Introduction on Bit Memory Systems: Approximation and Stabilization

Koji Tsumura

***Abstract*— In this paper, we introduce recent results on bit memory systems. The bit memory systems are operators from analog inputs to discretized outputs and their memories are elements of discrete numbers. Their dynamics is represented by the time evolution of the bit memories and output equations. In this paper, we consider an approximation problem or a stabilization problem of ordinary linear time invariant systems by the bit memory systems. Then, we consider the minimization problem of the bit length of the memories with which the bit memory systems attain a given error bound for the approximation problem or the stability of the closed loop systems for the stabilization problem. We show several upper and lower bounds of the bit length of memories in both of the problems.**

I. Introduction

In a series of literatures [15], [16], [17], we introduced bit memory systems in which the state variable takes the value in a countable set of a finite number of the elements. The motivations of considering such systems are much related to the recent active research fields of control theory on networked control systems as explained below.

In the last few years, networked control problems have been actively investigated and one of interesting problems is stabilization under constraints of channel capacities for transmitting signals [21], [22], [1], [8], [2], [9], [19], [20], [10], [14]. A key technique for this problem is quantization of analog signals and the result shows necessary complexity of signals for stabilization.

We recognize the importance of the results, however, the focus of the research is to decrease the complexity of signals on channels between subsystems and the complexity of dynamical systems which generate the coded signal has been often ignored. As a result, the optimal coders usually become highly complex systems and they do not consist with the original purpose of the networked control where computation power is naturally supposed to be limited on the side of the controlled plants. This fact indicates a necessity of measuring the computation complexity of dynamical systems and the stability problem should be replaced by one in which the complexities of the coders/decoders are also constrained.

Another issue related on the complexity of dynamical systems is in distributed control systems. Distributed control systems having communication networks appear in several situations such as multi-agent systems or fault-tolerant systems. In order to realize such systems, preferable distribution of computation loads to the subsystems is a significant issue. It is reasonable to locate less-complex subsystems to the end of the network structure and complex subsystems to the junction in the networks. For discussing such problem, the notion of the complexity of dynamical systems is necessary.

The final example, embeded control, is more directly. The embedded control systems are supposed to be realized on chips in which the computation power or the memories are strictly limited. A stabilization problem under the limitation of the computation power is the challenge in such research field.

Motivated by those issues, we introduced a notion of complexity of dynamical systems [15], [16], [17]. We have explained that ordinary degree is not necessarily enough to represent the complexity of dynamical systems for the above mentioned purpose because in the networked control problem, the variation of “discrete-valued” signals plays an important role and degree is not proper for dealing with it directly. In the series of our papers, we focus on the memory of systems and consider cases that it takes discrete numbers. We call such dynamical system as a bit memory system, and then, we regard the bit length of the memory as the complexity of systems.

In [18], [15], [17], we considered to find the necessary bit length of the memory to attain the given performance for the approximation and gave the lower/upper bounds of the necessary bit length. Another control problem considered in the papers [16] is stabilization by the bit memory systems and we have tried to find the necessary bit length for the memory to attain the stability. This paper is the tutorial on those results.

The much related results are [4], [12], [11], [13] where the focus of the research is on the algorithm to derive the discrete abstraction for the original analog-valued systems. However, unfortunately they do not concern with the the necessary complexity which is required for the networked control problem. The other related work in the field of digital filter is on the optimal realization problem with respect to the quantization error (round off noise) of system variables [7], [5].

II. Approximation Problem

A. Bit memory system and formulation

In general, approximation problem is to derive a less complex model which imitates the input/output behavior of the original system. In this section, we consider the simple case that the original is the following discrete time SISO

The author is with Department of Information Physics and Computing, The University of Tokyo, Hongo 7–3–1, Bunkyo-ku, Tokyo 113–8656, Japan, tel: +81–3–5841–6891, fax: +81–3–5841–6886, e-mail: tsumura@i.u-tokyo.ac.jp

ISBN 978-963-311-370-7

n-th degree linear system P:

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t), \end{aligned} \quad (1)$$

where $x \in \mathcal{R}^n$, $u \in \mathcal{R}$, $y \in \mathcal{R}$. We also assume that A is asymptotically stable, A and B are controllable and A and C are observable.

Ordinary approximation methods are to derive a model which reduces the degree of the state variable x. Although this methodology has been well-investigated and it is widely used (e.g., see [6], [3]), we can also point out some drawbacks. The first one is that degree of systems is too coarse to manage the resolution of approximated systems. The second one is that on-line tuning of resolution of systems is preferable for application, however, it is usually difficult. The final one is that the degree of systems does not necessarily consist with information theoretic complexity of signals.

