Technical Notes and Correspondence

Robust Convergence of Low-Data Rate-Distributed Controllers

Wing Shing Wong and Chi Wan Sung

***Abstract*—This note is focused on control systems in which the observer is not co-located with the controller and the communication channel between them is of limited bandwidth. Due to the low data rate requirement, the control law must be of simple form. As a result, a tristate distributed controller is examined. The control algorithm is proven to be convergent if the system structure satisfies certain technical conditions. The convergence is robust in the sense that the system parameters and even the structure of the underlying system dynamics need not to be completely known.**

***Index Terms*—Communication constaints, distributed controllers, performance-target tracking, robust convergence.**

I. Introduction

Many challenging control problems in networking and communication systems can be cast in the framework of controlling a complex system with a large number of decision-makers distributed over a wide area. Examples that readily come to mind include the power control problem in wireless communication (see, for example, [12], [23], and [31]) and the congestion control problem in the Internet (see, for example, [10] and [17]).

Although the application nature of these problems may vary, they all have three quintessential aspects, which in our view make them also interesting from a theoretical point of view. First of all, the distributed nature of the system implies that observation information is highly distributed and the decision-makers have to determine their control in a distributed manner, based on limited, partial information. Secondly, since the observer, decision-maker, and physical controller are not necessarily co-located, measurement data and control information need to be communicated over a data network. In order not to impose a heavy overhead demand, it is desirable to design algorithms that require low communication data rate. Finally, in many of these systems, system parameters and even the structure of the underlying dynamics may not be known or are at best only partially known. As a result, the controlling algorithms are required to be robust so that errors in knowledge about system parameters or structures do not seriously affect algorithm performance.

In this note, we consider a class of problems that deals with performance-target tracking. Our motivation comes from power control problems in wireless communication. However, these models also occur in other application contexts involving multiple distributed users such as a peer-to-peer network. A class of distributed feedback control algorithms, called tristate algorithms, is proposed. These algorithms are closely related to relay feedback (see, for example, [11] and [15]) and require only one ternary symbol for coding the control decision. Our objective is to ensure performance functions of the system be controlled within a pre-assigned target range. Hence, the objective bears some resemblance to output regulation, in particular robust regulation, (see, for example, [4], [9], [13], [14], and [16]), but without disturbances and with trivial system dynamics.

Manuscript received February 5, 2002; revised October18, 2002 and June 7, 2003. Recommended by Associate Editor F. M. Callier. This work was supported in part by a grant from the Research Grants Council of the Hong Kong Special Administrative Region. (Project CUHK4222/00E).

W. S. Wong is with the Department of Information Engineering, Chinese University of Hong Kong, Shatin, Hong Kong (e-mail: wswong@ie.cuhk.edu.hk).

C. W. Sung is with the Department of Computer Engineering and Information Technology, City University of Hong Kong (e-mail: itcwsung@cityu.edu.hk).

Digital Object Identifier 10.1109/TAC.2003.821418

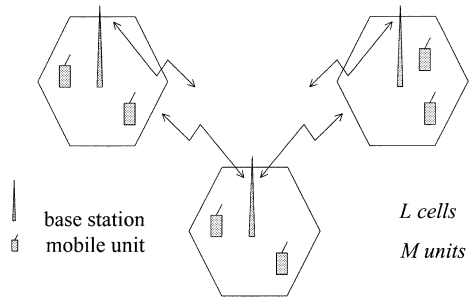

Fig. 1. Schematic of a cellular network.

It should be pointed out that our work differs from these classical results in two major aspects. First of all, in our model the information structure and the control decisions are highly distributed. Each player only sees its individual set of observations and there is no direct coordination among users. Second, although the underlying dynamics of our model is basically trivial, the structure of the performance functions can be quite general and is not completely known. As a result, exact nature of the interaction among users is only known to the individual users through output measurements. This precludes much of the detailed analysis approach used in classical results. Due to the distributed nature of users, we further assume that individual output has to be coded and transmitted over a low speed data channel to its corresponding user for decision-making. This last aspect makes connects between this work and other research work on low data rate controllers, discussed in [2], [5], [28], and [29], as well as with more recent work on the subject such as [1], [3], [6], [7], [19], [21], and [26]. Note that, instead of examining the relationship between data rate and stability or related issues, our goal here is to examine properties of a specific class of feedback control algorithms with low data rate property.

Specifically, in this note we show that a tristate algorithm converges if the output functions satisfying both a Lipschitz condition as well as a dominant-diagonal-like condition. System parameters need not be known to the distributed controllers. In a sense, this class of algorithms is robust by allowing uncertainty in the system structure (see [30] for the issue of uncertainty allowed by feedback). This is a desirable property for many wireless network and Internet control problems, where it is not cost effective or simply infeasible to perform detailed system parameter estimates.

II. Motivation From the Power Control Problem

The performance-target-tracking problem is modeled after the Quality-of-Service (QoS) tracking problem in wireless communication [8]. Consider a cellular network consisting of M mobile units distributed over L cells as depicted in Fig. 1. Each mobile unit intends to communicate with the base station controlling the cell to which it belongs. The communication channels are usually duplex, that is, bidirectional. For simplicity, we will focus on the uplink channel, that is, the communication channel from mobile units to the base stations. Due to propagation characteristics of the electromagnetic wave, the

0018-9286/04$20.00 © 2004 IEEE

signal received for each communication channel is corrupted by interference noises coming from other mobile units.

