Proceedings of the 27th Chinese Control Conference
July 16–18, 2008, Kunming, Yunnan, China

Quantization for Feedback Control and Estimation

Fu Minyue
School of Electrical Engineering and Computer Science, University of Newcastle, Australia
E-mail: minyue.fu@newcastle.edu.au

Abstract: This paper discusses a number of technical issues in a new area of research in control systems, namely, quantized feedback control and estimation. This area is motivated by the increasing need of incorporating communication networks in a control system. In such a framework, feedback information needs to be transmitted over a digital network, which results in a number of new challenges for control design. The focus of this article is on how to design quantizers for the purposes of control design and state estimation.

Key Words: Quantized feedback control, Networked control, Quantized estimation, Quantization

1 INTRODUCTION

The concept of feedback is the heart of the modern control theory. We use feedback to validate system models, predict system behaviors and drive controllers to deliver desired performances. Control designs rely heavily on the availability and accuracy of the feedback signal. Even in the case when feedback signals contain various types of noises, we typically require precise noise models, either deterministic or stochastic, to be available in design and analysis. Naturally, when additional errors occur in the feedback signal, the expected performance can no longer be guaranteed.

Quantization is a common source of errors which may cause the system performance to deteriorate. To simplify the design and analysis processes, quantization errors are often ignored or treated using simple noise models. This approach is valid only when the quantization errors are insignificant. However, there are two important scenarios where such treatment may be inadequate. One scenario is when the system requires high performance while high-precision sensors may not be available. For example, a high-resolution optical encoder may be too expensive. Another scenario is in the so-called networked control when the measured signal must be transmitted over a digital communication link with a restrictive data rate. A wireless sensor network is a typical example of this kind, where only low data rates are possible due to power and bandwidth constraints.

There are two types of research problems for control and estimation with regard to quantization errors. The first type assumes that that the quantizer is given and the problem is to work out the best strategy to reduce the impact of the quantization error. For example, a angular position of a rotational motor may be measured by a given optical encoder which has a fixed resolution. We must design an optimal controller or optimal estimator based on the given quantized signal. The second type allows the quantizer to be designed in conjunction with a controller or estimator. This type of problem arises in networked control where the measured signal can be regarded as a high-precision signal but it must be quantized for digital transmission. In this case, we are allowed to choose the rules for quantization. This paper will discuss both types of problems.

2 LINEAR QUANTIZATION

Quantizers can be either static or dynamic. We will discuss dynamic quantizers in Section 5. A static quantizer is a nonlinear function described by

$$v = Q(y) \tag{1}$$

where $y \in \mathbb{R}$ is the input signal and v is the output signal taking values in

$$\mathcal{V} = \{\pm\mu_i : i = 0, \pm1, \pm2, \cdots\}, \tag{2}$$

Linear quantizers are very common and they have

$$Q(y) = i\varepsilon + d, \text{ if } i\varepsilon \leqslant y < (i+1)\varepsilon,\ i = 0, \pm1, \dots \tag{3}$$

where ε is the quantization step size and d is the offset. It is clear that $\mu_i = i\varepsilon + d$ in (2). We assume $d = \varepsilon/2$, which makes $Q(y)$ an even function. In practice, the quantized output is saturated to yield a finite-level quantizer.
Linear quantizers are common because of its simplicity in construction and the fact that the quantizer can be modeled as a linear mapping with an additive quantization noise, i.e.,

$$v = y + \Delta(y) \tag{4}$$

with $\Delta(y) = Q(y) - y$ which has the property that $|\Delta(y)| \leqslant \varepsilon/2$.
The performance of a quantizer depends on how "small" the quantization noise is. This, in turn, depends on the type of input signal to the quantizer. The following result shows that linear quantization is optimal when the input signal is uniformly distributed in a given interval and the quantization noise is measured in average power (mean squares).