With the above discussion, in [18], [15], [17] we introduced a discrete time SISO bit memory system $\tilde{P}$ for approximating the dynamics of (1) as

$$\begin{aligned} \mathcal{I}(t+1) &= F(\mathcal{I}(t), u(t)), \\ \tilde{y}(t) &= H(\mathcal{I}(t)), \end{aligned} \quad (2)$$

where $u \in \mathcal{R}$, $\tilde{y} \in \mathcal{R}$, and $\mathcal{I} \in \{0,1\}^L$. The signal $\mathcal{I}$ is the state variable of the system and it is called "bit memory". The function F or H is a static mapping from the pair of $\mathcal{I}(t)$ and $u(t)$ to $\mathcal{I}(t+1)$ or from $\mathcal{I}(t)$ to the output $\tilde{y}(t)$, respectively, and therefore, (2) is a time invariant system.

The quantity L is called "bit length" of memory. The large number of L implies the variety of the bit memory and it also means the complexity of the behavior of the dynamics. More directly, when the bit memory system is implemented on a computer, the bit memory is realized on a memory or a register of the computer and the bit length means the binary word length. The computation complexity can be directly represented by the word length of the memory or the registers. From above, the bit memory system and the notion of the bit length of memory are appropriate for representing the complexity of dynamical systems in information theoretic view point.

Then we consider the next approximation problem:

Problem 2.1: For the systems (1) with a given subset $\Omega_x \subseteq \mathcal{R}^n$ and $\{u\}$, find a bit memory system (2) which minimizes L such that

$$|y(t) - \tilde{y}(t)| \leq \frac{\epsilon}{2}, \ \ \forall y(t) \text{ where } x(t) \in \Omega_x. \quad (3)$$

This problem can be used in several ways. The first one is simply for deriving a simple model of the original system. The second one is that we can measure the complexity of the dynamics of the original system (1) with a tolerance ϵ from a view point of computation complexity. When the complexity of the dynamics is high, the necessary bit length with a tolerance ϵ in the above problem naturally becomes large and vice versa.

A subclass of (2), which is explicitly intended to approximate (1), is given by

$$\begin{aligned} \mathcal{X}(t+1) &= Q(A\mathcal{X}(t) + Bu(t)) \\ \mathcal{Y}(t) &= C\mathcal{X}(t), \\ \tilde{y}(t) &= J(\mathcal{Y}(t)), \end{aligned} \quad (4)$$

where $x(0) \in \mathcal{X}(0) \subseteq \Omega_x$, $\Omega_x = \cup_{i=1,\dots,2^L}\mathcal{X}_i$, $\mathcal{X}_i \in \{\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_{2^L}\}$, $\mathcal{X}_i \subset \mathcal{R}^n$, Q: quantizer, $\mathcal{I}(t)$: $\mathcal{Z} \to \{1, 2, \dots, 2^L\}$, which represents the index of $\mathcal{X}_i$ when $\mathcal{X}(t) = \mathcal{X}_i$, $u, \tilde{y} \in \mathcal{R}$, $J(\bullet)$: center of the set, and the set-valued operations are defined by

$$\begin{aligned} M\mathcal{X} &:= \{Mx \mid x \in \mathcal{X}\}, \ M \in \mathcal{R}^{\bullet \times n}, \\ \mathcal{X} + w &:= \{x + w \mid x \in \mathcal{X}\}, \ w \in \mathcal{R}^n. \end{aligned} \quad (5)$$

Note that Q is a memory-less and time invariant quantizer and it can be regarded as a mapping from the index i of $\mathcal{X}(t)$ and $u(t)$ to the index i of $\mathcal{X}(t+1)$. From above, the system (4) can be regarded as a dynamical system which maps the initial bit memory $\mathcal{X}(0)$ and the input sequence $\{u(t)\}$ to the output sequence $\{\tilde{y}(t)\}$.

Obviously the systems (4) is just a subclass of (2), however we can show a kind of equivalence between those classes in the following lemma.

Lemma 2.1 ([17]): For each bit memory system (2), there exists a system (4) which conserves the same input/output behavior and has the same bit length of memory.

