An Improved Closed-Loop Stability Related Measure for Finite-Precision Digital Controller Realizations

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

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

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

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

[4] Department of Electronic and Computer Engineering
Brunel University, Uxbridge, Middlesex UB8 3PH, UK

Abstract

The pole-sensitivity approach is employed to investigate the stability issue of the discrete-time control system, where a digital controller, implemented with finite word length (FWL), is used. A new stability related measure is derived, which is more accurate in estimating the closed-loop stability robustness of an FWL implemented controller than some existing measures for the pole-sensitivity analysis. This improved stability measure thus provides a better criterion to find the optimal realizations for a generic 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—digital controller, 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. It is well-known that a designed stable control system may achieve a lower than predicted performance or even become unstable when the controller is implemented with a finite-precision device. It has been noted that a controller design can be implemented with different realizations and that the FWL effect on the closed-loop stability depends on the controller realization structure. This property

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

can be utilized to select controller realization in order to improve the robustness of closed-loop stability under controller perturbations. Currently, two approaches exist for determining the optimal controller realizations under different criteria, namely pole-sensitivity measures [1]-[5] and complex stability radius measures [6],[7].

In the first approach, the pole sensitivity measures based on a 2-norm [2] and a 1-norm [3] are used to quantify the FWL effect, leading to a non-convex and non-smooth optimization problem in finding an optimal FWL controller realization. Efficient global optimization techniques to solve for this optimization problem are readily available [4],[5],[8]. Fialho and Georgiou [7] used the complex stability radius measure to formulate an optimal FWL controller realization problem that can be represented as a special H_∞ norm minimization problem and solved for with the method of linear matrix inequality [9],[10]. In this second approach, the FWL perturbations are assumed to be complex-valued. Although this assumption is somewhat artificial, the approach has certain attractive features and requires further investigation.

The contribution of this note is two fold. Firstly, a generic controller structure is considered that includes output-feedback and observer-based controllers. Secondly, adopting the pole-sensitivity approach, a new stability related measure is proposed for the unified controller structure and an optimisation procedure is developed to find the optimal controller realization that maximises this new measure. Through theoretical analysis and numerical results, it is shown that this improved measure is less conservative in estimating the FWL closed-loop stability robustness of a controller realization than the existing pole-sensitivity measures of [2],[3].

2 Problem formulation

Consider the discrete-time closed-loop control system depicted in Fig. 1, where the linear time-invariant plant P is described by

$$\left\{\begin{aligned} \mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{e}(k) \\ \mathbf{y}(k) &= \mathbf{C}\mathbf{x}(k) \end{aligned}\right. \tag{1}$$

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

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

with $\mathbf{F} \in \mathcal{R}^{m\times m}$, $\mathbf{G} \in \mathcal{R}^{m\times q}$, $\mathbf{J} \in \mathcal{R}^{p\times m}$, $\mathbf{M} \in \mathcal{R}^{p\times q}$ and $\mathbf{H} \in \mathcal{R}^{m\times p}$. The output-feedback and observer-based controllers can be unified in this general structure: C is an output-feedback controller when $\mathbf{H} = \mathbf{0}$; a full-order observer-based controller when $\mathbf{F} = \mathbf{A} - \mathbf{GC}$, $\mathbf{M} = \mathbf{0}$ and $\mathbf{H} = \mathbf{B}$; a reduced-order observer-based controller, otherwise [11],[12].

Assume that a realization $(\mathbf{F}_0, \mathbf{G}_0, \mathbf{J}_0, \mathbf{M}_0, \mathbf{H}_0)$ of C has been designed. It is well-known that the realizations of C are not unique. All the realizations of C form the realization set:

$$\mathcal{S} \triangleq \{(\mathbf{F}, \mathbf{G}, \mathbf{J}, \mathbf{M}, \mathbf{H}) : \mathbf{F} = \mathbf{T}^{-1}\mathbf{F}_0\mathbf{T}, \mathbf{G} = \mathbf{T}^{-1}\mathbf{G}_0, \mathbf{J} = \mathbf{J}_0\mathbf{T}, \mathbf{M} = \mathbf{M}_0, \mathbf{H} = \mathbf{T}^{-1}\mathbf{H}_0\} \tag{3}$$

