

Proc. 1<sup>st</sup> European Control Conf.,  
Grenoble, July 1991, pp. 544-549,

C57

## Minimization of Finite Worldlength Effects in Compensator Design

G. Li and M. Gevers

Technical Report  
91.68

This paper presents research results of the Belgian Programme on Interuniversity Poles of attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. The scientific responsibility rests with its authors.

is derived. It is shown that the FWL noise gain is the sum of a sensitivity and a roundoff noise gain of the closed loop system. Another contribution is to find the optimal compensator realizations in terms of minimizing the degradation of the closed loop performance due to the FWL implementation of the compensator. The optimal design in terms of minimizing the FWL noise gain is formulated under the usual dynamic range constraint on the compensator states. Due to the complexity of the FWL noise gain minimization problem, this gain is replaced by an upper bound, which is easy to minimize. An algorithm is given, which leads to an optimal compensator realization that minimizes this upper bound. The validity of this optimal design technique is confirmed with a numerical example.

An outline of this paper is as follows. In Section 2 we present a general control law design and formulate the finite precision compensator realization problem. Our first new contribution is in Section 3, where we give a synthetic analysis of the FWL effects on the desired plant output and derive a computable expression of what we call FWL noise gain. In Section 4 we formulate the optimal compensator design problem. A procedure for solving this minimization problem is given, from which the optimal compensator realizations can be found. Finally a computational example in Section 5 confirms the validity of this optimal FWL compensator design. Several realizations are compared. Some concluding remarks are given in Section 6.

## 2. STATE-SPACE DESIGN OF A COMPENSATOR

The aim of this paper is to analyze the effects of the FWL errors in the controller implementation on the overall performance of the controlled system, and to propose a controller implementation that minimizes the performance degradation due to these FWL errors. We shall therefore not be concerned with the design of the ideal (i.e. infinite precision) controller transfer function. We shall from now on assume that some design procedure (e.g. model reference, or LQG, or  $H_\infty$ ) has led to the specification of a 2-degree of freedom controller  $C_1(z)$ ,  $C_2(z)$  as represented in Fig 2.1. Here  $C_1(z)$  and  $C_2(z)$  denote the ideal (infinite precision) controller transfer functions obtained as a result of the control design calculations. The diagram of this control law is given as follows:

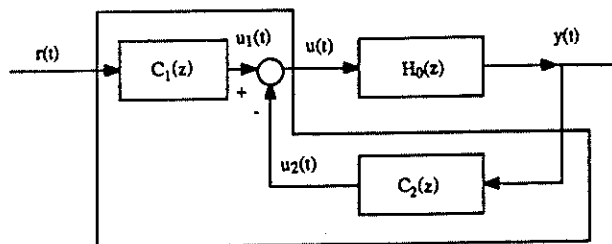


Fig.2-1

The designed closed loop transfer function is

$$H_c(z) = \frac{C_1(z)H_0(z)}{1 + H_0(z)C_2(z)} \quad (2.1)$$

The compensator is a MISO system:

$$u(t) = C_1(z) r(t) - C_2(z) y(t) \stackrel{\Delta}{=} C(z) \begin{bmatrix} r(t) \\ y(t) \end{bmatrix} \quad (2.2)$$

If  $(A_r, B_r, C_r, D_r)$  is a realization of the compensator  $C(z)$ , we have the following state-space model which must now be implemented in finite precision:

$$x_r(t+1) = A_r x_r(t) + B_r [r(t) \ y(t)]^T \quad (2.3a)$$

$$u(t) = C_r x_r(t) + D_r [r(t) \ y(t)]^T \quad (2.3b)$$

where  $A_r \in \mathbb{R}^{m \times m}$ ,  $B_r \in \mathbb{R}^{m \times 2}$ ,  $C_r \in \mathbb{R}^{1 \times m}$ ,  $D_r \in \mathbb{R}^{1 \times 2}$ ,  $m$  is the order of  $C(z)$ , and

$$C(z) = D_r + C_r(zI - A_r)^{-1}B_r \quad (2.4)$$

where  $B_r = [B_1 \ B_2]$  and  $D_r = [d_1 \ 0]$ .

It is easy to understand that there is a set of equivalent realizations for this compensator, which is given by  $S_c = \{(A_r, B_r, C_r, D_r) : (A_r, B_r, C_r, D_r) \text{ satisfies (2.4)}\}$ .