This lemma supports substituting Problem 2.1 by the next more tractable problem:

Problem 2.2: For the systems (1) and (4), find Q which minimizes L such that

$$|y(t) - \tilde{y}(t)| \leq \frac{\epsilon}{2}, \ \ \forall y(t) \text{ where } x(t) \in \Omega_x. \quad (6)$$

B. Upper and lower bounds of L for n-th degree systems

For Problem 2.2, we gave an upper bound of the minimum L in [18], [15], by restricting the class of (4). In [18], [15], we assume that the quantized subsections $\{\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_{2^L}\}$ are restricted to be time invariant, uniform and orthogonal to the basis $\{x_i\}$. Moreover, the mapping of sets Q is defined by

$$Q(\mathcal{S}) := \mathcal{X}_i \supseteq \mathcal{S}. \quad (7)$$

The bit memory $\mathcal{I}(t)$ is defined as the index for $\mathcal{X}_i$.

Before describing assumptions, we define a notation for a matrix $M = m_{ij}$:

$$\overline{M} := |m_{ij}|. \quad (8)$$

Assumption 2.1 ([18], [15]): The matrix A satisfies the following conditions:

1)
$$0 < |\alpha_i| < 1, \ \forall i, \ \alpha_i : \text{ eigenvalue of } \overline{A} \quad (9)$$

2) there exists vector x satisfying

$$\overline{x} > \overline{A}\,\overline{x} \quad (10)$$

Then, we get the following results:

Theorem 2.1 ([18], [15]): Under Assumption 2.1, for the systems (1) and (4), an upper bound of the bit length L at $t \to \infty$ for Problem 2.2 is

$$\log_2 \frac{m(\Omega_x)}{D}, \tag{11}$$

$$D := \frac{\epsilon^n \prod_j (1-\alpha_j)^n}{2^n n^n} \prod_{i=1}^{n} \frac{1}{\overline{C}(\mathrm{adj}(I-\overline{A}))_i}, \tag{12}$$

$m(\bullet)$ denotes the volume of $\bullet$ and $(\mathrm{adj}(I-\overline{A}))_i$ denotes the i-th column vector respectively.

We next consider to derive the lower bounds for the minimum bit length L of any types of (4) satisfying (6). We get the following result:

Theorem 2.2 ([17]): For the systems (1) and (4), let $u(t) \in [-\nu/2, \nu/2]$ and Ω_x be the subset of $x(n)$ which is reachable from $x(0) = 0$ by $u(0), u(1), \ldots, u(n-1) \in [-\nu/2, \nu/2]$. A lower bound of the bit length L for Problem 2.1 is

$$\left\lceil \log_2 \frac{|\det \Gamma| \nu^n}{(1-\prod |a_i|)\epsilon^n} \right\rceil, \tag{13}$$

where

$$\Gamma = \begin{bmatrix} y_1^I & y_2^I & y_3^I & \cdots & y_n^I \\ y_2^I & y_3^I & y_4^I & \cdots & y_{n+1}^I \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ y_n^I & y_{n+1}^I & y_{n+2}^I & \cdots & y_{2n}^I \end{bmatrix}, \tag{14}$$

y_i^I, $i = 1, \ldots,$ are the impulse response of the original system (1) and a_i is the eigenvalue of the matrix A.

In the case of (1) of the first degree (i.e., $n = 1$), we can get a more strong result on the lower bound:

Corollary 2.1 ([17]): For the systems (1) of $n = 1$ and (4), let $u(t) \in [-\nu/2, \nu/2]$ and Ω_x be the subset of $x(n)$ which is reachable from $x(0) = 0$ by $u(0) \in [-\nu/2, \nu/2]$. The minimum bit length of (4) which solves Problem 2.2, i.e. Problem 2.1, is given by

$$\left\lceil \log_2 \frac{|bc|\nu}{(1-|a|)\epsilon} \right\rceil. \tag{15}$$

Remark 2.1: The lower bound (13) is essentially given only by the system properties, that is, the poles of the original system P and a kind of Hankel singular values. The magnitude of the poles closes to 1, the lower bound of the minimum bit length becomes large. On the other hand, Hankel singular values represent the intensity of the controllability and the observability of the original n-th degree systems. This means when the original n-th degree systems are fully complex in the whole dimension of the state space, then, the complexity in the sense of the bit memory system is also high. This result represent an interesting connection between the minimum bit length of memory and the degree of the original systems.

C. Numerical example

In order to illustrate the results, we show some numerical examples for Theorem 2.1 and Corollary 2.1. At first, consider the system (1) of the first degree for Corollary 2.1 where

$$A = 0.8, B = 1, C = 1$$

and set $\epsilon = 1$ in Problem 2.2. The minimum bit length (15) with $\nu = 2$ can be calculated as

$$\left\lceil \log_2 \frac{|bc|\nu}{(1-|a|)\epsilon} \right\rceil = \lceil 2.3219 \rceil = 3. \tag{16}$$

We can also derive a bit memory system given in [17], which has this minimum bit length, and demonstrate the response of $x(t)$ and $\mathcal{X}(t)$ for $u \in [-1, 1]$. In Fig. 1, we can find that the true response $x(t)$ (marked by '$\times$') is successfully captured by $\mathcal{X}(t)$ (marked by '$|$') width 1 at each time step.

