Infimum Data Rates for Stabilising Markov Jump Linear Systems

G. N. Nair, S. Dey and R. J. Evans
University of Melbourne
Dept. Electrical & Electronic Engineering
VIC 3010, Australia
{g.nair, s.dey, r.evans}@ee.mu.oz.au

Abstract—In recent years, the problem of stabilising linear dynamical systems with low feedback data rates has been intensively investigated. A particular focus has been the characterisation of the infimum data rate for stabilisability, which specifies the smallest rate, in bits per unit time, at which information can circulate in a stable feedback loop. This paper extends this line of research to the case of fully-observed, finite-dimensional, linear systems without process noise but with control-independent, Markov parameters. Unlike previous formulations, the coding alphabet is permitted to be random and time-varying via a possible dependence on the observed Markov modes. Using quantisation techniques and real Jordan forms, it is shown that the smallest asymptotic mean data rate for stabilisability in r-th absolute output moment, over all coding and control schemes, is given by an exponent which measures the asymptotic mean growth rate of unstable eigenspace volumes. An explicit formula for it is obtained in the case of antistable dynamics. For scalar systems, this expression is quite different from an earlier one derived assuming a constant alphabet, in particular being independent of the moment order.

I. INTRODUCTION

One of the issues that has begun to emerge in a number of application areas, ranging from micro-electromechanical systems to large-scale irrigation networks, is how to control a system consisting of numerous components without consuming excessive communication resources. Although the total communication capacity may be large, the size of such systems means that the bit rate available *per component* may be small and thus have a large negative effect on control loop performance, particularly with unstable plants. This has motivated a considerable amount of research on the interrelationship between communications and feedback control.

For the most part, the focus has been on the the basic problem in which a single dynamical system transmits feedback information to a controller over a noiseless digital link with finite data rate. In recent years significant progress has been made with regard to this set-up, particularly for linear systems. Beginning with [5] and continuing with [19], [1], [4], [2], [3], [6], [18], [17], [13], [15], [14], [16], [9], [10], [11], [7], a number of different coding and control schemes have been proposed and analysed. A feature of many of these articles (e.g. [5], [19], [2], [17], [13], [15], [14], [7]) was the derivation of tight data rate bounds for stabilisability. At data rates below these bounds stability is impossible to achieve by any means, but above them it becomes possible to construct coding and control schemes that stabilise the

plant in various senses. For the case of finite-dimensional LTI systems, despite significant differences in formulation this lower bound is in all cases equal to the sum of the base-2 logarithms of the unstable open-loop eigenvalue magnitudes, underlining the fundamental nature of this quantity.

In this paper, we extend this line of research to the class of fully-observed, finite-dimensional, noiseless linear systems governed by Markov parameters, often encountered in telecommunications, manufacturing and defence. As in [15], our objective is to characterise the infimum data rate which permits stability in rth absolute output moment. *A priori*, we may expect this rate to reflect the level of open-loop instability, as quantified by some exponent that measures the 'rate of growth' of the system. However, what distinguishes this situation from the LTI case is that, due to the random parameters, there are many, non-equivalent ways to define growth rates. We demonstrate here that the one which is appropriate for the purpose of data-rate-limited stabilisation governs the evolution of unstable subspace volumes.

In the next section, the problem is formulated precisely and the main result, Theorem 2.1, is stated. Unlike our preliminary results in [12] on scalar Markov jump linear systems, we permit the coding alphabet to be stochastic here, via a possible dependence on the observed Markov modes. This extra latitude leads to a more natural expression for the infimum data rate, one which does not display the somewhat counter-intuitive r-dependence exhibited by the expression derived in [12]. In section III, we exploit a quantisation argument to show that all stabilising schemes must operate at or above this data rate. We then construct a coding and control scheme in the subsequent section, and prove that it achieves exponential stability at data rates arbitrarily close to the lower bound.

II. FORMULATION AND STATEMENT OF RESULT

Before formulating the problem, we define some of the notation used. Sequences $\{a_j\}_{j=0}^k$ are denoted $\bar{a}_k$, with $\bar{a}_{-1}$ denoting the null sequence. The $m \times n$ zero matrix is written $0_{m \times n}$, the $m \times n$ matrix with 1's on the upper left corner diagonal and 0's elsewhere is written $I_{m \times n}$, and $\sigma(\cdot)$ denotes the spectrum of a square matrix . The symbol $\| \cdot \|$ represents the Euclidean norm and $\|p\|_\theta$ the $\mathscr{L}_\theta$-norm $\left(\int \|p(\mathbf{x})\|^\theta d\lambda(\mathbf{x})\right)^{1/\theta}$ of a function p with respect to Lebesgue measure λ. Random variables are written in uppercase and their realisations are indicated by the corresponding

lower-case letters. All random variables are assumed to exist in a common probability space, the probability density of a random variable $\mathbf{X}$ w.r.t. Lebesgue measure is written $p_{\mathbf{X}}$, $P\{S = s\}$ denotes the probability of an event $S = s$ and E_s represents conditional expectation with respect to an event $S = s$.

Consider the noiseless, fully observed jump linear system

$$\mathbf{x}_{k+1} = \mathbf{A}(z_k)\mathbf{x}_k + \mathbf{B}(z_k)\mathbf{u}_k, \quad \forall k \in \mathbb{Z}_+, \quad (1)$$

where $\mathbf{x}_k \in \mathbb{R}^n$ is the output, $\mathbf{u}_k \in \mathbb{R}^m$ the control signal and $z_k \in \mathscr{Z}_k[1, 2, \ldots, \zeta]$ the discrete mode at time k. We assume that

A1 $\{z_k\}_{k \in \mathbb{Z}_+}$ is a realisation of a control-independent Markov chain s.t.

$$P\{Z_{k+1} = i | Z_k = j\} \overset{\Delta}{=} t_{ij} > 0, \forall i, j \in [1, \ldots, \zeta], k \in \mathbb{Z}_+,$$

where $\mathbf{T} \in \mathbb{R}_+^{\zeta \times \zeta}$ is the transition probability matrix with t_{ij} in the ith row and jth column.

