53rd IEEE Conference on Decision and Control
December 15-17, 2014. Los Angeles, California, USA

Control with Minimum Communication Cost per Symbol

Justin Pearson, João P. Hespanha, Daniel Liberzon

***Abstract*— We address the problem of stabilizing a continuous-time linear time-invariant process under communication constraints. We assume that the sensor that measures the state is connected to the actuator through a finite capacity communication channel over which an encoder at the sensor sends symbols from a finite alphabet to a decoder at the actuator. We consider a situation where one symbol from the alphabet consumes no communication resources, whereas each of the others consumes one unit of communication resources to transmit. This paper explores how the imposition of limits on an encoder's bit-rate and average resource consumption affect the encoder/decoder/controller's ability to keep the process bounded. The main result is a necessary and sufficient condition for a bounding encoder/decoder/controller which depends on the encoder's bit-rate, its average resource consumption, and the unstable eigenvalues of the process.**

I. Introduction

This paper addresses the problem of stabilizing a continuous-time linear time-invariant process under communication constraints. As in [1–7], we assume that the sensor that measures the state is connected to the actuator through a finite capacity communication channel. At each sampling time, an encoder sends a symbol through the channel. The problem of determining whether or not it is possible to bound the state of the process under this type of encoding scheme is not new; it was established in [2–4] that a necessary and sufficient condition for stability can be expressed as a simple relationship between the unstable eigenvalues of A and the communication bit-rate.

We expand upon this result by considering the notion that encoders can effectively save communication resources by not transmitting information, while noting that the absence of an explicit transmission nevertheless conveys information. To capture this, we suppose that one symbol from the alphabet consumes no communication resources to transmit, whereas each of the others consumes one unit of communication resources. We then proceed to define the *average cost per symbol* of an encoder, which is essentially the average fraction of non-free symbols emitted. This paper's main technical contribution is a necessary and sufficient condition for the existence of an encoder/decoder/controller that bounds the state of the process. This condition depends on the channel's bit-rate, the encoder's average cost per symbol, and the unstable eigenvalues of A.

Our result extends [2] in the sense that as the constraint on the average cost per symbol is allowed to increase (becomes looser), our necessary and sufficient condition becomes the condition from [2]. As with [2–4], our result is constructive in the sense that we describe a family of encoder/decoder pairs that bound the process when our condition holds. A counterintuitive corollary to our main result shows that if the process may be bounded with bit-rate r, then there exists a bounding encoder/decoder/controller with bit-rate r which uses no more than 50% non-free symbols in its symbol-stream.

The remainder of this paper is organized as follows. Section III contains the main negative result of the paper, namely that boundedness is not possible when our condition does not hold. To prove this result we actually show that it is not possible to bound the process with a large class of encoders — which we call M-of-N encoders — that includes all the encoders with average cost per symbol not exceeding a given threshold. Section IV contains the positive result of the paper, showing that when our condition *does* hold, there is an encoder/decoder pair that can bound the process; we provide the encoding scheme.

II. Problem Statement

Consider a stabilizable linear time-invariant process

$$\dot{x} = Ax + Bu, \qquad x \in \mathbb{R}^n, u \in \mathbb{R}^m, \tag{1}$$

for which it is known that $x(0)$ belongs to a known bounded set $\mathcal{X}_0 \subset \mathbb{R}^n$. A sensor that measures the state $x(t)$ is connected to the actuator through a finite-data-rate, error-free, and delay-free communication channel. An *encoder* collocated with the sensor samples the state once every T time units, and from this sequence of measurements $\{x(kT) : k \in \mathbb{N}_{>0}\}$ causally constructs a sequence of symbols $\{s_k \in \mathcal{A} : k \in \mathbb{N}_{>0}\}$ from a nonempty finite alphabet $\mathcal{A}$. Without loss of generality, $\mathcal{A} = \{0, 1, \ldots, S\}$ with $S := |\mathcal{A}| - 1$. The encoder sends this symbol sequence through the channel at the rate of 1 symbol every T time units to a *decoder/controller* collocated with the actuator, which causally constructs the control signal $u(t)$, $t \geqslant 0$ from the sequence of symbols $\{s_k \in \mathcal{A} : k \in \mathbb{N}_{>0}\}$ that arrive at the decoder.

The positive time $T \in \mathbb{R}_{>0}$ between successive samplings is called the *sampling period*. The non-negative *bit-rate* $r \in \mathbb{R}_{\geqslant 0}$ of the channel is the rate of transmitted information in units of bits per time unit. For a channel capable of

J. Pearson and J. P. Hespanha are with the Department of Electrical and Computer Engineering, University of California, Santa Barbara, U.S.A., {jppearson,hespanha}@ece.ucsb.edu. This material is based upon work supported by the Institute for Collaborative Biotechnologies through grant W911NF-09-D-0001 from the U.S. Army Research Office.

