

©2008 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE

- [6] H. Ito, "Overlapping decomposition for multirate decentralized control," in *Preprints 8th IFAC/IFORS/IMACS/IFIP Symp. Large Scale Systems: Theory and Applications*, vol. 1, Patras, Greece, 1980, pp. 259–264.
- [7] K. Malinowski and M. Singh, "Controllability and observability of expanded systems with overlapping decompositions," *Automatica*, vol. 21, pp. 203–208, 1985.
- [8] J. M. Rossell, "Contribution to decentralized control of large-scale systems via overlapping models," Ph.D. dissertation (in Spanish), Tech. Univ. Catalunya, Barcelona, Spain, 1998.
- [9] D. D. Šiljak, *Decentralized Control of Complex Systems*. New York: Academic, 1991.
- [10] D. D. Šiljak, S. M. Mladenović, and S. Stanković, "Overlapping decentralized observation and control of a platoon of vehicles," in *Proc. Amer. Control Conf.*, San Diego, CA, 1999, pp. 4522–4526.
- [11] E. D. Sontag, *Mathematical Control Theory*. New York: Springer-Verlag, 1990.
- [12] S. S. Stanković, S. Mladenović, and D. D. Šiljak, "Headway control of a platoon of vehicles based on the inclusion principle," in *Complex Dynamic Systems with Incomplete Information*. Aachen, Germany: Shaker Verlag, 1998, pp. 153–167.
- [13] S. S. Stanković, X. B. Chen, and D. D. Šiljak, "Stochastic inclusion principle applied to decentralized control automatic generation control," *Int. J. Control*, vol. 72, no. 3, pp. 276–288, 1999.
- [14] S. S. Stanković and D. D. Šiljak, "Contractibility of overlapping decentralized control," in *Proc. Amer. Control Conf.*, Chicago, IL, 2000, pp. 811–818.

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

J. Wu, S. Chen, G. Li, R. H. Istepanian, and J. Chu

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.

Index Terms—Closed-loop stability, digital controller, finite word length, optimization.

I. INTRODUCTION

The current controller design methodology often assumes that the controller is implemented exactly, even though in reality a control law

Manuscript received May 15, 2000; revised October 5, 2000 and January 30, 2001. Recommended by Associate Editor T. Chen. The work of the first two authors was supported by the UK Royal Society under a KC Wong fellowship (RL/ART/CN/XFI/KCW/11949).

J. Wu and J. Chu are with the National Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou, 310027, P. R. China.

S. Chen is with the Department of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K. (e-mail: sqc@ecs.soton.ac.uk).

G. Li is with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore.

R. H. Istepanian is with the Department of Electronic and Computer Engineering, Brunel University, Uxbridge, Middlesex UB8 3PH, UK.

Publisher Item Identifier S 0018-9286(01)06599-0.

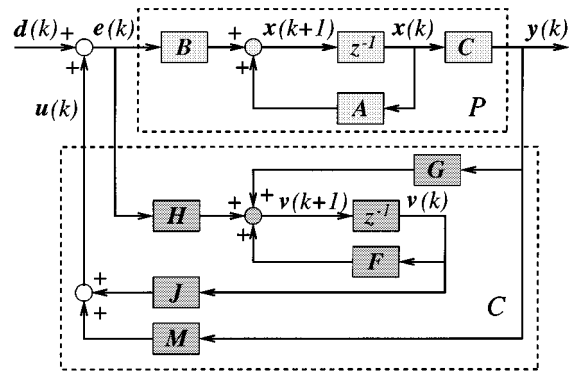


Fig. 1. Discrete-time closed-loop system with a generic digital controller.

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 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 nonconvex and nonsmooth 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 twofold. First, a generic controller structure is considered that includes output-feedback and observer-based controllers. Second, adopting the pole-sensitivity approach, a new stability related measure is proposed for the unified controller structure and an optimization procedure is developed to find the optimal controller realization that maximizes 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].

