

## THE APPROPRIATE QUANTIZATION FOR DIGITAL FINITE COMMUNICATION BANDWIDTH (DFCB) CONTROL

K. LI<sup>1</sup>     J. BAILLIEUL<sup>1</sup>

ABSTRACT. The problem of DFCB control has come to the attention of the research community in connection with a growing interest in the development of distributed and/or networked control systems. In these systems, actuators, sensors and other components are connected via data-rate constrained links such as wireless radio, etc.

In this document, we consider a model of DFCB control that accomodates time-varying data rate constraint, which might be the case with intermittent network congestion, and asynchronism of sampling and control actuation. Because of the possibly unpredictable fluctuation of data rate, we are interested in feedback control designs that will tolerate significantly constrained data-rates on feedback loops, while providing acceptable performance when such data rate constraints are not in force. In light of a very basic notion of acceptable performance, we found that control designs with different number of quantization levels tolerate constrained data rates differently. This leads to the conclusion that binary control represents the appropriate control quantization under data rate constraint. The advantage margin of binary control is further investigated numerically with and without the sampling/control asynchronism consideration. We end the document by showing that the advantage margin is more substantial with the asynchronism consideration.

### 1. INTRODUCTION

The DFCB control problem is a generalization of those situations where information is exchanged between the decision maker (the controller) and the controlled dynamic object (the plant) through a channel that can transmit only a finite number of states in any given period of time. In this problem, the observations and reactions made by the decision maker are quantized in both time and space with finite precision. Being digital means the uncertainty in this information has a very special structure, which results from the quantization in the coding process. This distinguishes it from the classical discrete-time-system-with-observation-noise problem.

An essential part of the present problem is the tradeoff between the resolution of temporal quantization and spatial quantization. This time-space tradeoff has been studied by a number of researchers in recent years. Wong and Brockett [9] considered the issues of coding, communication protocol and delays explicitly. A

---

<sup>1</sup>Dept. of Aerospace and Mechanical Engineering, Boston University, Boston, MA 02215. E-mail: johnb@bu.edu Both authors gratefully acknowledge support from the Army Research Office under the ODDR&E MURI97 Program Grant No. DAAG55-97-1-0114 to the Center for Dynamics and Control of Smart Structures (through Harvard University) and from ODDR&E MURI01 Program Grant Number DAAD19-01-1-0465 to Boston University and the Center for Networked Communicating Control Systems.

model with a memoryless and time-invariant coder-controller, stop-and-wait communication scheme and impulsive control actuation was considered. The notion of *containability* was introduced as the appropriate notion of stability. Both necessary and sufficient conditions for containability are given in the form of bounds on the system data rate. Soon after this work was published, Baillieul [8] reported a tight bound on the data rate requirement in order to stabilize a real right-half-plane pole. The article showed that even for the coarsest quantization with only two admissible control values, which is referred to as binary control, the bound is achievable. At about the same time Tatikonda *et al.* presented results on the optimal LQG control of LTI systems with noisy feedback channels. Incorporating a “cost of information” with the usual cost of control, it was shown in the scalar case that the average optimal cost would be finite only if the data-rate inequality was strictly satisfied. In [7], Nair and Evans discussed a coder with infinite memory. Uncertainty is assumed to originate from initial state only and grow because of the unstable dynamics of the plant. The optimal coder is claimed to be a causally reformulated optimal quantizer for the initial state. In [6], Tatikonda and Mitter classified the encoder of the feedback in terms of their knowledge of the control actuation as well as their memory. The encoders are time-varying and operate on a finite region in the state space for each sample. The data rate requirement is determined by insuring the coding resolution for each new sample is no worse than the previous. Asymptotic stability is established consequently. All these recent results considered control of inherently unstable plants through data rate constrained communication. The core reasoning is: As the uncertainty of plant state evolves through the unstable plant dynamics, enough amount of information must be communicated between plant and controller to keep this uncertainty bounded. In particular, the common result in [3], [6], [8], [10] is that the minimum data rate needed to stably close a feedback loop for a system with open-loop pole  $\lambda > 0$  is  $R > \lambda \log e$ , where the base of logarithm depends on the communication scheme in various ways.

Given the recent progress, there is a lack of design tools for DFCB control which devote consideration to practical communication issues explicitly. Here, we focus on an essential design question: How fine should the control quantization be? We will answer this question in light of some very basic practical considerations. The answer to this question turns out to simplify the design task significantly. Previous theoretical results tell us that the data rate requirement for stabilizing an unstable plant does not rely on the number of quantization levels. However, we will show that when some practical considerations come into play, very coarse quantization such as binary control might be needed to operate near the theoretical minimum data rate.

In what follows, we first introduce a fairly general model for memoryless time-invariant DFCB control. Similar to previous literature, we assume the communication bandwidth can be characterized as some kind of data rate, i.e., the ratio of a certain delay to the data length of control is constant. However, our definition of data rate is slightly different because our model allows non-uniform communication delays and asynchronous sampling and actuation arrangement. By introducing this model and corresponding concepts, we pay attention to the control design scenario where the available communication bandwidth fluctuates in an unpredictable fashion except that a tight lower bound of the data rate is known. Consequently, we are interested in control designs that work reasonably well under both high data rate and low data rate. With high data rate, it should drive the system state to

the origin while with low data rate it should guarantee the state trajectory remain bounded. Under this thinking, we introduce a very basic notion for acceptable performance of a dynamic system controlled by time-invariant quantized control. This notion requires that all trajectories specified by the control design must converge to the origin. This notion is important regardless of data rate constraint. Thus, our scope of search is reduced to the control designs that are acceptable in this sense. Next, we compare the data rates with which different “acceptable” control designs can guarantee bounded state trajectory given certain initial conditions. Possible asynchronism of sampling and control actuation will also be taken into account. Our main result states that: Among all “acceptable” control designs, binary control, and binary control only, can operate on the lowest data rate possible. And this data rate is in the form of the theoretically minimum data rate that previous literature gives. This statement is based on the assumption that data rate does not depend on message length. We end the paper by showing numerically that even when the data rate does depend on message length, it is still likely that binary control can operate on lower data rates. This conclusion obviously simplifies the control design because binary control is easy to design and analyze.

