REFERENCES

[1] V. Chellaboina, W. M. Haddad, D. S. Bernstein, and D. A. Wilson, "Induced convolution operator norms of linear dynamical systems," *Math. Control, Signals, Syst.*, vol. 13, pp. 216–239, 2000.

[2] I. J. Fialho and T. T. Georgiou, "A variational approach to $\mathcal{L}_\infty$-gain analysis of nonlinear systems," in *Proc. Conf. Decision Control*, 1995, pp. 823–828.

[3] M. R. James and S. Yuliar, "Numerical approximation of the $\mathcal{H}_\infty$ norm for nonlinear systems," *Automatica*, vol. 31, no. 8, pp. 1075–1086, 1995.

[4] H. K. Khalil, *Nonlinear Systems*, 3rd ed. Upper Saddle River, NJ: Prentice Hall, 2002.

[5] H. J. Marquez, *Nonlinear Control Systems, Analysis and Design*. New York: Wiley, 2003.

[6] *"Multi-Tank System User's Manual"* INTECO, Ltd. [Online]. Available: www.inteco.com.pl

[7] B. G. Romanchuk, "On the computation of the induced $\mathcal{L}_2$ norm of single-input linear systems with saturation," *IEEE Trans. Automat. Control*, vol. 43, no. 2, pp. 262–268, Feb. 1998.

[8] V. Zahedzadeh, H. J. Marquez, and T. Chen, "On the robust stability of unforced nonlinear systems," in *Proc. 45th IEEE Conf. Decision Control*, San Diego, CA, Dec. 2006, pp. 343–348.

Finite-Level Quantized Feedback Control for Linear Systems

Minyue Fu and Lihua Xie

***Abstract*—This technical note studies quantized output feedback control of discrete-time linear systems using a finite-level quantizer. The main objective is to find a quantization strategy which is easily implementable and achieves asymptotic stabilization. Based on a known logarithmic quantization scheme, we introduce a simple dynamic scaling method for the quantizer. A suboptimal approach for the optimization of the number of quantization levels and the design of a corresponding quantized dynamic output feedback controller is given. The robustness of the dynamic quantization scheme with respect to input disturbances is also examined.**

***Index Terms*—Linear time-invariant (LTI), single-input single-output (SISO).**

I. INTRODUCTION

There has been a lot of new interest in quantized feedback control where the feedback signal is quantized and coded for transmission [4], [9]. A fundamental problem is how to design a feedback controller and a quantizer jointly in order to achieve a given control objective.

The research on quantized feedback control can be categorized depending on whether the quantizer is static or dynamic. A *static quantizer* is a memoryless nonlinear function, whereas a *dynamic quantizer* uses memory and thus can be much more complex and potentially more powerful. Existing work using static quantizers includes, e.g., [1]–[3]. For quadratic stabilization of a linear system using state feedback, it is shown in [1] that the coarsest quantizer is logarithmic. This result is generalized in [3] to a number of output feedback problems using a sector bound approach.

When a dynamic quantizer is allowed, it is shown in [4] that stabilization of a single-input single-output (SISO) linear time-invariant (LTI) system (in some stochastic sense) can be achieved using only a finite number of quantization levels, and the minimum number of quantization levels (also known as the minimum *feedback information rate*) is explicitly related to the unstable poles of the system. In this setting, the dynamic quantizer effectively consists of two parts: an *encoder* at the output end and a *decoder* at the input end. Although it is shown in [4] that stabilization of a linear system can be achieved by feeding back only a finite number of bits per sample and this number is typically very small, the encoder-decoder pair used for proving this result is impractical and could be non-robust against noises.

Another type of dynamic quantizers uses *dynamic scaling* in conjunction with a static quantizer. That is, the input signal is pre-scaled so that its range is more suitable for quantization. The scaling parameter is dynamically adjusted (i.e., adjusted online). Noticeable work along this line includes [5]–[9]. In [5], it is pointed out that if a system is not excessively unstable, by employing a quantizer with various sensitivity a feedback strategy can be designed to bring the closed-loop state arbitrarily close to zero for an arbitrarily long time. This idea is extended in [6] where it is shown that there exists a dynamic adjustment of the quantizer sensitivity and a quantized state feedback that asymptotically stabilizes the system. In the case of output feedback, a local (or semi-global) stabilization result is obtained. [7] studies the destabilizing effect of a quantizer on closed-loop system stability. Given a stabilizing state feedback controller designed under no quantization, the state of the closed-loop system is shown to enter a set characterized by the worst-case quantization error. The quantizer design problem is then reduced to a so-called multicenter problem.

In this technical note, we propose a simple dynamic scaling method for a logarithmic quantizer based output feedback controller. A dynamic *scaling factor* is simply adjusted up or down depending whether the input signal to the quantizer is "too small" or "too large" in magnitude. Using this dynamic scaling method, we show that a linear system can be asymptotically stabilized using a logarithmic quantizer with only a finite number of quantization bits. We also show that the proposed scheme is robust in the sense it can tolerate additive noises in the system effectively. Unlike [7], [8] where quantized state feedback is considered, we investigate a quantized output feedback control problem. In our work, we give a suboptimal approach for minimizing the number of quantization levels of a logarithmic quantizer and the design of a corresponding quantized output feedback controller for stabilization.

