

Stability issues of finite precision state estimate feedback controller realizations for discrete time systems¹

Jun Wu

National Lab. of Industrial Control Tech.
Institute of Advanced Process Control
Zhejiang University
Hangzhou, 310027, P. R. China
jwu@iipc.zju.edu.cn

Gang Li

The school of EEE
Nanyang Technological University
Singapore
egli@ntu.edu.sg

Jian Chu

National Lab. of Industrial Control Tech.
Institute of Advanced Process Control
Zhejiang University
Hangzhou, 310027, P. R. China
chuj@iipc.zju.edu.cn

Sheng Chen

Dep. of Electronics and Computer Sci.
University of Southampton
Highfield, Southampton SO17 1BJ, UK
sqc@ecs.soton.ac.uk

Robert Habib Istepanian

Dep. of Electrical and Computer Eng.
Ryerson Polytechnic University
Toronto, Ontario, Canada M5B 2K3
ristepan@ee.ryerson.ca

James F. Whidborne

Department of Mechanical Engineering
King's College London
London WC2R 2LS, UK
james.whidborne@kcl.ac.uk

Abstract

This paper present a new effective algorithm for the optimal realization of state estimate feedback controller structures for discrete time systems subject to Finite-Word-Length (FWL) constraints. The problem is formulated as a nonlinear programming to provide an easy and efficient optimization tool to solve such complex problem. Simulation results of the optimum realizations of state estimate feedback controller structures are presented to illustrate the effectiveness of the proposed strategy.

Index Terms—Finite word length, stability, discrete time system.

1 Introduction

The recent advances in fixed-point implementation of digital controllers such as the design of dedicated fixed-point Digital Signal Processors (DSP) and new Digital Control Processors (DCP) architectures made Finite Word Length (FWL) implementation an important issue in modern digital control engineering design appli-

cations. Improved control performance and increased levels of integration are especially important in many areas such as consumer electronic products, automotive and electro-mechanical control systems. This is because hardware controller implementation with fixed-point arithmetic offer the advantages of speed, memory space, cost and simplicity over floating-point arithmetic.

The FWL effects have been well studied in digital signal processing, especially in digital filter implementation since the 1970's [1]. The results have recently been extended to the study of FWL effects of digital controller on control systems. [2] studied the effects of FWL implemented digital controller on the degradation of an LQG cost function from a statistical point of view. [3] analyzed the effects of FWL implemented digital controller on the stability and performance of sampled data systems. A FWL stability measure μ_0 was presented in [3], but computing explicitly this measure seems very hard and is still an open problem. Based on the first order approximation, [4] and [5] developed two tractable FWL stability measures which are the lower bounds of μ_0 respectively.

In all these studies of FWL effects of digital controllers, the controllers are output feedback controllers.

¹This work was supported by Zhejiang Provincial Natural Science Foundation of China under Grant 699085

When $\mu(\Delta w) < \mu_1(w)$, from (6)–(8), we have

$$\begin{aligned} |\lambda_i(\bar{A}(w + \Delta w))| &\leq |\lambda_i(\bar{A}(w))| + |\Delta \lambda_i| \\ &\leq |\lambda_i(\bar{A}(w))| + \mu_1(w) \sum_{j=1}^N \left| \frac{\partial \lambda_i}{\partial w_j} \right| < 1 \end{aligned} \quad (9)$$

which means that the closed-loop system remains stable under the FWL error Δw . In other words, for a given realization w , the closed-loop stability can tolerate those FWL perturbations Δw , whose elements have magnitudes less than $\mu_1(w)$. Hence $\mu_1(w)$ can be taken as an FWL stability measure: the larger $\mu_1(w)$ is, the larger FWL errors the closed-loop system can tolerate.

For computation of $\mu_1(w)$, the following theorem is important.