II. PROBLEM FORMULATION

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

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

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

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

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 $\mathbf{F} = \mathbf{A} - \mathbf{G}\mathbf{C}$, $\mathbf{M} = \mathbf{0}$ 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

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

where $\mathbf{T} \in \mathcal{R}^{m \times m}$ is any real-valued nonsingular matrix, called a similarity transformation. Let $\mathbf{w}_F = \text{Vec}(\mathbf{F})$, where $\text{Vec}(\cdot)$ 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

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

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

$$\begin{aligned} \bar{\mathbf{A}}(\mathbf{w}) &= \begin{bmatrix} \mathbf{A} + \mathbf{B}\mathbf{M}\mathbf{C} & \mathbf{B}\mathbf{J} \\ \mathbf{G}\mathbf{C} + \mathbf{H}\mathbf{M}\mathbf{C} & \mathbf{F} + \mathbf{H}\mathbf{J} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{T}^{-1} \end{bmatrix} \bar{\mathbf{A}}(\mathbf{w}_0) \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{T} \end{bmatrix}. \end{aligned} \quad (5)$$

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

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

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 $\pm\epsilon/2$,

$$\|\Delta\mathbf{w}\|_{\max} \triangleq \max_{i \in \{1, \dots, N\}} |\Delta w_i| \leq \epsilon/2. \quad (7)$$

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

$$\epsilon = 2^{-B_f}. \quad (8)$$

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]

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

From the definition of $\mu_0(\mathbf{w})$, it is obvious that

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

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

$$\hat{B}_{s_0}^{\min} = B_i + \text{Int}[-\log_2(\mu_0(\mathbf{w}))] - 1 \quad (10)$$

where the integer $\text{Int}[x] \geq x$. It can easily be seen that the closed-loop system remains stable if \mathbf{w} is implemented with a fixed-point processor of $\hat{B}_{s_0}^{\min}$. 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 less conservative the estimation will be. The pole sensitivity measures [2], [3] can be regarded as such lower bounds.

III. 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 $|\lambda_i(\bar{\mathbf{A}}(\mathbf{w}))|$ are to 1 and how sensitive they are to the controller parameter perturbations. We propose the following FWL stability related measure¹

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

with

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

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

$$\|\mathbf{x}\|_1 \triangleq \sum_{i=1}^s |x_i| \quad (13)$$

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

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

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

$$\begin{aligned} \mathcal{P}(\mathbf{w}) &\triangleq \{\Delta\mathbf{w}: |\lambda_i(\bar{\mathbf{A}}(\mathbf{w} + \Delta\mathbf{w}))| - |\lambda_i(\bar{\mathbf{A}}(\mathbf{w}))| \\ &\leq \|\Delta\mathbf{w}\|_{\max} \bar{\sigma}_i(\mathbf{w}), \quad \forall i\} \end{aligned} \quad (15)$$

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

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

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: $\mathcal{Q}(\mathbf{w}) \triangleq \{\Delta\mathbf{w}: \mu(\Delta\mathbf{w}) < \mu_0(\mathbf{w})\}$, i.e., those $\Delta\mathbf{w}$ that will not cause the closed-loop instability. Similar to [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

$$\rho(\mathcal{P}(\mathbf{w})) \triangleq \inf_{\Delta\mathbf{w} \notin \mathcal{P}(\mathbf{w})} \|\Delta\mathbf{w}\|_{\max}. \quad (16)$$

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. In practice, it is very difficult to verify the sufficient condition $\rho(\mathcal{P}(\mathbf{w})) > \mu_0(\mathbf{w})$, as this would require to know $\mu_0(\mathbf{w})$. However, the conditions for Proposition 2 are verifiable.

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

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

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

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

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

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

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

with T denoting the transpose operator and

$$\mathbf{L}_i(\mathbf{w}) = \frac{\text{Re} [\lambda_i^*(\overline{\mathbf{A}}(\mathbf{w})) \mathbf{y}_i^*(\overline{\mathbf{A}}(\mathbf{w})) \mathbf{x}_i^T(\overline{\mathbf{A}}(\mathbf{w}))]}{|\lambda_i(\overline{\mathbf{A}}(\mathbf{w}))|} \quad (22)$$

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

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

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 defined as [3]

$$\mu_1(\mathbf{w}) \triangleq \min_{i \in \{1, \dots, m+n\}} \frac{1 - |\lambda_i(\overline{\mathbf{A}}(\mathbf{w}))|}{\sigma_i(\mathbf{w})} \quad (24)$$

with

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

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

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

The key difference between $\mu_{1I}(\mathbf{w})$ and $\mu_1(\mathbf{w})$ is that the former considers the sensitivity of $|\lambda_i(\overline{\mathbf{A}}(\mathbf{w}))|$ while the latter considers the sensitivity of $\lambda_i(\overline{\mathbf{A}}(\mathbf{w}))$. 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