Theorem 2.1 Suppose y is a scalar random noise with uniform distribution in $[-1,\ 1)$ and $v = Q(y)$ is an N-level quantizer with an even N. Let the quantization noise be measured by

$$E = \mathcal{E}\{\Delta(y)^2\} \tag{5}$$

where $\mathcal{E}(\cdot)$ is the expectation operator. Then, the optimal structure for $Q(\cdot)$ which minimizes E is a linear quantizer

$$Q(y) = i\varepsilon + d, \text{ if } i\varepsilon \leqslant y < (i+1)\varepsilon,\ y \in [-1,\ 1) \tag{6}$$

with $\varepsilon = 2/N$ and $d = \varepsilon/2$.

Proof: Let the quantization intervals be given by $[\alpha_k, \alpha_{k+1}),\ k = 0, 1, \ldots, N-1$ with $\alpha_0 = -1, \alpha_N = 1$ and $\alpha_k < \alpha_{k+1}$ for all $0 \leqslant k < N$. Our first step is to show that if y falls into $[\alpha_k, \alpha_{k+1})$, the optimal $Q(y)$ should take the midpoint, i.e., $Q(y) = (\alpha_k + \alpha_{k+1})/2$. Indeed, denoting $Q(y)$ by v_k and the probability density by $p = 1/2$, the contribution of such y to E in (5) is given by

$$E_k = \int_{\alpha_k}^{\alpha_{k+1}} p(y - v_k)^2 dy = \frac{p}{3}[(\alpha_{k+1} - v_k)^3 - (\alpha_k - v_k)^3]$$

Differentiating E_k with respect to v_k and setting the derivative to zero yields $v_k = (\alpha_k + \alpha_{k+1})/2$, which gives the minimum

$$E_k = \frac{p}{12}(\alpha_{k+1} - \alpha_k)^3$$

Next, we prove by contradiction that the quantization intervals must be equal in length for E to reach minimum. Indeed, if this is not the case, then there must be at least two adjacent intervals, say $[\alpha_k, \alpha_{k+1})$ and $[\alpha_{k+1}, \alpha_{k+2})$, which are not equal in length, we argue that the sum of E_k and E_{k+1} can be reduced by shifting α_{k+1} to the midpoint of α_k and α_{k+2}. More precisely,

$$E_k + E_{k+1} = \frac{p}{12}[(\alpha_{k+1} - \alpha_k)^3 + (\alpha_{k+2} - \alpha_{k+1})^3$$

It is again straightforward that minimizing the above with respect to α_{k+1} yields $\alpha_{k+1} = (\alpha_k + \alpha_{k+2})/2$. Since $E = E_0 + E_1 + \ldots + E_{N-1}$, the above leads to a contradiction, which implies that the quantization intervals must be equal in length for E to reach minimum. Finally, $\varepsilon = 2/N$ and $d = \varepsilon$ follow from the size of each interval and that the quantized value is the midpoint of the interval.

3 LOGARITHMIC QUANTIZATION

Although linear quantizers have a number of advantages as explained above, it is not an ideal choice in many applications. In this section, we consider several cases where a logarithmic quantizer is more appropriate. A logarithmic quantizer is described by

$$\mathcal{V} = \{\mu_i = \rho^i \mu_0 :\ i = 0, \pm 1, \pm 2, \cdots\} \cup \{0\},\ \mu_0 > 0, \quad (7)$$

where $\rho \in (0,\ 1)$ and

$$Q(y) = \begin{cases} \rho^i \mu_0, & \text{if } \frac{1}{1+\delta}\rho^i \mu_0 < y \leqslant \frac{1}{1-\delta}\rho^i \mu_0, \\ 0, & \text{if } y = 0, \\ -Q(-y), & \text{if } y < 0, \end{cases} \quad (8)$$

where

$$\delta = \frac{1-\rho}{1+\rho}. \quad (9)$$

A pictorial representation is given in Fig. 1. The description above is for an infinite-level logarithmic quantizer. In practice, it is truncated when the input is too large (by a saturator) or too small (by a dead zone) in magnitude.

The first case where logarithmic quantization is superior to linear quantization is in quantized feedback control where the objective is to drive the output or the state to the origin but the control signal or measurement signal need to be quantized [1, 2]. This arises in stabilization, tracking and disturbance attenuation. The reason is that logarithmic quantization gives a multiplicative quantization error, which reduces as the input signal becomes small. As a tradeoff, the quantization error becomes large when the input signal is large, but this does not create problems.