where $\mathbf{T} \in \mathcal{R}^{m\times m}$ is any real-valued non-singular matrix, called a similarity transformation. Let $\mathbf{w}_F = \mathrm{Vec}(\mathbf{F})$, where $\mathrm{Vec}(\cdot)$ denotes the column stacking operator. The vectors $\mathbf{w}_{F_0}$, $\mathbf{w}_G$, $\mathbf{w}_{G_0}$, $\mathbf{w}_J$, $\mathbf{w}_{J_0}$, $\mathbf{w}_M$, $\mathbf{w}_{M_0}$, $\mathbf{w}_H$ and $\mathbf{w}_{H_0}$ are similarly defined. Denote

$$\mathbf{w} = \begin{bmatrix} w_1 \\ \vdots \\ w_N \end{bmatrix} \triangleq \begin{bmatrix} \mathbf{w}_F \\ \mathbf{w}_G \\ \mathbf{w}_J \\ \mathbf{w}_M \\ \mathbf{w}_H \end{bmatrix} \quad \text{and} \quad \mathbf{w}_0 \triangleq \begin{bmatrix} \mathbf{w}_{F_0} \\ \mathbf{w}_{G_0} \\ \mathbf{w}_{J_0} \\ \mathbf{w}_{M_0} \\ \mathbf{w}_{H_0} \end{bmatrix} \tag{4}$$

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

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

All the different realizations $\mathbf{w}$ have 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

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

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 $\pm\epsilon/2$,

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

The value of ϵ is determined as follows. For a fixed point processor of B_s bits, let $B_s = B_i + B_f$, where 2^{B_i} is the smallest normalization factor that makes the absolute value of each element of $2^{-B_i}\mathbf{w}$ no larger than 1. Thus, B_i are bits needed for the integer part of a number and B_f are bits for implementing the fractional part of a number. It is easy to see

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

With the perturbation $\Delta\mathbf{w}$, $\lambda_i(\bar{\mathbf{A}}(\mathbf{w}))$ is moved to $\lambda_i(\bar{\mathbf{A}}(\mathbf{w} + \Delta\mathbf{w}))$. If an eigenvalue of $\bar{\mathbf{A}}(\mathbf{w} + \Delta\mathbf{w})$ 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 ultimately means that we would like to know the largest open "sphere" in the controller perturbation space, within which the closed-loop remains stable. The size or radius of this "sphere" is defined by [6]:

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

From the definition of $\mu_0(\mathbf{w})$, 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 that, 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 $\mathrm{Int}[x] \geq x$. It can easily be seen that the closed-loop system remains stable if $\mathbf{w}$ is implemented with a fixed-point processor of $\hat{B}_{s0}^{\min}$. As $\mu_0(\mathbf{w})$ is a function of the controller realization $\mathbf{w}$, an optimal realization can be found that maximizes $\mu_0(\mathbf{w})$. The difficulty however is that computing the value of $\mu_0(\mathbf{w})$ is an unsolved open problem. A practical solution is to consider a lower bound of the stability measure $\mu_0(\mathbf{w})$ in some sense, which is computationally tractable. Obviously, the closer such a lower bound is to $\mu_0(\mathbf{w})$, the better. The pole sensitivity measures [2],[3] can be regarded as such lower bounds.

3 A new FWL stability related 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 related measure[1]

$$\mu_{1I}(\mathbf{w}) \triangleq \min_{i\in\{1,\cdots,m+n\}} \frac{1-|\lambda_i(\bar{\mathbf{A}}(\mathbf{w}))|}{\bar{\sigma}_i(\mathbf{w})} \tag{11}$$

with

$$\bar{\sigma}_i(\mathbf{w}) \triangleq \sum_{\mathbf{X}=\mathbf{F},\mathbf{G},\mathbf{J},\mathbf{M},\mathbf{H}} \delta(\mathbf{w}_\mathbf{X}) \left\| \frac{\partial\,|\lambda_i(\bar{\mathbf{A}}(\mathbf{w}))|}{\partial \mathbf{w}_\mathbf{X}} \right\|_1 \tag{12}$$

[1]This measure, as shown later, is an improved version of the existing measure μ_1 and hence is denoted with μ_{1I}.

