

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.

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 Fig. 1. This is an issue of relevance in the design of cyberphysical systems [14].

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 slowly varying wireless fading channels for which the channel conditions can be learnt with a short training sequence.

Following this model, Tatikonda and Mitter [27] examined the special case where the rate process is constant in time and the system has bounded disturbances. Nair and Evans [19] studied the case where the disturbances have unbounded support, maintaining the rate constant. Martin *et al.* [17] analyzed the case of a time-varying independent and identically distributed (i.i.d.) rate process but bounded system disturbances.

L. Coviello and M. Franceschetti are with the Department of Electrical and Computer Engineering, University of California, San Diego CA, 92093, USA. (e-mail: lcoviell@ucsd.edu, massimo@ece.ucsd.edu)

P. Minero is with the Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN, 46556 USA (e-mail: pminero@nd.edu).

This work was partially supported by NSF grants CNS-0916778, CNS 0546235, and ARO W911NF-10-1-0468.

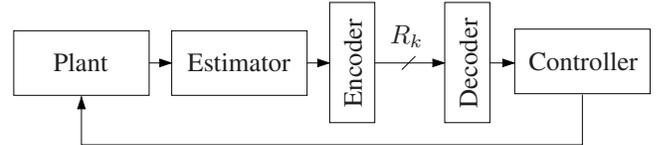


Fig. 1. Feedback loop model. The estimated state is quantized, encoded and sent to a decoder over a digital channel. The rate process $\{R_k\}_{k \geq 0}$ evolves as a Markov chain.

The work in [18] allowed for both a time-varying i.i.d. rate process and unbounded disturbances. Finally, You and Xie [29] 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 Fig. 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 λ , 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 $\{R_k\}_{k \geq 0}$. Intuitively, this equivalent MJLS describes the evolution of the mean-square estimation error at the decoder, which at every time step k increases by $|\lambda|^2$ because of the system dynamics and is reduced by 2^{2R_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

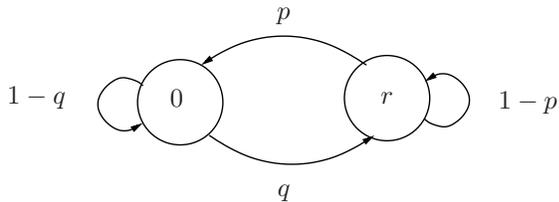


Fig. 2. A two-state Markov chain modeling a bursty packet erasure channel.

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 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 [18], [29], 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 [29] for the special case of a Markov erasure channel. As a remark, in [30] a necessary and sufficient condition for the mean-square stabilization of a vector systems with scalar input is derived in the case of a channel with finite rate and *iid* packet losses.

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 [20]. Broadly speaking, authors have followed two distinct approaches. The information-theoretic approach followed by the present paper and by [1], [4], [11], [17], [19], [27]–[31], 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], [22], [25] 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 [18] 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 $r \rightarrow \infty$. The same bridge also holds for our work that generalizes [18]. 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 $r \rightarrow \infty$.

The theory of MJLS has been previously used to investigate control problems over communication channels. Seiler and Sengupta [23] 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.* [16] 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.* [21] 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 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

$$x_{k+1} = Ax_k + Bu_k + v_k, y_k = Cx_k + w_k, \quad k = 0, 1, \dots \quad (1)$$

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. This assumption is made for ease of presentation but in the case of a system with both stable and unstable modes, all our results hold by replacing A by the submatrix corresponding to the unstable subspace of the plant. We assume that the distributions of the initial condition x_0 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. x_0 , v_k and w_k , $k \geq 0$, have uniformly bounded $(2 + \epsilon)$ th absolute moments for some $\epsilon > 0$.
- A2. The system disturbance v_k has a continuous probability density function f of finite differential entropy $h(v_k) = -\int f(x) \log_2 f(x) dx$, so there exists a constant $\beta > 0$ such that $e^{2h(v_k)/d} > \beta$ for all k .

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. 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

$$\mathcal{R} = \{r_1, \dots, r_n\},$$

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

$$p_{ij} = \mathbb{P}\{R_{k+1} = r_j | R_k = r_i\}$$

for all $k \in \mathbb{N}$ and $i, j \in \{1, \dots, n\}$. 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 $s_k = s_k(y_0, \dots, y_k)$ that maps all past and present measurements into the set $\{1, \dots, 2^{R_k}\}$, the digital link is the identity function on the set $\{1, \dots, 2^{R_k}\}$, and the control u_k is a function $\hat{x}_k(s_0, \dots, s_k)$ 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,

$$\sup_k \mathbb{E} [\|x_k\|^2] < \infty, \quad (2)$$

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

A. Markov Jump Linear Systems.

Consider the scalar non-homogeneous MJLS defined by

$$z_{k+1} = \frac{\lambda}{2^{R_k}} z_k + c, \quad (3)$$

where $z_k \in \mathbb{R}$ with $z_0 < \infty$, $c \geq 0$ is a constant, $\{R_k\}_{k \geq 0}$ is the Markov rate process described above. Let $H^{(m)}$, $m = 1, 2$, be the $n \times n$ matrix with nonnegative real elements

$$h_{ij}^{(m)} = \frac{1}{2^m r_j} p_{ji}, \quad (4)$$

for all $i, j \in \{1, \dots, n\}$. The following (known) lemma states the necessary and sufficient condition for m -th moment stability, $m = 1, 2$, 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 m -th moment stability of the system (3), i.e., $\sup_k \mathbb{E} [|z|^m] < \infty$, is that

$$|\lambda|^2 < \frac{1}{\rho(H^{(m)})}, \quad (5)$$

for $m = 1, 2$. If $c = 0$, then “ $<$ ” in (5) is replaced by “ \leq ”.

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 $H^{(m)}$ in determining the dynamics of the m -th moment of z_k . ■

The above claim is sufficient for our purposes, but in general a stronger results holds. The condition in (5) 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 m -th moment stable but also that the m -th moment of the system converge to a steady state value independent of the initial condition [3].

III. SCALAR SYSTEM

Consider the special case of a scalar system

$$x_{k+1} = \lambda x_k + u_k + v_k, \quad y_k = x_k + w_k, \quad k = 0, 1, \dots, \quad (6)$$

where $|\lambda| \geq 1$. Let H be the $n \times n$ matrix with nonnegative real elements

$$h_{ij} = \frac{1}{2^{2r_j}} p_{ji}, \quad (7)$$

for all $i, j \in \{1, \dots, n\}$.

Theorem 1: Under assumptions A0-A2, there exists a control scheme that stabilizes the scalar system (6) in mean-square sense if and only if the MJLS (3) is mean-square stable, that is, if and only if

$$|\lambda|^2 < \frac{1}{\rho(H)}. \quad (8)$$

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 (8) reduces to the well known data rate theorem¹ condition [19], [26],

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

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 $p_i = \mathbb{P}\{R = r_i\}$, $r_i \in \mathcal{R}$. In this case

$$H = (p_1, \dots, p_n)^T (2^{-2r_1}, \dots, 2^{-2r_n}) = p h^T$$

is a rank-one matrix whose only nonzero eigenvalue is $h^T p$. Therefore, Theorem 1 yields the result in [17], [18],

$$\begin{aligned} |\lambda|^2 \rho(H) &= |\lambda|^2 (2^{-2r_1}, \dots, 2^{-2r_n}) (p_1, \dots, p_n)^T \\ &= \mathbb{E} \left[\frac{|\lambda|^2}{2^{2R}} \right] < 1. \end{aligned}$$

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

c) Two-state Markov process. In the special case illustrated in Fig. 2 in which $n = 2$, $p_{12} = q$, and $p_{21} = p$ for some

¹In the case of a deterministically varying rate process, r denotes the infinite horizon time-average of the process.

$0 < p, q < 1$, we have

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

and the condition in Theorem 1 reduces to

$$|\lambda|^2 \rho(H) = \frac{|\lambda|^2}{2} \text{tr}(H) + \frac{|\lambda|^2}{2} \sqrt{\text{tr}(H)^2 - 4\det(H)} < 1, \quad (9)$$

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 [29] where, by following a different approach, the authors proved that necessary and sufficient condition for stabilization is that

$$\mathbb{E} \left[\frac{|\lambda|^{2T}}{2^{2r}} \right] < 1. \quad (10)$$

Here T denotes the random ‘‘excursion time’’ between two consecutive visits of state r , and the expectation is taken with respect to T . 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 (10) simplifies to

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

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 (9) is equivalent to (10). Hence our result, that is derived via a reduction to a MJLS problem, recovers the one in [29] and extend it to the case of a general finite-state Markov rate process, for which the stability condition (10) derived via a random ‘‘excursion time’’ approach does not seem to generalize in a straightforward manner.

Taking the limit $r \rightarrow \infty$, the above inequalities simplify 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 (6) is lower bounded by the first moment of a MJLS. The proof is based on information-theoretic bounds and follows similar steps as in the proof of necessity in [29].