A2 $(\mathbf{A}(z), \mathbf{B}(z))$ is controllable, $\forall z \in [1, \ldots, \zeta]$.

A3   The initial output $\mathbf{x}_0$ is a realisation of a random variable $\mathbf{X}_0$ which is independent of $Z_0, Z_1, \ldots$, possesses an absolutely continuous distribution on $\mathbb{R}^n$ and has $(r + \varepsilon)$th absolute moment $E\|\mathbf{X}\|^{r+\varepsilon} < \infty$ for some $r, \varepsilon > 0$.

Suppose the sensor observing the outputs and Markov modes of this plant is connected to the controller by a noiseless digital channel which can accept only one discrete-valued symbol s_k from a coding alphabet $\mathscr{S}_k(\tilde{z}_k)$ of $m_k \overset{\Delta}{=} \mu_k(\tilde{z}_k)$ elements during each sampling interval, where $E\{\log_2 M_k\} = o(k)$. The (asymptotic mean) data rate of the channel may then be defined as

$$R \overset{\Delta}{=} \liminf_{t \to \infty} \frac{1}{t} \sum_{k=0}^{t-1} E\{\log_2 M_k\}. \quad (2)$$

This definition is more general than the one adopted in previous work on scalar Markov jump linear systems [12], in that the alphabet sizes may be time-varying and also stochastic. As remarked in the Introduction, this extra freedom leads to a rather different expression for the infimum stabilising data rate.

Ignoring computational restrictions, each symbol the coder transmits may generally depend on all past and present outputs and modes. Hence the coding law can without loss of generality be written

$$s_k = \gamma_k(\tilde{\mathbf{x}}_k, \tilde{z}_k) \in \mathscr{S}_k(\tilde{z}_k), \quad \forall k \in \mathbb{Z}_+, \quad (3)$$

where γ_k is the (possibly time-varying) coder mapping at time k. Assuming a fixed transmission delay d and no channel errors, each symbol transmitted at time k is received by the controller at the other end of the channel at time $k + d$. Further assuming that the controller is also able to observe

the Markov modes directly, the signal it generates will most generally be of the form

$$\mathbf{u}_k = \upsilon_k(\tilde{s}_{k-d}, \tilde{z}_k) \in \mathbb{R}^m, \quad \forall k \in \mathbb{Z}_+, \quad (4)$$

where υ_k is the controller mapping at time k.

Define the *coder-controller* as the triple of alphabet, coder and controller mapping sequences $(\mathscr{S}_\infty, \bar{\gamma}_\infty, \bar{\upsilon}_\infty)$ and call the closed-loop plant stable in rth absolute output moment if

$$\sup_{k \in \mathbb{Z}_+} E\|\mathbf{X}_k\|^r < \infty, \quad (5)$$

where the expectation is over both the initial output and the Markov modes. The aim of this paper is to characterise the critical data rate above which a coder-controller exists which can stabilise the plant in the sense above, and below which no such coder-controller exists.

From previous work on linear systems, it is natural to expect this data rate to be determined by the rate at which the open-loop plant output grows. However, what distinguishes Markov jump linear systems is that their random dynamics lead to many, non-equivalent ways to define growth rates. For instance, one notion is the *top Lyapunov exponent*,

$$\lambda_1 \overset{\Delta}{=} \lim_{k \to \infty} k^{-1} E \ln \|\mathbf{A}(Z_{k-1})\mathbf{A}(Z_{k-2}) \cdots \mathbf{A}(Z_0)\|.$$

However, there is no *a priori* reason not to interchange the expectation and logarithm and raise the norm to some power, yielding different measures of open-loop growth rate. Clearly, the correct measure of rate of growth will depend on the problem at hand. The main result of this paper is stated below:

Theorem 2.1: Under assumptions A1–A3, any coder-controller (3)-(4) which stabilises the Markov jump linear plant (1) in the rth moment sense (5) must have a data rate (2) satisfying

$$R \geq \liminf_{k \to \infty} \frac{1}{k} E \left\{ \sum_{\eta \in \sigma(\mathbf{F}_k), |\eta| \geq 1} \log_2 |\eta| \right\} =: H, \quad (6)$$

where

$$\mathbf{F}_k \overset{\Delta}{=} \mathbf{A}(Z_{k-1}) \cdots \mathbf{A}(Z_0) \quad (7)$$

Furthermore, it is possible to construct coder-controllers that take $E\|\mathbf{X}\|^r \to 0$ exponentially fast at data rates arbitrarily close to the lower bound (6).

The quantity H above is similar to a Lyapunov exponent, but measures the asymptotic mean growth rate not of lengths but of unstable eigenspace volumes. By this theorem, H specifies the smallest rate at which data can circulate in a stable feedback loop for a Markov jump linear plant, over a very broad class of coding and control laws. More fundamentally, it can be interpreted as the *feedback entropy* of the plant, i.e. the rate at which the plant generates stability information. Note that it is independent of the notion of stability as specified by the moment order r.

