Optimal Finite-Precision Digital Controller Realizations Based on an Improved Closed-Loop Stability Measure

S. Chen [1] *, J. Wu [2], R.H. Istepanian [3] and G. Li [4]

[1] Department of Electronics and Computer Science
University of Southampton, Southampton SO17 1BJ, U.K.

[2] National Laboratory of Industrial Control Technology
Zhejiang University, Hangzhou, 310027, P. R. China

[3] Department of Electrical and Computer Engineering
Ryerson Polytechnic University, Toronto, Canada, M5B 2K3

[4] School of Electrical and Electronic Engineering
Nanyang Technological University, Singapore

Abstract

This paper investigates the stability issue of the discrete-time control system, where a digital controller, implemented with finite word length (FWL), is used. A new tractable stability measure is derived, which is much less conservative than some existing measures in estimating the closed-loop stability robustness of an FWL implemented controller. An optimisation procedure is developed based on this improved measure to find the optimal realization for a general controller structure that includes output-feedback and observer-based controllers. A numerical example is used to verify the theoretical analysis and to illustrate the design procedure.

Keywords—finite word length, closed-loop stability, optimization.

1 Introduction

The current controller design methodology often assumes that the controller is implemented exactly, even though in reality a control law can only be realized in finite precision. The justification of this assumption is usually on the ground that the plant uncertainty is the most significant source of uncertainty. However, researchers have realized that the controller uncertainty has significant influence on the performance of the control system. A stable control system may achieve a lower than predicted performance or even become unstable when the control law is implemented with a finite-precision device. This is highlighted in the so-called fragility puzzles [1]–[3]: some robust optimal controllers are fragile. Ironically, these controllers have been designed to tolerate uncertainty in the plant, and yet small perturbations on the controller parameters may cause the closed-loop system to go unstable.

The FWL effect on the closed-loop stability depends on the controller realization structure. This property can be utilized to "select" controller realization in order to improve the "robustness" of closed-loop stability under controller perturbations. Several metrics, such as the LQG measure [4], the stability radius and $\mathcal{H}_\infty$ based measure [5], and the pole-sensitivity measure based on an l_2-norm [6], have been proposed to quantify the FWL effect on closed-loop stability. The approach of [6] is attractive because the design procedure for searching for optimal FWL controller realizations that maximize the proposed measure was developed. Recently, another tractable measure based on an l_1-norm [7] has been proposed, which provides a less conservative estimate of the FWL closed-loop stability robustness than the measure of [6].

The contribution of this paper is two fold. Unlike the most existing works which only consider output-feedback controller structure, we consider a general controller structure that includes output-feedback and observer-based controllers. We derive a new tractable stability measure for the unified controller structure and develop an optimisation procedure for finding the optimal controller realization that maximises this new measure. Through theoretical analysis and numerical results, it is shown that this improve measures is much less conservative in estimating the FWL closed-loop stability robustness of a controller realization than the measure given in [7].

2 Problem formulation

Consider the discrete-time closed-loop control system shown in Fig. 1, where the linear time-invariant plant $\hat{P}$ is described by

*Contact author. Tel/Fax: +44 (0)23 8059 6660/4508; Email: sqc@ecs.soton.ac.uk

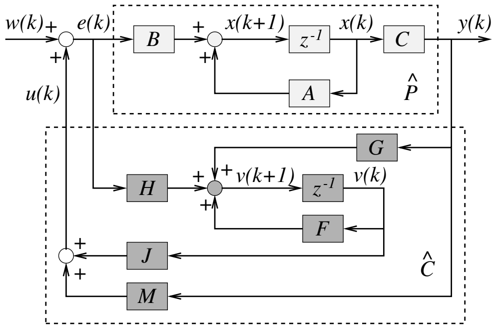

Figure 1: A discrete-time closed-loop system with a digital controller.

$$\begin{cases} \mathbf{x}(k+1) = A\mathbf{x}(k) + B\mathbf{e}(k) \\ \mathbf{y}(k) = C\mathbf{x}(k) \end{cases} \tag{1}$$

which is completely state controllable and observable with $A \in \mathcal{R}^{n\times n}$, $B \in \mathcal{R}^{n\times p}$ and $C \in \mathcal{R}^{q\times n}$; and the digital stabilizing controller $\hat{C}$ is described by

$$\begin{cases} \mathbf{v}(k+1) = F\mathbf{v}(k) + G\mathbf{y}(k) + H\mathbf{e}(k) \\ \mathbf{u}(k) = J\mathbf{v}(k) + M\mathbf{y}(k) \end{cases} \tag{2}$$

with $F \in \mathcal{R}^{m\times m}$, $G \in \mathcal{R}^{m\times q}$, $J \in \mathcal{R}^{p\times m}$, $M \in \mathcal{R}^{p\times q}$ and $H \in \mathcal{R}^{m\times p}$. The output-feedback and observer-based controllers can be unified in this general structure: $\hat{C}$ is an output-feedback controller when $H = 0$; a full-order observer-based controller when $F = A - GC$, $M = 0$ and $H = B$; a reduced-order observer-based controller, otherwise [8],[9].

