2011 American Control Conference
on O'Farrell Street, San Francisco, CA, USA
June 29 - July 01, 2011

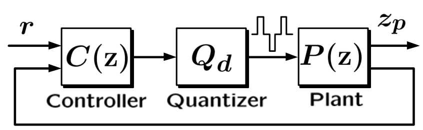
Synthesis of dynamic quantizers for quantized feedback systems within invariant set analysis framework

Kenji Sawada and Seiichi Shin

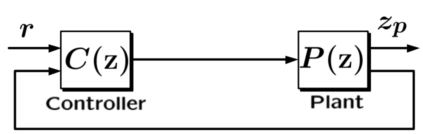
Abstract— **This paper focuses on quantized feedback systems which consist of the linear system and the dynamic quantizer. From the perspective of the invariant set analysis, we discuss the effectiveness and the limitation of our quantizer synthesis condition considering several performances. If the all transmission zeros of the given system are stable, the proposed quantizer gives an optimal output approximation property. Otherwise, our method can design a stable suboptimal quantizer guaranteed with infinite time control performance. The numerical examples show that our quantizer can achieve the coarsely-quantized signal, while avoiding the excess performance deterioration.**

I. INTRODUCTION

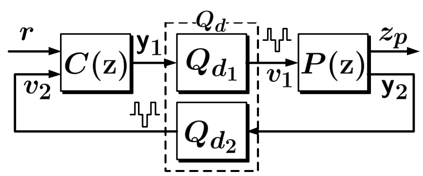
Recently, one of the most active control studies is the discrete-valued control theory in which networked systems, hybrid systems, embedded devices with D/A·A/D converters and ON/OFF actuators are addressed as systems containing discrete-valued signals. Since the theory leads to various practical applications, this topic has been studied from various perspectives [1]–[5] so far.



(a) Quantized feedback system.



(b) Usual feedback system.



(c) Quantized feedback system with I/O quantizers.

Fig. 1. Three control systems.

As a study focusing on optimality of systems controlled by the discrete-valued signals, it is well known that the optimal dynamic quantizer framework in [6]–[9] is useful. When a plant $P(\mathbf{z})$ and a controller $C(\mathbf{z})$ are given in the usual feedback system in Fig. 1 (b), the framework can provide a "dynamic" quantizer Q_d such that the quantized feedback system in Fig. 1 (a) "optimally" approximates the system in Fig. 1 (b) in the sense of the input-output relation. However, when the given systems have unstable zeros, the optimal dynamic quantizer in [6], [7] becomes unstable [8]. Although the numerical design method in [9] can provide a stable optimal dynamic quantizer, its infinite time control performance is not always guaranteed and the order of the obtained quantizer is basically (in some cases much) higher than that of the given system.

K. Sawada is with the Department of Mechanical Engineering and Intelligent Systems, The University of Electro-Communications, Chofu, Tokyo, 182-8585, Japan, sawada@se.uec.ac.jp

S. Shin is with the Department of Mechanical Engineering and Intelligent Systems, The University of Electro-Communications, Chofu, Tokyo, 182-8585, Japan, shin@se.uec.ac.jp

Motivated by the above, the authors have reconsidered the dynamic quantizer design. Our approach is based on the invariant set analysis [10]–[12] and the linear matrix inequality (LMI) technique [13]. The framework can synthesize the dynamic quantizers for SISO quantized feedback systems [14] in Fig. 1 (a) or the quantization intervals of the dynamic quantizers for feedforward-type networked systems with the communication rate constraints [15], while guaranteeing infinite quantization time and stability. However, in networked control, the quantized sensor information is often transmitted over communication channels as shown in Fig 1 (c). Also, it is often necessary to consider several performances (such as approximation performance, stability and signal coarseness) simultaneously due to the low capacity of the communication channels. Then it is important to consider more general dynamic quantizer design problems.

Therefore this paper deals with the generalized quantized feedback system in Fig. 2 (b) that includes the various systems such as Fig. 1 (a) and (c). We discuss the effectiveness and the limitation of our quantizer synthesis condition considering approximation performance, stability and signal coarseness. If the all transmission zeros of the given systems are stable, the proposed quantizer gives an optimal output approximation property. Otherwise, our method can design a stable suboptimal quantizer such that the order of the quantizer is exactly the same as that of the given system and the infinite time control performance is always guaranteed. In addition, this paper clarifies a performance relation between the static quantizer and the dynamic quantizer under some circumstances. Finally, the numerical examples show that our quantizer can achieve the coarsely-quantized signal, while attenuating the excess performance deterioration.

Notation: The set of $n \times m$ (positive) real matrices is denoted by $\mathbb{R}^{n\times m}$ ($\mathbb{R}_+^{n\times m}$). The set of $n \times m$ (positive) integer matrices is denoted by $\mathbb{N}^{n\times m}$ ($\mathbb{N}_+^{n\times m}$). $0_{n\times m}$ and I_m (or for simplicity of notation, 0 and I) denote the $n \times m$

zero matrix and the $m \times m$ identity matrix, respectively. $\lfloor a \rfloor$ denotes the floor of $a \in \mathbb{R}_+$. For a matrix M, M^{T}, $\rho(M)$, $\sigma_{\max}(M)$ and $\mathrm{abs}(M)$ denote its transpose, its spectrum radius, its maximum singular value and the matrix composed of the absolute values of its elements, respectively. For a vector x, x_i is the i-th entry of x. For a symmetric matrix X, $X > 0$ ($X \geq 0$) means that X is positive (semi) definite. For a full row rank matrix M, $M^{\dagger}$ denotes its pseudo inverse matrix which is given by $M^{\dagger} = M^{\mathrm{T}}(MM^{\mathrm{T}})^{-1}$. For a matrix X, $\|X\|_2$ denotes its 2-norm. For a vector x and a sequence of vectors $X := \{x_1, x_2, ...\}$, $\|x\|$ and $\|X\|$ denote their ∞-norms, respectively. Finally, we use the "packed" notation for transfer functions: $\left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right) := C(\mathbf{z}I - A)^{-1}B + D$.

II. PRELIMINARIES

Consider the linear time invariant (LTI) discrete-time system given by

$$\xi(k+1) = \mathcal{A}\xi(k) + \mathcal{B}w(k) \tag{1}$$

where $\xi \in \mathbb{R}^n$ and $w \in \mathbb{R}^m$ denote the state vector and disturbance input, respectively. We define the invariant set.