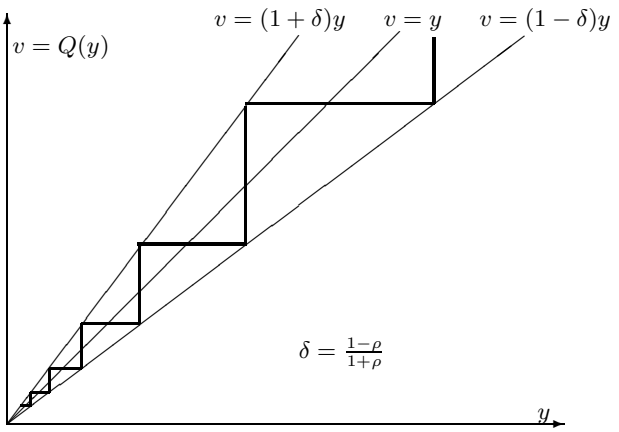

Fig. 1 Logarithmic Quantizer

The second case where logarithmic quantization is superior to linear quantization is in quantized state estimation where the state of a system needs to be estimated using quantized information [3]. If the measured signal is quantized directly, logarithmic quantization may not be appropriate because the measurement may be persistently large. However, one may quantize the estimation error instead. In doing so, logarithmic quantization is better because we want a small quantization error when the estimation error becomes small and we can tolerate a large quantization error when the estimation error is large.

Another case where logarithmic quantization is advantageous is when the signal to be quantized already has a multiplicative noise. Many sensors have the feature that measurement errors are specified using a relative error. For example, positions are often measured by range (distance) and most range sensors have accuracies specified by relative errors. Recall that logarithmic quantization also introduces a multiplicative error. When it is combined with a multiplicative noise, it is simply magnified without changing the noise structure.

It is interesting to note that most control and estimation settings deal with additive noises. We note here that this is indeed done mainly for mathematical convenience because multiplicative noises are somewhat more difficult to deal with; see [4].

3.1 Quantized Feedback Control

Consider the following system:

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k), \\ y(k) &= Cx(k), \end{aligned} \quad (10)$$

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 $G(z)$. We assume that A is unstable and (A, B, C) is a minimal realization.

The quantized feedback control problem is is depicted in Fig. 2, i.e., is to design a feedback quantizer

$$v(k) = Q(y(k)), \tag{11}$$

and a feedback controller of the form

$$\begin{aligned} \hat{x}(k+1) &= A_c\hat{x}(k) + B_c v(k), \quad \hat{x}(0) = 0, \\ u(k) &= C_c\hat{x}(k) + D_c v(k), \end{aligned} \tag{12}$$

with $\hat{x}(k) \in \mathbb{R}^n$, such that the closed-loop system is stable and that the so-called quantization density [1] is coarsest. The quantization density of $Q(\cdot)$ is defined as follows:

$$\eta_Q = \limsup_{\epsilon \to 0} \frac{\#g[\epsilon]}{-\ln \epsilon}, \tag{13}$$

where $\#g[\epsilon]$ denotes the number of quantization levels in the interval $[\epsilon,\ 1/\epsilon]$.

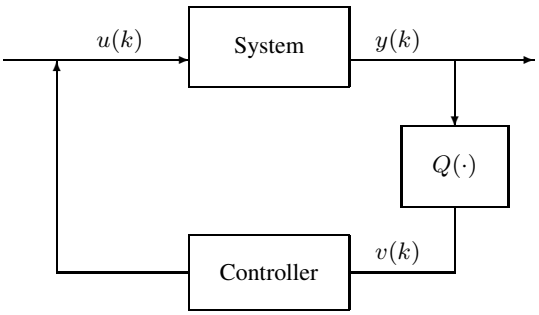

Fig. 2 Quantized Feedback Control

It was shown in [2] that the optimal quantizer structure for the quadratic stabilization of (10) is given by logarithmic quantization. Moreover, under quadratic stabilization, quantized feedback control is equivalent to robust control with sector bounded uncertainty, and the coarsest quantization density (which is equivalent to the smallest ρ) can be found by standard H_∞ optimization as detailed below.