II. FINITE-LEVEL QUANTIZED FEEDBACK STABILIZATION

Consider the following system:

$$x_{k+1} = Ax_k + Bu_k \tag{1}$$

$$y_k = Cx_k \tag{2}$$

where $x_k \in \mathbb{R}^n$ is the state, $u_k \in \mathbb{R}$ is the control input, $y_k \in \mathbb{R}$ is the measured output, $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times 1}$ and $C \in \mathbb{R}^{1\times n}$ are given. We will denote the transfer function from u_k to y_k by $G(z)$. Without loss of generality, we assume that A is unstable and (A, B, C) is a minimal realization.

The quantized output feedback stabilization problem is to design the quantizer, $v_k = Q(y_k)$, and a feedback controller of the form

$$\hat{x}_{k+1} = A_c\hat{x}_k + B_c v_k, \quad \hat{x}_0 = 0 \tag{3}$$

$$u_k = C_c\hat{x}_k + D_c v_k \tag{4}$$

Manuscript received February 28, 2008. Current version published May 13, 2009. Recommended by Associate Editor D. Henrion.

M. Fu is with School of Electrical Engineering and Computer Science, University of Newcastle, Callaghan, N.S.W. 2308, Australia (e-mail: minyue.fu@newcastle.edu.au).

L. Xie is with School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore (e-mail: elhxie@ntu.edu.sg).

Digital Object Identifier 10.1109/TAC.2009.2017815

with $\hat{x}_k \in \mathbb{R}^n$, such that the closed-loop system is stable. It has been shown in [1], [3] that under quadratic stability the quantizer that achieves the coarsest quantization is logarithmic and is given by

$$Q(y) = \begin{cases} \rho^i \mu_0, & \text{if } \frac{1}{1+\delta}\rho^i\mu_0 < y \le \frac{1}{1-\delta}\rho^i\mu_0 \\ 0, & \text{if } y = 0 \\ -Q(-y), & \text{if } y < 0 \end{cases} \tag{5}$$

where $\rho \in (0,1)$, μ_0 is a scaling constant, $i = 0, \pm 1, \pm 2, \cdots$, and

$$\delta = \frac{1-\rho}{1+\rho}. \tag{6}$$

Note that a smaller ρ corresponds to a coarser quantization density and the smallest ρ for which the system (1), (2) can be quadratically stabilized via a quantized feedback controller (5), (3), (4) is given by [3]

$$\rho_{\inf} = \frac{1-\delta_{\sup}}{1+\delta_{\sup}} \tag{7}$$

$$\delta_{\sup}^{-1} = \inf_{H(z)} \left\| (1 - H(z)G(z))^{-1} H(z)G(z) \right\|_\infty \tag{8}$$

where $H(z)$ is the transfer function of the controller.

However, a logarithmic quantizer (5) has an infinite number of quantization levels and is not implementable practically. One simple approach is to truncate the quantizer using a large saturator and a small dead zone. That is, we use a $2N$-level logarithmic quantizer with quantization density $\rho > \rho_{\inf}$

$$Q(y) = \begin{cases} \rho^i \mu_0, & \text{if } \frac{1}{1+\delta}\rho^i\mu_0 < y \le \frac{1}{1-\delta}\rho^i\mu_0 \\ & \quad 0 < i < N-1 \\ \rho^{N-1}\mu_0, & \text{if } 0 \le y \le \frac{1}{1-\delta}\rho^{N-1}\mu_0 \\ \mu_0, & \text{if } y > \frac{1}{1+\delta}\mu_0 \\ -Q(-y), & \text{if } y < 0. \end{cases} \tag{9}$$

This quantization scheme will allow the state of the system to converge to a small neighborhood, provided that the initial state is within a known bound.

Our main objective here is to show that it is possible to dynamically scale the input-output signals of the quantizer so that asymptotic stabilization can be achieved using a finite-level logarithmic quantizer, even without knowing the bound for the initial state.

The basic idea of dynamic scaling is very simple: When the signal y_k is outside of the quantization range, we scale it back by a *scaling factor* (or *gain*) $g_k > 0$ before quantization. The quantized signal is then scaled back by g_k^{-1}. That is, we use

$$v_k = g_k^{-1} Q(g_k y_k). \tag{10}$$

The key problem with dynamic scaling is how to design g_k. The main technical difficulty is that there is no separate feedback channel to communicate the gain value. One approach is that both sides of the feedback channel compute the same g_k independently. This is possible only when the gain g_k can be computed using only the quantized signal because this signal is available to both sides of the feedback channel, assuming no packet losses and transmission errors. In the sequel, we introduce a very simple dynamic scaling method.

The closed-loop system of (1), (2), (3), (4) and (10) is given by

$$\xi_{k+1} = \bar{A}\xi_k + \bar{B} g_k^{-1} Q(g_k \bar{C} \xi_k) \tag{11}$$

where $\xi = [x^T \ \hat{x}^T]^T$,

$$\bar{A} = \begin{bmatrix} A & BC_c \\ 0 & A_c \end{bmatrix}, \ \bar{B} = \begin{bmatrix} BD_c \\ B_c \end{bmatrix}, \ \bar{C} = [C \ 0]. \tag{12}$$

For the moment, we assume that an infinite-level logarithmic quantizer with density $\rho > \rho_{\inf}$ is adopted. Then, g_k has no effect. Following the sector bound approach [3], we can write (11) as

$$\xi_{k+1} = \bar{A}(\Delta_k)\xi_k \tag{13}$$

where $\bar{A}(\Delta_k) = \bar{A} + \bar{B}(1+\Delta_k)\bar{C}$ and Δ_k represents the quantization error defined by $\Delta_k y_k = Q(y_k) - y_k$ with $-\delta \le \Delta_k \le \delta$.