Assume that a realization $(F_0, G_0, J_0, M_0, H_0)$ of $\hat{C}$ has been designed. It is well-known that the realizations of $\hat{C}$ are not unique. All the realizations of $\hat{C}$ form the realization set:

$$\begin{aligned} \mathcal{S} = \{(F,G,J,M,H) : F = T^{-1}F_0T, G = T^{-1}G_0, \\ J = J_0T, M = M_0, H = T^{-1}H_0\} \end{aligned} \tag{3}$$

where $T \in \mathcal{R}^{m\times m}$ is any real-valued non-singular matrix, called a similarity transformation. Let $\mathbf{w}_F = \mathrm{Vec}(F)$, where $\mathrm{Vec}(\cdot)$ denotes the column stacking operator. The vectors $\mathbf{w}_{F0}$, $\mathbf{w}_G$, $\mathbf{w}_{G0}$, $\mathbf{w}_J$, $\mathbf{w}_{J0}$, $\mathbf{w}_M$, $\mathbf{w}_{M0}$, $\mathbf{w}_H$ and $\mathbf{w}_{H0}$ are similarly defined. Denote

$$\mathbf{w} = \begin{bmatrix} w_1 \\ \vdots \\ w_N \end{bmatrix} \triangleq \begin{bmatrix} \mathbf{w}_F \\ \mathbf{w}_G \\ \mathbf{w}_J \\ \mathbf{w}_M \\ \mathbf{w}_H \end{bmatrix}, \quad \mathbf{w}_0 \triangleq \begin{bmatrix} \mathbf{w}_{F0} \\ \mathbf{w}_{G0} \\ \mathbf{w}_{J0} \\ \mathbf{w}_{M0} \\ \mathbf{w}_{H0} \end{bmatrix} \tag{4}$$

where $N = (m+p)(m+q) + mp$. We also refer to $\mathbf{w}$ as a realization of $\hat{C}$. The stability of the closed-loop system in Fig. 1 depends on the poles of the matrix

$$\begin{aligned} \bar{A}(\mathbf{w}) &= \begin{bmatrix} A + BMC & BJ \\ GC + HMC & F + HJ \end{bmatrix} \\ &= \begin{bmatrix} I & 0 \\ 0 & T^{-1} \end{bmatrix} \bar{A}(\mathbf{w}_0) \begin{bmatrix} I & 0 \\ 0 & T \end{bmatrix} \end{aligned} \tag{5}$$

All the different realizations $\mathbf{w}$ achieve exactly the same set of closed-loop poles if they are implemented with infinite precision. Since the closed-loop system is designed to be stable, the eigenvalues

$$|\lambda_i(\bar{A}(\mathbf{w}))| = |\lambda_i(\bar{A}(\mathbf{w}_0))| < 1,\ \forall i \in \{1, \dots, m+n\} \tag{6}$$

When a $\mathbf{w}$ is implemented with a fixed-point processor, it is perturbed into $\mathbf{w} + \Delta\mathbf{w}$ due to the FWL effect. Each element of $\Delta\mathbf{w}$ is bounded by $\pm\epsilon/2$,

$$\|\Delta\mathbf{w}\|_\infty \triangleq \max_{i\in\{1,\cdots,N\}} |\Delta w_i| \le \epsilon/2 \tag{7}$$

For a fixed point processor of B_s bits, let $B_s = B_i + B_f$, where 2^{B_i} is a "normalization" factor to make the absolute value of each element of $2^{-B_i}\mathbf{w}$ no larger than 1. Thus, B_i are bits required for the integer part of a number and B_f are bits used to implement the fractional part of a number. It can be seen that

$$\epsilon = 2^{-B_f}\,, \tag{8}$$

With the perturbation $\Delta\mathbf{w}$, $\lambda_i(\bar{A}(\mathbf{w}))$ is moved to $\lambda_i(\bar{A}(\mathbf{w} + \Delta\mathbf{w}))$. If an eigenvalue of $\bar{A}(\mathbf{w} + \Delta\mathbf{w})$ is outside the open unit disk, the closed-loop system, designed to be stable, becomes unstable with an FWL implemented $\mathbf{w}$. It is, therefore, critical to know when the FWL error will cause the closed-loop instability. This means to compute the following stability measure [5]:

$$\mu_0(\mathbf{w}) \triangleq \inf\{\|\Delta\mathbf{w}\|_\infty : \bar{A}(\mathbf{w} + \Delta\mathbf{w}) \text{ is unstable}\} \tag{9}$$

From this definition, it is obvious that:

Proposition 1 $\bar{A}(\mathbf{w} + \Delta\mathbf{w})$ is stable if $\|\Delta\mathbf{w}\|_\infty < \mu_0(\mathbf{w})$.

