Stabilization Over Markov Feedback Channels:
The general case.

Lorenzo Coviello, Paolo Minero, and Massimo Franceschetti

Abstract

The problem of mean-square stabilization of a discrete-time linear dynamical system over a Markov time-varying digital feedback channel is studied. In the scalar case, it is shown that the system can be stabilized if and only if a Markov jump linear system describing the evolution of the estimation error at the decoder is stable —videlicet if and only if the product of the unstable mode of the system and the spectral radius of a second moment matrix that depends only on the Markov feedback rate is less than one. This result generalizes several previous data rate theorems that appeared in the literature, quantifying the amount of instability that can be tolerated when the estimated state is received by the controller over a noise free digital channel. In the vector case, a necessary condition for stabilizability is derived and a corresponding control scheme is presented, which is tight in some special cases and which strictly improves on a previous result on stability over Markov erasure channels. The proof techniques rely on both information-theoretic and control-theoretic tools.

Keywords: Quantized control, communication theory, entropy, Markov channels, Markov jump linear systems.

I. INTRODUCTION

We consider the problem of stabilization of a dynamical system where the estimated signal is transmitted for control over a time-varying communication channel, as depicted in Figure 1. This arises, for example, in pursuit evasion games where the state of the evader is estimated by...
distributed sensors and is communicated over a wireless fading channel to automatically control the pursuer [23].

The mathematical abstraction is that of an unstable linear, discrete-time, dynamical system whose state is observed, quantized, encoded, and sent to a decoder over a noiseless digital link that supports the transmission of $R_k$ bits at any given time step $k$, where $R_k$ evolves according to a Markov chain representing the current state of the channel. Based on the decoded message, the control signal is computed and applied to the system. Both the encoder and the decoder are assumed to have causal knowledge of the channel state information, a legitimate assumption for slow wireless fading channels for which the channel conditions can be learnt with a short training sequence.

Following this model, Tatikonda and Mitter [26] examined the special case where the rate process is constant in time and the system has bounded disturbances. Nair and Evans [18] studied the case where the disturbances have unbounded support, maintaining the rate constant. Martin et al. [16] analyzed the case with time-varying independent and identically distributed (i.i.d.) rate but bounded system disturbances. The work in [17] allowed both time-varying i.i.d. rate process and unbounded disturbances. Finally, You and Xie [27] considered the case of unbounded disturbances and the time-varying channel rate taking values 0 or $r$ according to the two-state Markov process depicted in Figure 2.

Results of these works analytically relate the speed of the dynamics of the plant to the information rate of the communication channel. They show that in order to guarantee stability, the rate must be large enough compared to the unstable modes of the system, so that it can compensate for the expansion of the state during the communication process. These kind of results are known in the literature as data rate theorems and there is interest in formulating them in the most general conditions.

In the present work, we consider a generalization that allows for unbounded system disturbances and models the time-varying rate of the channel as a homogeneous positive-recurrent Markov chain that takes values in a finite subset of the nonnegative integers. This allows arbitrary temporal correlations of the channel variations and includes all previous models mentioned above. Our results recover and sometime improve the ones mentioned in the above papers as special cases. Previous analysis methods cannot be applied directly and our analysis technique follows a different approach.
We derive the necessary and sufficient condition for mean-square stabilization of the scalar system in terms of the stability of a Markov Jump Linear System (MJLS) [3] whose evolution depends on both the system’s unstable mode and the Markov rate process. In the case of a plant with unstable mode $\lambda$, we establish that stabilizing the system is equivalent to stabilizing the MJLS having state dynamics $\lambda/2^{R_k}$, i.e., $z_{k+1} = \lambda/2^{R_k} z_k$, and transition jumps given by the Markov rate process. Intuitively, this equivalent MJLS describes the evolution of the estimation error at the decoder, which at every time step $k$ increases by $\lambda$ because of the system dynamics and is reduced by $2^{R_k}$ because of the information sent across the channel. A tight condition for second-moment stability is then expressed in terms of the spectral radius of an augmented matrix describing the dynamics of the second moment of this MJLS.

A similar approach is followed to provide stability conditions for the case of vector systems. Necessary conditions are derived by proceeding in two steps. First, we assume that a “genie” helps the channel decoder by stabilizing a subset of the unstable states. Then, we relate the stability of the reduced vector system to the one of a scalar MJLS whose evolution depends on the remaining unstable modes. By considering all possible subsets of unstable modes, we obtain a family of conditions that relate the degree of instability of the system to the parameters governing the rate process.

On the other hand, a sufficient condition for mean-square stability is given using a control scheme in which each unstable component of the system is quantized using a separate scalar quantizer. A bit allocation function determines how the bits available for communication over the Markov feedback channel are distributed among the various unstable sub-systems. Given a bit allocation function, the sufficient condition is then given as the intersection of the stability
conditions for the scalar jump linear systems that describe the evolution of the estimation error for each unstable mode. This control scheme yields optimal performance only in some special cases, but it can be easily analyzed using our result for scalar systems. Similar schemes have been presented in the related literature [17], [27], however our scheme is more general than these because it allows the bit allocation function to change over time. In this way we can strictly improve upon the sufficient condition obtained in [27] for the special case of a Markov erasure channel.

We now wish to spend some additional words on the related literature. Recent surveys on the theory of control with communication constraints appear in [8] and [19]. Broadly speaking, authors have followed two distinct approaches. The information-theoretic approach followed by the present paper and by [1], [4], [16], [18], [11], [26], [27], [28], aims to derive data-rate theorems quantifying how much rate is needed to construct a stabilizing quantizer/controller pair. The network-theoretic approach followed by the works [6], [7], [21], [24] aims to determine the critical packet dropout probability above which the system cannot be stabilized by any control scheme. In this context, a packet models a real number, carrying an unbounded amount of information in its binary expansion of infinite precision. The work [17] created a bridge between the two approaches, as it recovers results of the packet loss model by assuming the rate $R_k$ takes values 0 or $r$, and by letting $r \to \infty$. The same bridge also holds for our work that generalizes [17]. Indeed, when $R_k$ is a Markov chain with state space $\{0, r\}$, we recover the results in [7] and [10] in the limit as $r \to \infty$.

The theory of MJLS has been previously used to investigate control problems over communication channels. Seiler and Sengupta [22] studied the mean-square stabilization of a dynamical system with random delays and used tools from MJLS to derive a linear matrix inequality condition for the existence of a stabilizing controller. A similar approach is taken by Kawka.
and Alleyne in [13] to identify the most efficient estimation strategy to compensate for losses when controlling a system over a Markov erasure channel. Liu et al. [15] used control theoretic techniques to design capacity achieving codes for the communication problem over a finite-state Markov channel with feedback, by studying the second moment stability of a MJLS with output feedback and perfect knowledge of the Markov state. Sahai et al. [20] studied anytime communication over a two-state Markov feedback channel using a linear encoder modeled with a MJLS.

The rest of the paper is organized as follows. Section II provides a description of the problem and introduces the necessary and sufficient condition for the stability of a MJLS. In Section III a tight condition for second-moment stability condition of scalar systems is stated and proved. Section IV is devoted to the multi-dimensional case, for which necessary and sufficient conditions are provided.

II. Problem formulation

Consider the linear dynamical system

\[ x_{k+1} = Ax_k + Bu_k + v_k, \quad y_k = Cx_k + w_k, \quad k \in \mathbb{N} \]  

where \( x_k \in \mathbb{R}^d \) represents the state variable of the system, \( u_k \in \mathbb{R}^m \) the control input, \( v_k \in \mathbb{R}^d \) an additive disturbance independent of the initial condition \( x_0 \), \( y_k \in \mathbb{R}^p \) the sensor measurement and \( w_k \in \mathbb{R}^p \) the measurement noise, that is supposed to be independent of all other random variables. It is assumed that \( A \) is uniquely composed by unstable modes, so the open loop system is unstable. We assume that the distributions of the initial condition of the system and of the disturbances can have unbounded support, but we do assume the following conditions:

A0. \((A,B)\) is reachable and \((A,C)\) is observable.

A1. \( \exists \epsilon > 0 \) such that \( x_0, v_k \) and \( w_k \) have uniformly bounded \((2 + \epsilon)\)th absolute moments for all \( k \).

A2. The distribution of the noise is such that \( e^{2h(v_k)/d} > 0 \) for all \( k \) and \( v_k \in \mathbb{R}^d \).

The state observer is connected to the actuator through a noiseless digital communication link that at each time \( k \) allows transmission without errors of \( R_k \) bits, see Fig. 1. The rate process \( \{R_k\}_{k \geq 0} \) is a homogeneous positive-recurrent Markov chain that takes values in a finite subset
of the nonnegative integers

\[ \mathcal{R} = \{ r_1, \ldots, r_n \}, \]

and whose evolution through one time step is described by the transition probabilities

\[ p_{ij} = P\{ R_{k+1} = r_j | R_k = r_i \} \]

for all \( k \in \mathbb{N} \) and \( i, j \in \{1, \ldots, n\} \). The rate process is independent of the other quantities describing the system and is causally known at observer and controller. For every \( k \geq 0 \), the coder is a function \( s_k = s_k(y_0, \ldots, y_k) \) that maps all past and present measurements into the set \( \{1, \ldots, 2^{R_k}\} \), the digital link is the identity function on the set \( \{1, \ldots, 2^{R_k}\} \), and the control \( u_k \) is a function \( \hat{x}_k(s_0, \ldots, s_k) \) that maps all past and present symbols sent over the digital link into \( \mathbb{R}^d \).