Consider the given plant specified by its transfer function  $H_0(z)$ . Even though the plant is not implemented in a computer, it will be useful to think of  $H_0(z)$  as being implemented by an infinite precision state-variable realization  $(A, B, C, d)$  in some coordinate system :

$$x(t+1) = A x(t) + B u(t)$$

$$y(t) = C x(t) + d u(t) \quad (2.5a)$$

$$\text{where } H_0(z) = d + C(zI - A)^{-1}B. \quad (2.5b)$$

The state model of the closed loop system is given by

$$\begin{aligned} \bar{x}(t+1) &= \bar{A} \bar{x}(t) + \bar{B} r(t) \\ y(t) &= \bar{C} \bar{x}(t) + \bar{d} r(t) \end{aligned} \quad (2.6)$$

where  $\bar{A}$ ,  $\bar{B}$ ,  $\bar{C}$  and  $\bar{d}$  are four explicit functions of  $(A, B, C, d)$  and  $(A_r, B_r, C_r, D_r)$ :

$$\bar{A} \stackrel{\Delta}{=} \begin{bmatrix} A & B C_r \\ B_2 C & A_r + d B_2 C_r \end{bmatrix}, \quad \bar{B} \stackrel{\Delta}{=} \begin{bmatrix} d_1 B \\ B_1 + d d_1 B_2 \end{bmatrix}$$

$$\bar{C} \stackrel{\Delta}{=} [C \ d C_r], \quad \bar{d} \stackrel{\Delta}{=} d d_1. \quad (2.7)$$

The desired transfer function is then given by

$$H_c(z) = \bar{d} + \bar{C}(zI - \bar{A})^{-1}\bar{B} \quad (2.8)$$

**Comment:** In practice, the compensator (2.3) has to be implemented with FWL, that is the ideal realization  $(A_r, B_r, C_r, D_r)$  of the compensator has to be replaced with its FWL version  $(A_r^*, B_r^*, C_r^*, D_r^*)$  in (2.8) and there will be a roundoff before or after arithmetical operations. As a result, the actual control signal and hence the actual output of the plant will differ from the desired ones.

The closed loop transfer function is a function of the realization  $(A_r, B_r, C_r, D_r)$  of the compensator. Clearly, if  $(A_r, B_r, C_r, D_r)$  belongs to  $S_c$ , so does  $(T^{-1}A_rT, T^{-1}B_r, C_rT, D_r)$  for any similarity transformation matrix  $T$ . In the infinite precision case, any compensator realization will yield the same closed loop transfer function, while this is no longer the case when the compensator is implemented with FWL. In fact, the degradation of the system performance depends strongly on the choice of coordinates for the state space realization of the compensator as will be shown later, and the optimal FWL state-space compensator design is to identify those realizations in the set  $S_c$ , which minimize some specified performance in the face of FWL errors

### 3. FWL EFFECT ANALYSIS

In this section, we first analyze the FWL effects of the compensator on the plant output. The optimal realization problem will be performed in the next section. With  $(A_r, B_r, C_r, D_r)$  replaced by its FWL version  $(A_r^*, B_r^*, C_r^*, D_r^*)$ , (2.3) becomes

$$x_r^*(t+1) = A_r^* x_r^*(t) + B_r^* [r(t) y^*(t)]^T \quad (3.1a)$$

$$u^*(t) = C_r^* x_r^*(t) + D_r^* [r(t) y^*(t)]^T \quad (3.1b)$$

where  $y^*(t)$  is the output  $y(t)$  of the plant with  $u(t)$  replaced by  $u^*(t)$ . When, additionally, the roundoff in arithmetic operations is also considered, the above model has to be modified. Now, let us consider a fixed-point implementation with roundoff before multiplication [12] which leads to the following actual computation model of the compensator:

$$x_r'(t+1) = A_r^* Q[x_r'(t)] + B_r^* [Q[r(t)] Q[y'(t)]]^T \quad (3.2a)$$

$$u'(t) = C_r^* Q[x_r'(t)] + D_r^* [Q[r(t)] Q[y'(t)]]^T \quad (3.2b)$$

where  $y'(t)$  is the actual output of the plant with  $u(t)$  replaced by  $Q[u(t)]$ ;  $Q[\cdot]$  denotes the quantization operation. If the signals and the parameters in  $(A_r^*, B_r^*, C_r^*, D_r^*)$  are implemented, respectively, with  $B_s$  and  $B_c$  bits, then the quantizer  $Q[\cdot]$  makes  $Q[x]$  have a  $B_s$  bit expression where  $x$  has  $(B_s + B_c)$  or more than  $(B_s + B_c)$  bits. Denote