The larger $\mu_0(\mathbf{w})$ is, the larger FWL error the closed-loop stability can tolerate. Let $B_s^{\min}$ be the smallest word length, when used to implement $\mathbf{w}$, can guarantee the closed-loop stability. $B_s^{\min}$ is generally unknown. An estimate of $B_s^{\min}$ can be obtained by

$$\hat{B}_{s0}^{\min} = B_i + \mathrm{Int}[-\log_2(\mu_0(\mathbf{w}))] - 1 \tag{10}$$

where the integer $\mathrm{Int}[x] \ge x$. It can easily be seen that the closed-loop system remains stable if $\mathbf{w}$ is implemented with a fixed-point processor of at least $\hat{B}_{s0}^{\min}$. Moreover, as the stability measure $\mu_0(\mathbf{w})$ is a function of the controller realization $\mathbf{w}$, we can search for an "optimal" realization that maximizes $\mu_0(\mathbf{w})$:

$$\mathbf{w}_{\mathrm{opt}} = \arg\max_{\mathbf{w}\in\mathcal{S}} \mu_0(\mathbf{w}) \tag{11}$$

The difficulty with this approach is that computing the value of $\mu_0(\mathbf{w})$ is still an unsolved open problem. Thus, the stability measure $\mu_0(\mathbf{w})$ and the optimization procedure (11) have limited practical value.

3 A new FWL stability measure

Roughly speaking, how easily the FWL error $\Delta\mathbf{w}$ can cause a stable control system to become unstable is determined by how close $|\lambda_i(\bar{A}(\mathbf{w}))|$ are to 1 and how sensitive they are to the controller parameter perturbations. We propose the following FWL stability measure

$$\mu_{1a}(\mathbf{w}) \triangleq \min_{i\in\{1,\cdots,m+n\}} \frac{1-|\lambda_i(\bar{A}(\mathbf{w}))|}{\sigma_{ai}(\mathbf{w})} \tag{12}$$

with

$$\begin{aligned}\sigma_{ai}(\mathbf{w}) \quad &\triangleq \quad \alpha_{Fi}(\mathbf{w})+\alpha_{Gi}(\mathbf{w})+\alpha_{Ji}(\mathbf{w})\\ &\quad\quad +\alpha_{Mi}(\mathbf{w})+\alpha_{Hi}(\mathbf{w})\end{aligned} \tag{13}$$

$$\alpha_{Fi}(\mathbf{w}) \triangleq \delta(\mathbf{w}_F)\left\|\frac{\partial|\lambda_i(\bar{A}(\mathbf{w}))|}{\partial\mathbf{w}_F}\right\|_1 \tag{14}$$

$$\alpha_{Gi}(\mathbf{w}) \triangleq \delta(\mathbf{w}_G)\left\|\frac{\partial|\lambda_i(\bar{A}(\mathbf{w}))|}{\partial\mathbf{w}_G}\right\|_1 \tag{15}$$

$$\alpha_{Ji}(\mathbf{w}) \triangleq \delta(\mathbf{w}_J)\left\|\frac{\partial|\lambda_i(\bar{A}(\mathbf{w}))|}{\partial\mathbf{w}_J}\right\|_1 \tag{16}$$

$$\alpha_{Mi}(\mathbf{w}) \triangleq \delta(\mathbf{w}_M)\left\|\frac{\partial|\lambda_i(\bar{A}(\mathbf{w}))|}{\partial\mathbf{w}_M}\right\|_1 \tag{17}$$

$$\alpha_{Hi}(\mathbf{w}) \triangleq \delta(\mathbf{w}_H)\left\|\frac{\partial|\lambda_i(\bar{A}(\mathbf{w}))|}{\partial\mathbf{w}_H}\right\|_1 \tag{18}$$

where, for a vector $\mathbf{x}=[x_1\cdots x_p]^{\mathcal{T}}$,

$$\delta(\mathbf{x}) = \begin{cases} 0, & \text{if } \mathbf{x} \text{ is a zero vector}\\ 1, & \text{otherwise}\end{cases} \tag{19}$$

$^{\mathcal{T}}$ denotes the transpose operator and the l_1-norm of $\mathbf{x}$ is given by

$$||\mathbf{x}||_1 \triangleq \sum_{j=1}^{p}|x_i| \tag{20}$$

Defining

$$\begin{aligned}\mathcal{P}(\mathbf{w}) \triangleq \{\Delta\mathbf{w}:\ &|\lambda_i(\bar{A}(\mathbf{w}+\Delta\mathbf{w}))|-|\lambda_i(\bar{A}(\mathbf{w}))|\\ &\le \|\Delta\mathbf{w}\|_\infty \max_{j\in\{1,\cdots,m+n\}}\sigma_{aj}(\mathbf{w}),\ \forall i\}\end{aligned} \tag{21}$$

we have the following proposition

Proposition 2 $\bar{A}(\mathbf{w}+\Delta\mathbf{w})$ is stable if $\Delta\mathbf{w}\in\mathcal{P}(\mathbf{w})$ and $\|\Delta\mathbf{w}\|_\infty < \mu_{1a}(\mathbf{w})$.