Given this system, the objective is to find conditions on \( A \) and on the rate process \( \{R_k\}_{k \geq 0} \) under which it is possible to design a control scheme, i.e., a sequence of coder/control functions, such that the closed loop system is mean-square stable, that is,

\[ \sup_k \mathbb{E} \left[ \| x_k \|^2 \right] < \infty, \tag{2} \]

where the expectation is taken with respect to the random initial conditions, the additive disturbance \( v_k \), and the rate process \( R_k \), and \( \| \cdot \|^2 \) denotes the \( L^2 \)-norm.

A. Markov Jump Linear Systems.

Consider the scalar non-homogeneous MJLS defined by

\[ z_{k+1} = \frac{\lambda}{2^{R_k}} z_k + c, \tag{3} \]

where \( z_k \in \mathbb{R} \) with \( z_0 < \infty \), \( c > 0 \) is a constant, \( \{R_k\}_{k \geq 0} \) is the Markov rate process described above. Let \( H \) be the \( n \times n \) matrix with nonnegative real elements

\[ h_{ij} = \frac{1}{2^{r_j}} p_{ji}, \tag{4} \]

for all \( i, j \in \{1, \ldots, n\} \). The following (known) lemma states the necessary and sufficient condition for the mean-square stability of the system (3) in terms of the unstable mode \( |\lambda| \) and
the spectral radius \( \rho(\cdot) \) of \( H \). The spectral radius of a matrix is the maximum among the absolute values of its eigenvalues.

**Lemma 1:** Necessary and sufficient condition for the mean-square stability of the system (3) is that
\[
|\lambda|^2 < \frac{1}{\rho(H)}.
\]

**Proof:** The claim follows from [3, Theorem 3.9] and [3, Theorem 3.33], but we provide in Appendix A a proof for our specific setting which highlights the role played by the augmented matrix \( H \) in determining the dynamics of the second moment of \( \{z_k\} \).

The above claim is sufficient for our purposes, but in general a stronger results holds. The above condition can be shown to be equivalent to the existence of a unique solution to a set of coupled Lyapunov equations and it guarantees not only that the system (3) is mean-square stable but also that the first two moments of the system converge in \( L^2 \) to steady state values independent of the initial conditions [3].

III. Scalar system

Consider the special case of a scalar system
\[
x_{k+1} = \lambda x_k + u_k + v_k, \quad y_k = x_k + w_k, \quad \forall k \in \mathbb{N}.
\]

(5)

where \( |\lambda| \geq 1 \).

**Theorem 1:** Under assumptions A0-A3, there exists a control that stabilizes the scalar system (5) in mean-square sense if and only if the MJLS (3) is mean-square stable, that is, if and only if
\[
|\lambda|^2 < \frac{1}{\rho(H)}.
\]

(6)

**Remark 1:** In the above condition the unstable mode of the system and the channel properties are decoupled. This means that for a given Markov rate process there exists a threshold above which the system cannot be stabilized by any control scheme.

Application of Theorem 1 yields the following results as special cases.
a) **Constant rate.** When the channel supports a constant rate $r$, we have that $H = 1/2^{2r}$ and thus (6) reduces to the well known data rate theorem condition [18], [25],

$$r > \log |\lambda|.$$  

b) **Independent rate process.** Consider the special case of an i.i.d. rate process $R_k$ where $R_k \sim R$ has probability mass function $p_i = P\{R = r_i\}, r_i \in \mathcal{R}$. In this case

$$H = (p_1, \ldots, p_n)^T (2^{-2r_1}, \ldots, 2^{-2r_n}) = ph^T$$

is a rank-one matrix whose only nonzero eigenvalue is $h^T p$. Therefore, Theorem 1 yields the result in [16], [17],

$$|\lambda|^2 \rho(H) = |\lambda|^2 (2^{-2r_1}, \ldots, 2^{-2r_n}) (p_1, \ldots, p_n)^T$$

$$= E \left[ \frac{|\lambda|^2}{2^R} \right] < 1.$$  

If we further specialize to the case $n = 2$, $r_1 = 0$, $r_2 = r$, and we let $r \to \infty$, then the stability condition $p_1 > 1/|\lambda|^2$ depends only on the erasure rate of the channel, i.e. we recover the packet loss model result in [6].

c) **Two-state Markov process.** In the special case illustrated in Figure 2 in which $n = 2$, $p_{12} = q$, and $p_{21} = p$ for some $0 < p, q < 1$, we have

$$H = \begin{pmatrix} \frac{1}{2^{2r_1}}(1 - q) & \frac{1}{2^{2r_2}}p \\ \frac{1}{2^{2r_1}}q & \frac{1}{2^{2r_2}}(1 - p) \end{pmatrix}$$

(7)

and the condition in Theorem 1 reduces to

$$|\lambda|^2 \rho(H) = \frac{|\lambda|^2}{2} \text{tr}(H) + \frac{|\lambda|^2}{2} \sqrt{\text{tr}(H)^2 - 4\det(H)} < 1,$$

(8)

where $\text{tr}(H)$ and $\det(H)$ denote the trace and determinant of $H$, respectively. The special case where $r_1 = 0$ and $r_2 = r$ was previously studied in [27] where, by following a different approach, the authors proved that necessary and sufficient condition for stabilization is that

$$E \left[ \frac{|\lambda|^{2r}}{2^{2r}} \right] < 1.$$  

(9)
Here $\tau$ denotes the “hitting time” of state $r$, i.e., the time between two consecutive visits of that state, and the expectation is taken with respect to $\tau$. Intuitively, this condition says that $r$ should be large enough to compensate for the expansion of the state during the time in which packets are erased and thus no information can be sent from the observer to the controller. It is possible to show that condition (9) simplifies to

$$|\lambda|^2 < \begin{cases} \frac{1}{\nu(H)} & \text{if } \det(H) = 0, \\ \frac{\text{tr}(H) - \sqrt{\text{tr}(H)^2 - 4\det(H)}}{2\det(H)} & \text{otherwise,} \end{cases}$$

where $H$ is given by (7) with $r_1 = 0$ and $r_2 = r$. On the other hand, observe that

$$\frac{\text{tr}(H) - \sqrt{\text{tr}(H)^2 - 4\det(H)}}{2\det(H)} = \frac{2}{\text{tr}(H) + \sqrt{\text{tr}(H)^2 - 4\det(H)}},$$

which means that condition (8) is equivalent to (9), and thus when specialized to the case where $r_1 = 0$ and $r_2 = r$ our result recovers the one in [27].

Taking the limit $r \to \infty$, the above inequalities simplifies to the condition $(1 - q)|\lambda|^2 < 1$, so the stability condition is only determined by the recovery rate $q$, i.e. we recover the packet loss model result in [7].

A. Necessity

The following lemma shows that the second moment of the state in (5) is lower bounded by the second moment of a homogeneous MJLS that evolves as (3).

**Lemma 2:** For every $k \geq 0$,

$$\mathbb{E}[|x_k|^2] \geq \frac{1}{2\pi} \mathbb{E}[|z_k|^2],$$

where $\{z_k\}$ is a homogeneous MJLS with dynamics $z_{k+1} = \lambda/2 R_k z_k$ and initial condition $z_0 = e^{h(x_0)}$.

**Proof:** See Appendix B.

It follows that the state cannot be stable if this MJLS is second moment unstable. Hence, the “only if” condition in Theorem 1 follows directly from Lemma 1.
B. Sufficiency

1) Noiseless systems with bounded initial condition: We first consider the special case of a fully observed, discrete time, unstable, scalar system defined by

\[ x_{k+1} = \lambda x_k + u_k, \quad k \geq 1, \]  

with initial condition \( |x_0| \leq M_0 \) for some bounded \( M_0 > 0 \). In this case, we have that

\[ |x_k| \leq z_k, \quad k \geq 0, \]  

where \( z_k \) is a homogeneous MJLS with dynamics

\[ z_{k+1} = \frac{|\lambda|}{2R_k} z_k, \quad k \geq 1 \]

and \( z_0 = M_0 \). To see this, consider the following inductive proof. By construction, \( |x_0| \leq M_0 = z_0 \). Assume that the claim holds for all times up to \( k \), so \( |x_k| \leq z_k \). Suppose that at time \( k \) the uncertainty set \([-z_k, z_k]\) is quantized using a \( R_k \)-bit uniform quantizer, and that the encoder communicates to the decoder the interval containing the state. Then, the decoder approximates the state by the centroid \( \hat{x}_k \) of this interval and sends to the plant the control input \( u_k = -\lambda \hat{x}_k \).

By construction \( |x_k - \hat{x}_k| \leq z_k/2R_k \), thus

\[ |x_{k+1}| = |\lambda||x_k - \hat{x}_k| \leq \frac{|\lambda|}{2R_k} z_k = z_{k+1}, \]

i.e., the claim holds at time \( k + 1 \) as well. It follows that \( x_k \) is stable if the MJLS \( z_k \) is second moment stable. Hence, the “if” condition in Theorem 1 follows from Lemma 1.

By combining (10) with (12) we can see that the second moment of the state can be upper and lower bounded by two homogeneous MJLSs that differ only in their initial conditions, and therefore share the same stability condition. What emerges from the analysis in this simple setting is that the MJLS \( \{z_k\}_{k \geq 0} \) that determines the stability of the system describes the evolution of the estimation error at the decoder, which at every time step \( k \) increases by \( \lambda \), because of the system dynamics, and is reduced by \( 2R_k \), because this is the best attainable accuracy of representation sending \( R_k \) bits across the channel.