Theorem 1: Suppose $n \times n$ square matrix $\bar{A} = M_1 X M_2 + M_3 X M_4$ is diagonalizable and has $\{\lambda_i\}$ as its eigenvalues, matrix $X \in R^{p \times q}$, matrix M_1, M_2, M_3 and M_4 has proper dimension respectively. Let x_i be a right eigenvector of \bar{A} corresponding to the eigenvalue λ_i . Denote $M_x = [x_1 \ x_2 \ \dots \ x_n]$ and $M_y = [y_1 \ y_2 \ \dots \ y_n] = M_x^{-H}$, where y_i is called the reciprocal left eigenvector corresponding to λ_i . Then

$$\begin{aligned} \frac{\partial \lambda_i}{\partial X} &= \begin{bmatrix} \frac{\partial \lambda_i}{\partial x_{11}} & \dots & \frac{\partial \lambda_i}{\partial x_{1q}} \\ \vdots & \dots & \vdots \\ \frac{\partial \lambda_i}{\partial x_{p1}} & \dots & \frac{\partial \lambda_i}{\partial x_{pq}} \end{bmatrix} \\ &= M_1^T y_i^H x_i^T M_2^T + M_3^T y_i^H x_i^T M_4^T \end{aligned} \quad (10)$$

where superscript 'H' denotes the transpose and conjugate operation, superscript 'T' denotes the transpose operation. ' y_i^H ' is conjugate to y_i .

Proof: Let α be a variable independent of M_1, M_2, M_3 and M_4 . It follows from $y_i^H x_i = 1$ that

$$\frac{\partial y_i^H}{\partial \alpha} x_i + y_i^H \frac{\partial x_i}{\partial \alpha} = 0 \quad (11)$$

Notice that $\bar{A} x_i = \lambda_i x_i$ and $\lambda_i = y_i^H \bar{A} x_i$. Hence,

$$\frac{\partial \lambda_i}{\partial \alpha} = \frac{\partial y_i^H}{\partial \alpha} \bar{A} x_i + y_i^H \frac{\partial \bar{A}}{\partial \alpha} x_i + y_i^H \bar{A} \frac{\partial x_i}{\partial \alpha} \quad (12)$$

It follows from (11) and $y_i^H \bar{A} = \lambda_i y_i^H$ that

$$\begin{aligned} \frac{\partial \lambda_i}{\partial \alpha} &= \left(\frac{\partial y_i^H}{\partial \alpha} \lambda_i x_i + \lambda_i y_i^H \frac{\partial x_i}{\partial \alpha} \right) + y_i^H \frac{\partial \bar{A}}{\partial \alpha} x_i \\ &= y_i^H \frac{\partial \bar{A}}{\partial \alpha} x_i \\ &= y_i^H M_1 \frac{\partial X}{\partial \alpha} M_2 x_i + y_i^H M_3 \frac{\partial X}{\partial \alpha} M_4 x_i \end{aligned} \quad (13)$$

Let $\alpha = x_{kj}$. Then,

$$\frac{\partial \lambda_i}{\partial \alpha} = (y_i^H M_1)_k (M_2 x_i)_j + (y_i^H M_3)_k (M_4 x_i)_j \quad (14)$$

where $(y_i^H M_1)_k, (M_2 x_i)_j, (y_i^H M_3)_k$ and $(M_4 x_i)_j$ is the k th element of $y_i^H M_1$, the j th element of $M_2 x_i$, the k th element of $y_i^H M_3$ and the j th element of $M_4 x_i$ respectively. This leads to (10).

Using Theorem 1, we have

$$\begin{aligned} \frac{\partial \lambda_i}{\partial A} &= [0 \ I_n] \begin{bmatrix} 0 & -J^T \\ 0 & I_n \end{bmatrix} y_i^H x_i^T \begin{bmatrix} 0 & 0 \\ -K^T & I_n \end{bmatrix} \begin{bmatrix} 0 \\ I_n \end{bmatrix} \\ &= [0 \ I_n] y_i^H x_i^T \begin{bmatrix} 0 \\ I_n \end{bmatrix} \end{aligned} \quad (15)$$

$$\begin{aligned} \frac{\partial \lambda_i}{\partial B} &= [0 \ I_n] \begin{bmatrix} 0 & -J^T \\ 0 & I_n \end{bmatrix} y_i^H x_i^T \begin{bmatrix} 0 & 0 \\ -K^T & I_n \end{bmatrix} \begin{bmatrix} I_p \\ 0 \end{bmatrix} \\ &= [0 \ I_n] y_i^H x_i^T \begin{bmatrix} 0 \\ -K^T \end{bmatrix} \end{aligned} \quad (16)$$