## 2. THE PROBLEM

Here we use a hybrid model in order to capture some practical issues. The controlled object is a continuous-time dynamic plant while sampling and control happen at discrete instants. The discrete-time model with uniform sampling interval will appear as a simplified version of our model and will be used in some of our discussions.

Consider the plant

$$(1) \quad \frac{d}{dt}x(t) = x(t) - u(t).$$

where the state  $x \in \Re$ . And  $x = 0$  is the desired value. Following the literature [3], [6] and [9], we are considering an inherently unstable plant. For simplicity, we present our argument with this non-dimensional scalar case. The results can be extended to higher dimensionalities.

For the control scheme, we make following assumptions:

1. The feedback control law is time-invariant and memoryless, and no estimation is made. I. e., the next control value depends explicitly on only the latest state feedback.
2. Control actuation is of the *update-hold* type. I. e., the present control value is applied on the plant and held until next control value is applied. An alternative is the *impulse control*, where the control is in effect only for an arbitrarily short duration [9]. In certain contexts, these two types of control actuation lead to the same theoretical results. But this is beyond the scope of this paper.
3. The initial control value is consistent with the plant state as designed, i.e., the controller is prepared [9].

The general features of digital finite communication bandwidth control together with the assumptions above result in the following constraints:

**Sampling:** The state of the plant is sampled only at discrete instants  $t_j, t_{j+1} \geq t_j, j = 0, 1, 2, \dots$

FIGURE 1. Sampling and control-updating instants

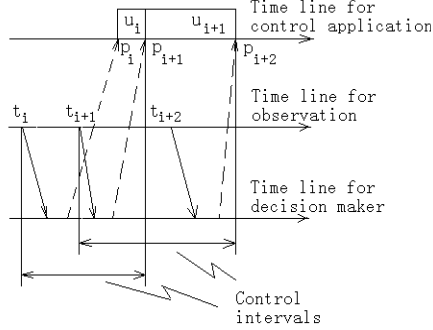
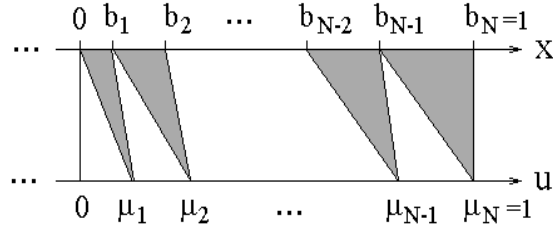


FIGURE 2. The Selection Function



**Delay:** Control value is updated only at instants  $p_j, p_{j+1} \geq p_j \geq t_j, j = 0, 1, 2, \dots, p_0 = t_0$ .  $p_j - t_j$  is the delay. Actually, a more specific name for this delay might be “*sampling-actuation-asynchronism*”. The relation between  $\{t_j\}$  and  $\{p_j\}$  is illustrated in figure 1. The quantity *control interval* in the figure will be discussed a little later.

**Digital:** The value of  $u(t)$  are restricted to a finite set  $U$ , called the *admissible control value set*. Thus the feedback control law becomes a *selection function* that maps the state space to  $U$ .

Thus,  $u(t)$  is piecewise constant and can be represented by a sequence  $u_j, j = 0, 1, 2, \dots$

$$u(t) = u_j = f(x(t_j)), t \in (p_j, p_{j+1}).$$

The design task is to choose a finite set  $U$  of control values and corresponding selection function  $f$ . We also call  $f$  a *control design* and consider  $U$  as a part of  $f$ . Without loss of generality, further assume that the selection function  $f$  is symmetric except for the origin and let

$$(2) \quad \begin{cases} U = \{-\mu_l, \mu_l\}_{l=1}^N \quad \text{where} \quad 1 = \mu_N > \mu_{N-1} > \dots > \mu_1 > 0 \\ f(x) = \mu_l, x \in [b_{l-1}, b_l], l = 1, \dots, N, \quad 1 = b_N > b_{N-1} > \dots > b_0 = 0 \\ f(-x) = -f(x), x \neq 0. \end{cases}$$

It is illustrated in figure 2.

**Definition 1.** Call  $(b_{l-1}, b_l)$  and  $(-b_l, -b_{l-1}), l = 1, \dots, N$  each a *control region* and  $(-b_N, b_N)$  the *working region* of the system.

Note that the values  $x = \pm\mu_N$  are unstable rest points for the system (1) given any design of  $U$  and  $f$ . It is clear that a state trajectory can not be bounded using any of the admissible control values once it has left the *working region*.

We believe our sampling and control arrangement uses the simplest possible set of assumptions which captures the essential features of the way in which communication delays affect control system performance. In particular, our assumptions allows non-uniform sampling intervals/delays and asynchronism of sampling and control actuation. Concerns in these two aspects are important motivation for the study presented below. It is also worth mentioning that the constraint of  $p_{j+1} \geq p_j$  preserves the possibility of lost packets while avoiding unnecessary complexity. To describe a lost packet, one might let  $p_j = p_{j+1}$  so that the duration of control value  $u_j$  is zero, which effectively means either  $x(t_j)$  or  $u_j$  is lost.

So far we have formulated a DFCB control system that can be built on a wide range of communication schemes. However, for the control system to work properly, the supporting communication must meet certain criteria. Now we move our attention to the characterization of the communication requirement. Under the preceding settings, we find the following quantity of particular interest.