In this specific setting, it is possible to show that condition (6) is the necessary and sufficient condition to ensure convergence of the system to zero, i.e. \( \lim_{k \to \infty} E[|x_k|^2] = 0 \), which is
stronger notion of stability than the one we have assumed in (2).

2) Unbounded noise and initial condition: Consider now the system (5) assuming that conditions A0-A3 hold. The main challenge in this setting is that the disturbance has unbounded support, so it is not possible to confine the state dynamics within a finite interval as in the case of noiseless systems. Consequently, the uniform quantizer used for stabilizing noiseless systems has to be replaced by a dynamic quantizer that follows a zoom-in zoom-out strategy [2], [14], where the range of the quantizer is increased (zoom-out phase) when atypically large disturbances affect the system, and decreased as the state reduces its size (zoom-in phase). In this section we build a stabilizing control scheme based on the adaptive quantizer defined in [18].

The main elements of our construction are as follows. We assume that coder and decoder share at each time $k$ an estimate of the state $\hat{x}_k$ that is recursively updated using the information sent through the channel. Time is divided into cycles of fixed duration $\tau$. In each cycle the coder sends information about the state of the system at the beginning of the cycle. At the end of each cycle the decoder updates the estimate of the state and sends a control signal to the plant. The key step in the analysis is to show that the dynamics of the mean-square estimation error (and thus of the state) can be bounded by a non-homogeneous MJLS, so the stability of this MJLS implies the stability of the state. Specifically, we prove that at times $\tau, 2\tau, 3\tau, \ldots$, i.e., at the beginning of each cycle, the second moment of the estimation error $x_{j\tau} - \hat{x}_{j\tau}$ satisfies

$$E[|x_{j\tau} - \hat{x}_{j\tau}|^2] \leq E[z_{j\tau}^2] + \text{const}, \quad j \geq 0,$$

where const is a suitable constant, and $z_{j\tau}$ is a process formed recursively as

$$z_{j\tau} = \phi^{|\lambda|\tau/2R_{(j-1)\tau} + \ldots + R_{\tau-1}} z_{(j-1)\tau} + \varsigma, \quad j \geq 1,$$

for some bounded initial condition $z_0$ and constants $\phi > 1, \varsigma > 0$. Notice that equation (13) represents the evolution at times $\tau, 2\tau, \ldots$ of a non-homogeneous MJLS with state evolution

$$z_k = \phi^{1/|\lambda|/2R_k} z_{k-1} + c, \quad k \geq 1,$$

for some appropriately chosen constant $c$. Thus, by Lemma 1 sufficient condition for the second moment stability of $\{z_k\}$ is that

$$\phi^{2/|\lambda|^2} \rho(H) < 1.$$
On the other hand, if the condition of Theorem 1 is satisfied, that is, if $|\lambda|^2 \rho(H) < 1$, then we can choose the duration of a cycle $\tau$ large enough to ensure that (14) holds and, as a consequence, (13) is stable, the second moment of the estimation error at the beginning of each cycle is bounded, and the state remains second-moment bounded.

In the remainder of this section we provide a detailed description of the control scheme and its analysis.

**Adaptive quantizer.** An $r$-bit quantizer is a function $q_r$ that maps the real line into a discrete set with cardinality $2^r$ and thus its inverse image defines a partition of the real line into $2^r$ disjoint intervals. For $r \geq 2$, the quantizer proposed in [18] induces the following partitions of the real line

- the set $[-1, 1]$ is divided into $2^{r-1}$ intervals of the same length,
- the sets $(\delta^{i-2}, \delta^{i-1}]$ and $(-\delta^{i-1}, -\delta^{i-2}]$ are divided into $2^{r-1-i}$ intervals of the same length, for each $i \in \{2, \ldots, r-1\}$,
- the leftmost and rightmost intervals are the semi-open sets $(-\infty, -\delta^{r-1}]$ and $(\delta^{r-1}, \infty)$.

Here $\delta > 1$ is a parameter that determines the concentration of intervals around the origin. We can see that the width of the quantization regions increases with $\delta$, so the partition becomes more spread out. Given a real number $x$, the output value of the quantizer $q_r(x)$ is the midpoint of the interval in the partition containing $x$. In the sequel we will also make use of the function $\kappa_r(x)$, which instead returns the half-length of such interval. If $x$ is in one of the two semi-open sets at the two extremes of the partition, then we set $q_r(x) = \text{sign}(x) \delta^r$ and $\kappa_r(x) = \delta^r - \delta^{r-1}$.

Observe that the partition of the $r$-bit quantizer can be obtained recursively from the one of the $(r-1)$-bit quantizer by dividing each bounded interval into two intervals of the same length and the two semi-open intervals $(\pm \infty, \pm \delta^{r-2}]$ into $(\pm \infty, \pm \delta^{r-1}]$ and $(\pm \delta^{r-1}, \pm \delta^{r-2}]$. Thus, $q_{r+r'}(x)$ can be computed recursively starting from $q_r(x)$ by repeating the above procedure $r'$ times. We will make use of this property in our control scheme, where we use the fact that if coder and decoder know $q_{R_k}(x)$ at time $k$, then the coder can communicate to the decoder $q_{R_k+R_{k+1}}(x)$ by sending $R_{k+1}$ bits at time $k+1$.

**Encoder.** At every time $k$, the encoder keeps two different estimators for the state $x_k$. The first is denoted by $\bar{x}_k$ and is the Kalman filter estimator of the state given the observations $y_0, \ldots, y_k$. 

DRAFT
This is computed recursively as

$$\bar{x}_{k+1} = \lambda \bar{x}_k + u_k + \varepsilon_k,$$  

(15)

where $\varepsilon_k$ is the product of the innovation $(y_k - \bar{x}_k)$ and the optimal Kalman filter gain, and by assumption A2 has bounded $(2 + \epsilon)$th absolute moment. The second estimator is denoted by $\hat{x}_k$ and is updated only based on the information sent across the digital link and therefore is available at both coder and decoder. The error between the two estimators $\bar{x}_k - \hat{x}_k$ is recursively encoded using the quantizer described above.

More precisely, time is divided into cycles of length $\tau$. At the beginning of the $j$th cycle, i.e., at time $j\tau$, the coder computes

$$q_{R_{j\tau}}\left((\bar{x}_{j\tau} - \hat{x}_{j\tau})/\xi_j\right),$$

(16)

where $\xi_j$ is a scaling factor updated at the beginning of each cycle, and communicates to the decoder the index $s_{j\tau} \in \{1, \ldots, 2^{R_{j\tau}}\}$ of the quantization interval containing the scaled estimation error. At time $j\tau + 1$, the rate $R_{j\tau+1}$ becomes available at coder and decoder, that divide the quantization interval into $2^{R_{j\tau+1}}$ subintervals according to the recursive procedure described above. The coder sets $s_{j\tau+1} \in \{1, \ldots, 2^{R_{j\tau+1}}\}$ equal to the subinterval containing $(\bar{x}_{j\tau} - \hat{x}_{j\tau})/\xi_j$, so the decoder can compute

$$q_{R_{j\tau}+R_{j\tau+1}}\left((\bar{x}_{j\tau} - \hat{x}_{j\tau})/\xi_j\right).$$

By repeating the same procedure for the rest of the cycle, at time $(j+1)\tau - 1$ the decoder knows $(\bar{x}_{j\tau} - \hat{x}_{j\tau})/\xi_j$ at the resolution provided by a quantizer with

$$R(j) = R_{j\tau} + \cdots + R_{(j+1)\tau-1}$$

levels. Before the beginning of the next cycle, the encoder updates the state estimator according to

$$\hat{x}_{(j+1)\tau} = \lambda^\tau \left(\hat{x}_{j\tau} + \xi_j q_{R(j)} \left(\frac{\bar{x}_{j\tau} - \hat{x}_{j\tau}}{\xi_j}\right) + \sum_{k=j\tau}^{(j+1)\tau-1} \lambda^{(j+1)\tau-1-k} K \hat{x}_k\right),$$

(17)

where

$$\hat{x}_{k+1} = (\lambda + K) \hat{x}_k, \text{ for } k = j\tau, \ldots, (j + 1)\tau - 2,$$  

(18)
and $K$ is the certainty-equivalent control such that $|\lambda + K| < 1$. The scaling factor used in (16) is updated according to

$$
\xi_{j+1} = \max\{\varphi, |\lambda|^t \xi_j \kappa_{R(j)} \left( \frac{\bar{x}_{j+1} - \hat{x}_{j+1}}{\xi_j} \right) \},
$$

(19)

with $l_0 = \varphi$, where $\varphi$ is a constant independent of the state.

**Decoder.** Encoder and decoder are synchronized and share the state estimator $\hat{x}_k$ and the scaling factor $\xi_j$. Synchronism is guaranteed because both encoder and decoder set $\hat{x}_0 = 0$ and $l_0 = \varphi$ and the channel is noiseless. At every time $k$ the decoder sends to the plant the control signal

$$
u_k = K \hat{x}_k,$$

(20)

where $\hat{x}_k$ is updated as in (17) at the beginning of each cycle and as in (18) within each cycle. 

**Analysis.** First, we prove that if (6) holds, then the second moment of the estimation error $x_{j\tau} - \hat{x}_{j\tau}$ at the beginning of each cycle is bounded.