where, for a vector $\mathbf{x} \in \mathcal{C}^s$, the 1-norm $\|\mathbf{x}\|_1$ is defined as

$$\|\mathbf{x}\|_1 \stackrel{\triangle}{=} \sum_{i=1}^{s} |x_i| \tag{13}$$

and the indicator function $\delta(\mathbf{x})$ is given by

$$\delta(\mathbf{x}) = \begin{cases} 0, & \text{if } \mathbf{x} \text{ is a zero vector} \\ 1, & \text{otherwise} \end{cases} \tag{14}$$

Defining a perturbation subset to the controller realization $\mathbf{w}$

$$\mathcal{P}(\mathbf{w}) \stackrel{\triangle}{=} \{\Delta\mathbf{w} : |\lambda_i(\bar{A}(\mathbf{w}+\Delta\mathbf{w}))| - |\lambda_i(\bar{A}(\mathbf{w}))| \leq \|\Delta\mathbf{w}\|_\infty \bar{\sigma}_i(\mathbf{w}),\ \forall i\} \tag{15}$$

we have the following proposition, the proof of which is straightforward.

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

Remarks: The requirement for $\Delta\mathbf{w} \in \mathcal{P}(\mathbf{w})$ is not too restricted. In practice, we will only be interested in those $\Delta\mathbf{w}$ that lie in the bounded region: $\mathcal{Q}(\mathbf{w}) \stackrel{\triangle}{=} \{\Delta\mathbf{w} : \mu(\Delta\mathbf{w}) < \mu_0(\mathbf{w})\}$, i.e. those $\Delta\mathbf{w}$ that will not cause the closed-loop instability. Similar to [5], 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

$$\rho(\mathcal{P}(\mathbf{w})) \stackrel{\triangle}{=} \inf_{\Delta\mathbf{w} \notin \mathcal{P}(\mathbf{w})} \|\Delta\mathbf{w}\|_\infty \tag{16}$$

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

It can be seen that $\mu_{1I}(\mathbf{w})$ 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 related measure $\mu_{1I}(\mathbf{w})$ is computationally tractable, as it can be shown that:

$$\frac{\partial\,|\lambda_i(\bar{\mathbf{A}}(\mathbf{w}))|}{\partial\mathbf{F}} = [\,\mathbf{0} \quad \mathbf{I}\,]\,\mathbf{L}_i(\mathbf{w}) \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} \tag{17}$$

$$\frac{\partial\,|\lambda_i(\bar{\mathbf{A}}(\mathbf{w}))|}{\partial\mathbf{G}} = [\,\mathbf{0} \quad \mathbf{I}\,]\,\mathbf{L}_i(\mathbf{w}) \begin{bmatrix} \mathbf{C}^T \\ \mathbf{0} \end{bmatrix} \tag{18}$$

$$\frac{\partial\,|\lambda_i(\bar{\mathbf{A}}(\mathbf{w}))|}{\partial\mathbf{J}} = [\,\mathbf{B}^T \quad \mathbf{H}^T\,]\,\mathbf{L}_i(\mathbf{w}) \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} \tag{19}$$

$$\frac{\partial\,|\lambda_i(\bar{\mathbf{A}}(\mathbf{w}))|}{\partial\mathbf{M}} = [\,\mathbf{B}^T \quad \mathbf{H}^T\,]\,\mathbf{L}_i(\mathbf{w}) \begin{bmatrix} \mathbf{C}^T \\ \mathbf{0} \end{bmatrix} \tag{20}$$

$$\frac{\partial\left|\lambda_i(\bar{\mathbf{A}}(\mathbf{w}))\right|}{\partial \mathbf{H}} = [\,\mathbf{0} \quad \mathbf{I}\,]\, \mathbf{L}_i(\mathbf{w}) \begin{bmatrix} \mathbf{C}^T\mathbf{M}^T \\ \mathbf{J}^T \end{bmatrix} \tag{21}$$

with ^{T} denoting the transpose operator and

$$\mathbf{L}_i(\mathbf{w}) = \frac{\mathrm{Re}\left[\lambda_i^*(\bar{\mathbf{A}}(\mathbf{w}))\mathbf{y}_i^*(\bar{\mathbf{A}}(\mathbf{w}))\mathbf{x}_i^T(\bar{\mathbf{A}}(\mathbf{w}))\right]}{\left|\lambda_i(\bar{\mathbf{A}}(\mathbf{w}))\right|} \tag{22}$$

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

$$\hat{B}_{s1I}^{\min} = B_i + \mathrm{Int}[-\log_2(\mu_{1I}(\mathbf{w}))] - 1 \tag{23}$$

Provided that the conditions of Proposition 2 and Corollary 1 are met, $\hat{B}_{s1I}^{\min} \geq \hat{B}_{s0}^{\min} \geq B_s^{\min}$. Unlike $\hat{B}_{s0}^{\min}$, however, $\hat{B}_{s1I}^{\min}$ can be computed easily.