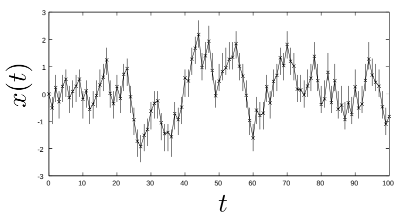

Fig. 1: Transition of $x(t)$ (marked by '$\times$') and $\mathcal{X}(t)$ (marked by '$|$')

We next explain the result for 2nd degree system:

$$A = \begin{bmatrix} 0.7416 & -0.1121 \\ 0.1121 & 0.7416 \end{bmatrix} \; B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \; C = \begin{bmatrix} 1 \\ 4 \end{bmatrix} \tag{17}$$

The poles of this system are 0.7416 $\pm$ i 0.1121. For this original system, we derive the lower bound of the bit length by using Theorem 2.2. With the performance index $\epsilon = 1$, $\nu = 2$, the lower bound (13) is calculated as

$$\left\lceil \log_2 \frac{|\det \Gamma| \nu^2}{(1-\prod |a_i|)\epsilon^2} \right\rceil = \lceil 8.1827 \rceil. \tag{18}$$

On the other hand, by employing the method in Theorem [18], [15], we derive a bit memory system which exactly takes the upper bound of the bit length (11) and guarantees the performance index $\epsilon = 1$.

We simulate $x(t)$ and $y(t)$ of the original system and $\mathcal{X}(t)$ and $\tilde{y}(t)$ of the derived bit memory system for a uniform random input $u(t) \in [-1, 1]$ with the initial condition $x(0) = [-0.5793 \;\; -1.8894]^{\mathrm{T}}$. Fig. 2 and Fig. 3 show their transitions during $t = 0-100$ or $t = 0-50$, respectively. The maximum of the output error during $t = 0-100$ is $\max |y(t) - \tilde{y}(t)| = 0.4026$ and we can confirm the error bound $\epsilon/2 = 0.5$ is guaranteed.

The bit length of this system is

$$\left\lceil \log_2 \frac{m(\Omega_x)}{D} \right\rceil = \lceil 10.9482 \rceil \text{ bit}, \tag{19}$$

and the difference between (18) and (19) is 2 bit. This example suggests that the difference between the upper bound and the lower bound is not so seriously large.

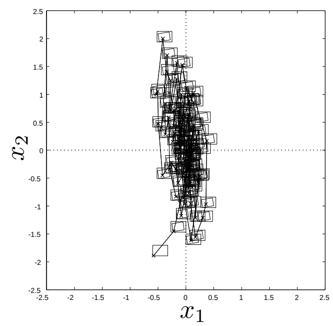

Fig. 2: Transition of $\mathcal{X}(t)$ during $t = 0-100$ ('x' and each rectangular show $x(t)$ and $\mathcal{X}(t)$ respectively)

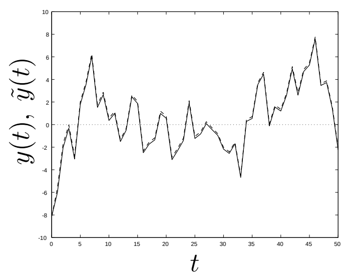

Fig. 3: Output responses $y(t)$ and $\tilde{y}(t)$ during $t = 0-50$ ($y(t)$: solid line, $\tilde{y}(t)$: dashed line)

III. Stabilization Problem

A. Motivation and formalation

In the last decade, a networked control problem of stabilization under constraints of channel capacity has been actively investigated [21], [22], [1], [8], [2], [9], [10], [14], [20]. The simplest situation of the problem is shown in Fig. 4, where the plant P and the controller K are connected over channels with finite capacities.

In order to attain small signal complexity of i and j, the coders and the decoders are required to be complex and a lot of memory is necessary because they use the past record of y or u, or i or j to compute their output signals. Such demand of the computation power especially for C_o or D_i, which are on the side of the plant, is not consistent with the original purpose of the networked control.

