

## Stochastic Stabilization of Partially Observed and Multi-Sensor Systems Driven by Unbounded Noise Under Fixed-Rate Information Constraints

Andrew P. Johnston and Serdar Yüksel

**Abstract**—We investigate the stabilization of unstable multidimensional partially observed single-sensor and multi-sensor (single-controller) discrete-time linear systems driven by unbounded noise and controlled over discrete noiseless channels. Stability is achieved under fixed-rate communication requirements that are asymptotically tight in the limit of large sampling periods. Through the use of similarity transforms, sampling and random-time drift conditions we obtain a coding and control policy leading to the existence of a unique invariant distribution and finite second moment for the sampled state. We obtain tight necessary and sufficient conditions for the general multi-sensor case under an assumption related to the Jordan form structure of such systems. In the absence of this assumption, we give sufficient conditions for stabilization.

**Index Terms**—Networked control systems, quantization, stability, stochastic systems.

### I. INTRODUCTION

Networked control systems are becoming increasingly commonplace, with recent focus on the design of such systems being on the settings with multiple decision makers connected over rate-limited information channels. The goal for the design of such systems is either stabilization or optimization given information limitations. For a detailed review of the literature for the design of such systems, see [1], [2]. This technical note investigates stabilization of such networked control systems where the system to be stabilized is driven by noise with unbounded support and the communication channel between the sensors and the controller is a discrete noiseless channel with rate constraints.

#### A. Problem Statement

In this technical note, we consider the class of multi-sensor LTI discrete-time systems with both plant and observation noise. The system equations are given by

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{B}\mathbf{u}_t + \mathbf{w}_t, \quad (1a)$$

$$\mathbf{y}_t^j = \mathbf{C}^j \mathbf{x}_t + \mathbf{v}_t^j, \quad 1 \leq j \leq M \quad (1b)$$

where  $\mathbf{x}_t \in \mathbb{R}^n$  and  $\mathbf{u}_t \in \mathbb{R}^m$  are the state and control action variables at time  $t \in \mathbb{N}$  respectively. The observation made by sensor  $j$  at time  $t$  is denoted by  $\mathbf{y}_t^j \in \mathbb{R}^{p_j}$ . The matrices  $\mathbf{A}, \mathbf{B}, \mathbf{C}^j$  and random vectors  $\mathbf{w}_t, \mathbf{v}_t^j$  are of compatible size.

We require that  $\{\mathbf{w}_t\}$  and each  $\{\mathbf{v}_t^j\}$  are sequences of i.i.d. random vectors drawn from a distribution  $\mathcal{Z}$ , with finite  $2 + \epsilon$  moments in each component for some  $\epsilon > 0$ , which admits a probability density that

Manuscript received September 07, 2012; revised March 11, 2013; accepted July 20, 2013. Date of publication August 28, 2013; date of current version February 19, 2014. This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC). This paper was presented in part at the IEEE Conference on Decision and Control, Maui, HI, December 2012. Recommended by Associate Editor P. Shi.

The authors are with the Department of Mathematics and Statistics, Queen's University, Kingston, ON K7L 3N6, Canada (e-mail: a.johnston@queensu.ca; yuksel@mast.queensu.ca).

Digital Object Identifier 10.1109/TAC.2013.2279912

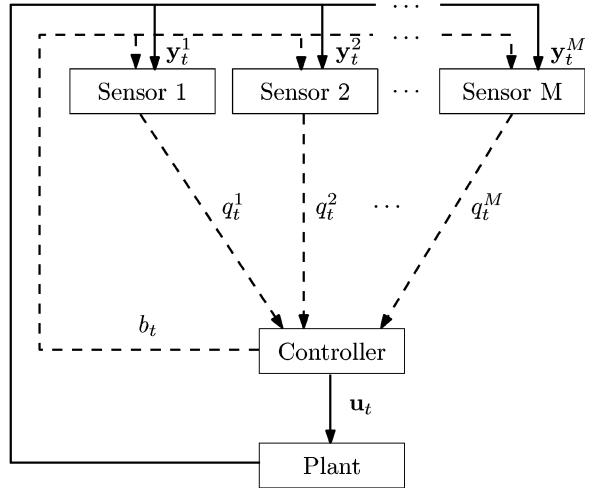


Fig. 1. Multi-sensor system with finite-rate communication channels.

is positive on every (non-empty) open set. At time  $t$ ,  $\mathbf{w}_t$  and each  $\mathbf{v}_t^j$  are independent of each other and the state  $\mathbf{x}_t$ . The initial state,  $\mathbf{x}_0$ , is drawn from the  $\mathcal{Z}$  distribution.

**Assumption 1.1:** We require controllability and joint observability. That is, the pair  $(\mathbf{A}, \mathbf{B})$  is controllable and the pair  $([(\mathbf{C}^1)^T \cdots (\mathbf{C}^M)^T]^T, \mathbf{A})$  is observable but the individual pairs  $(\mathbf{C}^j, \mathbf{A})$  may not be observable.

The setup is depicted in Fig. 1. The observations are made by a set of  $M$  sensors and each sensor sends information to the controller through a finite capacity channel. At each time stage  $t$ , we allow sensor  $j \in \{1, \dots, M\}$  to send an encoded value  $q_t^j \in \{1, 2, \dots, N_t^j\}$  for some  $N_t^j \in \mathbb{N}$ . In addition, the controller can send a feedback value  $b_t \in \{0, 1\}$  at times  $t = Ts$ , where  $T$  is the period of our coding policy and  $s \in \mathbb{N}$ . The value  $b_t$  is seen by all sensors at time  $t + 1$ . We define the rate at time  $t$  as  $R_t = \sum_{j=1}^M \log_2(N_t^j)$ . The coding scheme is applied periodically with period  $T$  and so the rate for all time stages is specified by  $\{N_0^j, \dots, N_{T-1}^j : 1 \leq j \leq M\}$ . The average rate is  $R_{\text{avg}} = (1/T)(M + \sum_{t=0}^{T-1} R_t)$ , accounting for the encoded and feedback signals. For the case with a single sensor,  $M$  can be taken to be zero in the rate expression, as we discuss further in the technical note.

**Information Structure:** For a process  $\{\mathbf{x}_t\}$  we define  $\mathbf{x}_{[a,b]} = \{\mathbf{x}_a, \mathbf{x}_{a+1}, \dots, \mathbf{x}_b\}$ . At time  $t$ , each sensor  $j$  maps its information  $I_t^{s,j} := \{\mathbf{y}_{[0,t]}^j, b_{[0,t-1]}\} \rightarrow q_t^j \in \{1, \dots, N_t^j\}$ . The controller maps its information  $I_t^c := \{q_{[0,t]}^1, \dots, q_{[0,t]}^M\} \rightarrow (\mathbf{u}_t, b_t) \in \mathbb{R}^m \times \{0, 1\}$ .

We denote the indicator function of an event  $E$  by  $1_E$ . Unless otherwise stated, all vectors are assumed to be column vectors. For any  $\mathbf{x} \in \mathbb{R}^n$  we write  $\mathbf{x} = [x^1 \cdots x^n]^T$  where  $x^i \in \mathbb{R}$  is the  $i$ th entry. We define the absolute value operation for vectors as the component-wise absolute value. That is,  $|\mathbf{x}| = [|x^1| \cdots |x^n|]^T$ . For a matrix  $\mathbf{A} \in \mathbb{R}^{n \times n}$ , we denote its transpose by  $\mathbf{A}^T$  and determinant by  $\det(\mathbf{A})$ . We let  $\Lambda(\mathbf{A})$  denote the set of eigenvalues of  $\mathbf{A}$ . The  $p$  norm is denoted by  $\|\cdot\|_p$  and defined as  $\|\mathbf{x}\|_p = \{\sum_{i=1}^n |x^i|^p\}^{1/p}$ .

**Definition 1.2:** For  $\mathbf{x} \in \mathbb{R}^n$  and  $\mathbf{y} \in \mathbb{R}_+^n$  we write  $\mathbf{x} \leq \mathbf{y}$  if  $|x_i| \leq y_i$  for all  $1 \leq i \leq n$ . We write  $\mathbf{x} \not\leq \mathbf{y}$  otherwise.

The observability matrix of sensor  $j$  is  $\mathcal{O}_{(\mathbf{C}^j, \mathbf{A})} = [(\mathbf{C}^j)^T (\mathbf{C}^j \mathbf{A})^T \cdots (\mathbf{C}^j \mathbf{A}^{n-1})^T]$ , the unobservable space is  $N^j = \text{Ker}(\mathcal{O}_{(\mathbf{C}^j, \mathbf{A})})$  and the observable subspace is defined to be  $O^j = (N^j)^\perp$  for  $1 \leq j \leq M$ .