An existing stability related measure, which is also computationally tractable, is define as [3]:

$$\mu_1(\mathbf{w}) \triangleq \min_{i\in\{1,\cdots,m+n\}} \frac{1-\left|\lambda_i(\bar{\mathbf{A}}(\mathbf{w}))\right|}{\sigma_i(\mathbf{w})} \tag{24}$$

with

$$\sigma_i(\mathbf{w}) \triangleq \sum_{\mathbf{X}=\mathbf{F},\mathbf{G},\mathbf{J},\mathbf{M},\mathbf{H}} \delta(\mathbf{w}_{\mathbf{X}}) \left\| \frac{\partial\lambda_i(\bar{\mathbf{A}}(\mathbf{w}))}{\partial \mathbf{w}_{\mathbf{X}}} \right\|_1 \tag{25}$$

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

$$\hat{B}_{s1}^{\min} = B_i + \mathrm{Int}[-\log_2(\mu_1(\mathbf{w}))] - 1 \tag{26}$$

The key difference between $\mu_{1I}(\mathbf{w})$ and $\mu_1(\mathbf{w})$ is that the former considers the sensitivity of $\left|\lambda_i(\bar{\mathbf{A}}(\mathbf{w}))\right|$ while the latter considers the sensitivity of $\lambda_i(\bar{\mathbf{A}}(\mathbf{w}))$. 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 $\mu_{1I}(\mathbf{w})$. This can strictly be verified. Noting

$$\frac{\partial\left|\lambda_i(\bar{\mathbf{A}}(\mathbf{w}))\right|}{\partial w_j} = \mathrm{Re}\left[\lambda_i^*(\bar{\mathbf{A}}(\mathbf{w}))\frac{\partial\lambda_i(\bar{\mathbf{A}}(\mathbf{w}))}{\partial w_j}\right] / \left|\lambda_i(\bar{\mathbf{A}}(\mathbf{w}))\right| \tag{27}$$

one has

$$\left|\frac{\partial\left|\lambda_i(\bar{\mathbf{A}}(\mathbf{w}))\right|}{\partial w_j}\right| \leq \frac{\left|\lambda_i^*(\bar{\mathbf{A}}(\mathbf{w}))\frac{\partial\lambda_i(\bar{\mathbf{A}}(\mathbf{w}))}{\partial w_j}\right|}{\left|\lambda_i(\bar{\mathbf{A}}(\mathbf{w}))\right|} = \left|\frac{\partial\lambda_i(\bar{\mathbf{A}}(\mathbf{w}))}{\partial w_j}\right| \tag{28}$$

which means that $\bar{\sigma}_i(\mathbf{w}) \leq \sigma_i(\mathbf{w})$. We conclude that $\mu_1(\mathbf{w}) \leq \mu_{1I}(\mathbf{w})$ and $\hat{B}_{s1}^{\min} \geq \hat{B}_{s1I}^{\min}$. Notice that $\mu_{1I}(\mathbf{w})$ is also superior in this sense than another measure based on a 2-norm [2], called $\mu_2(\mathbf{w})$, since it has been shown that under the similar conditions $\mu_2(\mathbf{w}) \leq \mu_1(\mathbf{w})$ [3].