**Lemma 3:** For all $j \geq 0$

$$
E[|\bar{x}_{j\tau} - \hat{x}_{j\tau}|^2] \leq E[z_{j\tau}^2],
$$

where $z_{j\tau}$ satisfies (13) for some bounded initial condition $z_0$ and constants $\phi > 1$ and $\varsigma > 0$. Furthermore, (6) is a sufficient condition to ensure that $\{z_{j\tau}\}$ is second moment stable and that

$$
sup_j E[|x_{j\tau} - \hat{x}_{j\tau}|^2] < \infty.
$$

**Proof:** See Appendix C

Next, by Minkowski inequality,

$$
(E[|\bar{x}_{j\tau+k} - \hat{x}_{j\tau+k}|^2])^{1/2} \leq |\lambda|^k (E[|\bar{x}_{j\tau} - \hat{x}_{j\tau}|^2])^{1/2} + \left( E\left[ \sum_{i=0}^{k-1} \lambda^{k-1-i} K z_{j\tau+k}^2 \right] \right)^{1/2}
$$

for all $k = 0, 1, \ldots, \tau - 1$. Since by assumption A2 $z_k$ has bounded second moment, it follows from Lemma 3 that the second moment of $|\bar{x}_{j\tau+k} - \hat{x}_{j\tau+k}|^2$ is bounded if (6) holds. Finally, rewriting (15) as

$$
\bar{x}_{k+1} = (\lambda + K) \bar{x}_k - K (\bar{x}_k - \hat{x}_k) + z_k, \ k \geq 0
$$

we conclude that if (6) holds, then $sup_k E[\bar{x}_k^2] < \infty$ and thus, from the orthogonality principle, the system (5) is stable in the mean-square sense (2).
IV. VECTOR SYSTEM

Consider the system (1), where $A \in \mathbb{R}^{d \times d}$ is uniquely composed by unstable modes. Let $\lambda_1, \ldots, \lambda_u \in \mathbb{C}$ be the distinct, non-conjugate eigenvalues of $A$ (if $\lambda_i, \lambda_i^*$ are complex conjugate eigenvalues, only one of them is considered). Let $\mu_i$ be the algebraic multiplicity of $\lambda_i$ and let $a_i = 1$ if $\lambda_i \in \mathbb{R}$ and $a_i = 2$ otherwise, such that $\sum_{i=1}^u a_i \mu_i = d$.

In order to decompose its dynamical modes, we express $A$ into real Jordan canonical form $J$ [9]. The matrix $J \in \mathbb{R}^{d \times d}$ is such that $A = T^{-1}JT$ for some similarity matrix $T$, and has the block diagonal structure $J = \text{diag}(J_1, \ldots, J_u)$ where $J_i \in \mathbb{R}^{a_i \mu_i \times a_i \mu_i}$ and $\det J_i = \lambda_i^{a_i \mu_i}$. Let $U = \{1, \ldots, u\}$ denote the index set of the subsystems corresponding to the modes $\lambda_1, \ldots, \lambda_u$.

Consider the recursion

$$ x_{k+1} = Jx_k + TBu_k + Tv_k, \quad y_k = CT^{-1}x_k + w_k. \quad (21) $$

As (21) and (1) are related only by the matrix $T$, we assume that the system evolves according to (21).

A. Necessity

To find necessary conditions for stability, we make the optimistic assumption that a “genie” helps the channel decoder by stabilizing part of the system, so that the information sent across the channel is only used to stabilize a subset of the unstable states. Formally, let $\nu_i \in \{0, \ldots, \mu_i\}$ denote the algebraic dimensionality of the $i$th unstable mode after the genie’s intervention, and let $\mathcal{V} = \{\nu : \nu_i \in \{0, \ldots, \mu_i\}\} \subset \mathbb{N}^u$ denote the space of all possible algebraic dimensionalities. For each $\nu \in \mathcal{V}$, let $x_k(\nu)$ denote the unstable system obtained by removing from (21) the components that have been stabilized by the genie. Observe that $x_k(\nu)$ is real valued and has total dimension

$$ d(\nu) = \sum_{i \in U} a_i \nu_i. \quad (22) $$

Next, we find necessary conditions for the second-moment stability of the reduced system $x_k(\nu)$. The following lemma shows that the second moment $x_k(\nu)$ is lower bounded by the second moment of a scalar homogeneous MJLS whose evolution depends on the unstable modes of $x_k(\nu)$ and their dimensionalities.
Lemma 4: For each \( \nu \in \mathcal{V} \), let \( z_k(\nu) \) be the scalar MJLS defined by

\[
z_{k+1}(\nu) = \left( \frac{\prod_{i \in \mathcal{U}} |\lambda_i|^{a_i \mu_i}}{2R_k} \right)^{1/d(\nu)} z_k(\nu),
\]

with \( z_0(\nu) = 2^{\frac{1}{d(\nu)} h(x_0(\nu))} \). Then, for every \( k \geq 0 \),

\[
E[\|x_k(\nu)\|_2^2] \geq E[|z_k(\nu)|^2].
\]

The proof of the above lemma is omitted since it closely follows the proof of Lemma 2 combined with standard information-theoretic techniques. It follows that \( x_k(\nu) \) cannot be stable if the MJLS \( \{z_k(\nu)\} \) is second moment unstable. Then, we can derive a set of necessary conditions in terms of the spectral radius of the matrix describing the dynamics of the second moment of \( z_k(\nu) \). Let \( H(\nu) \) be the \( n \times n \) matrix with nonnegative real elements

\[
h_{ij}(\nu) = \frac{1}{2^{\frac{1}{d(\nu)} r_j}} p_{ji}
\]

for all \( i, j \in \{1, \ldots, n\} \). Combining Lemma 1 with Lemma 4 it then follows that

Theorem 2: Under assumptions A0-A3, necessary condition for the stabilization of the system (21) in the mean-square sense (2) is that, for every \( \nu \in \mathcal{V} \),

\[
|\lambda(\nu)|^{2/d(\nu)} < \frac{1}{\rho(H(\nu))}
\]

where \( \lambda(\nu) = \prod_{i \in \mathcal{U}} |\lambda_i|^{a_i \mu_i} \), \( d(\nu) \) is as in (22), and \( H(\nu) \) is given by (23).

Remark 2: Application of Theorem 2 yields the following results as special cases.

a) Scalar systems. In the special case of a scalar system, we have that \( \lambda(\nu) = \lambda \), \( d(\nu) = 1 \), and \( H(\nu) \) is equal to (4), thus the above condition reduces to the necessary condition in Theorem 1.

b) Constant rate. When the channel supports a constant rate \( r \), we have that \( H(\nu) = 2^{\frac{1}{\rho(\nu)}} r \) and thus (24) reduces to the well known data rate theorem condition [18], [25],

\[
r > \sum_{i \in \mathcal{U}} a_i \mu_i \log |\lambda_i|.
\]

b) Independent rate process. In the special case of an i.i.d. rate process \( R_k \) where \( R_k \sim R \).
has probability mass function \( p_i = P\{R = r_i\}, r_i \in \mathcal{R} \)

\[
\rho(H(\nu)) = \rho\left((p_1, \ldots, p_n)^T (2^{-\frac{2}{\nu_1} r_1}, \ldots, 2^{-\frac{2}{\nu_n} r_n})\right) \\
= (2^{-\frac{2}{\nu_1} r_1}, \ldots, 2^{-\frac{2}{\nu_n} r_n}) (p_1, \ldots, p_n)^T \\
= E\left[2^{-\frac{2}{\nu} R}\right],
\]

and (24) simplifies to

\[
\sum_{i \in U} a_i \nu_i \log |\lambda_i| < -\frac{d(\nu)}{2} \log E\left[2^{-\frac{2}{\nu} R}\right], \tag{25}
\]

and our condition reduces to the one in [17]. It can be shown that in this case it suffices to choose \( \nu_i \in \{0, \mu_i\} \) as the remaining inequalities are redundant. If we further specialize to the case \( n = 2, r_1 = 0, r_2 = r \), and we let \( r \to \infty \), then the stability condition

\[
\max_{i \in U} |\lambda_i|^2 < \frac{1}{p_1}
\]

depends only on the erasure rate of the channel.

c) Two-state Markov process. In the special case\(^1\) illustrated in Fig. 2, (24) simplifies to

\[
\log |\lambda(\nu)| < -\frac{d(\nu)}{2} \log \left(\frac{1}{2} \text{tr}(H(\nu)) + \frac{1}{2} \sqrt{\text{tr}(H(\nu))^2 - 4 \det(H(\nu))}\right),
\]

where

\[
\text{tr}(H(\nu)) = (1-q)2^{-\frac{2}{\nu_1} r_1} + (1-p)2^{-\frac{2}{\nu_2} r_2}
\]

and

\[
\det(H(\nu)) = (1-q-p)2^{-\frac{2}{\nu}(r_1+r_2)}.
\]

It can be shown that when specialized to the case where \( r_1 = 0 \) and \( r_2 = r \) our condition reduces to the one derived in [27]. Also, it can be easily seen that if \( r_1 = 0 \), then as \( r_2 = r \to \infty \) we have that \( \det(H(\nu)) \to 0 \) and \( \text{tr}(H(\nu)) \to 1 - q \), so the necessary condition reduces to

\[
\max_{i \in U} |\lambda_i|^2 < \frac{1}{(1-q)}.
\]

\(^1\)Here we assume \( p \neq 1 - q \), otherwise the rate process becomes i.i.d. and the necessary condition simplifies to (25).
B. Sufficiency

A sufficient condition for mean-square stability is given extending the control scheme described in [17]. The main challenge in studying vector systems is designing a vector quantizer that dynamically adapts to the time-varying communication rate. The solution adopted in [17] for the case of an i.i.d. rate process is to quantize the state using one scalar quantizer per each state component, so the performance can be easily analyzed using the result for scalar systems. In this paper, we follow a similar method of proof and assume that at each time step the bits available for transmission are distributed among the various unstable modes of the system. Given a Markov channel and a system matrix $J$, let $A$ denote the set of functions $\alpha : \mathcal{R} \rightarrow [0, 1]^u$ having the properties that

$$\frac{1}{\alpha_i(r)}r \in \mathbb{N}, \quad i = 1, \ldots, u,$$

and

$$\alpha_1(r) + \cdots + \alpha_u(r) \leq 1$$

for all $r \in \mathcal{R}$. An element $\alpha \in A$ is referred to as a bit allocation function.