$$\begin{aligned} & \frac{\partial |\lambda_i(\overline{\mathbf{A}}(\mathbf{w}))|}{\partial w_j} \\ &= \text{Re} \left[\lambda_i^*(\overline{\mathbf{A}}(\mathbf{w})) \frac{\partial \lambda_i(\overline{\mathbf{A}}(\mathbf{w}))}{\partial w_j} \right] / |\lambda_i(\overline{\mathbf{A}}(\mathbf{w}))| \end{aligned} \quad (27)$$

one has

$$\begin{aligned} & \left| \frac{\partial |\lambda_i(\overline{\mathbf{A}}(\mathbf{w}))|}{\partial w_j} \right| \\ & \leq \frac{\left| \lambda_i^*(\overline{\mathbf{A}}(\mathbf{w})) \frac{\partial \lambda_i(\overline{\mathbf{A}}(\mathbf{w}))}{\partial w_j} \right|}{|\lambda_i(\overline{\mathbf{A}}(\mathbf{w}))|} = \left| \frac{\partial \lambda_i(\overline{\mathbf{A}}(\mathbf{w}))}{\partial w_j} \right| \end{aligned} \quad (28)$$

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].

IV. 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}_{opt} that maximizes $\mu_{1I}(\mathbf{w})$, since the controller implemented with \mathbf{w}_{opt} can tolerate a maximum FWL error. This optimal realization problem is formally defined as

$$v \triangleq \max_{\mathbf{w} \in \mathcal{S}} \mu_{1I}(\mathbf{w}). \quad (29)$$

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

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

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

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

From (17)–(21), we have

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

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

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

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

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

where

$$\begin{aligned} \mathbf{L}_{i,j,l}(\mathbf{w}_0) &= \frac{\text{Re} [\lambda_i^*(\overline{\mathbf{A}}(\mathbf{w}_0)) \mathbf{y}_{i,j}^*(\overline{\mathbf{A}}(\mathbf{w}_0)) \mathbf{x}_{i,l}^T(\overline{\mathbf{A}}(\mathbf{w}_0))]}{|\lambda_i(\overline{\mathbf{A}}(\mathbf{w}_0))|}, \\ & j, l = 1, 2. \end{aligned} \quad (37)$$

Define the following cost function:

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

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

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

Although $f(\mathbf{T})$ is nonsmooth and nonconvex, efficient global optimization methods exist for solving for this kind of optimization 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}_{opt} of the problem (39). With \mathbf{T}_{opt} , we can obtain the optimal realization \mathbf{w}_{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].

V. AN ILLUSTRATIVE EXAMPLE

A numerical example is used to illustrate the design procedure and verify the theoretical results given in Section III. The plant model used is a modification of the plant studied in [2] which was a single-input-single-output (SISO) 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 (40), shown at the bottom of the page. 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

$$\begin{aligned} \mathbf{F}_0 &= \begin{bmatrix} 0 & 1 \\ -9.3303e-01 & 1.9319e+00 \end{bmatrix} \\ \mathbf{G}_0 &= \begin{bmatrix} 4.1814e-02 & 2.7132e+02 \\ 3.9090e-02 & 1.0167e+03 \end{bmatrix} \\ \mathbf{J}_0 &= [3.0000e-04 \quad 5.0000e-04] \\ \mathbf{M}_0 &= [0 \quad 6.1250e-01] \\ \mathbf{H}_0 &= \begin{bmatrix} 7.8047e+01 \\ 7.3849e+01 \end{bmatrix}. \end{aligned}$$

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 optimization problem (39) was then formed with $\mathbf{T} \in \mathcal{R}^{2 \times 2}$. The ASA algorithm was used to find a \mathbf{T}_{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}_{opt} , the corresponding optimal controller realization \mathbf{w}_{opt} was determined

$$\begin{aligned} \mathbf{F}_{\text{opt}} &= \begin{bmatrix} 9.8677e-01 & 1.4943e-02 \\ -2.9047e-02 & 9.4511e-01 \end{bmatrix} \\ \mathbf{G}_{\text{opt}} &= \begin{bmatrix} 1.7066e-03 & -1.8080e+03 \\ 5.2084e-04 & 8.3794e+02 \end{bmatrix} \\ \mathbf{J}_{\text{opt}} &= [1.1208e-02 \quad 2.4887e-02] \\ \mathbf{M}_{\text{opt}} &= [0 \quad 6.1250e-01] \\ \mathbf{H}_{\text{opt}} &= \begin{bmatrix} 1.0691e+00 \\ 1.9430e+00 \end{bmatrix}. \end{aligned}$$

