{x}_k$ is updated as in (10). At the end of the each channel block in the jth cycle, the decoder receives estimates of the states in the way described above.

At time $(j+1)n\tau - 1$, once computed $q_{\nu_j}(f_{jn\tau}/l_j)$ the decoder updates the estimator $\hat{x}_{(j+1)n\tau}$ using (9). Synchronism between coder and decoder is ensured by the fact that the initial value $\hat{x}_0$ is set equal to zero at both coder and decoder, and by the fact that the digital link is noiseless.

4) Analysis: In this section it is shown that the coder-decoder pair described above ensures that the second moment of $\bar{x}$ is bounded if (5) is satisfied.

The analysis is developed in three steps. First, we show that f_k is bounded for all times $k = jn\tau, j \in \mathbb{N}$, i.e. the beginning of each cycle. Next, the analysis is extended to all $k \in \mathbb{N}$. Finally, the stability of f_k for all $k \in \mathbb{N}$ is shown to imply that $\bar{x}$ (and so x) is bounded.

First we show that the coder error f_k is bounded in the mean square sense for all times $k = jn\tau, j \in \mathbb{N}$. Instead of looking at $\mathbb{E}[|f_{jn\tau}|^2]$, it is more convenient to consider the functional $M_\epsilon[X, L]$ defined in (6), with $X = f_{jn\tau}$ and $L = l_j$. Thus, let

$$\theta_j := M_\epsilon[f_{jn\tau}, l_j] \equiv \mathbb{E}\left[l_j^2 + f_{jn\tau}^{2+\epsilon} l_j^{-\epsilon} \right].$$

Equation (7) implies that $\mathbb{E}[f_{jn\tau}^2] \leq \theta_j$. Therefore, it suffices to show that $\sup_{j \in \mathbb{N}} \theta_j < \infty$.

Substituting (13) into (8), and iterating over the duration of a cycle, we have

$$\bar{x}_{(j+1)n\tau} = \lambda^{n\tau} \bar{x}_{jn\tau} + g_j + \sum_{k=jn\tau}^{(j+1)n\tau-1} \lambda^{(j+1)n\tau-1-k}(P\hat{x}_k) \quad (14)$$

where g_j is defined in (12). Subtracting (9) from (14), we have

$$\begin{aligned} f_{(j+1)n\tau} &= \bar{x}_{(j+1)n\tau} - \hat{x}_{(j+1)n\tau} \\ &= \lambda^{n\tau} \left[f_{jn\tau} - l_j q_{\nu_j}(f_{jn\tau}/l_j) \right] + g_j. \end{aligned}$$

Notice that, by assumption **A2.**, the $(2+\epsilon)$th moment of $f_{jn\tau}$ is bounded for any finite $j \in \mathbb{N}$. Next, $f_{(j+1)n\tau}$ is used to derive an expression for θ_{j+1}. From the inequality $(|x| + |y|)^\alpha \leq 2^{\alpha-1}(|x|^\alpha + |y|^\alpha) \forall \alpha > 0$, we obtain

$$f_{(j+1)n\tau}^{2+\epsilon} \leq \phi \left(|\lambda^{n\tau}|^{2+\epsilon} \left| f_{jn\tau} - l_j q_{\nu_j}(f_{jn\tau}/l_j) \right|^{2+\epsilon} + g_j^{2+\epsilon} \right)$$

with $\phi = 2^{1+\epsilon}$. Dividing by l_{j+1}^ϵ, taking expectations and using (11), we have

$$\mathbb{E}\left[f_{(j+1)n\tau}^{2+\epsilon} l_{j+1}^{-\epsilon} \right]$$
$$\leq \phi \left(|\lambda^{n\tau}|^{2+\epsilon} \mathbb{E}\left[\frac{|f_{jn\tau} - l_j q_{\nu_j}(f_{jn\tau}/l_j)|^{2+\epsilon}}{l_{j+1}^\epsilon} \right] + \mathbb{E}\left[\frac{g_j^{2+\epsilon}}{l_{j+1}^\epsilon} \right] \right)$$
$$\leq \phi \left(|\lambda^{n\tau}|^{2+\epsilon} \mathbb{E}\left[\frac{|f_{jn\tau} - l_j q_{\nu_j}(f_{jn\tau}/l_j)|^{2+\epsilon}}{[l_j |\lambda^{n\tau}| \kappa_{\nu_j}(\omega_{(j+1)n\tau-n})]^\epsilon} \right] + \mathbb{E}\left[\frac{g_j^{2+\epsilon}}{\sigma^\epsilon} \right] \right)$$
$$= \phi \left(\lambda^{2n\tau} \mathbb{E}\left[\frac{|f_{jn\tau} - l_j q_{\nu_j}(f_{jn\tau}/l_j)|^{2+\epsilon}}{[l_j \kappa_{\nu_j}(\omega_{(j+1)n\tau-n})]^\epsilon} \right] + \mathbb{E}\left[\frac{g_j^{2+\epsilon}}{\sigma^\epsilon} \right] \right) \quad (15)$$

Next, observe that

$$\mathbb{E}\left[l_{j+1}^2 \right] \leq \sigma^2 + \lambda^{2n\tau} \mathbb{E}\left[|l_j \kappa_{\nu_j}(\omega_{(j+1)n\tau-n})|^2 \right]. \quad (16)$$

Summing (15) and (16), using $\mathbb{E}[g_j^{2+\epsilon}] \leq \sigma^{2+\epsilon}$ and the definition of θ_j, we have

$$\begin{aligned} \theta_{j+1} &\leq \phi \left(2\sigma^2 + \right. \\ &+ \left. \lambda^{2n\tau} \mathbb{E}\left[\frac{|f_{jn\tau} - l_j q_{\nu_j}(f_{jn\tau}/l_j)|^{2+\epsilon}}{[l_j \kappa_{\nu_j}(\omega_j)]^\epsilon} + |l_j \kappa_{\nu_j}(\omega_j)|^2 \right] \right) \\ &= \phi \left(2\sigma^2 + \lambda^{2n\tau} \mathbb{E}_{\nu_j}\left[M_\epsilon\left[f_{jn\tau} - l_j q_{\nu_j}\left(\frac{f_{jn\tau}}{l_j} \right), l_j \kappa_{\nu_j}(\omega_j) | \nu_j \right] \right] \right) \\ &\leq \phi 2\sigma^2 + \phi \lambda^{2n\tau} \mathbb{E}_{\nu_j}\left[\frac{\zeta}{2^{2\nu_j}} M_\epsilon[f_{jn\tau}, l_j] \right] \\ &= \phi 2\sigma^2 + \phi \zeta \left(\mathbb{E}\left[\frac{\lambda^{2n}}{2^{2nR}} \right] \right)^\tau \theta_j \end{aligned}$$

where the second inequality follows from Lemma (4.3), and the last equality uses the fact that the rate process is i.i.d. and that f_{jn} and l_j are independent of $R_{j\tau}, R_{j\tau+1}, \ldots, R_{(j+1)\tau-1}$

because of the causality constraint. Therefore, θ_j evolves according to the following recursive equation:

$$\theta_{j+1} \le \phi 2\sigma^2 + \phi\zeta \left(\mathbb{E} \left[\left(\frac{\lambda^2}{2^{2R}} \right)^n \right] \right)^\tau \theta_j.$$

It follows that if $\mathbb{E}[(\lambda^2/2^{2R})^n] < 1$, then by making τ sufficiently large we can ensure that the coefficient of θ_j is strictly less than 1. Thus we have established that θ_j remains bounded in the limit of j going to infinity and therefore $\sup_{j\in\mathbb{N}} \theta_j < \infty$. Hence, from (7) it follows that $\sup_{j\in\mathbb{N}} \mathbb{E}[f_{jnr}^2] < \infty$.