For LTI plants, H reduces to the sum of all positive, base-2 Lyapunov exponents previously derived, but in general a formula for H is difficult to obtain. However, if all the matrices $\mathbf{A}(z)$ have eigenvalues on or outside the unit circle, a more explicit expression becomes possible. In this case, the eigenvalues of $\mathbf{F}_k$ are automatically never less than 1 in magnitude, so we can simply write

$$
\begin{aligned}
H &= \liminf_{k\to\infty} \frac{1}{k} \mathrm{E}\left\{ \sum_{\eta\in\sigma(\mathbf{F}_k)} \log_2|\eta| \right\}, \\
&\equiv \liminf_{k\to\infty} \frac{1}{k}\mathrm{E}\{\log_2|\det\mathbf{F}_k|\} = \liminf_{k\to\infty} \frac{1}{k}\sum_{h=0}^{k-1}\mathrm{E}\{\log_2|\det\mathbf{A}(Z_k)|\}, \\
&= \sum_{z=1}^{\zeta} \mathrm{P}\{Z^* = z\} \log_2|\det\mathbf{A}(z)|
\end{aligned}
$$

by Perron's Theorem [8], where Z^* is the stationary distribution obtained by solving $\mathbf{Tp} = \mathbf{p}$. For the special case of a scalar system with $|a(z)| \geq 1$, H is strictly smaller than the expression derived in [12], since we permit the coding alphabet to depend on the observed Markov modes here.

In the next section, a recent quantisation lower bound is applied to establish the necessity of (6). We demonstrate sufficiency in the subsequent section by constructing an exponentially stabilising coder-controller with data rate arbitrarily close to H.

III. NECESSITY OF DATA RATE BOUND

We prove the necessity of (6) by first showing that the stabilisation problem can be reduced to one of initial condition quantisation. A quantiser error moment inequality is then used to derive a lower bound on the rth absolute output moment.

Suppose that stability in the sense (5) is achieved by some coder-controller $(\mathscr{S}_\infty, \bar{\gamma}_\infty, \bar{\upsilon}_\infty)$. Let $\mathbf{V}_k$ be a real, orthonormal similarity transformation which converts $\mathbf{F}_k$ as defined in (7) into a real Jordan form $\mathbf{J}_k$ [8], i.e. $\mathbf{F}_k = \mathbf{V}_k^\mathrm{T} \mathbf{J}_k \mathbf{V}_k$. Without loss of generality, the diagonal blocks of $\mathbf{J}_k$ may be assumed to be ordered according to descending eigenvalue magnitudes. Given a mode sequence $\tilde{z}_{k-1}$, let the number of eigenvalues greater than or equal to 1 in magnitude be denoted by n_k, including multiplicity, and define $\mathbf{x}_k^\mathrm{u} \in \mathbb{R}^{n_k}$ by

$$
\mathbf{x}_k^\mathrm{u} \triangleq \mathbf{I}_{n_k \times n} \mathbf{V}_k \mathbf{x}_k \equiv \mathbf{V}_k^\mathrm{u} \mathbf{x}_k. \tag{8}
$$

By the orthonormality of $\mathbf{V}_k$, $\|\mathbf{x}_k^\mathrm{u}\| \leq \|\mathbf{I}_{n_k \times n}\|\|\mathbf{V}_k\|\|\mathbf{x}_k\| = \|\mathbf{x}_k\|$, $\forall k \in \mathbb{Z}_+$, so $\mathrm{E}\|\mathbf{X}_k^\mathrm{u}\|^r$ is also uniformly bounded. Let

$$
\mathbf{y}_k \triangleq \mathbf{I}_{n_k \times n} \mathbf{V}_k \mathbf{x}_0 \in \mathbb{R}^{n_k} \tag{9}
$$

and observe that by the block diagonal structure of $\mathbf{J}_k$,

$$
\mathbf{V}_k^\mathrm{u} \mathbf{F}_k \mathbf{x}_0 = (\mathbf{I}_{n_k \times n} \mathbf{V}_k)(\mathbf{V}_k^\mathrm{T} \mathbf{J}_k \mathbf{V}_k)\mathbf{x}_0 = \mathbf{J}_k^\mathrm{u} \mathbf{y}_k, \tag{10}
$$

where $\mathbf{J}_k^\mathrm{u} \in \mathbb{R}^{n_k \times n_k}$ is the block diagonal matrix consisting of all the real Jordan blocks of $\mathbf{J}_k$ with eigenvalue greater than

or equal to 1 in magnitude. Expanding (1) out, multiplying by $\mathbf{V}_k^\mathrm{u}$ and using the fact above,

$$
\mathbf{x}_k^\mathrm{u} = \mathbf{J}_k^\mathrm{u} \mathbf{y}_k + \mathbf{V}_k^\mathrm{u} \sum_{j=0}^{k-1} \left(\mathbf{A}(z_{k-1})\cdots\mathbf{A}(z_{j+1})\right) \mathbf{B}(z_j)\upsilon_j(\tilde{s}_{j-d},\tilde{z}_j). \tag{11}
$$

For every mode sequence $\tilde{z}_{k-1}$, the sum on the RHS is a function of the symbol sequence $\tilde{s}_{k-d-1}$, which is in turn completely determined by $\mathbf{x}_0$ for a given coder-controller. Thus $\forall k \in \mathbb{Z}_+, \tilde{z}_{k-1}$, a function $q_{k-1}(\cdot, \tilde{z}_{k-1}): \mathbb{R}^n \to \mathbb{R}^{n_k}$ may be defined by

$$
\begin{aligned}
&q_{k-1}(\mathbf{x}_0, \tilde{z}_{k-1}) \\
&\triangleq -\mathbf{V}_k^\mathrm{u} \sum_{j=0}^{k-1}\left(\mathbf{A}(z_{k-1})\cdots\mathbf{A}(z_{j+1})\right)\mathbf{B}(z_j)\upsilon_j(\tilde{s}_{j-d},\tilde{z}_j),
\end{aligned}
$$