D. Liberzon is with the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, U.S.A., liberzon@uiuc.edu. Supported by NSF grants CNS-1217811 and ECCS-1231196 and by the Korean National Research Foundation grant NRF-2011-220-D00043.

transmitting one symbol from an alphabet $\mathcal{A}$ every T time units, the bit-rate is given by

$$r := \frac{\log_2 |\mathcal{A}|}{T} = \frac{\log_2(S+1)}{T}. \tag{2}$$

We assume that the symbol $0 \in \mathcal{A}$ can be transmitted without consuming any communication resources, but the other S symbols each require one unit of communication resources per transmission. One can think of the "free" symbol 0 as the absence of an explicit transmission. The "communication resources" at stake may be energy, time, or any other resource that may be consumed in the course of the communication process. In order to capture the average rate at which an encoder consumes communication resources, we define the *average cost per symbol* of an encoder as follows: We say an encoder has *average cost per symbol not exceeding* $\gamma_{\max}$ if there exists a non-negative integer N_0 such that for every symbol sequence $\{s_k\}$ the encoder may generate, we have

$$\frac{1}{N_2} \sum_{k=N_1}^{N_1+N_2-1} I_{s_k \neq 0} \leqslant \gamma_{\max} + \frac{N_0}{N_2} \tag{3}$$

for all positive integers N_2, N_1, where $I_{s_k \neq 0} := 1$ if the kth symbol is not the free symbol, and 0 if it is. The summation in (3) captures the total resources spent transmitting symbols $s_{N_1}, s_{N_1+1}, \ldots, s_{N_1+N_2-1}$. Motivating this definition of average cost per symbol is the observation that the lefthand side has the intuitive interpretation of the average cost per transmitted symbol between symbols s_{N_1} and $s_{N_1+N_2-1}$. As $N_2 \to \infty$, the rightmost term vanishes, leaving $\gamma_{\max}$ as an upper bound on the average long-term cost per symbol of the symbol sequence. Note that the average cost per symbol $\gamma_{\max}$ of any encoder always satisfies $\gamma_{\max} \in [0,1]$ and does not depend on the sampling period T.

Whereas the bit-rate r only depends on the symbol alphabet $\mathcal{A}$ and sampling period T, the average cost per symbol of an encoder/decoder pair depends on every possible symbol sequence it may generate, and therefore may in general depend on the encoder/decoder pair, the controller, the process (1), and the initial condition $x(0)$.

The specific question considered in this paper is: under what conditions on the bit-rate and average cost per symbol does there exist a controller and encoder/decoder pair that keep the state of the process (1) bounded?

III. Necessary condition for boundedness with limited-communication encoders

It is well known from [2–4] that it is possible to construct a controller and encoder/decoder pair that bounds the process (1) with bit-rate r only if

$$r \ln 2 \geqslant \sum_{i:\Re\lambda_i[A] \geqslant 0} |\lambda_i[A]|, \tag{4}$$

where $\ln$ denotes the base-e logarithm, and the summation is over all eigenvalues of A with nonnegative real part. The following result characterizes the necessary bit-rate when one poses constraints on the encoder's average cost per symbol $\gamma_{\max}$. Specifically, when $\gamma_{\max} < S/(S+1)$ a bit-rate larger than (4) is necessary, but provided that $\gamma_{\max} \geqslant S/(S+1)$ the condition reduces to (4).

Theorem 1: Consider a channel with a sampling period T and an alphabet with S nonfree symbols and one free symbol, and an encoder/decoder pair with average cost per symbol not exceeding $\gamma_{\max}$. If this pair keeps the state of the process (1) bounded for every initial condition $x_0 \in \mathcal{X}_0$, then we must have

$$r\, f(\gamma_{\max}, S) \ln 2 \geqslant \sum_{i:\Re\lambda_i[A] \geqslant 0} \lambda_i[A], \tag{5}$$

where the bit-rate r is related to S and T via Equation (2), the function $f : [0,1] \times [0,\infty) \to [0,\infty)$ is defined as

$$f(\gamma, S) := \begin{cases} \frac{H(\gamma)+\gamma \log_2 S}{\log_2(S+1)} & 0 \leqslant \gamma \leqslant \frac{S}{S+1} \\ 1 & \frac{S}{S+1} < \gamma \leqslant 1, \end{cases} \tag{6}$$

and $H(p) := -p \log_2(p) - (1-p)\log_2(1-p)$ is the base-2 entropy of a Bernoulli random variable with parameter p. □