Next, for any $k \in \{0, \ldots, n-1\}$ the triangle inequality implies $|f_{jn+k}| \le |\lambda|^k |f_{jn}| + \sum_{i=0}^{k-1} |\lambda^{k-1-i} P| |z_{jn+k}|$, so the error f_k is bounded for all $k \in \mathbb{N}$. Finally, by rewriting (8) as $\bar{x}_{k+1} = (\lambda + P)\bar{x}_k - Pf_k + z_k$, the fact that f_k and z_k are bounded and that $|\lambda + P| < 1$ ensures that $\mathbb{E}[\bar{x}_k^2] < \infty$ for all $k \in \mathbb{N}$.

V. VECTOR SYSTEMS

In this section, we consider the case of multi-dimensional unstable linear systems. A necessary condition for stabilizability is derived using an information-theoretic approach. It is proved that the stabilizability region is contained inside a polytope with a polymatroid structure. A sub-optimal coder-decoder construction is provided and its optimality is shown in some limiting cases. The main difficulty in stabilizing a multi-dimensional system over time-varying channels consists of allocating optimally the bits to each unstable sub-system. The scheme proposed can be applied to *any* rate distribution. For some specific rate distributions, however, it is possible to design more efficient schemes. We illustrate this point at the end of this section, studying the specific problem of stabilization over a binary erasure channel, for which a better scheme is proposed.

A. Real Jordan Form

As usual, it is convenient to put $\mathbf{A}$ into real Jordan canonical form [10] so as to decouple its dynamical modes. Denote the system matrix in real Jordan canonical form as $\mathbf{J}$. The matrices $\mathbf{J}$ and $\mathbf{A}$ are related via a similarity matrix $\mathbf{T}$ such that $\mathbf{T}^{-1}\mathbf{J}\mathbf{T} = \mathbf{A}$. Let $\lambda_1, \ldots, \lambda_u \in \mathbb{C}$ be the distinct unstable eigenvalues (if λ_i is non-real, we exclude from this list the complex conjugates λ_i^*) of $\mathbf{A}$, and let m_i be the algebraic multiplicity of each λ_i. The real Jordan canonical form $\mathbf{J}$ then has the block diagonal structure $\mathbf{J} = \operatorname{diag}(\mathbf{J}_1, \ldots, \mathbf{J}_u) \in \mathbb{R}^{d\times d}$, where the block $\mathbf{J}_i \in \mathbb{R}^{\mu_i \times \mu_i}$ and $\det \mathbf{J}_i = \lambda_i^{\mu_i}$, with

$$\mu_i = \begin{cases} m_i & \text{if } \lambda_i \in \mathbb{R} \\ 2m_i & \text{otherwise.} \end{cases}$$

As $\mathbf{A}$ is uniquely composed by unstable systems, we have that $\sum_{i=1}^u \mu_i = d$. Let $\mathcal{U} := [1, \ldots, u]$ denote the index set of unstable systems. Then, the dynamical system equation can be written as

$$\mathbf{x}_{k+1} = \mathbf{J}\mathbf{x}_k + \mathbf{T}\mathbf{B}\mathbf{u}_k + \mathbf{T}\mathbf{v}_k \in \mathbb{R}^d, \quad \mathbf{y}_k = \mathbf{C}\mathbf{T}^{-1}\mathbf{x}_k + \mathbf{w}_k \in \mathbb{R}^p \tag{17}$$

with $\mathbf{x}_k = [\mathbf{x}_k^{(1)}, \ldots, \mathbf{x}_k^{(u)T}]^T \in \mathbb{R}^d$, and where each subsystem $\mathbf{x}_k^{(i)}$ evolves according to

$$\mathbf{x}_{k+1}^{(i)} = \mathbf{J}_i\mathbf{x}_k^{(i)} + (\mathbf{T}\mathbf{B}\mathbf{u}_k)^{(i)} + (\mathbf{T}\mathbf{v}_k)^{(i)} \in \mathbb{R}^{\mu_i}, \quad i \in \mathcal{U}.$$

As the states of (17) and (2) are related through the transformation matrix $\mathbf{T}$, in the following we will assume that the system evolves according to (17).

B. Necessity

Theorem 5.1: Under assumptions **A0.–A3.** above, necessary condition for stabilizability of the system in (17) in the mean square sense (3) is that $(\log_2 |\lambda_1|, \ldots, \log_2 |\lambda_u|) \in \mathbb{R}_+^u$ satisfy, for all $s_i \in \{0, \ldots, m_i\}$ and $i \in \mathcal{U}$

$$\sum_{i\in\mathcal{U}} a_i s_i \log_2 |\lambda_i| < -\frac{\mu'(\mathbf{s})}{2n} \log_2 \mathbb{E}\left[2^{-\frac{2n}{\mu'(\mathbf{s})}R} \right] \tag{18}$$

wherein $\mu'(\mathbf{s}) \equiv \sum_{i\in\mathcal{U}} a_i s_i$, and $a_i = 1$ if $\lambda_i \in \mathbb{R}$, and $a_i = 2$ otherwise.

The following example highlights the special geometric structure of the region defined by (18):

Example 5.1: Consider a two-mode system with two distinct eigenvalues (λ_1, λ_2), where λ_1 is complex and has dimensionality $m_1 = 1$ (so $\mu_1 = 2$) while λ_2 is real and has dimensionality $m_2 = \mu_2 = 1$. Suppose that the digital channel in the feedback link is an erasure channel. Computing the bounds in (18) we obtain the following necessary conditions on $(\log_2 |\lambda_1|, \log_2 |\lambda_2|)$ for stabilizability:

$$2\log_2 |\lambda_1| < -\frac{1}{n} \log_2 \left[p + (1-p)2^{-nr} \right]$$

$$\log_2 |\lambda_2| < -\frac{1}{2n} \log_2 \left[p + (1-p)2^{-2nr} \right]$$

$$2\log_2 |\lambda_1| + \log_2 |\lambda_2| < -\frac{3}{2n} \log_2 \left[p + (1-p)2^{-2n/3r} \right]. \tag{19}$$

In general, these three bounds define a pentagon in the $(\log_2 |\lambda_1|, \log_2 |\lambda_2|)$ domain. In Fig. 2 the boundaries of this pentagon are plotted as dashed lines in the case $n = 6$, $r = 1$ and $p = 1/3$. In some limiting cases, however, the pentagon reduces to a square or a triangle. On the one hand, in the limit of r going to infinity the third constraint in (19) becomes inactive and the pentagonal region reduces to the square determined by the first two inequalities. On the other hand, in the limit of p going to zero the only active constraint is the third inequality, and the region determined by (19) is triangular.