Definition 1: Define the invariant set of the system (1) to be a set $\mathcal{X}$ which satisfies

$$\xi \in \mathcal{X}, \quad w \in \mathcal{W} \quad \Rightarrow \quad \mathcal{A}\xi + \mathcal{B}w \in \mathcal{X} \tag{2}$$

where $\mathcal{W} := \{w \in \mathbb{R}^m : w^{\mathrm{T}}w \leq 1\}$.

The analysis condition can be expressed in terms of matrix inequalities as summarized in the following proposition [11].

Proposition 1: Consider the system (1). For a matrix $0 < \mathcal{P} \in \mathbb{R}^{n \times n}$, the ellipsoid $\mathbb{E}(\mathcal{P}) := \left\{\xi \in \mathbb{R}^n : \xi^{\mathrm{T}}\mathcal{P}\xi \leq 1\right\}$ is an invariant set if and only if there exists a scalar $\alpha \in [0, 1 - \rho(\mathcal{A})^2]$ satisfying

$$\begin{bmatrix} \mathcal{A}^{\mathrm{T}}\mathcal{P}\mathcal{A} - (1-\alpha)\mathcal{P} & \mathcal{A}^{\mathrm{T}}\mathcal{P}\mathcal{B} \\ \mathcal{B}^{\mathrm{T}}\mathcal{P}\mathcal{A} & \mathcal{B}^{\mathrm{T}}\mathcal{P}\mathcal{B} - \alpha I_m \end{bmatrix} \leq 0. \tag{3}$$

The necessity of (3) has been proved in [11]. The all ellipsoidal invariant sets are parameterized by Proposition 1. Also, the ellipsoidal invariant set allows us to approximate the reachable set from outside since the former covers the latter. We consider a criterion for the approximation of $\mathbb{E}(\mathcal{P})$ to the reachable set. Since the matrix $\mathcal{P}$ determines the ellipsoid, we denote all the criterion for the above approximation by $f(\mathcal{P})$ similar to [11]. $f(\mathcal{P})$ has the monotonical decreasingness in the sense that its value for the set of inside is less than that of outside. When α is fixed in (3), reference [11] clarifies that the infimum of $f(\mathcal{P})$ does not change even if $\mathcal{P}$ is restricted to $\mathcal{P}(\alpha)$ given by

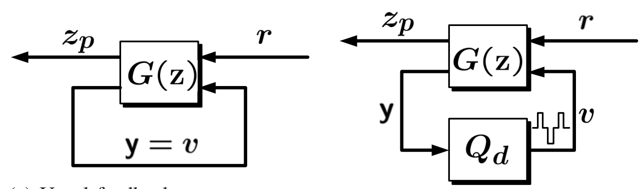
$$\mathcal{P}(\alpha)^{-1} = \sum_{k=0}^{\infty} \frac{1}{\alpha(1-\alpha)^k}\mathcal{A}^k\mathcal{B}\mathcal{B}^{\mathrm{T}}(\mathcal{A}^{\mathrm{T}})^k \tag{4}$$

where $\alpha \in (0, 1 - \rho(\mathcal{A})^2)$. Therefore, the invariant sets in (3) can be parameterized by $\alpha \in (0, 1-\rho(\mathcal{A})^2)$. Also, $f(\mathcal{P})$ can be replaced by $f(\mathcal{P}(\alpha))$. Denote by $\xi(k, \xi(0), w)$ the state trajectory of the system (1) at the k-th time. For the set $\mathbb{E}(\mathcal{P})$ characterized by Proposition 1, the property

$$\lim_{k \to \infty} \inf_{\xi \in \mathbb{E}(\mathcal{P})} \|\xi(k, \xi(0), w) - \xi\| = 0 \tag{5}$$

also holds (the proof is given by [12]).

III. PROBLEM FORMULATION



(a) Usual feedback system.

(b) Quantized feedback system.

Fig. 2. Generalized quantized and unquantized systems.

Consider the quantized feedback system depicted in Fig. 2 (b), which consists of the LTI discrete-time generalized plant $G(\mathbf{z})$ and the dynamic quantizer Q_d. The system $G(\mathbf{z})$ is represented by

$$\begin{bmatrix} x(k+1) \\ z_p(k) \\ \mathsf{y}(k) \end{bmatrix} = \begin{bmatrix} A & B_1 & B_2 \\ C_1 & D_{11} & 0 \\ C_2 & D_{21} & D_{22} \end{bmatrix} \begin{bmatrix} x(k) \\ r(k) \\ v(k) \end{bmatrix} \tag{6}$$

where $x \in \mathbb{R}^{n_g}$, $z_p \in \mathbb{R}^q$, $r \in \mathbb{R}^p$, $v \in \mathbb{R}^m$ and $\mathsf{y} \in \mathbb{R}^m$ denote the state vector, the controlled output, the exogenous input, the measured input and output, respectively. Considering the relation $\mathsf{y} = v$, we assume that the matrix $A + B_2(I - D_{22})^{-1}C_2$ is stable in the discrete domain, that is, the usual feedback system in Fig. 2 (a) is stable.

The formulation in (6) covers the various systems. Consider the LTI plant $P(\mathbf{z})$ with the state $x_p \in \mathbb{R}^{n_p}$ and the LTI controller $C(\mathbf{z})$ with the state $x_c \in \mathbb{R}^{n_c}$ given by

$$\begin{bmatrix} z_p \\ \mathsf{y}_2 \end{bmatrix} = \left(\begin{array}{c|c} A_p & B_p \\ \hline C_{p1} & 0 \\ C_{p2} & 0 \end{array}\right) v_1, \quad \mathsf{y}_1 = \left(\begin{array}{c|cc} A_c & B_{c1} & B_{c2} \\ \hline C_c & D_{c1} & D_{c2} \end{array}\right) \begin{bmatrix} v_2 \\ r \end{bmatrix}$$