Remark 1: It is worth making three observations regarding the function f. First, by inspection, $f(\gamma, S)$ is continuous in γ for any fixed S. Second, $f(\gamma, S)$ is monotone nonincreasing in S for any fixed $\gamma \in [0,1]$, which implies that smaller alphabets are preferable to large ones when trying to satisfy (5) with a given fixed bit-rate. The reasoning is as follows: for a particular fixed bit-rate, an encoder may either rapidly transmit symbols from a small alphabet or slowly transmit symbols from a large alphabet. In the former case, since the free symbol occupies a larger fraction of the alphabet, it will tend to be used more frequently, resulting in lower resource consumption. The third observation is that the average cost per time unit, which is γ/T, can be made arbitrarily small while still satisfying (5). Depending on which one of the parameters γ, T, S is fixed, the others can be chosen in the following ways:

1) If the average cost per symbol γ is fixed, one could pick T very large to make γ/T as small as desired. Then, leveraging the fact that

$$r f(\gamma, S) = \begin{cases} \frac{H(\gamma)+\gamma \log_2 S}{T} & 0 \leqslant \gamma \leqslant \frac{S}{S+1} \\ \frac{\log_2(S+1)}{T} & \frac{S}{S+1} < \gamma \leqslant 1 \end{cases} \tag{7}$$

is monotone increasing in S for fixed γ, pick S large enough to satisfy (5). This approach has two downsides: First, with a large sampling period, the state, although remaining bounded, can grow quite large between transmissions. Second, large S means that the encoder/decoder pair must store and process a large symbol library, adding complexity to the pair's implementation.
2) If the sampling period T is fixed, one picks γ small so as to make γ/T as small as desired, then increases S as in the previous case to satisfy (5). Like the previous case, this approach requires a large symbol library.
3) If the symbol alphabet size S is fixed, one can choose sequences γ_k, T_k for which $\gamma_k/T_k \to 0$ but for which

(5) is satisfied for large enough k. For example, the sequences

$$\gamma_k := e^{-k}, \qquad T_k := e^{-k}\sqrt{k}, \qquad k \in \mathbb{N}_{>0}$$

have the property that $\gamma_k \to 0$, $T_k \to 0$, and $\gamma_k/T_k \to 0$, but $H(\gamma_k)/T_k \to \infty$, so leveraging (7) we conclude that $r_k f(\gamma_k, S) \ln 2 \to \infty$ (where $r_k := \log_2(S+1)/T_k$). This means that one can find $k \in \mathbb{N}_{>0}$ to make the average cost per time unit γ_k/T_k arbitrarily small, and also satisfy the necessary condition (5). The drawback of this approach is that to achieve a small sampling period T in practice requires an encoder/decoder pair with a very precise clock.

Remark 2: The addition of the "free" symbol effectively increases the bit-rate without increasing the rate of resource consumption, as seen by the following two observations:

- Without the free symbols, the size of the alphabet would be S and the bit-rate would be $\log_2(S)/T$. It could happen that this bit-rate is too small to bound the plant, yet after the introduction of the free symbol, the condition (5) is satisfied.
- Since $\gamma_{\max}$ is the fraction of non-free symbols, the quantity $r\gamma_{\max}$ is the number of bits per time unit spent transmitting non-free symbols. But since $f(\gamma, S) \geqslant \gamma$, again we see that the free symbols help satisfy (5). To see that $f(\gamma, S) \geqslant \gamma$, observe that for any $S \in \mathbb{N}_{>0}$, $f(\cdot, S)$ is concave and reaches 1 before the identity function does, hence it's everywhere above the identity function on $(0, 1)$, and it matches the identity function at the endpoints 0 and 1.

A. *Proof of Theorem 1*

We lead up to the proof of Theorem 1 by first establishing three lemmas centered around a restricted large class of encoders called M-of-N encoders. We first define M-of-N encoders, which essentially partition their symbol sequences into N-length *codewords*, each with M or fewer non-free symbols. Lemma 1 demonstrates that every encoder with a bounded average cost per symbol is an M-of-N encoder for appropriate N and M. Next, in Lemma 2 we establish a relationship between the number of codewords available to an M-of-N encoder and the function f as defined in (6). Then, in Lemma 3 we leverage previous work to establish a necessary condition for an M-of-N encoder to bound the state of the process. Finally, the proof of Theorem 1 leverages these three results.