Theorem 3.1 Consider the system (10). For a given quantization density $\rho > 0$, the system is quadratically stabilizable via a quantized controller (11) if and only if the following auxiliary system:

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ v(k) &= (1+\Delta)Cx(k), \ \ |\Delta| \leqslant \delta \end{aligned} \tag{14}$$

is quadratically stabilizable via:

$$\begin{aligned} x_c(k+1) &= A_c x_c(k) + B_c v(k) \\ u(k) &= C_c x_c(k) + D_c v(k) \end{aligned} \tag{15}$$

where δ, which is the sector bound produced by the quantization error, and ρ are related by (8).
The largest sector bound $\delta_{\sup}$ (which gives $\rho_{\inf}$) is given by

$$\delta_{\sup} = (\inf_{H(z)} \|\bar{G}_c(z)\|_\infty)^{-1} \tag{16}$$

where $\bar{G}_c(z) = (1 - H(z)G(z))^{-1}H(z)G(z)$ and $H(z) = D_c + C_c(zI - A_c)^{-1}B_c$.

The result builds a fundamental bridge between quantized feedback control and robust control, paving way for a lot of further research on networked control.

3.2 Quantized State Estimation

Consider the following linear system:

$$\begin{aligned} x(k+1) &= Ax(k) + Bw(k), \ \ x(0) = x_0 \\ y(k) &= Cx(k) + v(k) \end{aligned} \tag{17}$$

where $w(k) \in \mathbb{R}^m$ is the process noise, $v(k) \in \mathbb{R}$ is the measurement noise. It is assumed that $x_0 \in \mathbb{R}^n$ is a random variable with mean $\bar{x}_0$ and covariance Σ_0, and w and v are uncorrelated zero-mean white noises with covariances Σ_w and Σ_v, respectively, and they are uncorrelated with x_0.
We study the problem of state estimation using quantized measurement transmitted over a digital communication channel with a limited data rate. It is desirable to know how to quantize the measured signal so that good state estimation can be achieved using limited information.

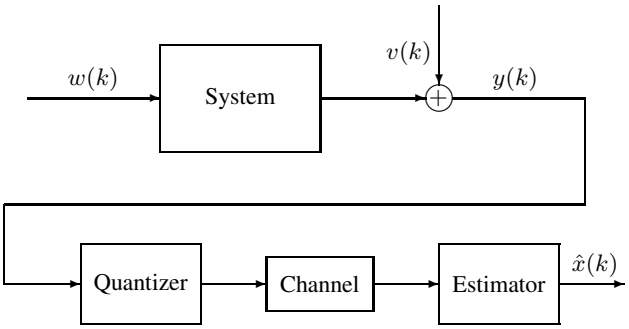

Fig. 3 Quantized State Estimation

The quantized estimator is shown in Fig. 3. Instead of quantizing the measured signal directly, we choose to quantize the prediction error of the estimator. The estimator is chosen to be

$$\begin{aligned} \hat{x}(k+1) &= A\hat{x}(k) + LQ(y(k) - \hat{y}(k)), \ \ \hat{x}(0) = \bar{x}_0 \\ \hat{y}(k) &= C\hat{x}(k)) \end{aligned} \tag{18}$$

where $\hat{x}(k) \in \mathbb{R}^n$ is the estimate of $x(k)$, $\hat{y}(k) \in \mathbb{R}$ is the estimate of $y(k)$ based on $\hat{x}(k)$, $Q(\cdot)$ is the quantizer, and L is the estimator gain.
Note in the above that state estimation is constructed only using the quantized prediction error. Therefore, under the ideal channel assumption, both sides of the channel can construct the same estimate using the quantized prediction error. In particular, the construction of $\hat{x}(k)$ on the transmission side does not require the estimated state to be transmitted back from the receiver side.
A logarithmic quantizer is used. Defining the estimation error

$$e(k) = x(k) - \hat{x}(k)$$

and its covariance matrix

$$E(k) = \mathcal{E}\{e(k)e^T(k)\}$$

the aim is to design both the filter gain L and the quantizer so that the trace of the asymptotic $E(k)$, i.e., $E = \lim_{k\to\infty}$, is to be minimized. Details can be found in [3].