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 I compares the values of the two stability related measures, corresponding estimated

TABLE I
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

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}_{opt}	$8.696940e-05$	$3.012354e-06$	24	29	21

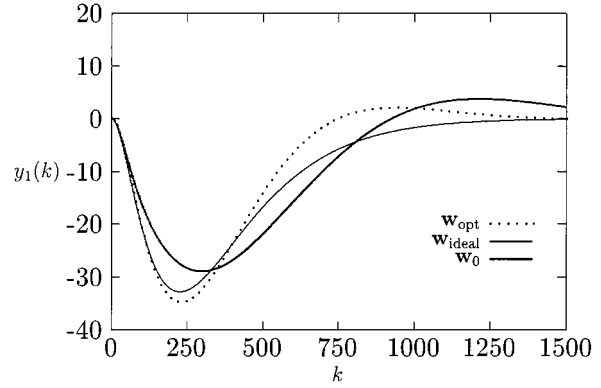


Fig. 2. Comparison of unit impulse response for the infinite-precision controller implementation $\mathbf{w}_{\text{ideal}}$ with those for the two 22-bit implemented controller realizations \mathbf{w}_0 and \mathbf{w}_{opt} .

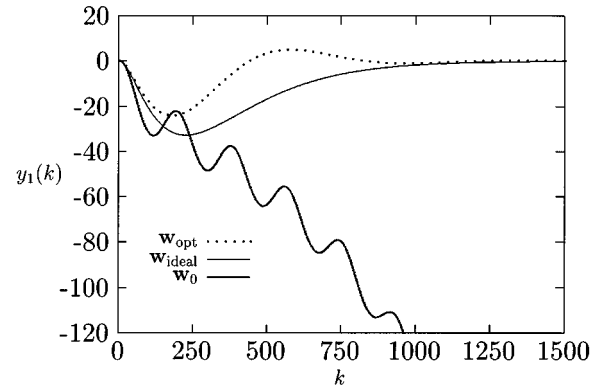


Fig. 3. Comparison of unit impulse response for the infinite-precision controller implementation $\mathbf{w}_{\text{ideal}}$ with those for the two 21-bit implemented controller realizations \mathbf{w}_0 and \mathbf{w}_{opt} .

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.

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} 3.2439e-01 & -4.5451e+00 & -4.0535e+00 & -2.7003e-03 & 0 \\ 1.4518e-01 & 4.9477e-01 & -4.6945e-01 & -3.1274e-04 & 0 \\ 1.6814e-02 & 1.6491e-01 & 9.6681e-01 & -2.2114e-05 & 0 \\ 1.1889e-03 & 1.8209e-02 & 1.9829e-01 & 1.0000e+00 & 0 \\ 6.1301e-05 & 1.2609e-03 & 1.9930e-02 & 2.0000e-01 & 1 \end{bmatrix} \\ \mathbf{B} &= \begin{bmatrix} 1.4518e-01 \\ 1.6814e-02 \\ 1.1889e-03 \\ 6.1301e-05 \\ 2.4979e-06 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 0 & 0 & 1.6188e+00 & -1.5750e-01 & -4.3943e+01 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}. \end{aligned} \quad (40)$$

We also computed the unit impulse response of the closed-loop control system when the controllers were the infinite-precision implemented w_0 and various FWL implemented realizations. Notice that any realization $w \in \mathcal{S}$, implemented in infinite precision, will achieve the exact performance of the infinite-precision implemented w_0 , which is the *designed* controller performance. For this reason, the infinite-precision implemented w_0 is referred to as the *ideal* controller realization w_{ideal} . Figs. 2 and 3 compares the unit impulse response of the first plant output $y_1(k)$ for the ideal controller w_{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 w_0 . However, the closed-loop system remained stable with the 21-bit implemented w_{opt} .

VI. CONCLUSION

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 maximizes the proposed measure. An efficient optimization strategy has been developed based on the ASA algorithm for optimizing a unified controller structure which includes output-feedback and observer-based controllers.