We now introduce the class of M-of-N encoders. For $N \in \mathbb{N}_{>0}$, we define an *N-symbol codeword* to be a sequence

$$\{s_{\ell N+1}, s_{\ell N+2}, \dots, s_{\ell N+N}\}$$

of N consecutive symbols starting at an index $k = \ell N + 1$, with $\ell \in \mathbb{N}_{\geqslant 0}$. For $M \in \mathbb{R}_{\geqslant 0}$ with $M \leqslant N$, we say an *M-of-N encoder* is an encoder for which every N-symbol codeword has M or fewer non-free symbols, i.e.,

$$\sum_{k=\ell N+1}^{\ell N+N} I_{s_k \neq 0} \leqslant M, \qquad \forall \ell \in \mathbb{N}_{\geqslant 0}. \tag{8}$$

The total number of distinct N-symbol codewords available to an M-of-N encoder is thus given by

$$L(N, M, S) := \sum_{i=0}^{\lfloor M \rfloor} \binom{N}{i} S^i, \tag{9}$$

where the ith term in the summation counts the number of N-symbol codewords with exactly i non-free symbols.

Note that in keeping with the problem setup, the M-of-N encoders considered here each draw their symbols from the symbol library $\mathcal{A} := \{0, 1, \dots, S\}$ and transmit symbols with sampling period T.

An intuitive property of M-of-N encoders is that have they average cost per symbol not exceeding M/N. The proof is omitted for brevity.

The fact that an M-of-N encoder refrains from sending "expensive" codewords effectively reduces its ability to transmit information. Indeed, since an M-of-N encoder takes NT time units to transmit one of $L(N, M, S)$ codewords, effectively the encoder can transmit merely $\frac{\log_2 L(N,M,S)}{NT}$ bits of useful information per time unit. For $M < N$, this is strictly less than the bit-rate r.

The first lemma, proved in the appendix, shows that the set of M-of-N encoders is "complete" in the sense that every encoder with average cost per symbol not exceeding a finite threshold $\gamma_{\max}$ is actually an M-of-N encoder for N sufficiently large and $M \approx \gamma_{\max} N$.

Lemma 1: For every channel with bit-rate r and any encoder/decoder pair with average cost per symbol not exceeding $\gamma_{\max} \in [0, 1]$, and every constant $\epsilon > 0$, there exist $M \in \mathbb{R}_{\geqslant 0}$ and $N \in \mathbb{N}_{>0}$ with $M < N\gamma_{\max}(1+\epsilon)$ such that the encoder/decoder pair is an M-of-N encoder. □

The next lemma establishes a relationship between the number of codewords $L(N, M, S)$ available to an M-of-N encoder and the function f defined in (6).

Lemma 2: For any $N \in \mathbb{N}_{>0}$, $S \in \mathbb{N}_{\geqslant 0}$ and $\gamma \in [0, 1]$, the function L defined in (9) and the function f defined in (6) satisfy

$$\frac{\log_2 L(N, N\gamma, S)}{N} \leqslant \log_2(S+1) f(\gamma, S), \tag{10}$$

with equality holding only when $\gamma = 0$ or $\gamma = 1$. Moreover, we have asymptotic equality in the sense that

$$\lim_{N\to\infty} \frac{\log_2 L(N, N\gamma, S)}{N} = \log_2(S+1) f(\gamma, S). \tag{11}$$

□

Proof of Lemma 2. Let $N \in \mathbb{N}_{>0}$ and $S \in \mathbb{N}_{\geqslant 0}$ be arbitrary. First we prove (10) for $\gamma \in \left(0, \frac{S}{S+1}\right]$. Applying the Binomial Theorem to the identity $1 = (\gamma + (1-\gamma))^N$, we obtain

$$1 = \sum_{i=0}^{N} \binom{N}{i} \gamma^i (1-\gamma)^{N-i}$$

Since each term in the summation is positive, keeping only the first $\lfloor N\gamma \rfloor$ terms yields the inequality

$$1 > \sum_{i=0}^{\lfloor N\gamma \rfloor} \binom{N}{i} \gamma^i (1-\gamma)^{N-i} \tag{12}$$