$$\begin{aligned} \frac{\partial \lambda_i}{\partial C} &= [I_q \ 0] \begin{bmatrix} 0 & -J^T \\ 0 & I_n \end{bmatrix} y_i^H x_i^T \begin{bmatrix} 0 & 0 \\ -K^T & I_n \end{bmatrix} \begin{bmatrix} 0 \\ I_n \end{bmatrix} \\ &= [0 \ -J^T] y_i^H x_i^T \begin{bmatrix} 0 \\ I_n \end{bmatrix} \end{aligned} \quad (17)$$

$$\frac{\partial \lambda_i}{\partial K} = [-B_0^T \ -B^T] y_i^H x_i^T \begin{bmatrix} 0 \\ I_n \end{bmatrix} \quad (18)$$

$$\frac{\partial \lambda_i}{\partial J} = [0 \ I_n] y_i^H x_i^T \begin{bmatrix} C_0^T \\ -C^T \end{bmatrix} \quad (19)$$

With $\frac{\partial \lambda_i}{\partial A}, \frac{\partial \lambda_i}{\partial B}, \frac{\partial \lambda_i}{\partial C}, \frac{\partial \lambda_i}{\partial K}$, and $\frac{\partial \lambda_i}{\partial J}$, $\mu_1(w)$ can be computed easily using (8). Based on $\mu_1(w)$, we compute

$$\hat{B}_{s1}^{\min} = \text{Int}(-\log_2 \mu_1(w) - 1 + B_X) \quad (20)$$

where $\text{Int}(x)$ rounds x to the nearest integer towards $+\infty$. From the analysis of this section, we know the closed loop system is still stable when w is implemented by a DCP of at least \hat{B}_{s1}^{\min} bits.

3 Optimal Realization

From the last section, we know that there are different realizations for a given $C(z)$, and the FWL stability measure $\mu_1(w)$ is a function of the realization w . Hence there is an interesting problem of finding out the realization such that $\mu_1(w)$ is maximized. This realization is called optimal realization in such a sense. The digital state-estimate feedback controller implemented with an optimal realization means the minimum hardware requirements in terms of less word length (i.e optimized controller data path hardware design) and such that the closed loop system remains stable.

Let x_{i0} be a right eigenvector of $\bar{A}(w_0)$ corresponding to the eigenvalue $\lambda_{i0} = \lambda_i(\bar{A}(w_0))$, y_{i0} be the reciprocal left eigenvector corresponding to x_{i0} . It is easy to see from (2) that $\forall i \in \{1, \dots, 2n\}$, the right eigenvector of

$\bar{A}(w)$ corresponding to the eigenvalue $\lambda_i(\bar{A}(w)) = \lambda_{i0}$ is $x_i = \begin{bmatrix} I_n & 0 \\ 0 & T^{-1} \end{bmatrix} x_{i0} \in C^{2n}$, and the reciprocal left eigenvector is $y_i = \begin{bmatrix} I_n & 0 \\ 0 & T^T \end{bmatrix} y_{i0} \in C^{2n}$. Applying (15)–(19), we have

$$\frac{\partial \lambda_i}{\partial A} = [0 \quad T^T] y'_{i0} x_{i0}^T \begin{bmatrix} 0 \\ T^{-T} \end{bmatrix} \quad (21)$$

$$\frac{\partial \lambda_i}{\partial B} = [0 \quad T^T] y'_{i0} x_{i0}^T \begin{bmatrix} 0 \\ -K_0^T \end{bmatrix} \quad (22)$$

$$\frac{\partial \lambda_i}{\partial C} = [0 \quad -J_0^T] y'_{i0} x_{i0}^T \begin{bmatrix} 0 \\ T^{-T} \end{bmatrix} \quad (23)$$

$$\frac{\partial \lambda_i}{\partial K} = [-B_0^T \quad -B_0^T] y'_{i0} x_{i0}^T \begin{bmatrix} 0 \\ T^{-T} \end{bmatrix} \quad (24)$$

$$\frac{\partial \lambda_i}{\partial J} = [0 \quad T^T] y'_{i0} x_{i0}^T \begin{bmatrix} C_0^T \\ -C_0^T \end{bmatrix}. \quad (25)$$