**Definition 2.** *For the above DFCB control system, if  $\Theta = \sup_{j=1 \rightarrow \infty} \{p_j - t_{j-1}\}$  exists, call it the **control interval** of the system.*

The control interval is the upper bound of the time it takes for the system to respond to an event. The response to an event concerns two delays: 1. the possible delay of detection,  $t_j - t_{j-1}$  and 2. the possible delay of decision making and actuation,  $p_j - t_j$ . So, the “control interval” captures the sum of these two delays. Consider the following situation: the plant state enters a new control region immediately after being sampled at  $t_{j-1}$ . This event will be first detected, if ever, at  $t_j$ . The corresponding action (changing the control value) cannot be applied until  $p_j$ . Thus, the time difference between the occurrence of the event and the application of control action is captured by the control interval. Neither the sampling interval nor the delay is enough by itself. In the above sense, control interval is the proper measure of the length of control cycle for our model.

Noticing the significance of control interval in a DFCB control system, we will choose a characterization of the supporting communication that is directly linked to control interval. And since we allow fluctuation of communication bandwidth, our characterization is essentially a bound.

Assume the largest possible delay for processing (generating, transmitting and receiving, etc.) each one bit data between the plant and controller (in either direction) is  $\delta$ . Following literature [3], [6] and [9], etc., we choose the characterization to be some kind of data rate. Here we define this quantity to be  $R_b = 1/\delta$  and call it the **base data rate** to distinguish it from the data rate for the uniform-sampling-interval-and-delay cases. The **base data rate** defined as such still bears the spirit of data rate in previous literature that the length of control cycle depends linearly on the message length in the feedback loop. Only the control cycle here is composed of more components. Like some important quantities used to characterize telecommunication channels (e.g., *throughput*), the **base data rate** defined here corresponds to the lumped effect of both the hardware capacity (e.g., speed of transceivers) and software specification (protocols, e.g., the choice of sampling instances in our case) of real world systems. Given the data rate, if the length of each message is uniformly  $L$  bits (depending on the chosen control quantization)

then a control cycle (a round trip — plant side generating a sample message and feeding it back to the controller; controller generating a control decision and applying it to the plant) can take as long as  $2L/R_b$ , but not longer. I.e., the control interval  $\Theta = 2L/R_b$ . Further assume that each sample or control message takes the minimum length that can support the selection function  $f$ . I.e., the length of each message is  $L = \lceil \log_2 N + 1 \rceil$  bits. ( $f$  is symmetric with  $2N$  levels, see (2).)

Now, we have a DFCB control system with a selection function  $f$ , whose design will be refined, and a sampling/actuation scheme. The communication constraint is characterized by  $R_b$ .  $R_b$  and the message length  $L$  together determines the control interval of the system. Given a design  $f$ , it may fail to deliver the expected performance if  $R_b$  is not high enough. Conversely, given a data rate  $R_b$ , only certain  $f$ 's, if any, can control the plant satisfactorily. So, there is a relation between  $f$  and  $R_b$  with respect to some performance criterion. Here we shall be concerned with the most basic stability requirement — boundedness of closed-loop trajectories.

**Definition 3.** For a given control design  $f$ , and control interval  $\Theta$ , suppose there is a **bounded response** in the sense that for some compact set  $S$ , any trajectory of the closed-loop system (1) that has initial state in  $S$  remains in  $S$  for all time. Define the **required base data rate** of the design  $f$  to be the greatest lower bound of all rates  $R_b$  for which a bounded response using this design is assured.

To close this section, we call the DFCB system described above **DFCBO** and summarize it as follows: (Notice that it is in a non-dimensionalized form.)

$$(3) \quad \begin{cases} \frac{d}{dt}x(t) = x(t) - u(t), & x \in \mathfrak{X}, u \in U \\ U = \{-\mu_l, \mu_l\}_{l=1}^N \text{ where } 1 = \mu_N > \mu_{N-1} > \dots > \mu_1 > 0 \\ t_{j+1} \geq t_j, & p_{j+1} \geq p_j \geq t_j, \quad j = 0, 1, \dots \\ p_0 = t_0 = 0 \\ u(t) = u_j, & t \in [p_j, p_{j+1}) \\ u_j = f(x(t_j)) \end{cases}$$

where

- $t_j$  : instant of sampling
- $p_j$  : instant of updating control value
- $f$  : control law in the form of a selection function  $f : \mathfrak{X} \rightarrow U$
- $U$  : set of admissible control values.

The communication between the plant and controller has following properties:

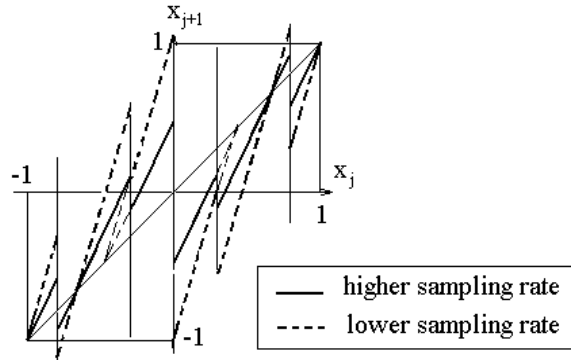
$$(4) \quad \begin{cases} L = \lceil \log_2 N + 1 \rceil, \\ R_b = 1/\delta, \\ \Theta = 2L/R_b, \\ \sup_{j=0 \rightarrow \infty} p_{j+1} - t_j = \Theta \end{cases}$$

in which

- $L$  : length of each sample or control message
- $\delta$  : largest possible delay for processing one single bit of data from the plant side to the controller or in the opposite direction.
- $R_b$  : base data rate
- $\Theta$  : control interval, the worst-scenario delay of reaction given  $L$  and  $R_b$ .

In addition, we defined *control region* and *working region* to facilitate the discussions about the selection function  $f$ . Finally, we defined the *required data rate* to represent the communication requirement of control designs.