We now demonstrate quantized state estimation by an example. The system model is given by (17) with

$$A = \begin{bmatrix} 2.4744 & -2.811 & 1.7038 & -.5444 & .0723 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 01 & 0 & 0 & \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix};$$

$$B^T = [1\ 0\ 0\ 0\ 0];$$

$$C = [0.245\ 0.236\ 0.384\ 0.146\ 0.035] \quad (19)$$

$\Sigma_w = 1$ and $\Sigma_v = 1/16$. The range of δ for the tests is chosen to be $[0,\ 0.3]$. For each δ, we try two estimator gains L, one taken as the Kalman gain designed by ignoring the quantization error and one being the robust gain computed by treating the quantization error as a multiplicative noise. Fig. 4 shows the simulated values of $\mathrm{Tr}(E)$. Also shown in the figure are the estimates of $\mathrm{Tr}(E)$ which we can ignore for this paper. We have two observations: 1) As the quantization becomes coarse (ρ becomes small or δ becomes large), the estimation error increases; 2) the robust gain outperforms the Kalman gain more significantly when the quantization becomes coarse.

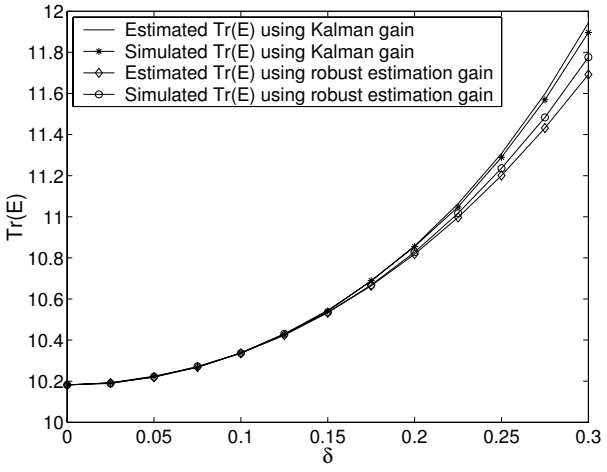

Fig. 4 Infinite-level Logarithmic Quantization

When the quantizer is truncated to a finite-level one, additional estimation error arises. In this case, apart from the ρ, the parameter μ_0 in the quantizer needs to be designed as well. As a result, with about $4 \sim 5$ bits of quantization, the quantized estimator has its estimation error variance only marginally larger than in the case without quantization. The details on the design of ρ and μ_0 can be found in [3]. Fig. 5 shows the result of estimation error vs. the number of quantization bits N_b.

4 NONLINEAR QUANTIZATION

Logarithmic quantization is a special type of nonlinear quantization. In this section, we consider a case when the input signal to the quantizer has a given Gaussian distribution and present an optimal nonlinear quantizer.

Consider an N-level quantizer acting on a random variable with Gaussian distribution $\mathcal{N}(0,1)$. An N-level quantizer is defined as

$$y = Q(x) = y_i, \ \text{ if } \ x_{i-1} < x \leqslant x_i \quad (20)$$

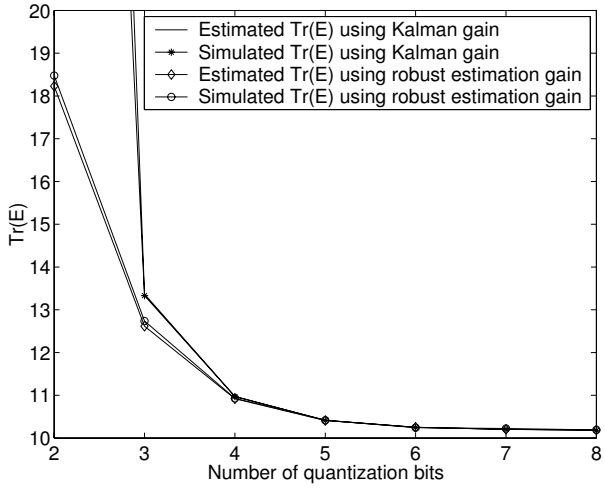

Fig. 5 16-level Quantization

where $x_0 < x_1 < \ldots < x_N$ with $x_0 = -\infty$ and $x_N = \infty$. We will call $[x_{i-1}, x_i)$ a quantization interval and y_i the associated quantization level.