Lemma 2: For every $k = 0, 1, \dots$,

$$\mathbb{E}[|x_k|^2] > \frac{1}{2\pi e} \mathbb{E}[|z_k|], \quad (11)$$

where $\{z_k\}$ is a non-homogeneous MJLS

$$z_{k+1} = \frac{|\lambda|^2}{2^{2R_k}} z_k + \beta, \quad (12)$$

with $z_0 = e^{2h(x_0)}$, where $\beta > 0$ is defined in Assumption A2.

Proof: See Appendix B. ■

It follows from Lemma 2 that the state cannot be stable if the MJLS in $\{z_k\}$ is first-moment unstable. Hence, the ‘‘only

if’’ condition in Theorem 1 follows directly from Lemma 1, since the augmented matrix $H^{(1)}$ for the MJLS (12) is equal to H .

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

$$x_{k+1} = \lambda x_k + u_k, \quad k = 0, 1, \dots, \quad (13)$$

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

$$|x_k| \leq z_k, \quad k = 0, 1, \dots, \quad (14)$$

where z_k is a homogeneous MJLS with dynamics

$$z_{k+1} = \frac{|\lambda|}{2^{R_k}} z_k, \quad k = 0, 1, \dots$$

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 $|x_k - \hat{x}_k| \leq z_k/2^{R_k}$, thus

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

i.e., the claim holds at time $k+1$ as well. It follows that x_k is stable if the homogeneous MJLS z_k is second-moment stable. Since the augmented matrix $H^{(2)}$ defined in (4) for this MJLS is equal to H , Lemma 1 implies that the plant is stable if $|\lambda|^2 \rho(H) \leq 1$. In this specific setting, it is possible to show that condition (8) is the necessary and sufficient condition to ensure convergence of the system to zero, i.e. $\lim_{k \rightarrow \infty} \mathbb{E}[|x_k|^2] = 0$, which is a stronger notion of stability than the one we have assumed in (2).

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 mean-square estimation error at the decoder, which at every time step k increases by $|\lambda|^2$ because of the system dynamics and is reduced by 2^{2R_k} because this is the best attainable accuracy of representation sending R_k bits across the channel.

2) *Unbounded noise and initial condition:* Consider now the system (6) assuming that conditions A0-A2 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], [15], 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 [19].

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 τ . 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 the second moment of a non-homogeneous MJLS, so the stability of this MJLS implies the stability of the state. Specifically, we prove that at times $\tau, 2\tau, 3\tau, \dots$, i.e., at the beginning of each cycle, the second moment of the estimation error $x_{j\tau} - \hat{x}_{j\tau}$ satisfies

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

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

$$z_{j\tau} = \phi \frac{|\lambda|^\tau}{2^{R_{(j-1)\tau} + \dots + R_{j\tau-1}}} z_{(j-1)\tau} + \varsigma, \quad j \geq 1, \quad (15)$$

for some bounded initial condition z_0 and constants $\phi > 1$, $\varsigma > 0$. It is immediate to verify that the process $z_\tau, z_{2\tau}, \dots$ is upper bounded by the process $w_\tau, w_{2\tau}, \dots$, defined as the evolution at times $\tau, 2\tau, \dots$ of the non-homogeneous MJLS with dynamics

$$w_k = \phi^{1/\tau} \frac{|\lambda|}{2^{R_{k-1}}} w_{k-1} + \text{const}, \quad k \geq 1.$$

By Lemma 1, this MJLS is second-moment stable if

$$\phi^{2/\tau} |\lambda|^2 \rho(H) < 1. \quad (16)$$

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 τ large enough to ensure that (16) holds and, as a consequence, (15) is stable, the second moment of the estimation error at the beginning of each cycle is bounded, and the state remains second-moment bounded. The choice of a larger τ translates into larger oscillations of the system state because, according to our quantization scheme, the system evolves in open loop during a cycle. However, this is not a problem as we are interested in the mean-square stabilization of the system.

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 [19] 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 $i \in \{2, \dots, r-1\}$,

- the leftmost and rightmost intervals are the semi-open sets $(-\infty, -\delta^{r-2}]$ and (δ^{r-2}, ∞) .

Here $\delta > 1$ is a parameter that determines the concentration of intervals around the origin. In particular, we set $\delta > 2^{2/\epsilon}$, for some ϵ such that assumption A1 holds. We can see that the width of the quantization regions increases with δ , so the partition becomes more spread out as δ increases. 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 into two intervals each. In particular, the interval (δ^{r-2}, ∞) is divided into the bounded interval $(\delta^{r-2}, \delta^{r-1}]$ and the semi-open interval (δ^{r-1}, ∞) , and similarly for the interval $(-\infty, -\delta^{r-2}]$. 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 Luenberger observer of the state. This is computed recursively as

$$\bar{x}_{k+1} = \lambda \bar{x}_k + u_k + L(y_k - \bar{x}_k), \quad (17)$$

where $\bar{x}_k = 0$, and L is a constant such that $|\lambda - L| < 1$. 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 τ . At the beginning of the j th cycle, i.e., at time $j\tau$, the coder computes

$$q_{R_{j\tau}}((\bar{x}_{j\tau} - \hat{x}_{j\tau})/\xi_j), \quad (18)$$

where ξ_j is a scaling factor updated at the beginning of each cycle, and communicates to the decoder the index $s_{j\tau} \in \{1, \dots, 2^{R_{j\tau}}\}$ 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 $s_{j\tau+1} \in \{1, \dots, 2^{R_{j\tau+1}}\}$ equal to the subinterval containing $(\bar{x}_{j\tau} - \hat{x}_{j\tau})/\xi_j$, so the decoder can compute

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

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

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

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

$$\hat{x}_{(j+1)\tau} = \lambda^\tau \left(\hat{x}_{j\tau} + \xi_j q_{R(j)} \left(\frac{\bar{x}_{j\tau} - \hat{x}_{j\tau}}{\xi_j} \right) \right) + \sum_{k=j\tau}^{(j+1)\tau-1} \lambda^{(j+1)\tau-1-k} K \hat{x}_k, \quad (19)$$

where

$$\hat{x}_{k+1} = (\lambda + K) \hat{x}_k, \text{ for } k = j\tau, \dots, (j+1)\tau - 2, \quad (20)$$

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

$$\xi_{j+1} = \max\left\{ \varphi, |\lambda|^\tau \xi_j^k q_{R(j)} \left(\frac{\bar{x}_{j\tau} - \hat{x}_{j\tau}}{\xi_j} \right) \right\}, \quad (21)$$

with $\xi_0 = \varphi$, where φ is a constant independent of the state that will be specified later.

Decoder. Encoder and decoder are synchronized and share the state estimator \hat{x}_k and the scaling factor ξ_j . Synchronism is guaranteed because both encoder and decoder set $\hat{x}_0 = 0$ and $\xi_0 = \varphi$ and the channel is noiseless. At every time k the decoder sends to the plant the control signal

$$u_k = K \hat{x}_k, \quad (22)$$

where \hat{x}_k is updated as in (19) at the beginning of each cycle and as in (20) within each cycle.

Analysis. First, we prove that if (8) 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 $j = 0, 1, \dots$

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

where $z_{j\tau}$ satisfies (15) for some bounded initial condition z_0 and constants $\phi > 1, \varsigma > 0$. Furthermore, (8) is a sufficient condition to ensure that $\{z_{j\tau}\}$ is second-moment stable and that

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

Proof: See Appendix C

By (17), (20) and (22) we have that

$$\begin{aligned} \bar{x}_{j\tau+k} - \hat{x}_{j\tau+k} &= \lambda^k (\bar{x}_{j\tau} - \hat{x}_{j\tau}) \\ &+ L \sum_{i=0}^{k-1} \lambda^{k-1-i} (y_{j\tau+i} - \bar{x}_{j\tau+i}) \end{aligned}$$

for all $j \geq 0$ and $k = 0, 1, \dots, \tau - 1$, where L is the constant of the Luenberger observer such that $|\lambda - L| < 1$. Therefore, by Minkowski inequality,

$$\begin{aligned} (\mathbb{E}[|\bar{x}_{j\tau+k} - \hat{x}_{j\tau+k}|^2])^{1/2} &\leq |\lambda|^k (\mathbb{E}[|\bar{x}_{j\tau} - \hat{x}_{j\tau}|^2])^{1/2} \\ &+ \left(\mathbb{E} \left[\left| \sum_{i=0}^{k-1} \lambda^{k-1-i} L (y_{j\tau+i} - \bar{x}_{j\tau+i}) \right|^2 \right] \right)^{1/2} \end{aligned}$$

for all $j \geq 0$ and $k = 0, 1, \dots, \tau - 1$.