Next, a calculation presented as Lemma 5 in the appendix reveals that

$$\gamma^i(1-\gamma)^{N-i} \geqslant 2^{-N\,H(\gamma)}\,\frac{S^i}{S^{N\gamma}} \tag{13}$$

for all $N, S \in \mathbb{N}_{>0}$, $\gamma \in \left(0, \frac{S}{S+1}\right]$, and $i \in [0, N\gamma]$. Using this in (12) and taking $\log_2$ of both sides yields

$$\frac{\log_2 L(N, N\gamma, S)}{N} < H(\gamma) + \gamma \log_2 S. \tag{14}$$

By the definition of f, we have $\log_2(S+1)f(\gamma, S) = H(\gamma) + \gamma\log_2 S$ when $\gamma \in \left[0, \frac{S}{S+1}\right]$. Thus, (14) proves the strict inequality in (10) for $\gamma \in \left(0, \frac{S}{S+1}\right]$. Next, suppose $\gamma \in \left(\frac{S}{S+1}, 1\right)$ and observe from (9) that $L(N, M, S)$ is a sum of positive terms whose index reaches $\lfloor M \rfloor$, hence $L(N, N\gamma, S)$ is strictly less than $L(N, N, S)$ for any $\gamma < 1$. We conclude that

$$\begin{aligned}\frac{\log_2 L(N, N\gamma, S)}{N} &< \frac{\log_2 L(N, N, S)}{N}\\ &= \log_2(S+1) \qquad (15)\\ &= \log_2(S+1)f(\gamma, S), \qquad (16)\end{aligned}$$

where the equality in (15) follows simply from the fact that $L(N, N, S)$ is the number of all possible codewords of length N and hence equals $(S+1)^N$, and (16) follows from the definition of f when $\gamma \in (\frac{S}{S+1}, 1)$. This concludes the proof of the strict inequality in (10) for $\gamma \in (0, 1)$. The proof of (10) for $\gamma = 0$ follows merely from inspection of (10), and the $\gamma = 1$ case follows from the equality in (15). The proof of the asymptotic result (11) appears in [8]. This concludes the proof of Lemma 2. ■

The following lemma provides a necessary condition for an M-of-N encoder to bound the process (1).

Lemma 3: Consider an M-of-N encoder/decoder pair using a channel with symbols $\{0, \ldots, S\}$ (with 0 the free symbol) and sampling period T. If the pair keeps the state of the process (1) bounded for every initial condition, then we must have

$$\frac{\ln L(N, M, S)}{NT} > \sum_{i:\Re\lambda_i[A]\geqslant 0} \lambda_i[A]. \tag{17}$$

□

Proof of Lemma 3. Consider an encoder/decoder/controller triple that bounds the state of the process (1) and whose encoder is an M-of-N encoder using symbols $\{0, \ldots, S\}$ and sampling period T. Since the encoder sends one of $L(M, N, S)$ codewords every NT time units, the bit-rate of the encoder is $\log_2 L(M, N, S)/NT$. Theorem 1 of [2] proves that if the state of the process (1) is bounded by an encoder/decoder/controller triple under the communication constraints described in our problem setup, then the pair's bit-rate must not be less than $\frac{1}{\ln 2}\sum_{i:\Re\lambda_i[A]\geqslant 0}\lambda_i[A]$. Hence, we have

$$\frac{\log_2 L(M, N, S)}{NT} \geqslant \frac{1}{\ln 2}\sum_{i:\Re\lambda_i[A]\geqslant 0} \lambda_i[A],$$

from which (17) follows. This proves the lemma. ■

Now we are ready to prove Theorem 1.

Proof of Theorem 1. By Lemma 1, for any $\epsilon > 0$ there exist $M \in \mathbb{R}_{\geqslant 0}$ and $N \in \mathbb{N}_{>0}$ with $M < N\gamma_{\max}(1+\epsilon)$ for which the encoder/decoder is an M-of-N encoder. Since the state of the process is kept bounded, by Lemma 3 we have

$$\sum_{i:\Re\lambda_i[A]\geqslant 0} \lambda_i[A] < \frac{\log_2 L(N, M, S)}{NT}\ln 2. \tag{18}$$

Since L is monotonically nondecreasing in its second argument and $M < N\gamma_{\max}(1+\epsilon)$, we have

$$\frac{\log_2 L(N, M, S)}{NT} \leqslant \frac{\log_2 L(N, N\gamma_{\max}(1+\epsilon), S)}{NT}. \tag{19}$$

Lemma 2, together with (2), implies that

$$\frac{\log_2 L(N, N\gamma_{\max}(1+\epsilon), S)}{NT} \leqslant rf(\gamma_{\max}(1+\epsilon), S). \tag{20}$$

Combining these and letting $\epsilon \to 0$, we obtain (5). This completes the proof of Theorem 1. ■

IV. Sufficient Condition for Boundedness With Limited-Communication Encoders

The previous section established a necessary condition (5) on the channel bit-rate and the average cost per symbol of an encoder/decoder pair in order to bound the state of the process (1). In this section, we show that this condition is also sufficient for a bounding encoder/decoder to exist. The proof is constructive in that we provide the encoder/decoder.