$\forall \mathbf{x}_0 \in \mathbb{R}^n$. As the symbol sequence $\tilde{s}_{k-d-1}$ can take

$$
\varphi_{k-d-1} \triangleq \prod_{j=0}^{k-d-1} m_j \tag{12}
$$

distinct values, $q_{k-1}(\cdot, \tilde{z}_{k-1})$ can only assume up to φ_{k-d-1} distinct values in $\mathbb{R}^n$, i.e. it can be viewed as a φ_{k-d-1}-point quantiser. Substituting this into (11) and taking rth moments,

$$
\mathrm{E}\|\mathbf{X}_k^\mathrm{u}\|^r = \mathrm{E}\left\|\mathbf{J}_k^\mathrm{u}\mathbf{Y}_k - q_{k-1}(\mathbf{X}_0, \tilde{Z}_{k-1})\right\|^r. \tag{13}
$$

The expression on the RHS resembles the standard mean rth power error (MrPE) criterion of quantisation theory. One complication is that the argument $\mathbf{x}_0$ of the quantiser $q_{k-1}(\cdot, \tilde{z}_{k-1})$ and the variable

$$
\mathbf{w}_k \triangleq \mathbf{J}_k^\mathrm{u}\mathbf{y}_k \tag{14}
$$

being quantised are different. However, for the purpose of obtaining a lower-bound, this is easily remedied. Denote the φ_{k-d-1} points of $q_{k-1}(\cdot, \tilde{z}_{k-1})$ by $\mathbf{t}_i \in \mathbb{R}^{n_k}$, $1 \leq i \leq \varphi_{k-d-1}$, and define the *nearest-neighbour* quantiser

$$
q_{k-1}^*(\mathbf{w}, \tilde{z}_{k-1}) \triangleq \arg\min_{\mathbf{t}_i} \|\mathbf{w} - \mathbf{t}_i\|, \ \forall \mathbf{w} \in \mathbb{R}^{n_k}.
$$

Consequently, $\|\mathbf{w} - q_{k-1}(\mathbf{x}_0, \tilde{z}_{k-1})\| \geq \|\mathbf{w} - q_{k-1}^*(\mathbf{w}, \tilde{z}_{k-1})\|$, $\forall \mathbf{w} \in \mathbb{R}^{n_k}, \mathbf{x}_0 \in \mathbb{R}^n$ and thus

$$
\mathrm{E}\|\mathbf{X}_k^\mathrm{u}\|^r \geq \mathrm{E}\|\mathbf{W}_k - q_{k-1}^*(\mathbf{W}_k, \tilde{Z}_{k-1})\|^r, \tag{15}
$$

a more familiar form of 'quantisation error'.

We now use the following inequality to lower-bound this term:

Lemma 3.1: Let $\mathbf{W} \in \mathbb{R}^\delta$ be a random variable with absolutely continuous distribution and $c_\nu : \mathbb{R}^\delta \to \mathbb{R}^\delta$ a quantiser with up to ν distinct values. Then $\forall \theta \in (\delta/[\delta+r], 1), \nu \in \mathbb{N}$,

$$
\mathrm{E}\|\mathbf{W} - c_\nu(\mathbf{W})\|^r \geq \frac{\beta}{\nu^{r/\delta}}\|p_\mathbf{W}\|_\theta^{r\theta/[\delta(1-\theta)]}, \tag{16}
$$

where $\beta(\delta) > 0$ is determined only by r, δ, and θ.

Proof: See [15].

1178

Using the absolute continuity of the distribution of $\mathbf{X}_0$ on $\mathbb{R}^n$, the mutual independence of $\tilde{Z}_{k-1}, \mathbf{X}_0$ and the invertibility of $\mathbf{V}_k$ and $\mathbf{J}_k^u$, it is straightforward to show that $\mathbf{W}_k \equiv \mathbf{J}_k^u \mathbf{I}_{n_k \times n} \mathbf{V}_k \mathbf{X}_0$ has an absolutely continuous distribution on $\mathbb{R}^{n_k}$, given the mode sequence $\tilde{z}_{k-1}$. We can thus apply the lemma above, with $\delta = n_k$, $p_{\mathbf{W}} = p_{\mathbf{W}_k | \tilde{z}_{k-1}}$, $v = \varphi_{k-d-1}$ and $c_v = q_{k-1}^*(\cdot, \tilde{z}_{k-1})$, to yield

$$\mathrm{E}_{\tilde{z}_{k-1}} \| \mathbf{W}_k - q_{k-1}^*(\mathbf{W}_k, \tilde{Z}_{k-1}) \|^r$$
$$\geq \frac{\beta(n_k)}{\varphi_{k-d-1}^{r/n_k}} \left(\int_{\mathbb{R}^{n_k}} p_{\mathbf{J}_k^u \mathbf{Y}_k | \tilde{z}_{k-1}}(\mathbf{w})^\theta \mathrm{d}\lambda(\mathbf{w}) \right)^{\frac{r}{n_k(1-\theta)}},$$
$$= \frac{\beta(n_k) |\det \mathbf{J}_k^u|^{r/n_k}}{\varphi_{k-d-1}^{r/n_k}} \left(\int_{\mathbb{R}^{n_k}} p_{\mathbf{Y}_k | \tilde{z}_{k-1}}(\mathbf{y})^\theta \mathrm{d}\lambda(\mathbf{y}) \right)^{\frac{r}{n_k(1-\theta)}} (17)$$

where the last equality follows from a change of integration variable. From the absolute continuity of the distribution of $\mathbf{X}_0$ on $\mathbb{R}^n$ given $\tilde{z}_{k-1}$ and the fact that $\mathbf{Y}_k = \mathbf{I}_{n_k \times n} \mathbf{V}_k \mathbf{X}_0$ with $\mathbf{V}_k$ orthonormal, it can be established that there is a constant $\rho > 0$ which lower-bounds the integral on the RHS, over all $k \in \mathbb{Z}_+$ and sequences $\tilde{z}_{k-1}$.

