STABILIZABILITY OF STOCHASTIC LINEAR SYSTEMS WITH FINITE FEEDBACK DATA RATES[*]

GIRISH N. NAIR[†‡] AND ROBIN J. EVANS[†§]

Abstract. Feedback control with limited data rates is an emerging area which incorporates ideas from both control and information theory. A fundamental question it poses is how low the closed loop data rate can be made before a given dynamical system is impossible to stabilize by *any* coding and control law. Analagously to source coding, this defines the smallest error-free data rate sufficient to achieve "reliable" control, and explicit expressions for it have been derived for linear time-invariant systems without disturbances. In this paper, the more general case of finite-dimensional linear systems with process and observation noise is considered, the object being mean square state stability. By inductive arguments employing the *entropy power inequality* of information theory, and a new quantizer error bound, an explicit expression for the infimum stabilizing data rate is derived, under very mild conditions on the initial state and noise probability distributions.

Key words. stochastic control, communication theory, source coding, entropy

AMS subject classifications. 93E15, 94A05, 94A17, 94A29

1. Introduction. Communications and control have traditionally been areas with little common ground. For the most part communications theory is concerned with the reliable transmission of information from one point to another, and is relatively indifferent to its specific purpose or whether it is eventually fed back to the source. Control theory, in contrast, is concerned mainly with using information in a feedback loop to achieve some performance objective, and usually assumes that limitations in the communications links are not significant enough to affect performance drastically.

The reasons usually given for this mutual indifference are firstly, that a communications system is generally used for a broad range of purposes and can rarely be designed to match a particular objective and, secondly, that to explicitly model communication limitations would complicate controller synthesis. However, in recent years emerging applications such as micro-electromechanical systems, mobile telephone power control, and networked industrial control systems have begun to cross the boundary between these disciplines. In these applications, the aim is to control a dynamical system consisting of many separate components connected by a digital communication network. Although the total available capacity in bits per second may be large, each component is effectively allocated only a small portion. This can introduce significant quantization errors and delays, due to the low resolution and finite transmission time of each discrete-valued, digital *symbol*. Quantization resolution can be improved at the expense of delay and vice-versa, but nonetheless there remains an upper bound on the amount of *information*, in some sense, that may be exchanged per unit time. Clearly, by designing coders and decoders that are matched to the dynamical system and controllers, a more economical use of communication

resources ought to be possible. Conversely, closed-loop performance should improve by matching the feedback laws to the specific coding and decoding schemes used.

The first step towards gaining a comprehensive understanding of these issues is to analyze the simplest possible network topology, consisting of one controller and dynamical system connected by a feedback loop with a given data rate in bits per unit time. In view of the limited communication resource, a natural question is: what is the smallest data rate above which there exists a coding and control law that stabilizes the system? This is analogous to Shannon's *source coding* theory, which seeks to determine the smallest data rate above which a given random process can be reliably communicated, i.e. with arbitrarily small error, by some coder and decoder [25, 7]. However, despite this analogy Shannon's theory has generally not been fruitful in real-time control systems since its reliance on arbitrarily long *block* coders entails arbitrarily long delays. While this can be overcome by recursive coders such as *delta* and *differential pulse code* modulators, stationarity or ergodicity are still assumed [17]. Although this may be justified in communications, it does not always suit the unstable dynamics often encountered in control.

In recent years, somewhat more progress on this topic has been made in the control literature. Beginning with the seminal paper [8] and continuing with [29, 2, 6, 10, 3, 23, 11, 16, 18], various schemes have been proposed, and proven to asymptotically or *practically* stabilize linear time-invariant (LTI) systems at sufficiently high data rates. The first rigorous results on minimum data rates were in [29, 2], where it was shown that a discretized scalar plant with parameter a was stabilizable iff the data rate was not less than $\log_2 |a|$ bits per sampling interval. Similar tight bounds were subsequently obtained for noiseless autoregressive moving average [20] and and linear state-space systems [27, 22, 14], using different formulations and techniques. With regard to stochastic plants, separation principles, causal *rate-distortion* theorems [28, 5] and the notion of feedback capacity [24] have been introduced.

This paper focuses on finite-dimensional, stochastic linear plants, under very mild assumptions on the noise and initial state probability distributions. In particular, the objective is to construct a coding and control scheme which achieves mean square state stability *while consuming as little data rate as possible*. The problem is formulated precisely in the next section, and the main result, which specifies the infimum stabilizing data rate, is stated. Somewhat counter-intuitively, it depends only on the unstable dynamics of the plant and not on the noise statistics.

The remainder of the paper essentially constitutes the proof. As the presence of noise makes it difficult to extend the asymptotic quantization approach of [20, 22], a completely different method is developed here. In section 4, the well-known *entropy power inequality* of information theory [7, 9] is used to derive a strict lower bound on the data rate of any stabilizing, causal coding and control scheme, regardless of structure. It is shown that as the feedback data rate approaches this bound from above, the mean square state norms become arbitrarily large. In section 5, a specific, finite-dimensional scheme is then proposed. By applying a new, finite-level quantizer error inequality, it is proven to achieve mean square stability at any data rate exceeding the critical bound.

2. Formulation. First, certain conventions are defined. Vectors are written in bold-face type, matrices in bold-face upper-case, random variables in upper-case and their realizations in corresponding lower-case letters. All random variables are assumed to exist on a common probability space with measure P. The probability density of random vector $\mathbf{X}$ in Euclidean space with respect to (w.r.t.) Lebesgue mea-

sure λ on the space is denoted by $p_{\mathbf{X}}$, the probability density conditioned on the σ-field generated by an event $A = a$ by $p_{\mathbf{X}|a}$, expectation by E, and expectation conditioned on $A = a$ by E_a. The *(differential) entropy* of $\mathbf{X}$ is written $\mathrm{H}\{\mathbf{X}\} \triangleq -\mathrm{E}\{\ln p_{\mathbf{X}}(\mathbf{X})\}$, the *conditional entropy* of $\mathbf{X}$ given $A = a$ as $\mathrm{H}_a\{\mathbf{X}\} \triangleq -\mathrm{E}_a\{\ln p_{\mathbf{X}|A}(\mathbf{X})\}$, and the *average conditional entropy* of $\mathbf{X}$ given $A = a, B = b$, and averaged over B, as $\mathrm{H}_a\{\mathbf{X}|B\} \triangleq -\mathrm{E}_a\{\mathrm{H}_{A,B}\{\mathbf{X}\}\}$. Sequences $\{a_j\}_{j=0}^k$ are denoted $\tilde{a}_k$ (defined as the empty sequence when $k < 0$), and $\|\cdot\|$ represents either the Euclidean norm on a real vector space or the matrix norm induced by it. The $d \times d$ identity matrix is written $\mathbf{I}_d$, the $m \times n$ null matrix $\mathbf{0}_{m \times n}$, real numbers $\mathbb{R}$, positive reals $\mathbb{R}_+$, complex numbers $\mathbb{C}$, integers $\mathbb{Z}$, positive integers $\mathbb{Z}_+$, and non-negative integers $\mathbb{W}$.

Consider the partially-observed, discrete-time, stochastic linear system

$$(2.1) \qquad \mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{v}_k, \quad \mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{w}_k, \ \forall k \in \mathbb{W},$$

with state $\mathbf{x}_k$ and process noise $\mathbf{v}_k \in \mathbb{R}^n$, control signal $\mathbf{u}_k \in \mathbb{R}^m$, and measurement $\mathbf{y}_k$ and measurement noise $\mathbf{w}_k \in \mathbb{R}^p$. It is assumed that

A1 $(\mathbf{A}, \mathbf{B})$ is reachable and $(\mathbf{C}, \mathbf{A})$, observable;

A2 $\mathbf{x}_0, \mathbf{v}_k$, and $\mathbf{w}_k$ are realizations of random variables $\mathbf{X}_0, \mathbf{V}_k$, and $\mathbf{W}_k$ respectively, where $\mathbf{X}_0, \mathbf{V}_k, \mathbf{W}_j$ are mutually independent, $\forall k, j \in \mathbb{W}$;

A3 $\exists \varepsilon > 0$ such that $\mathbf{X}_0, \mathbf{V}_k, \mathbf{W}_k$ have uniformly bounded $(2+\varepsilon)$th absolute moments over $k \in \mathbb{W}$;

A4 the probability distribution of each random variable $\mathbf{V}_k$ is absolutely continuous with respect to Lebesgue measure λ on $\mathbb{R}^n$;

A5 $\inf_{k \in \mathbb{W}} \mathrm{H}\{\mathbf{V}_k^{\mathrm{u}}\} > -\infty$, where $\mathbf{V}_k^{\mathrm{u}} \in \mathbb{R}^{f \times n}$ is the process noise seen by the $f \geq 1$ unstable eigenvectors of $\mathbf{A}$; i.e. the process noise injects a minimum amount of uncertainty into the unstable dynamics.

Suppose that the sensor producing the measurements is connected to the controller via a digital channel, onto which one symbol s_k from a finite alphabet $\mathcal{S}_k$, of possibly time-varying size $\mu_k \geq 1$, is transmitted during the $(k+1)$th sampling interval. It is assumed that each transmitted symbol is received without error, as in Shannon source coding, after a delay of d intervals. The *(asymptotic average) data rate* of the channel may then be defined as

$$(2.2) \qquad R \triangleq \liminf_{t \to \infty} \frac{1}{t} \sum_{k=0}^{t-1} \log_2 \mu_k.$$

This is a more general definition than in [21, 22], in which the alphabet size μ_k is constant. In particular, it permits the alphabet $\mathcal{S}_k$ to vary periodically. For technical reasons, it is also assumed that $\mu_k/k \to 0$ as $k \to \infty$.

As the symbols in the channel are discrete-valued but the plant measurements are continuous-valued, analog-to-digital conversion, or coding, is required. In practice constraints such as complexity and finite memory may be important but, in the spirit of source coding, such limitations will be largely ignored here to concentrate on the communication aspect of the problem. Each transmitted symbol may thus depend on all past and present measurements and past symbols,

$$(2.3) \qquad s_k = \gamma_k(\tilde{\mathbf{y}}_k, \tilde{s}_{k-1}), \quad \forall k \in \mathbb{W},$$

where $\gamma_k : \mathbb{R}^{p \times (k+1)} \times \tilde{\mathcal{S}}_{k-1} \to \mathcal{S}_k$ is the coder mapping at time k. Note in particular that s_k does not necessarily correspond to a quantized version of the latest measurement alone. At time k the controller has the symbols $s_0, \ldots, s_{k-d}$ available to it and

can then generate a control signal of the general form

$$(2.4) \qquad \mathbf{u}_k = \delta_k(\tilde{s}_{k-d}), \quad \forall k \in \mathbb{W},$$

where $\delta_k : \tilde{\mathcal{S}}_{k-d} \to \mathbb{R}^m$ is the controller mapping at time k. As $\tilde{s}_k$ is the empty sequence $\{\}$ when $k < 0$, the first d control signals $\mathbf{u}_0, \ldots, \mathbf{u}_{d-1}$ are just preset inputs. Similarly, in the coder equation (2.3) at time $k = 0$, s_0 is a function only of $\mathbf{y}_0$.

Now, define the *coder-controller* as the triple of alphabet, coder and controller mapping sequences $(\tilde{\mathcal{S}}_\infty, \tilde{\gamma}_\infty, \tilde{\delta}_\infty)$. The objective here is to construct a coder-controller which stabilizes the plant in the mean square sense

$$(2.5) \qquad \sup_{k \in \mathbb{W}} \mathrm{E}\|\mathbf{X}_k\|^2 < \infty,$$

while using as small a data rate as possible. The main result of this paper is now stated:

THEOREM 2.1. *Given assumptions A1–A5, any coder-controller which stabilizes the plant (2.1) in the mean square sense (2.5) must have a data rate R (2.2) strictly satisfying*

$$(2.6) \qquad R > \sum_{|\eta_j| \geq 1} \log_2 |\eta_j| =: H,$$

where $\eta_1, \ldots, \eta_n$ are the open-loop eigenvalues. As R approaches this bound from above, the supremum mean square state norm (2.5) approaches ∞.

This inequality is also tight. I.e. for any number $R' > H$, a mean-square-stabilizing coder-controller at data rate $R \leq R'$ can be constructed, which furthermore is finite-dimensional with periodic alphabet.

This result assumes nothing about the coding and control laws but causality, and imposes only mild requirements on the noise distributions. It thus draws a fundamental line of demarcation between what is and is not achievable with stochastic linear systems when communication rates are limited. In this sense, H plays a role similar to source entropy in errorless Shannon source coding, and can be taken as a measure of the rate at which information is generated by an unstable, stochastic linear plant. Hence (2.6) states that to achieve stability, the channel must transport data as fast as it is produced.

A more physical insight can be gained by rewriting the inequality above as $2^R > \prod_{|\eta_j| \geq 1} |\eta_j|$. The right-hand side (RHS) is simply the factor by which a volume in the unstable subspace increases at each time step due to the plant dynamics, while the left-hand side (LHS) is the asymptotic average number of disjoint regions into which the coder can partition it. In other words, the system is stabilizable if and only if the dynamical increase in "uncertainty volume" due to unstable dynamics is outweighed by the partitioning induced by the coder.

Note also that the data rate bound above is completely independent of the noise distributions and link delay, a consequence of the weak notion of stability used here. Increasing the noise variances (or delay) would obviously increase the the mean square state norms, but as long as (2.6) is satisfied it remains possible to keep the state uniformly bounded in a mean square sense. The reason for this is that the noise increases state uncertainty volumes in an additive rather than multiplicative fashion but is averaged out exponentially, in effect, by the coder. However, as the data rate approaches the critical limit, this exponential averaging becomes increasingly weak,

leading to an unbounded increase in uncertainty volumes and hence mean square states.