Remarks: The requirement for $\Delta\mathbf{w}\in\mathcal{P}(\mathbf{w})$ is not over restricted. In practice, we will only be interested in those $\Delta\mathbf{w}$ that lie in the bounded region: $\mathcal{Q}(\mathbf{w}) \triangleq \{\Delta\mathbf{w}:\ \mu(\Delta\mathbf{w}) < \mu_0(\mathbf{w})\}$, i.e. those $\Delta\mathbf{w}$ that will not cause the closed-loop instability. Similar to [10], it can be shown that $\mathcal{P}(\mathbf{w})$ exists and at least a large part of $\mathcal{Q}(\mathbf{w})$ is covered by $\mathcal{P}(\mathbf{w})$. Define

$$\rho(\mathcal{P}(\mathbf{w})) \triangleq \inf_{\Delta\mathbf{w}\notin\mathcal{P}(\mathbf{w})}\|\Delta\mathbf{w}\|_\infty \tag{22}$$

Corollary 1 $\mu_{1a}(\mathbf{w}) \le \mu_0(\mathbf{w})$ if $\rho(\mathcal{P}(\mathbf{w})) > \mu_0(\mathbf{w})$.

It can be seen that $\mu_{1a}(\mathbf{w})$ is a lower bound of $\mu_0(\mathbf{w})$, provided that $\mu_0(\mathbf{w})$ is small enough. The assumption of small $\mu_0(\mathbf{w})$ is generally valid, and most of digital control systems do have a small stability robustness, especially when fast sampling is applied.

The stability measure $\mu_{1a}(\mathbf{w})$ is computationally tractable, as it can be shown that:

$$\frac{\partial|\lambda_i(\bar{A}(\mathbf{w}))|}{\partial F} = [\,0 \quad I\,]\,L_i(\mathbf{w})\begin{bmatrix}0\\ I\end{bmatrix} \tag{23}$$

$$\frac{\partial|\lambda_i(\bar{A}(\mathbf{w}))|}{\partial G} = [\,0 \quad I\,]\,L_i(\mathbf{w})\begin{bmatrix}C^{\mathcal{T}}\\ 0\end{bmatrix} \tag{24}$$

$$\frac{\partial|\lambda_i(\bar{A}(\mathbf{w}))|}{\partial J} = [\,B^{\mathcal{T}} \quad H^{\mathcal{T}}\,]\,L_i(\mathbf{w})\begin{bmatrix}0\\ I\end{bmatrix} \tag{25}$$

$$\frac{\partial|\lambda_i(\bar{A}(\mathbf{w}))|}{\partial M} = [\,B^{\mathcal{T}} \quad H^{\mathcal{T}}\,]\,L_i(\mathbf{w})\begin{bmatrix}C^{\mathcal{T}}\\ 0\end{bmatrix} \tag{26}$$

$$\frac{\partial|\lambda_i(\bar{A}(\mathbf{w}))|}{\partial H} = [\,0 \quad I\,]\,L_i(\mathbf{w})\begin{bmatrix}C^{\mathcal{T}}M^{\mathcal{T}}\\ J^{\mathcal{T}}\end{bmatrix} \tag{27}$$

with

$$L_i(\mathbf{w}) = \frac{\mathrm{Re}\left[\lambda_i^*(\bar{A}(\mathbf{w}))\mathbf{y}_i^*(\bar{A}(\mathbf{w}))\mathbf{x}_i^{\mathcal{T}}(\bar{A}(\mathbf{w}))\right]}{|\lambda_i(\bar{A}(\mathbf{w}))|} \tag{28}$$

where $\mathbf{x}_i(\bar{A}(\mathbf{w}))$ and $\mathbf{y}_i(\bar{A}(\mathbf{w}))$ are the right and reciprocal left eigenvectors related to the $\lambda_i(\bar{A}(\mathbf{w}))$, respectively, and * denotes the conjugate operation. Similar to (10), an estimate of $B_s^{\min}$ can be provided with $\mu_{1a}(\mathbf{w})$ by

$$\hat{B}_{s1a}^{\min} = B_i + \mathrm{Int}[-\log_2(\mu_{1a}(\mathbf{w}))] - 1 \tag{29}$$

Provided that the conditions of Proposition 2 and Corollary 1 are met, $\hat{B}_{s1a}^{\min} \ge \hat{B}_{s0}^{\min} \ge B_s^{\min}$. Unlike $\hat{B}_{s0}^{\min}$, however, $\hat{B}_{s1a}^{\min}$ can be computed easily.