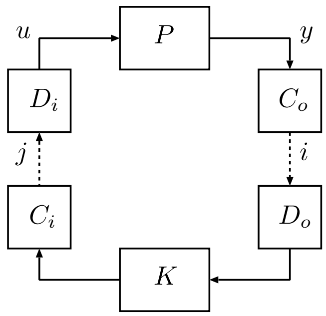

Fig. 4: Closed loop system with data transmission

This result shows a fact of trade-off between the complexities of the transmitted signals and the subsystems, and as a whole the controller including the coders and decoders becomes complex. Motivated by this fact, we considered a problem to stabilize a plant with a controller which has a constraint of computation complexity [16], i.e., stabilization via a bit memory system with constraints on the bit length. In the following of this section, we introduce the results in [16].

The plant is the following discrete time SISO n-th degree linear system P:

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t), \end{aligned} \tag{20}$$

where $x \in \mathcal{R}^n$, $u \in \mathcal{R}$, $y \in \mathcal{R}$, $x(0)$ is assumed to be in a closed subset $\mathcal{X}_o$ with a size $\nu_i := m(\mathcal{X}_{oi})$, m denotes the volumes, A and B are controllable and A and C are observable. The controller K is the following discrete time SISO bit memory system:

$$\begin{aligned} \mathcal{I}(t+1) &= F(\mathcal{I}(t), y(t)) \\ u(t) &= H(\mathcal{I}(t)) \end{aligned} \tag{21}$$

where $\mathcal{I} \in \{0,1\}^L$ is the bit memory of bit length L. Note that the time evolution of the bit memory or the output is defined only by the bit memory or the input at that time and it does not use their past records.

Fig. 5 shows a block diagram of the closed loop system with a bit memory controller where the bit memory $\mathcal{I}(t)$ is transfered from subsystem F to H. When we regard the path from F to H as transmission channel in a networked control system, then a control problem with the bit length limitation corresponds one with data rate constraint.

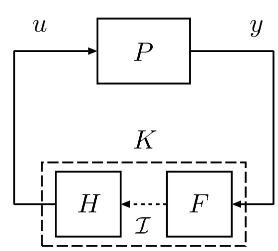

Fig. 5: Control system with bit memory controller

The bit memory controllers of finite bit lengths can not attain the asymptotic stability and then we define another notion of stability as follows.

Definition 3.1: A dynamical system is said to be practically stable with ϵ when its output $y(t)$ satisfies

$$|y(t)| \le \frac{\epsilon}{2} \tag{22}$$

at $t \to \infty$.

Then we introduce a problem:

Problem 3.1: For the plant (20), find a bit memory controller (21) of L bit length satisfying the practically stability with ϵ of the closed loop system.

In [16], we gave the upper bounds of the bit length k in Problem 3.1 by showing examples of stabilizing bit memory controllers. In the following, we show them.

B. Bit memory observer approach

In this subsection, we give a bit memory controller for Problem 3.1 with the following assumption:

Assumption 3.1: The matrix A in (20) is invertive.

Here we introduce a observer-based bit memory controller K as follows.

$$\mathcal{X}^\star(t+1) = Q(A(\mathcal{X}^\star(t) \cap \mathcal{W}_0(y(t),C)) + Bu(t)) \quad (23)$$

$$\mathcal{U}(t) = K(\mathcal{X}^\star(t) \cap \mathcal{W}_0(y(t),C)) \quad (24)$$

$$u(t) = J(\mathcal{U}(t)) \quad (25)$$

where

$$\mathcal{W}_\epsilon(y,C) := \left\{ x \,|\, y - \frac{\epsilon}{2} \le Cx \le y + \frac{\epsilon}{2} \right\}.$$

The state variable $\mathcal{I}(t)$ of the bit memory controller is the index assigned to $\mathcal{X}^\star(t)$ and note that the memory in the above operations (23)–(25) is only $\mathcal{I}(t)$.

We can derive the estimation error between $x'(t) \in \mathcal{X}^\star(t)$ and the true state variable $x(t)$ of the plant:

Lemma 3.1: When $t \to \infty$, the following inequality holds.

$$|Kx' - Kx(t)| \le \tilde{\epsilon}, \ \forall x' \in \mathcal{X}^\star(t) \cap \mathcal{W}_0(y(t),C) \quad (26)$$

$$\tilde{\epsilon} := \overline{K\,A^n\,\mathcal{O}^{-1}} \cdot \mathbf{1} \cdot \overline{CA^{-1}}\, d$$

$$\mathcal{O} := \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}$$

$$\mathbf{1} := [1\ 1\ \cdots\ 1]^{\mathrm{T}}$$

With this lemma, we can derive the following theorem:

Theorem 3.1: Consider the closed loop system composed of the plant (20) and the controller (23)–(25). Then, an upper bound of the bit length L at $t \to \infty$ is

$$\log_2 \frac{\prod_{i=1}^n \kappa_i}{D}, \quad (27)$$

where

$$\begin{aligned}
\kappa_i &:= \left\| ((z-(A+BK))^{-1})^i \begin{bmatrix} \nu_o & \tilde{\epsilon}B \end{bmatrix} \right\|_1, \\
D &:= \frac{\tilde{\epsilon}^n}{n^n h^n \prod_i \overline{g}_i}, \\
\tilde{\epsilon} &:= \overline{K\,A^n\,\mathcal{O}^{-1}} \cdot \mathbf{1} \cdot \overline{CA^{-1}}\, d^*, \\
g &:= \begin{bmatrix} g_1 & g_2 & \ldots & g_n \end{bmatrix} := CA^{-1}, \\
d_i^* &= \frac{\overline{K\,A^n\,\mathcal{O}^{-1}} \cdot \mathbf{1} \cdot \epsilon}{nh\overline{g}_i}, \\
h &:= \overline{KA^n\mathcal{O}^{-1}} \cdot \mathbf{1},
\end{aligned}$$

$((z-(A+BK))^{-1})^i$ denotes its i-th row vector and $\|\cdot\|_1$ means ℓ_1 norm.

The condition (27) implies that when the poles of the plant is large, the upper bound of the bit length becomes large and vice versa. This results consist with that of stabilization under constraints on channel capacities [21], [22], [1], [8], [2], [9], [19], [20], [10], [14].

C. Controller approximation approach

In this subsection, we show another approach giving a bit memory controller for Problem 3.1.

Under the assumption of the controllability and observability of the plant (20), there always exits an LTI stabilizing controller K:

$$\begin{aligned} x_K(t+1) &= A_K x_K(t) + B_K y(t), \\ u(t) &= C_K x_K(t), \end{aligned} \quad (28)$$

where $x_K(t) \in \mathcal{R}^m$. Here we assume the following:

Assumption 3.2: The controller K is stable.

This condition is called strongly stabilizability [23] and the necessary and sufficient condition, called pip condition, is known.

Then, by employing the discussion in Section II, we can introduce a bit memory controller $\tilde{K}$:

$$\begin{aligned} \mathcal{X}(t+1) &= Q(A_K\mathcal{X}(t) + B_K y(t)) \\ \mathcal{U}(t) &= C_K\mathcal{X}(t), \\ u(t) &= J(\mathcal{U}(t)), \end{aligned} \quad (29)$$

which approximates K such as described in Section II.

Hereafter, we assume that the quantized subsections $\{\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_{2^L}\}$ are restricted to be time invariant, uniform and orthogonal to the basis $\{x_i\}$. Moreover, the mapping of sets Q is defined by (7) as mentioned in Section II.

Now, we assume the following:

Assumption 3.3: The matrix A_K satisfies the following conditions:

1)

$$0 < |\alpha_i| < 1, \ \forall i, \ \alpha_i : \text{ eigenvalue of } \overline{A}_K \quad (30)$$

2) there exists vector x satisfying

$$\overline{x} < \overline{A}_K \overline{x} \quad (31)$$

Then, we get the following result:

Theorem 3.2: For the systems (20) and (29), an upper bound of the bit length L at $t \to \infty$ for Problem 3.1 is

$$\log_2 \frac{\prod_{i=1}^m \kappa_i}{D}, \quad (32)$$

where

$$\begin{aligned}
\kappa_i &:= \left\| ((zI - A_K)^{-1})^i B_K \cdot \frac{1}{1+PK} \begin{bmatrix} \epsilon & \overline{C} \cdot \nu_o \end{bmatrix} \right\|_1 \\
\nu_o &:= [\nu_1\ \nu_2\ \ldots\ \nu_n]^{\mathrm{T}} \\
D &:= \frac{\epsilon^m \prod_j (1-\alpha_j)^m}{2^m m^m} \prod_{i=1}^m \frac{1}{\overline{C}_K(\text{adj}(I-\overline{A}_K))_i},
\end{aligned}$$

$((zI - A_K)^{-1})^i$ and $(\text{adj}(I-\overline{A}_K))_i$ denote their i-th row vector and i-th column vector, respectively, and $\|\cdot\|_1$ means ℓ_1 norm.

The condition (32) corresponds to the condition (27) and the similar discussion as mentioned in Subsection III-B is applicable. On the other hand, we can observe that the assumptions for this approximation method are strong compared to the case of the observer-based approach.