$$\Rightarrow \mathrm{E}_{\tilde{z}_{k-1}} \| \mathbf{W}_k - q_{k-1}^*(\mathbf{W}_k, \tilde{Z}_{k-1}) \|^r$$
$$\geq \frac{\beta(n_k) |\det \mathbf{J}_k^u|^{r/n_k}}{\varphi_{k-d-1}^{r/n_k}} \rho^{\frac{r}{n_k(1-\theta)}} \geq \psi \frac{|\det \mathbf{J}_k^u|^{r/n_k}}{\varphi_{k-d-1}^{r/n_k}},$$

where $\psi \triangleq \min_{\delta \in [1,...,n]} \beta(\delta) \rho^{\frac{r}{\delta(1-\theta)}}$. Taking expectations,

$$\mathrm{E} \| \mathbf{W}_k - q_{k-1}^*(\mathbf{W}_k, \tilde{Z}_{k-1}) \|^r \geq \psi \mathrm{E} \left\{ |(\det \mathbf{J}_k^u)/\Phi_{k-d-1}|^{r/N_k} \right\}. \quad (18)$$

The following lemma is now useful:

Lemma 3.2: Let $M \in [1,...,n]$ and $A > 0$ be random variables. If $\mathrm{E}\{A^{1/M}\} \in (0, b]$, then

$$\mathrm{E}\{\ln A\} \leq \ln b + n \ln \left(\frac{1-b}{1-b^{1/n}} \right). \quad (19)$$

Proof: Omitted. Briefly, it proceeds by first applying Jensen's inequality to an expectation conditioned on M, exploiting the fact that M is finite-valued, and then applying Jensen's inequality again.

As the LHS and hence RHS of (18) is uniformly bounded above by some $b > 0$, this lemma may be applied with $A = |(\det \mathbf{J}_k^u)/\Phi_{k-d-1}|^r$ and $M = N_k$ to yield

$$r \mathrm{E}\{\log_2 |\det \mathbf{J}_k^u| - \log_2 \Phi_{k-d-1}\} \leq c, \quad \forall k \in \mathbb{Z}_+,$$

where c is given by the RHS of (19). Hence $\forall k \in \mathbb{Z}_+$,

$$\mathrm{E}\{\log_2 \Phi_{k-d-1}\} = \sum_{j=0}^{k-d-1} \mathrm{E}\{\log_2 M_j\} \geq \mathrm{E}\{\log_2 |\det \mathbf{J}_k^u|\} - \frac{c}{r}$$

and thus

$$R = \liminf_{k \to \infty} \frac{1}{k} \sum_{j=0}^{k-d-1} \mathrm{E}\{\log_2 M_j\} \geq \liminf_{k \to \infty} \frac{1}{k} \mathrm{E}\{\log_2 |\det \mathbf{J}_k^u|\},$$

$$\equiv \liminf_{k \to \infty} \frac{1}{k} \mathrm{E} \left\{ \sum_{\eta \in \sigma(\mathbf{F}_k): |\eta| \geq 1} \log_2 |\eta| \right\}. \square$$

We remark that for exponential rth absolute moment stability, the inequality above must hold strictly. Note also that we have not used anywhere in the above the fact that $\{Z_k\}$ is a finite-state Markov chain, implying that (6) is necessary for any linear plant with random jumps.

IV. A STABILISING SCHEME

In this section, we prove sufficiency by constructing and analysing a coder-controller which exponentially stabilises the plant at a data rate that can be made arbitrarily close to the lower bound (6). For clarity we assume in this section that the initial condition is bounded by $\|\mathbf{x}_0\| \leq l_0/2$. This allows us to use simple, uniform quantisers to illustrate the main concepts and avoid complications originating from the tails of the output distribution. The argument below can in any case be extended to deal with infinite-support distributions with finite $(r+\varepsilon)$th absolute moments by the methods of [14].

As in the LTI case, the basic technique we use is to convert the plant into real Jordan form and quantise each subsystem component using a number of levels proportional to or greater than the magnitude of its governing eigenvalue. To average out the round-off arising from the integer codeword size, we 'down-sample' the system and encode the outputs once every τ instants. The chief distinction between this problem and the LTI situation is that the real Jordan form and hence the quantisation scheme depend on the Markov chain realisation, which changes stochastically with time. As neither coder nor controller know in advance what the modes are going to be, an additional delay is incorporated into the coder so that it waits $\tau - d$ intervals to observe the modes $z_{j\tau}, \ldots, z_{j\tau+\tau-d}$. It then calculates the real Jordan form corresponding to the product of the dynamical matrices over that period of time and uses it to encode the old output $\mathbf{x}_{j\tau}$. As the propagation time is d, the transmitted codeword is received by time $(j+1)\tau$. For large τ this extra delay will have a significant effect on transient performance, but nevertheless the rth absolute output moment will approach 0 exponentially fast.