An existing stability measure, which is also computationally tractable, is define as [7]:

$$\mu_1(\mathbf{w}) \triangleq \min_{i\in\{1,\cdots,m+n\}} \frac{1-|\lambda_i(\bar{A}(\mathbf{w}))|}{\sigma_i(\mathbf{w})} \tag{30}$$

with

$$\sigma_i(\mathbf{w}) \triangleq \beta_{Fi}(\mathbf{w}) + \beta_{Gi}(\mathbf{w}) + \beta_{Ji}(\mathbf{w}) + \beta_{Mi}(\mathbf{w}) + \beta_{Hi}(\mathbf{w}) \quad (31)$$

$$\beta_{Fi}(\mathbf{w}) \triangleq \delta(\mathbf{w}_F) \left\| \frac{\partial \lambda_i(\bar{A}(\mathbf{w}))}{\partial \mathbf{w}_F} \right\|_1 \quad (32)$$

$$\beta_{Gi}(\mathbf{w}) \triangleq \delta(\mathbf{w}_G) \left\| \frac{\partial \lambda_i(\bar{A}(\mathbf{w}))}{\partial \mathbf{w}_G} \right\|_1 \quad (33)$$

$$\beta_{Ji}(\mathbf{w}) \triangleq \delta(\mathbf{w}_J) \left\| \frac{\partial \lambda_i(\bar{A}(\mathbf{w}))}{\partial \mathbf{w}_J} \right\|_1 \quad (34)$$

$$\beta_{Mi}(\mathbf{w}) \triangleq \delta(\mathbf{w}_M) \left\| \frac{\partial \lambda_i(\bar{A}(\mathbf{w}))}{\partial \mathbf{w}_M} \right\|_1 \quad (35)$$

$$\beta_{Hi}(\mathbf{w}) \triangleq \delta(\mathbf{w}_H) \left\| \frac{\partial \lambda_i(\bar{A}(\mathbf{w}))}{\partial \mathbf{w}_H} \right\|_1 \quad (36)$$

An estimate of $B_s^{\min}$ is provided with $\mu_1(\mathbf{w})$ by

$$\hat{B}_{s1}^{\min} = B_i + \mathrm{Int}[-\log_2(\mu_1(\mathbf{w}))] - 1 \quad (37)$$

The key difference between $\mu_{1a}(\mathbf{w})$ and $\mu_1(\mathbf{w})$ is that the former considers the sensitivity of $|\lambda_i(\bar{A}(\mathbf{w}))|$ while the latter considers the sensitivity of $\lambda_i(\bar{A}(\mathbf{w}))$. It is well known that the stability of a linear discrete-time system depends only on the moduli of its eigenvalues. As $\mu_1(\mathbf{w})$ includes the unnecessary eigenvalue arguments in consideration, it is reasonable to believe that $\mu_1(\mathbf{w})$ is conservative in comparison with $\mu_{1a}(\mathbf{w})$. We can strictly verify this by:

$$\left| \frac{\partial |\lambda_i(\bar{A}(\mathbf{w}))|}{\partial w_j} \right| \leq \frac{\left| \lambda_i^*(\bar{A}(\mathbf{w})) \frac{\partial \lambda_i(\bar{A}(\mathbf{w}))}{\partial w_j} \right|}{|\lambda_i(\bar{A}(\mathbf{w}))|} = \left| \frac{\partial \lambda_i(\bar{A}(\mathbf{w}))}{\partial w_j} \right| \quad (38)$$

which means that $\sigma_{ai}(\mathbf{w}) \leq \sigma_i(\mathbf{w})$. This leads to:

Theorem 1 $\mu_1(\mathbf{w}) \leq \mu_{1a}(\mathbf{w})$ and $\hat{B}_{s1}^{\min} \geq \hat{B}_{s1a}^{\min}$.

4 Optimization procedure

As different realizations $\mathbf{w}$ yield different values of $\mu_{1a}(\mathbf{w})$, it is of practical importance to find a realization $\mathbf{w}_{\mathrm{opt}}$ that maximises $\mu_{1a}(\mathbf{w})$. Such a realization is optimal, since the digital controller implemented with $\mathbf{w}_{\mathrm{opt}}$ can tolerate a maximum FWL error. This optimal realization problem is formally defined as

$$\upsilon \triangleq \max_{\mathbf{w} \in \mathcal{S}} \mu_{1a}(\mathbf{w}) \quad (39)$$