Let G_{ij} represent the channel gain between the jth transmitter and the base receiver of the cell to which user i belongs. Note that $G_{ij} > 0$. Define the M-by-M channel gain matrix by $\mathbf{G} = (G_{ij})$, (we use the convention that bold face denotes vectors or matrices). Mathematically, we can represent the quality of a channel by the signal-to-noise ratio defined by

$$\Gamma_i(\mathbf{p}) = \frac{G_{ii}p_i}{\sum\limits_{j\neq i} G_{ij}p_j + \eta_i} \tag{2.1}$$

where $\mathbf{p} = (p_1, \ldots, p_M)$ denotes the power vector and $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_M)$ denotes the thermal noise vector; both of them are nonnegative. (For vector $\mathbf{x}$, $\mathbf{x} \geq \mathbf{0}$ if and only if all components of $\mathbf{x}$ are nonnegative).

Notice that in this model, the observer, the base station measuring the signal-to-noise ratio and the controller, the mobile unit, are not co-located. Hence, a communication channel is needed between the observer and the controller.

Given a set of QoS targets, defined by signal-to-noise ratio levels, γ_i, the objective of the target-tracking problem is to adjust the vector $\mathbf{p}$ to satisfy the goal

$$\Gamma_i \geq \gamma_i \tag{2.2}$$

for all i. If the channel gain matrix and the noise vector are completely known, the problem can be easily solved by classical theory of nonnegative matrices (see, for example, [22]). However, this classical result cannot be used to provide a robust and decentralized algorithm to the target-tracking problem because it assumes that a central decision-maker knows the channel gain matrix, computes the solutions and conveys to all the mobile units. In fact, it is almost impossible to know the exact value of the gain channel. Moreover, due to physical phenomena known as fading effects, the gain matrix behaves as a stochastic process, so any parameter identification has to be a continuous process.

In [8], Foschini and Miljanic proposed an algorithm that does not require any information on the value of $\mathbf{G}$. Furthermore, the algorithm is distributed in the sense that in order to update its power level, each mobile unit only needs to communicate with its base station independently. The only information needed is its latest signal-to-noise ratio. This algorithm and its asynchronous version [18] provide a more practical approach for the target-tracking problem in regard to our stated criterions. However, a major shortcoming, due to the requirement of updating based on precise signal-to-noise ratios, is that a high data rate communication channel between the base station and the mobile unit is needed. Inevitably, it is necessary to quantize the observation data. The granularity of quantization levels is closely tied with the communication data rate requirement and fundamentally affects the convergence issue. In [24], Sung and Wong proposed a quantized version of the Foschini–Miljanic algorithm that requires only 1 ternary symbol of data per unit cycle per mobile unit. The *tristate algorithm* is simple to describe

$$\begin{aligned} p_i^{(n+1)} &= \omega(\mathbf{p}^{(n)}, \gamma_i) = \begin{cases} \delta p_i^{(n)} & \Gamma_i^{(n)} < \delta^{-1}\gamma_i \\ \delta^{-1} p_i^{(n)} & \Gamma_i^{(n)} > \delta\gamma_i \\ p_i^{(n)}, & \text{otherwise} \end{cases} \\ p_i^{(0)} &= \text{initial power setting.} \end{aligned} \tag{2.3}$$

Here, $\mathbf{p}^{(n)} = \left(p_1^{(n)}, \ldots, p_N^{(n)}\right)$ and $\delta > 1$ represents the basic multiplicative quantization unit of the power levels. That is, for user i the power level is of the form $p_i^{(0)}\delta^n$, for some integer n. For this problem, the following convergence property was established.

Theorem 1 [24]: Suppose there exists a positive vector $\mathbf{p}$ that satisfies (2.2). Given any positive initial point, algorithm (2.3) converges to a solution, $\hat{\mathbf{p}}$ with the property

$$\delta^{-1}\gamma_i \leq \Gamma_i(\hat{\mathbf{p}}) \leq \delta\gamma_i. \tag{2.4}$$

We call the region between $\delta^{-1}\gamma_i$ and $\delta\gamma_i$ the output convergence zone. One can construct a *bi-state* algorithm (relay feedback) by removing the output convergence zone. In this case, the algorithm will not converge for the target-tracking problem. With the right technical assumptions, it can be shown to oscillate periodically around the target. In fact, this is basically the power control mechanism adopted in the IS95 CDMA standard (see, for example, [20]).

III. Performance-Target-Tracking Model

To fix ideas for subsequent discussions, a performance-target-tracking model is defined in this section. Consider a discrete time system, Σ, described by the equations

$$\begin{cases} x_i^{(n)} = x_i^{(n-1)} + u_i^{(n-1)} & x_i^{(n)} \in \mathbf{R} \\ y_i^{(n)} = h_i(x_1^{(n)}, \ldots, x_M^{(n)}), & y_i^{(n)} \in \mathbf{R} \\ u_i^{(n)} = u_i(y_i^{(n)}) \end{cases} \tag{3.1}$$

where $x_i^{(n)}$, $y_i^{(n)}$, and $u_i^{(n)}$ represent the state, the performance index, and the control of player i at iteration n, respectively; all of them are scalars. Given a performance-target vector, $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_M)$, the objective of the performance-target-tracking problem is to find a feedback controller that guarantees the performance-targets can be achieved and maintained. Note that the power control model is a special case of (3.1) provided that we define $x_i^{(n)} = \log P_i^{(n)}$. This model also can be applied to some peer-to-peer network control problems. For example in [27], the concept of a *fairness index* was proposed. Roughly speaking, the fairness index of a node in a peer-to-peer network is defined by the ratio

$$\Gamma_i = \frac{\sum\limits_{j=1}^{N} M_{ij}p_j}{\sum\limits_{j=1}^{N} L_{ij}p_j} \tag{3.2}$$

where p_i here represents the traffic intensity controlled by node i and (M_{ij}) and (L_{ij}) are matrices defined by the routing and the traffic distribution of the network. This index basically measures how much a node benefits from the network in proportion to the total amount of traffic it handles for the network. It can be used as a performance index for network control.