Finally, it is remarked that in the problem formulation above there is no explicit communication constraint between the controller and actuator. This is reasonable if they are co-located, but even otherwise the formulation above is applicable, since from a plant output-to-input perspective the location of the controller is purely nominal. The symbols that would be transmitted by it over an additional digital link to the actuator would have to be converted once again into inputs, making intermediate calculations redundant. In other words the "bottle-neck" link determines the effective data rate, as expected, and Theorem 2.1 still applies. This is stated below more precisely:

PROPOSITION 2.2. *Suppose that two cascaded digital links connect the sensor to the actuator, with associated mappings*

$$(2.7) \qquad s_k^1 = \gamma_k^1(\tilde{\mathbf{y}}_k, \tilde{s}_{k-1}^1) \in \mathcal{S}_k^1, \qquad \textit{(link-1 coder)}$$

$$(2.8) \qquad s_k^2 = \gamma_k^2(\tilde{s}_{k-d_1}^1) \in \mathcal{S}_k^2, \qquad \textit{(link-2 coder)}$$

$$(2.9) \qquad \mathbf{u}_k = \delta_k^2(\tilde{s}_{k-d_2}^2) \in \mathbb{R}^m. \qquad \textit{(actuator)}$$

Let R_1 be the data rate (2.2) of the first link and R_2 that of the second. Then this coding and control scheme can be expressed as a single-link coder-controller of the form (2.3)–(2.4), with delay $d = d_1 + d_2$ and data rate $R = \min\{R_1, R_2\}$.

Conversely, any single-link coder-controller with periodic alphabet, data rate R, and delay d can always be expressed as a two-link coding and control scheme of the form above, with periodic alphabet sizes, link data rates both equal to R, and arbitrary delays $d_1, d_2 \in \mathbb{W}$ such that $d_1 + d_2 = d$.

Proof: See Appendix A.

The remainder of this paper is devoted to proving Theorem 2.1, in three stages. In the next section, the system dynamics are transformed into a simpler form. In section 4, the strict necessity of (2.6) is established via an inductive argument using the *entropy power inequality* of information theory [9]. Finally, its sufficiency is demonstrated in section 5 by constructing coder-controller, and using a new quantizer error result to recursively bound the mean square state norms.

3. Real Jordan Form. Before proceeding, it is convenient to transform the system so as to decouple its dynamical modes. The obvious approach of putting the matrix $\mathbf{A}$ into *Jordan canonical form* generally requires a transformation matrix with complex elements. As this would complicate the analysis somewhat, the *real Jordan canonical form* [15] is used here.

Let $\lambda_1, \ldots, \lambda_b$ be the *distinct* eigenvalues of $\mathbf{A} \in \mathbb{R}^{n \times n}$, ordered by non-increasing magnitude with conjugates excluded, and let the algebraic multiplicity of each λ_i be m_i. The real Jordan canonical form $\mathbf{J}$ then has the block diagonal structure

$$(3.1) \qquad \mathbf{J} \equiv \operatorname{diag}(\mathbf{J}_1, \ldots, \mathbf{J}_b) \in \mathbb{R}^{n \times n},$$

where the block $\mathbf{J}_i \in \mathbb{R}^{n_i \times n_i}$ with

$$(3.2) \qquad n_i \triangleq \begin{cases} m_i & \text{if } \lambda_i \in \mathbb{R} \\ 2m_i & \text{otherwise} \end{cases}.$$

More detail regarding the structure of each block can be found in e.g. [15](pp. 150–3) or [26]. For the purposes of this paper, the most important fact is that each block $\mathbf{J}_i$

is *similar* to the block-diagonal matrix of all *standard* Jordan blocks corresponding to λ_i, λ_i^*. Hence $\mathbf{J}_i$ has either exactly one distinct eigenvalue λ_i or a pair of complex conjugate eigenvalues λ_i, λ_i^*, each with multiplicity m_i.

The real Jordan canonical form is related to the matrix $\mathbf{A}$ via a real similarity matrix $\mathbf{T} \in \mathbb{R}^{n \times n}$ such that $\mathbf{T}^{-1}\mathbf{J}\mathbf{T} = \mathbf{A}$. Defining the transformed state

$$(3.3) \qquad\qquad\qquad \acute{\mathbf{x}}_k \overset{\Delta}{=} \mathbf{T}\mathbf{x}_k, \ \forall k \in \mathbb{W},$$

the system equations (2.1) can then be written

$$(3.4) \ \ \acute{\mathbf{x}}_{k+1} = \mathbf{J}\acute{\mathbf{x}}_k + \mathbf{T}\mathbf{B}\mathbf{u}_k + \mathbf{T}\mathbf{v}_k \in \mathbb{R}^n, \quad \mathbf{y}_k \overset{\Delta}{=} \mathbf{C}\mathbf{T}^{-1}\acute{\mathbf{x}}_k + \mathbf{w}_k \in \mathbb{R}^p, \ \forall k \in \mathbb{W}.$$

By partitioning the transformed state vector into the vectors $\acute{\mathbf{x}}_k^{(1)}, \ldots, \acute{\mathbf{x}}_k^{(b)}$ corresponding to each subsystem, the dynamical equation above can be rewritten more explicitly as

$$(3.5) \qquad \acute{\mathbf{x}}_{k+1}^{(i)} = \mathbf{J}_i\acute{\mathbf{x}}_k^{(i)} + (\mathbf{T}\mathbf{B}\mathbf{u}_k)^{(i)} + (\mathbf{T}\mathbf{v}_k)^{(i)} \in \mathbb{R}^{d_i}, \ \ \forall k \in \mathbb{W}, i \in [1, \ldots, b],$$

where $(\cdot)^{(i)}$ denotes that portion of the vector argument that feeds into the ith subsystem.

The original system has thus been decomposed into b real subsystems, with dynamics characterized by either a single eigenvalue or a pair of complex conjugate eigenvalues, possibly repeated. As $\mathbf{T}$ is invertible, it follows that the problems of stabilizing (3.4) and (2.1) are equivalent.

4. Proof of Necessity. The first step towards proving Theorem 2.1 is to establish the necessity of (2.6) for mean square stability. In order to do so, a recursive lower bound for $\mathrm{E}\|\acute{\mathbf{X}}_k\|^2$ shall be sought which is independent of the coder-controller, and easier to analyze in terms of dynamics and data rate. More precisely, this bound will be sought for the state vector corresponding to subsystems with eigenvalue $|\lambda_i| \geq 1$.

If a strict inequality in (2.6) was not desired, a lower bound could quickly be obtained by observing that since the noise terms are independent, the mean square state norm cannot increase if they are all suppressed. In other words, the mean square state norm is bounded below by that of the plant with a random initial state but no noise. This is precisely the situation explored in [22] and, by a slight modification of the quantization argument used there, the nonstrict version of (2.6) is easily seen to be necessary for mean square stability.

However, this reduction to a noiseless system does not reveal that stability is in fact impossible at a data rate equal to the critical bound H. More importantly, it states nothing about behaviour *near* H, in particular the fact that as the data rate approaches it the supremum mean square state norm (2.5) becomes arbitrarily large, drastically degrading performance, regardless of the coder-controllers used. This is made apparent by the entropy-based analysis below.

Denote the index set of unstable subsystems and their total dimension respectively by

$$(4.1) \qquad\qquad \mathcal{U} \overset{\Delta}{=} \{i : |\lambda_i| \geq 1\}, \quad f \overset{\Delta}{=} \sum_{i \in \mathcal{U}} n_i$$

and stack the unstable subsystem states $\acute{\mathbf{x}}_k^{(i)}$, $i \in \mathcal{U}$, to construct

$$\mathbf{x}_k^{\mathrm{u}} := \left[\acute{\mathbf{x}}_k^{(1)\mathrm{T}} \cdots \acute{\mathbf{x}}_k^{(|\mathcal{U}|)\mathrm{T}} \right]^{\mathrm{T}} \equiv \mathbf{R}\acute{\mathbf{x}}_k \in \mathbb{R}^f, \ \forall k \in \mathbb{W},$$

$$(4.2) \qquad\qquad \text{where } \mathbf{R} \overset{\Delta}{=} \begin{bmatrix} \mathbf{I}_f & \mathbf{0}_{f \times (n-f)} \end{bmatrix} \in \mathbb{R}^{f \times n}.$$

Now, suppose that the coder-controller $(\tilde{\mathcal{S}}_\infty, \tilde{\gamma}_\infty, \tilde{\delta}_\infty)$ stabilizes (2.1), and hence (3.4), in mean square state norm. As $\|\mathbf{x}_k^u\| \le \|\hat{\mathbf{x}}_k\|$, it follows that $\{\mathbf{X}_k^u\}_{k\in\mathbb{W}}$ is also bounded in mean square norm. Following the usual definition of *entropy power* (see, e.g. [9]), let the *conditional entropy power* of a random variable $\mathbf{X} \in \mathbb{R}^f$ given an event $A = a$ be

$$(4.3) \qquad \mathrm{N}_a\{\mathbf{X}\} \triangleq (2\pi\mathrm{e})^{-1}\mathrm{e}^{2\mathrm{H}_a\{\mathbf{X}\}/f}.$$

In connexion with the uncertainty volume interpretation of Theorem 2.1, $(\mathrm{N}_a\mathbf{X})^{f/2}$ can be regarded as the volume of the effective support set of $p_{\mathbf{X}|a}$. Furthermore,

$$(4.4) \qquad \mathrm{N}_a\{\mathbf{X}\} \le \mathrm{e}^{1/f-1}\mathrm{E}_a\|\mathbf{X}\|^2,$$

with equality if and only if $\mathbf{X}$ is symmetric Gaussian with zero mean when conditioned on $A = a$ (see Appendix B). This is essentially a statement of the well-known entropy-maximizing property of Gaussian distributions.

By analogy with the notation for average conditional entropy, denote the *average conditional entropy power* of $\mathbf{X}$ given $A = a$, averaged over A, by

$$(4.5) \qquad \mathrm{N}\{\mathbf{X}|A\} \triangleq \mathrm{E}\left\{\mathrm{N}_A\{\mathbf{X}\}\right\}.$$

Setting $\mathbf{X} = \mathbf{X}_k^u$ and $A = \tilde{S}_{k-d-1}$ the random variable associated with the symbol sequence $\tilde{s}_{k-d-1}$,

$$n_k := \mathrm{N}\{\mathbf{X}_k^u|\tilde{S}_{k-d-1}\} = \mathrm{E}\left\{\mathrm{N}_{\tilde{S}_{k-d-1}}\{\mathbf{X}_k^u\}\right\},$$
$$(4.6) \qquad \le \mathrm{e}^{1/f-1}\mathrm{E}\left\{\mathrm{E}_{\tilde{S}_{k-d-1}}\|\mathbf{X}_k^u\|^2\right\} = \mathrm{e}^{1/f-1}\mathrm{E}\|\mathbf{X}_k^u\|^2, \quad \forall k \in \mathbb{W},$$

so that $\{n_k\}_{k\in\mathbb{W}}$ must also be bounded. Note here that n_k can be interpreted as the average unstable subspace uncertainty volume given the symbol sequence $\tilde{s}_{k-d-1}$.

Another important property of conditional entropy power is its super-additivity for summed independent random variables, i.e.

$$(4.7) \qquad \mathrm{N}_a\{\mathbf{X} + \mathbf{Y}\} \ge \mathrm{N}_a\{\mathbf{X}\} + \mathrm{N}_a\{\mathbf{Y}\},$$

where $\mathbf{X}, \mathbf{Y} \in \mathbb{R}^f$ are mutually independent when conditioned on an event $A = a$ (see, e.g. [7, 9]). By means of a recursive argument that employs this so-called *entropy power inequality*, it shall be shown that any stabilizing data rate must satisfy (2.6) strictly. First, observe from (4.2) that

$$(4.8) \qquad \mathbf{R}\mathbf{J} = \mathbf{J}^u\mathbf{R}, \quad \text{where } \mathbf{J}^u \triangleq \mathrm{diag}(\mathbf{J}_1, \ldots, \mathbf{J}_{|\mathcal{U}|}) \in \mathbb{R}^{f\times f}.$$

Left-multiplying (3.4) by $\mathbf{R}$ and using (2.4), the dynamical equation for $\mathbf{x}_k^u$ is then

$$(4.9) \qquad \mathbf{x}_{k+1}^u = \mathbf{J}^u\mathbf{x}_k^u + \mathbf{R}\mathbf{T}\mathbf{v}_k + \mathbf{R}\mathbf{T}\mathbf{B}\delta_k(\tilde{s}_{k-d}).$$
$$\Rightarrow \mathrm{N}_{\tilde{S}_{k-d}}\{\mathbf{X}_{k+1}^u\} = \mathrm{N}_{\tilde{S}_{k-d}}\{\mathbf{J}^u\mathbf{X}_k^u + \mathbf{R}\mathbf{T}\mathbf{V}_k + \mathbf{R}\mathbf{T}\mathbf{B}\delta_k(\tilde{S}_{k-d})\},$$
$$(4.10) \qquad = \mathrm{N}_{\tilde{S}_{k-d}}\{\mathbf{J}^u\mathbf{X}_k^u + \mathbf{R}\mathbf{T}\mathbf{V}_k\},$$
$$(4.11) \qquad \ge \mathrm{N}_{\tilde{S}_{k-d}}\{\mathbf{J}^u\mathbf{X}_k^u\} + \mathrm{N}_{\tilde{S}_{k-d}}\{\mathbf{R}\mathbf{T}\mathbf{V}_k\} = \mathrm{N}_{\tilde{S}_{k-d}}\{\mathbf{J}^u\mathbf{X}_k^u\} + \mathrm{N}\{\mathbf{R}\mathbf{T}\mathbf{V}_k\},$$
$$(4.12) \qquad = |\det\mathbf{J}^u|^{2/f}\mathrm{N}_{\tilde{S}_{k-d}}\{\mathbf{X}_k^u\} + \mathrm{N}\{\mathbf{R}\mathbf{T}\mathbf{V}_k\}, \quad \forall k \in \mathbb{W}.$$

The equality in (4.10) is due to the property that $H_a\{\mathbf{X} + g(A)\} = H_a\mathbf{X}$ for any function g (translation invariance). The inequality (4.11) uses the mutual independence of $\mathbf{V}_k$ and $\mathbf{X}_k, \tilde{S}_{k-d}$, the latter both being determined by $\mathbf{X}_0, \mathbf{V}_j, \mathbf{W}_j, j \leq k-1$, to apply the entropy power inequality (4.7) and finally (4.12) expresses the effect of an invertible linear transformation on differential entropy; see corollary 9.6.4 in [7].