Given $\mathbf{w}_0$, $\forall i \in \{1, \cdots, m+n\}$, we partition

$$\mathbf{x}_i(\bar{A}(\mathbf{w}_0)) = \begin{bmatrix} \mathbf{x}_{i,1}(\bar{A}(\mathbf{w}_0)) \\ \mathbf{x}_{i,2}(\bar{A}(\mathbf{w}_0)) \end{bmatrix} \quad (40)$$

$$\mathbf{y}_i(\bar{A}(\mathbf{w}_0)) = \begin{bmatrix} \mathbf{y}_{i,1}(\bar{A}(\mathbf{w}_0)) \\ \mathbf{y}_{i,2}(\bar{A}(\mathbf{w}_0)) \end{bmatrix} \quad (41)$$

where $\mathbf{x}_{i,1}(\bar{A}(\mathbf{w}_0)), \mathbf{y}_{i,1}(\bar{A}(\mathbf{w}_0)) \in \mathcal{C}^n$, $\mathbf{x}_{i,2}(\bar{A}(\mathbf{w}_0))$, $\mathbf{y}_{i,2}(\bar{A}(\mathbf{w}_0)) \in \mathcal{C}^m$. It is easily seen from (5) that

$$\mathbf{x}_i(\bar{A}(\mathbf{w})) = \begin{bmatrix} \mathbf{x}_{i,1}(\bar{A}(\mathbf{w}_0)) \\ T^{-1}\mathbf{x}_{i,2}(\bar{A}(\mathbf{w}_0)) \end{bmatrix} \quad (42)$$

$$\mathbf{y}_i(\bar{A}(\mathbf{w})) = \begin{bmatrix} \mathbf{y}_{i,1}(\bar{A}(\mathbf{w}_0)) \\ T^{\mathcal{T}}\mathbf{y}_{i,2}(\bar{A}(\mathbf{w}_0)) \end{bmatrix} \quad (43)$$

From (23)–(27), we have

$$\frac{\partial |\lambda_i(\bar{A}(\mathbf{w}))|}{\partial F} = T^{\mathcal{T}} L_{i,2,2}(\mathbf{w}_0) T^{-\mathcal{T}} \quad (44)$$

$$\frac{\partial |\lambda_i(\bar{A}(\mathbf{w}))|}{\partial G} = T^{\mathcal{T}} L_{i,2,1}(\mathbf{w}_0) C^{\mathcal{T}} \quad (45)$$

$$\frac{\partial |\lambda_i(\bar{A}(\mathbf{w}))|}{\partial J} = \left(B^{\mathcal{T}} L_{i,1,2}(\mathbf{w}_0) + H_0^{\mathcal{T}} L_{i,2,2}(\mathbf{w}_0)\right) T^{-\mathcal{T}} \quad (46)$$

$$\frac{\partial |\lambda_i(\bar{A}(\mathbf{w}))|}{\partial M} = \left(B^{\mathcal{T}} L_{i,1,1}(\mathbf{w}_0) + H_0^{\mathcal{T}} L_{i,2,1}(\mathbf{w}_0)\right) C^{\mathcal{T}} \quad (47)$$

$$\frac{\partial |\lambda_i(\bar{A}(\mathbf{w}))|}{\partial H} = T^{\mathcal{T}} \left(L_{i,2,1}(\mathbf{w}_0) C^{\mathcal{T}} M_0^{\mathcal{T}} + L_{i,2,2}(\mathbf{w}_0) J_0^{\mathcal{T}}\right) \quad (48)$$

where

$$L_{i,j,l}(\mathbf{w}_0) = \frac{\mathrm{Re}\left[\lambda_i^*(\bar{A}(\mathbf{w}_0)) \mathbf{y}_{i,j}^*(\bar{A}(\mathbf{w}_0)) \mathbf{x}_{i,l}^{\mathcal{T}}(\bar{A}(\mathbf{w}_0))\right]}{|\lambda_i(\bar{A}(\mathbf{w}_0))|}, \quad j,l = 1,2 \quad (49)$$

Define the following cost function:

$$f(T) \triangleq \min_{i \in \{1,\cdots,m+n\}} \frac{1 - |\lambda_i(\bar{A}(\mathbf{w}_0))|}{\sigma_{ai}(\mathbf{w})} = \mu_{1a}(\mathbf{w}) \quad (50)$$

The optimal realization problem (39) can then be posed as the following optimisation problem:

$$\upsilon \triangleq \max_{\substack{T \in \mathcal{R}^{m\times m} \\ \det(T) \neq 0}} f(T) \quad (51)$$

Although $f(T)$ is non-smooth and non-convex, efficient global optimisation methods exist for solving for this kind of optimisation problem. The adaptive simulated annealing (ASA) [11],[12] is such an algorithm and is adopted in this study to search for a true global optimum T_{opt} of the problem (51). With T_{opt}, we can obtain the optimal realization $\mathbf{w}_{\mathrm{opt}}$.