### B. Brief Literature Review

Due to space limitations, we are unable to give a fair account of the literature. We refer the reader to [1], [2] and [3] where this note builds on ([3] also contains the proofs of some of the results which are not included here). Of particular importance, [4] and [5] obtained a fundamental lower bound on the average rate of the information transmission for quadratic (mean-square) stabilization as we discuss in the following. For the system (1), letting  $\{\lambda_i\}$  be the set of eigenvalues of  $\mathbf{A}$ , this bound is

$$R_{\min} = \sum_{|\lambda_i| > 1} \log_2 (|\lambda_i|). \quad (2)$$

Various publications have studied the characterization of minimum information requirements for multi-sensor and multi-controller linear systems with various topologies of decentralization and the fundamental bounds have been extensively studied in [1], [6]–[20]. These contributions do not consider the noisy case with multi-sensor settings, as well as ergodicity properties for multi-dimensional systems. When a linear system is driven by unbounded noise, the analysis is particularly difficult since a bounded quantizer range leads to a transient state process (see for example Theorem 7.3.1 in [2] or Proposition 5.1 in [4]). In such a noisy setup, for single-sensor systems, a stability result of the form  $\limsup_{t \rightarrow \infty} E[\|\mathbf{x}_t\|_2^2] < \infty$  was given for noisy systems with unbounded support in [4], which uses a variable-rate quantizer. A fixed-rate scheme was presented in [15] for a scalar noisy system using martingale theory, which achieved the lower bound plus an additional symbol required for encoding. The existence of an invariant distribution was established under the coding and control policy presented, along with a finite second moment of the state; that is,  $\lim_{t \rightarrow \infty} E[\|\mathbf{x}_t\|_2^2] < \infty$ . Authors in [21] considered a general random-time stochastic drift criteria for Markov chains and applied it to erasure channels in a similar spirit. Variable-rate coding schemes for time-varying channels were considered in [17]. Authors in [22] studied conditions for stabilization when the control actions are uniformly bounded, the controlled multi-dimensional system is marginally stable and is driven by noise with unbounded support. The multi-sensor setting with unbounded noise has not been studied to our knowledge.

In view of the literature, the contributions of this note are as follows: (i) The case where the system is multi-dimensional and driven by unbounded noise over a noiseless discrete-channel has not been studied to our knowledge, regarding the existence of an invariant distribution, ergodicity and finite moment properties. (ii) We give sufficient conditions for multi-sensor systems with both system noise and observation noise with unbounded support, which has not been treated previously (see [1] for a review). Even though our approach builds on the program in [15] and [21], a more general class of stopping times is introduced in this technical note, as more tedious constructions are needed for the vector, partially observed and decentralized settings.

## II. SINGLE-SENSOR SYSTEMS

Consider the class of single-sensor LTI discrete-time systems with both plant and observation noise. The system equations are given by

$$\mathbf{x}_{t+1} = \mathbf{Ax}_t + \mathbf{Bu}_t + \mathbf{w}_t, \quad \mathbf{y}_t = \mathbf{Cx}_t + \mathbf{v}_t \quad (3)$$

where  $\mathbf{x}_t \in \mathbb{R}^n$ ,  $\mathbf{u}_t \in \mathbb{R}^m$  and  $\mathbf{y}_t \in \mathbb{R}^p$  are the state, control action and observation at time  $t$  respectively. The matrices  $\mathbf{A}, \mathbf{B}, \mathbf{C}$  and the noise vectors  $\mathbf{w}_t, \mathbf{v}_t$  are of compatible size. The initial state,  $\mathbf{x}_0$ , has a finite  $2 + \epsilon$  moment as in the noise variables. We label the eigenvalues of  $\mathbf{A}$  as  $\lambda_1, \dots, \lambda_n$ . Without loss of generality, we assume that  $\mathbf{A}$  is in real Jordan normal form and that  $|\lambda_i| > 1$  for all  $1 \leq i \leq n$ .

*Assumption 2.1:* The pair  $(\mathbf{A}, \mathbf{B})$  is controllable and the pair  $(\mathbf{C}, \mathbf{A})$  is observable.

The observations are made by the sensor and sent to the controller through a finite capacity channel. At each time stage  $t$ , we allow the sensor to send an encoded value  $q_t \in \{1, \dots, N_t\}$  for some  $N_t \in \mathbb{N}$ . We define the rate of our system at time  $t$  as  $R_t = \log_2(N_t)$ . Now, suppose that the channel is used periodically, every  $T$  time stages. The rate for all time stages is then specified by  $\{N_0, \dots, N_{T-1}\}$ . The average rate is

$$R_{\text{avg}} = \frac{1}{T} \sum_{t=0}^{T-1} R_t. \quad (4)$$

*Information Structure:* At time  $t$ , the sensor maps its information  $I_t^s := \mathbf{y}_{[0,t]} \rightarrow q_t \in \{1, \dots, N_t\}$ . The controller maps its information  $I_t^c := q_{[0,t]} \rightarrow \mathbf{u}_t \in \mathbb{R}^m$ .

Our main result for single-sensor systems is the following:

*Theorem 2.2:* There exists a coding and control policy with average rate  $R_{\text{avg}} \leq 1/(T2n) \sum_{i=1}^n \log_2(|\lambda_i|^{T2n} + \epsilon) + 1$  for some  $\epsilon > 0$  which gives:

- (a) the existence of a unique invariant distribution for  $\{\mathbf{x}_{2nt}\}$ ;
- (b)  $\lim_{t \rightarrow \infty} E[\|\mathbf{x}_{2nt}\|_2^2] < \infty$ ,  $\limsup_{t \rightarrow \infty} E[\|\mathbf{x}_t\|_2^2] < \infty$ .

*Theorem 2.3:* The average rate in Theorem 2.2 achieves the minimum rate (2) asymptotically for large sampling periods. That is,  $\lim_{T \rightarrow \infty} R_{\text{avg}} = R_{\min}$ .

### A. Coding and Control Policy

For now, assume that  $\mathbf{A}$  has only one eigenvalue  $\lambda$ ; that is, this eigenvalue is repeated. We will see later that this assumption is without loss of generality. Let  $K = \lceil |\lambda| + \epsilon \rceil$  for some parameter  $\epsilon > 0$  and consider the following scalar  $(K+1)$ -bin uniform quantizer. Assuming that  $K$  is even, this is defined for  $k \in \{1, 2, \dots, K\}$  as

$$Q_K^\Delta(x) = \begin{cases} \left(\frac{-(K+1)}{2} + k\right)\Delta, & \text{if } x \in \left[\left(\frac{-K}{2} + k - 1\right)\Delta, \left(\frac{-K}{2} + k\right)\Delta\right] \\ \frac{K-1}{2}\Delta, & \text{if } |x| = \frac{K}{2}\Delta \\ 0, & \text{if } |x| > \frac{K}{2}\Delta \end{cases}$$

where  $\Delta \in \mathbb{R}_+$  is the bin size. The set  $[-(K/2)\Delta, (K/2)\Delta]$  is called the granular region while the set  $(-\infty, -(K/2)\Delta) \cup ((K/2)\Delta, \infty)$  is called the overflow region. If the state is in the granular region, that is if  $|x| \leq (K/2)\Delta$  then we say the quantizer is *perfectly-zoomed*. Otherwise, we say it is *under-zoomed* [15].