where $\mathsf{y}_1 \in \mathbb{R}^{m_1}$, $\mathsf{y}_2 \in \mathbb{R}^{m_2}$, $v_1 \in \mathbb{R}^{m_1}$, and $v_2 \in \mathbb{R}^{m_2}$ denote the controller output, the plant measured output, the plant input, and the controller input, respectively. For example, by defining vectors: $x := [\ x_p^{\mathrm{T}} \ \ x_c^{\mathrm{T}}\ ]^{\mathrm{T}} \in \mathbb{R}^{n_g}$ ($n_g := n_p + n_c$), $\mathsf{y} := [\ \mathsf{y}_1^{\mathrm{T}} \ \ \mathsf{y}_2^{\mathrm{T}}\ ]^{\mathrm{T}}$, $v := [\ v_1^{\mathrm{T}} \ \ v_2^{\mathrm{T}}\ ]^{\mathrm{T}}$, and matrices:

$$A := \begin{bmatrix} A_p & 0 \\ 0 & A_c \end{bmatrix}, \quad B_1 := \begin{bmatrix} 0 \\ B_{c2} \end{bmatrix}, \quad B_2 := \begin{bmatrix} B_p & 0 \\ 0 & B_{c1} \end{bmatrix},$$

$$C_1 := [\ C_{p1} \ \ 0\ ], \quad D_{11} := 0, \quad C_2 := \begin{bmatrix} 0 & C_c \\ C_{p2} & 0 \end{bmatrix},$$

$$D_{21} := \begin{bmatrix} D_{c2} \\ 0 \end{bmatrix}, \quad D_{22} := \begin{bmatrix} 0 & D_{c1} \\ 0 & 0 \end{bmatrix},$$

one gets the control system with I/O quantizers in Fig. 1 (c).

For the system, we consider the dynamic quantizer $v = Q_d(\mathsf{y})$ with the state vector $x_q \in \mathbb{R}^{n_q}$. The system Q_d consists of the static quantizer $Q_{st} : \mathbb{R}^m \to d\mathbb{N}^m$, i.e.,

$$v = Q_{st}(u), \quad u := u_q + \mathsf{y} \tag{7}$$

and the dynamic compensator $Q(\mathbf{z})$

$$\begin{bmatrix} x_q(k+1) \\ u_q(k) \end{bmatrix} = \begin{bmatrix} A_q & B_q \\ C_q & 0 \end{bmatrix} \begin{bmatrix} x_q(k) \\ e_q(k) \end{bmatrix} \tag{8}$$

where $e_q := v - \mathsf{y}$. Note that Q_{st} is of the nearest-neighbor type toward $-\infty$ with the quantization interval $d \in \mathbb{R}_+$ and the initial state is given by $x_q(0) = 0$ for the drift-free of

Q_d [6]–[9]. One such static quantizer is the midtread type quantizer in Fig. 3. Also, the quantizer Q_d is said to be stable if its matrix $A_q + B_qC_q$ is stable in the discrete domain [8].

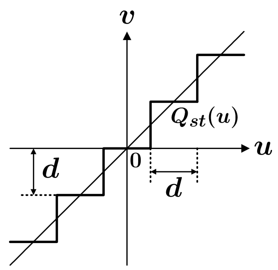


Fig. 3. Midtread type quantization.

We define the following matrices:

$$\begin{gathered}\mathsf{D} := (I - D_{22})^{-1}, \quad \mathsf{C}_2 := \mathsf{D}C_2, \quad \mathsf{D}_{21} := \mathsf{D}D_{21}, \\ \mathsf{D}_{22} := \mathsf{D}D_{22}, \quad \mathsf{A} := A + B_2\mathsf{C}_2, \quad \mathsf{B}_1 := B_1 + B_2\mathsf{D}_{21}, \\ \mathsf{B}_2 := B_2\mathsf{D}, \quad \mathcal{A} := \begin{bmatrix} \mathsf{A} & \mathsf{B}_2C_q \\ 0 & A_q + B_qC_q \end{bmatrix}, \quad \mathcal{B}_1 := \begin{bmatrix} \mathsf{B}_2 \\ B_q \end{bmatrix}, \\ \mathcal{B}_2 := \begin{bmatrix} \mathsf{B}_1 \\ 0 \end{bmatrix}, \quad \mathcal{C}_1 := \begin{bmatrix} \mathsf{C}_2 & \mathsf{D}C_q \end{bmatrix}, \quad \mathcal{D}_{11} := \mathsf{D}_{22}, \\ \mathcal{D}_{12} := \mathsf{D}_{21}, \quad \mathcal{C}_2 := \begin{bmatrix} C_1 & 0 \end{bmatrix}, \quad \mathcal{D}_{22} := D_{11}. \end{gathered} \tag{9}$$

For the system in Fig. 2 (b), the system $G(\mathbf{z})$ with the static quantizer Q_{st} seen by the linear compensator $Q(\mathbf{z})$ can be recast as the linear fractional transformation (LFT) of a generalized plant $\mathsf{G}(\mathbf{z})$:

$$\begin{bmatrix} x(k+1) \\ u(k) \\ z_p(k) \\ e_q(k) \end{bmatrix} = \begin{bmatrix} \mathsf{A} & \mathsf{B}_1 & \mathsf{B}_2 & \mathsf{B}_2 \\ \mathsf{C}_2 & \mathsf{D}_{21} & \mathsf{D}_{22} & \mathsf{D} \\ C_1 & D_{11} & 0 & 0 \\ 0 & 0 & I & I \end{bmatrix} \begin{bmatrix} x(k) \\ r(k) \\ e(k) \\ u_q(k) \end{bmatrix} \tag{10}$$

and the quantization error Q_e:

$$e = Q_e(u), \quad Q_e(u) := Q_{st}(u) - u \tag{11}$$

where the signal $e \in [-d/2, d/2]^m$. In this case, the control system in Fig. 2 (b) can be described as a LFT (Fig. 4) of the quantization error Q_e and a LTI system $H(\mathbf{z})$ represented by

$$e = Q_e(u), \quad \begin{bmatrix} \xi(k+1) \\ u(k) \\ z_p(k) \end{bmatrix} = \begin{bmatrix} \mathcal{A} & \mathcal{B}_1 & \mathcal{B}_2 \\ \mathcal{C}_1 & \mathcal{D}_{11} & \mathcal{D}_{12} \\ \mathcal{C}_2 & 0 & \mathcal{D}_{22} \end{bmatrix} \begin{bmatrix} \xi(k) \\ e(k) \\ r(k) \end{bmatrix} \tag{12}$$

where $\xi := [\ x^{\mathrm{T}} \ \ x_q^{\mathrm{T}}\ ]^{\mathrm{T}} \in \mathbb{R}^n (n := n_g + n_q)$ and the system $H(\mathbf{z})$ is the feedback connection of $\mathsf{G}(\mathbf{z})$ and $Q(\mathbf{z})$.