For complex matrix $X = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \cdots & \vdots \\ x_{m1} & \cdots & x_{mn} \end{bmatrix} \in C^{m \times n}$, denote

$$\|X\|_S = \sum_{i=1}^m \sum_{j=1}^n |x_{ij}| \quad (26)$$

We can describe the optimal FWL realization problem of state-estimate feedback controller as the optimization problem:

$$\begin{aligned} v &= \frac{1}{\max_w \mu_1(w)} = \min_{\substack{T \in R^{n \times n} \\ \det(T) \neq 0}} \max_{i \in \{1, \dots, 2n\}} \frac{\sum_{j=1}^N |\frac{\partial \lambda_i}{\partial w_j}|}{1 - |\lambda_{i0}|} \\ &= \min_{\substack{T \in R^{n \times n} \\ \det(T) \neq 0}} \max_{i \in \{1, \dots, 2n\}} g(A, B, C, K, J, i) \end{aligned} \quad (27)$$

where

$$g(A, B, C, K, J, i) = \frac{\|\frac{\partial \lambda_i}{\partial A}\|_S + \|\frac{\partial \lambda_i}{\partial B}\|_S + \|\frac{\partial \lambda_i}{\partial C}\|_S + \|\frac{\partial \lambda_i}{\partial K}\|_S + \|\frac{\partial \lambda_i}{\partial J}\|_S}{1 - |\lambda_{i0}|} \quad (28)$$

From (21)–(25), we can define

$$f(T) = \max_{i \in \{1, \dots, 2n\}} g(A, B, C, K, J, i) \quad (29)$$

which is a function of T . Then the optimal state-estimate feedback controller realization problem can be posed as

$$v = \min_{\substack{T \in R^{n \times n} \\ \det(T) \neq 0}} f(T) \quad (30)$$

the above problem is a nonconvex nonlinear programming problem. We intend to search for the minimum of problem (30) with iterative optimization methods, i.e. a sequence $\{T_0, T_1, T_2, \dots\}$ which converges to the minimum T_{opt} is generated. In the iterative procedure

we can neglect the constraint $\det T \neq 0$, i.e. we solve the problem

$$v = \min_{T \in R^{n \times n}} f(T) \quad (31)$$

with iterative methods. There are two reasons for us to do so:

- $\Omega = \{T \mid \det T = 0, T \in R^{n \times n}\}$ is a very small set in space $R^{n \times n}$. Hence the case is rare that the iterate T_i moves into Ω when we search the space $R^{n \times n}$ for $T_{opt} \notin \Omega$ by an iterative sequence from the start point $T_0 \notin \Omega$.
- Even if it happens that T_i moves into Ω in the iterative procedure, we can add a small perturbation τI_n to T_i such that $T_i + \tau I_n \notin \Omega$. This small perturbation would not affect the convergence of the iterative sequence to T_{opt} .

In this paper, the simplex search method is applied to solve problem (31) which is an unconstrained non-convex nonlinear programming problem. There are many existing optimization software which uses the simplex search method, for example, the *fmins* function in MATLAB Ver5.1 optimization toolbox. It is well known, even if for a nonconvex nonlinear programming problem, the simple search algorithm can always converge to a locally optimal point. In order to “globalize” the algorithm, we repeatedly run the algorithm starting from various initial point to obtain “randomized” solutions for v ; then pick the smallest solution obtained.

4 Illustrative Example

To show how the optimization approach presented in this paper can be used efficiently for the parameterization issues of optimal FWL state-estimate feedback controller realization with improved stability bounds and minimum word-length requirements. We consider an example given in [6] to confirm our theoretical results.

The discrete time plant is given by

$$P(z) = \frac{0.0022(z+1)^2}{(z-0.9588)(z-0.9231)(z-0.8763)}$$

a state space description of $P(z)$ is

$$\begin{aligned} A_0 &= \begin{bmatrix} 2.7582 & -2.5342 & 0.7756 \\ 1.0000 & 0 & 0 \\ 0 & 1.0000 & 0 \end{bmatrix}, B_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \\ C_0 &= [0.0022 \quad 0.0044 \quad 0.0022] \end{aligned}$$

Given the initial realization of the controller $C(z)$

$$A_{ini} = A_0, B_{ini} = B_0, C_{ini} = C_0,$$

$$K_{ini} = [0.4761 \quad -0.8183 \quad 0.3506],$$

$$J_{ini} = \begin{bmatrix} 118.2995 \\ 101.0891 \\ 81.8859 \end{bmatrix}$$

The corresponding transition matrix \bar{A} can then be formed using (2), from which the poles of the ideal closed loop system can be computed and given as