Observe that the error $\varepsilon_k = x_k - \bar{x}_k$ of the Luenberger observer (17) evolves as $\varepsilon_{k+1} = (\lambda - L)\varepsilon_k - Lw_k + v_k$ and is mean-square stable since $|\lambda - L| < 1$ and w_k and v_k are second-moment bounded by assumption A1. This implies that $y_k - \bar{x}_k$ is second-moment bounded since $(y_k - \bar{x}_k)^2 \leq 2[(x_k - \bar{x}_k)^2 + w_k^2]$. Hence, it follows from Lemma 3 that the second moment of $\bar{x}_{j\tau+k} - \hat{x}_{j\tau+k}$ is bounded for all $j \geq 0$ and $k = 0, 1, \dots, \tau - 1$ if (8) holds.

Finally, we can rewrite (17) as

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

where K is the certainty-equivalent control such that $|\lambda + K| < 1$. It follows that $\sup_k \mathbb{E}[\bar{x}_k^2] < \infty$ and therefore the system (6) is stable in the mean-square sense (2) since $x_k^2 \leq 2[\bar{x}_k^2 + \varepsilon_k^2]$.

IV. VECTOR SYSTEM

Consider the system (1), with $A \in \mathbb{R}^{d \times d}$. Let $\lambda_1, \dots, \lambda_u \in \mathbb{C}$ be the distinct, non-conjugate eigenvalues of A (if λ_i, λ_i^* are complex conjugate eigenvalues, only one of them is considered). Let μ_i be the algebraic multiplicity of λ_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 $J = \text{diag}(J_1, \dots, J_u)$ 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 $\mathcal{U} = \{1, \dots, u\}$ denote the index set of the subsystems corresponding to the modes $\lambda_1, \dots, \lambda_u$. Consider the recursion

$$x_{k+1} = Jx_k + TBu_k + Tv_k, \quad y_k = CT^{-1}x_k + w_k. \quad (23)$$

As (23) and (1) are related only by the matrix T , we assume that the system evolves according to (23).

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. For every $i = 1, \dots, u$, let $\nu_i \in \{0, \dots, \mu_i\}$ denote the algebraic dimensionality of the i th unstable mode after the genie’s intervention and let $\mathcal{V} = \{\nu : \nu_i \in \{0, \dots, \mu_i\}\} \subset \mathbb{N}^u$ denote the space of all possible algebraic dimensionalities. It is known that for every Jordan block J_i and for every $\nu_i \in \{1, \dots, \mu_i\}$ there exists an invariant real subspace of dimension $a_i \nu_i$. Given $\nu \in \mathcal{V}$, let $x_k(\nu)$ be equal to the Cartesian product of the invariant real subspaces associated to each block $J_i, i \in \mathcal{U}$. Observe that the reduced system $x_k(\nu)$ is real valued and has total dimension

$$d(\nu) = \sum_{i \in \mathcal{U}} a_i \nu_i. \quad (24)$$

Next, we find necessary conditions for the second-moment stability of $x_k(\nu)$. The following lemma shows that the second

moment of $x_k(\nu)$ is lower bounded by the second moment of a MJLS whose evolution depends on the unstable modes of $x_k(\nu)$ and their dimensionalities.

Lemma 4: For each $\nu \in \mathcal{V}$ and every $k \geq 0$,

$$\mathbb{E}[\|x_k(\nu)\|^2] \geq c_1 \mathbb{E}[|z_k(\nu)|],$$

where $z_k(\nu)$ is a MJLS defined by

$$z_{k+1}(\nu) = \left(\frac{\prod_{i \in \mathcal{U}} |\lambda_i|^{a_i \nu_i}}{2^{R_k}} \right)^{2/d(\nu)} z_k(\nu) + c_2,$$

with $z_0(\nu) = 2^{\frac{2}{d(\nu)} h(x_0(\nu))}$ and suitable constants $c_1, c_2 > 0$.

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 first-moment unstable. Then, we can derive a set of necessary conditions in terms of the spectral radius of the matrix $H^{(1)}$ describing the dynamics of the first moment of $z_k(\nu)$. Let $H(\nu)$ be the $n \times n$ matrix with nonnegative real elements

$$h_{ij}(\nu) = \frac{1}{2^{\frac{2}{d(\nu)} r_j}} p_{ji} \quad (25)$$

for all $i, j \in \{1, \dots, n\}$. Combining Lemma 1 with Lemma 4, the following result follows.

Theorem 2: Under assumptions A0-A2, necessary condition for the stabilization of the system (23) in the mean-square sense (2) is that, for every $\nu \in \mathcal{V}$,

$$|\lambda(\nu)|^{2/d(\nu)} < \frac{1}{\rho(H(\nu))} \quad (26)$$

where $\lambda(\nu) = \prod_{i \in \mathcal{U}} |\lambda_i|^{a_i \nu_i}$, $d(\nu)$ is as in (24), and $H(\nu)$ is given by (25).

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 for $\nu \neq 0$ $\lambda(\nu) = \lambda$, $d(\nu) = 1$, and $H(\nu)$ has elements given by (7), 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 $H(\nu) = 2^{-\frac{2}{d(\nu)} r}$ and thus (26) reduces to the well known data rate theorem [19], [26],

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

c) 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 $p_i = \mathbb{P}\{R = r_i\}$, $r_i \in \mathcal{R}$

$$\begin{aligned} \rho(H(\nu)) &= \rho((p_1, \dots, p_n)^T (2^{-\frac{2}{d(\nu)} r_1}, \dots, 2^{-\frac{2}{d(\nu)} r_n})) \\ &= (2^{-\frac{2}{d(\nu)} r_1}, \dots, 2^{-\frac{2}{d(\nu)} r_n}) (p_1, \dots, p_n)^T \\ &= \mathbb{E} \left[2^{-\frac{2}{d(\nu)} R} \right], \end{aligned}$$

and (26) simplifies to

$$\sum_{i \in \mathcal{U}} a_i \nu_i \log |\lambda_i| < -\frac{d(\nu)}{2} \log \mathbb{E} \left[2^{-\frac{2}{d(\nu)} R} \right], \quad (27)$$

and our condition reduces to the one in [18]. 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 $r \rightarrow \infty$, then the stability condition

$$\max_{i \in \mathcal{U}} \lambda_i^2 < \frac{1}{p_1}$$

depends only on the erasure rate of the channel.

d) Two-state Markov process. In the special case illustrated in Fig. 2, (26) simplifies to

$$\begin{aligned} \log |\lambda(\nu)| &< -\frac{d(\nu)}{2} \log \left(\frac{1}{2} \text{tr}(H(\nu)) \right) \\ &+ \frac{1}{2} \sqrt{\text{tr}(H(\nu))^2 - 4 \det(H(\nu))}, \end{aligned}$$

where

$$\text{tr}(H(\nu)) = (1 - q) 2^{-\frac{2}{d(\nu)} r_1} + (1 - p) 2^{-\frac{2}{d(\nu)} r_2}$$

and

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

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 [29]. Also, it can be easily seen that if $r_1 = 0$, then as $r_2 = r \rightarrow \infty$ we have that $\det(H(\nu)) \rightarrow 0$ and $\text{tr}(H(\nu)) \rightarrow 1 - q$, so the necessary condition reduces to

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

B. Sufficiency

A sufficient condition for mean-square stability is given extending the control scheme described in [18]. 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 [18] 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 \mathcal{A} denote the set of functions $\alpha : \mathcal{R} \rightarrow [0, 1]^u$ having the properties that

$$\frac{1}{a_i \mu_i} \alpha^{(i)}(r) r \in \mathbb{N}, \quad i = 1, \dots, u, \quad (28)$$

and

$$\alpha^{(1)}(r) + \dots + \alpha^{(u)}(r) \leq 1 \quad (29)$$

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

The operational meaning of $\alpha \in \mathcal{A}$ is as follows. At every time k , $\frac{1}{a_i \mu_i} \alpha^{(i)}(R_k) R_k$ bits are used to quantize each scalar component of the i th unstable state of the system². Condition (28) enforces that each subsystem is quantized using

²We 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.

an integer number of bits, while (29) ensures that we do not use more bits than what is available for transmission. Following the approach in the proof of the sufficiency part in [19], it can be shown that the i th subsystem is stable if the non-homogeneous MJLS with dynamics

$$z_{k\tau} = \phi \frac{|\lambda_i|^\tau}{2^{\frac{1}{a_i \mu_i} \sum_{j=(k-1)\tau}^{k\tau-1} \alpha^{(i)}(R_j) R_j}} z_{(k-1)\tau} + \varsigma,$$

for some bounded initial condition z_0 and constants $\phi > 1$, $\varsigma > 0$ 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 $H_i(\alpha)$, $i = 1, \dots, u$, as the $n \times n$ real matrix with (l, j) th entry equal to

$$2^{-2 \frac{\alpha^{(i)}(r_j)}{a_i \mu_i} r_j} p_{jl}, \quad (30)$$

and conclude that, given a bit allocation function α , sufficient condition for second-moment stability of the system is that

$$|\lambda_i|^2 < \frac{1}{\rho(H_i(\alpha))} \quad (31)$$

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-A2 above, a sufficient condition for the stabilization of the system (23) in the mean-square sense is that

$$|\lambda_i|^2 < \frac{1}{\rho(H_i(\alpha_1) H_i(\alpha_2) \cdots H_i(\alpha_m))^{1/m}} \quad \text{for all } i \in \mathcal{U}, \quad (32)$$

for some $m \geq 1$ and $\alpha_1, \dots, \alpha_m \in \mathcal{A}$.