We write our quantizer as the composite function  $Q_K^\Delta(x) = D_K^\Delta(E_K^\Delta(x))$ . The encoder  $E_K^\Delta : \mathbb{R} \rightarrow \{0, 1, \dots, K\}$  and decoder  $D_K^\Delta : \{0, 1, \dots, K\} \rightarrow \mathcal{C}$  for  $k \in \{1, 2, \dots, K\}$  are

$$E_K^\Delta(x) = \begin{cases} k, & \text{if } x \in \left[\left(\frac{-K}{2} + k - 1\right)\Delta, \left(\frac{-K}{2} + k\right)\Delta\right) \\ K, & \text{if } x = \frac{K}{2}\Delta \\ 0, & \text{if } |x| > \frac{K}{2}\Delta \end{cases}$$

$$D_K^\Delta(x) = \begin{cases} \left(-\frac{K+1}{2} + x\right)\Delta, & \text{if } x \neq 0 \\ 0, & \text{otherwise.} \end{cases}$$

At time  $t$ , we associate with each component  $x_t^i$  a bin size  $\Delta_t^i$ . Let  $q_t^i = E_K^{\Delta_t^i}(y_t^i)$ . We will be applying our control policy to system (8) where  $\mathbf{y}_s$  is a meaningful estimate of the state  $\mathbf{x}_s$ . Let our fixed rate be  $N_t = K^n + 1$  for all  $t \in \mathbb{N}$ . Choose any invertible function  $f : \{1, \dots, K\}^n \rightarrow \{1, \dots, K^n\}$ . We then choose the encoded value

$$q_t = \begin{cases} f(q_1^1, \dots, q_t^n), & \text{if } q_t^i \neq 0 \text{ for all } 1 \leq i \leq n \\ 0, & \text{otherwise.} \end{cases}$$

Upon receiving  $q_t \neq 0$ , the controller knows  $q_t^1, \dots, q_t^n$ . The controller forms the estimate  $\hat{\mathbf{x}}_t$  as  $\hat{\mathbf{x}}_t = [\hat{x}_t^1 \dots \hat{x}_t^n]^T$ , where

$$\hat{x}_t^i = \begin{cases} D_K^{\Delta_t^i}(q_t^i), & \text{if } q_t^i \neq 0 \\ 0, & \text{otherwise.} \end{cases}$$

We assume without loss of generality that  $\mathbf{A}$  is a Jordan block with eigenvalue  $\lambda$ . From the real Jordan canonical form, we know that  $\mathbf{A}$  can be written as

$$\begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix} \text{ or } \begin{bmatrix} D & I & & \\ & D & \ddots & \\ & & \ddots & I \\ & & & D \end{bmatrix}$$

when  $\lambda \in \mathbb{R}$  and  $\lambda \in \mathbb{C}$ , respectively. In the complex case we write  $\lambda = a + ib$  for some  $a, b \in \mathbb{R}$  and define

$$\mathbf{D} = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}.$$

The update equations are

$$\Delta_{t+1} = \bar{Q}(q_t, \Delta_t) \Delta_t \quad (5a)$$

$$\bar{Q}(q_t, \Delta_t) = \begin{cases} \rho |\lambda|, & \text{if } q_t = 0 \\ \beta(\Delta_t), & \text{otherwise} \end{cases} \quad (5b)$$

for some  $\rho > 1$  and with

$$\beta(\Delta_t) = \text{diag}(\beta_1(\Delta_t^1), \dots, \beta_n(\Delta_t^n)) \quad (6a)$$

$$\beta(\Delta_t^i) = \begin{cases} 1, & \text{if } \Delta_t^i \leq L^i, \\ \frac{|\lambda|}{|\lambda| + c - \eta}, & \text{otherwise} \end{cases} \quad (6b)$$

for some  $0 < \eta < \epsilon$  and  $\mathbf{L} \in \mathbb{R}_+^n$ . Note that if we define  $\bar{\mathbf{L}} = \mathbf{L}|\lambda|/(|\lambda| + \epsilon - \eta)$  then  $\Delta_t^i > \bar{\mathbf{L}}^i$  for all  $1 \leq i \leq n$  and all  $t \in \mathbb{N}$ .

*Bin Ordering:* We set  $\mathbf{L} = c\Delta_0$ , for some  $0 < c \leq 1$ . First let  $\lambda \in \mathbb{R}$ . For any  $\delta > 0$  we can choose  $\Delta_0^i$  and  $\Delta_0^{i+1}$  such that  $\Delta_0^{i+1} \leq \delta\Delta_0^i$  for all  $1 \leq i \leq n-1$ . With our update equations and our choice of  $\mathbf{L}$  we get that the ordering is preserved over all time stages. That is,  $\Delta_t^{i+1} \leq \delta\Delta_t^i$  for all  $1 \leq i \leq n-1$  and  $t \in \mathbb{N}$ .

Now let  $\lambda \in \mathbb{C}$ . We choose  $\Delta_0^i = \Delta_0^{i+1}$  for all  $i$  odd. Thus, we have divided the complex modes into their conjugate pairs and set their initial bin sizes to be equal. Our initial condition implies that  $\Delta_t^i = \Delta_t^{i+1}$  for all  $i$  odd and  $t \in \mathbb{N}$ . For any  $\delta > 0$  we can choose  $\Delta_0^i$  and  $\Delta_0^{i+2}$  such that  $\Delta_t^{i+2} \leq \delta\Delta_t^i$  for all  $1 \leq i \leq n-2$  and  $t \in \mathbb{N}$ .

*Control Action:* Under our information structure, the update (5) can be applied at the sensor and the controller. At time  $t$  the controller has access to  $\dot{\mathbf{x}}_t$  and the control action is picked as  $\mathbf{u}_t = -\mathbf{A}\dot{\mathbf{x}}_t$ .

## B. Outline of Proof of Theorem 2.2

In this subsection, we outline the supporting results and key steps in proving our main result for single-sensor systems, Theorem 2.2.

*Lemma 2.4:* We can sample every  $2n$  time stages and apply a similarity transform to  $\mathbf{x}_t$  in (3) to obtain  $\bar{\mathbf{x}}_s = \mathbf{P}\mathbf{x}_{2ns}$  with  $s \in \mathbb{N}$  for some invertible matrix  $\mathbf{P}$ . This new state satisfies the following system of equations:

$$\bar{\mathbf{x}}_{s+1} = \bar{\mathbf{A}}\bar{\mathbf{x}}_s + \bar{\mathbf{u}}_s + \bar{\mathbf{w}}_s, \quad \bar{\mathbf{y}}_s = \bar{\mathbf{x}}_s + \bar{\mathbf{v}}_s. \quad (7)$$

The control action  $\bar{\mathbf{u}}_s \in \mathbb{R}^n$  is chosen arbitrarily by the controller and the elimination of the  $\mathbf{B}$  matrix can be justified by sampling. The estimate  $\bar{\mathbf{y}}_s \in \mathbb{R}^n$  at time  $s$  is known by the sensor.

At time  $s$ ,  $\bar{\mathbf{w}}_s$  and  $\bar{\mathbf{v}}_s$  are independent of  $\bar{\mathbf{x}}_s$  but may be correlated with each other. For  $s_1 \neq s_2$ , the vectors  $\bar{\mathbf{w}}_{s_1}$  and  $\bar{\mathbf{v}}_{s_2}$  are independent. The matrix  $\bar{\mathbf{A}}$  is in real Jordan normal form and has eigenvalues  $\lambda_1^{2n}, \dots, \lambda_n^{2n}$ .

By a slight abuse of notation, we will rewrite system (7) as

$$\mathbf{x}_{s+1} = \mathbf{A}\mathbf{x}_s + \mathbf{u}_s + \mathbf{w}_s, \quad \mathbf{y}_s = \mathbf{x}_s + \mathbf{v}_s \quad (8)$$

where  $\mathbf{x}_s \in \mathbb{R}^n$ ,  $\mathbf{u}_s \in \mathbb{R}^n$  and  $\mathbf{y}_s \in \mathbb{R}^n$  are the state, control action and observation at time  $s$ , respectively.

*Remark 2.5:* In case of multiple Jordan blocks, we can apply our control policy to each Jordan block. In all remaining theorems of this section, we will work with system (8). Where necessary, we will distinguish between the real and complex eigenvalue cases.

*Lemma 2.6:* The process  $\{(\mathbf{x}_s, \Delta_s)\}$  is Markov.

Section II-A gives our control policy in terms of the parameters  $\rho, \epsilon$  and  $\eta$ .

*Lemma 2.7:* For appropriate choices of  $\rho, \epsilon$  and  $\eta$ , we can form a countable state space  $\mathcal{S}$  for  $\{\Delta_s\}$ . The process  $\{(\mathbf{x}_s, \Delta_s)\}$  is an irreducible and aperiodic Markov chain on  $\mathbb{R}^n \times \mathcal{S}$ , see [21] for a similar construction.

Define the sequence of stopping times, with  $\tau_0 = 0$  and,

$$\tau_{z+1} = \min \left\{ s > \tau_z : |\mathbf{y}_s| = |\mathbf{x}_s + \mathbf{v}_s| \leq \frac{K}{2} \Delta_s \right\}.$$

These are the times when all quantizers are perfectly-zoomed. We assume that this is satisfied at time  $s = 0$ . This technical condition is justified by showing that the process  $\{(\mathbf{x}_s, \Delta_s)\}$  moves to such a perfectly zoomed state in a random time which is dominated by a geometric distribution (see the proof of Proposition 3.2 in [21]).

*Lemma 2.8:* If  $K$  is even, then the following hold.

- (a) For any  $r > 0$  and any polynomial of finite degree  $Q(k)$  there exists a sufficiently large  $H$  such that

$$Q(k)P(\tau_{z+1} - \tau_z > k | \mathbf{x}_{\tau_z}, \Delta_{\tau_z}) \leq r^{-k}$$

for all  $k > H$  and for all  $z \in \mathbb{N}$ .