Definition: The performance target, $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_M)$, is feasible if there exists a state $(x_1^*, \ldots, x_M^*)$ such that

$$y_i = h_i(x_1^*, \ldots, x_M^*) = \gamma_i \tag{3.3}$$

for all i.

Assuming that a performance target is feasible, a natural question is how to ensure the target can be achieved. Due to the distributed nature of the players, we assume the sensor is not co-located with the controller. The picture can be schematically described in Fig. 2.

In the first scheme, the controller is co-located with the dynamical system and it requires a high data rate channel to receive information from the sensor. For systems with low data rate connection between the sensor and the controller, the scheme depicted in Fig. 2(b) is more desirable. In this scheme, control output takes value from a small finite set. Moreover, high-level decisions, such as setting the target levels, is separated from the low-level decisions, such as target-tracking for a given performance target. In this note, our focus is on the latter scheme.

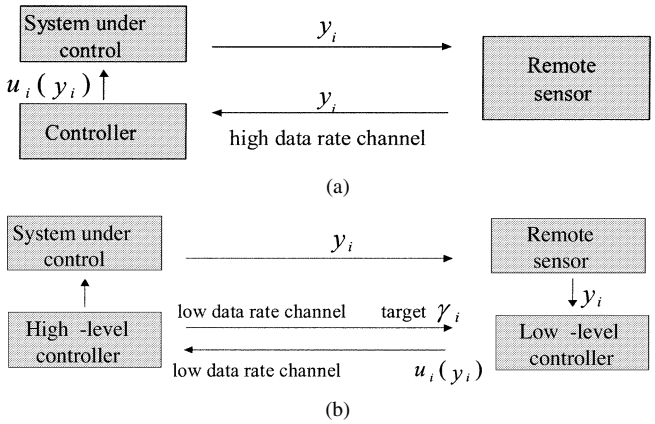

Fig. 2. (a) Feedback control scheme. (b) Alternative feedback control scheme.

For simplicity, all communication channels considered in this note are assumed to be error-free.

For systems using discrete-value controllers, it is unreasonable to expect infinite precision in controlling the performance-targets. Instead, a modified concept of convergence is needed. Let δ and ε be the quantization unit of the controller and the performance-target respectively. Here, the values are quantized linearly. That is, the values of the states $x_i^{(n)}$ are of the form $x_i^{(0)} + j\delta$ for some integer j. These states are called the *quantized states*. Correspondingly, for the performance-target vector, $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_M)$, the performance-index values are quantized into the form $\gamma_i + j\varepsilon$.

Definition: Given performance targets, $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_M)$, and quantization unit, ε, the system Σ is said to be convergent if for any initial state, there exists an N, such that

$$\left| y_i^{(n)} - \gamma_i \right| \leq \varepsilon \tag{3.4}$$

for all i and $n \geq N$.

In this note, the convergence property of a class of tristate algorithms is studied. The algorithm is defined with respect to two quantization units introduced. In particular, given quantization levels δ and ε, we define *the positive tristate feedback control*, ω^+ by

$$u_i^{(n)} = \omega^+\left(y_i^{(n)}, \gamma_i\right) = \begin{cases} \delta & y_i^{(n)} < \gamma_i - \varepsilon \\ -\delta & y_i^{(n)} > \gamma_i + \varepsilon \\ 0 & \text{otherwise.} \end{cases} \tag{3.5}$$

Similarly, one defines the *negative tristate feedback control*, ω^-, by reversing the inequality signs in (3.5). The focus of this note is to investigate the convergent issue of these classes of tristate controllers.

IV. Counter-Example

The convergence result stated in Theorem 1 is robust in the sense that the tristate controller converges for any channel gain matrix for which a feasible solution exists. However, the power control problem has a specific structure, which is essential to the proof. Namely, in the power control problem, whenever a mobile unit raises its power level to improve its performance target the rest of the mobile units suffer a detrimental effect. For problems in which a player may simultaneously improve its own performance-target as well as other players, the convergence property may not hold for the tristate controller.