Proof: Consider the system in (17). Notice that each block $\mathbf{J}_i$ has an invariant real subspace of dimension $a_i s_i$, for *any* $s_i \in \{0, \ldots, m_i\}$. Consider the subspace $\mathcal{A}$ formed by taking the product of any of the invariant real subspaces for each real Jordan block. The total dimension of $\mathcal{A}$ is $\mu'(\mathbf{s}) = \sum_{i=1}^u a_i s_i$, for some $s_i \in \{0, \ldots, m_i\}$. Denote by $S = \{e_1, \ldots, e_n\} \subseteq \{1, 2, \ldots, d\}$ the index set of the components of $\mathbf{x}$ belonging to $\mathcal{A}$.

Suppose that a genie helps the decoder by stabilizing all the unstable states that are not in $\mathcal{A}$. Thus, stack the remaining unstable subsystem states to construct a new state $\mathbf{x}_k^S \in \mathbb{R}^{\mu'(\mathbf{s})}$

$$\mathbf{x}_k^S = \left[\mathbf{x}_k^{(e_1)T}, \ldots, \mathbf{x}_k^{(e_n)T} \right]^T := \mathbf{R}\mathbf{x}_k \in \mathbb{R}^{\mu'(\mathbf{s})}$$

where $\mathbf{R}$ is some transformation matrix. Observe that $\mathbf{x}_k^S$ evolves as follows:

$$\mathbf{x}_{k+1}^S = \mathbf{J}^S\mathbf{x}_k^S + \mathbf{R}\mathbf{T}\mathbf{u}_k + \mathbf{R}\mathbf{v}_k, \text{ where } \det \mathbf{J}^S = \prod_{i\in S} \lambda_i^{a_i s_i}.$$
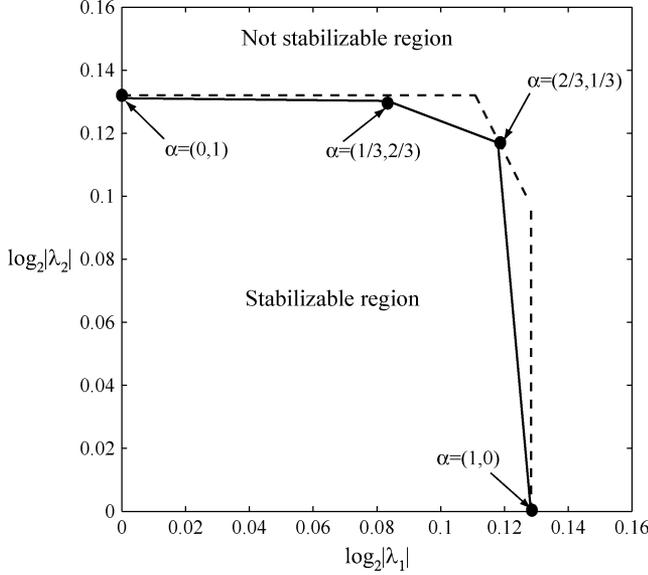
Fig. 2. Stabilizability region for the system described in Example 5.2.

In order to prove the statement, we find a lower bound for $\mathbb{E}[\|\mathbf{x}_k^S\|^2]$, and show that (18) is a necessary condition for the lower bound to be finite. As in Theorem 4.1, the lower bound is given by $n_j^S = (1/2\pi e)\mathbb{E}_{\bar{S}_0^{k-1}}[e^{(2/\mu'(\mathbf{s}))h(\mathbf{x}_k^S|\bar{S}_0^{k-1}=\bar{\mathbf{s}}_0^{k-1})}]$.

Proceeding as in Theorem 4.1, one can derive a recursive formula for n_j^S of the form

$$n_{j+1}^S \geq \mathbb{E}\left[\frac{|\det \mathbf{J}^S|^{\frac{2n}{\mu'(\mathbf{s})}}}{2^{\frac{2n}{\mu'(\mathbf{s})}R}}\right] n_j^S + \gamma^S$$

for some constant $\gamma^S < \infty$. Therefore, $|\prod_{i\in S}|\lambda_i|^{|a_is_i|(2n/\mu'(\mathbf{s}))}\mathbb{E}[2^{-(2n/\mu^S)R}] \geq 1$ implies that $\sup_{j\in\mathbb{N}} n_j^S = \infty$. ∎

The region determined in Theorem 5.1 has a special combinatorial structure. The polytope (18) is defined by the set function $f(S) := -(|S|/2n)\log_2 \mathbb{E}[2^{-2n/|S|R}]$, $\forall S \subset E := \{1, \ldots, d\}$. It is shown in the following proposition, proved in the Appendix, that this set function defines a polymatroid.

Proposition 5.2: The polytope defined by (18) is a polymatroid.

Remarks:

1) When the rate process is constant, the constraints in (18) reduce to the well known condition [16], [21]

$$\sum_{|\eta_i|\geq 1} \log_2 |\eta_i| \equiv \sum_{i\in\mathcal{U}} \mu_i \log_2 |\lambda_i| < R$$

and the stabilizability is contained in the region in the positive orthant strictly inside the hyperplane $\sum_{i\in\mathcal{U}} \mu_i \log_2 |\lambda_i| = R$.

2) Notice that the right hand side of (18) can be rewritten as

$$-\frac{\mu'(\mathbf{s})}{2n}\log_2 2^{-\frac{2n}{\mu'(\mathbf{s})}r_{min}}\mathbb{E}\left[2^{-\frac{2n}{\mu'(\mathbf{s})}(R-r_{min})}\right]$$

$$= r_{min} - \frac{\mu'(\mathbf{s})}{2n}\log_2 \mathbb{E}\left[2^{-\frac{2n}{\mu'(\mathbf{s})}(R-r_{min})}\right]$$

$$\to r_{min}$$

as $n \to \infty$. Thus, in the limit of n going to infinity, (18) reduce to

$$\sum_{i\in\mathcal{U}} \mu_i \log_2 |\lambda_i| < r_{min} \qquad (20)$$

and the stabilizability region is determined uniquely by r_{min}. The intuitive justification of this latter fact is that the digital link supports the same rate for an arbitrarily long time interval, so stability has to be guaranteed under the worst possible rate. In the limit, stabilization is not possible for those channels where $r_{min} = 0$ (e.g. erasure channels).

3) In an erasure channel, for a fixed n, as r goes to infinity the stabilizability reduces to the n-dimensional cube described by

$$\log_2 |\lambda_i| < \frac{1}{2n}\log_2 \frac{1}{p} \quad \forall i \in \mathcal{U}. \qquad (21)$$

In other words, the system in (17) cannot be stabilized if the erasure probability is such that

$$p \geq \frac{1}{\max_{i\in\mathcal{U}} \lambda_i^{2n}}.$$

In the case $n = 1$, this is the same condition derived in [9] in the context of the LQG problem with erasures.

C. Sufficiency

We now present a sufficient condition for mean-square stabilizability of the multi-dimensional system (17). The scheme is based on the adaptive quantizer introduced in Section IV-B-1. We introduce a rate allocation vector which indicates what fraction of the available rate is allocated to each unstable sub-system.