Theorem 2: Assume that A is diagonalizable. For every $S \in \mathbb{N}_{\geqslant 0}$, $T > 0$, and $\gamma_{\max} \in [0, 1]$ satisfying

$$rf(\gamma_{\max}, S)\ln 2 > \sum_{i:\Re\lambda_i[A]\geqslant 0} \lambda_i[A], \tag{21}$$

where r is defined in (2) and the function f is defined in (6), there exists a channel using S nonfree symbols, sampling period T, and an encoder/decoder pair with average cost per symbol not exceeding $\gamma_{\max}$, which keeps the state of the process (1) bounded for every initial condition $x_0 \in \mathcal{X}_0$. □

The proof of Theorem 2 relies on the following lemma, which provides a sufficient condition for the existence of an M-of-N encoder to bound the state of the process (1).

Lemma 4: Assume that A is diagonalizable. For every $T \in \mathbb{R}_{>0}$, $N \in \mathbb{N}_{>0}$, $S \in \mathbb{N}_{\geqslant 0}$, and $M \in \mathbb{R}_{\geqslant 0}$ with $N \geqslant M$ satisfying

$$\frac{\ln L(N, M, S)}{NT} > \sum_{i:\Re\lambda_i[A]\geqslant 0} \lambda_i[A], \tag{22}$$

there exists an M-of-N encoder using alphabet $\{0, \ldots, S\}$ with sampling period T that keeps the state of the process (1) bounded for every initial condition. □

Proof of Lemma 4. This proof builds on Theorem 2 from [2], which provides sufficient conditions on an encoder's bit-rate to ensure the existence of a stabilizing controller and encoder/decoder pair. The result states that if an encoder's bit-rate r satisfies

$$r \geqslant \frac{1}{\ln 2} \sum_{i:\Re\lambda_i[A]\geqslant 0} \lambda_i[A], \tag{23}$$

where A is the continuous-time process matrix, then there exists a controller and encoder/decoder pair that bound the state of the process. An outline of the proof of this result from [2] follows. The encoder works by placing a bounding rectangle around the volume where the state is known to lie, and partitions it into two pieces along its largest axis. Then the encoder transmits a 0 or 1, depending on which sub-rectangle the state lies in. Since the decoder can calculate the bounding rectangles with its knowledge of $\mathcal{X}_0$ and the process dynamics in (1), it estimates the state to be the centroid of whichever sub-rectangle the received symbol corresponds to. The decoder transmits this state estimate to the controller, which can be any stabilizing linear state-feedback controller.

We now provide a summary of the proof of Lemma 4, and refer the reader to [8] for the detailed proof. Suppose T, N, M, S satisfy the constraints in the statement of the lemma as well as (22). Theorem 2 from [2] guarantees the existence of a stabilizing encoder/decoder/controller triple with bit-rate $\log_2 L(N, M, S)/NT$. All that remains is to adapt this encoder/decoder pair into our framework, i.e., an M-of-N encoder that uses $S+1$ symbols with sampling period T. We do this by building an encoder which runs a copy of the encoder from [2] internally. The outer encoder feeds samples of the state to the inner encoder, from which the inner encoder generates a string of 1's and 0's. Due to the bit-rate being $\log_2 L(N, M, S)/NT$, after NT time units the inner encoder has generated a string of length $\log_2 L(N, M, S)$. There are $L(N, M, S)$ possible such strings, and so the outer encoder can map this string uniquely to one of the $L(N, M, S)$ possible N-length codewords with M or fewer non-free symbols from the alphabet $\{0, \ldots, S\}$. The outer encoder transmits this codeword across the channel to the decoder, which performs the inverse mapping to recover the string of $\log_2 L(N, M, S)$ 1's and 0's, which it delivers to its inner decoder. From this bit-string and knowledge of the process dynamics, the inner decoder computes a state estimate, which it delivers to the stabilizing linear state-feedback controller. This process repeats each NT time units. Since the inner encoder/decboder pair bound the process, the outer M-of-N encoder described here does as well. This proves the lemma. ∎

Now we are ready to prove Theorem 2.