Because (13) is quadratically stable, we have a quadratic Lyapunov function $V(\xi) = \xi^T P \xi$ with $P = P^T > 0$ such that [3]

$$\bar{A}(\Delta)^T P \bar{A}(\Delta) < P, \quad \forall |\Delta| \le \delta. \tag{14}$$

Using the continuity argument, the above is equivalent to

$$\bar{A}(\Delta)^T P \bar{A}(\Delta) \le (1-\eta)P, \quad \forall |\Delta| \le \delta \tag{15}$$

for some $0 < \eta < 1$.

We now assume that a $2N$-level logarithmic quantizer with the same density ρ and dynamic scaling (9), (10) is applied instead. Let γ_1 and γ_2 be two positive scaling factors such that $0 < \gamma_1 < 1$, $\sqrt{1-\eta} < \gamma_2 < 1$, and

$$\gamma_1^2 \bar{A}^T P \bar{A} < (1-\eta)P. \tag{16}$$

Note that (16) is always possible by taking γ_1 sufficiently small.

We initialize g_0 to be any positive value and define g_{k+1} for any $k \ge 0$ as follows:

$$g_{k+1} = \begin{cases} g_k \gamma_1, & \text{if } |Q(g_k y_k)| = \mu_0 \\ g_k/\gamma_2, & \text{if } |Q(g_k y_k)| = \rho^{N-1}\mu_0 \\ g_k, & \text{otherwise.} \end{cases} \tag{17}$$

Because of the flexibility in g_0, we can normalize $\mu_0 = 1$ without loss of generality. We will also denote $\varepsilon = \rho^{N-1}$. The choice of g_0 does not affect stabilizability, but choosing it according to an estimate of $\|x_0\|$ helps improve the transient performance; see Example 2 in Section V.

Consider the scaled state defined by

$$z_k = g_k \xi_k \tag{18}$$

and the associated Lyapunov function $V(z) = z^T P z$. We have the following result:

Lemma 2.1: Consider the closed-loop system (11) with a scaled N-level logarithmic quantizer (9) and (17), where ρ in (9) is such that $\rho > \rho_{\inf}$, and γ_1, γ_2 and η are chosen according to (15), (16). Then, for any initial state x_0 and any $k \ge 0$

$$V(z_{k+1}) \le \begin{cases} (1-\eta)V(z_k), & \text{if } \varepsilon < \left|Q(\bar{C}z_k)\right| \le 1 \\ (1-\eta_1)V(z_k) + \eta_2 \varepsilon^2, & \text{if } \left|Q(\bar{C}z_k)\right| = \varepsilon \end{cases} \tag{19}$$

where

$$\eta_1 = 1 - \gamma_2^{-2}(1+\tau)(1-\eta) \tag{20}$$

$$\eta_2 = \gamma_2^{-2}(1+\tau^{-1})\bar{B}^T P \bar{B} \tag{21}$$

with τ being any positive constant satisfying $\eta_1 > 0$.

Proof: The result for the case of $\varepsilon < |Q(\bar{C}z_k)| < 1$ follows directly from (13), (15) and $g_{k+1} = g_k$. For the case of $|Q(\bar{C}z_k)| = 1$, $g_{k+1} = g_k\gamma_1$. It follows that:

$$V(z_{k+1}) = \gamma_1^2(\bar{A}z_k + \bar{B}\sigma_k)^T P (\bar{A}z_k + \bar{B}\sigma_k)$$

where $\sigma_k = \mathrm{sign}(\bar{C}z_k)$. Denote

$$f(u) = \gamma_1^2(\bar{A}z_k + \bar{B}u)^T P(\bar{A}z_k + \bar{B}u).$$

From (16), $f(0) \leq (1-\eta)V(z_k)$. Since $|Q(\bar{C}z_k)| = 1$, we have $\sigma_k = \theta u_1$ for some $0 < \theta \leq 1$, where $u_1 = (1+\Delta_k)\bar{C}z_k$ with $|\Delta_k| \leq \delta$ is the unsaturated output of the quantizer. Also from (15), we get

$$f(u_1) = \gamma_1^2 z_k^T \bar{A}(\Delta_k)^T P\bar{A}(\Delta_k)z_k \leq \gamma_1^2(1-\eta)V(z_k).$$

Since $f(u)$ is quadratic and convex (because $f(u) \to \infty$ when $|u| \to \infty$), it is clear that

$$V(z_{k+1}) = f(\sigma_k) \leq \max\{f(0), f(u_1)\} \leq (1-\eta)V(z_k).$$

For the case of $|Q(\bar{C}z_k)| = \varepsilon$, $g_{k+1} = g_k/\gamma_2$. From (9) and (13), we can write

$$\xi_{k+1} = \bar{A}(\Delta_k)\xi_k + \bar{B}g_k^{-1}\varepsilon_k$$

where $|\varepsilon_k| \leq \varepsilon$. It follows that:

$$\begin{aligned}
V(z_{k+1}) &= \gamma_2^{-2}\left(\bar{A}(\Delta_k)z_k + \bar{B}\varepsilon_k\right)^T P\left(\bar{A}(\Delta_k)z_k + \bar{B}\varepsilon_k\right)\\
&= \gamma_2^{-2} z_k^T \bar{A}(\Delta_k)^T P\bar{A}(\Delta_k)z_k\\
&\quad + \gamma_2^{-2}\left(2\varepsilon_k \bar{B}^T P\bar{A}(\Delta_k)z_k + \varepsilon_k^2 \bar{B}^T P\bar{B}\right)\\
&\leq \gamma_2^{-2}(1+\tau)z_k^T \bar{A}(\Delta_k)^T P\bar{A}(\Delta_k)z_k\\
&\quad + \gamma_2^{-2}(1+\tau^{-1})\varepsilon^2 \bar{B}^T P\bar{B}\\
&\leq \gamma_2^{-2}(1+\tau)(1-\eta)z_k^T P z_k + \eta_2\varepsilon^2\\
&= (1-\eta_1)V(z_k) + \eta_2\varepsilon^2.
\end{aligned}$$