Now, denote the conditional entropy power of $\mathbf{X}$ given $A = a, S = s$ and averaged only over S by $N_a\{\mathbf{X}|S\} \triangleq E_a\{N_{A,S}\{\mathbf{X}\}\}$. The next step utilizes the result below.

LEMMA 4.1. *Let* $\mathbf{X} \in \mathbb{R}^f$ *and* $S \in \mathcal{S}$ *a finite alphabet be random variables conditioned on an event* $A = a$. *Then*

$$(4.13) \qquad N_a\{\mathbf{X}|S\} \geq |\mathcal{S}|^{-2/f} N_a\{\mathbf{X}\}.$$

Proof: See Appendix C.

As $(N_a\{\mathbf{X}|S\})^{f/2}$ can be viewed as the average uncertainty volume of $\mathbf{X}$ given S conditioned on $A = a$, this inequality states that knowledge of a correlated random variable $S \in \mathcal{S}$ with $|\mathcal{S}|$ distinct values reduces the average uncertainty volume of $\mathbf{X}$ by at most a factor of $|\mathcal{S}|$. In a sense, this is an extension of deterministic volume partitioning to a stochastic setting.

Setting $\mathbf{X} = \mathbf{X}_k^{\mathrm{u}}$, $A = \tilde{S}_{k-d-1}$, $S = S_{k-d}$, and averaging (4.12),

$$\begin{aligned}
n_{k+1} &\equiv E\left\{N_{\tilde{S}_{k-d}}\{\mathbf{X}_{k+1}^{\mathrm{u}}\}\right\}, \\
&\geq |\det\mathbf{J}^{\mathrm{u}}|^{2/f} E\left\{E_{\tilde{S}_{k-d-1}}\left\{N_{\tilde{S}_{k-d}}\{\mathbf{X}_k^{\mathrm{u}}\}\right\}\right\} + E\{N\{\mathbf{RTV}_k\}\}, \\
&= |\det\mathbf{J}^{\mathrm{u}}|^{2/f} E\left\{N_{\tilde{S}_{k-d-1}}\{\mathbf{X}_k^{\mathrm{u}}|S_{k-d}\}\right\} + N\{\mathbf{RTV}_k\}, \\
&\geq \left|\frac{\det\mathbf{J}^{\mathrm{u}}}{|\mathcal{S}_{k-d}|}\right|^{2/f} E\left\{N_{\tilde{S}_{k-d-1}}\{\mathbf{X}_k^{\mathrm{u}}\}\right\} + N\{\mathbf{RTV}_k\},
\end{aligned}$$

$$(4.14) \qquad \equiv \left|\frac{\det\mathbf{J}^{\mathrm{u}}}{\mu_{k-d}}\right|^{2/f} n_k + N\{\mathbf{RTV}_k\}, \quad \forall k \in \mathbb{W}.$$

By assumption A5 in §2, $\exists \theta \in \mathbb{R}$ s.t. $H\{\mathbf{RTV}_j\} \geq \theta, \forall j \in \mathbb{W}$. Hence

$$(4.15) \qquad N\{\mathbf{RTV}_j\} \geq (2\pi)^{-1} e^{2\theta/f - 1} =: \beta > 0, \quad \forall j \in \mathbb{W},$$

by definition (4.3). Substituting this into (4.14) and for convenience setting $\mu_i = 1$ when $i < 0$,

$$\begin{aligned}
n_{k+1} &\geq \left|\frac{\det\mathbf{J}^{\mathrm{u}}}{\mu_{k-d}}\right|^{2/f} n_k + \beta \geq \beta \sum_{j=0}^{k} \prod_{i=j+1}^{k} \left|\frac{\det\mathbf{J}^{\mathrm{u}}}{\mu_{i-d}}\right|^{2/f}, \\
&\equiv \beta \sum_{j=0}^{k} \frac{h_k}{h_j}, \quad \text{where } h_k := \prod_{i=0}^{k} \left|\frac{\det\mathbf{J}^{\mathrm{u}}}{\mu_{i-d}}\right|^{2/f}.
\end{aligned}$$

$$(4.16) \quad \Rightarrow \quad \infty > \alpha \triangleq \frac{1}{\beta} \sup_{k \in \mathbb{Z}_+} n_k \geq \sum_{j=0}^{k} \frac{h_k}{h_j} \quad \Leftrightarrow \quad \frac{\alpha - 1}{h_k} \geq \sum_{j=0}^{k-1} \frac{1}{h_j}, \quad \forall k \in \mathbb{Z}_+.$$

As $h_k > 0$, $\alpha > 1$. By upward induction on k it can be verified that

$$h_k \leq \alpha(1 - 1/\alpha)^k h_0, \quad \forall k \in \mathbb{Z}_+.$$

$$\Rightarrow\ 0 > \log_2(1 - 1/\alpha) \ge \frac{1}{k} \log_2\left(\frac{h_k}{\alpha}\right),$$

$$= \frac{2(k+1)}{fk} \log_2|\det\mathbf{J}^{\mathrm{u}}| - \frac{2}{fk} \sum_{i=0}^{k} \log_2 \mu_{i-d} - \frac{\log_2 \alpha}{k}, \quad \forall k \in \mathbb{Z}_+,$$

$$\ge \frac{2}{f} \log_2|\det\mathbf{J}^{\mathrm{u}}| - \liminf_{k\to\infty}\left[\frac{2}{fk}\sum_{i=0}^{k}\log_2\mu_{i-d} - \frac{2\log_2|\det\mathbf{J}^{\mathrm{u}}|}{fk} + \frac{\log_2\alpha}{k}\right],$$

$$(4.17) \qquad \equiv (2/f)\left(\log_2|\det\mathbf{J}^{\mathrm{u}}| - R\right) \equiv (2/f)(H - R).$$

This proves the strict necessity of (2.6). Furthermore, after rearranging the inequalities above and using (4.6),

$$(4.18) \qquad \sup_{k\in\mathbb{Z}_+} \mathrm{E}\|\mathbf{X}_k^{\mathrm{u}}\|^2 \ge \mathrm{e}^{1-1/f} \sup_{k\in\mathbb{Z}_+} n_k = \mathrm{e}^{1-1/f}\beta\alpha \ge \frac{\mathrm{e}^{1-1/f}\beta}{1 - 2^{-2(R-H)/f}}.$$

As β (4.15) depends only on process noise statistics, this is a universal lower bound on the supremum mean square state of *all* coder-controllers with data rate R. Hence as $R \searrow H$, the supremum mean square state becomes arbitrarily large. $\qquad\square$

Note that the argument above can be adapted to deal with stability in the sense of uniform boundedness with bounded disturbances. The idea is to replace the average conditional entropy power n_k with the maximum state uncertainty volume given past symbols and then use the deterministic analogue of the entropy power inequality, the *Brunn-Minkowski inequality* (see, e.g. [7]). This states that, given two λ-measurable regions $X, Y \subset \mathbb{R}^f$, the volume of the set sum $X + Y \stackrel{\Delta}{=} \{x + y | x \in X, y \in Y\}$, satisfies $\lambda(X+Y)^{1/f} \ge \lambda(X)^{1/f} + \lambda(Y)^{1/f}$. It then follows that (2.6) is also strictly necessary for uniformly bounded stabilizability.

5. Achievability of Data Rate Bound. The final step in proving Theorem 2.1 is to establish that (2.6) is attainable, i.e. that the system (2.1) can in principle be stabilized at any data rate arbitarily close to but greater than the critical bound H. The general, entropy-based argument of the preceding section does not offer many clues as to how to prove this, so in this section a completely different approach is taken. Based on a semi-heuristic line of reasoning, a finite-dimensional coder-controller with periodically-varying coding alphabet is constructed in §5.2. By means of a new quantizer bound, it is then demonstrated in §5.3 that it achieves means square stability at data rates arbitrarily close to H.

The chief complications in the design and analysis of this scheme arise from the unbounded support of the noise terms. With uniformly bounded noise, any coding and control law which achieves asymptotic contraction *without* disturbances, in the sense that $\exists \gamma \in (0,1)$ such that for all sufficiently large k, $\|\mathbf{x}_k\| < \gamma r, \forall r > 0, \|\mathbf{x}_0\| \le r$, can easily be modified to achieve uniformly bounded stability. The idea is to recast such a law as an equivalent open-loop scheme which generates symbols according to the initial state alone. Assuming no noise, $\exists$ a sufficiently large $\tau \in \mathbb{Z}_+$ such that any state with norm $\le r$ will after τ steps have norm $\le \gamma r$. The effect of bounded disturbances then boosts the radius of this worst-case region by an additive constant $c \cdot \tau$, so the open-loop scheme can then be reapplied with $r^{\mathrm{new}} = \gamma r + c\tau$. As the recursion $r \mapsto \gamma r + c\tau$ is stable, a uniform bound on the state at times $0, \tau, 2\tau, \ldots$ is guaranteed, and trivially leads to a bound over all integer times.

The situation is quite different when dealing with unbounded stochastic disturbances, because of the impossibility of a coder-controller which uniformly con-

tracts mean square norms in an analogous sense. Briefly, the reason for this is that even though a distribution may have finite second absolute moment, the tail integral $\int_{\|\mathbf{x}\| \geq t} \|\mathbf{x}\|^2 \mathrm{d}P_{\mathbf{X}}(\mathbf{x})$ can approach 0 arbitrarily slowly with large t. In §5.3 this difficulty is overcome by dealing with a functional M_ε (5.9) instead of the mean square state norm. Before proceeding to the construction and analysis of the stabilizing coder-controller, several structural issues are first discussed below.

5.1. Structural Issues. In order to make the analysis tractable, a certain amount of structure will need to be imposed on the general coder-controller equations (2.3)–(2.4). It is known (see, e.g. [28]) that for a linear, Gauss-Markov system under a mean quadratic cost, there exist optimal coding and control schemes with the following form:

1. Prior to coding, a Kalman filter is applied to recursively calculate the linear minimum variance prediction $\bar{\mathbf{x}}_{k+d|k}$ of $\acute{\mathbf{x}}_{k+d}$ given the measurement and control sequences $\tilde{\mathbf{y}}_k, \tilde{\mathbf{u}}_{k+d-1}$. Note that the control signals are not observed directly by the coder, but inferred from knowledge of the symbol sequence $\tilde{s}_{k-1}$ and the controller mappings.

2. Based on the past symbol sequence $\tilde{s}_{k-1}$, the latest prediction $\bar{\mathbf{x}}_{k+d}$ is recursively (and possibly non-uniformly) quantized to yield a coded estimate

$$\hat{\mathbf{x}}_{k+d} \equiv Q_k(\bar{\mathbf{x}}_{k+d|k}, \tilde{s}_{k-1}) \equiv \varpi_k(s_k)$$

with μ_k possible values. The index s_k of the selected quantizer point $\varpi_k(s_k)$ is transmitted.

3. Upon receiving s_k at time $k+d$, the controller uses it and the previous symbols to regenerate $\hat{\mathbf{x}}_{k+d}$ and applies a certainty-equivalent linear control law $\mathbf{u}_{k+d} \equiv \mathbf{L}\hat{\mathbf{x}}_{k+d}$. Although no Gaussian assumptions are made in this paper, it is convenient to use a modified version of this tri-stage structure as a basis to construct a stabilizing scheme.

Considering the first stage, recall that the linear minimum variance predictions satisfy the separation principle

$$\mathrm{E}\|\acute{\mathbf{X}}_k\|^2 = \mathrm{E}\|\acute{\mathbf{X}}_k - \bar{\mathbf{X}}_{k|k-d}\|^2 + \mathrm{E}\|\bar{\mathbf{X}}_{k|k-d}\|^2, \quad \forall k \in \mathbb{W},$$

even with non-Gaussian noise. The first term on the RHS is uniformly bounded, by the observability and $(2+\varepsilon)$th moment assumptions A1 and A3 in section 2, and independent of the control law (see e.g. [1]). Hence the mean square stability of the partially observed system (3.4) is equivalent to that of the fully observed filter process. Furthermore, this process satisfies a recursive equation of the same form as (3.4), i.e.

$$(5.1) \qquad \bar{\mathbf{x}}_{k+1+d|k+1} = \mathbf{J}\bar{\mathbf{x}}_{k+d|k} + \mathbf{TB}\mathbf{u}_k + \mathbf{z}_{k+1}, \quad \forall k \in \mathbb{W},$$

where, by assumption A3, the $(2+\varepsilon)$th absolute moments of the *innovation* $\mathbf{Z}_{k+1}$ can be shown to be uniformly bounded over $k \in \mathbb{W}$.