5 A numerical example

This section presents a numerical example to illustrate the design procedure and verify our theoretical results. The plant model used was a modification of the plant studied in [6], which was a single-input single-output system. We had added one more output that is the first state in the original plant model. The state-space model of this modified plant was given by

$$A = \begin{bmatrix} 3.2439e-1 & -4.5451e+0 & -4.0535e+0 & -2.7003e-3 & 0 \\ 1.4518e-1 & 4.9477e-1 & -4.6945e-1 & -3.1274e-4 & 0 \\ 1.6814e-2 & 1.6491e-1 & 9.6681e-1 & -2.2114e-5 & 0 \\ 1.1889e-3 & 1.8209e-2 & 1.9829e-1 & 1.0000e+0 & 0 \\ 6.1301e-5 & 1.2609e-3 & 1.9930e-2 & 2.0000e-1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1.4518e-1 \\ 1.6814e-2 \\ 1.1889e-3 \\ 6.1301e-5 \\ 2.4979e-6 \end{bmatrix}$$

$$C = \begin{bmatrix} 0 & 0 & 1.6188e+0 & -1.5750e-1 & -4.3943e+1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The closed-loop poles as given in [6] were used in design, and the designed reduced-order observer-based controller obtained using a standard design procedure [9] had the form:

$$F_0 = \begin{bmatrix} 0 & 1 \\ -9.3303e-1 & 1.9319e+0 \end{bmatrix}$$

$$G_0 = \begin{bmatrix} 4.1814e-2 & 2.7132e+2 \\ 3.9090e-2 & 1.0167e+3 \end{bmatrix}$$

$$J_0 = [\,3.0000e-4 \quad 5.0000e-4\,]$$

$$M_0 = [\,0 \quad 6.1250e-1\,], \quad H_0 = \begin{bmatrix} 7.8047e+1 \\ 7.3849e+1 \end{bmatrix}$$

With this initial controller realization $\mathbf{w}_0$, the transition matrix $\bar{A}(\mathbf{w}_0)$ was formed using (5), from which the poles and eigenvectors of the ideal closed-loop system were computed. The optimisation problem (51) was then formed with $T \in \mathcal{R}^{2\times 2}$. The ASA algorithm was used to find a T_{opt}, which was:

$$T_{\mathrm{opt}} = \begin{bmatrix} 1.4714e+1 & 3.2071e+1 \\ 1.3588e+1 & 3.0531e+1 \end{bmatrix}$$

From T_{opt}, the corresponding optimal controller realization $(F_{\mathrm{opt}}, G_{\mathrm{opt}}, J_{\mathrm{opt}}, M_{\mathrm{opt}}, H_{\mathrm{opt}})$ was determined

$$F_{\mathrm{opt}} = \begin{bmatrix} 9.8677e-1 & 1.4943e-2 \\ -2.9047e-2 & 9.4511e-1 \end{bmatrix}$$

$$G_{\mathrm{opt}} = \begin{bmatrix} 1.7066e-3 & -1.8080e+3 \\ 5.2084e-4 & 8.3794e+2 \end{bmatrix}$$

$$J_{\mathrm{opt}} = [\,1.1208e-2 \quad 2.4887e-2\,]$$

$$M_{\mathrm{opt}} = [\,0 \quad 6.1250e-1\,], \quad H_{\mathrm{opt}} = \begin{bmatrix} 1.0691e+0 \\ 1.9430e+0 \end{bmatrix}$$

For the initial and optimal controller realizations, the true minimal bit lengths $B_s^{\min}$ that can guarantee the closed-loop stability were also determined using a computer simulation method. Table 1 compares the values of the two stability measures, corresponding estimated minimum bit lengths and true minimum bit lengths for the initial and optimal controller realizations. The results clearly show that the new measure μ_{1a} is much less conservative than the existing measure μ_1 in estimating the true minimum bit length.

realization	$\mathbf{w}_0$	$\mathbf{w}_{\mathrm{opt}}$
B_i	10	11
μ_{1a}	$2.556877e-6$	$8.696940e-5$
$\hat{B}_{s1a}^{\min}$	28	24
μ_1	$4.050854e-7$	$3.012354e-6$
$\hat{B}_{s1}^{\min}$	31	29
$B_s^{\min}$	22	21

Table 1: Comparison of the two stability measures, corresponding estimated minimum bit lengths and true minimum bit lengths for the two reduced-order observer-based controller realizations.

We also computed the unit impulse response of the closed-loop control system when the controllers were the infinite-precision implemented $\mathbf{w}_0$ and various FWL implemented realizations. Notice that any realization $\mathbf{w} \in \mathcal{S}$, implemented in infinite precision, will achieve the exact performance of the infinite-precision implemented $\mathbf{w}_0$, which is the *designed* controller performance. For this reason, the infinite-precision implemented $\mathbf{w}_0$ is referred to as the *ideal* controller realization $\mathbf{w}_{\mathrm{ideal}}$. Figs. 2 and 3 compares the unit impulse response of the first plant output for the ideal controller $\mathbf{w}_{\mathrm{ideal}}$ with those of various 22-bit and 21-bit implemented realizations, respectively. It can be seen that the closed-loop became unstable with a 21-bit implemented controller realization $\mathbf{w}_0$. However, the closed-loop system remained stable with the 21-bit implemented $\mathbf{w}_{\mathrm{opt}}$.