A. Coder

Divide times $k \in \mathbb{Z}_+$ into cycles $[j\tau, \ldots, (j+1)\tau - 1]$, $j \in \mathbb{Z}_+$, of uniform integer duration $\tau \in \mathbb{N}$. At time $k = j\tau$, let $\hat{\mathbf{x}}_{j\tau} \in \mathbb{R}^n$ be an internal estimate of $\mathbf{x}_{j\tau}$ and $l_j/2 \in \mathbb{R}_+$ be an upper bound on the error norm $\|\mathbf{x}_{j\tau} - \hat{\mathbf{x}}_{j\tau}\|$. The *coder-controller state* is defined as $\psi_j \triangleq (\hat{\mathbf{x}}_{j\tau}, l_j)$, with $\psi_0 \triangleq (\mathbf{0}_{n \times 1}, l_0)$.

Given a channel propagation delay d, let

$$\mathbf{G}_j \triangleq \mathbf{A}(z_{(j+1)\tau-d}) \cdots \mathbf{A}(z_{j\tau}) \quad (20)$$

and $\mathbf{S}_j \in \mathbb{R}^{n \times n}$ be a real, orthonormal matrix such that

$$\mathbf{D}_j \triangleq \mathbf{S}_j \mathbf{G}_j \mathbf{S}_j^T \equiv \mathrm{diag}\left(\mathbf{D}_j^{(1)}, \ldots, \mathbf{D}_j^{(d_j)} \right) \quad (21)$$

1179

is a real Jordan form. Each real Jordan block $\mathbf{D}_j^{(i)} \in \mathbb{R}^{f_{ij} \times f_{ij}}$ is governed by a single eigenvalue $\eta_j^{(i)}$ or a pair of complex conjugate eigenvalues $\eta_j^{(i)}, \overline{\eta_j^{(i)}}$, and is thus similar to a conventional Jordan block. Thus

$$\|\mathbf{D}_j^{(i)}\| \leq |\eta_j^{(i)}| + 1, \tag{22}$$

a fact which is useful later. Further define

$$h_j^{(i)} \triangleq \lfloor \xi^{\tau-d+1} \max\{|\eta_j^{(i)}|, 1\} \rfloor + 1, \tag{23}$$

where $\xi > 1$ is a parameter to be selected. At time $k = (j+1)\tau - d$, encode the transformed error vector $\mathbf{S}_j(\mathbf{x}_{j\tau} - \hat{\mathbf{x}}_{j\tau})$ by dividing each of its components by l_j and passing it through a uniform quantiser on $[-0.5, 0.5]$ with $h_j^{(i)}$ levels, i denoting the index of the real Jordan block which governs the component. The indices of the selected quantiser levels are then transmitted at time $j\tau + \tau - d$ as a composite codeword $s_{j\tau+\tau-d}$ having $m_{j\tau+\tau-d} \triangleq \prod_{i=1}^{d_j} (h_j^{(i)})^{f_{ij}}$ possible values. Note that $m_{j\tau+\tau-d}$ is stochastic via a fixed dependence on the Markov modes $z_{j\tau}, \ldots, z_{j\tau+\tau-d}$ which determine $\mathbf{D}_j$. At time $j\tau + \tau - 1$, the coder state is then updated via

$$\hat{\mathbf{x}}_{(j+1)\tau} = \left(\mathbf{A}(z_{(j+1)\tau-1}) + \mathbf{B}(z_{(j+1)\tau-1})\mathbf{K}(z_{(j+1)\tau-1})\right)\hat{\mathbf{x}}_{(j+1)\tau-1} + l_j \left(\mathbf{A}(z_{(j+1)\tau-1}) \cdots \mathbf{A}(z_{j\tau})\right) \mathbf{S}_j^{\mathsf{T}} Q_j \left(l_j^{-1}(\mathbf{S}_j[\mathbf{x}_{j\tau} - \hat{\mathbf{x}}_{j\tau}])\right), \tag{24}$$

where $\forall k \in [j\tau, \ldots, j\tau + \tau - 2]$,

$$\hat{\mathbf{x}}_{k+1} \triangleq (\mathbf{A}(z_k) + \mathbf{B}(z_k)\mathbf{K}(z_k))\hat{\mathbf{x}}_k, \tag{25}$$

$$l_{j+1} = 2\sqrt{n}c^{d-1}\xi^{-\tau+d-1}l_j. \tag{26}$$

The function $Q_j(\cdot)$ in (24) is a quantiser which uniformly partitions the unit hypercube $[-0.5, 0.5]^n$ into $m_{j\tau+\tau-d}$ subcubes with sides of length $1/h_j^{(i)}$ along every coordinate axis governed by $\mathbf{D}_j^{(i)}$. In addition, $\mathbf{K}(\cdot)$ is a mode-dependent control gain chosen so $\mathbf{A}(z) + \mathbf{B}(z)\mathbf{K}(z)$ has distinct eigenvalues all sufficiently small that $\|\mathbf{A}(z) + \mathbf{B}(z)\mathbf{K}(z)\| < 1, \forall z \in [1, \ldots, \zeta]$. Finally, the constant c in (26) is any uniform upper bound on $\|\mathbf{A}(z)\|$.

B. Controller

Assume that at some time $k = j\tau$ the controller knows the internal state ψ_j. At times $k \in [j\tau, \ldots, j\tau + \tau - 1]$, it applies the control $\mathbf{u}_k = \mathbf{K}(z_k)\hat{\mathbf{x}}_k$, using equation (25) to generate the $\hat{\mathbf{x}}_k$. At time $k = (j+1)\tau$, it receives the symbol $s_{j\tau+\tau-d}$ and then updates its internal state using (24) and (26).