The operational meaning of $\alpha \in A$ is as follows. At every time $k$, $\frac{1}{\alpha_i(R_k)}R_k$ bits are used to quantize each scalar component of the $i$th unstable state of the system. Condition (26) enforces that each subsystem is quantized using an integer number of bits, while (27) ensures that we do not use more bits than what is available for transmission. From the result on scalar systems in Section III, the $i$th subsystem is stable if the MJLS with dynamics

$$z_{k\tau} = \phi \frac{|\lambda_i|^\tau}{2^\alpha_i(R_k)\sum_{j=(k-1)r}^{k\tau-1} \alpha_j(R_j)R_j} z_{(k-1)r} + \varsigma,$$

for some bounded initial condition $z_0$, and constants $\phi > 1$ and $\varsigma > 0$, is stable. Notice that the exponent at denominator in the above equation contains the total number of bits that are allocated to each component of the $i$th unstable mode during the $k$th cycle. Then, paralleling the analysis for the scalar system, we can define $H_i(\alpha), i = 1, \ldots, u$, as the $n \times n$ real matrix with $(l, j)$th entry equal to

$$2^{-\alpha_i(R_i)}p_{jl},$$

for different allocations to different components of the same sub-system.

\footnote{We restrict to this case for ease of presentation, but it is straightforward to generalize the definition to allow for different allocations to different components of the same sub-system.}
and conclude that, given a bit allocation function $\alpha$, sufficient condition for second moment stability of the system is that
\[ |\lambda_i|^2 < \frac{1}{\rho(H_i(\alpha))} \tag{29} \]
for every $i \in \mathcal{U}$. A more general sufficient condition can be given by allowing the bit allocation function to change over time.

**Theorem 3:** Under the assumptions A0-A3 above, a sufficient condition for the stabilization of the system (21) in the mean-square sense is that
\[ |\lambda_i|^2 < \frac{1}{\rho(H_i(\alpha_1)H_i(\alpha_2)\cdots H_i(\alpha_m))^{\frac{1}{m}}} \tag{30} \]
for some $m \geq 1$ and $\alpha_1, \ldots, \alpha_m \in \mathcal{A}$.

**Proof:** Fix $m \geq 1$ and $\alpha_1, \alpha_2, \ldots, \alpha_m \in \mathcal{A}$, non necessarily distinct. The main idea of the proof is to use these bit allocations functions in a cyclic repeated order via a time-sharing protocol. As in the proof of Lemma 3, divide time into cycles of length $\tau$ and further divide each cycle into $q$ blocks of length $m$ each, such that $\tau = qm$. Within each cycle, use the bit allocation function $\alpha_j$ in the $j$th channel use of each block, $j = 1, \ldots, m$. Let
\[ R(k) = \frac{1}{\alpha_i \mu_i} \sum_{l=0}^{q-1} \sum_{j=1}^{m} \alpha_{j,i}(R(k-1)\tau+lm+j-1)R(k-1)\tau+lm+j-1 \]
denote the total number of bits that are used to quantize each scalar component of the $i$th unstable mode during the $k$th cycle. Then, following similar steps as in the proof of Lemma 3, we can show that the $i$th sub-system is stable if the non-homogeneous MJLS with dynamics
\[ z_{k\tau} = \phi \frac{|\lambda_i|^\tau}{2R(k)} z_{(k-1)\tau} + \varsigma, \]
for some bounded initial condition $z_0$, and constants $\phi > 1$, $\varsigma > 0$, is second-moment stable. Equation (31) represents the evolution at times $qm, 2qm, \ldots$ of the MJLS
\[ z_{lm} = \phi^{1/q} \frac{|\lambda_i|^m}{2\alpha_i \mu_i} \sum_{j=1}^{m} \alpha_{j,i}(R(k-1)\tau+j-1)R(k-1)\tau+j-1 z_{(l-1)m} + c, \quad l \geq 1. \tag{32} \]
for some suitable constant $c$. Thus, if condition (30) is satisfied, then we can choose the number of blocks $q$ within each cycle, and hence the duration of a cycle $\tau$, large enough to ensure that
the MJLS (31) is stable. The rest of the proof is as in the proof of Theorem 1, hence is omitted. ■

Given a Markov feedback rate process and a system matrix \( J \), Theorem 2 and Theorem 3 provide inner and outer bounds, respectively, to the set of unstable modes

\[
(\log |\lambda_1|, \ldots, \log |\lambda_u|)
\]

that can be stabilized in second moment sense by some control scheme. These bounds are amenable to the following geometric interpretation. On the one hand, by taking the \( \log \) of both sides in (24), we can see that (24) defines an open half-space in the domain of the unstable modes for each \( \nu \in \mathcal{V} \). Thus, Theorem 2 states that it is not possible to stabilize systems whose unstable rates lie outside the intersection of all such open half-spaces. In general some of the inequalities in Theorem 2 are redundant, because implied by other inequalities, and in some special cases it is possible to find the minimal set of non-redundant inequalities, e.g., in the constant rate case discussed above. On the other hand, by taking the \( \log \) on both sides in (30), we can see that (30) defines an open hypercube in the domain of the unstable rates for each choice of \( m \) and of \( \alpha_1, \ldots, \alpha_m \in \mathcal{A} \). Then, Theorem 2 states that it is possible to stabilize all those systems whose unstable rates lie inside the union of all such open hypercubes. Observe that the resulting region is not computable because there is no upper bound on the value of \( m \). However, the next proposition shows that the region is convex.

**Proposition 4:** For every \( m \geq 1 \) and \( \alpha_1, \ldots, \alpha_m \in \mathcal{A} \), let \( \mathcal{C}(m, \alpha_1, \ldots, \alpha_m) \) denote the set of rates satisfying (30). Then, the closure of

\[
\bigcup_{m \geq 1} \bigcup_{\alpha_1, \ldots, \alpha_m \in \mathcal{A}} \mathcal{C}(m, \alpha_1, \ldots, \alpha_m)
\]

is a convex set.

**Proof:** See Appendix D. ■

It follows from the above proposition that the set of rates satisfying the condition in Theorem 2 can be approximated by fixing a bound \( M \) for \( m \) and taking the convex hull of the union of the hypercubes over all \( \alpha_1, \ldots, \alpha_m \in \mathcal{A} \) and \( m \leq M \). In particular, by setting \( M = 1 \) we obtain
that the inner bound contains the set of rates inside

$$\text{conv} \bigcup_{\alpha \in \mathcal{A}} \mathcal{C}(1, \alpha),$$

i.e., the convex hull of the union of the rates that can be stabilized using the same bit allocation function $\alpha \in \mathcal{A}$ at every instant of time. In general the inclusion is strict, as choosing $M > 1$ might yield a strictly larger region, but there are some special cases where the two sets are equal.

As an application of Theorem 3 and Proposition 4, consider the following special cases:

a) Independent rate process. In the special case of an i.i.d rate process, for every $m \geq 1$ and bit allocation functions $\alpha_1, \ldots, \alpha_m \in \mathcal{A}$, (30) simplifies to

$$\log |\lambda_i| < -\sum_{j=1}^{m} \frac{1}{2m} \log E \left[ 2^{-2\alpha_{j,i}(R)R} \right], \text{ for all } i \in \mathcal{U},$$

and therefore our condition recovers the sufficient condition given in [17]. It can be shown that in this case the set of rates satisfying Theorem 2 is exactly equal to (33).

b) Two-state Markov erasure process. Consider the Markov rate process illustrated in Figure 2 in the special case where $r_1 = 0$ and $r_2 = r$. For this problem, a sufficient condition for mean-square stabilization was previously given in [27, Theorem 3]. Restating their result in our notation, You and Xie proved that it is possible to stabilize all rates inside (33). We claim that Theorem 2 above yields a strictly larger inner bound to the set of stabilizable rates. To see this, consider the special case where $r = 6$ and the transition probabilities are $p_{12} = 0.1$ and $p_{21} = 0.7$. Let the system matrix $J$ have two distinct real eigenvalues $\lambda_1$ and $\lambda_2$ of multiplicity $\mu_1 = 2$ and $\mu_2 = 3$, respectively. Figure IV-B plots inner and outer bounds to the set of stabilizable rates $(\log |\lambda_1|, \log |\lambda_2|)$. The outer bound is delimited by the solid thick black line and is obtained by evaluating the bounds in Theorem 2. For the inner bounds, observe that in this setting there are 12 possible bit allocation functions, i.e., $|\mathcal{A}| = 12$, and therefore the inner bound in [27, Theorem 3] (delimited by the solid thin black line) is obtained taking the convex hull of 12 hypercubes. On the other hand, by evaluating Theorem 3 for $m \leq 8$ and every choice of $\alpha_1, \ldots, \alpha_m \in \mathcal{A}$ and by taking the convex hull of the union of the resulting hypercubes, we obtain the region enclosed by the dashed black line, which strictly includes the inner bound in [27, Theorem 3].
V. CONCLUDING REMARKS