FIGURE 3. Transition Graph of DFCB1



### 3. A SIMPLIFIED VERSION — THE DISCRETE TIME MODEL

With the additional assumptions that  $p_j = t_j$  (sampling and control actuation are synchronous) and  $t_{j+1} - t_j \equiv h$  (uniform *sampling rate/interval*),  $j = 0, \dots, \infty$ , DFCB0 can take the form of the commonly used discrete time model

$$(5) \quad x_{j+1} = \alpha x_j - \beta u_j$$

where  $x_j = x(t_j)$ ,  $u_j = f(x_j)$ ,  $\alpha = e^h$ ,  $\beta = e^h - 1$ . The *sampling interval*  $h$  takes the place of control interval and  $R = \frac{2L}{h}$  is called the data rate. Also define the required data rate of control designs following the definition of required base data rate (Definition 3). In this paper, we call this simplified version **DFCB1**. All of our major results apply to DFCB0. But for simplicity, we present some of the discussion with DFCB1. It is quite convenient to describe the behavior of DFCB1 graphically. Figure 3 illustrates the state transition of two DFCB1 systems with the same controller (and plant) but different sampling rate.

**Remarks 1.** • For DFCB1 with  $2N$  admissible control levels (symmetric about the origin), the close-loop transition function consists of  $2N$  parallel line segments, one for each control region.

- For the control region  $[b_{l-1}, b_l]$  (see (2)), whose assigned control value is  $\mu_l$ , its line segment (or the extension) always pass through the point  $(\mu_l, \mu_l)$  regardless of  $h$ .
- The slope of the line segments increase with  $h$ .

### 4. REGULARITY OF CONTROL DESIGNS

Starting from this section, we try to find the “appropriate” DFCB control design, assuming that the communication constraint can be characterized by (base) data rate. It is well known that for quantized discrete-time linear control systems, asymptotically stable response can only be achieved in an approximate sense. The degree to which it can be achieved will depend in large part on communication data-rate constraints that limit the sampling frequency needed to close the feedback loop. Recent work has provided a detailed picture of how closed-loop performance is degraded when severe data-rate constraints are encountered. (See e.g. [2] and [4]) In the present paper, we study the setting in which data-rate constraints are time-varying (e.g., as would be the case with intermittent network congestion.)

We are interested in feedback control designs which will tolerate significantly constrained data rates in feedback loops, while provide acceptable performance when such data-rate constraints are not in force.

Our main argument suggests that binary control is the appropriate control quantization. The comparison of binary control with all other control designs is presented in four steps:

1. Eliminate the control designs that could result in degraded performance even when the communication constraint does not exist. (The degraded performance mentioned here will be explained below.)
2. For the rest control design candidates, group them by number of quantization levels. Find the minimum *required data rate* within each group.
3. Compare the minimum *required data rate* of the groups in step 2. This will show that binary control require the strictly lowest minimum data rate, which is equal to the data rate limit agreed by previous literature [3], [6], [8], [9], etc.
4. Use numerical analysis to show that there can be substantial margins between the data rate required by binary control and those of others.

For steps 1, 2 and the first part of 3 and 4, we present the discussion in terms of DFCB1 (Section 3) because these discussions consider high- and low-data-rate scenarios separately; and they do not consider sampling/actuation asynchronism. For the second part of step 3 and 4, however, the discussions will address the particular issues in DFCB0. This section is the first step of the four.

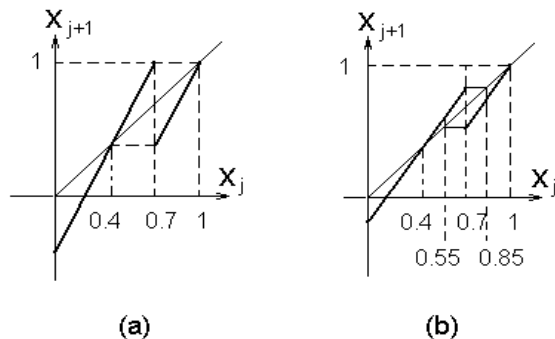
In previous literature, the data rate requirement of DFCB control system using time-invariant control with finite number of admissible control values is derived by requiring that the system have a bounded response, e.g., [3]. In this setting the data rate limit is independent of the number of control quantization levels. However, we noticed that the data rate limit is sometimes achieved by control designs that are not acceptable as regulators in the classical sense (without data rate constraint) of diminishing residual error. So, we eliminate these control designs as possible candidates and reexamine the data rate requirements. In our case the degraded performance of these unacceptable designs of the selection function  $f$  can be illustrated in the following example.

**Example 1.** *Suppose the set  $U$  of control values has four elements, and the selection function is given by*

$$f = \begin{cases} -1 & : x \leq -0.7 \\ -0.4 & : -0.7 < x < 0 \\ -0.4 & : 0 \leq x < 0.7 \\ 1 & : x \geq 0.7. \end{cases}$$

*Since the discrete-time model DFCB1 is considered, we plot the close-loop state transition in figure 4. In figure 4, (a), where the sampling interval is  $h = \ln 2$ , one can see that the region  $[0.4, 1]$  is invariant. (Both the regions  $[0.4, 0.7]$  and  $[0.7, 1]$  maps to  $[0.4, 1]$  after one transition.) The state trajectory will present spurious limit cycle behavior once it enters this region. Notice that this is not a desired invariant set for the controlled trajectory.*

FIGURE 4. Unacceptable Control Design



In figure 4, (b), where  $h$  is reduced to  $\ln 1.5$ , the region  $[0.55, 0.85]$  is invariant. Although the state trajectory converges better locally after  $h$  is reduced, this invariant set is still away from the origin. Actually, the state trajectory will always be trapped around the point 0.7 no matter how small  $h$  is.

This example shows a type of control design that needs to be avoided. We next discuss the characteristics of quantized control designs which do not exhibit these pathologies as  $h \rightarrow 0$ .