For the system in Fig. 2 (b) without the exogenous signal ($r(k) = 0\ \forall k$), $z_p(k, x_0)$ and $v(k, x_0)$ denote the outputs of z_p and v at the k-th time for the initial state $x_0 := x(0)$. In this case, this paper considers the following cost function:

$$L(Q_d) := \sup_{x_0 \in \mathbb{R}^{n_g}} \limsup_{k \to \infty} \|z_p(k, x_0)\|$$

and the constraint:

$$V(Q_d) := \sup_{x_0 \in \mathbb{R}^{n_g}} \limsup_{k \to \infty} \|v(k, x_0)\| \le \psi$$

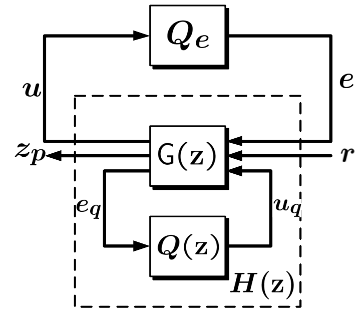


Fig. 4. Feedback system with quantization error.

where the constraint parameter $\psi \in \mathbb{R}_+$ is given.

For the sufficiently small $L(Q_d)$ case, it is expected that the quantizer minimizes the effect of the quantization error on the controlled output z_p in a neighborhood of the origin [14]. On the other hand, it is also important to attenuate the excess output gain of Q_d that the small $L(Q_d)$ leads to. Since $V(Q_d)$ specifies the output amplitude of Q_d around the origin, it is expected that the smaller value ψ leads to the coarser output of Q_d around the origin. Then our first objective is to solve the following dynamic quantizer synthesis problem **(L)**: *For the system (12) without the exogenous signal, suppose that the quantization interval $d \in \mathbb{R}_+$, the performance level $\gamma \in \mathbb{R}_+$ and the constraint parameter $\psi \in \mathbb{R}_+$ are given. Characterize a stable dynamic quantizer Q_d (i.e., find parameters (n_q, A_q, B_q, C_q)) achieving $L(Q_d) \le \gamma$ and $V(Q_d) \le \psi$ based on Proposition 1.*

Next, let $T \in \mathbb{N}_+ \cup \{\infty\}$ be the period over which we consider the quantizer performance. For the system in Fig. 2 (b) with the exogenous signal sequence $R_T := \{r(0), r(1), ..., r(T-1)\} \in l_\infty^p$, $z_p(k, x_0, R_T)$ denotes the output of z_p at the k-th time for the initial state x_0. Also, for the system in Fig. 2 (a) without the quantizer, $z_p^*(k, x_0, R_T)$ denotes its output at the k-th time for the initial state x_0. $Z_p(x_0, R_T)$ and $Z_p^*(x_0, R_T)$ denote the vector sequence of $z_p(k, x_0, R_T)$ and $z_p^*(k, x_0, R_T)$ for $k = 1, ..., T$, respectively. This paper considers the following cost function:

$$E_T(Q_d) := \sup_{(x_0, R_T) \in \mathbb{R}^{n_g} \times l_\infty^p} \|Z_p^*(x_0, R_T) - Z_p(x_0, R_T)\|$$

which is discussed in [6]–[9]. For the case $T = \infty$, note that we consider $R = (r(0), r(1), ...) \in l_\infty^p$ instead of $R_T \in l_\infty^p$. Along with this, $Z_p(x_0, R_T)$ and $Z_p^*(x_0, R_T)$ are replaced by $Z_p(x_0, R)$ and $Z_p^*(x_0, R)$, respectively.

Our second objective is to solve the following dynamic quantizer synthesis problem **(E)**: *For the system (12) with the exogenous signal sequence $R_T := \{r(0), r(1), ..., r(T-1)\} \in l_\infty^p$, suppose that the quantization period $T \in \mathbb{N}_+ \cup \{\infty\}$, the quantization interval $d \in \mathbb{R}_+$ and the performance level $\gamma \in \mathbb{R}_+$ are given. Characterize a stable dynamic quantizer Q_d (i.e., find parameters (n_q, A_q, B_q, C_q)) achieving $E_T(Q_d) \le \gamma$ based on Proposition 1.*

If the minimum value of γ is sufficiently small, the system in Fig. 2 (b) "optimally" approximates the usual system in Fig. 2 (a) in the sense of the input-output relation.

IV. Main result

A. Quantizer analysis

Suppose that the stable quantizer Q_d is given. For the set $\mathcal{E} := \{e \in \mathbb{R}^m : e \text{ satisfies } (11)\}$, the relation $\mathcal{E} \subseteq \frac{md^2}{4}\mathcal{W}$ clearly holds (namely, $\frac{md^2}{4}\mathcal{W}$ is the set $\mathcal{E}$ ignores the relation (11)). That is, the reachable set of $(2,1)$ block of $H(\mathbf{z})$ in (12) is no larger than that of $(2,1)$ block of $H(\mathbf{z})$ with an independent bounded disturbance $e \in \frac{md^2}{4}\mathcal{W}$. Considering the reachable set to estimate the influences of the quantization error, this paper utilizes the ellipsoidal invariant set which covers the reachable set from outside. Define

$$\mathcal{A} := \mathcal{A}, \quad \mathcal{B} := \mathcal{B}_1\frac{d\sqrt{m}}{2}, \quad \hat{\mathcal{D}}_{11} := \mathcal{D}_{11} + I_m. \tag{13}$$