Consider a system with four players and utilization functions of the form

$$\begin{aligned} y_1 &= x_1 + x_2 - x_3, \quad y_2 = x_1 + x_2 - x_4 \\ y_3 &= x_3 + x_4 - x_1, \quad y_4 = x_3 + x_4 - x_2. \end{aligned} \tag{4.1}$$

In a sense, players 1 and 2 are in one clique with mutual interests and player 3 and 4 in another. Let $\varepsilon = \delta$ and set the targets for all players to 0. This target is obviously feasible. Suppose the initial points are given by

$$x_1^{(0)} = 2\delta \quad x_2^{(0)} = -2\delta, \quad x_3^{(0)} = 2\delta, \quad x_4^{(0)} = -2\delta. \tag{4.2}$$

Since

$$y_1^{(0)} = -2\delta = -2\varepsilon, \quad y_2^{(0)} = 2\varepsilon, \quad y_3^{(0)} = -2\varepsilon, \quad y_4^{(0)} = 2\varepsilon \tag{4.3}$$

it follows that

$$x_1^{(1)} = 3\delta, \quad x_2^{(1)} = -3\delta, \quad x_3^{(1)} = 3\delta, \quad x_4^{(1)} = -3\delta. \tag{4.4}$$

In general,

$$\begin{aligned} &x_1^{(n)} = (2+n)\delta, \quad x_2^{(n)} = -(2+n)\delta, \quad x_3^{(n)} = (2+n)\delta \\ &x_4^{(n)} = -(2+n)\delta. \end{aligned} \tag{4.5}$$

Hence, the algorithm is divergent. This counter-example may seem counter-intuitive. The algorithm fails since players in a clique are uncertain about their control directions due to a lack of direct communication. One may label this as the "too many cooks syndrome."

V. Convergence Results

In this section, the convergence issue of the tristate controller is discussed. In particular, a sufficient condition for convergence is presented. The result depends on two technical conditions on the structure of the system.

Technical Assumption 1: The observation functions, h_i, admit the following decomposition:

$$h_i(x_1, \ldots, x_M) = c_i x_i + d_i(x_1, \ldots, x_M) \tag{5.1}$$

where d_i satisfies a Lipschitz condition

$$\begin{aligned} &|d_i(a_1, \ldots, a_M) - d_i(b_1, \ldots, b_M)| \\ &\qquad \leq L \left\| (a_1 - b_1, \ldots, a_M - b_M) \right\|_\infty \end{aligned} \tag{5.2}$$

for some L, with

$$|c_i| \geq L. \tag{5.3}$$

Note that for linear observation functions, this technical condition translates into a simple dominant diagonal condition.

Technical Assumption 2: The quantization units satisfy the relation

$$\varepsilon \geq \frac{|c_i| + L}{2}\delta \tag{5.4}$$

for all i.

As noted earlier, it is reasonable to expect that in order to control the output with a high accuracy, the controller should have a fine quantization level. In particular, note that according to (3.5) there is an output convergence zone with a width 2ε such that the controller takes a zero value. Suppose the system is in a state with output slightly below this zone and suppose the controller step size, δ, is large enough, one can easily construct an example in which the output of the system oscillates around the convergence zone. As a result, such a system cannot converge. However, if (5.4) holds, this possibility is ruled out.

Suppose the performance targets, $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_M)$, are feasible. That is, there exists a state, $(x_1^*, \ldots, x_M^*)$ such that

$$h_i(x_1^*, \ldots, x_M^*) = \gamma_i, \text{ for } i = 1, 2, \ldots, M. \tag{5.5}$$

In general, $(x_1^*, \ldots, x_M^*)$ may not be a quantized state. However, the following property holds.

Proposition 1: Suppose the targets are feasible. If technical assumptions 1 and 2 hold, then for any initial state and any $\varepsilon > 0$, there

exists a sufficiently small δ, such that there exists a quantized state, $(\hat{x}_1, \ldots, \hat{x}_M)$, satisfying the property

$$|y_i(\hat{x}_1, \ldots, \hat{x}_M) - \gamma_i| \leq \varepsilon. \tag{5.6}$$

Proof: Let the initial state be $\left(x_1^{(0)}, \ldots, x_M^{(0)}\right)$. There exists, $(\hat{x}_1, \ldots, \hat{x}_M)$, of the form $x_i^{(0)} + j\delta$ with the property

$$x_i^* - \frac{\delta}{2} \leq \hat{x}_i < x_i^* + \frac{\delta}{2}. \tag{5.7}$$

Using the decomposition stated in technical assumption 1

$$\begin{aligned} &|y_i(\hat{x}_1, \ldots, \hat{x}_M) - \gamma_i| = |y_i(\hat{x}_1, \ldots, \hat{x}_M) - y_i(x_1^*, \ldots, x_M^*)| \\ &\leq |c_i\hat{x}_i - c_i x_i^*| + |d_i(\hat{x}_1, \ldots, \hat{x}_M) - d_i(x_1^*, \ldots, x_M^*)| \\ &\leq \frac{|c_i|\delta}{2} + \frac{L\delta}{2} \leq \varepsilon. \end{aligned} \tag{5.8}$$

The last inequality follows from technical assumption 2. ■

We call such a quantized state a *quantized solution*. We are now ready to state the main theorem.

Theorem 2: Consider the system, Σ, satisfying technical assumptions 1 and 2. For each player, a tristate controller is assigned according to the rule that if in the decomposition of y_i defined by assumption 1, the term c_i is positive, then let $u_i = \omega^+$, otherwise let $u_i = \omega^-$. Then the resulting system is convergent to any feasible set of targets.

The proof of this theorem is related to the approach in [24] and [25]. The proof can be broken into two major steps. Namely, under the technical assumptions stated, one can show that the algorithm defines a bounded trajectory. Since the feasible states are discrete, this implies that the algorithm either converges to a fixed point or exhibits a periodic trajectory. However, it can be shown that no cycle can exist in the current model. Before proving these results, note that by mapping x_i to $-x_i$ if necessary, there is no loss in generality in assuming that c_i is positive and that $u_i = \omega^+$.