- (b) Let  $\Delta_{\tau_z} \rightarrow \infty$  be equivalent to stating that  $\Delta_{\tau_z^i} \rightarrow \infty$  for all  $1 \leq i \leq n$ . Then

$$\lim_{\Delta_{\tau_z} \rightarrow \infty} P(\tau_{z+1} - \tau_z > 1 | \mathbf{x}_{\tau_z}, \Delta_{\tau_z}) = 0$$

uniformly in  $\mathbf{x}_{\tau_z}$ .

We define the compact sets

$$S = S_{\mathbf{x}} \times S_{\Delta},$$

$$S_{\Delta} = \left\{ \Delta \in \mathbb{R}_+^n : \Delta^i \leq F, 1 \leq i \leq n \right\},$$

$$S_{\mathbf{x}} = \left\{ \mathbf{x} \in \mathbb{R}^n : |x^i| \leq \frac{K}{2} F, 1 \leq i \leq n \right\}$$

for some  $F > L^1$  where  $L^1$  is a component of  $\mathbf{L}$  as described in Section II-A. Note that at the stopping time  $\tau_z$ , if  $\Delta_{\tau_z} \in S_{\Delta}$  then  $|x_{\tau_z}^i| \leq (K/2)\Delta_{\tau_z}^i \leq (K/2)F$ , for all  $1 \leq i \leq n$ , and thus  $\mathbf{x}_{\tau_z} \in S_{\mathbf{x}}$  and  $(\mathbf{x}_{\tau_z}, \Delta_{\tau_z}) \in S$ .

*Lemma 2.9:* For some  $\gamma > 0, b < \infty$ , the following drift condition holds:

$$\begin{aligned} \gamma E \left[ \sum_{s=\tau_z}^{\tau_{z+1}-1} (\Delta_s^1)^2 \mid \mathbf{x}_{\tau_z}, \Delta_{\tau_z} \right] &\leq (\Delta_{\tau_z}^1)^2 \\ &- E \left[ (\Delta_{\tau_z+1}^1)^2 \mid \mathbf{x}_{\tau_z}, \Delta_{\tau_z} \right] + b \mathbf{1}_{\{(\mathbf{x}_{\tau_z}, \Delta_{\tau_z}) \in S\}}. \end{aligned} \quad (9)$$

For  $\lambda \in \mathbb{C}$ , the above also holds with  $\Delta^2$  in place of  $\Delta^1$ .

For  $\mathbf{x} \in \mathbb{R}^n$ , we say that  $x^i$  and  $x^{i+1}$  are a *conjugate pair* if  $i$  is odd. To simplify notation in the complex eigenvalue case we find it convenient to define for any  $\mathbf{x} \in \mathbb{R}^n$ , the set of vectors  $\tilde{\mathbf{x}}^i = [x^i \ x^{i+1}]^T$ , if  $i$  is odd,  $\tilde{\mathbf{x}}^i = [x^{i-1} \ x^i]^T$ , if  $i$  is even, for  $1 \leq i \leq n$ . Note that  $\tilde{\mathbf{x}}^i = \tilde{\mathbf{x}}^{i+1}$  for  $i$  odd. We are only concerned with the case when  $n$  is even.

*Theorem 2.10:* Let  $\lambda \in \mathbb{R}$ . For  $i = n$ , there exists a  $\kappa > 0$  such that

$$E \left[ \sum_{s=\tau_z}^{\tau_{z+1}-1} (\tilde{\mathbf{x}}_s^i)^2 | \mathbf{x}_{\tau_z}, \Delta_{\tau_z} \right] \leq \kappa (\Delta_{\tau_z}^1)^2. \quad (10)$$

If  $\lim_{s \rightarrow \infty} E[(x_s^k)^2] < \infty$  then the above holds for  $i = k - 1$ .

For  $\lambda \in \mathbb{C}$ , with  $i = n - 1$ , there exists a  $\kappa > 0$  such that

$$E \left[ \sum_{s=\tau_z}^{\tau_{z+1}-1} (\tilde{\mathbf{x}}_s^i)^T \tilde{\mathbf{x}}_s^i | \mathbf{x}_{\tau_z}, \Delta_{\tau_z} \right] \leq \kappa (\tilde{\Delta}_{\tau_z}^1)^T \tilde{\Delta}_{\tau_z}^1.$$

If  $\lim_{s \rightarrow \infty} E[(\tilde{\mathbf{x}}_s^k)^T \tilde{\mathbf{x}}_s^k] < \infty$  then the above holds for  $i = k - 2$ .

*Proof of Theorem 2.2:* [(a)] We know from Lemmas 2.6 and 2.7 that the process  $\{(\mathbf{x}_s, \Delta_s)\}$  is an irreducible and aperiodic Markov chain. The set  $S$  is small (see [21]). Using Lemma 2.9 we can apply Theorem 5.1 with  $a = 1$ , the Markov chain  $\{(\mathbf{x}_s, \Delta_s)\}$  and the functions  $V(\mathbf{x}_s, \Delta_s) = (\Delta_s^1)^2$ ,  $\beta(\mathbf{x}_s, \Delta_s) = 1$  and  $b$  as given in Lemma 2.9 to get that  $\{(\mathbf{x}_s, \Delta_s)\}$  is positive Harris recurrent and has a unique invariant distribution.

[(b)] Suppose that  $\lambda \in \mathbb{R}$ . We will apply Theorem 5.1 with  $a = 0$ , the Markov chain  $\{(\mathbf{x}_s, \Delta_s)\}$  and the functions  $V(\mathbf{x}_s, \Delta_s) = (\Delta_s^1)^2$ ,  $\beta(\mathbf{x}_s, \Delta_s) = \gamma(\Delta_s^1)^2$ ,  $f(\mathbf{x}_s, \Delta_s) = (\gamma/\kappa)(x_s^n)^2$ . From Lemma 2.9, we get

$$\begin{aligned} & E[V(\mathbf{x}_{\tau_{z+1}}, \Delta_{\tau_{z+1}}) | \mathcal{F}_{\tau_z}] \\ &= E[(\Delta_{\tau_{z+1}}^1)^2 | \mathbf{x}_{\tau_z}, \Delta_{\tau_z}] \\ &\leq (\Delta_{\tau_z}^1)^2 - \gamma E \left[ \sum_{s=\tau_z}^{\tau_{z+1}-1} (\Delta_s^1)^2 | \mathbf{x}_{\tau_z}, \Delta_{\tau_z} \right] \\ &\quad + b1_{\{(\mathbf{x}_{\tau_z}, \Delta_{\tau_z}) \in S\}} \\ &\leq (\Delta_{\tau_z}^1)^2 - \gamma(\Delta_{\tau_z}^1)^2 + b1_{\{(\mathbf{x}_{\tau_z}, \Delta_{\tau_z}) \in S\}} \\ &= V(\mathbf{x}_{\tau_z}, \Delta_{\tau_z}) - \beta(\mathbf{x}_{\tau_z}, \Delta_{\tau_z}) + b1_{\{(\mathbf{x}_{\tau_z}, \Delta_{\tau_z}) \in S\}} \end{aligned}$$

Lemma 2.10 holds for  $\{x_s^n\}$  and thus

$$\begin{aligned} E \left[ \sum_{s=\tau_z}^{\tau_{z+1}-1} f(\mathbf{x}_s, \Delta_s) | \mathcal{F}_{\tau_z} \right] &= \frac{\gamma}{\kappa} E \left[ \sum_{s=\tau_z}^{\tau_{z+1}-1} (x_s^n)^2 | \mathbf{x}_{\tau_z}, \Delta_{\tau_z} \right] \\ &\leq \gamma (\Delta_{\tau_z}^2)^2 = \beta(\mathbf{x}_{\tau_z}, \Delta_{\tau_z}), \end{aligned}$$

where we have used the ordering of bin sizes as described in Section II-A.

Thus,  $\lim_{s \rightarrow \infty} (\gamma/\kappa) E[(x_s^n)^2] < \infty$  by Theorem 5.1 and so  $\lim_{s \rightarrow \infty} E[(x_s^n)^2] < \infty$ . This implies that Theorem 2.10 holds for  $\{x_s^{n-1}\}$  as mentioned in the proof and theorem statement. The finite second moment of all components then follows by induction.