In this case, the ellipsoidal invariant set $\mathbb{E}(\mathcal{P})$ for the system (12) can be estimated by Proposition 1. If there exists the set $\mathbb{E}(\mathcal{P})$, there exists a scalar $\gamma \in \mathbb{R}_+$ satisfying

$$\max_i \sup_{\xi\in\mathbb{E}(\mathcal{P})} |c_{2i}^{\mathrm{T}}\xi| = \gamma \quad \Leftarrow \quad \begin{bmatrix} \mathcal{P} & \mathcal{C}_2^{\mathrm{T}} \\ \mathcal{C}_2 & \gamma^2 I_q \end{bmatrix} \geq 0 \tag{14}$$

where c_{2i}^{T} is the i-th entry of $\mathcal{C}_2$ [10]. Furthermore, the following relation holds:

$$\max_i \sup_{\xi\in\mathbb{E}(\mathcal{P}),\ e\in\frac{md^2}{4}\mathcal{W}} |c_{1i}^{\mathrm{T}}\xi + \hat{d}_i^{\mathrm{T}}e| \leq \psi$$
$$\Leftarrow \begin{bmatrix} \mathcal{P} & {\mathcal{C}_1}^{\mathrm{T}} \\ \mathcal{C}_1 & (\psi - \frac{\sigma(\hat{\mathcal{D}}_{11})d\sqrt{m}}{2})^2 \end{bmatrix} \geq 0 \tag{15}$$

where c_{1i}^{T} and $\hat{d}_i^{\mathrm{T}}$ are the i-th entries of $\mathcal{C}_1$ and $\hat{\mathcal{D}}_{11}$, respectively. From the property (5) of the invariant set and the definition (11), for any initial state $\xi(0) \in \mathbb{R}^n$ (equivalently, $x_0 \in \mathbb{R}^{n_g}$), the performance level γ in (14) satisfies

$$L(Q_d) \leq \gamma \tag{16}$$

and the scalar ψ in (15) satisfies

$$V(Q_d) \leq \psi. \tag{17}$$

For Proposition 1 with (13) and the given ψ, we have the optimization problem **(Aop)**:

$$\min_{\mathcal{P}>0,1-\rho(\mathcal{A})^2>\alpha>0,\gamma>0} \gamma \quad \text{s.t.} \quad (3),\ (14) \text{ and } (15).$$

Focusing on the left side of (14), we see that γ is corresponding to the value of $f(\mathcal{P}(\alpha))$ in terms of the approximation of $\mathbb{E}(\mathcal{P}(\alpha))$ to the reachable set. From the parameterization $\mathcal{P}(\alpha)$ in (4), the infimum of γ can be expressed by the following lemma.

Lemma 1: For the feedback system (12), suppose that the quantization interval $d \in \mathbb{R}_+$ is given. Consider the relaxed problem **(Aop')** *in the sense that the constraint (15) is removed from the original problem* **(Aop)** *with (13). The infimum of γ is given by*

$$\inf \gamma = \inf_\alpha \frac{d\sqrt{m}}{2\sqrt{\alpha}}\sigma_\alpha, \quad 0 < \alpha < 1 - \rho(\mathcal{A})^2, \tag{18}$$
$$\sigma_\alpha := \sigma_{\max}\left(\textstyle\sum_{k=0}^{\infty} \frac{1}{(1-\alpha)^k}\mathcal{C}_2\mathcal{A}^k\mathcal{B}_1\mathcal{B}_1^{\mathrm{T}}(\mathcal{A}^{\mathrm{T}})^k\mathcal{C}_2^{\mathrm{T}}\right).$$

Lemma 1 points out that Proposition 1 characterizes the infimum of the upper bound of the cost function $L(Q_d)$.

B. Quantizer synthesis

The problem **(Aop)** with (13) suggests that the quantizer synthesis problem **(L)** reduces to the search for the quantizer parameters satisfying condition (3) in Proposition 1, (14) and (15) as summarized in the following theorem.

Theorem 1: For the feedback system (12), suppose that the quantization interval $d \in \mathbb{R}_+$, the performance level $\gamma \in \mathbb{R}_+$ and the constraint parameter $\psi \in \mathbb{R}_+$ are given. For a scalar $\alpha \in (0,1)$, there exists a stable dynamic quantizer Q_d achieving (16) and (17) if one of the following equivalent statements holds.
(i) There exist a matrix $0 < \mathcal{P} \in \mathbb{R}^{n\times n}$ and a dynamic quantizer Q_d satisfying (3), (14) and (15).
(ii) There exist matrices $0 < X \in \mathbb{R}^{n_g\times n_g}$, $0 < Y \in \mathbb{R}^{n_g\times n_g}$, $F \in \mathbb{R}^{m\times n_g}$, $W \in \mathbb{R}^{n_g\times n_g}$, and $U \in \mathbb{R}^{n_g\times m}$ satisfying

$$\begin{bmatrix} (1-\alpha)\Xi_{\mathsf{P}} & 0 & \Xi_{\mathsf{A}}^{\mathrm{T}} \\ 0 & \frac{4\alpha}{md^2}I_m & \Xi_{\mathsf{B}}^{\mathrm{T}} \\ \Xi_{\mathsf{A}} & \Xi_{\mathsf{B}} & \Xi_{\mathsf{P}} \end{bmatrix} \geq 0, \quad \begin{bmatrix} \Xi_{\mathsf{P}} & \Xi_{\mathsf{C}_1}^{\mathrm{T}} \\ \Xi_{\mathsf{C}_1} & \gamma^2 I_q \end{bmatrix} \geq 0, \tag{19}$$

$$\begin{bmatrix} \Xi_{\mathsf{P}} & \Xi_{\mathsf{C}_2}^{\mathrm{T}} \\ \Xi_{\mathsf{C}_2} & (\psi-\phi)^2 \end{bmatrix} \geq 0 \tag{20}$$

where

$$\Xi_{\mathsf{P}} := \begin{bmatrix} X & I \\ I & Y \end{bmatrix}, \qquad \Xi_{\mathsf{A}} := \begin{bmatrix} X\mathsf{A} & W \\ \mathsf{A} & \mathsf{A}Y + \mathsf{B}_2F \end{bmatrix},$$
$$\Xi_{\mathsf{B}} := \begin{bmatrix} U^{\mathrm{T}} & \mathsf{B}_2^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}, \quad \Xi_{\mathsf{C}_1} := \begin{bmatrix} C_1 & C_1Y \end{bmatrix},$$
$$\Xi_{\mathsf{C}_2} := \begin{bmatrix} \mathsf{C}_2 & \mathsf{C}_2Y + \mathsf{D}F \end{bmatrix}, \quad \phi := \frac{\sigma(\mathsf{D}_{22}+I)d\sqrt{m}}{2}.$$

In this case, one such quantizer is given by

$$B_q = Z^{-1}(U - X\mathsf{B}_2), \quad C_q = -FY^{-1}, \quad n_q = n_g,$$
$$A_q = Z^{-1}(X\mathsf{A}Y + UF - W)Y^{-1} \tag{21}$$

where $Z = X - Y^{-1}$.