Our objective is to choose the quantization intervals and quantization levels so that the quantization error has the minimum variance. That is, we want to minimize

$$\Sigma = \mathcal{E}(x - Q(x))^2 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (x - Q(x))^2 e^{-x^2/2} dx \quad (21)$$

We can rewrite Σ as

$$\Sigma = \sum_{i=1}^{N} \Sigma_i = \sum_{i=1}^{N} \frac{1}{\sqrt{2\pi}} \int_{x_{i-1}}^{x^i} (x - y_i)^2 e^{-x^2/2} dx \quad (22)$$

Lemma 4.1 Defining

$$F(\alpha, \beta, \gamma) = \int_a^{\gamma} (x - \alpha)^2 e^{-x^2/2} dx + \int_{\gamma}^{b} (x - \beta)^2 e^{-x^2/2} dx \quad (23)$$

for some fixed a and b with $0 \leqslant a < b$. Then, $F(\alpha, \beta, \gamma)$ is minimized when

$$\begin{aligned} \alpha &= f(a, \gamma); \\ \beta &= f(\gamma, b) \\ \gamma &= \frac{1}{2}(f(a, \gamma) + f(\gamma, b)) \end{aligned} \quad (24)$$

where

$$f(x, y) = \frac{e^{-x^2/2} - e^{-y^2}}{\sqrt{2\pi}(Q(x) - Q(y))} \quad (25)$$

with

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-x^2/2} dx \quad (26)$$

Moreover, (24) has a unique solution and can be numerically solved by iterating

$$\gamma_{k+1} = \frac{1}{2}(f(a, \gamma_k) + f(\gamma_k, b)), \ \ \gamma_0 = \frac{1}{2}(a + b) \quad (27)$$

i.e., $\gamma = \lim_{k \to \infty} \gamma_k$.

Proof: Partially differentially $F(\alpha,\beta,\gamma)$ with respect to α and setting it to zero gives

$$\begin{aligned}&\frac{\partial}{\partial\alpha}F(\alpha,\beta,\gamma)\\ =\;&\frac{2}{\sqrt{2\pi}}\int_a^\gamma(\alpha-x)e^{-x^2/2}dx\\ =\;&2\alpha(Q(a)-Q(\gamma))-\frac{2}{\sqrt{2\pi}}(e^{-a^2/2}-e^{-\gamma^2/2})=0\end{aligned}$$

which gives the solution to α. The solution to β is given similarly by partially differentiating $F(\alpha,\beta,\gamma)$ with respect to β. Now, partially differentially $F(\alpha,\beta,\gamma)$ with respect to γ and setting it to zero gives

$$\frac{\partial}{\partial\gamma}F(\alpha,\beta,\gamma)=\frac{1}{\sqrt{2\pi}}\{(\gamma-\alpha)^2-(\gamma-\beta)^2)e^{-\gamma^2/2}=0$$

which gives $\gamma=(\alpha+\beta)/2$.
Next, we show that (24) has a unique solution and it is a minimizer. We have $f(y,x)=f(x,y)$,

$$\begin{aligned}\lim_{y\to x}f(x,y)&=\lim_{y\to x}\frac{\frac{d}{dy}(e^{-x^2/2}-e^{-y^2/2})}{\sqrt{2\pi}\frac{d}{dy}(Q(x)-Q(y))}\\ &=\lim_{y\to x}\frac{ye^{-y^2/2}}{e^{-y^2/2}}=x\end{aligned}$$