Proof: Fix $m \geq 1$ and $\alpha_1, \alpha_2, \dots, \alpha_m \in \mathcal{A}$, non necessarily distinct. The main idea of the proof is to use these bit allocation functions in a cyclic repeated order via a time-sharing protocol. As in the proof of Lemma 3, divide time into cycles of length τ and further divide each cycle into q blocks of length m each, such that $\tau = qm$. Within each cycle, use the bit allocation function α_j in the j th channel use of each block, $j = 1, \dots, m$. Let

$$R_i(k) = \frac{1}{a_i \mu_i} \sum_{l=0}^{q-1} \sum_{j=1}^m \alpha_j^{(i)}(R_{(k-1)\tau+lm+j-1}) R_{(k-1)\tau+lm+j-1}$$

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. The proof follows the same arguments in the proof of sufficiency of [19] and in the proof of Lemma 3, and therefore we only provide a sketch. For $i \in \mathcal{U}$ and $j \geq 0$, let $e_{j\tau}^{(i)}$ denote the estimation error vector relative to the i th subsystem at time $j\tau$ (of dimension $n_i = a_i \mu_i$). For $j \geq 0$, let $e_{j\tau}$ be the d -dimensional vector obtained by stacking the vectors $e_{j\tau}^{(i)}$, for each $i \in \mathcal{U}$. Notice that

$$\|e_{j\tau}\|^{2+\varepsilon} = \left(\sum_{i \in \mathcal{U}} \|e_{j\tau}^{(i)}\|^2 \right)^{1+\varepsilon/2}.$$

For $j \geq 0$ and $r_\ell \in \mathcal{R}$, let

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

After some computation and the application of Lemma 5.2 in [19], it is possible to show that

$$\theta_{j,\ell} \leq \mathbb{E} \left[\sum_{i \in \mathcal{U}} \|z_{j\tau}^{(i)}\|^2 1_{\{R_{j\tau}=r_\ell\}} \right],$$

for all $j \geq 0$ and $r_\ell \in \mathcal{R}$, and that the second moment of the error vector is upper bounded by

$$\mathbb{E} \left[\|e_{j\tau}\|^2 \right] \leq \sum_{i \in \mathcal{U}} \mathbb{E} \left[\|z_{j\tau}^{(i)}\|^2 \right],$$

for all $j \geq 0$, where

$$z_{j\tau}^{(i)} = \phi_i \frac{|\lambda_i|^\tau}{2^{R_i(j-1)}} z_{(j-1)\tau}^{(i)} + \varsigma \quad (33)$$

for some bounded initial condition $z_0^{(i)}$ and constants $\phi_i > 1$, $\varsigma > 0$, for all $i \in \mathcal{U}$. The dynamics of the process (33) is upper bounded by the one of the non-homogenous MJLS

$$z_{lm} = \frac{\phi_i^{1/q} |\lambda_i|^m}{2^{\frac{1}{a_i \mu_i} \sum_{j=1}^m \alpha_j^{(i)}(R_{(k-1)\tau+j-1}) R_{(k-1)\tau+j-1}}} z_{(l-1)m} + \varsigma, \quad (34)$$

for $l \geq 1$. Thus, if condition (32) is satisfied, then we can choose the number of blocks q within each cycle, and hence the duration of a cycle τ , large enough to ensure that the MJLS (33) 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*

$$(\log |\lambda_1|, \dots, \log |\lambda_u|)$$

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 (26), we can see that (26) 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. There are special cases, however, for which 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, notice that for every choice of m and of $\alpha_1, \dots, \alpha_m \in \mathcal{A}$ the set

$$\mathcal{C}(m, \alpha_1, \dots, \alpha_m) = \{ (\log |\lambda_1|, \dots, \log |\lambda_u|) : (32) \text{ holds} \}$$

is an open hypercube in the domain of the unstable modes. Then, Theorem 3 states that it is possible to stabilize all systems whose unstable modes lie inside

$$\bigcup_{m \geq 1} \bigcup_{\alpha_1, \dots, \alpha_m \in \mathcal{A}} \mathcal{C}(m, \alpha_1, \dots, \alpha_m). \quad (35)$$

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: The closure of (35) is a convex set.

Proof: See Appendix D. ■

It follows from the above proposition that the set of system's modes satisfying the condition in Theorem 3 can be approximated by fixing a bound M for m and taking the *convex hull* of the union of the hypercubes over all $\alpha_1, \dots, \alpha_m \in \mathcal{A}$ and $m \leq M$. In particular, by setting $M = 1$ we obtain that the inner bound contains the set of rates inside

$$\text{conv} \bigcup_{\alpha \in \mathcal{A}} \mathcal{C}(1, \alpha), \quad (36)$$

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.

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

a) *Scalar systems.* In the special case of a scalar system, taking $m = 1$ and $\alpha_1 = 1$, Theorem 3 reduces to the sufficient condition in Theorem 1.

b) *Constant rate.* When the channel supports a constant rate r , we have that $H_i(\alpha) = 2^{-2 \frac{\alpha^{(i)}}{a_i \mu_i} r}$ for all $i \in \mathcal{U}$, and thus (32) reduces to

$$\frac{r}{m} (\alpha_1^{(i)} + \dots + \alpha_m^{(i)}) > a_i \mu_i \log |\lambda_i|$$

for all $i \in \mathcal{U}$ and some $m \geq 1, \alpha_1, \dots, \alpha_m$ (where $\alpha_k^{(i)}$ is the fraction of the rate allocated to the i -th subsystem under bit allocation α_k). Choosing $\sum_{i \in \mathcal{U}} \alpha_k^{(i)} = 1$ for all $k = 1, \dots, m$, and summing over all $i \in \mathcal{U}$, we obtain the data rate theorem [19], [26],

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

c) *Independent rate process.* In the special case of an i.i.d rate process, for every $m \geq 1$ and bit allocation functions $\alpha_1, \dots, \alpha_m \in \mathcal{A}$, it can be shown that (32) simplifies to

$$\log |\lambda_i| < - \sum_{j=1}^m \frac{1}{2m} \log \mathbb{E} \left[2^{-2 \alpha_j^{(i)} (R) R} \right], \quad \text{for all } i \in \mathcal{U}, \quad (37)$$

and therefore our condition recovers the sufficient condition given in [18]. It can be shown that in this case the set of rates satisfying Theorem 3 is exactly equal to (36).

d) *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 [29, Theorem 3]. Restating their result in our notation, You and Xie proved that it is possible to stabilize all rates inside (36). We claim that Theorem 3 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

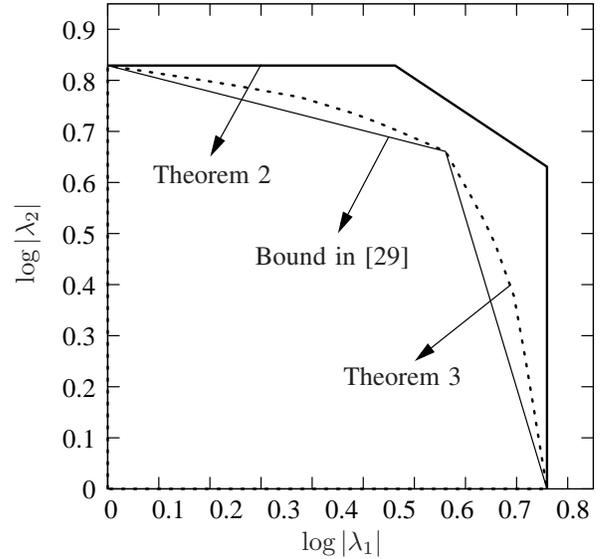


Fig. 3. Inner and outer bound on the set of stabilizable rates over a Markov erasure channel.

are $p_{12} = 0.1$ and $p_{21} = 0.7$. Let the system matrix J have two distinct real eigenvalues λ_1 and λ_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 [29, 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 $\alpha_1, \dots, \alpha_m \in \mathcal{A}$ 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 [29, 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" [24]. 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 a MJLS problem 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