The second stage above is not quite so straightforward, since the optimal quantizer $Q_k(\cdot, \cdot)$ is generally time-varying and stores all past symbols. As the objective here is not optimality but stability, it is natural to investigate if simpler quantizer structures will suffice. One possibility is a static, memoryless coder,

$$(5.2) \qquad \hat{\mathbf{x}}_{k+d} \equiv Q(\bar{\mathbf{x}}_{k+d|k}) \equiv \varpi(s_k) \in \mathbb{R}^n,$$

where Q is a fixed quantizer with points $\varpi(0), \varpi(1), \ldots, \varpi(\mu-1)$. Another option is a *finite-state, predictive quantizer* (see, e.g. [12]), in which the latest coded state

estimate is stored and the *prediction error* is recursively coded according to a finite-valued internal variable $\iota_k \in \mathcal{I}$,

(5.3) $\quad Q\left(\bar{\mathbf{x}}_{k+d|k} - (\mathbf{J} + \mathbf{TBL})\hat{\mathbf{x}}_{k-1+d|k-1}, \iota_k\right) \equiv \varpi(s_k),$

(5.4) $\qquad\qquad\qquad\qquad\qquad \hat{\mathbf{x}}_{k+d} \equiv (\mathbf{J} + \mathbf{TBL})\hat{\mathbf{x}}_{k-1+d} + \varpi(s_k),$

$$\iota_{k+1} \equiv g(\iota_k, s_k),$$

where $\mathbf{L}$ is a certainty-equivalent control gain such that $\mathbf{J} + \mathbf{TBL}$ is stable. The symbol s_k, corresponding to the index of the selected quantizer point, is then used to update the finite state, $(\iota_k, s_k) \mapsto \iota_{k+1}$. Examples are *differential pulse code* and *delta modulation* in speech processing.

For noise distributions with compact support, it can be shown that either type of coder can achieve stability. It may seem as if this should also hold in the case of infinite support, since if stability has been achieved then the states and prediction errors remain with high probability in some bounded region, which could then be quantized without memory. However, this somewhat circular argument fails drastically if the plant is strictly unstable and either the initial state or a process noise term has infinite support in all directions:

PROPOSITION 5.1. *Suppose that the plant (2.1) has at least one open-loop eigenvalue with magnitude strictly greater than 1, and that, for any non-zero* $\mathbf{h} \in \mathbb{R}^n$, *either*

(5.5) $\quad \mathrm{P}\{\mathbf{h}^{\mathrm{T}}\mathbf{X}_0 > \theta\} > 0, \ \forall \theta \in \mathbb{R}, \ or \ \exists t \in \mathbb{W} \ s.t. \ \mathrm{P}\{\mathbf{h}^{\mathrm{T}}\mathbf{V}_t > \theta\} > 0, \ \forall \theta \in \mathbb{R}.$

Then for any static memoryless coder (5.2) or finite-state predictive quantizer (5.3)–(5.4), the r*th absolute state moments are unbounded with time,* $\forall r > 0$, *regardless of the number of quantization points.*

Proof: See Appendix D.

This distinguishes the stochastic, communication-limited stabilization problem from the deterministic, bounded disturbance version, for which either memoryless or finite-state quantization suffice. The reason for the difference is basically that the finite range of the quantizer causes controller saturation. If the initial state or process noise has infinite support, there is consequently a finite chance that at some time k, the propagated state $\mathbf{A}\mathbf{x}_k$ is beyond reach of the control signal. The unstable plant dynamics then amplify this short-fall, causing the same phenomenon to occur with increasing probability at subsequent times, and inevitably leading to instability.

An obvious solution is to use an adaptive quantizer with possibly unbounded range, thereby allowing the control signal to "catch up" with the state. One simple approach is to use a predictive scheme with a scaling factor $l_k > 0$ which is recursively adjusted according to the symbols transmitted:

$$Q\left(\frac{\bar{\mathbf{x}}_{k+d|k} - (\mathbf{J} + \mathbf{TBL})\hat{\mathbf{x}}_{k-1+d|k-1}}{l_k}\right) \equiv \varpi(s_k),$$

$$\hat{\mathbf{x}}_{k+d} \equiv (\mathbf{J} + \mathbf{TBL})\hat{\mathbf{x}}_{k-1+d} + l_k\varpi(s_k),$$

$$l_{k+1} \equiv g(l_k, s_k) \in \mathbb{R}_+.$$

This approach, similar to [6, 19], is adopted in §5.2.

Another characteristic of the coder-controller constructed here is that its symbol alphabet varies periodically with time, a point of departure from the time-invariant, constant data rate schemes in [3] and elsewhere. If the symbol alphabet was fixed then

the data rate in bits/interval could only take the discrete values $\log_2 \mu$, $\mu = 1, 2, \ldots$, making it impossible in general to attain data rates arbitrarily close to H. In [3] this is not an issue, since the underlying plant is in continuous-time and the corresponding data rate bound, in bits/*second*, can be approached by increasing the sampling period. In the scenario considered in this paper, it is assumed that the sampling interval is not adjustable. However, by using a symbol alphabet which varies periodically, *average* data rates as close as desired to H can be achieved with sufficiently large cycle lengths, similar to the way that irrationals are approximated by rationals.

A periodic alphabet is also suggested by a consideration of the plant dynamics. By the block structure of the real Jordan form $\mathbf{J}$, the filter process (5.1) consists of b subsystems with decoupled dynamical matrices $\mathbf{J}_i \in \mathbb{R}^{n_i \times n_i}$, $i = 1, \ldots, b$. The speed at which the ith subsystem grows in any direction is determined by the eigenvalue λ_i and, intuitively, a larger $|\lambda_i|$ necessitates using a higher data rate. A natural approach is to cycle through the subsystems and encode each at a rate determined by the corresponding level of instability.

Notwithstanding the discussion above, the implementation of time-varying alphabets can be difficult. In the coder-controller presented below, the alphabet size is in fact kept constant for the initial part of each cycle and no data is transmitted for the remainder. The subsystems are then allocated different effective data rates by means of a time-sharing protocol. More explicitly, time is divided into cycles of sufficiently large duration $\tau \in \mathbb{Z}_+$ and within each cycle, the components of the unstable subsystem states $\bar{\mathbf{x}}_k^{(i)} \in \mathbb{R}^{n_i}$ are allocated transmission slots of fixed length τ_i, roughly proportional to $\log_2 |\lambda_i|$. During each slot a fixed alphabet of size $\mu_k \equiv \mu \geq 2$ is used to quantize the corresponding subsystem state component with a total of μ^{τ_i} levels. Towards the end of each cycle, there is then a quiet slot during which no information is transmitted, i.e. $\mu_k = 1$.

5.2. Stabilizing Coder-Controller. The coder-controller to be applied is defined below and analyzed in subsection 5.3. First however, the static quantizer which is its basis is constructed, and a key lemma which stochastically bounds the quantizer errors is presented.

5.2.1. Quantizer. Recall that the floating point representation of a number $x \geq 1$ can be generated recursively by means of the following algorithm:

1. At iteration i, let $x \in J_i$, where $J_0 = [1, \infty)$.
2. At iteration $i + 1$, if J_i was the semi-infinite interval $[10^i, \infty)$ then partition it into 9 contiguous, disjoint subintervals of length 10^i and 1 semi-infinite interval $[10^{i+1}, \infty)$. If J_i was bounded however, then partition it into 10 equally long subintervals. In both cases set J_{i+1} to the subinterval containing x.
3. Repeat step 2 until $i = $ some predefined $\nu \geq 1$.
4. Approximate x by the lower limit of J_ν.

At termination, each interval $[10^{i-1}, 10^i)$, $i = 1, \ldots, \nu$, has been partitioned into $9 \times 10^{\nu-i}$ subintervals of equal length, and each x in it approximated as a floating point number with $\nu - i$ significant figures. Including the semi-infinite 'overload' subinterval $[10^\nu, \infty)$, there are 10^ν subintervals in total and so this algorithm can be viewed as a non-uniform quantization of $x \geq 1$ with 10^ν points.

The scalar quantizer which underpins the coder-controller constructed here is basically an extension of this floating-point scheme to $x \in \mathbb{R}$, with a base which may be non-decimal. First, select $\varrho > 1$ and let

$$(5.6) \qquad\qquad r_i \stackrel{\Delta}{=} \varrho^{i-1}, \quad i \in \mathbb{Z}_+.$$

Then select a base $\mu \geq 2$ and for any integer $\nu \geq 2$ generate μ^ν disjoint, exhaustive intervals symmetrically about the origin by

(i) partitioning $[-r_1, r_1] = [-1, 1]$ into $(\mu^2 - 2)\mu^{\nu-2}$ intervals of length $2/[(\mu^2 - 2)\mu^{\nu-2}]$,

(ii) partitioning $(r_{i-1}, r_i]$ and $[-r_i, -r_{i-1})$ each into $(\mu - 1)\mu^{\nu-i}$ intervals of length $(\varrho^{i-1} - \varrho^{i-2})/[(\mu - 1)\mu^{\nu-i}]$, $\forall i \in [2, 3, \ldots, \nu]$,

(iii) leaving (r_ν, ∞) and $(-\infty, -r_\nu)$ as the right- and left-most intervals.

In general, there are precisely $(\mu^2 - 2)\mu^{\nu-2} + 2 + \sum_{2 \leq i \leq \nu} 2(\mu - 1)\mu^{\nu-i} = \mu^\nu$ intervals. Label them $I(0), I(1), \ldots, I(\mu^\nu - 1)$, from left- to right-most, and $\forall x \in \mathbb{R}$ let

$$(5.7)\quad \kappa_\nu(\omega) \triangleq \begin{cases} \text{half-length of } I(\omega) & \text{if } 1 \leq \omega \leq \mu^\nu - 2, \\ 0.5(1 - 1/\mu)^{-1}(r_{\nu+1} - r_\nu) & \text{if } \omega = \mu^\nu - 1, \\ -0.5(1 - 1/\mu)^{-1}(r_{\nu+1} - r_\nu) & \text{if } \omega = 0. \end{cases},$$

$$(5.8)\quad q_\nu(x) := \varpi_\nu(\omega) := \begin{cases} \text{midpoint of } I(\omega) & \text{if } 1 \leq \omega \leq \mu^\nu - 2, \\ r_\nu + \kappa_\nu(\omega) & \text{if } \omega = \mu^\nu - 1, \\ -r_\nu - \kappa_\nu(\omega) & \text{if } \omega = 0. \end{cases} \text{ if } x \in I(\omega),$$

The precise form of the equations above is immaterial, some of the constants being selected solely to simplify subsequent analysis. Observe that like the floating point quantizer, the intervals of $q_{\nu+1}$ can be generated recursively by partitioning each interval of q_ν into μ subintervals. Furthermore, as $\nu \to \infty$ the range $[-r_\nu, r_\nu]$ covered by finite I's becomes unbounded, and at the same time any number x is eventually captured in an interval with length $\to 0$. Both these competing properties are necessary for the quantization error of a random variable with infinite support to approach zero pointwise as ν increases.

However, to show the attainability of (2.6) convergence in a stronger sense will be required; in particular, the mean square quantization error should diminish like the inverse square of the number of levels, $\mu^{-2\nu}$. Guaranteeing this for fat-tailed distributions is the real motivation for the exponential form of (5.6). If the distribution being quantized had exponentially-decaying tails, then it can be shown that (5.6) leads to a waste of quantizer levels on regions of very low probability and, consequently, a shortage of levels in high-probability regions. However, as the tails may decay according to a power law, a sufficiently large ϱ ensures that both the low- and high-probability mean square error contributions die off like $\mu^{-2\nu}$.

In fact, a slightly stronger result can be proven. Before stating it formally, for any random variable $L \in \mathbb{R}_+$ and random vector $\mathbf{X}$ define the functional

$$(5.9)\quad \mathrm{M}_\varepsilon\{\mathbf{X}|L\} := \mathrm{E}\{L^2\} + \mathrm{E}\{\|\mathbf{X}\|^{2+\varepsilon}L^{-\varepsilon}\}.$$

This cannot be smaller than the mean square norm of $\mathbf{X}$, since

$$\mathrm{E}\|\mathbf{X}\|^2 = \mathrm{E}\left\{\|\mathbf{X}\|^2 \left[\chi(\|\mathbf{X}\| \leq L) + \chi(\|\mathbf{X}\| > L)\right]\right\},$$
$$\leq \mathrm{E}\{L^2\} + \mathrm{E}\left\{\|\mathbf{X}\|^2(\|\mathbf{X}\|/L)^\varepsilon \chi(\|\mathbf{X}\| > L)\right\},$$
$$(5.10)\qquad \leq \mathrm{E}\{L^2\} + \mathrm{E}\left\{\|\mathbf{X}\|^{2+\varepsilon}L^{-\varepsilon}\right\} \equiv \mathrm{M}_\varepsilon\{\mathbf{X}|L\},$$

where χ is the usual indicator function. The mean square quantizer errors generated by q_ν can then be bounded as follows:

LEMMA 5.2. *Let $X \in \mathbb{R}$, $L > 0$ be random variables with $\mathrm{E}|X|^{2+\varepsilon} < \infty$ for some $\varepsilon > 0$. If the quantizer parameter ϱ of (5.6) and the base $\mu \in [2, 3, \ldots)$ are selected so*

that $\varrho > \mu^{2/\varepsilon}$, then the quantizer errors $X - Lq_\nu(X/L)$ satisfy

$$(5.11) \qquad \mathrm{M}_\varepsilon \left\{ X - Lq_\nu \left(\frac{X}{L} \right) \middle| L\kappa_\nu(\Omega) \right\} \leq \frac{\zeta}{\mu^{2\nu}} \mathrm{M}_\varepsilon\{X|L\}, \quad \forall \nu \in [2,3,\ldots)$$

where q_ν, κ_ν are defined in (5.8)–(5.7), $\Omega \in [0,1,\ldots\mu^\nu-1]$ is the index of the quantizer level $q_\nu(X/L)$, and $\zeta > 0$ is determined only by $\varepsilon, \mu,$ and ϱ.