Proof of Theorem 2. Assume that S, T, and $\gamma_{\max}$ satisfy (21), so that

$$\epsilon := r f(\gamma_{\max}, S) \ln 2 - \sum_{i:\Re\lambda_i[A]\geqslant 0} \lambda_i[A] > 0. \tag{24}$$

Equation (11) establishes that $\frac{\ln L(N, N\gamma_{\max}, S)}{NT}$ gets arbitrarily close to $r f(\gamma_{\max}, S) \ln 2$ as we increase N, so we pick N sufficiently large to satisfy

$$r f(\gamma_{\max}, S) \ln 2 - \frac{\ln L(N, N\gamma_{\max}, S)}{NT} < \epsilon. \tag{25}$$

By (9), we have $L(N, \lfloor N\gamma_{\max}\rfloor, S) = L(N, N\gamma_{\max}, S)$ for every N, $\gamma_{\max}$, and S. Setting $M := N\gamma_{\max}$, by (24) and (25) we have found N and M satisfying $\frac{\ln L(N,M,S)}{NT} > \sum_{i:\Re\lambda_i[A]\geqslant 0} \lambda_i[A]$. Hence by Lemma 4, there exists an M-of-N encoder/decoder which bounds the state of the process (1). Since all M-of-N encoder/decoders have average cost per symbol not exceeding M/N and for this encoder/decoder we have $M/N = \gamma_{\max}$, we conclude that this encoder has an average cost per symbol not exceeding $\gamma_{\max}$. This concludes the proof, since we have found the desired encoder. ∎

An unexpected consequence of Theorems 1 and 2 is that when it is possible to keep the state of the process bounded with a given bit-rate $r := \log_2(S+1)/T$, one can always find M-of-N encoders that bound it for (essentially) the same bit-rate and average cost per symbol not exceeding $S/(S+1)$, i.e., approximately a fraction $1/(S+1)$ of the symbols will not consume communication resources. In the most advantageous case, the encoder/decoder use the alphabet $\{0, 1\}$ and the encoder's symbol stream consumes no more than 50% of the communication resources. The price paid for using an encoder/decoder with average cost per symbol closer to $S/(S+1)$ than 1 is that it may require prohibitively long codewords (large N) as compared to an encoder with higher average cost per symbol. To see this, note that $f(\gamma, S) = 1$ when $\gamma \in [S/(S+1), 1]$ and recall that $\ln L(N, N\gamma, S)/NT$ is monotonically nondecreasing in γ and N. Hence, with r given (so S and T are fixed), one can decrease γ from 1 toward $S/(S+1)$ and still satisfy (25) by increasing N.

Corollary 1: If the process (1) can be bounded with an encoder/decoder pair using a channel with symbol alphabet $\{0, 1, \ldots, S\}$ and sampling period T, then for arbitrary $\epsilon \in (0, T)$ there exists an M-of-N encoder that uses that channel with average cost per symbol not exceeding $S/(S+1)$ that bounds its state. □

Proof of Corollary 1. Recalling $r := \log_2(S+1)/T$ and the definition of f, we have

$$\frac{\log_2(S+1)}{T-\epsilon} f\left(\frac{S}{S+1}, S\right) > r\, f\left(\frac{S}{S+1}, S\right) = r f(1, S) \ln 2 \geqslant \sum_{i:\Re\lambda_i[A]\geqslant 0} \lambda_i[A], \tag{26}$$

where the last inequality follows from Theorem 1. Applying Theorem 2 to (26), we conclude that there exists an encoder/decoder pair using symbol alphabet $\{0, \ldots, S\}$, sampling period $T-\epsilon$, and with average cost per symbol not exceeding $S/(S+1)$ which bounds the state of the process (1). By Lemma 1, this encoder is an M-of-N encoder for appropriately chosen M and N. ∎

V. Conclusion and Future Work

In this paper we considered the problem of bounding the state of a continuous-time linear process under communication constraints. We considered constraints on both the channel bit-rate and the encoding scheme's average cost per symbol. Our main contribution was a necessary and sufficient condition on the process and constraints for which a bounding encoder/decoder/controller exists. In the absence of a limit on the average cost per symbol, the conditions recovered previous work. A surprising corollary to our main result was the observation that one may impose a constraint on the average cost per symbol without necessarily needing to loosen the bit-rate constraint. Specifically, we proved that if a process may be bounded with a particular bit-rate, then there exists a (possibly very complex) encoder/decoder that can bound it with that same bit-rate, while using no more than 50% non-free symbols on average. One would expect that the prohibition of some codewords would require that the encoder necessarily compensate by transmitting at a higher bit-rate, but this not the case.

We observed in Remark 1 that smaller alphabets incur a smaller penalty on the conditions for boundedness in Theorems 1 and 2. This suggests that encoding schemes with small alphabets may be able to bound the state of the process with bit-rates and average costs not far above the minimum theoretical bounds as established in Theorems 1 and 2. *Event-based* control strategies comprise one such class of encoders; they use a small number of non-free symbols to notify the decoder/controller about certain state-dependent events. We have preliminary results showing that event-based encoders can indeed be used to produce encoder/decoder pairs that almost achieve the minimum achievable average cost per symbol that appears in Theorems 1 and 2.