Proposition 2: Under the assumptions stated in Theorem 2, the trajectory of the algorithm is bounded.

Proof: Let $(\hat{x}_1, \ldots, \hat{x}_M)$ be a vector consisting of components of the form $x_i^{(0)} + j\delta$ that satisfies (5.6). Such a vector exists according to Proposition 1. Define $a(i,n)$ by

$$x_i^{(n)} = \hat{x}_i + a(i,n). \tag{5.9}$$

Due to the nature of the tristate algorithm

$$|a(i,n+1) - a(i,n)| \leq \delta. \tag{5.10}$$

Let $K(n) = \max_i |a(i,n)|$. If $K(n) = 0$, then for all i and for all $m \geq n$, $u_i^{(m)} = 0$, the trajectory is clearly bounded. So assume that $K(n) \neq 0$. The proposition clearly holds if $K(n)$ is a nonincreasing function of n whenever $K(n)$ has nonzero value. To show this, assume player i achieves the maximum at time n. There are two possibilities, either

1) $a(i,n) = K(n) > 0$;
2) $a(i,n) = -K(n) < 0$.

Assume that the first condition holds. Then

$$\begin{aligned} &h_i^{(n)}(x_1^{(n)}, \ldots, x_M^{(n)}) \\ &\quad = h_i^{(n)}(\hat{x}_1 + a(1,n), \ldots, \hat{x}_M + a(M,n)) \\ &\quad = c_i(\hat{x}_i + a(i,n)) \\ &\qquad + d_i(\hat{x}_1 + a(1,n), \ldots, \hat{x}_M + a(M,n)) \\ &\quad \geq c_i\hat{x}_i + d_i(\hat{x}_1, \ldots, \hat{x}_M) + c_i K(n) \\ &\qquad - L \max_j |a(j,n)| \\ &\quad \geq \gamma_i - \varepsilon + (c_i - L)K(n) \geq \gamma_i - \varepsilon. \end{aligned} \tag{5.11}$$

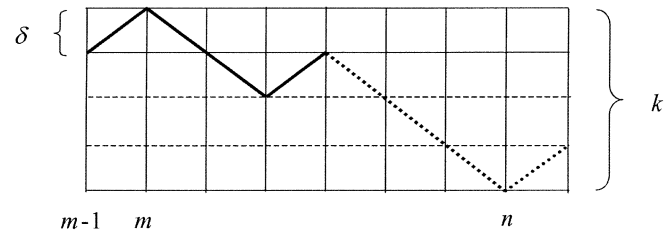

Fig. 3. Example of a peak-slide of length k.

Hence, $u_i^{(n)} \in \{-\delta, 0\}$, that is $u_i^{(n)} \leq 0$. But $a(i,n+1) = a(i,n) + u_i^{(n)} \leq a(i,n)$. Similarly, for the second case

$$\begin{aligned} &h_i^{(n)}(x_1^{(n)}, \ldots, x_M^{(n)}) \\ &\quad = h_i^{(n)}(\hat{x}_1 + a(1,n), \ldots, \hat{x}_M + a(M,n)) \\ &\quad = c_i(\hat{x}_i + a(i,n)) \\ &\qquad + d_i(\hat{x}_1 + a(1,n), \ldots, \hat{x}_M + a(M,n)) \\ &\quad \leq c_i\hat{x}_i + d_i(\hat{x}_1, \ldots, \hat{x}_N) - c_i K(n) \\ &\qquad + L \max_j |a(j,n)| \\ &\quad \leq \gamma_i + \varepsilon - (c_i - L)K(n) \leq \gamma_i + \varepsilon. \end{aligned} \tag{5.12}$$

Hence, $u_i^{(n)} \in \{0, \delta\}$, that is $u_i^{(n)} \geq 0$. But $0 \geq a(i,n+1) = a(i,n) + u_i^{(n)} \geq a(i,n)$. Therefore, for any player i that achieves the maximum value at iteration n

$$|a(i,n+1)| \leq |a(i,n)|. \tag{5.13}$$

For player j where $|a(j,n)| < K(n)$

$$|a(j,n+1)| \leq |a(j,n)| + \delta \leq K(n). \tag{5.14}$$

Hence, $K(n)$ is nonincreasing as a function of n. ■

Proposition 2 shows that for given any initial state, if the performance-target is feasible, the trajectory is bounded. Consider a sequence of vectors, $\mathbf{x}^{(n)}$, $n = 0, 1, \ldots$, the sequence is said to be *asymptotically periodic* if there exists integers, $N > 0$ and $T \geq 1$ such that for all $n \geq N$

$$\mathbf{x}^{(n)} = \mathbf{x}^{(n+T)}. \tag{5.15}$$

Since the transition of the algorithm depends only on the current state and is deterministic, the fact that the trajectory is finite implies that it either converges to a stationary point, or the trajectory is asymptotically periodic with a minimum period larger than 1. For systems satisfying the conditions of Theorem 2, we claim the second case cannot hold.

An arbitrary player, say i, is said to undergo a peak-slide of length $k (k \geq 1)$, from iteration m to n, if there is an integer $n > m$ so that:

$$x_i^{(m-1)} = x_i^{(m)} - \delta \quad x_i^{(m)} = x_i^{(n)} + k\delta. \tag{5.16}$$

Pictorially, a peak-slide is depicted in Fig. 3.