The above holds for any $\tau > 0$. Since $\sqrt{1-\eta} < \gamma_2 < 1$, we can choose τ sufficiently small to ensure $\eta_1 > 0$. □

From Lemma 2.1, it is clear that $V(z_k)$ converges to a bounded region. This bound can be computed by solving $-\eta_1 V_\infty + \eta_2\varepsilon^2 = 0$, which gives

$$V_\infty = \eta_1^{-1}\eta_2\varepsilon^2. \tag{22}$$

Lemma 2.1 leads to the following result:

Corollary 2.1: Suppose the scaled $2N$-level logarithmic quantizer (9), (10) and (17) is applied. Then, for any initial state x_0, $z_k = g_k\xi_k$ converges exponentially to the ellipsoid

$$Z_\infty = \left\{z : z \in \mathbb{R}^{2n}, V(z) \leq V_\infty\right\}. \tag{23}$$

From (22) and the corollary above, it is clear that we can choose N to be sufficiently large so that, when k is sufficiently large, $Q(\bar{C}z_k)$ will no longer be saturated. This is achieved by choosing N such that

$$|\bar{C}z| < 1, \; \forall z^T P z \leq \eta_1^{-1}\eta_2\rho^{2(N-1)}.$$

Since $\bar{C}z$ is a scalar, the above implies that

$$\bar{C}zz^T\bar{C}^T < 1, \; zz^T \leq \eta_1^{-1}\eta_2\rho^{2(N-1)}P^{-1}.$$

By substituting the second matrix inequality into the first one, we obtain $N > N_0$, where

$$N_0 = 1 + \frac{\log\left(\eta_1^{-1}\gamma_2^{-2}(1+\tau^{-1})\bar{B}^T P\bar{B}\bar{C}P^{-1}\bar{C}^T\right)}{2\log(\rho^{-1})}. \tag{24}$$

The analysis above yields the following main result:

Theorem 2.1: Suppose the scaled $2N$-level logarithmic quantizer (9), (10) and (17) is applied with $N > N_0$ in (24). Then, the state ξ_k converges to zero asymptotically.

Proof: From Corollary 2.1, z_k converges to Z_∞ exponentially. This property and the choice of N_0 imply that $Q(\bar{C}z_k)$ will no longer be saturated after a finite number of steps, say k_0 steps. This means that g_k will be non-decreasing for $k \geq k_0$. Note that whenever $g_{k+1} = g_k$, $V(z_k)$ decreases exponentially. If this continues for enough number of steps, $|\bar{C}z_k|$ will be less than ε, forcing g_{k+1} to increase by factor of $1/\gamma_2$. This means that g_k cannot converge to a constant. Hence, $g_k \to \infty$ as $k \to \infty$. Since z_k is bounded for $k > k_0$, we conclude that $\xi_k \to 0$ as $k \to \infty$. □

Remark 2.1: A typical behavior of the system is as follows: If the initial state is very large, the feedback signal tends to be saturated, forcing g_k to decrease fast. This would result in a period of overshoot. Once g_k is sufficiently small, saturation will stop and the state decays exponentially. When the state is sufficiently small, g_k will increase gradually, causing the quantizer to bounce back and forth between the dead zone and logarithmic region. During this phase, the state also decays exponentially, but at a lower rate.

III. Number of Quantization Levels

In this section, we try to analyze the number of quantization levels needed for stabilization. Recall that for a given controller (3), (4) with an infinite-level logarithmic quantizer with density $\rho > \rho_{\inf}$ that quadratically stabilizes the system (1), (2), a sufficient number of quantization levels is given by N_0 of (24). However, this formula is complicated because N_0 depends on a number of design parameters (η, γ_2, ρ, P, τ, and the controller). In the sequel, we consider how to choose these parameters.

We first minimize N_0 of (24) with respect to τ by assuming that other parameters are fixed. From (24), it is clear that minimizing N_0 is equivalent to

$$\min_{\tau>0} \eta_1^{-1}\gamma_2^{-2}(1+\tau^{-1}) \tag{25}$$

where η_1 is given in (20). The solution to (25) is simply given by

$$\tau = \frac{\gamma_2}{\sqrt{1-\eta}} - 1, \; \eta_1 = 1 - \gamma_2^{-1}\sqrt{1-\eta} \tag{26}$$

$$\min_{\tau>0} \eta_1^{-1}\gamma_2^{-2}(1+\tau^{-1}) = (\gamma_2 - \sqrt{1-\eta})^{-2} \tag{27}$$

Applying the above to (24) and noting $\log(\rho^{-1}) = -\log(\rho)$

$$N_0 = 1 + \frac{2\log(\gamma_2 - \sqrt{1-\eta}) - \log(\bar{B}^T P\bar{B}\bar{C}P^{-1}\bar{C}^T)}{2\log(\rho)}. \tag{28}$$

We next discuss the effect of γ_2 on N_0. Since $\gamma_2 < 1$ is required, it is clear from (28) that N_0 is minimized by taking γ_2 very close to 1, which, however, makes g_k increase very slowly, as seen from (17), resulting in that ξ_k converges to 0 very slowly. A good choice for γ_2 should balance the convergence rate of ξ_k and the number of quantization levels; see Example 1 in Section V.