**Definition 4.** A control design (selection function)  $f$  of system (3) is **regular** if for any given neighborhood  $\mathcal{N}$  of the origin and any subregion  $\Sigma$  of the working region with nonzero measure, there exist positive numbers  $\mathcal{R}$  and  $T$  such that if the system data rate is greater than  $\mathcal{R}$  and  $x(0) \in \Sigma$ , then  $x(t) \in \mathcal{N}$ , for all  $t > T$ .

The definition of *regularity* here is constructed specifically to exclude the spurious limit cycle behavior around an undesirable state as illustrated in Example 1. Regular control designs include a wide range of selection functions (e.g., evenly or logarithmically [1] quantized linear control with sufficiently high gain). The close-loop state transition corresponding to a possible regular control design is illustrated in figure 5. One might notice in this figure that the line segments representing transition mapping in the right-half-plane are all below the  $x_{j+1} = x_j$  line (conversely with the left-half-plane). This is the key difference between regular and unregular selection functions (also see figure 4). More precisely, the lemma below points out the parametric characteristic of regular selection functions.

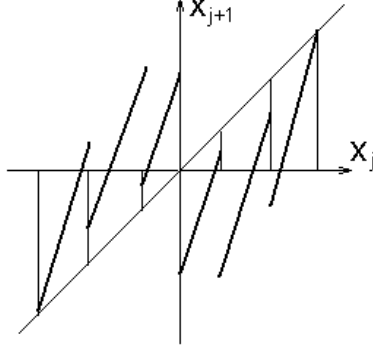
**Lemma 1.** For the system (3), a control design  $f$  is regular if and only if  $b_l \leq \mu_l$ . (Notice that  $f$  is symmetric and  $b_l, \mu_l (l = 1, \dots, N)$  are positive numbers that specify  $f$  on the right half axis.)

*Proof.* For simplicity, we present the proof in terms of DFCB1. Similar proof can be constructed for DFCB0.

For necessity, if  $b_l > \mu_l$  and  $b_{l+1} < \mu_{l+1}$  for a certain  $l \neq 0$ , then when we increase data rate (equivalently, reduce  $h$ ) in the hope of diminishing residual error, the state trajectory can actually be trapped in an arbitrarily small neighborhood of  $b_j$  instead of a small neighborhood of the origin as illustrated in Example 1 ( $b_l > \mu_l$  for all  $l$ 's is impossible because  $b_N < \mu_N$ .)

Next we prove sufficiency. From equation (5), the state transition of DFCB1 is  $x_{j+1} = e^h x_j - (e^h - 1)f(x_j) = e^h x_j - (e^h - 1)\mu_l$  for  $x_j \in [b_{l-1}, b_l), l = 1, \dots, N$

FIGURE 5. Regular Control Design



(We only need to consider  $x_j \geq 0$  because of symmetry). So,

$$(6) \quad x_{j+1}^2 = e^{2h} x_j^2 - 2e^h (e^h - 1) \mu_l x_j + (e^h - 1)^2 \mu_l^2.$$

$\forall \xi > 0, x_j^2 > \xi, \exists H_1 > 0$  such that for  $0 < h < H_1$ , if the  $\mu_l$  in (6) is replaced by another value, the change of the second term dominates that of the third because of the factor  $(e^h - 1)$  being small. Then, the condition that  $\mu_l \geq b_l > x_j$  lead to

$$(7) \quad x_{j+1}^2 < e^{2h} x_j^2 - 2e^h (e^h - 1) x_j^2 + (e^h - 1)^2 x_j^2 = x_j^2.$$

On the other hand, for arbitrarily small  $x_j, \forall \epsilon > 0, \exists H_2 > 0$  such that for  $0 < h < H_2$ ,

$$(8) \quad x_{j+1}^2 - x_j^2 = (e^{2h} - 1) x_j^2 + (e^h - 1)^2 \mu_l^2 - 2e^h (e^h - 1) \mu_l x_j < \epsilon.$$

To summarize, choose  $0 < h < \min(H_1, H_2)$ , for any  $x_j$  in the working region  $(-1, 1)$  and  $x_j^2 > \xi$ , we have  $x_{j+1}^2 < x_j^2$ ; and for any  $x_j$  that  $x_j^2 \leq \xi$ , we have  $x_{j+1}^2 - x_j^2 < \epsilon$ . So,  $\{x | x^2 < \xi + \epsilon\}$  is the unique attractor in the working region given sufficiently small  $h$ . This shows that the motion is consistent with the definition of regularity, and sufficiency is proved.  $\square$

Note: In this section, we used DFCB1 to illustrate the ideas, but the concepts of regularity and its condition (Lemma 1) are not confined to DFCB1. We consider regularity a very basic performance requirement. Only regular control designs are acceptable. Next, we examine the data rate requirements among regular control designs.

## 5. MINIMUM REQUIRED DATA RATE OF REGULAR CONTROL DESIGNS GIVEN THE NUMBER OF QUANTIZATION LEVELS

Given the number of control quantization levels and enforcing regularity, we will show below what control designs  $f$  require the minimum data rate. Our first theorem is for DFCB1 only.

**Theorem 1.** *For DFCB1, given  $N$  (half the number of control levels since the control is assumed symmetric), the regular control design (selection function  $f$ )*

that achieves the minimum data rate requirement is given by the set of equations

$$(9) \quad \begin{cases} b_N = 1, b_0 = 0 \\ (b_l - b_{l-1})e^{h_m} = (b_l + 1), l = 1, \dots, N \\ \mu_l = b_l, \quad l = 1, \dots, N \end{cases}$$

where  $h_m$  (also determined by this set of equations) is the largest tolerable sampling interval and the minimum required data rate is  $R_m = 2L/h_m$ ,  $L = \lceil \log_2 N + 1 \rceil$ .