and

$$\begin{aligned}\frac{d}{dy}f(x,y)&=\frac{1}{\sqrt{2\pi}}\int_x^y(y-t)e^{-t^2/2}dt\\ &\quad\cdot\frac{e^{-y^2/2}}{\sqrt{2\pi}(Q(x)-Q(y))^2}\end{aligned}\tag{28}$$

which is positive for $y>x$. It follows that

$$a<f(a,\gamma)<f(\gamma,b)<b,\;\text{ for }\;a<\gamma<b\tag{29}$$

Define

$$g(\gamma)=\gamma-\frac{1}{2}(f(a,\gamma)+f(\gamma,b))$$

Then,

$$g(a)=\frac{1}{2}(a-f(a,b))<0;\;g(b)=\frac{1}{2}(b-f(a,b))>0$$

We claim that $g(\gamma)$ is strictly monotonically increasing in $(a,\,b)$. Indeed,

$$\frac{d}{d\gamma}g(\gamma)=1-\frac{1}{2}(\frac{d}{d\gamma}f(a,\gamma)+\frac{d}{d\gamma}f(b,\gamma))$$

Using (28), we get

$$\begin{aligned}&\frac{d}{dy}f(x,y)\\ <\;&(y-x)\frac{1}{\sqrt{2\pi}}\int_x^y e^{-t^2/2}dt\frac{e^{-y^2/2}}{\sqrt{2\pi}(Q(x)-Q(y))^2}\\ =\;&\frac{(y-x)e^{-y^2/2}}{\sqrt{2\pi}(Q(x)-Q(y))}=\frac{\frac{1}{\sqrt{2\pi}}\int_x^y e^{-y^2/2}dt}{Q(x)-Q(y)}<1\end{aligned}$$

Hence, $dg(\gamma)/d\gamma>0$ for all $a<\gamma<b$. Hence, our claim holds. It then follows that (24) has a unique solution γ^*. Because $F(\alpha,\beta,\gamma)$ can be made arbitrarily large by choosing α or β arbitrarily large, the uniqueness of (24) means that the solution is indeed a minimizing one.
Finally, the reason for $\gamma_k\to\gamma^*$ as $k\to\infty$ is because $\gamma_{k+1}-\gamma_k=g(\gamma_k)$ which is positive when $\gamma_k<\gamma^*$ or negative when $\gamma_k>\gamma^*$. (More thoughts needed here)
Based on Lemma 4.1, we propose the following algorithm:

Step 1 : Choose any $x_1<x_2<\ldots<x_{N-1}$. One good empirical choice is to choose them to be uniformly distributed in $[-2\;2]$.

Step 2 : For $i=1:N-1$, let $a=x_{i-1}$ and $b=x_{i+1}$ be fixed and optimize $\gamma=x_i$ by iterating (27) until γ_k is sufficiently stable.

Step 3 : Repeat Step 2 until $\{x_i\}$ are sufficiently stable.

Step 4 : Compute $y_i=f(x_{i-1},x_i)$ for all i.

This algorithm has the feature that in each iteration (Step 2), all the points of $\{x_i\}$ are updated. We have no proof yet this algorithm converges or it converges to a global optimal solution. One obvious observation is that the quantization error variance reduces in each iteration. Simulations seem to suggest that the convergence is very fast and the global optimal solution is always found.

5 DYNAMIC QUANTIZATION

A *dynamic quantizer* uses memory, i.e., it can use the past input-output values of the quantizer to determine how to quantize a current input value, and thus is more complex and potentially more powerful.
One 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]- [8]. n [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. The idea of quantizer with sensitivity 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. In [9], a simple dynamic scaling method has been studied. This method employs a finite-level logarithmic quantizer $Q(\cdot)$ in conjunction with the following scaling:

$$v_k=g_k^{-1}Q(g_ky_k).\tag{30}$$

where the scaling gain g_k is adjusted by

$$g_{k+1}=\begin{cases}g_k\gamma_1, & \text{if } |Q(g_ky_k)|=\mu_0,\\ g_k/\gamma_2, & \text{if } |Q(g_ky_k)|=\rho^{N-1}\mu_0,\\ g_k, & \text{otherwise.}\end{cases}\tag{31}$$

with some initial $g_0>0$, where $\gamma_1,\gamma_2\in(0,\,1)$ are design parameters. The basic idea is to scale down (resp. up) the