Lemma 1: If there is a peak-slide of length $k (k \geq 1)$, from iteration m to n for player i, then there is a player j, and an integer t, $m \leq t < n$ such that

$$\left|x_j^{(t)} - x_j^{(m-1)}\right| \geq (k+1)\delta. \tag{5.17}$$

Proof: Since $x_i^{(m)} = x_i^{(n)} + k\delta$, and the algorithm changes by at most one δ at a time, there exists an integer t, $m \leq t < n$, such that $x_i^{(t)} = x_i^{(n)} + \delta$ and $y_i^{(t)} > \gamma_i + \varepsilon$. Therefore

$$x_i^{(m-1)} = x_i^{(m)} - \delta = x_i^{(t)} + (k-2)\delta = x_i^{(n)} + (k-1)\delta. \tag{5.18}$$

Note that $y_i^{(m-1)} < \gamma_i - \varepsilon$. Hence

$$\begin{aligned}\gamma_i + \varepsilon < y_i^{(t)} &= c_i x_i^{(t)} + d(\mathbf{x}^{(t)}) \\ &= c_i x_i^{(m-1)} - c_i(k-2)\delta + d(\mathbf{x}^{(m-1)}) \\ &\quad - d(\mathbf{x}^{(m-1)}) + d(\mathbf{x}^{(t)}) \\ &\le y_i^{(m-1)} - c_i(k-2)\delta + L\|\mathbf{x}^{(t)} - \mathbf{x}^{(m-1)}\|_\infty \\ &> \gamma_i - \varepsilon - c_i(k-2)\delta \\ &\quad + L\|\mathbf{x}^{(t)} - \mathbf{x}^{(m-1)}\|_\infty. \end{aligned} \tag{5.19}$$

It follows that

$$2\varepsilon + c_i(k-2)\delta < L\|\mathbf{x}^{(t)} - \mathbf{x}^{(m-1)}\|_\infty. \tag{5.20}$$

According to technical assumption 2, $2\varepsilon \ge (|c_i| + L)\delta$. So

$$\|\mathbf{x}^{(t)} - \mathbf{x}^{(m-1)}|_\infty > \left[1 + \frac{c_i}{L}(k-1)\right]\delta \ge k\delta. \tag{5.21}$$

Hence, there exists an integer j such that

$$\left|x_j^{(t)} - x_j^{(m-1)}\right| > k\delta. \tag{5.22}$$

Since the difference of the two states are multiples of δ, the Proposition holds. ■

Proposition 3: Under the assumption stated in Theorem 2, the trajectory of the algorithm cannot be asymptotically periodic with minimum period larger than 1 if the performance targets are feasible.

Proof: Suppose that the trajectory is asymptotically periodic and let $N > 0$ and $T > 1$ be integers such that for all $n \ge N$, $\mathbf{x}^{(n)} = \mathbf{x}^{(n+T)}$, where T is the minimum period. It follows that there exists a peak-slide starting after N with length larger than or equal to 1. Since the trajectory is bounded, there is a peak-slide with maximum length k. According to Lemma 1, there exists an integer j such that

$$\left|x_j^{(t)} - x_j^{(s)}\right| \ge (k+1)\delta \tag{5.23}$$

where $N \le s \le t$. Since the trajectory is periodic after time N, (5.23) implies that there is a peak-slide with length larger than k, a contradiction. ■

Proof of Theorem 2: Since the trajectory is bounded, it must be asymptotically periodic. According to Proposition 3, the minimum period is equal to 1. That is, the trajectory converges to a stationary state. ■

VI. Conclusion

In this note, we investigated the convergence properties of a class of distributed feedback control algortihms, the tristate feedback control. This class of controls is intended to be used for systems where the sensors and controllers are separated by a low data rate communication channel. The collection of ideas presented here is motivated by wireless communication problem and may find application in other networking systems. Although the tristate algorithm is robust in the sense that the same controller works for systems with different observation functions with trivial dynamics, the convergent rate may be slow. More complicated types of controllers (e.g., [3], [6], [19], [26], and [29]) may provide faster convergence at the price of requiring a higher communication data rate. This and other issues are open questions waiting to be answered.

References

[1] J. Baillieul, "Feedback coding for information-based control: Operating near the data-rate limit," in *Proc. 41st IEEE Conf. Decision Control*, Las Vegas, NV, Dec. 10–13, 2002, pp. 3229–3236.

[2] V. Borkar and S. Mitter, "LQG control with communication constraints," in *Communications, Computation, Control, and Signal Processing: A Tribute to Thomas Kailath*. Norwell, MA: Kluwer, 1997.

[3] R. W. Brockett and D. Liberzon, "Quantized feedback stabilization of linear systems," *IEEE Trans. Automat. Contr.*, vol. 45, pp. 1279–1289, July 2000.

[4] E. J. Davison, "The robust control of a servomechanism problem for linear time-invariant multivariable systems," *IEEE Trans. Automat. Contr.*, vol. AC-21, pp. 25–34, Jan. 1976.

[5] D. F. Delchamps, "Stabilizing a linear system with quantized state feedback," *IEEE Trans. Automat. Contr.*, vol. 35, pp. 916–924, Aug. 1990.

[6] N. Elia and S. K. Mitter, "Stabilization of linear system with limited information," *IEEE Trans. Automat. Contr.*, vol. 46, pp. 1384–1400, Sept. 2001.