D. Numerical example

In this subsection, we show a numerical simulation of stabilization by a bit memory controller, which is given by the controller approximation approach. The plant is given by

$$P(z) = \frac{z^2 - 0.398z + 0.04}{z^3 - 0.398z^2 - 3.96z + 0.816}$$

and the poles are 0.2040, 2.0992, and -1.9053. An ordinary stabilizing controller is

$$A_K = \begin{bmatrix} 0.1990 & -0.0200 \\ 0.0200 & 0.1990 \end{bmatrix}, \ B_K = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \ C_K = \begin{bmatrix} 1 \\ 4 \end{bmatrix}.$$

We next derive a bit memory controller (29) which approximates above. In the case of $\epsilon = 3$, the upper bound of the bit length given in Theorem 3.2 is $L = \lceil 7.1814 \rceil$ bit.

The transition of $\mathcal{X}(t)$ of the bit memory controller and that of output $y(t)$ are shown in Fig. 6 and Fig.7, respectively:

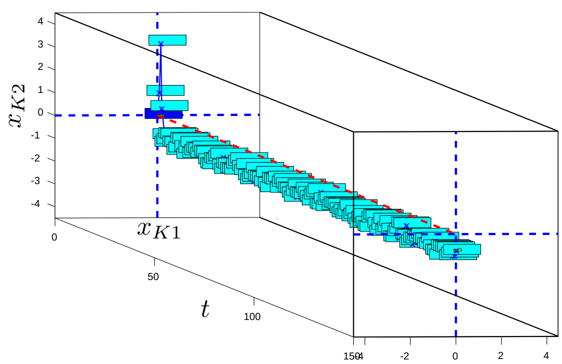

Fig. 6: Transition of $\mathcal{X}(t)$ ('x' and each rectangular show $x_K(t)$ and $\mathcal{X}(t)$, respectively)

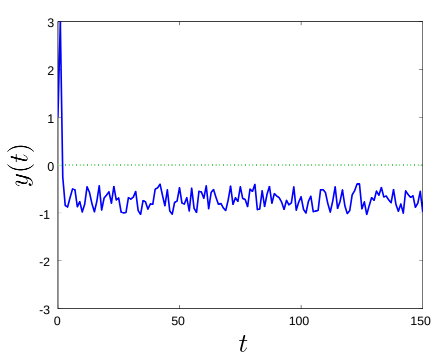

Fig. 7: Time response of $y(t)$

E. Discussion

1) Bit length vs performance: From the condition (27) and also (32), roughly speaking, the relationship between the control performance ϵ and the necessary bit length L is

$$L \propto \alpha + \beta \log_2 \nu^n + \gamma \log_2 \epsilon^{-n}, \tag{33}$$

or

$$L \propto \alpha + \beta \log_2 \nu^m + \gamma \log_2 \epsilon^{-m}, \tag{34}$$

where the ratios α, β and γ depend on the system property such as poles.

2) Plants with observation noise: In the case of Subsection III-B, consider a case that $y(t)$ is added by observation noise bounded by η. In this case, by the substitution $\frac{\epsilon}{2} - \eta \to \frac{\epsilon}{2}$, we can also construct the controller in the similar way o (23)–(25).

On the other hand, in the case of Subsection III-C, we can also derive the similar manipulation. Here we assume that the observation noise $v(t)$ is bounded by η, then, the magnitude of its influence on the true output $y(t)$ is bounded by

$$\eta \left\| \frac{PK}{1+PK} \right\|_1 =: \xi, \tag{35}$$

where K is the original nominal controller. Therefore, we substitute $\frac{\epsilon}{2} - \xi \to \frac{\epsilon}{2}$ and construct the controller, we can directly get the similar results in Theorem 3.2.

3) Unbounded noise or set of initial state: In the case of unbounded observation noise or set of the initial state of the plant, we no longer expect finite bit length of the memory. However, when the distribution density of the states is enough concentrated around the origin, by using variable bit length coding as in the information theory, we may get a finite expected bit length of the memory. Such method is practically useful as in the information theory. Case of the unbounded set of the initial state of the plant is also in the similar situation.

IV. Conclusion

In this paper, we introduced the recent results on approximation and stabilization problems with bit memory systems. We regard the bit length of memory as the complexity of the dynamical systems, and showed the relationship between it, the property of the plant, and the attained performance. In this paper, we omit the details of the proofs for the lemmas or the theorems. Refer to the relevant literature [15], [16], [17]. We also suggested that the lower bounds of the necessary bit length for the stabilization problem have much relationship with the results on the necessary channel capacities for the stabilization in [21], [22], [1], [8], [2], [9], [19], [20], [10], [14]. This issue will be appear in near future.