For the case where the quantizer is clearly inactive ($C_2 = 0$, $D_{21} = 0$ and $D_{22} = 0$ in (6)), the statement of Theorem 1 still holds since $L(Q_d)$ and $V(Q_d)$ of such a system become zero and then the both of (16) and (17) hold. As a synthesis problem minimizing γ of (16) and achieving the constraint (17), we have the optimization problem **(Sop)**:

$$\min_{X>0,Y>0,F,W,U,1>\alpha>0,\gamma>0} \gamma \quad \text{s.t.} \quad (19) \text{ and } (20).$$

In synthesis, the parameters (A_q, B_q, C_q) to be designed lead to $\alpha \in (0,1)$. When scalar α is fixed, the conditions in Theorem 1 are linear matrix inequalities (LMIs) in terms of the other variables. Using standard LMI software in combination with the line search of α for **(Sop)**, we can obtain a stable dynamic quantizer, numerically.

Denote by the synthesis problem **(L')** the relaxed problem in the sense that the constraint (17) is removed from the original synthesis problem **(L)**. Under some circumstances, Proposition 1 gives a closed-form solution to **(L')**.

Theorem 2: Consider the following non-convex optimization problem **(OP)** *with (13):*

$$\min_{\mathcal{P}>0,A_q,B_q,C_q,0>\alpha>1-\rho(\mathcal{A})^2,\gamma>0} \gamma \quad \text{s.t.} \quad (3) \text{ and } (14).$$

Suppose that the matrix $C_1\mathsf{A}^\tau B_2$ is full row rank for the smallest integer $\tau \in \{0\}\cup\mathbb{N}_+$ satisfying $C_1\mathsf{A}^\tau\mathsf{B}_2 \neq 0$ for the system (6). An optimal solution of (A_q, B_q, C_q) and its *infimum of $\gamma \in \mathbb{R}_+$ to the problem* **(OP)** *are given by*

$$A_q = \mathsf{A}, \quad B_q = \mathsf{B}_2, \quad C_q = -(C_1\mathsf{A}^\tau \mathsf{B}_2)^\dagger C_1 \mathsf{A}^{\tau+1} \quad (22)$$

and

$$\inf \gamma = (d\sqrt{m}\|C_1\mathsf{A}^\tau \mathsf{B}_2\|_2)/(2\rho(\mathcal{A})^\tau \sqrt{1-\rho(\mathcal{A})^2}) \quad (23)$$

if the matrix $A_q + B_q C_q$ defined in (22) is stable.

In the case of $m = p$, $(C_1\mathsf{A}^\tau \mathsf{B}_2)^\dagger$ becomes $(C_1\mathsf{A}^\tau \mathsf{B}_2)^{-1}$. In this case, the stable $A_q + B_q C_q$ in (22) implies that the all transmission zeros of the system $G(\mathbf{z})$ are stable [8]. If $G(\mathbf{z})$ is minimum phase and the matrix $\mathcal{P}(\alpha)$ defined by (4), (9), (13) and (22) satisfies (15) for the given ψ, that is, Theorem 2 gives a closed-form solution to **(L)**.

Also, the following corollary provides an analytical relation between the quantizers Q_d and Q_{st}.

Corollary 1: Consider the relaxed problem **(L')** *and denote by γ_{st} and γ_d upper bounds of $L(Q_{st})$ and $L(Q_d)$, respectively. In the case of $n_g = 1$,*

$$\inf \frac{\gamma_{st}}{\gamma_d} = \frac{1}{\sqrt{1-\mathsf{A}^2}} \quad (24)$$

holds.

Since the matrix A is stable, $\mathsf{A}^2 \in [0, 1)$ holds in (24). Corollary 1 guarantees that the quantizer Q_d improves $L(Q_d)$ compared with the quantizer Q_{st} in terms of the infimum of the upper bound ratio of the cost functions.

C. Relation to the optimal dynamic quantizer

The usual feedback system in Fig. 2 (a) is given by

$$\begin{bmatrix} x^*(k+1) \\ \mathsf{y}^*(k) \\ z_p^*(k) \end{bmatrix} = \begin{bmatrix} \mathsf{A} & \mathsf{B}_1 \\ \mathsf{C}_2 & \mathsf{D}_{21} \\ C_1 & D_{11} \end{bmatrix} \begin{bmatrix} x^*(k) \\ r(k) \end{bmatrix}$$

where $x^* \in \mathbb{R}^{n_g}$, $\mathsf{y}^* \in \mathbb{R}^m$, $z_p^* \in \mathbb{R}^q$ denote its state vector, measured output, and controlled output respectively, and $x^*(0) = x(0)$. Define the signals as follows:

$$\xi := [\ x^\mathsf{T} - x^{*\mathsf{T}} \quad x_q^\mathsf{T}\ ]^\mathsf{T}, \quad \nu := v - \mathsf{y}^*, \quad z := z_p - z_p^*.$$

The difference between $z_p^*(k, x_0, R_T)$ and $z_p(k, x_0, R_T)$ is generated by the following error system $\mathcal{H}(\mathbf{z})$:

$$\begin{bmatrix} \xi(k+1) \\ \nu(k) \\ z(k) \end{bmatrix} = \begin{bmatrix} \mathcal{A} & \mathcal{B}_1 \\ \mathcal{C}_1 & \hat{\mathcal{D}}_{11} \\ \mathcal{C}_2 & 0 \end{bmatrix} \begin{bmatrix} \xi(k) \\ e(k) \end{bmatrix}, \quad \xi(0) = 0$$

where $e \in \mathbb{R}^m$ is given by (11) and the matrices $\mathcal{A}$, $\mathcal{B}_1$, $\mathcal{C}_1$, $\mathcal{C}_2$, $\hat{\mathcal{D}}_{11}$ are defined by (9) and (13). Therefore, by using the same procedure of the subsection IV-A, Proposition 1 characterizes the cost function $E_\infty(Q_d)$ as follows:

$$E_\infty(Q_d) \le \gamma,\ \gamma := \min_{\mathcal{P},\alpha,A_q,B_q,C_q} \max_i \sup_{\xi \in \mathbb{E}(\mathcal{P})} |c_{2i}^\mathsf{T}\xi| \text{ s.t. } (3).$$

The quantization period T is ∞ within the invariant set analysis framework. From Theorems 1 and 2, solutions to the synthesis problem **(E)** is given by the following theorem.