With $\gamma_2 < 1$ chosen, we shall now minimize N_0 with respect to ρ, η, the controller and its associated P. Observe from (28) that N_0 can be reduced by increasing η and δ (or decreasing ρ). However, we can see from (15) that a larger η requires δ to be small. Furthermore, the choice of δ and η affect P and the controller. This implies that η and δ need to be optimized jointly. To this end, we return to (15) and provide the following relationship between δ and η.

Theorem 3.1: For any given $0 < \delta < \delta_{\sup}$, $0 < \eta < 1$ and $\sqrt{1-\eta} < \gamma_2 < 1$, N_0 in (28) is minimized by solving the following optimization problem:

$$\lambda_{\min} = \arg \min_{X,Y,R,S,W,D_c,\lambda_1,\lambda_2} \lambda_1\lambda_2 \tag{29}$$

subject to the following linear matrix inequalties:

$$\begin{bmatrix} (\eta-1)Y & * & * & * & * & * \\ (\eta-1)I & (\eta-1)X & * & * & * & * \\ AY+BW & A+BD_cC & -Y & * & * & * \\ R & XA+SC & -I & -X & * & * \\ 0 & 0 & D_c^TB^T & S^T & -1 & * \\ CY & C & 0 & 0 & 0 & -\delta^{-2} \end{bmatrix} < 0 \tag{30}$$

$$\begin{bmatrix} -Y & -I & BD_c \\ -I & -X & S \\ D_c^TB^T & S^T & -\lambda_1 \end{bmatrix} < 0,\ CYC^T < \lambda_2 \tag{31}$$

where $X = X^T, Y = Y^T \in \mathbb{R}^{n\times n}$, $R \in \mathbb{R}^{n\times n}$, $W \in \mathbb{R}^{1\times n}$, $S \in \mathbb{R}^{n\times 1}$ and λ_1 and λ_2 are scalars. The optimal N_0 is given by

$$N_0 = 1 + \frac{2\log(\gamma_2 - \sqrt{1-\eta}) - \log(\lambda_{\min})}{2\log(\rho)} \tag{32}$$

and the optimal controller (3), (4) is given by the solution of D_c together with

$$C_c = (W - D_cCY)\Psi^{-T} \tag{33}$$

$$B_c = M^{-1}(S - XBD_c) \tag{34}$$

$$A_c = M^{-1}(R - XAY - XBD_cCY - MB_cCY)\Psi^{-T} - M^{-1}XBC_c \tag{35}$$

where M and Ψ are any nonsingular matrices solving

$$M\Psi^T = I - XY \tag{36}$$

with M being the free parameter determining the state space realization of the controller.

Proof: First, we apply the S-Procedure [11] and standard technique of change of variables [10] to (15) to obtain (30). Then, observe from (28) that for the given ρ, η and γ_2, minimizing N_0 with respect to the controller and the matrix P is equivalent to minimizing $\lambda_1\lambda_2$, where $\bar{B}^TP\bar{B} < \lambda_1$ and $\bar{C}P^{-1}\bar{C}^T < \lambda_2$, which is equivalent to (31), following Schur complement and the above change of variables. The detail is omitted due to space limit. □

We now explain how to solve the optimization problem in Theorem 3.1. We assume that $\gamma_2 < 1$ is pre-specified (say, e.g., $\gamma_2 = 0.9$). Now, given η and δ (or ρ), (29) is bilinear in λ_1 and λ_2. Note that when $\lambda_1 = 1$, the first inequality in (31) is a part of (30). This means that the minimum of $\lambda_1\lambda_2$ is achieved at some $0 < \lambda_1 < 1$. Then, the minimum of N_0 can be found by numerically searching over a three-dimensional set Ω defined by

$$\Omega = \left\{(\delta,\eta,\lambda_1) : 0 < \delta < \delta_{\sup}, 1-\gamma_2^2 < \eta < 1, 0 < \lambda_1 < 1\right\}$$

and solving (29) for each chosen candidate in Ω. A simple brute force method, which also works well, is to discretize Ω uniformly in each dimension and solve (29) at each grid point. Since we only need an integer solution for the number of bits $N_b = \log_2(2N_0)$, the discretization can typically be done coarsely.

IV. Robustness Against Additive Noises

Next, we consider the scenario where the system (1), (2) is subject to some bounded additive noise, i.e., we consider the following system instead:

$$x_{k+1} = Ax_k + Bu_k + w_k \tag{37}$$

$$y_k = Cx_k \tag{38}$$

where $\|w_k\| \le \bar{w}$ for some constant $\bar{w} > 0$. The corresponding closed-loop system becomes

$$\xi_{k+1} = \bar{A}\xi_k + \bar{B}g_k^{-1}Q(g_k\bar{C}_k\xi_k) + \bar{I}w_k \tag{39}$$

where $\bar{I} = [I\ 0]^T$. Using the scaled state $z_k = g_k\xi_k$, (39) becomes

$$z_{k+1} = g_{k+1}^{-1}g_k\left(\bar{A}z_k + \bar{B}Q(\bar{C}_kz_k) + g_k\bar{I}w_k\right). \tag{40}$$

In this case, we want to drive the state to a bounded region. To do so, we first generalize Lemma 2.1 as follows:

Lemma 4.1: Consider the system (37), (38) and the dynamically scaled logarithmic controller as given before. Then, the scaled state $z_k = g_k\xi_k$ is bounded as follows:

$$V(z_{k+1}) \le \begin{cases} (1+\alpha)(1-\eta))V(z_k) \\ \quad +(1+\alpha^{-1})\|P_{11}\|g_{k+1}^2\bar{w}^2, & \text{if } \varepsilon < \left|Q(\bar{C}z_k)\right| \le 1 \\ (1+\alpha)\gamma_2^{-1}\sqrt{1-\eta}V(z_k) + (1+\alpha)\eta_2\varepsilon^2 \\ \quad +(1+\alpha^{-1})\|P_{11}\|g_{k+1}^2\bar{w}^2, & \text{if } \left|Q(\bar{C}z_k)\right| = \varepsilon \end{cases} \tag{41}$$

for any $\alpha > 0$, where $P_{11} = \bar{I}^TP\bar{I}$.