In the complex case, we have that the drift condition (9) in Lemma 2.9 also holds with  $\Delta_s^2$  in place of  $\Delta_s^1$  since they are equal. Choosing the functions  $V(\mathbf{x}_s, \Delta_s) = (\Delta_s^1)^2 + (\Delta_s^2)^2$ ,  $\beta(\mathbf{x}_s, \Delta_s) = \gamma((\Delta_s^1)^2 + (\Delta_s^2)^2)$ ,  $f(\mathbf{x}_s, \Delta_s) = (\gamma/\kappa)(\tilde{\mathbf{x}}_s^n)^T \tilde{\mathbf{x}}_s^n$ , we obtain the result.  $\square$

### III. MULTI-SENSOR SYSTEMS

This is the main problem of the technical note, as stated in Section I-A. Let us label the Jordan blocks of  $\mathbf{A}$  as  $\mathbf{J}_1, \dots, \mathbf{J}_\ell$ . Let  $V_i$  be the (possibly generalized) eigenspace corresponding to  $\mathbf{J}_i$ . That is, if  $\mathbf{v}_{i,1}, \dots, \mathbf{v}_{i,d_i}$  are the (possibly generalized) eigenvectors associated with  $\mathbf{J}_i$  then  $V_i = \text{span}\{\mathbf{v}_{i,1}, \dots, \mathbf{v}_{i,d_i}\}$  and has dimension  $d_i$ .

#### A. Case With a Technical Assumption on Observability

*Assumption 3.1:* Each eigenspace is observed by some sensor. That is, for each  $1 \leq i \leq \ell$  there exists a  $1 \leq j \leq M$  such that  $V_i \subseteq O^j$ .

The following is the main result of this subsection:

*Theorem 3.2:* Under Assumption 3.1, there exists a coding and control policy with average rate

$$R_{\text{avg}} \leq 1/(T2n) \left( M + \sum_{i=1}^n \log_2 \left( \lceil |\lambda_i|^{T2n} + \epsilon \rceil + 1 \right) \right)$$

for some  $\epsilon > 0$  which gives:

- (a) the existence of a unique invariant distribution for  $\{\mathbf{x}_{2nt}\}$ ;
- (b)  $\lim_{t \rightarrow \infty} E[\|\mathbf{x}_{2nt}\|_2^2] < \infty$ ,  $\limsup_{t \rightarrow \infty} E[\|\mathbf{x}_t\|_2^2] < \infty$ .

*Theorem 3.3:* The average rate in Theorem 3.2 achieves the minimum rate (2) asymptotically for large sampling periods. That is,  $\lim_{T \rightarrow \infty} R_{\text{avg}} = R_{\min}$ .

The proof of Theorem 3.2 is basically an application of the Jordan normal transformation together with Assumption 3.1.

*Proof of Theorem 3.2:* Under Assumption 3.1, we can assign each eigenspace  $V_i \subseteq O^j$  to some sensor  $j$ . Let  $V_{j,1}, \dots, V_{j,m_j}$  denote the eigenspaces assigned to sensor  $j$  and let us write  $V_{j,i} = \text{span}\{\mathbf{v}_{j,i,1}, \dots, \mathbf{v}_{j,i,d_{j,i}}\}$  where each  $\mathbf{v}_{j,i,h} \in \mathbb{R}^{1 \times n}$  is a generalized eigenvector. We put  $\mathbf{Q}_{j,i} = [(\mathbf{v}_{j,i,1})^T \cdots (\mathbf{v}_{j,i,d_{j,i}})^T]^T$ ,  $\mathbf{Q}_j = [\mathbf{Q}_{j,1}^T \cdots \mathbf{Q}_{j,m_j}^T]^T$  and  $\mathbf{Q} = [(\mathbf{Q}_M)^T \cdots (\mathbf{Q}_1)^T]^T$ . We apply the similarity transform  $\bar{\mathbf{x}}_t = \mathbf{Q}\mathbf{x}_t$  to (1) and define  $\bar{\mathbf{A}} = \mathbf{Q}\mathbf{A}\mathbf{Q}^{-1}$ ,  $\bar{\mathbf{B}} = \mathbf{Q}\mathbf{B}$  and  $\bar{\mathbf{w}}_t = \mathbf{Q}\mathbf{w}_t$  to get

$$\bar{\mathbf{x}}_{t+1} = \bar{\mathbf{A}}\bar{\mathbf{x}}_t + \bar{\mathbf{B}}\bar{\mathbf{u}}_t + \bar{\mathbf{w}}_t. \quad (11)$$

We now look at the estimation of the state by the sensors. For convenience, let us write  $\bar{\mathbf{x}}_t = [(\bar{\mathbf{x}}_t^M)^T \cdots (\bar{\mathbf{x}}_t^1)^T]^T$  where  $\bar{\mathbf{x}}_t^j = [(\bar{\mathbf{x}}_t^{j,1})^T \cdots (\bar{\mathbf{x}}_t^{j,m_j})^T]^T$  and  $\bar{\mathbf{x}}_t^{j,i} = [\bar{x}_t^{j,i,1} \cdots \bar{x}_t^{j,i,d_{j,i}}]^T$  with  $\bar{x}_t^{j,i,h} \in \mathbb{R}$ . Let us write  $\mathcal{O}_{(\mathbf{C}^j, \mathbf{A})} = [(\mathbf{o}_{j,1})^T \cdots (\mathbf{o}_{j,n_{p,j}})^T]^T$  where each  $\mathbf{o}_{j,i} \in \mathbb{R}^{1 \times n}$ . With our construction above, under Assumption 3.1, we have for each  $j, i, h$  that  $\mathbf{v}_{j,i,h} = \sum_{\ell=1}^{n_{p,j}} k_{\ell}^{j,i,h} \mathbf{o}_{j,\ell}$  for some real coefficients  $\{k_{\ell}^{j,i,h}\}$ . Consider the first  $n$  time stages. By putting  $\mathbf{k}^{j,i,h} = [k_1^{j,i,h} \cdots k_{n_{p,j}}^{j,i,h}]$ , it follows that:

$$\begin{aligned} & \mathbf{k}^{j,i,h} \left[ \left( \mathbf{y}_0^j \right)^T \cdots \left( \mathbf{y}_{n-1}^j \right)^T \right]^T \\ &= \mathbf{k}^{j,i,h} \mathcal{O}_{(\mathbf{C}^j, \mathbf{A})} \mathbf{x}_0 + \bar{v}_0^{j,i,h} \\ &= \sum_{\ell=1}^{n_{p,j}} k_{\ell}^{j,i,h} \mathbf{o}_{j,\ell} \mathbf{x}_0 + \bar{v}_0^{j,i,h} = \mathbf{v}_{j,i,h} \mathbf{x}_0 + \bar{v}_0^{j,i,h} = \bar{x}_0^{j,i,h} + \bar{v}_0^{j,i,h}, \end{aligned}$$

where  $\bar{v}_0^{j,i,h}$  is some noise term. We will use the same notation for  $\bar{\mathbf{v}}_0^j$  that we use for  $\bar{\mathbf{x}}_t^j$ . As in Lemma 2.4 for the single-sensor case, we can use the next  $n$  time stages to apply a control action. We then apply the above scheme repeatedly and sample every  $2n$  time stages. Furthermore, since  $\mathbf{Q}$  is the Jordan normal transformation matrix, it follows that  $\bar{\mathbf{A}}^{2n} = \text{diag}(\bar{\mathbf{J}}_1^{2n}, \dots, \bar{\mathbf{J}}_\ell^{2n})$  where each  $\bar{\mathbf{J}}_i \in \mathbb{R}^{d_i \times d_i}$  is a Jordan block. Since we can apply another Jordan transformation to this sampled system, we can assume without loss of generality that  $\bar{\mathbf{A}}^{2n}$  is actually in Jordan form and each  $\bar{\mathbf{J}}_i^{2n}$  is actually a Jordan block.

To simplify notation, we write  $\mathbf{A} := \bar{\mathbf{A}}^{2n} = \text{diag}(\mathbf{J}_1, \dots, \mathbf{J}_\ell)$  where each  $\mathbf{J}_j \in \mathbb{R}^{d_i \times d_i}$  is a Jordan block,  $\mathbf{x}_s = [(\mathbf{x}_s^1)^T \cdots (\mathbf{x}_s^\ell)^T]^T$  where  $\mathbf{x}_s^i \in \mathbb{R}^{d_i}$  and similarly for  $\mathbf{u}_s, \mathbf{w}_s, \mathbf{y}_s$  and  $\mathbf{v}_s$ . From the above, we can also see that for each  $i$ , there exists a  $j$  such that  $\mathbf{x}_s^i + \mathbf{v}_s^i$  is known by sensor  $j$  at time  $s$ .