Theorem 3: For the feedback system (12), suppose that the quantization interval $d \in \mathbb{R}_+$ is given and $m = p$. If the all transmission zeros of $G(\mathbf{z})$ in (6) are stable, a stable dynamic quantizer achieving $E_\infty(Q_d) \le \gamma$ and its performance are given by (22) and (23). If $G(\mathbf{z})$ has unstable zeros, a stable dynamic quantizer achieving $E_\infty(Q_d) \le \gamma$ and its performance are characterized by (19) and (21).

When the matrix C_2 is full row rank, reference [6] has present an optimal dynamic quantizer Q_d^{op} given by (22) with

$$E_\infty(Q_d^{op}) = \|\mathrm{abs}(C_1\mathsf{A}^\tau \mathsf{B}_2)\| \frac{d}{2}$$

where $\tau \in \{0\} \cup \mathbb{N}_+$ is the minimum integer satisfying $C_1\mathsf{A}^\tau \mathsf{B}_2 \neq 0$. It is striking that the structure of their quantizer is equivalent to our proposed one based on Proposition 1 even if the former performance evaluation is less conservative than the latter one. That is, Theorem 3 points out that *the proposed quantizer is also optimal in the sense that the quantizer gives an optimal output approximation property.*

When the quantization period $T \in \mathbb{N}_+$ is given, reference [9] has provided the numerical design method in which the stable and optimal quantizer synthesis problem is recast as the following optimization problem:

$$\min_{A_q,B_q,C_q,\mathsf{P}} \left\| \sum_{k=0}^{T-1} \mathrm{abs}(\mathcal{C}_2\mathcal{A}^k\mathcal{B}_1) \right\| \frac{d}{2} \,(= E_T(Q_d))$$
$$\text{s.t.} \quad (A_q + B_q C_q)^\mathsf{T}\mathsf{P}(A_q + B_q C_q) < \mathsf{P}, \quad \mathsf{P} > 0.$$

The order of the quantizer in [9] is given by $\lfloor T/2 \rfloor + 1$ or $(\lfloor T/2 \rfloor + 1)T$. When T is set to be large, then we need the reduction technique. On the other hand, our method has the following properties: *When the generalized plant $G(\mathbf{z})$ has unstable zeros, (i) the order is $n_q = n_g$, (ii) the infinite time control performance is always guaranteed, and (iii) the method provides a suboptimal dynamic quantizer in the sense that the upper bound of $E_\infty(Q_d)$ is minimized.*

The quantizer satisfying properties $(i)\backsim(iii)$ is obtained from the optimization problem **(Sop')**:

$$\min_{X>0,Y>0,F,W,U,1>\alpha>0,\gamma>0} \gamma \quad \text{s.t.} \quad (19).$$

On the other hand, the quantizer Q_d considering its approximation performance $E_\infty(Q_d)$ and its signal coarseness can be obtained from **(Sop)**. By denoting by $\nu(k, x_0, R)$ the output of ν at the k-th time for x_0 and R, (20) leads to

$$\sup_{(k,x_0,R)\in\mathbb{N}_+\times\mathbb{R}^n\times l_\infty^p} \|\nu(k, x_0, R)\| \le \psi.$$

That is, it is expected that the constraint (20) can adjust the output signal coarseness of the dynamic quantizer Q_d for $E_\infty(Q_d)$. Next section examines the above.

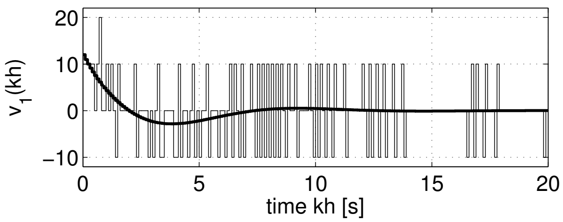
V. Numerical example

Consider the system in Fig. 1 (c). The plant $P(\mathbf{z})$ is the discretized system of the unstable non-minimum phase continuous-time LTI system

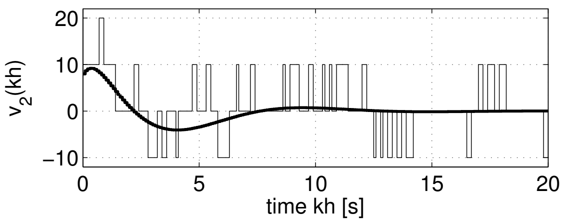
$$\begin{bmatrix} \dot{x}_p(t) \\ \mathsf{y}_2(t) \end{bmatrix} = \left[\begin{array}{cc:c} 0 & 1 & 0 \\ 1 & -1 & 1 \\ \hdashline -2 & 1 & 0 \end{array}\right] \begin{bmatrix} x_p(t) \\ v_1(t) \end{bmatrix}, \quad z_p(t) = \mathsf{y}_2(t)$$

with the sampling time $h = 0.1$ and zero-order hold. Its eigenvalues are $\{1.064, 0.857\}$ and its unstable zero is $\{1.224\}$. The stabilizing controller $C(\mathbf{z})$ is given by

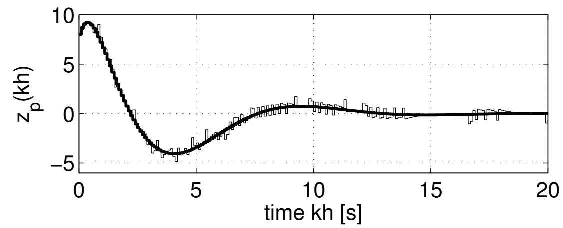
$$\begin{bmatrix} x_c(k+1) \\ \mathsf{y}_1(k) \end{bmatrix} = \left[\begin{array}{c:c} 0.741 & 0.086 \\ \hdashline -1 & 1 \end{array}\right] \begin{bmatrix} x_c(k) \\ v_2(k) \end{bmatrix}.$$



(a) v_1 (the input of $P(\mathbf{z})$).