Proof: The proof is simply extended from the proof of Lemma 2.1. The detail is omitted. □

From Lemma 4.1, we see that, by choosing α sufficiently small, the scaled state converges to a bounded set when g_k has an upper bound $\bar{g}$. There are 2 steady state bounding sets for $V(z_k)$ from (41), associated with the three cases of $|Q(\bar{C}z_k)|$, and they are given by

$$Z_{\infty,1} = \left\{z : V(z) \le \frac{1+\alpha^{-1}}{1-(1+\alpha)(1-\eta)}\|P\|\bar{g}^2\bar{w}^2\right\} \tag{42}$$

$$Z_{\infty,2} = \left\{z : V(z) \le \frac{(1+\alpha)\eta_2}{1-(1+\alpha)\gamma_2^{-1}\sqrt{1-\eta}}\varepsilon^2 + \frac{1+\alpha^{-1}}{1-(1+\alpha)\gamma_2^{-1}\sqrt{1-\eta}}\|P\|\bar{g}^2\bar{w}^2\right\} \tag{43}$$

It is straightforward to minimize $Z_{\infty,1}$ with respect to α and the result is given by

$$Z_{\infty,1} = \left\{ z : V(z) \le \frac{1}{(1-\sqrt{1-\eta})^2} \|P_{11}\| \bar{g}^2 \bar{w}^2 \right\}. \quad (44)$$

Noting that the overall minimization of $Z_{\infty,2}$ is difficult, we choose to minimize the term associated with $\bar{w}$. It is easy to verify that the result is given by

$$Z_{\infty,2} = \left\{ z : V(z) \le \frac{\eta_2}{\sqrt{1-\eta_1}(1-\sqrt{1-\eta_1})} \varepsilon^2 + \frac{1}{(1-\sqrt{1-\eta_1})^2} \|P_{11}\| \bar{g}^2 \bar{w}^2 \right\} \quad (45)$$

where $1-\eta_1 = \gamma_2^{-1}\sqrt{1-\eta}$.

Theorem 4.1: Consider the system (37), (38) and the dynamically scaled logarithmic controller as given before. We require $N > N_0$ with N_0 given by (28) and modify the scaling factor g_k by saturating it at some $\bar{g}$. Then, both the closed-loop system state ξ_k and the scaled state $z_k = g_k \xi_k$ are bounded when $k \to \infty$.

Proof: The asymptotic boundedness of z_k follows easily from (44), (45). To show the asymptotic boundedness of ξ_k, it suffices to show that g_k has some lower bound $\underline{g}$ asymptotically. To do so, we define

$$Z_{\infty,3} = \left\{ z : z^T P z \le (\bar{C} P^{-1} \bar{C}^T)^{-1} \right\}.$$

It is easy to check that $|\bar{C}z| \le 1$ for all $z \in Z_{\infty,3}$. Since $N > N_0$, we know that $Z_\infty = \lambda Z_{\infty,3}$ for some $0 < \lambda < 1$, where Z_∞ is defined in (23) and $\lambda Z_{\infty,3} = \{z : \lambda^{-1} z \in Z_{\infty,3}\}$.

Next, we note that if $g_{k+1} \to 0$, we can take $\alpha = g_{k+1}$ so that (41) becomes (19). Using $Z_\infty = \lambda Z_{\infty,3}$, the above means that there exist $\breve{g}$ and η_3, both positive and sufficiently small, such that, if $g_{k+1} \le \breve{g}$, then

$$V(z_{k+1}) \le (1-\eta_3) V(z_k), \ \forall z_k \notin Z_{\infty,3} \text{ or } |\bar{C} z_k| > \varepsilon. \quad (46)$$

The exponential convergence rate above implies that there exists some integer $\kappa > 0$ such that for any initial $z_k \in Z_{\infty,2}$ in (45), it takes at most κ steps for z_k to reach $Z_{\infty,3}$, provided that g_{k+1} can be kept below $\breve{g}$ all the way. Once $z_k \in Z_{\infty,3}$, it will stay there until $g_k > \breve{g}$ again. Since the exponential decay in (46) continues to happen as long as $|\bar{C} z_k| > \varepsilon$, z_k will decay sufficiently to allow g_k to grow back until $g_k > \breve{g}$.