*Proof.* First, equation set (9) has  $2N+2$  equations and  $2N+2$  variables ( $b_0, \dots, b_N, \mu_1, \dots, \mu_N, h_m$ ). To prove the existence of a meaningful solution, relax the condition that  $b_0 = 0$  and let  $h_m$  vary in the range  $(0, +\infty)$ . The modified equation set in the unknowns  $b_l$  and  $\mu_l$  is linear and always has a unique solution (easy to prove). Further,  $b_0$  is a continuous function of  $h_m$ .  $b_0 \rightarrow 1$  when  $h_m \rightarrow +\infty$  and  $b_0 \rightarrow -1$  when  $h_m \rightarrow 0$ . So, there must be a value of  $h_m$  such that  $b_0 = 0$ . This  $h_m$  together with the corresponding solution of the modified equation set is the solution of (9).

Next, we prove that the solution of (9) specifies the desired selection function  $f$ . This is a constrained optimization problem. The variables are  $b_l, l = 1, \dots, N-1$  and  $\mu_l, l = 1, \dots, N-1$  ( $b_0, b_N, \mu_N$  are preset, see (2)). The tolerable sampling interval (denoted  $\hat{h}$ ) is to be maximized. One can show that the system DFCB1 has bounded response under  $\hat{h}$  if and only if  $(\mu_l - b_{l-1})e^{\hat{h}} \leq (\mu_l + 1)$  for  $l = 1, \dots, N$ . (Trajectory starting at  $b_{l-1}$ , for which the control value is  $\mu_l$ , will not travel beyond  $-1$  after one sampling interval, so that the working region  $(-1, 1)$  is invariant.) So,

$$(10) \quad \hat{h} = \ln \min_{l=1, \dots, N} \frac{\mu_l + 1}{\mu_l - b_{l-1}}.$$

The constraints of the problem mainly come from the regularity condition, i.e.,  $b_l \leq \mu_l, l = 1, \dots, N$ . From (10),  $\hat{h}$  is a monotonically increasing function of all  $b_l$ . So,  $b_l$ 's must satisfy the constraints with equality, i.e.,  $b_l = \mu_l$  for all possible  $l$ . Then

$$(11) \quad \hat{h} = \ln \min_{l=1, \dots, N} \frac{b_l + 1}{b_l - b_{l-1}}.$$

Rewrite (11) as

$$(12) \quad \hat{h} = \ln \left( \min \left\{ \frac{1}{b_1} + 1, \frac{b_2 + 1}{b_2 - b_1}, \dots, \frac{b_{N-1} + 1}{b_{N-1} - b_{N-2}}, \frac{2}{1 - b_{N-1}} \right\} \right).$$

We have already shown in the beginning of this proof that the set of equations

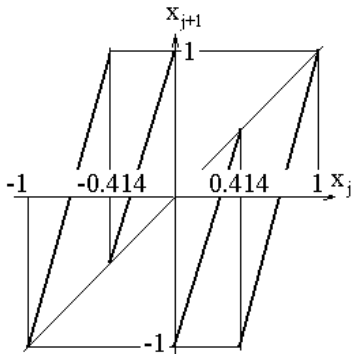
$$(13) \quad e^h = \frac{b_l + 1}{b_l - b_{l-1}}, \quad l = 1, \dots, N$$

has a solution. Write the solution as  $b_l^*, l = 1, \dots, N-1$  and  $h^*$ . Suppose a collection of  $b_l$ 's gives  $\hat{h} > h^*$ . According to (12),

$$(14) \quad \frac{b_l + 1}{b_l - b_{l-1}} \geq \hat{h} > h^* = \frac{b_l^* + 1}{b_l^* - b_{l-1}^*}, \quad l = 1, \dots, N.$$

If  $b_l > b_l^*$  for one  $l$ , to satisfy (14),  $b_{l-1} > b_{l-1}^*$  must hold. This would eventually lead to  $b_1 > b_1^*$  and  $\frac{1}{b_1} + 1 < \frac{1}{b_1^*} + 1 = h^*$ . A contradiction with (12) occurs. So  $b_l > b_l^*$  is impossible. Similarly,  $b_l < b_l^*$  is also impossible. So, the solution of (13) together with  $\mu_l = b_l, l = 1, \dots, N, b_0 = 0$  and  $b_N = \mu_N = 1$  is the optimal solution for maximizing tolerable sampling interval (minimizing required data rate).  $\square$

FIGURE 6. 4-level regular control that requires minimum data rate



- Remarks 2.**
- In designs given in Theorem 1, each control region has points that are mapped to the boundary of the working region..
  - Because of the regularity condition, only the outmost two control regions get mapped to the whole working region. Thus the inner control regions need to be smaller in volume than the outmost two. Actually, any inner control regions have smaller volumes than the outer ones. So, the resultant partition of working region is not uniform.

**Example 2.** Consider the four-level control, i.e.,  $N = 2$ . The regular control design that requires the minimum data rate turns out to be (figure 6)

$$f = \begin{cases} -1 & : x \leq -0.414 \\ -0.414 & : -0.414 < x < 0 \\ -0.414 & : 0 \leq x < 0.414 \\ 1 & : x \geq 0.414. \end{cases}$$

Note: in the graph, the line segments of state transition in each control region touch the  $x_{j+1} = x_j$  line at one end (because  $\mu_l = b_l$ ) and touch the  $x_{j+1} = 1$  or  $-1$  line at the other end (transition to the border of working region).

## 6. THE ABSOLUTELY LOWEST DATA RATE REQUIREMENT AND BINARY CONTROL

In the previous section, we found a way to calculate the minimum required data rate for groups of regular control designs with given number of quantization levels. In this section, we show that binary control achieves the strictly lowest data rate requirement among all regular control designs.

The following theorem follows directly from Theorem 1. Here we still consider DFCB1.

**Theorem 2.** For DFCB1, the binary selection function require strictly lower data rate than any other regular selection functions.