[7] F. Fagnani and S. Zampieri, "Stabilizing quantized feedback with minimal information flow: The scalar case," in *Proc. 5th Int. Symp. Mathematical Theory Networks Systems*, Aug. 2002, pp. 12–16.

[8] G. J. Foschini and Z. Miljanic, "A simple distributed autonomous power control algorithm and its convergence," *IEEE Trans. Veh. Technol.*, vol. 42, pp. 641–646, Nov. 1993.

[9] B. A. Francis and W. M. Wonham, "The internal model principle of control theory," *Automatica*, vol. 12, pp. 457–465, 1976.

[10] R. J. Gibbens and F. P. Kelly, "Resource pricing and the evolution of congestion control," *Automatica*, vol. 35, pp. 1969–1985, 1999.

[11] J. M. Goncalves, A. Megretski, and M. A. Dahleh, "Global stability of relay-feedback systems," *IEEE Trans. Veh. Technol.*, vol. 46, pp. 550–562, Sept. 2001.

[12] S. A. Grandhi, R. Vijayan, and D. J. Goodman, "Distributed power control in cellular radio systems," *IEEE Trans. Commun.*, vol. 42, pp. 57–62, Jan. 1994.

[13] J. Huang, "Remarks on the robust output regulation problem for nonlinear systems," *IEEE Trans. Automat. Contr.*, vol. 46, pp. 2028–2031, 2001.

[14] A. Isidori and C. I. Byrnes, "Output regulation of nonlinear systems," *IEEE Trans. Automat. Contr.*, vol. 35, pp. 131–140, Jan. 1990.

[15] K. H. Johansson, A. E. Barabanov, and K. J. Astrom, "Limit cycles with chattering in relay feedback systems," *IEEE Trans. Automat. Contr.*, vol. 47, pp. 1414–1423, Sept. 2002.

[16] H. Khalil, "Robust servomechanism ouptut feedback controllers for feedback linearizable systems," *Automatica*, vol. 30, pp. 1587–1599, 1994.

[17] V. Misra, W. Gong, and D. Towsley, "Fluid-based analysis of a network of AQM routers supporting TCP flows with an application to RED," in *Proc. ACM/SIGCOMM*, 2000.

[18] D. Mitra, "An asynchronous distributed algorithm for power control in cellular radio systems," in *Proc. 4th WINLAB Workshop Third Generation Wireless Inform. Net.*, 1993, pp. 249–259.

[19] G. N. Nair and R. J. Evans, "Stabilization with data-rate-limited feedback: Tightest attainable bounds," *Syst. Control Lett.*, vol. 41, no. 1, pp. 49–56, 2000.

[20] T. S. Rappaport, *Wireless Communications: Principles and Practice*. Upper Saddle River, NJ: Prentice-Hall, 1996.

[21] A. V. Savkin and I. R. Petersen, "Set-valued state estimation via a limited capacity communication channel," *IEEE Trans. Automat. Contr.*, vol. 48, pp. 676–680, Apr. 2003.

[22] E. Seneta, *Non-Negative Matrices and Markov Chain*, 2nd ed. New York: Springer-Verlag, 1981.

[23] C. W. Sung and W. S. Wong, "Mathematical aspects of the power control problem in mobile communication systems," in *Lectures on Systems, Control, and Information: Lectures at the Morningside Center of Mathematics*, L. Guo and S. S.-T. Yau, Eds. Providence, RI: ACM/International Press, 2000.

[24] ——, "A distributed fixed-step power control algorithm with quantization and active link quality protection," *IEEE Trans. Veh. Technol.*, vol. 48, pp. 553–562, Mar. 1999.

[25] C. W. Sung, K. K. Leung, and W. S. Wong, "A quality-based fixed-step power control algorithm with adaptive target threshold," *IEEE Trans. Veh. Technol.*, vol. 49, pp. 1430–1439, Apr. 2000.

[26] S. Tatikonda and S. K. Mitter, "Control under communication constraints," in *Proc. IEEE Conf. Decision Control*, Las Vegas, NV, Dec. 2002, pp. 10–13.

[27] W. S. Wong, "Fairness issues in peer-to-peer networks," in *Proc. IEEE 2002 Conf. Decision Control*, Las Vegas, NV, Dec. 2002, pp. 2005–2010.

[28] W. S. Wong and R. W. Brockett, "Systems with finite communication bandwidth constraints I: State estimation problems," *IEEE Trans. Automat. Contr.*, vol. 42, pp. 1294–1298, Sept. 1997.

[29] ——, "Systems with finite communication bandwidth constraints II: Stabilization with limited information feedback," *IEEE Trans. Automat. Contr.*, vol. 44, pp. 1049–1052, May 1999.

[30] L. L. Xie and L. Guo, "How much uncertainty can be dealt with by feedback?," *IEEE Trans. Automat. Contr.*, vol. 45, pp. 2203–2217, Dec. 2000.

[31] J. Zander, "Distributed cochannel interference control in cellular radio systems," *IEEE Trans. Veh. Technol.*, vol. 41, pp. 305–311, Aug. 1992.