Now we define $\underline{g} = \gamma_1^{\kappa+1} \breve{g}$ and proceed to prove that $g_k \ge \underline{g}$ asymptotically. We assume, on the contrary, that there exists an increasing sequence of k_i, $i = 1, 2, \ldots$, such that $k_i \to \infty$ as $i \to \infty$ and $g_{k_i} < \underline{g}$ for all i. Since $z_k \to Z_{\infty,2}$ in (45) as $k \to \infty$, we may assume that k_1 is so large that $z_{k_1 - \kappa} \in Z_{\infty,2}$. From the definition of $\underline{g}$, we know that $g_k \le \breve{g}$ for all $k_1 - \kappa < k < k_1$. Hence, from our earlier discussion, we know that $z_{k_1} \in Z_{\infty,3}$ and that g_k will stop decaying when $k > k_1$ and will eventually grow back to $g_k > \breve{g}$ while keeping $z_k \in Z_{\infty,2}$. Once this happens, g_k can not decay down to $\underline{g}$ again because as soon as $g_k < \breve{g}$ (but with $g_k > \gamma_1 \breve{g}$), it takes at most κ steps for z_k to reach $Z_{\infty,3}$ again while keeping $g_k < \breve{g}$, and g_k can not go below $\underline{g}$ in κ steps. This conclusion contradicts the assumption made on the sequence $\{k_i\}$. Hence, $g_k \ge \underline{g}$ asymptotically. □

V. Illustrative Examples

In this section, we use two examples to illustrate the proposed dynamic scaling method.

Example 1: We consider a first order system

$$x_{k+1} = a x_k + u_k, \ \ y_k = x_k \quad (47)$$

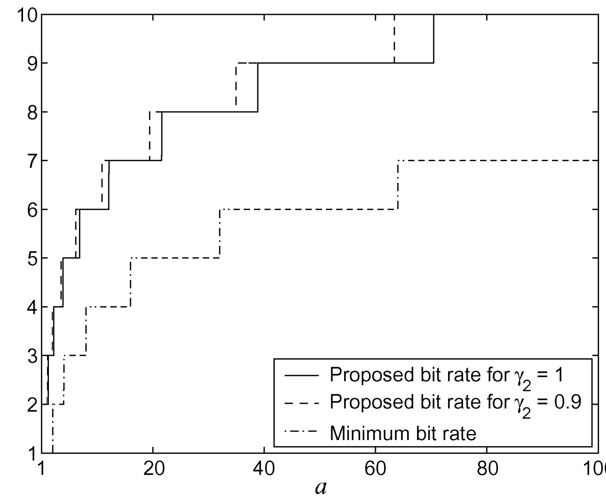

Fig. 1. Bit rate comparison for a first order system.

where $a > 1$. It turns out that we can have a relatively simple expression for N_0. Indeed, to stabilize the system using a logarithmic quantizer (5) with density ρ, the controller $H(z) = h$, where h is a constant, because of full state feedback. The closed-loop system is given by

$$x_{k+1} = (a + h(1+\Delta_k)) x_k, \ |\Delta_k| \le \delta$$

where δ relates to ρ as in (6). Since it is a first order system, we take $V(x_k) = x_k^2$, which gives

$$V(x_{k+1}) = (a + h(1+\Delta_k))^2 x_k^2 \le (|a+h| + \delta|h|)^2 x_k^2$$

with the right-hand side being the worst-case value. Minimizing it gives $h = -a$ and $V(x_{k+1}) \le \delta^2 a^2 V(x_k)$. This gives the upper bound for δ to be a^{-1}.

Now, for any $\delta < a^{-1}$, η in (15) is given by $\eta = 1 - \delta^2 a^2$. Applying it to (28), we obtain

$$N_0 = 1 + \frac{\log(\gamma_2 a^{-1} - \delta)}{\log(1-\delta) - \log(1+\delta)}, \ \delta < a^{-1} \quad (48)$$

which can be minimized numerically. The result is shown in Fig. 1, where two curves for the required bit rate, one for $\gamma_2 = 1$ and another for $\gamma_2 = 0.9$, are compared with the minimum bit rate $\lceil \log_2(a) \rceil$ given in [4]. We see that the difference is only a few bits even when a is taken up to 100.

Example 2: The second example we consider aims at demonstrating the convergence rate and robustness of the dynamic scaling method. Consider the system (1), (2) with

$$A = \begin{bmatrix} 2.7 & -2.41 & 0.507 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \ B = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$
$$C = [1 \quad -0.5 \quad 0.04].$$

The system is unstable with two unstable open-loop poles at $1.2 \pm i0.5$ but without unstable zero and the relative degree is 1. It follows from [3] that

$$\delta_{\sup} = |1.2 \pm i0.5|^{-2} = 0.5917, \quad \rho_{\inf} = 0.2565.$$

By applying the search mentioned in Section IV, we obtain the optimal values

$$\delta = 0.201, \ \eta = 0.561, \ \lambda_1 = 0.7223, \ \lambda_2 = 8.5046.$$

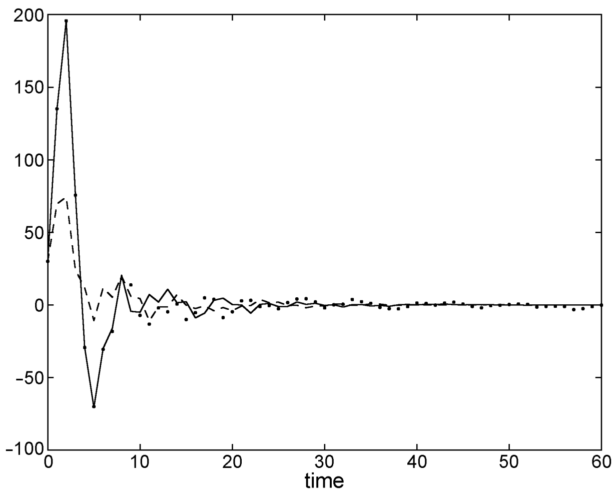

Fig. 2. Responses of x_1 of the closed-loop system with $N = 8$.