Proof: See Appendix E.

By (5.10), the LHS exceeds the mean square quantizer error $\mathrm{E}|X - Lq_\nu(X/L)|^2$, so this result upper-bounds the latter and guarantees that it decreases as fast as the square of the number of quantizer levels. The condition $\varrho > \mu^{2/\varepsilon}$ is crucial here, as it relates the speed at which the quantizer range increases with ν to the fatness of the distribution tails, and thereby ensures that the contribution of the overload regions decays faster than $\mu^{-2\nu}$, for any fixed integer $\mu \geq 2$.

However, the real utility of this lemma lies in the appearance of M_ε on both sides, together with the independence of the constant ζ of the distribution of X. These two facts permit (5.11) to be used recursively when proving coder-controller stability in subsection 5.3. In contrast, a similar inequality relating the mean square norms of the error and X, via a density-independent factor decaying like the number of levels squared, is impossible.[1] What can be done instead is to upper-bound the mean square error by some higher moment of X as in Lemma 6.6 of [13], an approach which does not permit recursive application.

5.2.2. Time-Sharing Protocol and Coder. The quantizer above will now be used to construct a coder-controller. The measurements are first passed through a Kalman filter to generate a fully-observed process of the form (5.1). In order to simplify the analysis, and reduce subscript clutter, a non-predictive filter with output $\bar{\mathbf{x}}_k := \bar{\mathbf{x}}_{k|k}$ is used. As discussed in §5.1, its mean-square stability is equivalent to that of the original system (2.1), and its innovations $\mathbf{z}_k$, $k \in \mathbb{Z}_+$, are uniformly bounded in $(2+\varepsilon)$th absolute moment.

Divide times $k \in \mathbb{W}$ into *cycles* $[j\tau, \ldots, (j+1)\tau - 1]$, $j \in \mathbb{W}$, of uniform integer duration $\tau \in \mathbb{Z}_+$. Let R' be any given number greater than H (2.6), and select any integer $\mu \geq 2^{R'}$. With $\mathcal{U}$ denoting the index set of unstable subsystems (4.1), subdivide each cycle into f *transmission slots* of duration τ_i, for each scalar component of $\bar{\mathbf{x}}_{j\tau}^{(i)} \in \mathbb{R}^{n_i}$, followed by a *quiet slot* of duration $\tau - \sum_{i \in \mathcal{U}} n_i\tau_i$, where

$$(5.12) \qquad \tau_i \overset{\Delta}{=} \lfloor \tau \log_\mu(\xi|\lambda_i|) \rfloor + 1, \quad \forall i \in \mathcal{U},$$

with $\lfloor \cdot \rfloor$ denoting rounding down. If the parameter ξ is chosen to satisfy

$$(5.13) \qquad 0 < f \log_2 \xi < R' - \sum_{i \in \mathcal{U}} n_i \log_2 |\lambda_i| \equiv R' - \sum_{|\eta_l| \geq 1} \log_2 |\eta_l|,$$

then the choice of transmission slot durations (5.12) is feasible, since

$$(5.14) \quad \sum_{i \in \mathcal{U}} n_i\tau_i \leq \sum_{i \in \mathcal{U}} n_i \left(\tau \frac{\log_2(\xi|\lambda_i|)}{\log_2 \mu} + 1 \right) \leq \tau \frac{\sum_{i \in \mathcal{U}} n_i \log_2 |\lambda_i| + f \log_2 \xi}{R'} + f.$$

---

[1]The reason for this is essentially that even if p_X has a finite second moment, $|x|^2 p_X(x)$ may decay so slowly with large x that the overload regions dominate the mean square error, making it decrease slower than the inverse square number of levels.

As the coefficient of τ is less than unity by (5.13), the sum of transmission slot durations is less than any sufficiently large cycle length τ.

Let the symbol alphabet $\mathcal{S}_k = \mathbb{Z}_\mu$ during transmission slots, and $= \{0\}$ during the quiet slot. Then by reasoning similar to the above, the asymptotic average data rate R of this periodic alphabet, equal to the average data rate over one cycle, satisfies

$$R = \frac{1}{\tau} \sum_{i \in \mathcal{U}} n_i \tau_i \log_2 \mu \leq \frac{\log_2 \mu}{\tau} \sum_{i \in \mathcal{U}} n_i \left(\tau \frac{\log_2(\xi|\lambda_i|)}{\log_2 \mu} + 1 \right),$$

$$= \sum_{i \in \mathcal{U}} n_i \log_2 |\lambda_i| + f \log_2 \xi + \frac{f \log_2 \mu}{\tau} < R',$$

for all sufficiently large τ. As R' is any number exceeding H, this confirms that the data rate of this protocol can be made arbitrarily close to H, leaving aside for now the question of stability.

Just before the start of a cycle at time $k = j\tau$, let $l_j \in \mathbb{R}_+$ be the adaptive quantizer scaling factor discussed in §5.1, and $\hat{\mathbf{x}}_{j\tau} \in \mathbb{R}^n$ be an estimate of $\bar{\mathbf{x}}_{j\tau}$ internal to the coder. The *coder state* is then defined as $\psi_j \triangleq (\hat{\mathbf{x}}_{j\tau}, l_j)$. Indexing the scalar components of vectors $\in \mathbb{R}^{n_i}$ by an additional superscript $h \in [1, \ldots, n_i]$, at the start of the transmission slot for $\bar{x}_{j\tau}^{(i,h)}$ let it be scaled and quantized via

$$(5.15) \qquad \varpi_{\tau_i}(\omega_j^{(i,h)}) \equiv q_{\tau_i}([\bar{x}_{j\tau}^{(i,h)} - \hat{x}_{j\tau}^{(i,h)}]/l_j) \quad \forall h \in [1, \ldots, n_i],$$

where ϖ_ν and q_ν are defined in (5.8). The index $\omega_j^{(i,h)} \in [0, \ldots, \mu^\tau - 1]$ of the selected quantizer level is then expanded as τ_i base-μ digits and transmitted. After this has been done for all unstable state components, the coder state is updated via

$$(5.16) \qquad \hat{\mathbf{x}}_{(j+1)\tau} = \mathbf{J}^\tau \left[\hat{\mathbf{x}}_{j\tau} + l_j \varpi(\omega_j) \right] + \sum_{k=j\tau}^{(j+1)\tau - 1} \mathbf{J}^{(j+1)\tau - 1 - k} \mathbf{TBL} \hat{\mathbf{x}}_k,$$

$$(5.17) \quad \text{where } \hat{\mathbf{x}}_{k+1} = (\mathbf{J} + \mathbf{TBL})\hat{\mathbf{x}}_k, \quad \forall k \in [j\tau, \ldots, j\tau + \tau - 2], \quad \hat{\mathbf{x}}_0 = \mathbf{0},$$

$$(5.18) \qquad l_{j+1} = \max_{i \in \mathcal{U}, h \in [1, \ldots, n_i]} \left\{ \sigma, \, l_j |\lambda_i|^\tau \kappa_{\tau_i} \left(\omega_j^{(i,h)} \right) \right\}, \quad \forall j \in \mathbb{W}, \quad l_0 = \sigma.$$

In the above $\varpi(\omega_j) \in \mathbb{R}^n$ is the vector with (i, h)th component $\varpi_{\tau_i}(\omega_j^{(i,h)})$ for $i \in \mathcal{U}$ and 0 for $i \notin \mathcal{U}$, $\mathbf{L} \in \mathbb{R}^{p \times n}$ is the certainty-equivalent controller gain matrix, and $\sigma^{2+\varepsilon}$ is a uniform upper bound on the $(2+\varepsilon)$th absolute moment of

$$(5.19) \qquad \mathbf{G}_j := \sum_{i=1}^{\tau} \mathbf{J}^{\tau - i} \mathbf{Z}_{j\tau + i}, \quad \forall j \in \mathbb{W}.$$

5.2.3. Controller. Similar to the coder, define a controller internal state $\psi_j^{\mathrm{con}} \triangleq (\hat{\mathbf{x}}_{j\tau}^{\mathrm{con}}, l_j^{\mathrm{con}}) \in \mathbb{R}^n \times \mathbb{R}_+$, initialized when $j = 0$ to $(\mathbf{0}, \sigma)$. At any time $k \in [j\tau, \ldots, j\tau + \tau - 1]$ in the cycle, a certainty-equivalent control signal

$$(5.20) \qquad \mathbf{u}_k = \mathbf{L}\hat{\mathbf{x}}_k^{\mathrm{con}}$$

is applied, where $\hat{\mathbf{x}}_k^{\mathrm{con}}$ is given by (5.17) (with superscripts 'con'), and $\mathbf{L}$ is the given gain matrix s.t. $\mathbf{J} + \mathbf{TBL}$ is strictly stable.

By the time-sharing protocol, the last transmission slot during this cycle ends at time $j\tau + c(\tau) - 1$, where $c(\tau) \triangleq \sum_{i \in \mathcal{U}} n_i \tau_i$. Recalling that the channel has a delay d, at time $k = j\tau + c(\tau) - 1 + d$ the controller then has available all the symbols $s_{j\tau}, \ldots, s_{j\tau+c-1}$, comprising the base-$\mu$ expansions of the quantizer level indices $\omega_j^{(i,h)}$, $h \in [1, \ldots, n_i]$, $i \in \mathcal{U}$. By reasoning similar to (5.14), $c(\tau) + d \leq \tau$ sufficiently large, so that these indices are guaranteed to be received before the beginning of the next cycle at time $(j+1)\tau$. The controller then updates its internal state via the same recursive equations (5.16)–(5.18) as the coder.

5.3. Analysis. A uniform bound on the mean square norms of the filter process (5.1) using the coder-controller above will now be derived, for a data rate arbitrarily close to the lower bound (2.6). First, it is shown that the coder error $\mathbf{F}_k \triangleq \bar{\mathbf{X}}_k - \hat{\mathbf{X}}_k$ is uniformly bounded in mean square norm over times $k = j\tau$, $j \in \mathbb{W}$, by using the functional M_ε defined in (5.9) and Lemma 5.2. The mean square stability of $\mathbf{F}_k$ over all integer times is then deduced, which in turn will be shown to imply that of $\bar{\mathbf{X}}_k$, $k \in \mathbb{W}$.

Observe that since the initial controller and coder internal states $\psi_0^{\mathrm{con}}, \psi_0$ are equal and, furthermore, the same update equations (5.16)–(5.18) are used for each, it follows that $\hat{\mathbf{x}}_k^{\mathrm{con}} = \hat{\mathbf{x}}_k$ and $l_j^{\mathrm{con}} = l_j$, $\forall j, k \in \mathbb{W}$. The superscript 'con' is thus dropped in the analysis. Substituting (5.20) into the filter recursion (5.1) and iterating over a cycle,

$$\bar{\mathbf{x}}_{(j+1)\tau} = \mathbf{J}^\tau \bar{\mathbf{x}}_{j\tau} + \sum_{k=j\tau}^{(j+1)\tau-1} \mathbf{J}^{(j+1)\tau-1-k} \left(\mathbf{TBL}\hat{\mathbf{x}}_k + \mathbf{z}_{k+1} \right),$$

$$= \mathbf{J}^\tau \bar{\mathbf{x}}_{j\tau} + \mathbf{g}_j + \sum_{k=j\tau}^{(j+1)\tau-1} \mathbf{J}^{(j+1)\tau-1-k} \mathbf{TBL}\hat{\mathbf{x}}_k,$$

where $\mathbf{g}_j \triangleq \sum_{k=1}^{\tau} \mathbf{J}^{\tau-k} \mathbf{z}_{j\tau+k}$ (5.19). Subtracting this from (5.16), and then exploiting the block-diagonal structure $\mathbf{J} \equiv \mathrm{diag}(\mathbf{J}_1, \ldots \mathbf{J}_b)$ of the real Jordan form, where $\mathbf{J}_i \in \mathbb{R}^{n_i \times n_i}$,

$$\mathbf{f}_{(j+1)\tau} = \mathbf{J}^\tau [\mathbf{f}_{j\tau} - l_j \varpi(\omega_j)] + \mathbf{g}_j,$$

$$(5.21) \quad \Leftrightarrow \mathbf{f}_{(j+1)\tau}^{(i)} = \mathbf{J}_i^\tau \left[\mathbf{f}_{j\tau}^{(i)} - l_j \varpi(\omega_j)^{(i)} \right] + \mathbf{g}_j^{(i)} \in \mathbb{R}^{n_i}, \quad \forall i \in [1, \ldots, b], j \in \mathbb{W}.$$

By (5.15) and the definition of $\varpi(\omega_j)$, $\varpi(\omega_j)^{(i)} \triangleq \mathbf{0}$, $\forall i \notin \mathcal{U}$, in which case the RHS above simply becomes the recursion $\mathbf{f}_{(j+1)\tau}^{(i)} = \mathbf{J}_i^\tau \mathbf{f}_{j\tau}^{(i)} + \mathbf{g}_j^{(i)}$. Recall that each block $\mathbf{J}_i$ has exactly either one real eigenvalue or two complex conjugate eigenvalues λ_i, λ_i^*. As $|\lambda_i| < 1$, $\forall i \notin \mathcal{U}$, and furthermore the noise term has a uniform moment bound $\mathrm{E}\|\mathbf{G}_j^{(i)}\|^2 \leq \sigma^2$, it immediately follows that $\mathrm{E}\|\mathbf{F}_{j\tau}^{(i)}\|^2$ must be uniformly bounded for all strictly stable subsystems.