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \\ \lambda_5 \\ \lambda_6 \end{bmatrix} = \begin{bmatrix} 0.9067 \\ 0.8437 \\ 0.4532 \\ 0.7523 \\ 0.5761 \\ 0.6231 \end{bmatrix}$$

the corresponding eigenvectors x_{i0}, y_{i0} can be computed and hence problem (31) can be constructed. For problem (31), we get the solution using the simplex search method:

$$T_{opt} = \begin{bmatrix} 19.2056 & 1.9810 & -1.2562 \\ 8.6287 & 1.4078 & 8.9271 \\ -2.2033 & -1.0608 & 21.3406 \end{bmatrix}$$

and $v = 1.0525 \times 10^4$. The optimal realization corresponding to T_{opt} is

$$A_{opt} = \begin{bmatrix} 1.4973 & -0.0311 & -0.4183 \\ 0.7003 & 0.9125 & -0.5379 \\ 0.5937 & 0.1081 & 0.3484 \end{bmatrix},$$

$$B_{opt} = \begin{bmatrix} 0.1089 \\ -0.5619 \\ -0.0167 \end{bmatrix},$$

$$C_{opt} = [0.0754 \quad 0.0082 \quad 0.0835],$$

$$K_{opt} = [1.3101 \quad -0.5808 \quad -0.4223],$$

$$J_{opt} = \begin{bmatrix} 5.8678 \\ 5.8305 \\ 4.7327 \end{bmatrix}$$

The results for the initial realization and optimal realization are summarized in Table 1:

Table 1. Measures and stabilized word lengths

Realization	$\mu_1(w)$	\bar{B}_{s1}^{\min}
w_{ini}	1.9810×10^{-5}	22
w_{opt}	9.5012×10^{-5}	16

The comparative results clearly show that optimal realization w_{opt} needs only 16 bits (including fractional part and part of integer) and provides larger stability measure while the initial realization w_{ini} requires 22 bits (including fractional part and part of integer) with lower stability bound.

5 Conclusions

In this paper we have presented an efficient approach for the stability measure of state estimate feedback

controllers with FWL consideration. It has also been shown that the optimal realization problem for state estimate feedback controller with FWL consideration can be interpreted as a nonlinear programming problem. The computation of the relevant FWL optimization problem was solved using the simplex search algorithm to illustrate that such problem can be efficiently and easily computed using existing mathematical programming techniques. The theoretical results were verified using a numerical example which illustrate that the optimum realization based on the optimization method presented here greatly improves the stability robustness of the relevant controller realizations with minimum word-length characteristics compared to non-optimal realizations.

References

- [1] Roberts, R. A., and Mullis, C. T., 1987, Digital Signal Processing, (Addison-Wesley).
- [2] Moroney, P., Willsky, A. S., and Houpt, P. K., 1980, The digital implementation of control compensators: The coefficient wordlength issue. IEEE Transactions on Automatic Control, 25, pp. 621-630.
- [3] Fialho, I. J., and Georgiou, T. T., 1994, On stability and performance of sampled data systems subject to word length constraint. IEEE Transactions on Automatic Control, 39, pp. 2476-2481.
- [4] Li, G., 1998, On the structure of digital controllers with finite word length consideration. IEEE Transactions on Automatic Control, 43, pp. 689-693.
- [5] Istepanian, R. H., Wu, J., Chu, J., and Whidborne, J. F., 1998, Maximizing lower bound stability measure of finite precision PID controller realization by nonlinear programming. Proceedings of 1998 American Control Conference, Philadelphia, USA, pp. 2596-2600.
- [6] Li, G., and Gevers, M., 1990, Optimal finite precision implementation of a state-estimate feedback controller. IEEE Transactions on Automatic Control, 37, pp. 1487-1498.
- [7] Gevers, M., and Li, G., 1993, Parametrizations in Control, Estimation and Filtering Problems: Accuracy Aspects, Communications and Control Engineering Series (London: Springer Verlag).
- [8] Istepanian, R. H., Wu, J., Whidborne, J. F., Yan, J., and Salcudean, S. E., 1998, Finite word length stability issues of a teleoperation motion-scaling control system. Proceedings of UKACC International Conference on Control'98, Swansea, UK, pp. 1676-1681.
- [9] Istepanian, R. H., Li, G., Wu, J., and Chu, J., 1998, Analysis of sensitivity measures of finite-precision digital controller structures with closed-loop stability bounds. Proceedings IEE Control Theory and Application, 145, pp. 472-478.