REFERENCES

- [1] M. Gevers and G. Li, *Parameterizations in Control, Estimation and Filtering Problems: Accuracy Aspects*. London, U.K.: Springer-Verlag, 1993.
- [2] G. Li, "On the structure of digital controllers with finite word length consideration," *IEEE Trans. Automat. Contr.*, vol. 43, pp. 689–693, May 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 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. Automat. Contr.*, vol. 44, pp. 2149–2153, Nov. 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. Automat. Contr.*, vol. 45, pp. 1550–1554, Aug. 2000.
- [6] I. J. Fialho and T. T. Georgiou, "On stability and performance of sampled data systems subject to word length constraint," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 2476–2481, Dec. 1994.
- [7] —, "Optimal finite wordlength digital controller realization," in *Proc. Amer. Control Conf.*, San Diego, CA, 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. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*. Philadelphia, PA: SIAM, 1994.
- [10] R. E. Skelton, T. Iwasaki, and K. M. Grigoriadis, *A Unified Algebraic Approach to Linear Control Design*. London, U.K.: Taylor and Francis, 1998.
- [11] T. Kailath, *Linear Systems*. Upper Saddle River, NJ: Prentice-Hall, 1980.
- [12] J. O'Reilly, *Observers for Linear Systems*. New York: Academic, 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," *Syst. Control Lett.*, 2000, submitted for publication.

Risk-Sensitive Decision-Theoretic Diagnosis

Mark A. Shayman and Emmanuel Fernandez-Gaucherand

Abstract—We consider the problem of determining the optimal sequence of tests for the discovery of a faulty component, where there is a random cost associated with testing a component. Our work is motivated by applications in telecommunications networks, e.g., location and isolation of faults (or intruders) in IP networks. A novel feature in our approach is that a risk-sensitive performance criterion is used in order to rank different competing schedules. Risk-sensitivity is incorporated through the use of an exponential utility function, and hence optimal schedules attain a trade-off between minimal expected costs and, e.g., a low variance about the achievable expected costs. We characterize optimal schedules both when the testing sequence is not subject to precedence constraints, and when it is subject to such constraints, given by an arbitrary partial order. For the case with precedence constraints, we show that our models can be analyzed via modular decompositions, as studied by Monma and Sidney

I. INTRODUCTION

The motivation for the work presented here comes from the problem of fault management for communication networks. An important element in many approaches to fault management is *sequential testing* [19]. Based on available network management data, a set of components (hardware or software) is identified as containing the potential root cause of the failure. Then the suspect components are tested sequentially until the defective component is identified. For the resulting scheduling problem, it is typically assumed that there is a single faulty component [the *mutually exclusive faults* (MEF) case], that the probability of component i being faulty is a known value p_i , and that there is a random cost C_i associated with testing it, and the goal is to minimize the expected sum of the testing costs. Under these assumptions, classical results apply and indicate that it is optimal to test in order of increasing ratios $E[C_i]/p_i$. This is sometimes referred to as the " C over p rule." There is a large literature on this problem and its extension to the case where there are precedence constraints on the testing sequence. See, e.g., [5], [35], [15], [26], [7], [11], [17], [1], [32], [32], [36], [16], and [24]. Analogous results are available on the problem in which the assumption of mutually exclusive faults is replaced by the assumption of independent faults, and a sequence of components are tested until the first faulty component is discovered at which time testing stops. This problem is referred to as the *independent faults* (INF) problem. See, e.g., [5], [27], [10], [22], [21], [13], [34], [23], and [28]. The " C over p " rule has been applied in network fault management in, e.g., [19], [3]. In the diagnosis problems we consider, a test either identifies a faulty component or eliminates it from suspicion. Diagnosis problems in which tests reveal only partial information concerning faults are considered in [8].

In the above approaches, the objective is to minimize the average sum of the testing costs. This may make sense for diagnostic problems that will be repeated many times under the same conditions—i.e., with the same model—such as the diagnosis of engine failures in a particular

Manuscript received May 1, 2000; revised January 28, 2001. Recommended by Associate Editor Q. Zhang. This work was supported in part by the National Science Foundation under Grant ECS-9626399, and in part by the Laboratory for Telecommunications Science under DoD Contract MDA90497C3015.

M. A. Shayman is with the Department of Electrical and Computer Engineering and Institute for Systems Research, University of Maryland, College Park, MD 20742 USA (e-mail: shayman@eng.umd.edu).

E. Fernandez-Gaucherand is with the Department of Electrical and Computer Engineering and Computer Science, University of Cincinnati, Cincinnati, OH 45221 USA (e-mail: emmanuel@eccecs.uc.edu).

Publisher Item Identifier S 0018-9286(01)06598-9.