4 Optimization procedure

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

$$\upsilon \triangleq \max_{\mathbf{w}\in\mathcal{S}} \mu_{1I}(\mathbf{w}) \tag{29}$$

Given $\mathbf{w}_0$, $\forall i \in \{1,\cdots,m+n\}$, partition $\mathbf{x}_i(\bar{\mathbf{A}}(\mathbf{w}_0))$ and $\mathbf{y}_i(\bar{\mathbf{A}}(\mathbf{w}_0))$:

$$\mathbf{x}_i(\bar{\mathbf{A}}(\mathbf{w}_0)) = \begin{bmatrix} \mathbf{x}_{i,1}(\bar{\mathbf{A}}(\mathbf{w}_0)) \\ \mathbf{x}_{i,2}(\bar{\mathbf{A}}(\mathbf{w}_0)) \end{bmatrix}, \quad \mathbf{y}_i(\bar{\mathbf{A}}(\mathbf{w}_0)) = \begin{bmatrix} \mathbf{y}_{i,1}(\bar{\mathbf{A}}(\mathbf{w}_0)) \\ \mathbf{y}_{i,2}(\bar{\mathbf{A}}(\mathbf{w}_0)) \end{bmatrix} \tag{30}$$

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

$$\mathbf{x}_i(\bar{\mathbf{A}}(\mathbf{w})) = \begin{bmatrix} \mathbf{x}_{i,1}(\bar{\mathbf{A}}(\mathbf{w}_0)) \\ \mathbf{T}^{-1}\mathbf{x}_{i,2}(\bar{\mathbf{A}}(\mathbf{w}_0)) \end{bmatrix}, \quad \mathbf{y}_i(\bar{A}(\mathbf{w})) = \begin{bmatrix} \mathbf{y}_{i,1}(\bar{\mathbf{A}}(\mathbf{w}_0)) \\ \mathbf{T}^T\mathbf{y}_{i,2}(\bar{\mathbf{A}}(\mathbf{w}_0)) \end{bmatrix} \tag{31}$$

From (17)–(21), we have

$$\frac{\partial\left|\lambda_i(\bar{\mathbf{A}}(\mathbf{w}))\right|}{\partial \mathbf{F}} = \mathbf{T}^T\mathbf{L}_{i,2,2}(\mathbf{w}_0)\mathbf{T}^{-T} \tag{32}$$

$$\frac{\partial\left|\lambda_i(\bar{\mathbf{A}}(\mathbf{w}))\right|}{\partial \mathbf{G}} = \mathbf{T}^T\mathbf{L}_{i,2,1}(\mathbf{w}_0)\mathbf{C}^T \tag{33}$$

$$\frac{\partial\left|\lambda_i(\bar{\mathbf{A}}(\mathbf{w}))\right|}{\partial \mathbf{J}} = \left(\mathbf{B}^T\mathbf{L}_{i,1,2}(\mathbf{w}_0) + \mathbf{H}_0^T\mathbf{L}_{i,2,2}(\mathbf{w}_0)\right)\mathbf{T}^{-T} \tag{34}$$

$$\frac{\partial\left|\lambda_i(\bar{\mathbf{A}}(\mathbf{w}))\right|}{\partial \mathbf{M}} = \left(\mathbf{B}^T\mathbf{L}_{i,1,1}(\mathbf{w}_0) + \mathbf{H}_0^T\mathbf{L}_{i,2,1}(\mathbf{w}_0)\right)\mathbf{C}^T \tag{35}$$

$$\frac{\partial\left|\lambda_i(\bar{\mathbf{A}}(\mathbf{w}))\right|}{\partial \mathbf{H}} = \mathbf{T}^T\left(\mathbf{L}_{i,2,1}(\mathbf{w}_0)\mathbf{C}^T\mathbf{M}_0^T + \mathbf{L}_{i,2,2}(\mathbf{w}_0)\mathbf{J}_0^T\right) \tag{36}$$

where

$$\mathbf{L}_{i,j,l}(\mathbf{w}_0) = \frac{\operatorname{Re}\left[\lambda_i^*(\bar{\mathbf{A}}(\mathbf{w}_0))\mathbf{y}_{i,j}^*(\bar{\mathbf{A}}(\mathbf{w}_0))\mathbf{x}_{i,l}^T(\bar{\mathbf{A}}(\mathbf{w}_0))\right]}{\left|\lambda_i(\bar{\mathbf{A}}(\mathbf{w}_0))\right|}, \quad j,l = 1,2 \tag{37}$$