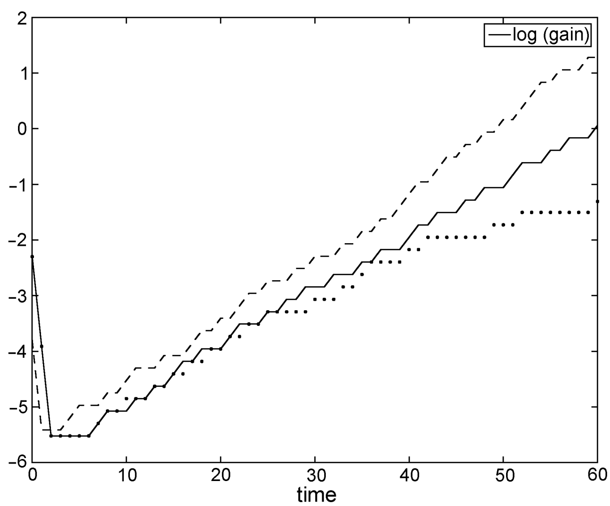

Fig. 3. Scaling factors g_k with $N = 8$.

The optimal controller is given by

$$\begin{aligned} A_c &= \begin{bmatrix} -255.6834 & 46.7502 & 217.854 \\ 616.3274 & -111.8387 & -523.9270 \\ -431.7862 & 79.0425 & 368.0348 \end{bmatrix} \\ B_c &= \begin{bmatrix} 5.8122 \\ -14.0003 \\ 9.8161 \end{bmatrix} \\ C_c &= [81.6699 \quad -15.0325 \quad -69.6715] \\ D_c &= -1.8594. \end{aligned}$$

Since γ_2 is lower bounded by $\sqrt{1-\eta} = 0.6626$, we choose $\gamma_2 = 0.8$. This gives $N_0 \approx 8$. We try $N = 8$ (4 b). Note that the minimal bit rate required for stabilizing this system is 1 bit [4].

Next, it can be easily verified that (16) is satisfied if $\gamma_1 \leq 0.25$. Thus, we take $\gamma_1 = 0.2$. Let the initial state of the controller be $\hat{x}_0 = [0\ 0\ 0]^T$ and $\mu_0 = 1$. The response of the first state variable of the closed-loop system with the initial state $x_0 = [30\ \ -30\ 0]^T$, $g_0 = 0.1$ and $N = 8$ is shown in Fig. 2 (solid line). Other state variables are not shown since they are similar. If x_0 is known, we may set $g_0 = 1/|Cx_0|$. The response of the first state variable under this situation is also given in Fig. 2 (dash line) which as expected, shows a much reduced overshoot. We also examine the robustness of the closed-loop system. Let w_k in (37) be a saturated Gaussian white noise with zero mean, covariance matrix $Q_w = 3I$ and $\bar{w} = 100$. For $N = 8$, $g_0 = 0.1$, $\mu_0 = 1$, and $\bar{g} = 0.3$, the response of the first state variable of the closed-loop system with $x_0 = [30\ \ -30\ 0]^T$ is also shown in Fig. 2 (dot line) for comparison. The corresponding scaling gains g_k for the above three cases are compared in Fig. 3.

VI. Conclusion

We have presented a simple dynamic scaling method for quantized output feedback control to achieve stabilization using a finite-level quantizer. The proposed control scheme is easily implementable and has nice convergence and robust properties. We emphasize that the concept of dynamic scaling can be applied to a much wider range of quantized feedback control problems, not just for the stabilization problem.

References

[1] N. Elia and K. Mitter, "Stabilization of linear systems with limited information," *IEEE Trans. Automat. Control*, vol. 46, no. 9, pp. 1384–1400, Sep. 2001.

[2] N. Elia, "Design of hybrid systems with guaranteed performance," in *Proc. 39th IEEE Conf. Decision Control*, 2000, pp. 993–998.

[3] M. Fu and L. Xie, "The sector bound approach to quantized feedback control," *IEEE Trans. Automat. Control*, vol. 50, no. 11, pp. 1698–1711, Nov. 2005.

[4] G. N. Nair and R. J. Evans, "Exponential stabilization of multi-dimensional linear systems," *Automatica*, vol. 39, pp. 585–593, 2003.

[5] D. F. Delchamps, "Stabilizing a linear system with qunatized state feedback," *IEEE Trans. Automat. Control*, vol. 35, no. 8, pp. 916–924, Aug. 1990.

[6] R. W. Brockett and D. Liberzon, "Quantized feedback stabilization of linear systems," *IEEE Trans. Automat. Control*, vol. 45, no. 7, pp. 1279–1289, Jul. 2000.

[7] F. Bullo and D. Liberzon, "Quantized cobtrol via locational optimization," *IEEE Trans. Automat. Control*, vol. 51, no. 1, pp. 2–13, Jan. 2006.

[8] D. Liberzon and D. Nesic, "Input-to-state stabilization of linear systems with quantized feedback," in *Proc. 44th IEEE Conf. Decision Control*, 2005, pp. 8197–8201.

[9] S. Tatikonda and S. Mitter, "Control under communication constraints," *IEEE Trans. Automat. Control*, vol. 49, no. 7, pp. 1056–1068, Jul. 2004.

[10] M. C. de Oliveira, J. C. Geromel, and J. Bernussou, "An LMI optimization approach to multiobjective controller design for discrete-time systems," in *Proc. 38th CDC*, Phoenix, AZ, Dec. 1999, pp. 3611–3616.

[11] L. Xie, "Output feedback H-infinity control of systems with parameter uncertainty," *Int. J. Control*, vol. 63, pp. 741–750, 1996.