Information and control theories have traditionally evolved as separate subjects. On the one hand, information theory almost religiously followed Shannon’s own words according to which: “the semantic aspects of communication are irrelevant to the engineering problem.” On the other hand, in control one is concerned with using information in a feedback loop to achieve some performance objective, and control theorists have assumed for decades that limitations in the communication links do not affect performance significantly. Emerging applications have shown that keeping such theoretical separation is not anymore desirable, nor possible. The engineer now aims to control one or more dynamical systems, using multiple sensors and actuators transmitting and receiving information over (wireless) digital communication links and bandwidth has become a scarce resource. In this scenario, the emerging area of control with limited data rates started incorporating ideas from both control and information theory to study the trade-offs arising from the desire to stabilize and the capability of actuating through communication. Our paper is placed in this framework. We studied the problem of mean-square stabilization of a linear, discrete-time, dynamical system over a time-varying, Markov digital feedback channel. We formulated a general version of the data rate theorem that extends previous formulations and provides a deeper understanding of the system’s dynamics over correlated channels. The result has been
obtained exploiting a reduction to MJLS and using the theory developed in this context, as well as some information-theoretic arguments. A missing point, that is still open for future research, is how results are modified in the presence of decoding errors, beside erasures.

**APPENDIX A**

**PROOF OF LEMMA 1**

For every $k \geq 0$, let $\mu_{k,i} = \mathbb{E} \left[ z_k 1_{\{R_k = r_i\}} \right]$ and $\sigma^2_{k,i} = \mathbb{E} \left[ z^2_k 1_{\{R_k = r_i\}} \right]$ denote the first and second moments of the MJLS $\{z_k\}$, respectively, in the event that the rate process is equal to $r_i$. Using the Markov property of the rate process $\{R_k\}$ and (3), it can be verified that the Markov chain $z_{k+1} \to R_k \to R_{k+1}$ holds and thus

$$\sigma^2_{k+1,j} = \sum_{i=1}^{n} \mathbb{E} \left[ |z_{k+1}|^2 | R_{k+1} = r_j, R_k = r_i \right] P\{R_k = r_i\} p_{ij}$$

$$= \sum_{i=1}^{n} \mathbb{E} \left[ \left( \frac{|\lambda|}{2} z_k + c \right)^2 \right] P\{R_k = r_i\} p_{ij}$$

$$= \sum_{i=1}^{n} \frac{|\lambda|^2}{2r_i} p_{ij} \sigma^2_{k,i} + 2c \sum_{i=1}^{n} \frac{|\lambda|}{2r_i} p_{ij} \mu_{k,i} + c^2 \sum_{i=1}^{n} p_{ij}.$$  

Similarly,

$$\mu_{k+1,j} = \sum_{i=1}^{n} \frac{|\lambda|}{2r_i} p_{ij} \mu_{k,i} + c \sum_{i=1}^{n} p_{ij}.$$  

We can see from the above equations that $\mu_k = (\mu_{k,1}, \ldots, \mu_{k,n})^T$ and $\sigma_k^2 = (\sigma^2_{k,1}, \ldots, \sigma^2_{k,n})^T$ evolve over time according to the following linear system:

$$\begin{bmatrix} \sigma^2_k \\ \mu_k \end{bmatrix}_{k+1} = \begin{bmatrix} |\lambda|^2 P D^2 & 2c |\lambda| P D \\ 0 & |\lambda| P D \end{bmatrix} \begin{bmatrix} \sigma^2_k \\ \mu_k \end{bmatrix}_k + \begin{bmatrix} c^2 p \\ c p \end{bmatrix},$$

(35)

where $P$ is the transition probability matrix of the Markov rate process, $D = \text{diag}(2^{-r_1}, \ldots, 2^{-r_n})$, and $p \in \mathbb{R}^n$ is vector whose $j$th element is equal to $\sum_{i=1}^{n} p_{ij}$. Now, by the law of total probability,

$$\mathbb{E} [z_k] = \sum_{i=1}^{n} \mu_{k,i}, \text{ and } \mathbb{E} [z^2_k] = \sum_{i=1}^{n} \sigma^2_{k,i}.$$  

It follows that the MJLS $\{z_k\}$ is second-moment stable if and only if the linear system (35) is second-moment stable. Observe that the eigenvalues of the block triangular matrix in (35) are
the union of the eigenvalues of the matrices $|\lambda|^2 P D^2$ and $|\lambda| P D$. The next lemma states that $\rho(|\lambda|^2 P D^2) < 1$ implies $\rho(|\lambda| P D) < 1$, so that $\rho(|\lambda|^2 P D^2) < 1$ is a necessary and sufficient condition for the second-moment stability of $\{z_k\}$.

**Lemma 5:** If $\rho(|\lambda|^2 P D^2) < 1$ then $\rho(|\lambda| P D) < 1$.

**Proof:** Assume that $\rho(|\lambda|^2 P D^2) < 1$. In order to show the claim, it suffices to consider the homogeneous system associated to the system (35), obtained by setting $c = 0$. Fix $j \in \{1, \ldots, n\}$ and assume $\mu_0 = \gamma e_j$, where $\gamma \in \mathbb{R}$ is a constant and $e_j$ is the column vector with the $j$th element equal to one and all other elements equal to zero. As a remark, $\mu_0 = \gamma e_j$ corresponds to the initial state of the rate process $R_0 = r_j$. Observe that $\mu_k$ evolves according to $\mu_k = (|\lambda| P D)^k \gamma e_j$ and that its 2-norm is upper bounded by

$$
\|((|\lambda| P D)^k \gamma e_j\|^2 = \sum_{i=1}^N |\mu_{k,i}|^2 = \sum_{i=1}^N \left(\mathbb{E}\left[z_k 1\{R_k = r_i\}\right]\right)^2
\leq \sum_{i=1}^N \mathbb{E}\left[|z_k|^2 1\{R_k = r_i\}\right] = \sum_{i=1}^N \sigma_{k,i}^2
= \|\sigma_k^2\|_1.
$$

For $R_0 = r_j$, we have that $\sigma_0^2 = \gamma' e_j$ for some $\gamma' \geq 0$, and that $\sigma_k^2$ evolves according to $\sigma_k^2 = (|\lambda|^2 P D^2)^k \gamma' e_j$. Therefore, if $\rho(|\lambda|^2 P D^2) < 1$ then $\|\sigma_k^2\|_1 \to 0$ and $\|(|\lambda| P D)^k \gamma e_j\| \to 0$ as $k \to \infty$, for all $j = 1, \ldots, n$ and $\gamma \in \mathbb{R}$. This implies that $\|((|\lambda| P D)^k \mu_0\| \to 0$ as $k \to \infty$ for every $\mu_0 \in \mathbb{R}^n$, which is equivalent to $\rho(|\lambda| P D) < 1$.

**APPENDIX B**

**PROOF OF LEMMA 2**

Let $S^k = \{s_0, \ldots, s_k\}$ denote the symbols transmitted over the digital link up to time $k$. By the maximum entropy theorem [5], the second moment of the state is lower bounded by the conditional entropy power of $x_k$, conditional on the event $S^{k-1} = s^{k-1}$

$$
\mathbb{E}\left[|x_k|^2\right] \geq \frac{1}{2\pi e} \mathbb{E}_{S^{k-1}} \left[e^{2h(x_k|S^{k-1} = s^{k-1})}\right]
$$

(36)
for all \( k \in \mathbb{N} \). We claim that the right hand side of the above inequality can be further lower bounded as follows

\[
E_{S^{k-1}} \left[ e^{2h(x_k|S^{k-1} = s^{k-1})} \right] \geq E \left[ \frac{|\lambda|^{2k}}{2^2(R_0 + \cdots + R_{k-1})} \right] e^{2h(x_0)},
\]  

(37)

for all \( k = 0, 1, \ldots \). We prove this claim by induction. For \( k = 0 \)

\[
E_{S_0} \left[ e^{2h(x_1|S_0 = s_0)} \right] = E_{S_0} \left[ e^{2h(\lambda x_0 + v_0|S_0 = s_0)} \right]
\geq |\lambda|^2 E_{S_0} \left[ e^{2h(x_0|S_0 = s_0)} \right] + e^{2h(v_0)}
\geq |\lambda|^2 E_{R_0} \left[ E_{S_0|R_0} \left[ e^{2h(x_0|S_0 = s_0)} \right] \right]
\geq |\lambda|^2 E_{R_0} \left[ E_{S_0|S_{-1}, R_0} \left[ e^{2h(x_0|S_0 = s_0)} \right] \right]
\geq E_{R_0} \left[ \frac{|\lambda|^2}{2^2R_0} \right] e^{2h(x_0)}
\]