Thus our system is equivalent to the following subsystems:

$$\mathbf{x}_{s+1}^i = \mathbf{J}_i \mathbf{x}_s^i + \mathbf{u}_s^i + \mathbf{w}_s^i, \quad \mathbf{y}_s^i = \mathbf{x}_s^i + \mathbf{v}_s^i \quad 1 \leq i \leq \ell$$

where for each  $1 \leq i \leq \ell$ , there exists a sensor  $j$  which knows  $\mathbf{y}_s^i$  at time  $s$  and  $\mathbf{u}_s^i$  is chosen by the controller.

As in Section II-A, we let  $\Delta_s$  be the vector of bin sizes at time  $s$  and define the sequence of stopping times

$$\tau_0 = 0, \tau_{z+1} = \min \{s > \tau_z : |\mathbf{y}_s| = |\mathbf{x}_s + \mathbf{v}_s| \leq \Delta_s\}.$$

The feedback value  $b_{2ns}$  is chosen as

$$b_{2ns} = \begin{cases} 1, & \text{if } s = \tau_z \text{ for some } z \in \mathbb{N}, \\ 0, & \text{otherwise,} \end{cases}$$

so that the policy as in Section II-A is implementable at the sensors and at the controller. This reduces the problem to the single-sensor case and we obtain the result.  $\square$

### B. Sufficient Conditions for the General Multi-Sensor Case

Here, we do not assume Assumption 3.1. In this case, the lower components of the state act as noise for the upper components. In particular, we need to bound these lower modes when all quantizers are perfectly-zoomed to achieve (b) of Lemma 2.8. To do this, we must have that the bin sizes of the lower modes are small compared with the upper ones. With many different eigenvalues, we cannot guarantee this in the general case. For this case, we obtain a sufficient condition. Toward this end, we have the following theorem which extends the classical observability canonical decomposition [23] to the decentralized case.

**Theorem 3.4:** [3] Under Assumption 1.1, there exists a matrix  $\mathbf{Q}$  such that if we define  $\bar{\mathbf{A}} = \mathbf{Q} \mathbf{A} \mathbf{Q}^{-1}$  and  $\bar{\mathbf{C}}^j = \mathbf{C}^j \mathbf{Q}^{-1}$  then

$$\bar{\mathbf{A}} = \begin{bmatrix} \bar{\mathbf{A}}_M & * & \cdots & * \\ & \bar{\mathbf{A}}_{M-1} & \cdots & * \\ & & \ddots & \\ & & & \bar{\mathbf{A}}_1 \end{bmatrix} \quad (12a)$$

$$\begin{bmatrix} \bar{\mathbf{C}}^M \\ \bar{\mathbf{C}}^{M-1} \\ \vdots \\ \bar{\mathbf{C}}^2 \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{C}}_O^M & * & \cdots & * \\ & \bar{\mathbf{C}}_O^{M-1} & \cdots & * \\ & & \ddots & \\ & & & \bar{\mathbf{C}}_O^2 \end{bmatrix} \quad (12b)$$

where the \*'s denote irrelevant submatrices, each  $\bar{\mathbf{A}}_j \in \mathbb{R}^{n_j \times n_j}$  and each  $\bar{\mathbf{C}}_O^j \in \mathbb{R}^{p_j \times n_j}$ .

Below, we give a sufficient rate and an alternative assumption for stability. For Theorem 3.5 below, let us write  $\Lambda(\bar{\mathbf{A}}_j) = \{\lambda_{j,1}, \dots, \lambda_{j,n_j}\}$  where  $\bar{\mathbf{A}}_j$  is given in (12).

**Theorem 3.5:** There exists a coding and control policy which gives:

- (a) the existence of a unique invariant distribution for  $\{\mathbf{x}_{2nt}\}$ ;
  - (b)  $\lim_{t \rightarrow \infty} E[\|\mathbf{x}_{2nt}\|_2^2] < \infty$ ,  $\limsup_{t \rightarrow \infty} E[\|\mathbf{x}_t\|_2^2] < \infty$ ,
- and with average rate in the limit of large sampling periods

$$\lim_{T \rightarrow \infty} R_{\text{avg}} = \sum_{j=1}^M \sum_{i=1}^{n_j} \log_2 (\max\{|\lambda_{j,i}|, |\lambda_{h,\ell}| : h < j, 1 \leq \ell \leq n_h\}).$$

**Proof of Theorem 3.5:** The proof follows that of Theorem 2.2. The main difference is that we define  $\lambda'_{j,i} = \max\{|\lambda_{j,i}|, |\lambda_{h,\ell}| : h < j, 1 \leq \ell \leq n_h\}$  and the bin numbers  $K_{j,i} = \lceil (\lambda'_{j,i})^{2n} + \epsilon \rceil$  for some  $\epsilon > 0$  and treat the lower components of the state as noise. We present the coding and the control policy in detail below.  $\square$

Clearly, we could also achieve (a) and (b) in Theorem 3.5 with  $\lim_{T \rightarrow \infty} R_{\text{avg}} = n \log_2 (\lambda_{\text{absmax}})$  where  $\lambda_{\text{absmax}} = \max_{j,i} \{|\lambda_{j,i}|\}$ .

**Theorem 3.6:** If the eigenvalues of  $\bar{\mathbf{A}}_M, \dots, \bar{\mathbf{A}}_1$  in (12) are ordered in decreasing magnitude then Theorem 3.2 holds without Assumption 3.1. That is, the theorem holds if for  $\lambda_i \in \Lambda(\bar{\mathbf{A}}_i)$  and  $\lambda_j \in \Lambda(\bar{\mathbf{A}}_j)$  we have that  $|\lambda_i| \leq |\lambda_j|$  when  $i < j$ .

**Proof of Theorem 3.6:** The proof follows directly from that of Theorem 3.5. Since the eigenvalues are ordered in decreasing magnitude, we can maintain the ordering of the bin sizes given in (13) without increasing the rate. Specifically, in the Proof of Theorem 3.5 we see that  $|\lambda'_{j,i}| = |\lambda_{j,i}|$  for all  $1 \leq j \leq M, 1 \leq i \leq m_j$ .  $\square$

**Coding and Control Policy for the General Case:** Consider the system (1). Sampling, observing and controlling as in the proof of Theorem 2.2 and applying the transformation  $\bar{\mathbf{x}}_t = \mathbf{Q} \mathbf{x}_t$  where  $\mathbf{Q}$  is given in Theorem 3.4, we obtain the system

$$\mathbf{x}_{s+1} = \mathbf{A} \mathbf{x}_s + \mathbf{u}_s + \mathbf{w}_s, \quad \mathbf{y}_s = \mathbf{x}_s + \mathbf{v}_s.$$

We do not relabel the variables (for example  $\bar{\mathbf{x}}_s$ ) by a slight abuse of notation. In the above,  $\mathbf{A}$  is block upper triangular with the blocks  $\mathbf{A}_M, \dots, \mathbf{A}_1$  descending along the diagonal and each  $\mathbf{A}_j \in \mathbb{R}^{n_j \times n_j}$  as in (12a) of Theorem 3.4. Since we can always apply a block transformation to  $\mathbf{A}$  in which each of the blocks is the Jordan transformation of  $\mathbf{A}_j$ , we can assume without loss of generality that each  $\mathbf{A}_j$  is in real Jordan normal form and we write  $\mathbf{A}_j = \text{diag}(\mathbf{A}_{j,1}, \dots, \mathbf{A}_{j,m_j})$  where each  $\mathbf{A}_{j,i} \in \mathbb{R}^{d_{j,i} \times d_{j,i}}$  is a Jordan block.

Let us write

$$\mathbf{x}_s = \left[ \left( \mathbf{x}_s^M \right)^T \cdots \left( \mathbf{x}_s^2 \right)^T \right]^T$$

where

$$\mathbf{x}_s^j = \left[ \left( \mathbf{x}_s^{j,1} \right)^T \cdots \left( \mathbf{x}_s^{j,m_j} \right)^T \right]^T$$

and  $\mathbf{x}_s^j \in \mathbb{R}^{n_j}$  with

$$\mathbf{x}_s^{j,i} = \left[ x_s^{j,i,1} \cdots x_s^{j,i,d_{j,i}} \right]^T$$

and each  $x_s^{j,i,h} \in \mathbb{R}$ . We will use the same notational convention for all relevant vectors in this section. Namely, we will follow this convention for  $\mathbf{u}_s, \mathbf{w}_s, \mathbf{y}_s, \mathbf{v}_s$  and for  $\Delta_s, \hat{\mathbf{x}}_s, \mathbf{L}$  which will be specified.

From the proof of Theorem 3.4, we know that the rows of  $\mathbf{Q}$  are taken from the row spaces of  $\{\mathcal{O}_{(\mathbf{C}^j, \mathbf{A})}\}$  and we can see how  $\mathbf{y}_s^j$  is known by sensor  $j$  at time  $s$ .