Finally, our problem setup considered merely whether there exists a bounding encoder/decoder/controller triple. It seems natural to extend this setup to finding stabilizing triples.

Appendix

Proof of Lemma 1. By the definition of average cost per symbol not exceeding $\gamma_{\max}$ in (3) there exists an integer $N_0 \in \mathbb{N}_{>0}$ such that for any symbol sequence $\{s_k\}$ that the encoder generates, we have

$$\sum_{i=1}^{N} I_{s_i \neq 0} \leqslant N_0 + N\gamma_{\max}, \quad \forall N \in \mathbb{N}_{>0}. \tag{27}$$

Pick $N \in \mathbb{N}_{>0}$ large enough to satisfy $N_0 + 2 < \epsilon N\gamma_{\max}$ and pick $M := \lfloor N_0 + 2 + N\gamma_{\max} \rfloor$. Combining these with (27), we obtain

$$\begin{aligned}\sum_{i=1}^{N} I_{s_i \neq 0} &< N_0 + N\gamma_{\max} < M \\ &\leqslant N_0 + 2 + N\gamma_{\max} < N\gamma_{\max}(1+\epsilon),\end{aligned}$$

which establishes that $M < N\gamma_{\max}(1+\epsilon)$. This completes the proof. ■

Lemma 5: The following inequality holds for all $N, S \in \mathbb{N}_{>0}$, $q \in (0, S/(S+1)]$, and $i \in [0, Nq]$:

$$q^i(1-q)^{N-i} \geqslant 2^{-N\,H(q)}\,\frac{S^i}{S^{Nq}} \tag{28}$$

where $H(q) := -q\log_2 q - (1-q)\log_2(1-q)$ is the the base-2 entropy of a Bernoulli random variable with parameter q.

Proof of Lemma 5. Let N, S, q, and i take arbitrary values from the sets described in the lemma's statement. Since $\log_2$ is a monotone increasing function, $\log_2(q/(1-q))$ for $q > 0$ is maximized at the right endpoint value, $q = S/(S+1)$, where it equals $\log_2 S$. This leads to

$$\log_2 q - \log_2(1-q) \leqslant \log_2 S \tag{29}$$

for all $S \in \mathbb{N}_{>0}$ and $q \in (0, S/(S+1)]$. Next, $i \in [0, Nq]$ by assumption, therefore $i - Nq \leqslant 0$. Multiplying (29) by $i - Nq$ and straightforward algebraic manipulation yields

$$\begin{aligned}&i\log_2 q + (N-i)\log_2(1-q) \\ &\leqslant Nq\log_2 q + N(1-q)\log_2(1-q) + (i-Nq)\log_2 S \\ &= -NH(q) + (i-Nq)\log_2 S,\end{aligned}$$

where the equality follows from the definition of $H(q)$. Raising 2 to the power of both sides, (28) follows. ■

References

[1] R. Brockett and D. Liberzon, "Quantized feedback stabilization of linear systems," *Automatic Control, IEEE Transactions on*, vol. 45, no. 7, pp. 1279–1289, Jul 2000.

[2] J. P. Hespanha, A. Ortega, and L. Vasudevan, "Towards the control of linear systems with minimum bit-rate," in *Proc. of the Int. Symp. on the Mathematical Theory of Networks and Syst.*, Aug. 2002.

[3] G. Nair and R. Evans, "Communication-limited stabilization of linear systems," in *Decision and Control, 2000. Proceedings of the 39th IEEE Conference on*, vol. 1, 2000, pp. 1005 –1010.

[4] S. Tatikonda and S. Mitter, "Control under communication constraints," *Automatic Control, IEEE Transactions on*, vol. 49, no. 7, pp. 1056 – 1068, july 2004.

[5] G. N. Nair and R. J. Evans, "Exponential stabilisability of finite-dimensional linear systems with limited data rates," *Automatica*, vol. 39, no. 4, pp. 585 – 593, 2003.

[6] A. Sahai and S. Mitter, "The necessity and sufficiency of anytime capacity for stabilization of a linear system over a noisy communication link — part i: Scalar systems," *Information Theory, IEEE Transactions on*, vol. 52, no. 8, pp. 3369–3395, Aug 2006.

[7] A. Matveev and A. Savkin, "Multirate stabilization of linear multiple sensor systems via limited capacity communication channels," *SIAM Journal on Control and Optimization*, vol. 44, no. 2, pp. 584–617, 2005.

[8] J. Pearson, J. P. Hespanha, and D. Liberzon, "Control with minimum communication cost per symbol," University of California, Santa Barbara, Tech. Rep., Mar. 2014. [Online]. Available: http://www.ece.ucsb.edu/~hespanha/techrep.html