Theorem 5.3: Under assumptions **A0.–A3.** above, sufficient condition for stabilizability of the system in (17) in the mean square sense (3) is that $(\log_2 |\lambda_1|, \ldots, \log_2 |\lambda_u|) \in \mathbb{R}_+^u$ are inside the convex hull of the region determined by

$$\log_2 |\lambda_i| < -\frac{1}{2n}\log_2 \mathbb{E}\left[2^{-\frac{2n}{\mu_i}\alpha_i(R)R}\right], \forall i \in \mathcal{U} \qquad (22)$$

for some rate allocation vector $\boldsymbol{\alpha}(R) := [\alpha_1(R), \ldots, \alpha_u(R)]^T$ satisfying

$$\begin{cases} \alpha_i(r) \in [0, 1] \\ \frac{1}{\mu_i}\alpha_i(r)nr \in \mathbb{N} \qquad \forall r \in \mathcal{R} \setminus \{0\}, i \in \mathcal{U}. \\ \sum_{i=1}^u \alpha_i(r) \leq 1 \end{cases} \qquad (23)$$

Suppose that transmission of r bits per channel use is supported on the digital link in a given block. The rate allocation vector $\boldsymbol{\alpha}(r)$ indicates what fraction of the total nr bits transmitted in a block is allocated to each sub-system. All μ_i modes in the ith sub-system are quantized using $\alpha_i(r)nr/\mu_i$ bits. Condition (23) requires that $\alpha_i(r)nr/\mu_i$ is an integer number for all $i \in \mathcal{U}$, and that the total number of bits used in each block should not exceed nr. Such conditions define finitely many rate allocation vectors, and for each allocation vector (22) defines a cube in the space of $(\log_2 |\lambda_1|, \cdots, \log_2 |\lambda_u|)$. By using a time-sharing protocol among different rate allocation vectors it is possible to stabilize those points inside the convex hull of the union of such

cubes. Before looking at the proof of the Theorem, consider the following Example:

Example 5.2: Consider the system in Example 5.1 and assume that $n = 6$ and $r = 1$. Under this channel model, (23) defines four allocations vectors, namely $\alpha_1(1) = 1 - \alpha_2(1) = j/6$, $j \in \{0, 2, 4, 6\}$. For each allocation vector, (22) defines a cube in the space of $(\log_2 |\lambda_1|, \log_2 |\lambda_2|)$, and the stability region defined by Theorem 5.3 is the convex hull of the union of such cubes. Fig. 2 shows the boundaries of the achievable stabilizability region in the case $p = 1/3$: vertexes of the cube defined by (22) are represented as dots, while the solid lines show the convex hull of the union of such cubes. Notice that the outer bounds defined by (19) are achieved in three points, two of which lie on the two axis and correspond to the case where only one of the two sub-systems is unstable. In these cases the optimal rate allocation consists of allocating all the available bits to the unstable mode. The third optimal point corresponds to the case where the two eigenvalues have the same magnitude, i.e. $|\lambda_1| = |\lambda_2|$, and the optimal allocation in this case is to allocate one bit to each unstable mode. We will see that a protocol that time-shares among these three points is optimal in the limit $r \to \infty$.

Proof: The proof is divided into two parts. First it is shown that the linear dynamical system in (17) is stabilizable if (22) holds for some rate allocation vector $\boldsymbol{\alpha}(R)$ satisfying (23). Second it is shown that, by using a time-sharing protocol, all the points in the convex hull can be stabilized.

The coder computes a minimum variance estimator $\bar{x}^{(i,h)}$ for the hth component of the ith unstable mode, $h \in \{1, \ldots, \mu_i\}$ and $i \in \mathcal{U}$. Similarly, coder and decoder compute an estimator $\hat{x}^{(i,h)}$. Define $f_k^{(i,h)} = \bar{x}_k^{(i,h)} - \hat{x}_k^{(i,h)}$ as the error between these two estimators at time k. Let the stacked vector of unstable subsystems errors be $\mathbf{f}_k = \hat{\mathbf{x}}_k - \bar{\mathbf{x}}_k$.

Suppose that coder and decoder agree, ahead of time, on some rate allocation $\boldsymbol{\alpha}(R)$ satisfying (23). As in the case of a scalar system, divide times $k \in \mathbb{N}$ into cycles of integer duration $n\tau$, $\tau \in \mathbb{Z}_+$. Let

$$R_k^{(i)} := \frac{\alpha_i(R_k) R_k}{\mu_i} \in \mathbb{N}, \quad i \in \mathcal{U}, k \in \mathbb{N}$$

denote the number of bits allocated to the transmission of $f_k^{(i,h)}$ during the kth channel block. By (23), $\sum_{i \in \mathcal{U}} \mu_i n R_k^{(i)} \leq n R_k \forall k \in \mathbb{N}$.

Therefore, at time $k = jn\tau$, the coder computes, for all $h \in \{1, \ldots, \mu_i\}$ and for all $i \in \mathcal{U}$

$$\bar{\omega}_{n R_{j\tau}^{(i)}} \left(\omega_{jn\tau}^{(i,h)} \right) = q_{n R_{j\tau}^{(i)}} \left(f_{jn\tau}^{(i,h)} / l_j \right).$$

The scaling factor l_j is updated at the beginning of each cycle as follows:

$$l_{j+1} = \max_{\substack{h \in [1, \ldots, \mu_i] \\ i \in \mathcal{U}}} \left\{ \sigma, l_j |\lambda_i|^{n\tau} \kappa_{\nu_j^{(i)}} \left(\omega_{(j+1)n\tau - n}^{(i,h)} \right) \right\}$$

where the random variable

$$\nu_j^{(i)} = \sum_{k=0}^{\tau-1} n R_{j\tau+k}^{(i)}$$

indicates the cumulative number of bits allocated to the ith sub-system during the jth cycle, and where $l_0 = \sigma$ and $\sigma^{2+\epsilon}$ is a uniform bound on the $(2 + \epsilon)$-moment of $\mathbf{g}_j := \sum_{i=0}^{n\tau-1} \mathbf{J}^{n\tau-1-i} \mathbf{z}_{jn\tau+i}$, $j \in \mathbb{N}$. After the first block in the cycle, the decoder identifies an uncertainty interval $I_{n R_{j\tau}^{(i)}}(\omega_{jn\tau}^{(i,h)})$ for each unstable sub-system. The remaining $(n - 1)\tau$ transmissions in the cycle are devoted to reducing the size of the uncertainty interval. After receiving the last $R_{(j+1)\tau-1}^{(i)}$ bits, the decoder can compute the final uncertainty interval $I_{\nu_j^{(i)}}(\omega_{(j+1)n\tau-n}^{(i,h)})$, corresponding to the uncertainty set formed by the quantizer $q_{\nu_j^{(i)}}(f_{jn\tau}^{(i,h)/l_j})$. For each unstable subsystem, the decoder sends to the plant a certainty-equivalent control signal $\mathbf{u}_k = P\hat{\mathbf{x}}_k$ as in (13).

Let $\theta_j := M_\epsilon[\|\mathbf{f}_{jn\tau}\|, l_j]$. Proceeding along the same lines as in the scalar case, it can be shown that θ_j evolves according to the following recursive equation:

$$\theta_{j+1} \leq 2\phi\sigma^2 + \phi \sum_{i \in \mathcal{U}} \mu_i |n\tau^{\mu_i - 1}|^{2+\epsilon} \zeta \left(\mathbb{E}\left[\frac{\lambda_i^{2n}}{2^{2n\alpha_i(R)R/\mu_i}} \right] \right)^\tau \theta_j.$$

Hence, if (22) is satisfied, by choosing a τ sufficiently large, the coefficient of θ_j can be made strictly less than 1. Therefore, the recursion above is stable and yields uniformly bounded θ_j. The same argument used for the scalar case applies *sic et simpliciter* and it is now straightforward to show that the system is second moment stable.