We prove the lemma in the case of $c > 0$. The proof in the case of $c = 0$ is similar. For every $k \geq 0$, let $\mu_{k,i} = \mathbb{E}[z_k \mathbf{1}_{\{R_k=r_i\}}]$ and $\sigma_{k,i}^2 = \mathbb{E}[z_k^2 \mathbf{1}_{\{R_k=r_i\}}]$ 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 $z_{k+1} \rightarrow R_k \rightarrow R_{k+1}$ holds and thus

$$\begin{aligned} \sigma_{k+1,j}^2 &= \sum_{i=1}^n \mathbb{E}[|z_{k+1}|^2 | R_{k+1} = r_j, R_k = r_i] \mathbb{P}\{R_k = r_i\} p_{ij} \\ &= \sum_{i=1}^n \mathbb{E}\left[\left(\frac{|\lambda|}{2^{r_i}} |z_k| + c\right)^2\right] \mathbb{P}\{R_k = r_i\} p_{ij} \\ &= \sum_{i=1}^n \frac{|\lambda|^2}{2^{2r_i}} p_{ij} \sigma_{k,i}^2 + 2c \sum_{i=1}^n \frac{|\lambda|}{2^{r_i}} p_{ij} \mu_{k,i} \\ &\quad + c^2 \sum_{i=1}^n p_{ij} \mathbb{P}\{R_k = r_i\}. \end{aligned}$$

Similarly,

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

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

$$\begin{bmatrix} \sigma^2 \\ \mu \end{bmatrix}_{k+1} = \begin{bmatrix} |\lambda|^2 PD^2 & 2c|\lambda|PD \\ 0 & |\lambda|PD \end{bmatrix} \begin{bmatrix} \sigma^2 \\ \mu \end{bmatrix}_k + \begin{bmatrix} c^2 p \\ cp \end{bmatrix}_k, \quad (38)$$

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

$$\mathbb{E}[z_k] = \sum_{i=1}^n \mu_{k,i}, \quad \text{and} \quad \mathbb{E}[z_k^2] = \sum_{i=1}^n \sigma_{k,i}^2.$$

It follows that the MJLS $\{z_k\}$ is first-moment stable if and only if the eigenvalues of $|\lambda|PD = |\lambda|H^{(1)}$ are strictly inside the unit circle. Similarly, $\{z_k\}$ is second-moment stable if and only if the linear system (38) is second-moment stable. Observe that the eigenvalues of the block triangular matrix in (38) are the union of the eigenvalues of the matrices $|\lambda|^2 PD^2$ and $|\lambda|PD$. The next lemma states that $\rho(|\lambda|^2 PD^2) < 1$ implies

$\rho(|\lambda|PD) < 1$, so that $\rho(|\lambda|^2 PD^2) < 1$ is a necessary and sufficient condition for the second-moment stability of $\{z_k\}$.

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

Proof: Assume that $\rho(|\lambda|^2 PD^2) < 1$. In order to show the claim, it suffices to consider the homogeneous system associated to the system (38), obtained by setting $c = 0$. Fix $j \in \{1, \dots, n\}$ 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 μ_k evolves according to $\mu_k = (|\lambda|PD)^k \gamma e_j$ and that its 2-norm is upper bounded by

$$\begin{aligned} \|(|\lambda|PD)^k \gamma e_j \|^2 &= \sum_{i=1}^n |\mu_{k,i}|^2 = \sum_{i=1}^n (\mathbb{E}[z_k \mathbf{1}_{\{R_k=r_i\}}])^2 \\ &\leq \sum_{i=1}^n \mathbb{E}[|z_k|^2 \mathbf{1}_{\{R_k=r_i\}}] = \sum_{i=1}^n \sigma_{k,i}^2 \\ &= \|\sigma_k^2\|_1. \end{aligned}$$

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

APPENDIX B PROOF OF LEMMA 2

Let $S^k = \{S_0, \dots, S_k\}$ denote the symbols transmitted over the digital link up to time k . By the law of total expectation and the maximum entropy theorem [5],

$$\mathbb{E}[|x_{k+1}|^2] \geq \frac{1}{2\pi e} \mathbb{E}_{S^k} [e^{2h(x_{k+1}|S^k=s^k)}]. \quad (39)$$

It follows that the second moment of the state is lower bounded by the average entropy power of x_k conditional on S^k . From the translation invariance property of the differential entropy, the conditional version of entropy power inequality [5], and Assumption A2, it follows that

$$\begin{aligned} &\mathbb{E}_{S^k} [e^{2h(x_{k+1}|S^k=s^k)}] \\ &= \mathbb{E}_{S^k} [e^{2h(\lambda x_k + \hat{x}(s^k) + v_k | S^k=s^k)}] \\ &\geq \mathbb{E}_{S^k} [e^{2h(\lambda x_k | S^k=s^k)}] + e^{2h(v_k)} \\ &\geq |\lambda|^2 \mathbb{E}_{S^k} [e^{2h(x_k | S^k=s^k)}] + \beta. \end{aligned} \quad (40)$$

We can further lower bound (40) making use Lemma 1 in [18], which states that for every $k \geq 0$,

$$\mathbb{E}_{S^k | S^{k-1}, R_k} [e^{2h(x_k | S^k=s^k)}] \geq \frac{1}{2^{2R_k}} e^{2h(x_k | S^{k-1}=s^{k-1})}, \quad (41)$$

with $S_{-1} = \emptyset$. By the tower rule of conditional expectation, it then follows that

$$\mathbb{E}_{S^k} [e^{2h(x_k | S^k=s^k)}] \geq \mathbb{E}_{S^{k-1}, R_k} \left[\frac{1}{2^{2R_k}} e^{2h(x_k | S^{k-1}=s^{k-1})} \right]. \quad (42)$$

Combining (42) and (40) gives

$$\begin{aligned} & \mathbb{E}_{S^k} \left[e^{2h(x_{k+1}|S^k=s^k)} \right] \\ & \geq \mathbb{E}_{R_k} \left[\frac{|\lambda|^2}{2^{2R_k}} \mathbb{E}_{S^{k-1}|R_k} \left(e^{2h(x_k|S^{k-1}=s^{k-1})} \right) \right] + \beta. \end{aligned} \quad (43)$$

Following similar steps and using the Markov chain $S^{k-1} \rightarrow (S^{k-2}, R_{k-1}) \rightarrow R_k$, we obtain

$$\begin{aligned} & \mathbb{E}_{S^{k-1}|R_k} \left[e^{2h(x_k|S^{k-1}=s^{k-1})} \right] \\ & \geq |\lambda|^2 \mathbb{E}_{S^{k-1}|R_k} \left[e^{2h(x_{k-1}|S^{k-1}=s^{k-1})} \right] + \beta \\ & \geq \mathbb{E}_{S^{k-2}, R_{k-1}|R_k} \left[\frac{|\lambda|^2}{2^{2R_{k-1}}} e^{2h(x_{k-1}|S^{k-2}=s^{k-2})} \right] + \beta \\ & = \mathbb{E}_{R_{k-1}|R_k} \left[\frac{|\lambda|^2}{2^{2R_{k-1}}} \mathbb{E}_{S^{k-2}|R_{k-1}, R_k} \left[e^{2h(x_{k-1}|S^{k-2}=s^{k-2})} \right] \right] \\ & \quad + \beta. \end{aligned} \quad (44)$$

Substituting (44) into (43) and re-iterating k times, it follows that $\mathbb{E}_{S^k} [e^{2h(x_{k+1}|S^k=s^k)}]$ is lower bounded by

$$\begin{aligned} & \mathbb{E}_{R_{k-1}, R_k} \left(\frac{|\lambda|^4}{2^{2(R_{k-1}+R_k)}} \mathbb{E}_{S^{k-2}|R_{k-1}, R_k} \left(e^{2h(x_{k-1}|S^{k-2}=s^{k-2})} \right) \right) \\ & \quad + \beta \left(1 + \mathbb{E}_{R_k} \left(\frac{|\lambda|^4}{2^{2R_k}} \right) \right) \\ & \geq \mathbb{E}_{R_1, \dots, R_k} \left(\frac{|\lambda|^{2k}}{2^{2(R_1+\dots+R_k)}} \mathbb{E}_{S_1|R_1, \dots, R_k} \left(e^{2h(x_1|S_0=s_0)} \right) \right) \\ & \quad + \beta \left(1 + \sum_{j=2}^k \mathbb{E}_{R_1, \dots, R_k} \left(\frac{|\lambda|^{2(k-j+1)}}{2^{2(R_j+\dots+R_k)}} \right) \right) \\ & = \mathbb{E} \left(\frac{|\lambda|^{2(k+1)}}{2^{2(R_1+\dots+R_k)}} \right) e^{2h(x_0)} + \beta \left(1 + \sum_{j=1}^k \mathbb{E} \left(\frac{|\lambda|^{2(k-j+1)}}{2^{2(R_j+\dots+R_k)}} \right) \right), \end{aligned} \quad (45)$$

where (45) uses the fact that the initial condition of the state x_0 is independent of the rate process R_k . Let $\{z_k\}$ be a non-homogeneous MJLS

$$z_{k+1} = |\lambda|^2 / 2^{2R_k} z_k + \beta, \quad (47)$$

with $z_0 = e^{h(x_0)}$. By taking the expectation on both sides of (47) and iterating k times, it is easy to see that the right hand side of (46) is equal to the first moment of z_{k+1} . Hence, combining (43)–(46) we conclude that $\mathbb{E}(|x_k|^2) > \frac{1}{2\pi e} \mathbb{E}(z_k)$, which is the claim.