Define the following cost function:

$$f(\mathbf{T}) \triangleq \min_{i\in\{1,\cdots,m+n\}} \frac{1-\left|\lambda_i(\bar{\mathbf{A}}(\mathbf{w}_0))\right|}{\bar{\sigma}_i(\mathbf{w})} = \mu_{1I}(\mathbf{w}) \tag{38}$$

The optimal realization problem (29) can then be posed as the following optimisation problem:

$$\upsilon \triangleq \max_{\substack{\mathbf{T}\in\mathcal{R}^{m\times m} \\ \det(\mathbf{T})\neq 0}} f(\mathbf{T}) \tag{39}$$

Although $f(\mathbf{T})$ is non-smooth and non-convex, efficient global optimisation methods exist for solving for this kind of optimisation problem. The adaptive simulated annealing (ASA) [8] is such an algorithm and is adopted in this study to search for a true global optimum $\mathbf{T}_{\text{opt}}$ of the problem (39). With $\mathbf{T}_{\text{opt}}$, we can obtain the optimal realization $\mathbf{w}_{\text{opt}}$.

An alternative optimal realization problem is based on the complex stability radius measure [7]. Space limitation precludes a comparison with this alternative approach. A detailed study on the pole-sensitivity and complex stability radius measure approaches for finite-precision digital controller realizations can be found in [13].

5 An illustrative example

A numerical example is used to illustrate the design procedure and verify the theoretical results given in Section 3. The plant model used is a modification of the plant studied in [2], which was a single-input single-output system. We have added one more output that is the first state in the original plant model. The state-space model of this modified plant is given by

$$\mathbf{A} = \begin{bmatrix} 3.2439e-01 & -4.5451e+00 & -4.0535e+00 & -2.7003e-03 & 0 \\ 1.4518e-01 & 4.9477e-01 & -4.6945e-01 & -3.1274e-04 & 0 \\ 1.6814e-02 & 1.6491e-01 & 9.6681e-01 & -2.2114e-05 & 0 \\ 1.1889e-03 & 1.8209e-02 & 1.9829e-01 & 1.0000e+00 & 0 \\ 6.1301e-05 & 1.2609e-03 & 1.9930e-02 & 2.0000e-01 & 1 \end{bmatrix},$$

$$\mathbf{B} = \begin{bmatrix} 1.4518e-01 \\ 1.6814e-02 \\ 1.1889e-03 \\ 6.1301e-05 \\ 2.4979e-06 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 0 & 0 & 1.6188e+00 & -1.5750e-01 & -4.3943e+01 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The closed-loop poles as given in [2] were used in design, and the designed reduced-order observer-based controller obtained using a standard design procedure [12] had the form:

$$\mathbf{F}_0 = \begin{bmatrix} 0 & 1 \\ -9.3303e-01 & 1.9319e+00 \end{bmatrix}, \quad \mathbf{G}_0 = \begin{bmatrix} 4.1814e-02 & 2.7132e+02 \\ 3.9090e-02 & 1.0167e+03 \end{bmatrix},$$

$$\mathbf{J}_0 = [\,3.0000e-04 \quad 5.0000e-04\,], \quad \mathbf{M}_0 = [\,0 \quad 6.1250e-01\,], \quad \mathbf{H}_0 = \begin{bmatrix} 7.8047e+01 \\ 7.3849e+01 \end{bmatrix}$$

With this initial controller realization $\mathbf{w}_0$, the corresponding transition matrix $\bar{\mathbf{A}}(\mathbf{w}_0)$ was formed using (5), from which the poles and the eigenvectors of the ideal closed-loop system were computed. The optimisation problem (39) was then formed with $\mathbf{T} \in \mathcal{R}^{2\times 2}$. The ASA algorithm was used to find a $\mathbf{T}_{\text{opt}}$, which was:

$$\mathbf{T}_{\text{opt}} = \begin{bmatrix} 1.4714e+01 & 3.2071e+01 \\ 1.3588e+01 & 3.0531e+01 \end{bmatrix}$$