C. Analysis

We first show by induction that $l_j/2$ is indeed a bound on the error $\mathbf{e}_{j\tau} \triangleq \mathbf{x}_{j\tau} - \hat{\mathbf{x}}_{j\tau}, \forall j \in \mathbb{Z}_+$. Suppose that this is true

for some $j \in \mathbb{Z}_+$. The τ-step error dynamics are given by

$$
\begin{aligned}
\mathbf{e}_{(j+1)\tau} &= \left(\mathbf{A}(z_{(j+1)\tau-1}) \cdots \mathbf{A}(z_{j\tau})\right) \left[\mathbf{e}_{j\tau} - l_j \mathbf{S}_j^{\mathsf{T}} Q_j \left(l_j^{-1} \mathbf{S}_j \mathbf{e}_{j\tau}\right)\right], \\
&= l_j \left(\mathbf{A}(z_{(j+1)\tau-1}) \cdots \mathbf{A}(z_{(j+1)\tau-d+1})\right) \mathbf{S}_j^{\mathsf{T}} \mathbf{D}_j \\
&\quad \times \left[l_j^{-1} \mathbf{S}_j \mathbf{e}_{j\tau} - Q_j \left(l_j^{-1} \mathbf{S}_j \mathbf{e}_{j\tau}\right)\right],
\end{aligned}
$$

hence

$$
\begin{aligned}
\|\mathbf{e}_{(j+1)\tau}\|^2 &\leq l_j^2 c^{2(d-1)} \left\|\mathbf{D}_j \left[l_j^{-1} \mathbf{S}_j \mathbf{e}_{j\tau} - Q_j \left(l_j^{-1} \mathbf{S}_j \mathbf{e}_{j\tau}\right)\right]\right\|^2, \\
&\leq l_j^2 c^{2(d-1)} \sum_{i=1}^{d_j} \left\|\mathbf{D}_j^{(i)}\right\|^2 \\
&\quad \times \sum_{v=1}^{f_{ij}} \left| \frac{(\mathbf{S}_j \mathbf{e}_{j\tau})^{(i,v)}}{l_j} - Q_j^{(i,v)}\left(\frac{(\mathbf{S}_j \mathbf{e}_{j\tau})^{(i,v)}}{l_j}\right)\right|^2, \\
&\leq l_j^2 c^{2(d-1)} \sum_{i=1}^{d_j} \left\|\mathbf{D}_j^{(i)}\right\|^2 \frac{f_{ij}}{4\lfloor \xi^{\tau-d+1} \max\{|\eta_j^{(i)}|, 1\} + 1 \rfloor^2}, \\
&\leq l_j^2 c^{2(d-1)} \sum_{i=1}^{d_j} \frac{f_{ij}(|\eta_j^{(i)}| + 1)^2}{4(\xi^{\tau-d+1} \max\{|\eta_j^{(i)}|, 1\})^2}, \\
&\leq l_j^2 c^{2(d-1)} \sum_{i=1}^{d_j} \frac{f_{ij}}{\xi^{2(\tau-d+1)}} = \frac{l_j^2 c^{2(d-1)} n}{\xi^{2(\tau-d+1)}} \equiv \frac{l_{j+1}^2}{4}, \tag{27}
\end{aligned}
$$

by (26), where the superscript (i, v) denotes the vth scalar component governed by the ith eigenvalue. By induction, $l_j/2$ is then an upper-bound on $\|\mathbf{e}_{j\tau}\|$ for all $j \in \mathbb{Z}_+$.

Further observe from (26) that for any $\xi > 1$, $2c^{d-1}\sqrt{n}\xi^{-\tau+d-1} < 1$ and thus $\|\mathbf{e}_{j\tau}\| \to 0$ exponentially fast with j, for all sufficiently large τ. Looking at the error at integer times $k \in [j\tau, \ldots, j\tau + \tau - 1]$, observe that

$$
\begin{aligned}
\mathbf{x}_{k+1} &= \mathbf{A}(z_k)\mathbf{x}_k + \mathbf{B}(z_k)\mathbf{u}_k, \quad \hat{\mathbf{x}}_{k+1} = \mathbf{A}(z_k)\hat{\mathbf{x}}_k + \mathbf{B}(z_k)\mathbf{u}_k. \\
&\Rightarrow \mathbf{e}_{k+1} = \mathbf{A}(z_k)\mathbf{e}_k = (\mathbf{A}_{k-1} \cdots \mathbf{A}(z_{j\tau}))\mathbf{e}_{j\tau} \\
&\Rightarrow \|\mathbf{e}_{k+1}\| \leq c^{\tau-1}\|\mathbf{e}_{j\tau}\| \to 0, \quad \text{as } j \to \infty.
\end{aligned}
$$

Now looking at the dynamical equation, observe that

$$
\begin{aligned}
\|\mathbf{x}_{k+1}\| &= \|(\mathbf{A}(z_k) + \mathbf{B}(z_k)\mathbf{K}(z_k))\mathbf{x}_k - \mathbf{B}(z_k)\mathbf{K}(z_k)\mathbf{e}_k\|, \\
&\leq \|\mathbf{A}(z_k) + \mathbf{B}(z_k)\mathbf{K}(z_k)\|\|\mathbf{x}_k\| + \|\mathbf{B}(z_k)\mathbf{K}(z_k)\|\|\mathbf{e}_k\|.
\end{aligned}
$$

As $\max_z \|\mathbf{A}(z) + \mathbf{B}(z)\mathbf{K}(z)\| < 1$ and $\|\mathbf{e}_k\| \to 0$ exponentially fast, it then follows that $\|\mathbf{x}_k\| \to 0$ exponentially fast as well.