Hence, only the unstable subsystems $i \in \mathcal{U}$ need be considered. For each such i, $\varpi(\omega_j)^{(i)} \in \mathbb{R}^{n_i}$ is defined to be the quantizer point vector with hth component $q_{\tau_i}(f_{j\tau}^{(i,h)}/l_j)$. Applying square-norms and the triangle-inequality to (5.21), $\forall i \in \mathcal{U}, j \in \mathbb{W}$,

$$\|\mathbf{f}_{(j+1)\tau}^{(i)}\|^2 \leq 2^2 \left[\|\mathbf{J}_i^\tau\|^2 \left\| \mathbf{f}_{j\tau}^{(i)} - l_j \varpi(\omega_j)^{(i)} \right\|^2 + \|\mathbf{g}_j^{(i)}\|^2 \right],$$

$$(5.22) \qquad = 4 \left[\| \mathbf{g}_j^{(i)} \|^2 + \| \mathbf{J}_i^\tau \|^2 \sum_{h=1}^{n_i} \left| f_{j\tau}^{(i,h)} - l_j q_{\tau_i} \left(f_{j\tau}^{(i,h)} / l_j \right) \right|^2 \right],$$

using (5.15) and the definition of $\varpi(\omega_j)$. As each $\mathbf{J}_i$ is *similar* to the block diagonal matrix of all Jordan blocks associated with λ_i, a trivial adaptation of a result in [15] (pg. 138) states that $\exists \zeta_0 > 0$ such that

$$(5.23) \qquad \| \mathbf{J}_i^\tau \| \leq \zeta_0 \tau^{n_i - 1} |\lambda_i|^\tau, \quad \forall i \in [1, \dots, b], \tau \in \mathbb{Z}_+.$$

Let the stacked vector of unstable subsystem errors be $\mathbf{f}_k^{\mathrm{u}} \triangleq [\mathbf{f}_k^{(1)\mathrm{T}}, \dots, \mathbf{f}_k^{(|\mathcal{U}|)\mathrm{T}}]^\mathrm{T}$ and define $\mathbf{g}_j^{\mathrm{u}}$ in a similar way. By summing (5.22) over $i \in \mathcal{U}$, applying the growth rate bound above, and twice using the trivial inequality $\left(\sum_{1 \leq l \leq r} |y_l| \right)^\alpha \leq r^\alpha \sum_{1 \leq l \leq r} |y_l|^\alpha$, $\forall r \in \mathbb{Z}_+, \alpha > 0$, it then follows that $\exists \phi \geq 1$ s.t.

$$\| \mathbf{f}_{(j+1)\tau}^{\mathrm{u}} \|^{2+\varepsilon} = \left(\sum_{i \in \mathcal{U}} \| \mathbf{f}_{(j+1)\tau}^{(i)} \|^2 \right)^{1+\varepsilon/2},$$

$$\leq \phi \left(\| \mathbf{g}_j^{\mathrm{u}} \|^{2+\varepsilon} + \sum_{i \in \mathcal{U}} |\tau^{n_i - 1} \lambda_i^\tau|^{2+\varepsilon} \sum_{h=1}^{n_i} \left| f_{j\tau}^{(i,h)} - l_j q_{\tau_i} \left(f_{j\tau}^{(i,h)} / l_j \right) \right|^{2+\varepsilon} \right).$$

Dividing by l_{j+1}^ε and taking expectations,

$$\mathrm{E} \left\{ \| \mathbf{F}_{(j+1)\tau}^{\mathrm{u}} \|^{2+\varepsilon} L_{j+1}^{-\varepsilon} \right\}$$

$$\leq \phi \left(\mathrm{E} \left\{ \frac{\| \mathbf{G}_j^{\mathrm{u}} \|^{2+\varepsilon}}{L_{j+1}^\varepsilon} \right\} + \sum_{i \in \mathcal{U}} \left| \tau^{n_i - 1} \lambda_i^\tau \right|^{2+\varepsilon} \sum_{h=1}^{n_i} \mathrm{E} \left\{ \frac{|F_{j\tau}^{(i,h)} - L_j q_{\tau_i}(F_{j\tau}^{(i,h)}/L_j)|^{2+\varepsilon}}{L_{j+1}^\varepsilon} \right\} \right),$$

$$\leq \phi \left(\mathrm{E} \left\{ \frac{\| \mathbf{G}_j^{\mathrm{u}} \|^{2+\varepsilon}}{\sigma^\varepsilon} \right\} + \sum_{i \in \mathcal{U}} \left| \tau^{n_i - 1} \lambda_i^\tau \right|^{2+\varepsilon} \sum_{h=1}^{n_i} \mathrm{E} \left\{ \frac{|F_{j\tau}^{(i,h)} - L_j q_{\tau_i}(F_{j\tau}^{(i,h)}/L_j)|^{2+\varepsilon}}{[L_j |\lambda_i|^\tau \kappa_{\tau_i}(\Omega_j^{(i,h)})]^\varepsilon} \right\} \right),$$

$$= \phi \left(\mathrm{E} \left\{ \frac{\| \mathbf{G}_j^{\mathrm{u}} \|^{2+\varepsilon}}{\sigma^\varepsilon} \right\} + \sum_{i \in \mathcal{U}} \tau^{(n_i - 1)(2+\varepsilon)} |\lambda_i|^{2\tau} \sum_{h=1}^{n_i} \mathrm{E} \left\{ \frac{|F_{j\tau}^{(i,h)} - L_j q_{\tau_i}(F_{j\tau}^{(i,h)}/L_j)|^{2+\varepsilon}}{[L_j \kappa_{\tau_i}(\Omega_j^{(i,h)})]^\varepsilon} \right\} \right),$$

$$(5.24)$$

where the second inequality is a consequence of the definition of l_{j+1} (5.18).

   Now, let

$$(5.25) \qquad \vartheta_j := \mathrm{M}_\varepsilon \{ \mathbf{F}_{(j+1)\tau}^{\mathrm{u}} | L_j \} \equiv \mathrm{E} \{ L_j^2 \} + \mathrm{E} \{ \| \mathbf{F}_{(j+1)\tau}^{\mathrm{u}} \|^{2+\varepsilon} L_j^{-\varepsilon} \}, \quad \forall j \in \mathbb{W}.$$

By (5.10), the RHS is never less than $\mathrm{E} \| \mathbf{F}_{(j+1)\tau}^{\mathrm{u}} \|^2$, so to establish the mean square boundedness of the errors it is sufficient to show that $\sup_{j \in \mathbb{W}} \vartheta_j < \infty$. Observe that

$$\mathrm{E} \{ L_{j+1}^2 \} \equiv \mathrm{E} \left\{ \max_{i \in \mathcal{U}, h \in [1, \dots, n_i]} \left\{ \sigma^2, L_j^2 |\lambda_i|^{2\tau} \kappa_{\tau_i} \left(\Omega_j^{(i,h)} \right)^2 \right\} \right\},$$

$$\leq \sigma^2 + \sum_{i \in \mathcal{U}} |\lambda_i|^{2\tau} \sum_{h=1}^{n_i} \mathrm{E} \left| L_j \kappa_{\tau_i} \left(\Omega_j^{(i,h)} \right) \right|^2.$$

Adding this to (5.24), noting that $\mathrm{E}\|\mathbf{G}_j^{\mathrm{u}}\|^{2+\varepsilon} \leq \sigma^{2+\varepsilon}$, $\tau \geq 1$, and using definition (5.25),

$$\vartheta_{j+1} \leq$$

$$\phi\left(\sigma^2 + \sum_{i \in \mathcal{U}} \tau^{(n_i-1)(2+\varepsilon)} |\lambda_i|^{2\tau} \sum_{h=1}^{n_i} \mathrm{M}_\varepsilon\left\{ \left. F_{j\tau}^{(i,h)} - L_j q_{\tau_i}\left(\frac{F_{j\tau}^{(i,h)}}{L_j}\right) \right| L_j \kappa_{\tau_i}\left(\Omega_j^{(i,h)}\right) \right\}\right)$$

(5.26)

Applying Lemma 5.2 to each term in the inner sum, with $X = F_{j\tau}^{(i,h)}$, $L = L_j$, $\nu = \tau_i$ and $\Omega = \Omega_j^{(i,h)}$, $\forall j \in \mathbb{W}$,

$$\vartheta_{j+1} \leq \phi\left(\sigma^2 + \sum_{i \in \mathcal{U}} \tau^{(n_i-1)(2+\varepsilon)} |\lambda_i|^{2\tau} \sum_{h=1}^{n_i} \frac{\zeta}{\mu^{2\tau_i}} \mathrm{M}_\varepsilon\left\{ \left. F_{j\tau}^{(i,h)} \right| L_j \right\}\right),$$

$$\leq \phi\left(\sigma^2 + \sum_{i \in \mathcal{U}} \tau^{(n_i-1)(2+\varepsilon)} |\lambda_i|^{2\tau} \sum_{h=1}^{n_i} \frac{\zeta}{\mu^{2\tau_i}} \mathrm{M}_\varepsilon\left\{ \left. \mathbf{F}_{j\tau}^{\mathrm{u}} \right| L_j \right\}\right),$$

(5.27) $$\equiv \phi\sigma^2 + \phi\zeta\left(\sum_{i \in \mathcal{U}} n_i \tau^{(n_i-1)(2+\varepsilon)} \frac{|\lambda_i|^{2\tau}}{\mu^{2\tau_i}}\right)\vartheta_j,$$

where the second inequality is obtained from the definition of M_ε (5.9) and the trivial fact that the magnitude of a vector is never less than the magnitude of any of its components.

The inequality above is a first order, sublinear recursion for ϑ_j with a forcing term. By (5.12) and the fact that $x - \lfloor x \rfloor < 1$, $\forall x \in \mathbb{R}$, $\tau_i > \tau \log_\mu(\xi|\lambda_i|)$, $\forall i \in \mathcal{U}$, $\tau \in \mathbb{Z}_+$. This is equivalent to $\xi^\tau |\lambda_i|^\tau < \mu^{\tau_i}$, which when substituted into the above yields

$$\vartheta_{j+1} \leq \phi\sigma^2 + \phi\zeta\left(\sum_{i \in \mathcal{U}} n_i \tau^{(n_i-1)(2+\varepsilon)} \frac{1}{\xi^{2\tau}}\right)\vartheta_j, \quad \forall j \in \mathbb{W}.$$

As $\xi > 1$ by the left inequality of (5.13), $\tau^{(n_i-1)(2+\varepsilon)}\xi^{-2\tau} \to 0$ as $\tau \to \infty$, $\forall i$. Hence by choosing a sufficiently large, finite cycle length τ, the coefficient of ϑ_j above can be be made strictly less than 1. As the τ-dependent noise term σ^2 is finite for any fixed τ, the recursion above is then stable and yields uniformly bounded ϑ_j. By definition (5.25) and the inequality (5.10), $\mathrm{E}\|\mathbf{F}_{j\tau}^{\mathrm{u}}\|^2$ is then also uniformly bounded over $j \in \mathbb{W}$. Recalling the discussion after (5.21), the overall error vector $\mathbf{F}_{j\tau}$ must be as well.

The rest of the proof is straightforward. Subtracting (5.17) from (5.1), iterating forward r steps from time $j\tau$, and taking norms, at any time $k \equiv j\tau + r$ with $r \in [0, \ldots, \tau-1]$,

$$\|\mathbf{f}_k\| \leq \|\mathbf{J}^r\|\|\mathbf{f}_{j\tau}\| + \sum_{l=0}^{r-1} \|\mathbf{J}^{r-l}\mathbf{TBL}\|\|\mathbf{Z}_{j\tau+l+1}\|.$$

As $\mathbf{F}_{j\tau}$ and $\mathbf{Z}_{j\tau+l+1}$ are uniformly bounded in mean square norm and r can only take a finite number of values, the RHS and hence LHS are also uniformly mean square bounded over $k \in \mathbb{W}$. Rewriting (5.1) as

$$\bar{\mathbf{x}}_{k+1} = (\mathbf{J}+\mathbf{TBL})\bar{\mathbf{x}}_k + \mathbf{TBL}(\hat{\mathbf{x}}_k - \mathbf{x}_k) + \mathbf{z}_{k+1} = (\mathbf{J}+\mathbf{TBL})\bar{\mathbf{x}}_k - \mathbf{TBLf}_k + \mathbf{z}_{k+1}, \quad \forall k \in \mathbb{W},$$

the strict stability of $\mathbf{J} + \mathbf{TBL}$ then ensures the uniform boundedness of the mean square filter outputs $\mathrm{E}\|\bar{\mathbf{X}}_{k+1}\|^2$ over $k \in \mathbb{W}$. This completes the proof that the coder controller constructed in subsection 5.2 stabilizes the system (2.1) at data rates arbitrarily close to, but exceeding, the critical bound (2.6). $\square$

In the foregoing analysis, the assumption that the coder and controller internal states have the same initial value is crucial. Even if true, real digital channels invariably introduce data errors, causing the coder and controller states to eventually differ. It is thus important to emphasize that the scheme presented here is not intended in the present form to be a practical solution to communication-limited stabilization problems, but is primarily a theoretical construct for demonstrating stabilizability in the limited sense (2.5). Nonetheless it does possess some attributes, such as finite dimensionality, which make implementation easy and may serve as a foundation for a more practical scheme. In this respect, an important and as yet open extension of this research is the *internal* stability of finite-dimensional, data-rate-limited control loops, i.e. ensuring that the plant and coder-controller internal states remain mean square stable with a random *overall* initial condition, channel errors, and process and measurement noise. It is easy to see that redundancy must be incorporated in the transmitted symbols to counteract channel noise, but it is not evident if an analogue of the well-known *source-channel separation theorem* of information theory [25] applies.