From $\mathbf{T}_{\text{opt}}$, the corresponding optimal controller realization $\mathbf{w}_{\text{opt}}$ was determined:

$$\mathbf{F}_{\text{opt}} = \begin{bmatrix} 9.8677e-01 & 1.4943e-02 \\ -2.9047e-02 & 9.4511e-01 \end{bmatrix}, \quad \mathbf{G}_{\text{opt}} = \begin{bmatrix} 1.7066e-03 & -1.8080e+03 \\ 5.2084e-04 & 8.3794e+02 \end{bmatrix},$$

$$\mathbf{J}_{\text{opt}} = [\, 1.1208e-02 \quad 2.4887e-02 \,], \quad \mathbf{M}_{\text{opt}} = [\, 0 \quad 6.1250e-01 \,], \quad \mathbf{H}_{\text{opt}} = \begin{bmatrix} 1.0691e+00 \\ 1.9430e+00 \end{bmatrix}$$

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 related 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_{1I} is much less conservative than the existing measure μ_1 in estimating the true minimum bit length.

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 $\mathbf{w}_{\text{ideal}}$. Figs. 2 and 3 compares the unit impulse response of the first plant output $y_1(k)$ for the ideal controller $\mathbf{w}_{\text{ideal}}$ 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 $\mathbf{w}_{\text{opt}}$.

6 Conclusions

We have applied the pole-sensitivity 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 related measure has been derived. It has been shown that this improved measure is a less conservative lower bound of the computationally intractable true stability measure than other existing measures for the pole-sensitivity method. As this new 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.

Acknowledgements

J. Wu and S. Chen wish to thank the support of the UK Royal Society under a KC Wong fellowship (RL/ART/CN/XFI/KCW/11949).

References

[1] M. Gevers and G. Li, *Parameterizations in Control, Estimation and Filtering Problems: Accuracy Aspects*. London: Springer Verlag, 1993.

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

[3] 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.

[4] S. Chen, J. Wu, R.H. Istepanian and J. Chu, "Optimizing stability bounds of finite-precision PID controller structures," *IEEE Trans. Automatic Control*, Vol.44, No.11, pp.2149–2153, 1999.

[5] J. Wu, S. Chen, G. Li and J. Chu, "Optimal finite-precision state-estimate feedback controller realization of discrete-time systems," *IEEE Trans. Automatic Control*, Vol.45, No.8, pp.1550–1554, 2000.

[6] 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.

[7] I.J. Fialho and T.T. Georgiou, "Optimal finite wordlength digital controller realization," in *Proc. American Control Conf.* (San Diego, USA), June 2-4, 1999, pp.4326-4327.

[8] 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.

[9] S. Boyd, L. EI Ghaoui, E. Feron and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*. SIAM, 1994.

[10] R.E. Skelton, T. Iwasaki and K.M. Grigoriadid, *A Unified Algebraic Approach to Linear Control Design*. London: Taylor and Francis, 1998.

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

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

[13] S. Chen, J. Wu and G. Li, "Two approaches based on pole sensitivity and stability radius measures for finite precision digital controller realizations," submitted to *Systems and Control Letters*, 2000.

realization	μ_{1I}	μ_1	$\hat{B}_{s1I}^{\min}$	$\hat{B}_{s1}^{\min}$	$B_s^{\min}$
$\mathbf{w}_0$	$2.556877e-06$	$4.050854e-07$	28	31	22
$\mathbf{w}_{\text{opt}}$	$8.696940e-05$	$3.012354e-06$	24	29	21

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

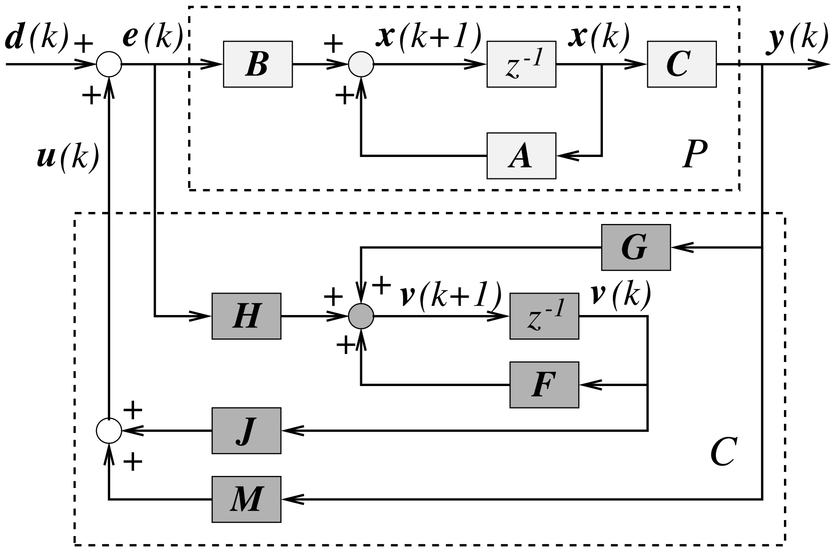

Figure 1: Discrete-time closed-loop system with a generic digital controller.