All that now remains is to verify that the data rate of this coder-controller can be brought arbitrarily close to the lower bound (6). First, observe that for a fixed τ the data rate is

$$
\begin{aligned}
R &= \liminf_{j \to \infty} \frac{1}{j\tau} \sum_{h=0}^{j} \sum_{z=1}^{\zeta} \mathrm{P}\{Z_{h\tau} = z\} \mathrm{E}_{Z_0=z}\{\log_2 M_{\tau-d}\}, \\
&\leq \frac{1}{\tau} \sum_{z=1}^{\zeta} \mathrm{E}_{Z_0=z}\{\log_2 M_{\tau-d}\} \limsup_{j \to \infty} \frac{1}{j} \sum_{h=0}^{j} \mathrm{P}\{Z_{h\tau} = z\},
\end{aligned}
$$

where the equality above holds since $m_{h\tau+\tau-d}$ is a fixed function of the Markov modes $z_{h\tau}, \ldots, z_{h\tau+\tau-d}$, and thus

1180

$E_{Z_{h\tau}=z}\{\log_2 M_{h\tau+\tau-d}\} = E_{Z_0=z}\{\log_2 M_{\tau-d}\}$. As the transition probability matrix is positive, Perron's theorem [8] states that the limit above is identical to the stationary distribution Z^* of the Markov chain. Hence

$$R \leq \frac{1}{\tau}\sum_{z=1}^{\zeta} P\{Z^* = z\}E_{Z_0=z}\{\log_2 M_{\tau-d}\},$$

$$= \frac{1}{\tau}E^*\left\{\sum_{\eta_0\in\sigma(\mathbf{G}_0)} \log_2\lfloor\xi^{\tau-d+1}\max\{|\eta|,1\}+1\rfloor\right\},$$

$$\leq \frac{1}{\tau}E^*\left\{\sum_{|\eta|\geq 1,\eta\in\sigma(\mathbf{G}_0)} \log_2|\eta|\right\} + n\log_2\xi + \frac{n}{\tau},$$

where E^* denotes expectation over $Z_0 \sim Z^*$. Hence

$$\liminf_{\tau\to\infty} R \leq \liminf_{\tau\to\infty}\frac{1}{\tau}E^*\left\{\sum_{|\eta|\geq 1,\eta\in\sigma(\mathbf{G}_0)} \log_2|\eta|\right\} + n\log_2\xi,$$

By the positivity of $\mathbf{T}$, the limit on the RHS can be shown to be independent of the initial mode distribution and thus equals H. The result the follows by choosing ξ arbitrarily close to 1 and τ sufficiently large with data rate R arbitrarily close to the inferior limit on the LHS. $\square$

V. CONCLUSION

In this paper the problem of stabilising a fully observed, noiseless, finite-dimensional Markov jump linear system over a finite data rate channel was investigated. By viewing the problem as one of initial condition quantisation, it was shown that the data rate of any rth moment-stabilising coding and control scheme cannot be less than an exponent characterising the asymptotic mean rate of unstable subspace volume growth. A scheme which exponentially stabilised the plant at a data rate arbitrarily close to this bound was then constructed, thus demonstrating its tightness. Similar techniques are currently being investigated for partially observed Markov jump linear systems with additive noise.

VI. REFERENCES

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

[2] J. Baillieul. Feedback designs in information-based control. In *Stoch. The. Contr. Proc. Worksh.*, pages 35–57, Uni. Kansas, USA, Oct 2001. Springer.

[3] J. Baillieul. Feedback coding for information-based control - operating near the data-rate limit. In *Proc. 41st IEEE Conf. Dec. Contr.*, pages 3229–36, Dec 2002.

[4] R. W. Brockett and D. Liberzon. Quantized feedback stabilization of linear systems. *IEEE Trans. Autom. Contr.*, 45(7):1279–89, 2000.

[5] D. F. Delchamps. Stabilizing a linear system with quantized state feedback. *IEEE Trans. Autom. Contr.*, 35:916–24, 1990.

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

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

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

[9] H. Ishii and B. A. Francis. Quadratic stabilization of sampled-data systems with quantization. *Automatica*, 39(10):1793–1800, 2003. Also published in Proc. 15th IFAC World Cong., 2002.

[10] D. Liberzon. Stabilization by quantized state or output feedback: a hybrid control approach. In *Proc. 15th IFAC World Congress*, Barcelona, Jul 2002.

[11] D. Liberzon. On stabilization of linear systems with limited information. *IEEE Trans. Autom. Contr.*, 48(2):304–7, 2003.

[12] G. N. Nair, S. Dey, and R. J. Evans. Communication-limited stabilisability of jump Markov linear systems. In *Proc. 15th Int. Symp. Math. The. Netw. Sys.*, U. Notre Dame, USA, Aug 2002. Available at *www.nd.edu/mtns/main.html*.

[13] G. N. Nair and R. J. Evans. Stabilization with data-rate-limited feedback: tightest attainable bounds. *Sys. Contr. Lett.*, 41(1):49–56, Sep 2000.

[14] G. N. Nair and R. J. Evans. Mean square stabilisability of stochastic linear systems with data rate constraints. In *Proc. 41st IEEE Conf. Dec. Contr.*, pages 1632–7, Las Vegas, USA, Dec 2002.

[15] G. N. Nair and R. J. Evans. Exponential stabilisability of finite-dimensional linear systems with limited data rates. *Automatica*, 39:585–93, 2003. Short version published in Proc. 15th IFAC World Congress, July 2002.

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

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

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

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