Let us denote the eigenvalue of  $\mathbf{A}_{j,i}$  by  $\lambda_{j,i}$  and define

$$|\lambda'_{j,i}| = \max \{|\lambda_{j,i}|, |\lambda_{h,\ell}| : 1 \leq h < j, 1 \leq \ell \leq m_h\}.$$

Let  $K_{j,i} = \lceil |\lambda'_{j,i}| + \delta + \epsilon \rceil$  for some  $\delta, \epsilon > 0$ . Let  $\mathbf{K}_j = \text{diag}(K_{j,1} \mathbf{I}, \dots, K_{j,m_j} \mathbf{I})$  where each  $\mathbf{I}$  is the identity matrix of appropriate size so that  $K_{j,i} \mathbf{I} \in \mathbb{R}^{d_{j,i} \times d_{j,i}}$ . Let  $\mathbf{K} = \text{diag}(\mathbf{K}_M, \dots, \mathbf{K}_1)$ .

Let  $\Delta_s^{j,i,h}$  be the bin size corresponding to the component  $x_s^{j,i,h}$  at time  $s$ . We let  $q_s^{j,i,h} = \mathcal{E}_{K_{j,i}}^{\Delta_s^{j,i,h}}(y_s^{j,i,h})$ . Let our fixed rate for sensor  $j$  be  $N^j = (\prod_{i=1}^{m_j} K_{j,i}^{d_{j,i}}) + 1$  for all  $s \in \mathbb{N}$ . Choose any invertible function

$$f_j : \prod_{i=1}^{m_j} \{1, \dots, K_{j,i}\}^{d_{j,i}} \rightarrow \left\{ 1, \dots, \prod_{i=1}^{m_j} K_{j,i}^{d_{j,i}} \right\}.$$

$$q_s^j = \begin{cases} f_j \left( q_s^{j,1,1}, \dots, q_s^{j,m_j,d_{j,i}} \right), & \text{if } q_s^{j,i,h} \neq 0 \text{ for all } 1 \leq i \leq m_j, 1 \leq h \leq d_{j,i}, \\ 0, & \text{otherwise.} \end{cases}$$

We then choose the encoded value shown in the equation at the top of the page. Upon receiving  $q_s^j \neq 0$ , the controller knows  $\{q_s^{j,i,h} : 1 \leq i \leq m_j, 1 \leq h \leq d_{j,i}\}$ . The controller forms the estimate  $\hat{\mathbf{x}}_s^j$  as  $\hat{\mathbf{x}}_s^j = [(\hat{\mathbf{x}}_s^{j,1})^T \dots (\hat{\mathbf{x}}_s^{j,m_j})^T]^T$ , where

$$\hat{\mathbf{x}}_s^{j,i} = \begin{bmatrix} \hat{x}_s^{j,i,1} & \dots & \hat{x}_s^{j,i,d_{j,i}} \end{bmatrix}^T$$

and

$$\hat{x}_s^{j,i,h} = \begin{cases} \mathcal{D}_{K_{j,i}}^{\Delta_s^{j,i,h}}(q_s^{j,i,h}), & \text{if } q_s^2, \dots, q_s^M \neq 0, \\ 0, & \text{otherwise.} \end{cases}$$

*Bin Ordering:* From (12a) of Theorem 3.4, we can write  $\mathbf{A} = [(\mathbf{A}_M^R)^T \dots (\mathbf{A}_1^R)^T]^T$  where each  $\mathbf{A}_j^R \in \mathbb{R}^{n_j \times n}$ . We can further write  $\mathbf{A}_j^R = [\mathbf{0} \ \mathbf{A}_j \ \mathbf{M}_j]$  where  $\mathbf{A}_j = \text{diag}(\mathbf{A}_{j,1}, \dots, \mathbf{A}_{j,m_j})$  and  $\mathbf{M}_j = [(\mathbf{M}_{j,1})^T \dots (\mathbf{M}_{j,m_j})^T]^T$  with  $\mathbf{M}_{j,i} = [\mathbf{M}_{j,i,j-1} \dots \mathbf{M}_{j,i,1}]$  and each  $\mathbf{M}_{j,i,h} \in \mathbb{R}^{d_{j,i} \times n_h}$ . Recall from the proof of Theorem 3.4 that  $n_h = \dim(O^h) - \dim(O^h \cap (\cup_{i=1}^{h-1} O^i))$ .

Let us denote the entries of  $\mathbf{M}_{j,i,h}$  by  $\{m_{j,i,h}^{k,\ell}\}$  and define the entry with maximum absolute value as  $\kappa_{j,i,h} = \max_{k,\ell} \{|m_{j,i,h}^{k,\ell}|\}$ . We set  $\mathbf{L} = c\Delta_0$ , for some  $0 < c \leq 1$ . For any  $\delta > 0$ , by our coding and control policy (and in particular the choice of  $\{|\lambda'_{j,i}|\}$ ) given above, we can choose  $\Delta_0$  such that the following ordering is maintained for all  $s \in \mathbb{N}$ :

$$\Delta_s^{j,i,h+1} \leq \frac{\delta}{j} \Delta_s^{j,i,h}, \quad 1 \leq j \leq M, 1 \leq i \leq m_j \\ 1 \leq h \leq d_{j,i} - 1 \quad (13)$$

$$\Delta_s^{j,k,1} \leq \Delta_s^{j,i,1}, \quad 1 \leq j \leq M, 1 \leq i \leq m_j - 1 \\ i < k \leq m_j \quad (14)$$

$$\Delta_s^{k,\ell,h} \leq \frac{\delta}{jn_K \kappa_{j,i,k}} \Delta_s^{j,i,d_{j,i}} \\ 2 \leq j \leq M, 1 \leq i \leq m_j, 1 \leq k \leq j - 1 \\ 1 \leq \ell \leq m_K, 1 \leq h \leq d_{k,\ell}. \quad (15)$$

Informally, we order the bins within Jordan blocks  $\mathbf{A}_{j,i}$ , within sensor blocks  $\mathbf{A}_j$ , and also between sensor blocks  $\mathbf{A}_j$ .

#### IV. CONCLUSION

We presented a coding and control policy which achieves the minimum rate asymptotically in the limit of large sampling periods for single-sensor and multi-sensor systems (under the assumption that each eigenspace is observed by some sensor) driven by unbounded noise having distributions with a finite  $2 + \epsilon$  moment. In the absence of the aforementioned observability assumption, we obtained sufficient conditions. We established the existence of a unique invariant distribution for the sampled state and a finite second moment of the state. Future work includes the extension to noisy channels as well as multi-controller systems, which introduce further intricacies in view of the presence of signaling among the controllers [2].

#### APPENDIX STOCHASTIC STABILITY

Let  $\phi = \{\phi_t, t \geq 0\}$  be a Markov chain with a complete separable metric state space  $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$ . Let  $P(\mathbf{x}, D) := P(\phi_{t+1} \in D | \phi_t = \mathbf{x})$  denote the transition probability from  $\mathbf{x}$  to the set  $D$ . For this chain, a probability measure  $\pi$  is *invariant* on the Borel space  $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$  if  $\pi(D) = \int_{\mathbb{X}} P(\mathbf{x}, D) \pi(d\mathbf{x})$ , for all  $D \in \mathcal{B}(\mathbb{X})$ .

In the following, let  $\mathcal{F}_t$  denote the filtration generated by the random sequence  $\{\phi_{[0,t]}\}$ . Define a sequence of stopping times  $\{\mathcal{T}_i : i \in \mathbb{N}_+\}$ , measurable on the filtration described above, which is assumed to be non-decreasing, with  $\mathcal{T}_0 = 0$ .

*1) Theorem 5.1: (Theorem 2.1 and Remark 2.1 of [21]):* Suppose that we have a  $\mu$ -irreducible and aperiodic Markov chain  $\phi$ . Suppose moreover that there are functions  $V : \mathbb{X} \rightarrow [a, \infty)$ ,  $\beta : \mathbb{X} \rightarrow [a, \infty)$ ,  $f : \mathbb{X} \rightarrow [a, \infty)$ , for some  $a \geq 0$ , small set  $C$  on which  $V$  is bounded, constant  $b \in \mathbb{R}$  and consider:

$$E[V(\phi_{\mathcal{T}_{p_z+1}}) | \mathcal{F}_{\mathcal{T}_{p_z}}] \leq V(\phi_{\mathcal{T}_{p_z}}) - \beta(\phi_{\mathcal{T}_{p_z}}) + b \mathbf{1}_{\{\phi_{\mathcal{T}_{p_z}} \in C\}} \quad (16)$$

$$E \left[ \sum_{k=\mathcal{T}_{p_z}}^{\mathcal{T}_{p_z+1}-1} f(\phi_k) | \mathcal{F}_{\mathcal{T}_{p_z}} \right] \leq \beta(\phi_{\mathcal{T}_{p_z}}), \quad z \geq 0. \quad (17)$$

If  $a = 1$  and (16) holds then  $\phi$  is positive Harris recurrent with a unique invariant distribution  $\pi$ . If  $a = 0$ , (16), (17) hold and  $\phi$  is positive Harris recurrent with some unique invariant distribution  $\pi$  then we get that  $\lim_{t \rightarrow \infty} E[f(\phi_t)] < \infty$ .