6. Conclusion. In this paper, the problem of stabilizing a general stochastic linear system in mean square state norm under a feedback data rate constraint was investigated. By employing information theoretic techniques and a new quantizer error bound, an expression was derived for the smallest data rate above which such a system is stabilizable by a coding and control law, without imposing any structural or computational constraints and with very mild conditions on the system noise. This infimum rate is determined only by the unstable eigenvalues of the dynamical matrix and it was demonstrated that as the data rate approaches it from above the mean square states become unbounded for any coder-controller. To establish the attainability of this bound, a finite-dimensional scheme was constructed and shown to achieve stability at data rates arbitrarily close to it. Extensions of these results to nonlinear systems, linear systems with Markov parameters, and decentralized control are being investigated.

Appendix A. Proof of Proposition 2.2. Suppose that $R_1 \leq R_2$. By direct substitution of (2.8) into (2.9), each input $\mathbf{u}_k$ depends (in a time-varying way) on the link-1 symbol sequence $\tilde{s}^1_{k-d_1-d_2}$,

$$\mathbf{u}_k \equiv \phi_k(\tilde{s}^1_{k-d_1-d_2}), \quad \forall k \in \mathbb{W}.$$

This mapping, (2.7), and the alphabet sequence $\tilde{\mathcal{S}}^1_\infty$ then constitute a coder-controller with data rate R_1 (2.2) and link delay $d = d_1 + d_2$.

Now suppose that $R_2 < R_1$. By (2.7), the link-1 symbol sequence $\tilde{s}^1_{k-d_1}$ is also a time-varying function of the measurement sequence $\tilde{\mathbf{y}}_{k-d_1}$. Hence (2.8) can be rewritten in the form

$$s^2_k \equiv \theta_k(\tilde{\mathbf{y}}_{k-d_1}).$$

Defining the d_1-step-ahead link-2 symbol $c_k \triangleq s^2_{k+d_1}$, $\forall k \in \mathbb{W}$, the expression above and the actuator mapping (2.9) become

$$c_k = \theta_k(\tilde{\mathbf{y}}_k) \in \mathcal{S}^2_{k+d_1}, \quad \mathbf{u}_k = \delta^2_k(\tilde{c}_{k-d_1-d_2}).$$

This is a single-link coder-controller with delay $d = d_1 + d_2$. As the asymptotic average data rate is independent of constant time shifts of the alphabet, its value remains R_2.

The proof of the second part is straightforward. Let the coder (2.7) for link-1, with delay d_1, be given by (2.3) and set the coder and actuator for link-2, with delay $d_2 = d - d_1$, to be

$$s_k^2 = s_{k-d_1}^1 \equiv s_{k-d_1}, \quad \mathbf{u}_k = \delta_k(\tilde{s}_{k-d_2}^2) \equiv \delta_k(\tilde{s}_{k-d_1-d_2}).$$

Evidently, with regard to the plant this is equivalent to the single-link coder-controller (2.3)–(2.4). Furthermore, the link-2 alphabet is obviously periodic if that of link-1 is, with the same average data rate. □

Appendix B. Proof of Inequality (4.4). The argument is essentially that of Lemma 5 in [9]. Denote the mean square norm of $\mathbf{X}$ given $A = a$ by σ^2 and let ϕ be the symmetric, f-dimensional Gaussian distribution with zero mean and variance σ^2. By the nonnegativity of the *Kullback-Leibler information distance D*,

$$0 \leq D(p_{\mathbf{X}|a}\|\phi) \triangleq \int_{\mathbb{R}^f} p_{\mathbf{X}|a}(\mathbf{x}) \ln \frac{p_{\mathbf{X}|a}(\mathbf{x})}{\phi(\mathbf{x})} d\lambda(\mathbf{x}),$$

$$= \int_{\mathbb{R}^f} p_{\mathbf{X}|a}(\mathbf{x}) \ln p_{\mathbf{X}|a}(\mathbf{x}) d\lambda(\mathbf{x}) - \int_{\mathbb{R}^f} p_{\mathbf{X}|a}(\mathbf{x}) \ln \phi(\mathbf{x}) d\lambda(\mathbf{x}),$$

$$= -\mathrm{H}_a\{\mathbf{X}\} - \int_{\mathbb{R}^f} p_{\mathbf{X}|a}(\mathbf{x}) \left(-\frac{f}{2} \ln(2\pi\sigma^2) - \frac{\|\mathbf{x}\|^2}{2\sigma^2} \right) d\lambda(\mathbf{x}),$$

$$= -\mathrm{H}_a\{\mathbf{X}\} + 0.5 f \ln(2\pi\sigma^2) + 0.5.$$

$$\Rightarrow \mathrm{N}_a\{\mathbf{X}\} \equiv (2\pi e)^{-1} e^{2\mathrm{H}_a\{\mathbf{X}\}/f} \leq e^{1/f-1}\sigma^2. \quad \square$$

Appendix C. Proof of Lemma 4.1. By standard properties of joint and average differential entropy [7, 9],

$$\mathrm{H}_a\{\mathbf{X}|S\} = \mathrm{H}_a\{\mathbf{X}, S\} - \mathrm{H}_a S \geq \mathrm{H}_a\{\mathbf{X}\} - \mathrm{H}_a S \geq \mathrm{H}_a\{\mathbf{X}\} - \ln|\mathcal{S}|.$$

Differential entropy is undefined for discrete-valued random variables, but the joint entropy above may be taken to denote $-\mathrm{E}_a \ln \left(p_{\mathbf{X}|S,A}(\mathbf{X})\mathrm{P}\{S|A\} \right)$ while $\mathrm{H}_a\{S\}$ represents the base-e discrete entropy of S given $A = a$. The first inequality arises from the fact that the entropy of joint random variables can never be smaller than the individual entropies, while the second inequality is a consequence of the fact that the base-e entropy of a random variable in a finite alphabet $\mathcal{S}$ is at most $\ln|\mathcal{S}|$. Using Jensen's inequality for convex functions [7] and the lower bound above,

$$2\pi e \mathrm{N}_a\{\mathbf{X}|S\} = \mathrm{E}_a\{e^{2\mathrm{H}_{S,A}\{\mathbf{X}\}/f}\} \geq e^{\mathrm{E}_a\{2\mathrm{H}_{S,A}\{\mathbf{X}\}/f\}},$$

$$\equiv e^{2\mathrm{H}_a\{\mathbf{X}|S\}/f} \geq e^{2(\mathrm{H}_a\{\mathbf{X}\}-\ln|\mathcal{S}|)/f} = |\mathcal{S}|^{-2/f} 2\pi e \mathrm{N}_a\mathbf{X}. \quad \square$$

Appendix D. Proof of Proposition 5.1. Consider the finite-state predictive quantizer (5.4). As there are a finite number $|\mathcal{I}|$ of possible internal variables ι_k, $Q(\cdot)$ is bounded. By the strict stability of $\mathbf{J} + \mathbf{TBL}$, it then follows from (5.4) that $\hat{\mathbf{x}}_k$ is bounded over k, and hence $\exists \rho > 0$ s.t. $\|\mathbf{u}_k\| = \|\mathbf{L}\hat{\mathbf{x}}_k\| \leq \rho$.

Now, convert (2.1) into *standard* Jordan form via a complex similarity matrix $\mathbf{S}$. There is then at least one scalar component $x_k \in \mathbb{C}$ of the transformed state vector that satisfies the scalar, decoupled recursion

$$x_{k+1} = \eta x_k + v_k + u_k = \sum_{j=-1}^{k} \eta^{k-j}(v_j + u_j) \in \mathbb{C},$$

where $|\eta| > 1$ and v_k, u_k are the corresponding scalar components of $\mathbf{Sv}_k$, $\mathbf{Su}_k$ respectively. For convenience, $u_{-1} \triangleq 0$ and $\mathbf{v}_{-1} \triangleq \mathbf{Sx}_0$. Evidently $V_{-1}, V_0, \ldots$ are still independent, and $|u_k| \leq \rho$. Defining

$$\beta_k := \rho \sum_{j=0}^{k} |\eta|^{k-j} \geq \left| \sum_{j=0}^{k} \eta^{k-j} u_j \right|,$$

$$g_k := \sum_{j=-1}^{k} \eta^{k-j} v_j, \quad \bar{v}_k \triangleq \sum_{-1 \leq j \leq k, j \neq t} \eta^{-j} v_j, \quad \forall k \in \mathbb{W},$$

where the time $t \geq -1$ is specified in (5.5), it follows that

$$P\{\|\mathbf{SX}_{k+1}\| \geq \beta_k\} \geq P\{|X_{k+1}| \geq \beta_k\} = P\left\{ \left| G_k - \sum_{j=0}^{k} \eta^{k-j} U_j \right| \geq \beta_k \right\},$$

$$\geq P\{|G_k| - \beta_k \geq \beta_k\} = P\{|G_k| \geq 2\beta_k\},$$

$$= P\left\{ \left| \sum_{j=-1}^{k} \eta^{k-j} V_j \right| \geq 2\rho \sum_{j=0}^{k} |\eta|^{k-j} \right\} = P\left\{ \left| \sum_{j=-1}^{k} \eta^{-j} V_j \right| \geq 2\rho \sum_{j=0}^{k} |\eta|^{-j} \right\},$$

$$\geq P\left\{ \left| \sum_{j=-1}^{k} \eta^{-j} V_j \right| \geq \theta \right\} \equiv P\{|\eta^{-t} V_t + \bar{V}_k| \geq \theta\} \geq P\left\{ \Re(\eta^{-t} V_t) + \Re(\bar{V}_k) \geq \theta \right\},$$

$$\geq P\left\{ \Re(\eta^{-t} V_t) \geq \alpha\theta, \Re(\bar{V}_k) \geq (1-\alpha)\theta \right\},$$

$$= P\left\{ \Re(\eta^{-t} V_t) \geq \alpha\theta \right\} P\left\{ \Re(\bar{V}_k) \geq (1-\alpha)\theta \right\}, \quad \forall \alpha \in \mathbb{R}, k \geq t,$$

(D.1)

where $\theta \triangleq 2\rho \sum_{j \in \mathbb{W}} |\eta|^{-j} = 2\rho/(1 - |\eta|^{-1})$ and the last step follows from the mutual independence of V_j, $j \geq -1$.

Furthermore, as $E|V_j|^2$ is uniformly bounded, it follows from Holder's inequality that

$$E|\Re(\bar{V}_k)|^2 \leq E \left| \sum_{j \geq -1, j \neq t} \eta^{-j} V_j \right|^2 \leq \sum_{j \geq -1} |\eta|^{-j} E\left\{ \sum_{j \geq -1} |\eta|^{-j} |V_j|^2 \right\} < \infty, \quad \forall k \in \mathbb{W}.$$

By theorem 22.6 in [4], $\Re(\bar{V}_k)$ then converges with probability 1, and thus in distribution, to a random variable $\bar{V}$. Hence $\exists \alpha_* \in \mathbb{R}, \epsilon > 0, k_* \in \mathbb{W}$ s.t. $\forall k \geq k_*$,

$$P\left\{ \Re(\bar{V}_k) \geq (1 - \alpha_*)\theta \right\} \geq P\left\{ \bar{V} \geq (1 - \alpha_*)\theta \right\} - \epsilon > 0.$$

In addition, since $\Re(\eta^{-t} V_t)$ is just a scalar linear function of $\mathbf{V}_t$, (5.5) implies that $P\{\Re(\eta^{-t} V_t) > \vartheta\} > 0$, $\forall \vartheta \in \mathbb{R}$. Applying this and the inequality above to (D.1),

$$P\{\|\mathbf{X}_{k+1}\| \geq \beta_k\} \geq P\left\{ \Re(\eta^{-t} V_t) \geq \alpha_*\theta \right\} \left[P\left\{ \bar{V} \geq (1 - \alpha_*)\theta \right\} - \epsilon \right] \equiv \nu > 0, \quad \forall k \geq k_*.$$

It then follows that $\forall r > 0$,

$$E\|\mathbf{X}_{k+1}\|^r \geq E\left\{ \|\mathbf{X}_{k+1}\|^r \chi(\|\mathbf{X}_{k+1}\| \geq \beta_k) \right\} \geq \beta_k^r P\{\|\mathbf{X}_{k+1}\| \geq \beta_k\} \geq \beta_k^r \nu \to \infty,$$

since $\beta_k \to \infty$. The same reasoning applies to static memoryless coding (5.2). $\square$

Appendix E. Proof of Lemma 5.2. Let $\phi = l\kappa_\nu(\omega)$, where $\omega \in \mathbb{Z}_{\mu^\nu}$ is the index of the selected quantizer point $q_\nu(x/l) \equiv \varpi_\nu(\omega)$ and κ_ν, ϖ_ν are defined by (5.8), (5.7) respectively. If $1 \le \omega \le \mu^\nu - 2$, then the interval $I(\omega)$ which contains x/l is bounded with length $2\kappa_\nu(\omega) \equiv 2\phi/l$ and midpoint $q_\nu(x/l) \overset{\Delta}{=} \varpi_\nu(\omega)$. In this case, $|x - lq_\nu(x/l)| < \phi$, and $\forall \omega \in [1, \ldots \mu^\nu - 2]$,

$$(E.1) \qquad \mathrm{E}_{\omega,l}\left\{|X - Lq_\nu(X/L)|^{2+\varepsilon}\Phi^{-\varepsilon}\right\} \le \mathrm{E}_{\omega,l}\{\Phi^{2+\varepsilon}\Phi^{-\varepsilon}\} = \phi^2.$$