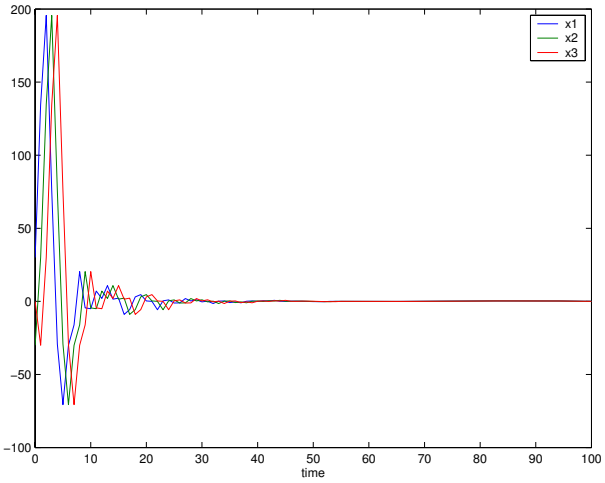

Fig. 6 Closed-loop Response with a 4-bit Quantizer

next input if the current input is too large (resp. small) in magnitude.

Note that g_{k+1} is determined based on when $Q(g_k y_k)$ (quantized information), no additional information needs to be passed on from the transmit side to the receive side for updating g_k, provided both sides start with the same g_0 and there is no transmission error for the quantized information. It is shown in [9] that it requires only a finite number of logarithmic quantization levels to quadratically stabilize a given linear system when the above dynamic scaling method is used. The detailed design of the dynamic quantizer and the controller are not discussed here.

Simulation results show that for most practical control systems, the number of quantization bits per time sample is very moderate [9]. To demonstrate this fact, we consider the system (10) with

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

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. Fig. 6 shows the state response of the closed-loop system with a 4-bit logarithmic quantizer.

6 CONCLUSION

In this paper, we have briefly discussed a number of quantization schemes for quantized feedback control and estimation. This is a relatively new area of research with many open and challenging questions. Although quantization is a well-studied subject in signal processing and digital communications, we caution that it is usually not appropriate to directly apply techniques in these areas to control problems. The main reason for this is that control systems involve feedback, which has two major implications: 1) The quantized signal re-enters the system; 2) The input signal to the quantizer is not known *a priori* to be bounded. Both of these implications make the analysis and design of quantizers much more difficult.

REFERENCES

[1] ELIA N, MITTER K. Stabilization of Linear Systems with Limited Information. IEEE Transactions on Automatic Control, 2001, 46(9): 1384-1400.

[2] FU M, XIE L. The sector bound approach to quantized feedback control. IEEE Transactions on Automatic Control, 2005, 50(11): 1698-1711.

[3] FU M, DE SOUZA C E. State Estimation of Linear Systems Using Quantized Information. IFAC World Congress, Seoul, Korea, 2008.

[4] GERSHON E, SHAKED U, YAESH U. Control and Estimation of State-Multiplicative Linear Systems, London: Springer-Verlag.

[5] DELCHAMPS D F. Stabilizing a linear system with qunatized state feedback. IEEE Transactions on Automatic Control, 1990, 35(8): 916-924.

[6] BROCKETT R W, LIBERZON D. Quantized feedback stabilization of linear systems. IEEE Transactions on Automatic Control, 2000, 45(7): 1279-1289.

[7] TATIKONDA S, MITTER S. Control under communication constraints. IEEE Transactions on Automatic Control, 2004, 49(7): 1056-1068.

[8] TATIKONDA S, MITTER S. Control over noisy channels. IEEE Transactions on Automatic Control, 2004, 49(7): 1196-1201.

[9] FU M, XIE L. Finite-Level Quantization Feedback Control for Linear Systems. IEEE Conf. Decision and Control, 2007.