It remains to show that, by time-sharing, all the points in the convex hull can be stabilized. Since the union of finite cubes in $\mathbb{R}^u$ is a connected compact set, by the Fenchel-Eggleston theorem [8, Theorem 18] each point in its convex closure can be represented as a convex combination of at most u points in the union, and thus each point is in the convex closure of the union of no more than u cubes in (22). Given u rate allocation vectors $\boldsymbol{\alpha}^{(l)}(R)$, $l = 1, \cdots, u$, satisfying (23), and any $\gamma_l \in [0, 1]$ such that $\sum_{l=1}^{u} \gamma_l = 1$, it suffices to construct a scheme that stabilizes all modes $(\log_2 |\lambda_1|, \ldots, \log_2 |\lambda_u|)$ inside the region

$$\log_2 |\lambda_i| < -\sum_{l=1}^{u} \frac{\gamma_l}{2n} \log_2 \mathbb{E}\left[2^{-2n\alpha_i^{(l)}(R)R/\mu_i} \right] \quad \forall i \in \mathcal{U}.$$

(24)

Divide times $k \in \mathbb{N}$ into cycles of duration τn, in such a way that $\gamma_l \tau n \in \mathbb{N}$ for $l = 1, \cdots, u$. During a fraction γ_l of the cycle allocate bits utilizing rate allocation vector $\boldsymbol{\alpha}^{(l)}(R)$. Repeating the analysis above, it can be proved that the crucial recursion for θ_j evolves as follows:

$$\theta_{j+1} \leq 2\phi\sigma^2 + \phi \sum_{i \in \mathcal{U}} \mu_i |n\tau^{\mu_i - 1}|^{2+\epsilon} \zeta \times$$
$$\left(\lambda_i^{2n} \prod_{l=1}^{u} \mathbb{E}\left[\frac{1}{2^{2n\alpha_i^{(l)}(R)R/\mu_i}} \right]^{\gamma_l} \right)^\tau \theta_j.$$

If (24) holds, we can choose τ large enough to make the recursion stable. Therefore, (24) are sufficient conditions for stabilizability. ∎

Remarks:

1) If $(rn/d) \in \mathbb{N}$ for all $r \in \mathcal{R}$, then the rate allocation $\alpha_i(r) = (\mu_i/d) \forall r \in \mathcal{R} \setminus \{0\}$ is optimal when $\lambda :=$

$\lambda_1 = \ldots = \lambda_u$. In fact, from (22) sufficient condition for stabilizability is that

$$\log_2|\lambda| < -\frac{1}{2n}\log_2\mathbb{E}\left[2^{-\frac{2n}{d}R}\right].$$

On the other hand, this condition is also necessary, as we can see from (18) by letting $s_i = m_i$ for all $i \in \mathcal{U}$. For example, in Example 5.2 we have that $nr/d = 2$, so the rate allocation $\boldsymbol{\alpha} = (2/3, 1/3)$ is optimal (See Fig. 2).

2) The scheme in Theorem 5.3 is optimal in the limit of n going to infinity, and the optimal coding scheme consists of a time-sharing protocol among the rate allocations $\boldsymbol{\alpha}^{(i)}(R) = \boldsymbol{e}_i$ for all $i \in \mathcal{U}$, where $\{\boldsymbol{e}_i\}_{i=1}^d$ are the canonical basis vectors of $\mathbb{R}^d$.

3) In an erasure channel, for a fixed n, as r goes to infinity the proposed achievable scheme is asymptotically optimal. The stabilizability region is given by the cube (21), and the optimal coding scheme consists of time-sharing among the rate distributions $\boldsymbol{\alpha}^{(i)}(r) = \boldsymbol{e}_i$ for all $i \in \mathcal{U}$ and the allocation given in Remark 1., i.e. $\alpha_i^{(u+1)}(r) = \mu_i/d$.

4) When the rate process is constant, Nair and Evans [16] showed that the necessary and sufficient conditions coincide. Once again, the optimal coding scheme consists of a time-sharing protocol among the rate distributions $\boldsymbol{\alpha}^{(i)}(R) = \boldsymbol{e}_i$ for all $i \in \mathcal{U}$.

5) A more general scheme is easily derived by allowing the rates allocated to each component of the same sub-system to be different. For ease of exposition, in Theorem 5.3 we assumed these rates to be equal.

D. Binary Erasure Channel

The stabilizing scheme proposed in the previous section provides an achievability result for stabilization over time-varying channels, and is optimal is some limiting cases. However, the scheme is not optimal in general. In this section, we improve the stabilizability region defined by Theorem 5.3 in the specific case of stabilization over a binary erasure channel. Before stating the result, we outline the main difference between the coding scheme used in this section and the construction in Theorem 5.3. In Theorem 5.3 time is divided into slots of fixed duration, and system state observations are quantized using a random number of bits dependent on the realization of the rate process. In this section, instead, we present a coder/decoder construction which is based on an alternative approach: state observations are quantized using a fixed number of bits per unstable mode; in turn, these are transmitted to the decoder over a random number of discrete time units which depends on the realization of the rate process. Based on this approach, it is possible to enlarge the set of feasible rate allocation vectors and, as a consequence, the stabilizability region. In this section, the following simplifying assumptions are made:

A4. The decoder has access to state feedback, i.e. in (2) we have that $C = I$ and $\mathbf{w}_k = 0$ for all $k \in \mathbb{N}$.

A5. $\exists W < \infty$ such that $|\mathbf{v}_k^{(i)}| \leq W$ uniformly in $k \in \mathbb{N}$, and $\mathbf{x}_0 \in [-(1/2); 1/2]^d$.

A6. The feedback digital link is a binary erasure channel, and the block length is $n = 1$.
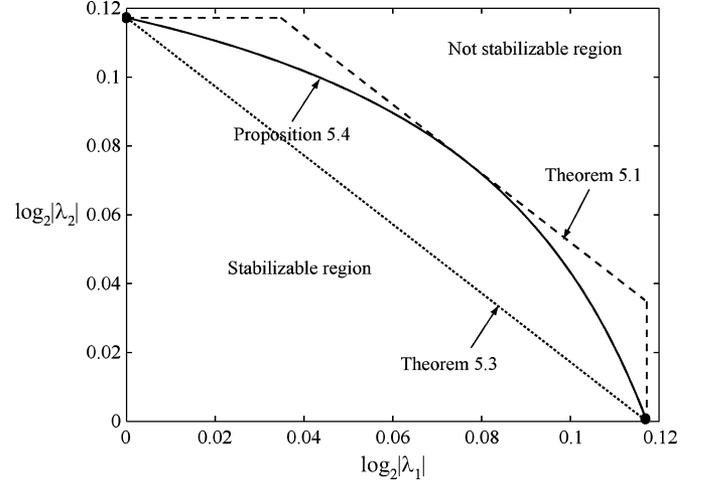
We have the following proposition:

Fig. 3. Stabilizability region for the system described in Example 5.3.