*Proof.* First, regularity is a trivial requirement for binary selection function. Second, we only need to compare the control designs given in Theorem 1, since they require the minimum data rates within the groups of different  $N$ 's.

For non-binary control, because  $b_l < b_N$ , we have

$$(15) \quad b_l - b_{l-1} = \frac{b_l + 1}{e^{h_m}} < \frac{b_N + 1}{e^{h_m}} = b_N - b_{N-1}, \quad l = 1, \dots, N-1.$$

Since  $\sum_{l=1}^N b_l - b_{l-1} = 1$ ,  $b_N - b_{N-1} \geq \frac{1}{N}$ . Equality holds only for the binary case. Then

$$(16) \quad e^{h_m} = \frac{2}{b_N - b_{N-1}} \leq 2N.$$

$$(17) \quad R_m = \frac{2[\log_2 N + 1]}{h_m} \geq \frac{2 \log_2 2N}{\ln 2N} = 2 \log_2 e.$$

Again, equality is achieved only by the binary case.  $\square$

**Remarks 3.** • *The above result follows quite easily from Theorem 1. Because of the regularity condition, the “data rate efficient” control designs given by Theorem 1 uses uneven quantization. This played a central role in the derivation of Theorem 2.*

- *It needs to be pointed out that by increasing the number of control levels, the system can tolerate longer sampling intervals. But, the above theorem indicates that the tolerable sampling interval grows slower than  $L$  the length of message needed to encode the sample and control values, assuming the actual delay grows proportionally with  $L$ . There are situations in reality that this assumption is not valid. This will be addressed in Section 7.*

Second, we show that for DFCB0, binary selection function still require lower data rate than others and the gap becomes even larger. Note: Although the theorem below compares the *required data rate* in DFCB1 and the *required base data rate* in DFCB0, the purpose is not the comparison itself. The comparison alone is meaningless because these two quantities concern different communication delay components. Rather, the purpose is to extend Theorem 2 to DFCB0.

**Theorem 3.** *The required base data rate of a regular control design in DFCB0 is greater or equal to its required data rate in DFCB1. Equality is achieved only when the design is binary. Thus, in DFCB0, binary control still require lower data rate than other regular control designs and the gap becomes even larger.*

*Proof.* Suppose a  $2N$  level regular control design  $f$  requires data rate  $R_1$  in DFCB1. The corresponding tolerable sampling interval is  $h_1 = \frac{2L}{R_1}$ . Note that  $f$  is not necessarily the most “data rate efficient” design in the sense of Theorem 1. We only need to show that when the same  $f$  is used in DFCB0 and the base data rate  $R_b$  takes the value of  $R_1$  — and the control interval  $\Theta$  takes the value of  $h_1$  — the state trajectory may leave the working region, thus become unbounded.

From the study of DFCB1, we know that

$$(18) \quad e^{h_1} = \min_{l=1, \dots, N} \frac{\mu_l + 1}{\mu_l - b_{l-1}}.$$

Suppose  $m$  is the minimizing index, i.e.,

$$(19) \quad e^{h_1} = \frac{\mu_m + 1}{\mu_m - b_{m-1}}.$$

This means that in the extreme case, a trajectory starting and sampled at  $b_m^+$  (a point in the control region  $[b_{m-1}, b_m)$  and arbitrarily close to the lower boundary)

will approach  $-1$  after one sampling interval  $h_1$ . Correspondingly in DFCB0, a trajectory starting and sampled at  $b_m^+$  will approach  $-1$  after  $\Theta$  units of time. For DFCB1, the control applied subsequently is  $f(-1) = -1$ , which will drive the trajectory back into the working region  $(-1, 1)$ . However, for DFCB0, the subsequent control value might not be  $-1$  since it could be based on an out-of-date sample. Actually, it can take any value in the admissible control value set  $U$  other than  $\mu_m, \dots, \mu_N$ . On the other hand, when the trajectory approaches  $-1$ , any control value that is not  $-1$  will push the trajectory out of the working region. So, generally, the tolerable  $\Theta$  needs to be smaller than  $h_1$ , and the required base data rate is higher than  $R_1$ , the required data rate in DFCB1. The only exception is binary control, where  $\pm 1$  are the only possible control values.  $\square$

The observation shown in the above proof will be further explored numerically in the next section.

## 7. THE ADVANTAGE MARGIN OF BINARY CONTROL

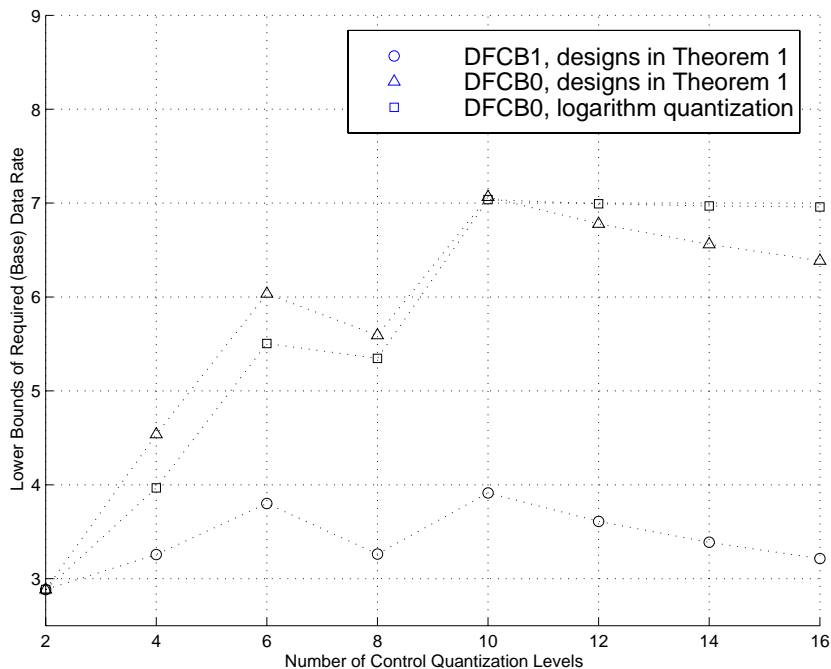
In this section, we show some quantitative details that support our argument. So far we have shown that regularity is a basic requirement for time-invariant control designs; and binary control is better than all other regular ones in the sense that it can tolerate strictly lower data rate and still achieve bounded response. The significance of the second conclusion is based on the assumption that the length of each control cycle depends linearly on the message length of sample and control. Otherwise the concept of data rate would not be an adequate characterization of the communication constraint. However, in the following two numerical studies, we show that the advantage margin of binary control is noticeable for DFCB1 and quite substantial for DFCB0.