6 Conclusions

In this paper, we have presented an approach to address the stability issue of the closed-loop discrete-time control system where a digital controller is implemented with a fixed-point processor. A new FWL closed-loop stability measure has been derived. It has been shown that this improved measure is a much less conservative lower bound of the computationally intractable true stability measure than other existing measures. As this new FWL stability measure is a function of the controller realization, it can be used as a cost function for obtaining an optimal controller realization that maximises the proposed measure. An efficient optimisation strategy has been developed based on the ASA algorithm for optimising a unified controller structure which includes output-feedback and observer-based controllers.

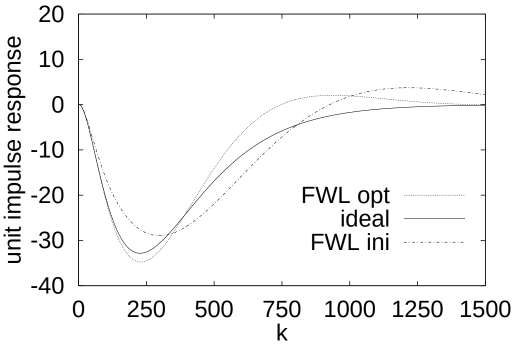

Figure 2: Comparison of unit impulse response for the infinite-precision controller implementation $\mathbf{w}_{\text{ideal}}$ with those for the two 22-bit implemented controller realizations $\mathbf{w}_0$ and $\mathbf{w}_{\text{opt}}$.

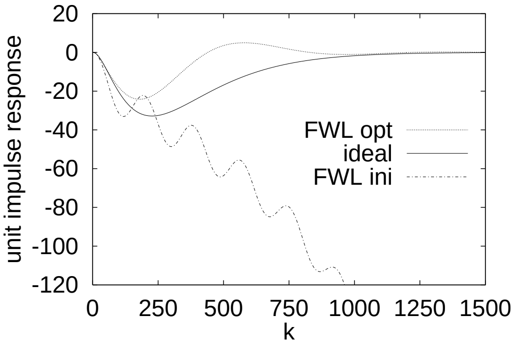

Figure 3: Comparison of unit impulse response for the infinite-precision controller implementation $\mathbf{w}_{\text{ideal}}$ with those for the two 21-bit implemented controller realizations $\mathbf{w}_0$ and $\mathbf{w}_{\text{opt}}$.

Acknowledgements

The first author, S. Chen, wishes to thank the support of the UK EPSRC under grant (GR/M16894).

References

[1] L.H. Keel and S.P. Bhattacharyya, "Robust, fragile, or optimal?" *IEEE Trans. Automatic Control*, Vol.42, No.8, pp.1098–1105, 1997.

[2] P.M. Mäkilä, "Comments on 'Robust, fragile, or optimal?'," *IEEE Trans. Automatic Control*, Vol.43, No.9, pp.1265–1267, 1998.

[3] P.M. Mäkilä, "Fragility and robustness puzzles," in *Proc. American Control Conf.* (San Diego, CA, USA), June 2-4, 1999, pp.2914–2919.

[4] P. Moroney, A.S. Willsky and P.K. Houpt, "The digital implementation of control compensators: the coefficient wordlength issue," *IEEE Trans. Automatic Control*, Vol.25, No.8, pp.621–630, 1980.

[5] I.J. Fialho and T.T. Georgiou, "On stability and performance of sampled data systems subject to word length constraint," *IEEE Trans. Automatic Control*, Vol.39, No.12, pp.2476–2481, 1994.

[6] G. Li, "On the structure of digital controllers with finite word length consideration," *IEEE Trans. Automatic Control*, Vol.43, pp.689–693, 1998.

[7] R.H. Istepanian, G. Li, J. Wu and J. Chu, "Analysis of sensitivity measures of finite-precision digital controller structures with closed-loop stability bounds," *IEE Proc. Control Theory and Applications*, Vol.145, No.5, pp.472–478, 1998.

[8] T. Kailath, *Linear Systems*. Prentice-Hall, 1980.

[9] J. O'Reilly, *Observers for Linear Systems*. London: Academic Press, 1983.

[10] S. Chen, J. Wu, R.H. Istepanian, J. Chu and J.F. Whidborne, "Optimizing stability bounds of finite-precision controller structures for sampled-data systems in the delta operator domain," *IEE Proc. Control Theory and Applications*, Vol.146, No.6, pp.517–526, 1999.

[11] L. Ingber, "Simulated annealing: practice versus theory," *Mathematical and Computer Modeling*, Vol.18, No.11, pp.29–57, 1993.

[12] S. Chen and B.L. Luk, "Adaptive simulated annealing for optimization in signal processing applications," *Signal Processing*, Vol.79, No.1, pp.117–128, 1999.