Use of the Kalman Filter for Inference in State-Space Models With Unknown Noise Distributions

John L. Maryak, James C. Spall, and Bryan D. Heydon

***Abstract*—The Kalman filter is frequently used for state estimation in state-space models when the standard Gaussian noise assumption does not apply. A problem arises, however, in that inference based on the incorrect Gaussian assumption can lead to misleading or erroneous conclusions about the relationship of the Kalman filter estimate to the true (unknown) state. This note shows how inequalities from probability theory associated with the probabilities of convex sets have potential for characterizing the estimation error of a Kalman filter in such a non-Gaussian (distribution-free) setting.**

***Index Terms*—Estimation theory, multivariate, non-Gaussian processes, state-space models, uncertainty.**

I. Introduction

The linear state-space model with random noises is widely used to model dynamic systems. In many practical settings, the noise terms will have unknown probability distributions, which contradicts the standard assumption of normality made in state-vector estimation. This assumption is usually made for reasons of mathematical tractability in algorithm derivation and uncertainty characterization; it is rarely motivated by the nature of the actual underlying random processes in the system under consideration. It is usual to make inferences about the system based on computed estimation uncertainties. Naturally, if the Gaussian assumption is made when the noise terms have unknown distributions, there is significant danger of drawing invalid inferences about the system.

The aim of this note is to present a computationally feasible and relatively easy-to-implement approach by which the system measurements can be used to make valid inference about the (unknown) state value in the non-Gaussian (distribution-free) setting. It will be shown that the Kalman filter can be used for this purpose. The Kalman filter is well known for state estimation; and widely used, efficient and numerically stable implementation forms exist for systems with many state variables. The essence of the approach is to show how uncertainty bounds (confidence regions) can be attached to the Kalman filter state estimate that properly account for the distribution-free setting.

The state-space model considered here has the standard discrete-time form

$$x_k = F_{k-1}x_{k-1} + w_k \text{ (State equation)} \tag{1a}$$

$$y_k = H_k x_k + v_k, \text{ (Measurement equation)} \tag{1b}$$

Manuscript received January 29, 2000; revised June 24, 2003. Recommended by Associate Editor R. Srikant. This work was supported in part by the JHU/APL Independent Research and Development Program and the U.S. Navy under Contract N00024-98-D-8124.

The authors are with the Applied Physics Laboratory, Johns Hopkins University, Laurel, MD 20723-6099 USA (e-mail: john.maryak@jhuapl.edu).

Digital Object Identifier 10.1109/TAC.2003.821415

$k = 1, \ldots, n$ where x_0, v_k, and w_k have (in general) unknown (non-Gaussian) distributions with known finite second moments, i.e., $v_k \sim (0, R_k)$, $w_k \sim (0, Q_k)$ (without loss of generality, we assume $x_0 = 0$), and F_k and H_k represent the state transition matrix and measurement matrix, which may, in general, be time varying. We assume that $x_0, v_1, w_1, v_2, w_2, \ldots$ are mutually independent. As with many state-estimation methodologies (e.g., [6] and [10]), we assume that the model parameters (F_k, H_k, etc.) are known.

The problem of state estimation in non-Gaussian settings has, of course, been extensively studied (e.g., see literature reviews and discussions by [5], [13], [17], [27], and other references listed later in this paragraph). The approach here is fundamentally different from existing methods in that we *do not require a detailed knowledge of the* w_k *and* v_k *distributions* (hence the term "distribution-free"). This approach recognizes the practical difficulties in specifying the exact form of the non-Gaussian distributions as required by other existing approaches, such as non-Gaussian filtering and Monte Carlo approaches; see, e.g., [12], [24], or [25]. The most popular Monte Carlo technique—Markov chain Monte Carlo, as in [9] and [23]—is an example of a non-Gaussian method requiring full knowledge of the distributions being sampled. Other non-Gaussian sampling approaches, such as [4], [6], and [10], also rely either on known noise distributions or on distributions that are conditionally Gaussian with the conditioning parameters being generated within the algorithm.

The approach to inference about the state vector presented here, based on attaching uncertainty bounds to the Kalman filter estimate, is motivated by the fact that the Kalman filter has been, and will continue to be, used in known non-Gaussian settings. The reasons for this include wide availability of software, computational efficiency and numerical stability of particular implementations, and general industry acceptance and familiarity. The complexity of implementing software for a full nonlinear filter is generally *much* higher than that required for a Kalman filter implementation.

There are several other approaches for quantifying the uncertainty in a Kalman filter estimate in general *distribution-free* settings. The first is to use a version of the Chebyshev inequality, which is very simple as it is based on a straightforward application of the error covariance matrix generated as part of the filter. This method suffers, however, in producing uncertainty bounds that are too large (i.e., too conservative) for many practical applications. A second approach is based on the Kantorovich inequality from linear algebra and probability theory. This idea, introduced in [22], has promise of producing tighter bounds than the Chebyshev inequality, but is currently only developed for the scalar state setting. A third approach, which applies to linear and nonlinear estimation problems, is presented in [14] and [21]; this method is based on asymptotic theory that requires that the noise terms become asymptotically Gaussian.

Section II and III describe the approach, which differs from all of the non-Gaussian and distribution-free approaches shown before. It is based on computing a bound for the probability of the Kalman filter estimation error exceeding certain threshold values in a manner analogous to the Chebyshev inequality. However, the assumptions being introduced on the general class in which the noise distributions lie are aimed at producing more informative (tighter) uncertainty bounds than those resulting from the Chebyshev inequality. Section IV demonstrates these ideas in an estimation problem taken from the engineering literature.

II. Establishing Estimation Uncertainty Bounds

The objective here is to compute uncertainty bounds for the state estimation error, $\tilde{x}_n = \hat{x}_n - x_n$, obtained from the Kalman filter [applied

0018-9286/04$20.00 © 2004 IEEE