(38)

where (38) follows from (5), the fact that the control input is a deterministic function of \( S^k \) and the shift invariance property of the conditional entropy, (39) is a consequence of the entropy power inequality [5], (40) makes use of assumption A2, i.e., \( e^{2h(v_k)} > 0 \) for all \( k \) and \( v_k \), and (41) follows from Lemma 1 in [17], which states that for every \( k \geq 0 \),

\[
E_{S^k|S^{k-1}, R_k} \left[ e^{2h(x_k|S^k = s^k)} \right] \geq \frac{1}{2^2R_k} e^{2h(x_k|S^{k-1} = s^{k-1})},
\]

(42)

with \( S_{-1} = \emptyset \). Thus, the claim holds for \( k = 0 \). Next, suppose by induction hypothesis that (10) holds up to \( k - 2 \). Then,

\[
E_{S^{k-1}} \left[ e^{2h(x_k|S^{k-1} = s^{k-1})} \right]
\geq |\lambda|^2 E_{S^{k-2}, R_{k-1}} \left[ E_{S^{k-1}|S^{k-2}, R_{k-1}} \left[ e^{2h(x_{k-1}|S^{k-1} = s^{k-1})} \right] \right]
\geq |\lambda|^2 E_{S^{k-2}, R_{k-1}} \left[ \frac{1}{2^2R_{k-1}} e^{2h(x_{k-1}|S^{k-2} = s^{k-2})} \right]
\geq |\lambda|^2 E_{R_{k-1}} \left[ \frac{1}{2^2R_{k-1}} E_{S^{k-2}|R_{k-1}} \left[ e^{2h(x_{k-1}|S^{k-2} = s^{k-2})} \right] \right]
\geq |\lambda|^2 E_{R_{k-1}} \left[ \frac{1}{2^2R_{k-1}} E_{S^{k-2}} \left[ e^{2h(x_{k-1}|S^{k-2} = s^{k-2})} \right] \right]
\geq E \left[ \frac{|\lambda|^{2k}}{2^2(R_0 + \cdots + R_{k-1})} \right] e^{2h(x_0)},
\]

(43)
where (43) follows from (42) and (44) from the induction hypothesis. It follows that the claim is true for time $k - 1$, hence the proof is complete.

Finally, observe that the right hand side of (37) is equal to the second moment of a MJLS $z_k$ that evolves as in (3), after setting $z_0 = e^{h(x_0)}$ and $c = 0$, and for which

$$E \left[ z_k^2 \right] = E \left[ \frac{\lambda^{2k}}{2^{2(R_0 + \cdots + R_{k-1})}} \right] e^{2h(x_0)}.$$  

Combining this with (36) and (37), the claim follows.

APPENDIX C
PROOF OF LEMMA 3

Let $e_k$ denote the error between the two estimators $\bar{x}$ and $\hat{x}$ at time $k$. Observe that, for any $\epsilon > 0$,

$$E[e_{j\tau}^2] = E[e_{j\tau}^2 (1_{\{|e_{j\tau}| > \xi_j\}} + 1_{\{|e_{j\tau}| \leq \xi_j\}})\leq E[e_{j\tau}^{2+\epsilon} \xi_j^{-\epsilon} + \xi_j^2] \triangleq \theta_j$$

(45)

where $\xi_j$ denotes the (non-negative) scaling factor (19) used at coder and controller. Thus, the second moment stability of the process $\{\theta_j\}$ implies the stability of the error at the beginning of each cycle. We investigate the stability conditions for $\{\theta_j\}$ by studying the dynamics of

$$\theta_{j,i} = E \left[ (\xi_j^2 + e_{j\tau}^{2+\epsilon} \xi_j^{-\epsilon}) 1_{\{R_{j\tau} = r_{i\tau}\}} \right].$$

for $i \in \mathcal{R}$.

First, observe that iterating (15) and making use of (20)

$$\bar{x}_{(j+1)\tau} = \lambda^{\tau-1} \bar{x}_{j\tau} + \sum_{k=j\tau}^{(j+1)\tau-1} \lambda^{(j+1)\tau-1-k} K \hat{x}_k + \eta_j,$$

where

$$\eta_j = \sum_{i=0}^{\tau-1} \lambda^{\tau-1-i} e_{j\tau+i}.$$  

(46)

has bounded $(2 + \epsilon)$th moment. Then, combining with (17) we obtain the following recursive
expression for the error
\[ e_{(j+1)\tau} = \bar{x}_{(j+1)\tau} - \hat{x}_{(j+1)\tau} \]
\[ = \lambda^\tau \left( e_{j\tau} - \xi_j q_{R(j)} \left( \frac{e_{j\tau}}{\xi_j} \right) \right) + \eta_j. \]  
(47)

From (47), the inequality \(|a| + |b| \leq 2^{1+\epsilon}(|a|^{2+\epsilon} + |b|^{2+\epsilon})\) yields
\[ e_{(j+1)\tau}^{2+\epsilon} \leq 2^{1+\epsilon} \left( |\lambda|^{\tau(2+\epsilon)} |e_{j\tau} - \xi_j q_{R(j)} \left( \frac{e_{j\tau}}{\xi_j} \right) |^{2+\epsilon} + \eta_j^{2+\epsilon} \right). \]

Next, suppose that the constant \(\varphi\) in (19) is chosen such that \(\varphi^{2+\epsilon}\) is an upper bound to the \((2 + \epsilon)\)th moment of (46). Then, multiplying the above inequality by \(\xi_j^{-\epsilon}\) and taking the expectation,
\[ E \left[ e_{(j+1)\tau}^{2+\epsilon} \xi_j^{-\epsilon} \right] \leq 2^{1+\epsilon} E \left[ \left( |\lambda|^{\tau(2+\epsilon)} |e_{j\tau} - \xi_j q_{R(j)} \left( \frac{e_{j\tau}}{\xi_j} \right) |^{2+\epsilon} + \eta_j^{2+\epsilon} \xi_j^{-\epsilon} \right) \right] \]
\[ \leq 2^{1+\epsilon} E \left[ \left( |\lambda|^{\tau(2+\epsilon)} |e_{j\tau} - \xi_j q_{R(j)} \left( \frac{e_{j\tau}}{\xi_j} \right) |^{2+\epsilon} \left( \xi_j |\lambda|^{\tau K_{R(j)} \left( \frac{e_{j\tau}}{\xi_j} \right)} \right)^{-\epsilon} + \varphi^2 \right) \right]. \]  
(48)

Similarly, by (19),
\[ E \left[ \xi_j^{2+\epsilon} \right] \leq 2^{1+\epsilon} E \left[ \left( \xi_j |\lambda|^{\tau K_{R(j)} \left( \frac{e_{j\tau}}{\xi_j} \right)} \right)^2 + \varphi^2 \right]. \]  
(49)

Adding (48) and (49), and multiplying both sides by \(1_{R(j+1)\tau = r_{j\tau}}\), we obtain the following
bound on the dynamics of $\theta_{j,i}$:

$$
\theta_{j+1,i} = E\left[\left(e_+^2 \xi_{j+1}^2 + \xi_{j+1}^2\right)1_{\{R_{j+1}=r_{i}\}}\right]
$$

$$
\leq 2^{1+\epsilon}|\lambda|^{2r}E\left[\left(|e_j^r - \xi_j^r q_{R(j)}\left(\frac{e_j^r}{\xi_j^r}\right)|^{2+\epsilon}\left(\xi_j^{K_{R(j)}}\left(\frac{e_j^r}{\xi_j^r}\right)\right)^{-\epsilon} + \left(\xi_j^{K_{R(j)}}\left(\frac{e_j^r}{\xi_j^r}\right)\right)^2\right)1_{\{R_{j+1}=r_{i}\}}\right]
$$

$$
+ 2^{1+\epsilon}\varphi^2,
$$

$$
= 2^{1+\epsilon}|\lambda|^{2r} \sum_{i_0,\ldots,i_{r-1}} E\left[\left(|e_j^r - \xi_j^r q_{R(j)}\left(\frac{e_j^r}{\xi_j^r}\right)|^{2+\epsilon}\left(\xi_j^{K_{R(j)}}\left(\frac{e_j^r}{\xi_j^r}\right)\right)^{-\epsilon} + \left(\xi_j^{K_{R(j)}}\left(\frac{e_j^r}{\xi_j^r}\right)\right)^2\right)|R_{(j+1)r-1} = r_{i_{r-1}}, \ldots, R_{jr} = r_{i_0}\right] P\{R_{jr} = r_{i_0}\} p_{i_{r-1},i_r} + 2^{1+\epsilon}\varphi^2,
$$

$$
\leq 2^{1+\epsilon}|\lambda|^{2r} \sum_{i_0,\ldots,i_{r-1}} E\left[\left(e_+^2 \xi_j^2 + \xi_j^2\right)|R_{jr} = r_{i_0}\right] P\{R_{jr} = r_{i_0}\} p_{i_{r-1},i_r} + 2^{1+\epsilon}\varphi^2,
$$

$$
\leq 2^{1+\epsilon} \xi \sum_{i_0} \left(\sum_{i_1,\ldots,i_{r-1}} \frac{|\lambda|^{2r}}{2^{2r_0+\ldots+r_{r-1}}} p_{i_0,i_1} \cdots p_{i_{r-1},i_r}\right) \theta_{j,i_0} + 2^{1+\epsilon}\varphi^2.
$$

(50)

Here step (a) follows from the law of total probability, (b) is justified by the fact that

$$(e_j^r, \xi_j, R(j)) \rightarrow (R_{jr}, \ldots, R_{(j+1)r-1}) \rightarrow R_{(j+1)r}$$

form a Markov chain, and (c) follows from the fact that $(e_j^r, \xi_j) \rightarrow R_{jr} \rightarrow (R_{jr+1}, \ldots, R_{(j+1)r-1})$ form a Markov chain and the following lemma which bounds the achievable distortion with the quantizer defined in [18]:

**Lemma 6:**  [18, Lemma 5.2] Let $e$ and $\xi > 0$ be random variables with $E[e^{2+\epsilon}]$ bounded for some $\epsilon > 0$. If $\delta > 2^{2/\epsilon}$, then for any positive rate $r$ the quantization error $e - \xi q_r(e/\xi)$ satisfies

$$
E\left[\left(e - \xi q_r(e/\xi)\right)^{2+\epsilon}\left(\xi q_r(e/\xi)\right)^{-\epsilon} + \left(-\xi q_r(e/\xi)\right)^2\right] \leq \delta \frac{\xi}{2^{2r}} E\left[e^{2+\epsilon} \xi^{-\epsilon} + \xi^2\right]
$$

for some constant $\zeta > 2$ only determined by $\epsilon$ and $\delta$. 