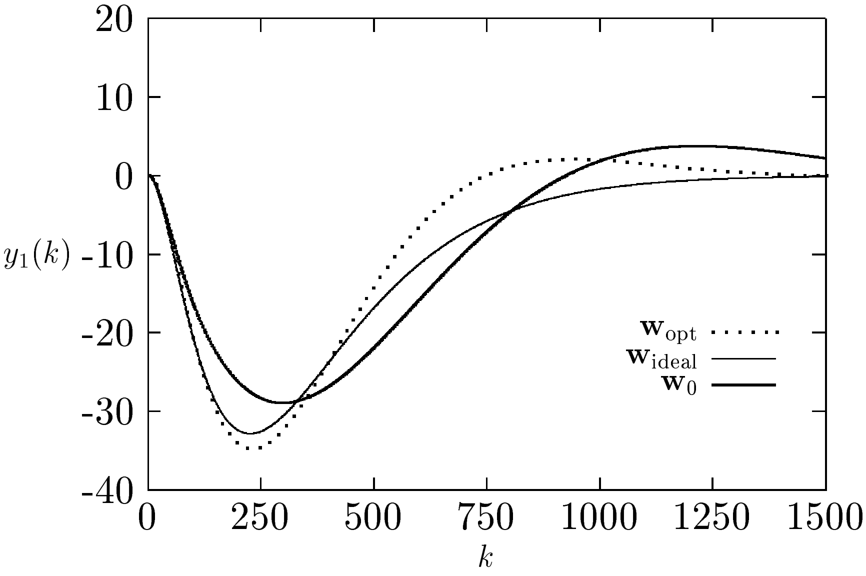

Figure 2: Comparison of unit impulse response for the infinite-precision controller implementation $\mathbf{w}_{\rm ideal}$ with those for the two 22-bit implemented controller realizations $\mathbf{w}_0$ and $\mathbf{w}_{\rm opt}$.

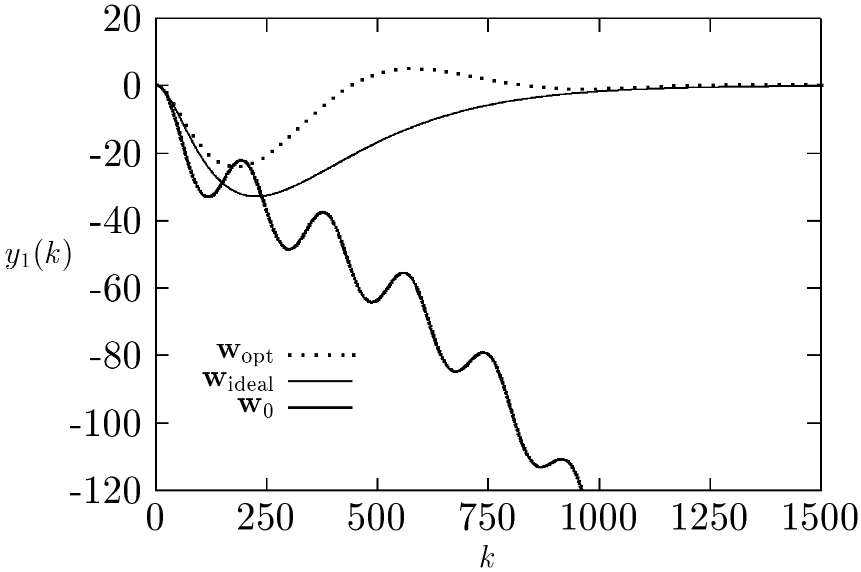

Figure 3: Comparison of unit impulse response for the infinite-precision controller implementation $\mathbf{w}_{\rm ideal}$ with those for the two 21-bit implemented controller realizations $\mathbf{w}_0$ and $\mathbf{w}_{\rm opt}$.