Proposition 5.4: Under assumptions **A0.–A6.** above, sufficient condition for stabilizability of the system in (2) in the mean square sense (3) is that $(\log_2|\lambda_1|, \ldots, \log_2|\lambda_u|)\mathbb{R}_+^u$ are inside the convex region determined by

$$\log_2|\lambda_i| < -\frac{1}{2}\log_2\mathbb{E}\left[2^{2\frac{\alpha_i}{\mu_i}R}\right], \quad i \in \mathcal{U} \qquad (25)$$

for some rate allocation vector $\boldsymbol{\alpha} := [\alpha_1, \ldots, \alpha_u]$ such that

$$\begin{cases} \alpha_i \in [0,1] \cap \mathbb{Q}, \ \forall i \in \mathcal{U} \\ \sum_{i=1}^u \alpha_i \leq 1. \end{cases} \qquad (26)$$

Comparing (23) and (26), notice that while in Theorem 5.3 only a *finite* number of rate allocation vectors satisfy (23), the region defined by Proposition 5.4 is given by the union of a *countable* number of u-dimensional cubes, each of which is defined by (25) for some rate allocation vector satisfying (26). We also notice that the stabilizability region defined by Proposition 5.4 is convex, so a time-sharing protocol among different rate allocation policies is not required.

Example 5.3: Consider a system with two distinct modes of dimensionality one, having unstable real eigenvalues λ_1 and λ_2, respectively. Fig. 3 shows the achievable stabilizability region under this channel model, assuming $p = 2/3$. The boundaries of the region defined by Proposition 5.4 are represented as a solid curve, and each point on this curve is obtained by (25) for some choice of the rate allocation vector. The region in Theorem 5.3 is delimited by a dotted line, which represents the convex combination of two points (bold dots), obtained by (23) with $\boldsymbol{\alpha}(1) = [1,0]$ and $\boldsymbol{\alpha}(1) = [0,1]$. Finally, the necessary conditions derived in Theorem 5.1 define a pentagon that is delimited by a dashed line. The region in Proposition 5.4 is optimal at the intersections with the two axis and at one point on the bisectrix $|\lambda_1| = |\lambda_2|$.

Proof: Fix an $\boldsymbol{\alpha} \in [0,1]^u$ satisfying (26) and $m \in \mathbb{Z}_+$ such that $\alpha_i m/\mu_i \in \mathbb{N}$ for all i. A renewal process $\{t_k\}_{k=1}^\infty$ determines the times at which the encoder quantizes the state observations. The random interarrival times of this renewal process are denoted by the sequence $\{\tau_k\}_{k=1}^\infty$, such that $t_k = \tau_1 + \cdots + \tau_k$ for all $k \in \mathbb{N}$.

The stability in the mean square sense of the system in (2) is proved by showing that for each unstable sub-system $x^{(i,h)}$, $h \in [0, \ldots, \mu_i]$ and $i \in \mathcal{U}$, there exists a mean square stable sequence $\{z_k^{(i,h)}\}_{k=0}^{\infty}$ such that $|x_{t_k}^{(i,h)}| \leq z_k^{(i,h)}$, for all $k \in \mathbb{N}$. We define $\{z_k^{(i,h)}\}_{k=0}^{\infty}$ recursively as follows:

$$
\begin{cases}
z_1^{(i,h)} = 1 + W \\
z_{k+1}^{(i,h)} = \frac{|\lambda_i|^{\tau_k}}{2^{\alpha_i m/\mu_i}} z_k^{(i,h)} + \eta
\end{cases} \tag{27}
$$

where $\eta = W/(1 - |\lambda_i|)$ if $|\lambda_i| > 1$ and $\eta = \tau_k W$ if $|\lambda_i| = 1$. At the random time t_k, the encoder partitions the interval $[-z_k^{(i,h)}; +z_k^{(i,h)}]$ into $2^{\alpha_i m/\mu_i}$ uniform intervals, and computes $\hat{x}_{t_k}^{(i,h)}$ as the center of the interval containing $x_{t_k}^{(i,h)}$. By construction, the approximation error satisfies

$$
\left| x_{t_k}^{(i,h)} - \hat{x}_{t_k}^{(i,h)} \right| \leq \frac{z_k^{(i,h)}}{2^{\alpha_i m/\mu_i}}. \tag{28}
$$

The time required for transmission of the cumulative m bits describing the quantized source symbols from coder to decoder is denoted by the interarrival time τ_k. We define $\tau := \inf\{k : \sum_{l=1}^{k} R_l = m\}$ as the time of the mth 'success' in the Bernoulli process $\{R_i\}_{i=1}^{\infty}$; for any $p > 0$, we have that $\Pr(\tau < \infty) = 1$, and τ has negative binomial distribution with parameters m and p. The interarrival times $\{\tau_k\}_{k=1}^{\infty}$ are independent non-negative random variables, identically distributed as τ.

At time $t_{k+1} = t_k + \tau_k$, upon reception of the m binary source symbols the decoder computes the control signal

$$
u_{t_{k+1}}^{(i,h)} = -\lambda_i^{\tau_k} \hat{x}_{t_k}^{(i,h)}. \tag{29}
$$

Making use of (27), (28) and (29), we have the following chain of inequalities

$$
\begin{aligned}
\left| x_{t_{k+1}}^{(i,h)} \right| &\leq \left| \lambda_i^{\tau_k} x_{t_k}^{(i,h)} + u_{t_{k+1}}^{(i,h)} \right| + \sum_{j=0}^{\tau_k - 1} |\lambda_i|^{\tau_k - 1 - j} |v_{t_k + j}| \\
&\leq |\lambda_i|^{\tau_k} \left| x_{t_k}^{(i,h)} - \hat{x}_{t_k}^{(i,h)} \right| + \eta \\
&\leq z_{k+1}^{(i,h)}.
\end{aligned} \tag{30}
$$

From (30) and proceeding by induction, it follows that $|x_{t_k}^{(i,h)}| \leq z_k^{(i,h)}$, for all $k \in \mathbb{N}$. Next, we show that (25) is a sufficient condition for the sequence $\{z_k^{(i,h)}\}_{k=0}^{\infty}$ to be mean square stable, i.e. $\sup_{k \in \mathbb{N}} \mathbb{E}[|z_k^{(i,h)}|^2] < \infty$, for all $h \in [0, \ldots, \mu_i]$ and $i \in \mathcal{U}$. From (27) and the triangle inequality, it follows that:

$$
\left(\mathbb{E}\left[\left| z_{k+1}^{(i,h)} \right|^2 \right] \right)^{\frac{1}{2}} \leq \left(\mathbb{E}\left[\frac{\lambda_i^{2\tau_k}}{2^{2\alpha_i m/\mu_i}} \right] \right)^{\frac{1}{2}} \left(\mathbb{E}\left| z_k^{(i,h)} \right|^2 \right)^{\frac{1}{2}} + \left(\mathbb{E}\left[|\eta|^2 \right] \right)^{\frac{1}{2}} \quad \forall i \in \mathcal{U} \quad (31)
$$

wherein $\mathbb{E}[|\eta|^2] < \infty$ as $\mathbb{E}[\tau^2] = (n(1-p)/p^2) < \infty$ for all $p \in (0,1)$. By writing explicitly the expectations in (25), we obtain that

$$
\lambda_i^2 \left[\frac{p}{2^{2\alpha_i/\mu_i}} + (1-p) \right] < 1 \quad \forall i \in \mathcal{U}. \tag{32}
$$

Making use of (32) simple algebra shows that

$$
\mathbb{E}\left[\frac{\lambda_i^{2\tau_k}}{2^{2\alpha_i m/\mu_i}} \right] = \left(\frac{1}{2^{2\alpha_i/\mu_i}} \frac{p|\lambda_i|^2}{1 - (1-p)e^{it}} \right)^m < 1 \quad \forall i \in \mathcal{U}. \tag{33}
$$

From (33) it follows that the recursive formula in (31) is stable. Therefore, (25) is a sufficient condition to ensure $\sup_k \mathbb{E}[|z_k^{(i,h)}|^2] < \infty$.