DRAFT
Next, we claim that, for every $k \geq 0$,

$$\theta_{j+1,i} \leq \mathbb{E}[z_{j+1}^2 1\{R_{(j+1)\tau} = r_i\}], \quad r_i \in \mathcal{R},$$

(51)

where the process $\{z_{j\tau}\}$ is formed recursively from $z_0 = \theta_0$ as

$$z_{j\tau} = \phi \frac{|\lambda|^\tau}{2^{R_{(j-1)\tau}+...+R_{(j-1)\tau-1}}} z_{(j-1)\tau} + \varsigma, \quad j \geq 1,$$

(52)

with $\phi = 2^{1+\varepsilon} \zeta > 1$ and $\varsigma = 2^{1+\varepsilon} \varphi^2 > 0$. To see this, consider the following inductive argument. By construction $z_0 = \theta_0$, hence the claim holds for $k = 0$. Now, suppose that the claim is true up to time $k$. Then, for any $r_i \in \mathcal{R}$,

$$\mathbb{E}[z_{(j+1)\tau}^2 1\{R_{(j+1)\tau} = r_i\}] = \phi \sum_{i_0,...,i_{\tau-1}} \frac{|\lambda|^{2\tau}}{2^{2(r_{i_0}+...+r_{i_{\tau-1}})}} p_{i_0,i_1}...p_{i_{\tau-1},i_\tau} \mathbb{E}[z_{j\tau}^2 1\{R_{j\tau} = r_{i_0}\}] + \varsigma \geq \phi \sum_{i_0,...,i_{\tau-1}} \frac{|\lambda|^{2\tau}}{2^{2(r_{i_0}+...+r_{i_{\tau-1}})}} p_{i_0,i_1}...p_{i_{\tau-1},i_\tau} \theta_{j,i_0} + \varsigma \geq \theta_{j+1,i_\tau}\]$$

where the first inequality follows from the induction hypothesis, while the second inequality uses (50). Hence, the claim holds at time $k + 1$ as well.

Summing both sides of (53) over $r_i \in \mathcal{R}$ and making use of (45), it follows that

$$\mathbb{E}[\varepsilon_{j\tau}^2] \leq \mathbb{E}[z_{j\tau}^2]$$

(53)

for all $j \geq 0$. Therefore, the second moment stability of $\{z_{j\tau}\}$ is a sufficient condition to the stability of the system, and the first statement of the Lemma is proved.

On the other hand, notice that (52) represents the evolution at times $\tau, 2\tau, \ldots$ of the MJLS with initial condition $z_0 = \theta_0$ and dynamics

$$z_k = \phi^{1/\tau} |\lambda| z_{k-1} + c, \quad k \geq 1.$$

for some appropriately chosen constant $c > 0$. Thus, by Lemma 1 sufficient condition for second moment stability of $\{z_k\}$ is that

$$\phi^{2/\tau} \rho (|\lambda|^2 H) < 1.$$  

(54)
Finally, if the condition of Theorem 1 is satisfied, that is, if $|\lambda|^2 \rho(H) < 1$, then we can choose the duration of a cycle $\tau$ large enough to ensure that (54) holds and, as a consequence, $\{e_{j\tau}\}$ is second-moment stable.

APPENDIX D

PROOF OF PROPOSITION 4

In order to prove the proposition, we first need to state two lemmas. The first lemma compares the spectral radius of a product of nonnegative matrices to the product of the spectral radii of each matrix.

Lemma 7 (Corollary 3 in [12]): Let $H_1, H_2, \ldots, H_m$ be positive square matrices. Let $x_i$ denote the eigenvector associated to the Perron eigenvalue of $H_i$, i.e., such that $H_ix_i = \rho(H_i)x_i$, $i = 1, \ldots, m$. Then, for any permutation $\sigma$ of $\{1, \ldots, m\}$ we have

$$c_{\min} \leq \frac{\rho(H_{\sigma_1}H_{\sigma_2} \cdots H_{\sigma_m})}{\rho(H_1)\rho(H_2) \cdots \rho(H_m)} \leq c_{\max} (55)$$

where $c_{\min} = \prod_{i=1}^{k} (\max_j x_{i,j} / x_j)^{-1}$, $c_{\max} = \prod_{i=1}^{k} (\max_j x_{i,j} / x_j)$, and $x_j = \min_i x_{i,j}$.

By the Perron-Frobenius theorem all the components of the Perron eigenvector $x_i$ are positive, $i = 1, \ldots, m$, so $0 < c_{\min} \leq c_{\max} < \infty$. The next lemma shows that equality holds in some asymptotic sense.

Lemma 8: Under the same assumptions as Theorem 7, we have

$$\lim_{k \to \infty} \left( \rho(H_{\sigma_1}^kH_{\sigma_2}^k \cdots H_{\sigma_m}^k) \right)^{\frac{1}{k}} = \rho(H_1)\rho(H_2) \cdots \rho(H_m). (56)$$

Proof: For any $k > 1$, $H_i^k$ has the same Perron eigenvector $x_i$ as $H_i$ and $\rho(H_i^k) = (\rho(H_i))^k$, i.e., $H_i^k x_i = (\rho(H_i))^k x_i$, $i = 1, \ldots, m$. Then, by Theorem 7

$$c_{\min}^k \leq \frac{\rho(H_{\sigma_1}^kH_{\sigma_2}^k \cdots H_{\sigma_m}^k)}{\rho(H_1)\rho(H_2) \cdots \rho(H_m)} \leq c_{\max}^k (57)$$

and the claim follows by taking the limit $k \to \infty$.

Equipped with the above lemmas, we can proceed to the proof of Theorem 3.
Proof: Let \( m, m' \geq 1, \alpha_1, \ldots, \alpha_m, \alpha'_1, \ldots, \alpha'_{m'} \in A \). To prove the claim it suffices to show that for any rate tuple \((\log |\lambda_1|, \ldots, \log |\lambda_u|)\) strictly inside the convex closure of

\[
C(m, \alpha_1, \ldots, \alpha_m) \cup C(m', \alpha'_1, \ldots, \alpha'_{m'})
\]

there exist an \( m'' \geq 1 \) such that

\[
(\log |\lambda_1|, \ldots, \log |\lambda_u|) \in C(m'', \alpha''_1, \ldots, \alpha''_{m''})
\]

for some \( \alpha''_1, \ldots, \alpha''_{m''} \in A \).

For every \( i \in U \), let \( H_i = \prod_{j=1}^{m} H_i(\alpha_j) \) and \( H'_i = \prod_{j=1}^{m'} H_i(\alpha'_j) \). Observe that any rate tuple \((\log |\lambda_1|, \ldots, \log |\lambda_u|)\) strictly inside the convex closure of

\[
C(m, \alpha_1, \ldots, \alpha_m) \cup C(m', \alpha'_1, \ldots, \alpha'_{m'})
\]

satisfies

\[
\log |\lambda_i| \leq -\frac{\gamma}{2m} \log \rho(H_i) - \frac{1 - \gamma}{2m'} \log \rho(H'_i) - \epsilon, \text{ for all } i \in U,
\]

for some \( \gamma \in [0, 1] \) and \( \epsilon > 0 \). Let \( 0 < \epsilon' < \epsilon \). By Lemma 8, for \( k \) large enough we have that for every \( i \in U \)

\[
-\frac{1}{2km} \log \left( \rho(H_i^{\gamma km'} H_i^{(1-\gamma)km}) \right) \\
\geq -\frac{\gamma}{2m} \log \rho(H_i) - \frac{1 - \gamma}{2m'} \log \rho(H'_i) - \epsilon' \\
> -\frac{\gamma}{2m} \log \rho(H_i) - \frac{1 - \gamma}{2m'} \log \rho(H'_i) - \epsilon \\
\geq \log |\lambda_i|.
\]

This shows that the rate tuple \((\log |\lambda_1|, \ldots, \log |\lambda_u|)\) can be stabilized by the control scheme in Theorem 3 because it is contained inside the hypercube \( C(km, \alpha''_1, \ldots, \alpha''_{km'}) \), where the sequence of bit allocation functions \( \{\alpha''_1, \ldots, \alpha''_{km'}\} \) is obtained by concatenating \( \gamma km' \) copies of \( \alpha_1, \ldots, \alpha_m \) followed by \( (1 - \gamma)km \) copies of the bit allocation functions \( \alpha'_1, \ldots, \alpha'_{m'} \).

\[
\Box
\]
REFERENCES