In the proof of Theorem 3, we have touched the issue of asynchronism of sampling and control — a newly applied control can be still out-of-date. Actually, this can happen consecutively for many control cycles in the extreme cases. By further exploring this issue, we calculated lower bounds of the required base data rate (in DFCB0) for some typical control designs. We omit the details but show the results directly in figure 7. We studied three cases here. The result for the first case — DFCB1 with control designs specified in Theorem 1 — represents the exact data rate limit for different control designs in DFCB1. For the second and third case, we picked two “typical” types of control designs to illustrate the data rate limit in DFCB0. (The one labeled “logarithm quantization” refers to logarithmically quantized linear control [1].) Because the sampling/actuation asynchronism, and consequent use of out-of-date control values makes the system behavior quite chaotic (except for when binary control is applied), the exact data rate requirements for non-binary controls in DFCB0 are not clear. However, this question seems to be unimportant. Comparing the lower bounds of required base data rate for non-binary controls with the exact required base data rate of binary control is already enough to demonstrate the substantial advantage of binary control in DFCB0. (Other non-binary controls, e.g., evenly quantized linear control, act similarly to the examples shown here.)

## 8. CONCLUSION

Above, we have shown that binary control might be the appropriate control considering the regularity condition (Section 4) and asynchronism of sampling and

FIGURE 7. Lower Bounds of Required (Base) Data Rates



control actuation (Section 6 and 7). Numerical study shows that the advantage margin binary control is substantial in terms of *data rate* (Section 2).

Now, we are exploring control synthesis based on binary control. A straightforward approach is to add a side channel (whose bandwidth can be arbitrarily low) to adjust the magnitude of binary control so that asymptotic stability can be achieved with fixed data rate.

## REFERENCES

- [1] Jialing Liu and Nicola Elia, 2002. "Quantized Control with Applications to Mobile Vehicles," Proceedings of the 41st IEEE Conference on Decision and Control, pp. 2391-2396
- [2] F. Fagnani Fabio, B. Scanavino, S. Zampieri, and R. Garello, 2002. "Some Results on Combined Parallel Concatenated Schemes with Trellis-Coded Modulation," IEEE International Symposium on Information Theory - Proceedings, pp 444
- [3] J. Baillieul, 2002. "Feedback Designs in Information-Based Control," Stochastic Theory and Control, (Proceedings of the Workshop held at the University of Kansas on October 18-20, 2001), B. Pasik-Duncan, Ed., LNCIS 280, 2002, Springer-Verlag Berlin Heidelberg, pp. 35-57.
- [4] J. Baillieul, 2002. "Feedback Coding for Information-Based Control: Operating Near the Data-Rate Limit," Proceedings of the 2002 IEEE Conference on Decision and Control, Las Vegas, NV, December, pp. 3229-3236.
- [5] Wei Zhang, Michael S. Branicky, & Stephen M. Phillips, 2001. "Stability of Networked Control Systems," IEEE Control Systems Magazine, February, pp. 84-99.
- [6] S. Tatikonda & S. K. Mitter, 2001. "Control Under Communication Constraints," Preprint.
- [7] G. N. Nair & R. J. Evans, 2000. "Stabilization with Data-rate-limited feedback: tightest attainable bounds," System & Control Letters, 41(1), pp. 49-56.
- [8] J. Baillieul, 1999. "Feedback Designs for Controlling Device Arrays with Communication Channel Bandwidth Constraints," ARO Workshop on Smart Structures, Penn. State University, August 16-18.

- [9] W.S. Wong & R.W. Brockett, 1999. "Systems with Finite Communication Bandwidth Constraints II: Stabilization with Limited Information Feedback," *IEEE Trans. Automatic Control*, 44, pp. 1049-53.
- [10] S. TATIKONDA, A. SAHAI, & S. MITTER, 1998. "Control of LQG Systems Under Communication Constraints," *Proceedings of the 37-th Conference on Decision & Control*, Tampa, FL, pp. 1165-1170.
- [11] W.S. Wong & R.W. Brockett, 1997. "Systems with Finite Communication Bandwidth Constraints I: State Estimation Problems," *IEEE Trans. Automatic Control*, 42, pp. 1294-1298.
- [12] D.F. Delchamps, 1990. "Stabilizing a Linear System with Quantized State Feedback," *IEEE Trans. Automatic Control*, 35, pp. 916-924.
- [13] D.F. Delchamps, 1989. "Extracting State Information from a Quantized Output Record," *System & Control Letters*, 13, pp. 365-372.
- [14] P. Moroney, 1983. "Issues in the Implementation of Digital Compensators," M.I.T. press.
- [15] D.P. Bertsekas & I.B. Rhodes, 1971. "Recursive State Estimation for a Set-Membership Description of Uncertainty," *IEEE Trans. Automatic Control*, AC-16, pp. 117-128.
- [16] R.E. Curry, 1970. "Estimation and Control with Quantized Measurements," M.I.T. press.
- [17] F.C. Schweppe, 1968. "Recursive State Estimation: Unknown but Bounded Errors and System Inputs," *IEEE Trans. Automatic Control*, AC-13, pp. 22-28.