Finally, the convexity of the region described by (25) follows from Jensen's inequality applied to the concave function $\phi(x) = -(1/2)\log_2((1-p) + p2^{-2x})$. $\blacksquare$

VI. CONCLUSION

Motivated by control problems over time-varying channels, we considered mean square stabilizability of a discrete-time, linear system with a noise free time-varying digital communication link. Process and observation disturbances were allowed to occur over an unbounded support. Necessary conditions were derived employing information-theoretic techniques, while a stabilization scheme based on an adaptive successively refinable quantizer was constructed. In the scalar case, this scheme was shown to be optimal. Furthermore, we have shown that in the vector case the necessary condition for stabilization has an interesting polymatroid structure, and have proposed a stabilization scheme that is optimal in some limiting regimes. An additional contribution is that we bridged the information-theoretic results of stabilization over rate limited channels, with the corresponding network-theoretic ones on critical dropout probabilities in systems with unbounded disturbances. We have done so by recovering the latter results as a special case of our analysis.

APPENDIX

A. Proof of Lemma 4.2

Proof: First, observe that the following chain of inequalities holds:

$$
\begin{aligned}
\mathbb{E}_{\bar{S}_j | \bar{S}_{j-1}, R_j} & h(x_{jn} | \bar{S}_{j-1} = \bar{s}_{j-1}, \bar{S}_j = \bar{s}_j, R_j) \\
&= h(x_{jn}, \bar{S}_j | \bar{S}_{j-1} = \bar{s}_{j-1}, R_j) - H(\bar{S}_j | \bar{S}_{j-1} = \bar{s}_{j-1}, R_j) \\
&\geq h(x_{jn} | \bar{S}_{j-1} = \bar{s}_{j-1}, R_j) - H(\bar{S}_j | \bar{S}_{j-1} = \bar{s}_{j-1}, R_j) \\
&\geq h(x_{jn} | \bar{S}_{j-1} = \bar{s}_{j-1}, R_j) - \ln 2^{nR_j} \\
&= h(x_{jn} | \bar{S}_{j-1} = \bar{s}_{j-1}) - \ln 2^{nR_j}
\end{aligned} \tag{34}
$$

where $h(x, A|B)$ with A discrete denotes $-\mathbb{E}[\ln(p_{A|B} f_{x|A,B})]$. The last inequality follows from the fact that, given R_j, the cardinality of $\{S_{jn}, \ldots, S_{(j+1)n-1}\}$ is 2^{nR_j}, and where the last equality follows from the fact that $x_{jn} \to \bar{S}_{j-1} \to R_j$ is a Markov chain. Then

$$
\begin{aligned}
\mathbb{E}_{\bar{S}_j | \bar{S}_{j-1}, R_j} & e^{2h(x_{jn} | \bar{S}_j = \bar{s}_j)} \\
&\geq \mathbb{E}_{\bar{S}_j | \bar{S}_{j-1}, R_j} e^{2h(x_{jn} | \bar{S}_{j-1} = \bar{s}_{j-1}, \bar{S}_j = \bar{s}_j, R_j)} \\
&\geq e^{2\mathbb{E}_{\bar{S}_j | \bar{S}_{j-1}, R_j} h(x_{jn} | \bar{S}_{j-1} = \bar{s}_{j-1}, \bar{S}_j = \bar{s}_j, R_j)} \\
&\geq e^{2[h(x_{jn} | \bar{S}_{j-1} = \bar{s}_{j-1}) - \ln 2^{nR_j}]} \\
&= \frac{1}{2^{2nR_j}} e^{2h(x_{jn} | \bar{S}_{j-1} = \bar{s}_{j-1})},
\end{aligned}
$$

where the first inequality follows from the fact that conditioning reduces the entropy; the second inequality follows from Jensen's inequality; finally, (34) implies the third inequality. ∎

B. Proof of Proposition 5.2

Proof: Let $E = \{1, \ldots, d\}$ and $f(S) = -\log_2 \left(\mathbb{E}[(2^{-R})^{2n/|S|}] \right)^{|S|/2n} 1_{|S|>0} = -\log \|2^{-R}\|_{2n/|S|} 1_{|S|>0}$. Following the definition in [6], in order for the polytope

$$\mathcal{B}(f) := \left\{ (x_1, \ldots, x_u) : \sum_{i \in S} x_i \le f(S) \forall S \subset E, x_i \ge 0 \forall i \right\}$$

to be a polymatroid, we need to show the following properties:
1) $f(\emptyset) = 0$: this is immediate from the definition of $f(\cdot)$.
2) $f(S) \le f(T)$ if $S \subset T$: this follows from the fact that $\|X\|_{1/m} \le \|X\|_{1/n}$ if $n \le m$.
3) $f(S) + f(T) \ge f(S \cup T) + f(S \cap T)$: this can be proved as follows. Note that $f(S)$ is a function only of $|S|$, i.e. $f(S) := g(|S|)$. W.l.o.g., assume that $j := |S| \le |T| =: k$. Let $i := |S| - |S \cap T|$ and note that this is never negative. Further note that $|S \cup T| = |S| + |T| - |S \cap T| = k + i$. The desired property is then that $g(j) - g(j - i) \ge g(k + i) - g(k)$ for all integers $i \le j \le k$. Now, from the fundamental theorem of calculus $\exists a \in [j - i, j] \cap \mathbb{R}$ and $\exists b \in [k, k + i] \cap \mathbb{R}$ such that $g(j) - g(j - i) = g'(a)i$ and $g(k + i) - g(k) = g'(b)i$. Thus proving the desired inequality is equivalent to proving that $g'(a) \ge g'(b)$ for all $0 < a \le b$. On the other hand, this inequality follows from the concavity of the function $g(x)$ for $x > 0$. ∎

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for detailed comments on the paper which helped to improve the clarity of the manuscript.