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 $\epsilon > 0$,

$$\begin{aligned} \mathbb{E}[e_{j\tau}^2] &= \mathbb{E}[e_{j\tau}^2 (1_{\{|e_{j\tau}| > \xi_j\}} + 1_{\{|e_{j\tau}| \leq \xi_j\}})] \\ &\leq \mathbb{E}[e_{j\tau}^{2+\epsilon} \xi_j^{-\epsilon} + \xi_j^2] \triangleq \theta_j \end{aligned} \quad (48)$$

where ξ_j denotes the (non-negative) scaling factor (21) used at the coder and the decoder. Thus, the second-moment stability of the sequence $\{\theta_j\}$ 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

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

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

First, observe that iterating (17) and making use of (20) and (22)

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

where

$$\eta_j = \sum_{i=0}^{\tau-1} \lambda^{\tau-1-i} L(y_{j\tau+i} - \bar{x}_{j\tau+i}). \quad (46)$$

Observe that the error $\varepsilon_k = x_k - \bar{x}_k$ of the Luenberger observer (17) evolves as $\varepsilon_{k+1} = (\lambda - L)\varepsilon_k - Lw_k + v_k$ and has bounded $(2 + \epsilon)$ th moment since $|\lambda - L| < 1$ and w_k and v_k have bounded $(2 + \epsilon)$ th moment by assumption A1. This implies that $y_k - \bar{x}_k$ has bounded $(2 + \epsilon)$ th moment since $|y_k - \bar{x}_k|^{2+\epsilon} \leq (|x_k - \bar{x}_k| + |w_k|)^{2+\epsilon} \leq 2^{1+\epsilon}(|x_k - \bar{x}_k|^{2+\epsilon} + |w_k|^{2+\epsilon})$, and hence η_j has bounded $(2 + \epsilon)$ th moment.

Then, combining with (19) we obtain the following recursive expression for the error

$$\begin{aligned} e_{(j+1)\tau} &= \bar{x}_{(j+1)\tau} - \hat{x}_{(j+1)\tau} \\ &= \lambda^\tau \left(e_{j\tau} - \xi_j q_{R(j)} \left(\frac{e_{j\tau}}{\xi_j} \right) \right) + \eta_j. \end{aligned} \quad (47)$$

From (47), the inequality $(|a| + |b|)^{2+\epsilon} \leq 2^{1+\epsilon}(|a|^{2+\epsilon} + |b|^{2+\epsilon})$ 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 φ in (21) 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 $\xi_{j+1}^{-\epsilon}$ and taking the expectation,

$$\begin{aligned} & \mathbb{E} \left[e_{(j+1)\tau}^{2+\epsilon} \xi_{j+1}^{-\epsilon} \right] \\ & \leq 2^{1+\epsilon} \mathbb{E} \left[\left(|\lambda|^{\tau(2+\epsilon)} |e_{j\tau} - \xi_j q_{R(j)} \left(\frac{e_{j\tau}}{\xi_j} \right)|^{2+\epsilon} \xi_j^{-\epsilon} \right. \right. \\ & \quad \left. \left. + \eta_j^{2+\epsilon} \xi_{j+1}^{-\epsilon} \right) \right] \\ & \leq 2^{1+\epsilon} \mathbb{E} \left[\left(|\lambda|^{\tau(2+\epsilon)} |e_{j\tau} - \xi_j q_{R(j)} \left(\frac{e_{j\tau}}{\xi_j} \right)|^{2+\epsilon} \right. \right. \\ & \quad \left. \left. \times (\xi_j |\lambda|^\tau \kappa_{R(j)} \left(\frac{e_{j\tau}}{\xi_j} \right))^{-\epsilon} + \varphi^2 \right) \right], \end{aligned} \quad (48)$$

Similarly, by (21),

$$\mathbb{E}[\xi_{j+1}^2] \leq 2^{1+\epsilon} \mathbb{E} \left[\left((\xi_j |\lambda|^\tau \kappa_{R(j)} \left(\frac{e_{j\tau}}{\xi_j} \right))^2 + \varphi^2 \right) \right]. \quad (49)$$

Adding (48) and (49), and multiplying both sides by $1_{\{R_{(j+1)\tau}=r_{i\tau}\}}$, we obtain the bound (52), at the top of the next page, on the dynamics of $\theta_{j,i}$.

In the derivation of (52), step (a) follows from the law of total probability, (b) is justified by the fact that

$$(e_{j\tau}, \xi_j, R(j)) \rightarrow (R_{j\tau}, \dots, R_{(j+1)\tau-1}) \rightarrow R_{(j+1)\tau}$$

form a Markov chain, and (c) follows from the fact that $(e_{j\tau}, \xi_j) \rightarrow R_{j\tau} \rightarrow (R_{j\tau+1}, \dots, R_{(j+1)\tau-1})$ form a Markov chain and the following lemma, which bounds the achievable

$$\begin{aligned}
\theta_{j+1,i_\tau} &= \mathbb{E} \left[\left(e_{(j+1)\tau}^{2+\epsilon} \xi_{j+1}^{-\epsilon} + \xi_{j+1}^2 \right) \mathbf{1}_{\{R_{(j+1)\tau} = r_{i_\tau}\}} \right] \\
&\leq 2^{1+\epsilon} |\lambda|^{2\tau} \mathbb{E} \left[\left(|e_{j\tau} - \xi_j q_{R(j)} \left(\frac{e_{j\tau}}{\xi_j} \right)|^{2+\epsilon} (\xi_j \kappa_{R(j)} \left(\frac{e_{j\tau}}{\xi_j} \right))^{-\epsilon} + (\xi_j \kappa_{R(j)} \left(\frac{e_{j\tau}}{\xi_j} \right))^2 \right) \mathbf{1}_{\{R_{(j+1)\tau} = r_{i_\tau}\}} \right] \\
&\quad + 2^{1+\epsilon} \varphi^2 \mathbb{E} \left[\mathbf{1}_{\{R_{(j+1)\tau} = r_{i_\tau}\}} \right], \\
&\stackrel{(a)}{=} 2^{1+\epsilon} |\lambda|^{2\tau} \sum_{i_0, \dots, i_{\tau-1}} \mathbb{E} \left[\left(|e_{j\tau} - \xi_j q_{R(j)} \left(\frac{e_{j\tau}}{\xi_j} \right)|^{2+\epsilon} (\xi_j \kappa_{R(j)} \left(\frac{e_{j\tau}}{\xi_j} \right))^{-\epsilon} + (\xi_j \kappa_{R(j)} \left(\frac{e_{j\tau}}{\xi_j} \right))^2 \right) \right. \\
&\quad \times \mathbf{1}_{\{R_{(j+1)\tau} = r_{i_\tau}, R_{(j+1)\tau-1} = r_{i_{\tau-1}}, \dots, R_{j\tau} = r_{i_0}\}} \left. + 2^{1+\epsilon} \varphi^2 \mathbb{E} \left[\mathbf{1}_{\{R_{(j+1)\tau} = r_{i_\tau}\}} \right] \right], \\
&\stackrel{(b)}{=} 2^{1+\epsilon} |\lambda|^{2\tau} \sum_{i_0, \dots, i_{\tau-1}} \mathbb{E} \left[\left(|e_{j\tau} - \xi_j q_{R(j)} \left(\frac{e_{j\tau}}{\xi_j} \right)|^{2+\epsilon} (\xi_j \kappa_{R(j)} \left(\frac{e_{j\tau}}{\xi_j} \right))^{-\epsilon} \right. \right. \\
&\quad \left. \left. + (\xi_j \kappa_{R(j)} \left(\frac{e_{j\tau}}{\xi_j} \right))^2 \right) \mathbf{1}_{\{R_{(j+1)\tau-1} = r_{i_{\tau-1}}, \dots, R_{j\tau} = r_{i_0}\}} \right] \mathbb{P}\{R_{j\tau} = r_{i_0}\} p_{i_0, i_1} \cdots p_{i_{\tau-1}, i_\tau} \\
&\quad + 2^{1+\epsilon} \varphi^2 \mathbb{E} \left[\mathbf{1}_{\{R_{(j+1)\tau} = r_{i_\tau}\}} \right], \\
&\stackrel{(c)}{\leq} 2^{1+\epsilon} |\lambda|^{2\tau} \sum_{i_0, \dots, i_{\tau-1}} \mathbb{E} \left[\frac{e_{j\tau}^{2+\epsilon} \xi_j^{-\epsilon} + \xi_j^2}{2^{2(r_{i_0} + \dots + r_{i_{\tau-1}})}} \mathbf{1}_{\{R_{j\tau} = r_{i_0}\}} \right] \mathbb{P}\{R_{j\tau} = r_{i_0}\} p_{i_0, i_1} \cdots p_{i_{\tau-1}, i_\tau} + 2^{1+\epsilon} \varphi^2 \mathbb{E} \left[\mathbf{1}_{\{R_{(j+1)\tau} = r_{i_\tau}\}} \right], \\
&\leq 2^{1+\epsilon} \zeta \sum_{i_0} \left(\sum_{i_1, \dots, i_{\tau-1}} \frac{|\lambda|^{2\tau}}{2^{2(r_{i_0} + \dots + r_{i_{\tau-1}})}} p_{i_0, i_1} \cdots p_{i_{\tau-1}, i_\tau} \right) \theta_{j, i_0} + 2^{1+\epsilon} \varphi^2 \mathbb{E} \left[\mathbf{1}_{\{R_{(j+1)\tau} = r_{i_\tau}\}} \right].
\end{aligned} \tag{52}$$

distortion with the quantizer defined in [19]:

Lemma 6: [19, Lemma 5.2] Let e and $\xi > 0$ be random variables with $\mathbb{E}[e^{2+\epsilon}]$ 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

$$\begin{aligned}
\mathbb{E} \left[\left(e - \xi q_r \left(\frac{e}{\xi} \right) \right)^{2+\epsilon} (\xi \kappa_r \left(\frac{e}{\xi} \right))^{-\epsilon} + (\xi \kappa_r \left(\frac{e}{\xi} \right))^2 \right] \\
\leq \frac{\zeta}{2^{2r}} \mathbb{E} \left[e^{2+\epsilon} \xi^{-\epsilon} + \xi^2 \right]
\end{aligned}$$

for some constant $\zeta > 2$ only determined by ϵ and δ .

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

$$\theta_{j+1, i} \leq \mathbb{E} \left[z_{(j+1)\tau}^2 \mathbf{1}_{\{R_{(j+1)\tau} = r_i\}} \right], \quad r_i \in \mathcal{R}, \tag{50}$$

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

$$z_{j\tau} = \phi \frac{|\lambda|^\tau}{2^{R_{(j-1)\tau} + \dots + R_{j\tau-1}}} z_{(j-1)\tau} + \varsigma, \quad j \geq 1, \tag{51}$$

with $\phi = \sqrt{2^{1+\epsilon} \zeta} > 1$ and $\varsigma = \sqrt{2^{1+\epsilon} \varphi^2} > 0$. 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 $r_{i_\tau} \in \mathcal{R}$,

$$\begin{aligned}
&\mathbb{E} \left[z_{(j+1)\tau}^2 \mathbf{1}_{\{R_{(j+1)\tau} = r_{i_\tau}\}} \right] \\
&= \mathbb{E} \left[\left(\phi \frac{|\lambda|^\tau}{2^{R_{j\tau} + \dots + R_{(j+1)\tau-1}}} z_{j\tau} + \varsigma \right)^2 \mathbf{1}_{\{R_{(j+1)\tau} = r_{i_\tau}\}} \right] \\
&\geq \mathbb{E} \left[\left(\phi \frac{|\lambda|^\tau}{2^{R_{j\tau} + \dots + R_{(j+1)\tau-1}}} z_{j\tau} \right)^2 \mathbf{1}_{\{R_{(j+1)\tau} = r_{i_\tau}\}} \right] \\
&\quad + \varsigma^2 \mathbb{E} \left[\mathbf{1}_{\{R_{(j+1)\tau} = r_{i_\tau}\}} \right] \\
&= \phi^2 \sum_{i_0, \dots, i_{\tau-1}} \frac{|\lambda|^{2\tau} p_{i_0, i_1} \cdots p_{i_{\tau-1}, i_\tau}}{2^{2(r_{i_0} + \dots + r_{i_{\tau-1}})}} \mathbb{E} \left[z_{j\tau}^2 \mathbf{1}_{\{R_{j\tau} = r_{i_0}\}} \right]
\end{aligned}$$

$$\begin{aligned}
&+ \varsigma^2 \mathbb{E} \left[\mathbf{1}_{\{R_{(j+1)\tau} = r_{i_\tau}\}} \right] \\
&\geq \phi^2 \sum_{i_0, \dots, i_{\tau-1}} \frac{|\lambda|^{2\tau}}{2^{2(r_{i_0} + \dots + r_{i_{\tau-1}})}} p_{i_0, i_1} \cdots p_{i_{\tau-1}, i_\tau} \theta_{j, i_0} + \\
&\quad \varsigma^2 \mathbb{E} \left[\mathbf{1}_{\{R_{(j+1)\tau} = r_{i_\tau}\}} \right] \\
&\geq \theta_{j+1, i_\tau}
\end{aligned}$$

where the first inequality follows from the fact that $(a+b)^2 \geq a^2 + b^2$ for all nonnegative numbers a and b , the second inequality uses the induction hypothesis, while the last inequality uses (52). Hence, the claim holds at time $k+1$ as well.

Summing both sides of (50) over $r_i \in \mathcal{R}$ and making use of (48), it follows that

$$\mathbb{E} \left[e_{j\tau}^2 \right] \leq \mathbb{E} \left[z_{j\tau}^2 \right] \tag{53}$$

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 $z_{j\tau} \leq w_{j\tau}$ for all j , where $w_\tau, w_{2\tau}, \dots$, are defined as the evolution at times $\tau, 2\tau, \dots$ of the non-homogeneous MJLS with dynamics

$$w_k = \phi^{1/\tau} \frac{|\lambda|}{2^{R_{k-1}}} w_{k-1} + c, \quad k \geq 1,$$

for $c = \varsigma \sum_{i=0}^{\tau-1} (|\lambda| \phi^{1/\tau})^i$. By Lemma 1, this MJLS is second-moment stable if

$$\phi^{2/\tau} |\lambda|^2 \rho(H) < 1. \tag{54}$$

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 τ 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, \dots, H_m be positive square matrices. Let x_i denote the eigenvector associated to the Perron eigenvalue of H_i , i.e., such that $H_i x_i = \rho(H_i) x_i$, $i = 1, \dots, m$. Then, for any permutation σ of $\{1, \dots, m\}$ we have

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

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

By the Perron-Frobenius theorem all the components of the Perron eigenvector x_i are positive, $i = 1, \dots, m$, 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 Lemma 7, we have

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

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., $H_i^k x_i = (\rho(H_i))^k x_i$, $i = 1, \dots, m$. Then, by Lemma 7

$$c_{\min}^{\frac{1}{k}} \leq \frac{(\rho(H_{\sigma_1}^k H_{\sigma_2}^k \cdots H_{\sigma_m}^k))^{\frac{1}{k}}}{\rho(H_1) \rho(H_2) \cdots \rho(H_m)} \leq c_{\max}^{\frac{1}{k}}, \quad (57)$$

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

Equipped with the above lemmas, we can proceed to the proof of Proposition 4.

Proof: Let $m, m' \geq 1$, $\alpha_1, \dots, \alpha_m, \alpha'_1, \dots, \alpha'_{m'} \in \mathcal{A}$. To prove the claim it suffices to show that for any rate tuple $(\log |\lambda_1|, \dots, \log |\lambda_u|)$ strictly inside the convex closure of

$$\mathcal{C}(m, \alpha_1, \dots, \alpha_m) \cup \mathcal{C}(m', \alpha'_1, \dots, \alpha'_{m'})$$

there exist an $m'' \geq 1$ such that

$$(\log |\lambda_1|, \dots, \log |\lambda_u|) \in \mathcal{C}(m'', \alpha''_1, \dots, \alpha''_{m''})$$

for some $\alpha''_1, \dots, \alpha''_{m''} \in \mathcal{A}$.

For every $i \in \mathcal{U}$, 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 $(\log |\lambda_1|, \dots, \log |\lambda_u|)$ strictly inside the convex closure of

$$\mathcal{C}(m, \alpha_1, \dots, \alpha_m) \cup \mathcal{C}(m', \alpha'_1, \dots, \alpha'_{m'})$$

satisfies

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

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 $i \in \mathcal{U}$

$$\begin{aligned} & -\frac{1}{2kmm'} \log \left(\rho(H_i^{\gamma km'} H_i'^{(1-\gamma)km}) \right) \\ & \geq -\frac{\gamma}{2m} \log \rho(H_i) - \frac{1-\gamma}{2m'} \log \rho(H'_i) - \epsilon' \\ & > -\frac{\gamma}{2m} \log \rho(H_i) - \frac{1-\gamma}{2m'} \log \rho(H'_i) - \epsilon \\ & \geq \log |\lambda_i|. \end{aligned}$$

This shows that the rate tuple $(\log |\lambda_1|, \dots, \log |\lambda_u|)$ can be stabilized by the control scheme in Proposition 4 because it is contained inside the hypercube $\mathcal{C}(kmm', \alpha''_1, \dots, \alpha''_{kmm'})$, where the sequence of bit allocation functions $\{\alpha''_1, \dots, \alpha''_{kmm'}\}$ is obtained by concatenating $\gamma km'$ copies of $\alpha_1, \dots, \alpha_m$ followed by $(1-\gamma)km$ copies of the bit allocation functions $\alpha'_1, \dots, \alpha'_{m'}$. ■