Acknowledgement: This research is supported by the Ministry of Education, Culture, Sports, Science and Technology, Japan, under Grant No. 21246067.

References

[1] R. W. Brockett and D. Liberzon, Quantized feedback stabilization of linear systems, *IEEE Trans. Automat. Control*, AC-45-7, 1279–1289, 2000.

[2] N. Elia and S. K. Mitter, Stabilization of linear systems with limited information, *IEEE Trans. on Automatic Control*, AC-46-9, 1384–1400, 2001.

[3] K. Glover, All optimal hankel-norm approximations of linear multivariable systems and their L^∞-error bounds, *Int. J. Cont.*, 39, 1115–1193, 1984.
[4] J. Lunze, Qualitative modelling of linear dynamical systems with quantized state measurements, *Automatica*, 30-3, 417–431, 1994.
[5] M. Kawamata, M. Iwatsuki and T. Higuchi, Balanced realizations as minimum sensitivity structures in linear systems, *Trans. of The Society of Instrument and Control Engineers*, 21-9, 900–906, 1985 (in Japanese).
[6] B. C. Moore, Principal component analysis in linear systems : Controllability, observability, and model reduction, *IEEE Trans. Automatic Control*, AC-26-1, 17–32, 1981.
[7] C. T. Mullis and R. A. Roberts, The use of second-order information on the approximation of discrete-time linear systems, *IEEE Trans. Acoustics, Speech, and Signal Processing*, ASSP-24-3, 226–238, 1976.
[8] G. N. Nair and R. J. Evans, Stabilization with data-rate-limited feedback: tightest attainable bounds, *Systems & Control Letters*, 41, 49–56, 2000.
[9] G. N. Nair and R. J. Evans, Mean square stabilisability of stochastic linear systems with data rate constraints, *Proceedings of the 41st IEEE Conference on Decision and Control*, 1632–1637, 2002.
[10] G. N. Nair and R. J. Evans, Stabilizability of stochastic linear systems with finite feedback data rates, *SIAM J. Control Optim.*, 43-2, 413–436, 2004.
[11] J. Raisch, Discrete abstractions of continuous systems –an input/output point of view, *Mathematical and computer modelling of dynamical systems*, 6-1, 6–29, 2000.
[12] J. Raisch and Siu D. O'Young, Discrete approximation and supervisory control of continuous systems, *IEEE Trans. Information Theory*, AC-43-4, 569–573, 1998.
[13] D. C. Tarraf, A. Megretski, and M. A. Dahleh, Finite automata approximations with error bounds for systems with quantized actuation and measurements: A case study, *Proceedings of the 43rd IEEE Conference on Decision and Control*, 1436–1441, 2004.
[14] S. Tatikonda and S. Mitter, Control under communication constraints, *IEEE Trans. Automat. Control*, AC-49-7, 1056–1068, 2004.
[15] K. Tsumura, Approximation of discrete time linear systems via bit length of memory, *MTNS2006*, 2006.
[16] K. Tsumura, Stabilization of linear systems by bit-memory controllers under constraints of bit-length, *Proceedings of the 45th IEEE Conference on Decision and Control*, 5507–5512, 2006.
[17] K. Tsumura, A Lower Bound of the Necessary Bit Length of Memory for Approximation of Discrete Time Systems. *Proceedings of the 46th IEEE Conference on Decision and Control*, 2241–2246, 2007.
[18] K. Tsumura, Complexity and resolution of approximated bit-memory systems. *SICE 6th Annual Conference on Control Systems*, 287–292, May 2006.
[19] K. Tsumura and J. Maciejowski, Optimal quantization of signals for system identification. *Technical Report of The Univ. Cambridge*, CUED/F-INFENG/TR445, 2002 (also in Proceedings of the Euromean Control Conference 2003, Cambridge, UK, 2003).
[20] K. Tsumura and J. Maciejowski, Stabilizability of SISO control systems under constraints of channel capacities. *Proceedings of the 42th Conference on Decision and Control*, 193–198, 2003.
[21] W. S. Wong and R. W. Brockett, Systems with finite communication bandwidth constraints – part I: State estimation problems, *IEEE Trans. Automat. Control*, AC-42-9, 1294–1299, 1997.
[22] W. S. Wong and R. W. Brockett, Systems with finite communication bandwidth constraints – II: Stabilization with limited information feedback, *IEEE Trans. Automat. Control*, AC-44-5, 1049–1053, 1999.
[23] D.C. Youla, J.J. Bongiorno, and C.N. Lu, Singla-loop feedback-stabilization of linear multivariable dynamical plants, *Automatica*, 10, 159–173, 1974.