#### REFERENCES

- [1] A. S. Matveev and A. V. Savkin, *Estimation and Control Over Communication Networks*. Boston, MA: Birkhäuser, 2008.
- [2] S. Yüksel and T. Başar, *Stochastic Networked Control Systems: Stabilization and Optimization Under Information Constraints*. Boston, MA: Birkhäuser, 2013.
- [3] A. Johnston, “Networked Control Systems With Unbounded Noise Under Information Constraints,” M.Sc. Thesis, Queen’s University, London, U.K., 2012.
- [4] G. N. Nair and R. J. Evans, “Stabilizability of stochastic linear systems with finite feedback data rates,” *SIAM J. Control and Optimiz.*, vol. 43, pp. 413–436, Jul. 2004.
- [5] S. Tatikonda and S. Mitter, “Control under communication constraints,” *IEEE Trans. Autom. Control*, vol. 49, no. 7, pp. 1056–1068, Jul. 2004.
- [6] N. Elia and S. K. Mitter, “Stabilization of linear systems with limited information,” *IEEE Trans. Autom. Control*, vol. 46, no. 9, pp. 1384–1400, Sep. 2001.
- [7] W. S. Wong and R. W. Brockett, “Systems with finite communication bandwidth constraints—Part ii: Stabilization with limited information feedback,” *IEEE Trans. Autom. Control*, vol. 42, pp. 1294–1299, Sep. 1997.
- [8] N. C. Martins and M. A. Dahleh, “Feedback control in the presence of noisy channels: ‘Bode-like’ fundamental limitations of performance,” *IEEE Trans. Autom. Control*, vol. 53, no. 8, pp. 1604–1615, Aug. 2008.
- [9] S. Tatikonda, “Some scaling properties of large distributed control systems,” in *Proc. 42nd IEEE CDC*, Dec. 2003, pp. 3142–3147.
- [10] G. N. Nair, R. J. Evans, and P. E. Cainfes, “Stabilising decentralised linear systems under data rate constraints,” in *Proc. IEEE Conf. Decision and Control*, Dec. 2004, pp. 3992–3997.

- [11] G. N. Nair and R. J. Evans, "Cooperative networked stabilisability of linear systems with measurement noise," in *Proc. 15th IEEE Mediterranean Conf. Control and Automation*, Athens, Greece, 2007.
- [12] A. S. Matveev and A. V. Savkin, "Multirate stabilization of linear multiple sensor systems via limited capacity communication channels," *SIAM J. Control and Optimiz.*, vol. 44, pp. 584–617, 2006.
- [13] A. S. Matveev and A. V. Savkin, "Decentralized stabilization of linear systems via limited capacity communication networks," in *Proc. 44th IEEE Conf. Decision Control (CDC-ECC'05)*, Seville, Spain, Dec. 2005, pp. 1155–1161.
- [14] V. Gupta, N. C. Martins, and J. S. Baras, "Optimal output feedback control using two remote sensors over erasure channels," *IEEE Trans. Autom. Control*, vol. 54, no. 7, pp. 1463–1476, Jul. 2009.
- [15] S. Yüksel, "Stochastic stabilization of noisy linear systems with fixed-rate limited feedback," *IEEE Trans. Autom. Control*, vol. 55, pp. 2847–2853, Dec. 2010.
- [16] S. Yüksel and T. Başar, "Optimal signaling policies for decentralized multi-controller stabilizability over communication channels," *IEEE Trans. Autom. Control*, vol. 52, no. 10, pp. 1969–1974, Oct. 2007.
- [17] P. Minero, M. Franceschetti, S. Dey, and G. N. Nair, "Data rate theorem for stabilization over time-varying feedback channels," *IEEE Trans. Autom. Control*, vol. 54, no. 2, pp. 243–255, Feb. 2009.
- [18] 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," *IEEE Trans. Inform. Theory*, vol. 52, no. 8, pp. 3369–3395, 2006.
- [19] E. I. Silva, M. S. Derpich, and J. Østergaard, "A framework for control system design subject to average data-rate constraints," *IEEE Trans. Autom. Control*, vol. 56, pp. 1886–1899, Aug. 2011.
- [20] A. S. Matveev and A. V. Savkin, "Decentralized stabilization of networked systems under data-rate constraints," in *Proc. Int. Federation of Automatic Control World Congress*, Seoul, Korea, 2008.
- [21] S. Yüksel and S. P. Meyn, "Random-time, state-dependent stochastic drift for Markov chains and application to stochastic stabilization over erasure channels," *IEEE Trans. Autom. Control*, vol. 58, no. 1, pp. 47–59, Jan. 2013.
- [22] F. Ramponi, D. Chatterjee, A. Milias-Argeitis, P. Hokayem, and J. Lygeros, "Attaining mean square boundedness of a marginally stable stochastic linear system with a bounded control input," *IEEE Trans. Autom. Control*, vol. 55, no. 11, pp. 2414–2418, Nov. 2010.
- [23] C. T. Chen, *Linear Systems Theory and Design*. Oxford, U.K.: Oxford University Press, 1999.

## Variable-Order Fractional Operators for Adaptive Order and Parameter Estimation

Milan R. Rapaić and Alessandro Pisano

**Abstract**—Adaptive parameter estimation schemes for linear fractional-order processes of commensurate type are illustrated in the present note. A gradient-based scheme for commensurate order estimation is outlined first, and a combined gradient/least-squares scheme for simultaneously estimating the commensurate order and the remaining process parameters is finally considered. A key feature of the proposed schemes is the utilization of appropriate variable-order fractional filters. Theoretical results are validated by means of thoroughly discussed simulation examples.

**Index Terms**—Adaptive estimation, fractional calculus, variable-order systems.

### I. INTRODUCTION

Mathematical models utilizing concepts and formalisms of fractional calculus have become increasingly popular in various areas of science and engineering during the last few decades [6], [12]. Fractional control design is an active field of research with several well-established linear and nonlinear methodologies (see, e.g., [6], [12]). Various identification techniques targeting fractional-order models, have recently emerged both in the time [1], [2], [9]–[11], [13] and in the frequency [5], [21] domain. Instrumental variable method has been utilized in [22]–[24], Kalman filtering in [18] and the use of modulating functions was considered in [7]. On-line methods for estimation of the commensurate order are given in [16].

Two distinct, yet closely related, problems are considered within the present work. A gradient-based approach to commensurate order estimation, assuming that all other process parameters are known, is presented first. This approach provides a continuous estimate of the system's commensurate order that eventually, under certain conditions, converges to the actual value. Some preliminary results in this regard were previously reported in [16]. The presented methodology involves the use of variable-order fractional operators [3], [17], [20]. A simultaneous order and parameter estimation procedure is subsequently illustrated, where a gradient-based algorithm is used to identify the commensurate order and the least-squares procedure is used to identify other process parameters. The two algorithms run in parallel in a synergic fashion. Note that a two-stage algorithm for simultaneous estimation of the process order and parameters, conceptually similar to that proposed in this technical note, has already been investigated in [22] and [23]. In comparison to the result presented there, formal conditions ensuring the convergence of the estimation procedure are provided in the present work. These conditions can be seen as generalizations of

Manuscript received July 13, 2012; revised December 23, 2012 and December 23, 2012; accepted July 15, 2013. Date of publication August 15, 2013; date of current version February 19, 2014. The work of M. Rapaić was supported by the Serbian Ministry of Science and Education under Grants 32018 and 33013. The work of A. Pisano was supported by the Italian Ministry of University and Research (MIUR) under PRIN project “Non integer order systems in modeling and control”, Grant 2009F4NZJP. Recommended by Associate Editor E. Weyer.

M. R. Rapaić is with the University of Novi Sad, Novi Sad, Serbia (e-mail: rapaja@uns.ac.rs).

A. Pisano is with the University of Cagliari, Cagliari, Italy (e-mail: pisano@diee.unica.it).

Digital Object Identifier 10.1109/TAC.2013.2278136