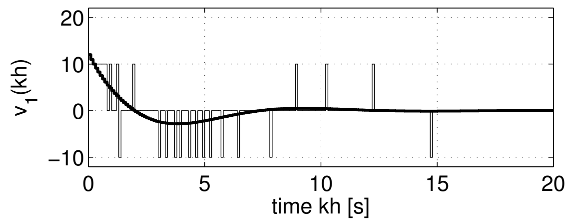


(b) v_2 (the input of $C(\mathbf{z})$).

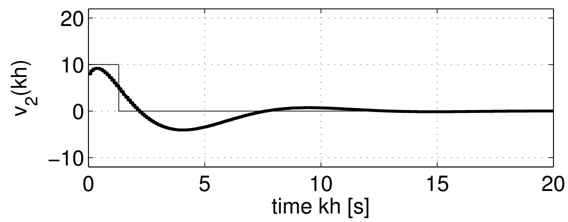


(c) z_p (the controlled output of $P(\mathbf{z})$).

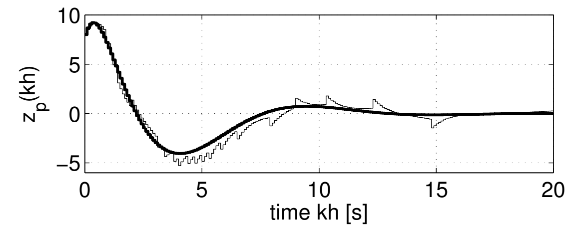
Fig. 5. Time responses of for **(Sop')**.



(a) v_1 (the input of $P(\mathbf{z})$).



(b) v_2 (the input of $C(\mathbf{z})$).



(c) z_p (the controlled output of $P(\mathbf{z})$).

Fig. 6. Time responses for **(Sop)**.

Consider $d = 10$ and $\psi = 45$ for the above systems ($\phi = 18.1$). The stable suboptimal quantizers with $\gamma = 2.09$ and $\gamma = 2.35$ are obtained from **(Sop')** and **(Sop)**, respectively.

Figures 5 and 6 illustrate the time responses of $v_1(kh)$, $v_2(kh)$ and $z_p(kh)$ for the quantizers via **(Sop')** and **(Sop)** with the initial state $x(0) = [\ -4\ \ 0\ \ -4\ ]^{\mathrm{T}}$, respectively. In Figs. 5 and 6, the thin lines and the thick lines illustrate the time responses of the usual feedback system in Fig. 1 (b) and the quantized feedback system in Fig. 1 (c), respectively. We see that the controlled output of Fig. 1 (c) approximates that of Fig. 1 (b) even if the discrete-valued signals v_1 and v_2 are applied. Especially, the quantizer of **(Sop)** can achieve the coarser discrete-valued signals such as $v_1 \in \{-10, 0, 10\}$ and $v_2 \in \{0, 10\}$ in Fig. 6 (a) and (b) than $v_1 \in \{-10, 0, 10, 20\}$ and $v_2 \in \{-10, 0, 10, 20\}$ in Fig. 5 (a) and (b), while attenuating the excess performance deterioration between both quantizers. Then this numerical example shows the effectiveness of our method and examines the validity of Theorem 3 in dynamic quantizer synthesis.

REFERENCES

[1] P. Antsaklis and J. Baillieul (eds.), "Special issue on networked control systems", *IEEE Trans. Auto. Contr.*, Vol. 49, No. 9, pp. 1421–1423, 2004.

[2] P. J. Antsaklis and J. Baillieul, Guest Editors, "Special Issue on the Technology of Networked Control Systems", *Proc. IEEE*, Vol. 95, No. 1, 2007.

[3] W. S. Wong and R. W. Brockett, "Systems with finite communication bandwidth constraints-Part II: Stabilization with limited information feedback", *IEEE Trans. on Auto. Contr.*, Vol. 44, No. 5, pp. 1049–1053, 1999.

[4] N. Elia and S. K. Mitter, "Stabilization of linear systems with limited information", *IEEE Trans. Auto. Contr.*, Vol. 46, No. 9, pp. 1384–1400, 2001.

[5] M. Fu and L. Xie, "The sector bound approach to quantized feedback control", *IEEE Trans. Auto. Contr.*, Vol. 50, No. 11, pp. 1698–1711, 2005.

[6] Y. Minami, S. Azuma and T. Sugie, "An optimal dynamic quantizer for feedback control with discrete-valued signal constraints", *Proc. of 46th IEEE CDC*, pp. 2259–2264, 2007.

[7] S. Azuma and T. Sugie, "Optimal dynamic quantizer for discrete-valued input control", *Automatica*, Vol. 44, No. 12, pp. 396–406, 2008.

[8] S. Azuma and T. Sugie, "Stability analysis of quantized feedback systems including optimal dynamic quantizers", *Proc. of 47th IEEE CDC*, pp. 3392–3397, 2008.

[9] S. Azuma and T. Sugie, "Synthesis of optimal dynamic quantizers for discrete-valued input control", *IEEE Trans. Auto. Contr.*, Vol. 53, No. 9, pp. 2064–2075, 2008.

[10] R. Sanchez-Peña and M. Sznaier, "Robust systems theory and applications", *John Wiley & Sons*, 1998.

[11] H. Shingin and Y. Ohta, "Optimal invariant sets for discrete-time systems: Approximation of reachable sets for bounded inputs", *Proc of 10th IFAC/IFORS/IMACS/IFIP Symposium on Large Scale Systems, Theory and Applications (LSS)*, pp. 401–406, 2004.

[12] T. Hu, "Nonlinear control design for linear differential inclusions via convex hull of quadratics", *Automatica*, Vol. 43, No. 4, pp. 685–692, 2007.

[13] S. Boyd, L. E. Ghaoui, E. Feron, and V. Balakrishnan, "Linear matrix inequalities in system and control theory", *SIAM*, 1994.

[14] K. Sawada and S. Shin, "Dynamic quantizer synthesis based on invariant set analysis for SISO systems with discrete-valued input", *Proc. of MTNS 2010*, pp. 1385–1390, 2010.

[15] H. Okajima, N. Matsunaga, and K. Sawada, "Optimal quantization interval design of dynamic quantizers which satisfy the communication rate constraints", *Proc. of 49th IEEE CDC*, pp. 4733–4739, 2010.