REFERENCES

- [1] J. H. Braslavsky, R. H. Middleton, and J. S. Freudenberg, "Feedback stabilization over signal-to-noise ratio constrained channels," *IEEE Transaction on Automatic Control*, vol. 52, no. 8, pp. 1391–1403, Aug. 2007.
- [2] R.W. Brockett and D. Liberzon, "Quantized feedback stabilization of linear systems," *IEEE Transaction on Automatic Control*, vol. 45, no.7, pp. 1279–1289, Jul. 2000.
- [3] O.L.V. Costa, M.D. Fragoso, and R.P. Marques, *Discrete-Time Markov Jump Linear Systems*, Springer, 2005.
- [4] D.F. Delchamps, "Stabilizing a linear system with quantized state feedback," *IEEE Transaction on Automatic Control*, vol. 35, no. 8, pp. 916–924, Aug. 1990.
- [5] A. El Gamal, and Y.H. Kim, *Network Information Theory*, Cambridge University Press, 2011.
- [6] V. Gupta, B. Hassibi and R.M. Murray, "Optimal LQG control across packet-dropping links," *Systems and Control Letters*, vol. 56, no. 6, pp. 439–446, Jun. 2007.
- [7] V. Gupta, N.C. Martins, and J.S. Baras, "Optimal Output Feedback Control Using Two Remote Sensors Over Erasure Channels", *IEEE Transactions on Automatic Control*, vol. 54, no. 7, pp. 1463–1476, July 2009.
- [8] J.P. Hespanha, P. Naghshtabrizi and Y. Xu, "A survey of recent results in networked control systems," *Proc. IEEE, Special issue on emerging technologies of networked control systems*, vol. 95, no. 1, pp. 138–162, Jan. 2007.
- [9] R.A. Horn, and C.R. Johnson, *Matrix Analysis*, Cambridge, UK: Cambridge University Press, 1985.
- [10] M. Huang, S. Dey, "Stability of Kalman filtering with Markov packet losses", *Automatica*, vol. 43, no. 4, pp. 598–607, Apr. 2007.
- [11] O. Imer, S. Yuksel, and T. Basar, "Optimal control of LTI systems over unreliable communication links," *Automatica*, vol. 42, pp. 1429–1439, 2006.
- [12] C. R. Johnson, R. Bru, "The spectral radius of a product of nonnegative matrices," *Linear Algebra and its Applications*, Vol. 141, pp 227–240, Nov. 1990.
- [13] P. Kawka, and A. Alleyne, "Stability and performance of packet-based feedback control over a Markov channel," *Proceedings of IEEE American Control Conference*, pp. 2807–2812, 2006.
- [14] P.R. Kumar, "Foundations of Cyberphysical Systems", *National CPS PI Meeting*, 2010.
- [15] D. Liberzon, "On stabilization of linear systems with limited information," *IEEE Transaction on Automatic Control*, vol. 48, no. 2, pp 304–307, Feb. 2003.
- [16] J. Liu, N. Elia, S. Tatikonda, "Capacity-achieving feedback scheme for Gaussian finite-state Markov channels with channel state information", submitted to *IEEE Transaction on Information Theory*, 2008,
- [17] N.C. Martins, M.A. Dahleh and N. Elia, "Feedback stabilization of uncertain systems in the presence of a direct link," *IEEE Transaction on Automatic Control*, vol. 51, no.3, pp. 438–447, Mar. 2006.

- [18] P. Minero, M. Franceschetti, S. Dey and G.N. Nair, "Data rate theorem for stabilization over time-varying feedback channels," *IEEE Transaction on Automatic Control*, vol. 54, no. 2, pp.243–255, Feb. 2009.
- [19] G.N. Nair and R.J. Evans, "Stabilizability of stochastic linear systems with finite feedback data rates," *SIAM Journal on Control and Optimization*, vol. 43, no. 2, pp. 413–436, Jul 2004.
- [20] G.N. Nair, F. Fagnani, S. Zampieri and R.J. Evans, "Feedback control under data rate constraints: an overview," *Proc. IEEE, Special issue on emerging technologies of networked control systems*, vol. 95, no. 1, pp. 108–137, Jan. 2007.
- [21] A. Sahai, S. Avestimehr, and P. Minero, "Anytime communication over the Gilbert-Eliot channel with noiseless feedback", in *Proc. IEEE International Symposium on Information Theory*, pp. 1783–1787, Sept. 2005.
- [22] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla and S.S. Sastry, "Foundations of control and estimation over lossy networks," *Proc. IEEE, Special issue on emerging technologies of networked control systems*, vol. 95, no. 1, pp.163–187, Jan. 2007.
- [23] P. Seiler, R. Sengupta, "Analysis of communication losses in vehicle control problems," *Proceedings of IEEE American Control of Conference*, pp. 1491–1496, 2001.
- [24] C.E. Shannon, "A mathematical theory of communication," *ACM SIG-MOBILE Mobile Computing and Communications Review*, vol. 5, no. 1, pp. 3–55, 2001.
- [25] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M.I. Jordan, and S.S. Sastry, "Kalman filtering with intermittent observations," *IEEE Transaction on Automatic Control*, vol. 49, no. 9, pp. 1453–1464, Sept. 2004.
- [26] S. Tatikonda and S. Mitter, "Control under communication constraints," *IEEE Transaction on Automatic Control*, vol. 49, no. 7, pp. 1056–1068, Jul. 2004.
- [27] S. Tatikonda and S. Mitter, "Control over noisy channels," *IEEE Transaction on Automatic Control*, vol. 49, no. 7, pp. 1196–1201, Jul. 2004.
- [28] W. S. Wong and R. W. Brockett, "Systems with finite communication bandwidth constraints-II: Stabilization with limited information feedback," *IEEE Transaction on Automatic Control*, vol. 44, no. 5, May 1999.
- [29] K. You, and L. Xie, "Minimum Data Rate for Mean Square Stabilizability of Linear Systems with Markovian Packet Losses," *IEEE Transaction on Automatic Control*, vol. 56, no. 4, pp. 772–785, April 2011.
- [30] K. You, and L. Xie, "Minimum Data Rate for Mean Square Stabilization of Discrete LTI Systems over Lossy Channel", *IEEE Transaction on Automatic Control*, vol. 55, no. 10, pp.2373-2378, October 2010.
- [31] S. Yüksel, "Stochastic Stabilization of Noisy Linear Systems With Fixed-Rate Limited Feedback," *IEEE Transaction on Automatic Control*, vol. 55, no. 12, pp. 2847–2853, December 2010.



Paolo Minero (M'11) received the Laurea degree (with highest honors) in electrical engineering from the Politecnico di Torino, Torino, Italy, in 2003, the M.S. degree in electrical engineering from the University of California at Berkeley in 2006, and the Ph.D. degree in electrical engineering from the University of California at San Diego in 2010. He is an Assistant Professor in the Department of Electrical Engineering at the University of Notre Dame. Before joining the University of Notre Dame, he was a postdoctoral scholar at the University of California at San Diego for six month. His research interests are in communication systems theory and include information theory, wireless communication, and control over networks. Dr. Minero received the U.S. Vodafone Fellowship in 2004 and 2005, and the Shannon Memorial Fellowship in 2008.



Massimo Franceschetti (M'98–SM'11) received the Laurea degree (magna cum laude) in computer engineering from the University of Naples, Naples, Italy, in 1997, and the M.S. and Ph.D. degrees in electrical engineering from the California Institute of Technology, Pasadena, CA, in 1999, and 2003, respectively. He is an Associate Professor in the Department of Electrical and Computer Engineering, University of California at San Diego (UCSD). Before joining UCSD, he was a postdoctoral scholar at the University of California at Berkeley for two years. He has held visiting positions at the Vrije Universiteit Amsterdam, the Ecole Polytechnique Federale de Lausanne, and the University of Trento. His research interests are in communication systems theory and include random networks, wave propagation in random media, wireless communication, and control over networks. Dr. Franceschetti is an Associate Editor for Communication Networks of the *IEEE TRANSACTIONS ON INFORMATION THEORY* (2009/2012) and has served as Guest Editor for two issues of the *IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATION*. He was awarded the C. H. Wilts Prize in 2003 for best doctoral thesis in electrical engineering at Caltech; the S.A. Schelkunoff Award in 2005 for best paper in the *IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION*; a National Science Foundation (NSF) CAREER award in 2006, an ONR Young Investigator Award in 2007; and the IEEE Communications Society Best Tutorial Paper Award in 2010.



Lorenzo Coviello (S'11) received the Laurea degree (with highest honors) in telecommunications engineering from the University of Padova, Padova, Italy, in 2006. He is currently pursuing a Ph.D. degree in Electrical and Computer Engineering at the University of California at San Diego, under the advice of Prof. Massimo Franceschetti. His research interests are in control systems, social networks, and distributed algorithms.