REFERENCES

[1] J. Baillieul, "Feedback designs in information-based control," in *Proc. Workshop Stochastic Theory Control*, B. Pasik-Duncan, Ed., Lawrence, Kansas, Oct. 2001, pp. 35–57.
[2] J. Baillieul and P. Antsaklis, "Control and communication challanges in networked real time systems," *Proc. IEEE, Special Iss. Emerg. Technol. Netw. Control Syst.*, vol. 95, no. 1, pp. 9–28, Jan. 2007.
[3] R. W. Brockett and D. Liberzon, "Quantized feedback stabilization of linear systems," *IEEE Trans. Automat. Control*, vol. 45, no. 7, pp. 1279–1289, Jul. 2000.
[4] T. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiley, 1987.
[5] D. F. Delchamps, "Stabilizing a linear system with quantized state feedback," *IEEE Trans. Automat. Control*, vol. 35, no. 8, pp. 916–924, Aug. 1990.
[6] J. Edmonds, "Submodular functions, matroids and certain polyhedra," in *Proc. Calgary Int. Conf. Combinatorial Struct. Appl.*, Calgary, AB, Canada, Jun. 1969, pp. 69–87.
[7] N. Elia, "Remote stabilization over fading channels," *Syst. Control Lett.*, vol. 54, no. 3, pp. 237–249, Mar. 2005.
[8] H. G. Eggleston, *Convexity*. Cambridge, U.K.: Cambridge University Press, 1963.
[9] V. Gupta, B. Hassibi, and R. M. Murray, "Optimal LQG control across packet-dropping links," *Syst. Control Lett.*, vol. 56, no. 6, pp. 439–446, Jun. 2007.
[10] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge University Press, 1985.
[11] J. P. Hespanha, P. Naghshtabrizi, and Y. Xu, "A survey of recent results in networked control systems," *Proc. IEEE, Special Iss. Emerg. Technol. Netw. Control Syst.*, vol. 95, no. 1, pp. 138–162, Jan. 2007.
[12] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, "Kalman filtering with intermittent observations," *IEEE Trans. Automat. Control*, vol. 49, no. 9, pp. 1453–1464, Sep. 2004.
[13] N. C. Martins, M. A. Dahleh, and N. Elia, "Feedback stabilization of uncertain systems in the presence of a direct link," *IEEE Trans. Automat. Control*, vol. 51, no. 3, pp. 438–447, Mar. 2006.
[14] A. S. Matveev and A. V. Savkin, "Comments on 'Control over noisy channels' and relevant negative results," *IEEE Trans. Automat. Control*, vol. 50, no. 12, pp. 2105–2110, Dec. 2005.
[15] A. S. Matveev and A. V. Savkin, "An analogous of Shannon information theory for networked control systems," in *Proc. IEEE Conf. Decision Control*, 2004, pp. 4491–4496.
[16] G. N. Nair and R. J. Evans, "Stabilizability of stochastic linear systems with finite feedback data rates," *SIAM J. Control Optim.*, vol. 43, no. 2, pp. 413–436, Jul. 2004.
[17] G. N. Nair, F. Fagnani, S. Zampieri, and R. J. Evans, "Feedback control under data rate constraints: An overview," *Proc. IEEE, Special Iss. Emerg. Technol. Netw. Control Syst.*, vol. 95, no. 1, pp. 108–137, Jan. 2007.
[18] A. Sahai and S. Mitter, "The necessity and sufficiency of anytime capacity for stabilization of a linear system over a noisy communication link Part I: Scalar Systems," *IEEE Trans. Inform. Theory*, vol. 52, no. 8, pp. 3369–3395, Aug. 2006.
[19] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. S. Sastry, "Foundations of control and estimation over lossy networks," *Proc. IEEE, Special Iss. Emerg. Technol. Netw. Control Syst.*, vol. 95, no. 1, pp. 163–187, Jan. 2007.
[20] C. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, 1948.
[21] S. Tatikonda and S. Mitter, "Control under communication constraints," *IEEE Trans. Automat. Control*, vol. 49, no. 7, pp. 1056–1068, Jul. 2004.
[22] S. Tatikonda and S. Mitter, "Control over noisy channels," *IEEE Trans. Automat. Control*, vol. 49, no. 7, pp. 1196–1201, Jul. 2004.
[23] S. Yuksel and T. Basar, "Control over noisy forward and feedback channels," *IEEE Trans. Automat. Control*, submitted for publication.
[24] W. S. Wong and R. W. Brockett, "Systems with finite communication bandwidth constraints, II: Stabilization with limited information feedback," *IEEE Trans. Automat. Control*, vol. 44, no. 5, pp. 1049–1053, May 1999.

Paolo Minero (S'05) received the Laurea degree (with highest honors) in electrical engineering from the Politecnico di Torino, Torino, Italy, in 2003, the M.S. degree in electrical engineering from the University of California at Berkeley in 2006, and is currently pursuing the Ph.D. degree in the Department of Electrical and Computer Engineering, University of California at San Diego.

His research interests are in communication systems theory and include information theory, wireless communication, and control over networks.

Mr. Minero received the U.S. Vodafone Fellowship in 2004 and 2005, and the Shannon Memorial Fellowship in 2008.

Massimo Franceschetti (M'98) received the Laurea degree (with highest honors) in computer engineering from the University of Naples, Naples, Italy, in 1997, and the M.S. and Ph.D. degrees in electrical engineering from the California Institute of Technology, Pasadena, in 1999 and 2003, respectively.

He is an Associate Professor in the Department of Electrical and Computer Engineering, University of California at San Diego. Before joining UCSD, he was a Post-Doctoral Scholar at the University of California at Berkeley for two years. He has held visiting positions at the Vrije Universiteit Amsterdam in the Netherlands, the Ecole Polytechnique Federale de Lausanne in Switzerland, and the University of Trento in Italy. His research interests are in communication systems theory and include random networks, wave propagation in random media, wireless communication, and control over networks.

Dr. Franceschetti received the C. H. Wilts Prize in 2003 for Best Doctoral Thesis in Electrical Engineering at Caltech; the S. A. Schelkunoff award in 2005 for Best Paper in the IEEE Transactions on Antennas and Propagation; an NSF CAREER Award in 2006, and an ONR Young Investigator Award in 2007. He was on the Guest Editorial Board of the IEEE Transactions on Information Theory, special issue on models, theory, and codes, for relaying and cooperation in communication networks; and several issues of the IEEE Journal on Selected Areas in Communications.

Subhrakanti Dey (S'94–M'96–SM'06) was born in Calcutta, India, in 1968. He received the B.Tech. and M.Tech. degrees from the Department of Electronics and Electrical Communication Engineering, Indian Institute of Technology, Kharagpur, in 1991 and 1993, respectively, and the Ph.D. degree from the Department of Systems Engineering, Research School of Information Sciences and Engineering, Australian National University, Canberra, Australia, in 1996.

He has been with the Department of Electrical and Electronic Engineering, University of Melbourne, Parkville, Australia, since 2000, where he is currently a full Professor. From September 1995 to September 1997 and September 1998 to February 2000, he was a postdoctoral Research Fellow with the Department of Systems Engineering, Australian National University. From September 1997 to September 1998, he was a post-doctoral Research Associate with the Institute for Systems Research, University of Maryland, College Park. His current research interests include networked control systems, wireless communications and networks, signal processing for sensor networks, and stochastic and adaptive estimation and control.

Dr. Dey currently serves on the Editorial Board of the IEEE Transactions on Signal Processing and *Elsevier Systems and Control Letters*. He was also an Associate Editor for the IEEE Transactions on Automatic Control from 2005 to 2007.

Girish N. Nair (S'97–M'99) was born in Petaling Jaya, Malaysia. He received the B.E. degree (with first class honors) in electronics, the B.Sc. degree in mathematics, and the Ph.D. degree in electrical engineering from the University of Melbourne, Melbourne, Australia, in 1994, 1995, and 2000, respectively.

He is currently an Associate Professor in the Department of Electrical and Electronic Engineering, University of Melbourne, and has also held visiting positions at the University of Padova, Padova, Italy and Boston University, Boston, MA. He serves as an Associate Editor for *SIAM Journal on Control and Optimization*, and as an Editorial Board Member of *IET Control Theory and Applications*. His research interests lie in the intersection of communications, information theory, and control.

Dr. Nair received several prizes, including the SIAM Outstanding Paper Prize 2006 and the Best Theory Paper Prize at the UKACC International Conference Control, Cambridge, U.K., 2000.