If $\omega = \mu^\nu - 1$, then x/l lies inside the semi-infinite interval $I(\mu^\nu - 1)$, defined as $(\varpi_\nu(\omega) - \phi/l, \infty)$. Hence

$$
\begin{aligned}
& \mathrm{E}_{\omega,l}\left\{|X - Lq_\nu(X/L)|^{2+\varepsilon}\Phi^{-\varepsilon}\right\} \\
&= \mathrm{E}_{\omega,l}\left\{|X - L\varpi_\nu(\Omega)|^{2+\varepsilon}\Phi^{-\varepsilon}\chi(|X - L\varpi_\nu(\Omega)| \le \Phi)\right\} \\
&\quad + \mathrm{E}_{\omega,l}\left\{[X - L\varpi_\nu(\Omega)]^{2+\varepsilon}\Phi^{-\varepsilon}\chi(X - L\varpi_\nu(\Omega) > \Phi)\right\}, \\
&\le \mathrm{E}_{\omega,l}\{\Phi^{2+\varepsilon}\Phi^{-\varepsilon}\} + \mathrm{E}_{\omega,l}\left\{X^{2+\varepsilon}\Phi^{-\varepsilon}\chi(X - L\varpi_\nu(\Omega) > \Phi)\right\}, \\
&= \phi^2 + \mathrm{E}_{\omega,l}\left\{|X|^{2+\varepsilon}[L\kappa_\nu(\omega)]^{-\varepsilon}\chi(X - L\varpi_\nu(\omega) > L\kappa_\nu(\omega))\right\}, \\
&\le \phi^2 + \kappa_\nu(\omega)^{-\varepsilon}\mathrm{E}_{\omega,l}\{|X|^{2+\varepsilon}L^{-\varepsilon}\}, \\
&= \phi^2 + \left[0.5(1 - 1/\mu)^{-1}(1 - \varrho^{-1})\varrho^\nu\right]^{-\varepsilon}\mathrm{E}_{\omega,l}\{|X|^{2+\varepsilon}L^{-\varepsilon}\}, \\
(E.2) \qquad &\le \phi^2 + \left[0.5(1 - 4^{-1/\varepsilon})\varrho^\nu\right]^{-\varepsilon}\mathrm{E}_{\omega,l}\{|X|^{2+\varepsilon}L^{-\varepsilon}\},
\end{aligned}
$$

since $1 - \varrho^{-1} > 1 - \mu^{-2/\varepsilon} > 1 - 4^{-1/\varepsilon}$ for $\mu \ge 2$. By the symmetry of the quantizer about the origin, the same bound applies if $\omega = 0$, corresponding to the other semi-infinite interval $I(0)$. Averaging this, (E.1) and (E.2) over Ω, L,

$$
\mathrm{E}\left\{\frac{|X - Lq_\nu(X/L)|^{2+\varepsilon}}{\Phi^\varepsilon}\right\} \le
$$

$$(E.3) \qquad \mathrm{E}\{\Phi^2\} + \left[0.5(1 - 4^{-1/\varepsilon})\right]^{-\varepsilon}\varrho^{-\varepsilon\nu}\mathrm{E}\left\{\frac{|X|^{2+\varepsilon}}{L^\varepsilon}\right\} =: \beta_\nu.$$

By the definitions of $\phi \equiv l\kappa_\nu(\omega)$ (5.7), and q_ν (5.8),

$\mathrm{E}_l\{\Phi^2\}$

$$
\begin{aligned}
&= \left[\frac{l}{(\mu^2 - 2)\mu^{\nu-2}}\right]^2 \mathrm{P}_l\left\{\frac{|X|}{L} \le r_1\right\} + \sum_{i=2}^\nu \left[\frac{(r_i - r_{i-1})l}{2(\mu - 1)\mu^{\nu-i}}\right]^2 \mathrm{P}_l\left\{r_{i-1} < \frac{|X|}{L} \le r_i\right\} \\
&\quad + \left[\frac{(r_{\nu+1} - r_\nu)l}{2(1 - 2/\mu)}\right]^2 \mathrm{P}_l\left\{\frac{|X|}{L} > r_\nu\right\}, \\
&= \left[\frac{l}{(\mu^2 - 2)\mu^{\nu-2}}\right]^2 \mathrm{P}_l\left\{\frac{|X|}{L} \le r_1\right\} \\
&\quad + \left[\frac{\varrho^{-1} - \varrho^{-2}}{2(\mu - 1)\mu^\nu}\right]^2 \left[\sum_{i=2}^\nu (\varrho\mu)^{2i}l^2\mathrm{P}_l\left\{r_{i-1} < \frac{|X|}{L} \le r_i\right\} + (\varrho\mu)^{2(\nu+1)}l^2\mathrm{P}_l\left\{\frac{|X|}{L} > r_\nu\right\}\right], \\
&\le \left[\frac{\mu^2 l}{(\mu^2 - 2)\mu^\nu}\right]^2 + \left[\frac{\varrho^{-1} - \varrho^{-2}}{2(\mu - 1)\mu^\nu}\right]^2 \sum_{i=2}^{\nu+1} (\varrho\mu)^{2i}l^2\mathrm{P}_l\left\{\frac{|X|}{L} \ge r_{i-1}\right\},
\end{aligned}
$$

$$\leq \left[\frac{2l}{\mu^\nu}\right]^2 + \left[\frac{1}{2(\mu-1)\varrho\mu^\nu}\right]^2 \sum_{i=2}^{\nu+1} (\varrho\mu)^{2i} l^2 P_l\left\{\frac{|X|}{L} \geq r_{i-1}\right\},$$

(E.4)

since $\mu^2/(\mu^2-2) \leq 2$ for $\mu \geq 2$ and $\varrho > 1$. By a Chebychev inequality type of argument,

$$
\begin{aligned}
l^2 P_l\{|X| > r_{i-1}L\} &\leq l^2 E_l\left\{[|X|/(r_{i-1}L)]^{2+\varepsilon} \chi(|X| > r_{i-1}L)\right\}, \\
&= r_{i-1}^{-2-\varepsilon} E_l\left\{|X|^{2+\varepsilon} L^{-\varepsilon} \chi(|X| > r_{i-1}L)\right\} \\
&\leq r_{i-1}^{-2-\varepsilon} E_l\{|X|^{2+\varepsilon} L^{-\varepsilon}\} = \varrho^{-(i-2)(2+\varepsilon)} E_l\{|X|^{2+\varepsilon} L^{-\varepsilon}\}, \; \forall i \geq 2.
\end{aligned}
$$

Substituting this into (E.4), averaging over L and letting $b \triangleq E\{|X|^{2+\varepsilon} L^{-\varepsilon}\}$,

$$E\Phi^2 \leq \frac{4EL^2}{\mu^{2\nu}} + \left[\frac{1}{2(\mu-1)\varrho\mu^\nu}\right]^2 \sum_{i=2}^{\nu+1} \frac{(\varrho\mu)^{2i}b}{\varrho^{(i-2)(2+\varepsilon)}} = \frac{4EL^2}{\mu^{2\nu}} + \left[\frac{\varrho^{1+\varepsilon}}{2(\mu-1)\mu^\nu}\right]^2 \sum_{i=2}^{\nu+1} \left(\frac{\mu^2}{\varrho^\varepsilon}\right)^i b.$$

As $\varrho > \mu^{2/\varepsilon}$, the geometric sum on the RHS is bounded with limit $\mu^4 \varrho^{-2\varepsilon}/(1-\mu^2\varrho^{-\varepsilon})$. Hence

$$E\Phi^2 \leq \frac{4EL^2}{\mu^{2\nu}} + \left[\frac{\varrho^{1+\varepsilon}}{2(\mu-1)\mu^\nu}\right]^2 \frac{\mu^4\varrho^{-2\varepsilon}}{1-\mu^2\varrho^{-\varepsilon}} b \equiv \frac{4EL^m + \zeta_0 b}{\mu^{2\nu}}.$$

Adding this to (E.3),

$$
\begin{aligned}
M_\varepsilon\{X - Lq_\nu(X/L)|\Phi\} &\leq E\Phi^2 + \beta_\nu \leq 2\frac{4EL^2 + \zeta_0 b}{\mu^{2\nu}} + \frac{[0.5(1-4^{-1/\varepsilon})]^{-\varepsilon}b}{\varrho^{\varepsilon\nu}}, \\
&\leq 2\frac{4EL^2 + \zeta_0 b}{\mu^{2\nu}} + \frac{[0.5(1-4^{-1/\varepsilon})]^{-\varepsilon}b}{\mu^{2\nu}}, \\
&\leq \max\left\{8, 2\zeta_0 + [0.5(1-4^{-1/\varepsilon})]^{-\varepsilon}\right\} \frac{EL^2 + b}{\mu^{2\nu}}. \qquad \square
\end{aligned}
$$

Note that virtually the same argument holds for the mean mth power quantization error, $m > 0$, by defining $M_{\varepsilon,m}\{X|L\} \triangleq EL^m + E\{|X|^{m+\varepsilon} L^{-\varepsilon}\}$ and setting $\varrho > \mu^{m/\varepsilon}$.

REFERENCES

[1] B. D. O. ANDERSON AND J. B. MOORE, *Optimal Filtering*, Prentice-Hall, 1979.

[2] J. BAILLIEUL, *Feedback designs for controlling device arrays with communication channel bandwidth constraints*, in ARO Workshop on Smart Structures, Penn. State U., USA, Aug 1999.

[3] ———, *Feedback designs in information-based control*, in Stochastic Theory and Control Proceedings of a Workshop held in Lawrence, Kansas, B. Pasik-Duncan, ed., Springer, Oct 2001, pp. 35–57.

[4] P. BILLINGSLEY, *Probability and Measure*, Wiley, 1995.

[5] V. S. BORKAR AND S. K. MITTER, *LQG control with communication constraints*, in Communications, Computation, Control and Signal Processing, Dordrecht, Boston, 1997, pp. 365–73.

[6] R. W. BROCKETT AND D. LIBERZON, *Quantized feedback stabilization of linear systems*, IEEE Trans. Autom. Contr., 45 (2000), pp. 1279–89.

[7] T. M. COVER AND J. A. THOMAS, *Elements of Information Theory*, Wiley, 1991.

[8] D. F. DELCHAMPS, *Stabilizing a linear system with quantized state feedback*, IEEE Trans. Autom. Contr., 35 (1990), pp. 916–24.

[9] A. Dembo, T. M. Cover, and J. A. Thomas, *Information theoretic inequalities*, IEEE Trans. Info. The., 37 (1991), pp. 1501–18.

[10] N. Elia and S. K. Mitter, *Stabilization of linear systems with limited information*, IEEE Trans. Autom. Contr., 46 (2001), pp. 1384–400.

[11] F. Fagnani and S. Zampieri, *Stability analysis and synthesis for scalar linear systems with a quantized feedback*, IEEE Trans. Autom. Contr., 48 (2003), pp. 1569–84.

[12] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, Kluwer, 1993.

[13] S. Graf and H. Luschgy, *Foundations of Quantization for Probability Distributions*, Springer, 2000.

[14] J. Hespanha, A. Ortega, and L. Vasudevan, *Towards the control of linear systems with minimum bit-rate*, in Proc. 15th Int. Symp. Math. The. Netw. Sys., U. Notre Dame, USA, Aug 2002.

[15] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.

[16] H. Ishii and B. A. Francis, *Quadratic stabilization of sampled-data systems with quantization*, Automatica, 39 (2003), pp. 1793–1800.

[17] J. C. Kieffer and J. G. Dunham, *On a type of stochastic stability for a class of encoding schemes*, IEEE Trans. Info. The., 29 (1983), pp. 703–17.

[18] D. Liberzon, *On stabilization of linear systems with limited information*, IEEE Trans. Autom. Contr., 48 (2003), pp. 304–7.

[19] G. N. Nair and R. J. Evans, *A finite-dimensional coder-estimator for rate-constrained state estimation*, in Proc. 14th IFAC World Congress, vol. I, Beijing, 1999, pp. 19–24.

[20] ———, *Stabilization with data-rate-limited feedback: tightest attainable bounds*, Sys. Contr. Lett., 41 (2000), pp. 49–56.

[21] ———, *Mean square stabilisability of stochastic linear systems with data rate constraints*, in Proc. 41st IEEE Conf. Dec. Contr., Las Vegas, USA, Dec 2002, pp. 1632–7.

[22] ———, *Exponential stabilisability of finite-dimensional linear systems with limited data rates*, Automatica, 39 (2003), pp. 585–93.

[23] I. R. Petersen and A. V. Savkin, *Multi-rate stabilization of multivariable discrete-time linear systems via a limited capacity communication channel*, in Proc. 40th IEEE Conf. Dec. Contr., 2001, pp. 304–9.

[24] A. Sahai, *Evaluating channels for control: capacity reconsidered*, in Proc. Am. Contr. Conf., Chicago, USA, 2000, pp. 2358–62.

[25] C. E. Shannon, *A mathematical theory of communication*, Bell Syst. Tech. Jour., (1948). Reprinted in 'Claude Elwood Shannon Collected Papers', IEEE Press, 1993.

[26] G. E. Shilov, *Linear Algebra*, Prentice-Hall, 1971.

[27] S. Tatikonda and S. Mitter, *Control under communication constraints*, in Proc. 38th Ann. Allerton Conf. Comm. Contr. Comp., Oct 2000, pp. 182–90.

[28] S. Tatikonda, A. Sahai, and S. K. Mitter, *Control of LQG systems under communication constraints*, in Proc. 37th IEEE Conf. Dec. Contr., 1998, pp. 1165–70.

[29] W. S. Wong and R. W. Brockett, *Systems with finite communication bandwidth constraints II: stabilization with limited information feedback*, IEEE Trans. Autom. Contr., 44 (1999), pp. 1049–53.