$$e_x(t) = x(t) - Q[x'(t)], e_r(t) = r(t) - Q[r(t)] \quad (3.3a)$$

$$e_y(t) = y(t) - Q[y'(t)], e_u(t) = u(t) - Q[u'(t)] \quad (3.3b)$$

which are four roundoff noise processes. In numerical analysis, these processes are modelled as independent white noise having zero-mean and a variance  $\sigma_n^2 = (1/12)2^{-2B_s}$  [5] and [10]. From (3.2) and (3.3), it follows

$$x_r'(t+1) = A_r^* x_r'(t) + B_r^* [r(t) y'(t)]^T - A_r^* e_x(t) - B_r^* [e_r(t) e_y(t)]^T \quad (3.4a)$$

$$u'(t) = C_r^* x_r'(t) + D_r^* [r(t) y'(t)]^T - C_r^* e_x(t) - D_r^* [e_r(t) e_y(t)]^T \quad (3.4b)$$

Now, the degradation of the output of the closed loop system due to a FWL implementation of the compensator and to a quantization of the signals in the compensator can be measured by the difference between the desired and the actual output of the plant:

$$\begin{aligned} \Delta y(t) &= y(t) - y'(t) = [y(t) - y^*(t)] + [y^*(t) - y'(t)] \\ &= \Delta y^*(t) + \Delta y'(t). \end{aligned} \quad (3.5)$$

Evidently, the overall output error process  $\Delta y(t)$  can be separated into two processes  $\Delta y^*(t) = y(t) - y^*(t)$  and  $\Delta y'(t) = y^*(t) - y'(t)$ , which stand for the FWL effects on coefficients and arithmetic operations, respectively. Note that  $y^*(t)$  is the output of the plant when the control signal  $u^*(t)$  is computed with (3.1) and that  $y'(t)$  is the plant output when the control signal  $u(t)$  is replaced by  $Q[u'(t)]$  with  $u'(t)$  computed in (3.2). With some manipulations, one gets the following equations:

$$E_x(t+1) = \bar{A}^* E_x(t) + N(t) \quad (3.6a)$$

$$\Delta y'(t) = \bar{C}^* E_x(t) + d[C_r^* e_x(t) + d_1^* e_r(t) + e_u(t)] \quad (3.6b)$$

where

$$E_x(t) = \begin{bmatrix} x^*(t) \\ x_r^*(t) \end{bmatrix}, N(t) = \begin{bmatrix} B[C_r^* e_x(t) + d_1^* e_r(t) + e_u(t)] \\ A_r^* e_x(t) + B_1^* e_r(t) + B_2^* e_y(t) \end{bmatrix} \quad (3.7)$$

So,  $\Delta y'(t)$  can be obtained from the above equations. Now, consider the term  $\Delta y^*(t)$ . Let  $\{h_i\}$  and  $\{h_i^*\}$  be the impulse responses of the closed loop system with infinite and finite coefficient implementation of the compensator, respectively, then

$$\Delta y^*(t) = y(t) - y^*(t) = \sum_{i=0}^{t-1} [h_i - h_i^*] r(t-1-i) \quad (3.8)$$

Since  $N(t)$  and  $r(t)$  are independent, the steady-state variance of the plant output error  $\Delta y(t)$  can be computed by the following expression:

$$\sigma^2 = \lim_{t \rightarrow \infty} E[\Delta y(t)^2] = \sigma_1^2 + \sigma_2^2 \quad (3.9)$$

where  $\sigma_1^2 = \lim_{t \rightarrow \infty} E[\Delta y^*(t)^2]$  and  $\sigma_2^2 = \lim_{t \rightarrow \infty} E[\Delta y'(t)^2]$ .

The second term  $\sigma_2^2$  can be obtained from (3.6) and (3.7). It can be shown in [14] (see the journal version of this paper) that  $\sigma_2^2$  can be separated into two parts:

$$\sigma_2^2 = \{\text{tr}(W) + \rho^2\} \sigma_n^2 \quad (3.10)$$

where

$$W = (d^2 + B^T w_{11} B)(C_r^*)^T C_r^* + (A_r^*)^T w_{12}^T B C_r^* + C_r^* B^T w_{12} A_r^* + (A_r^*)^T w_{22} A_r^* \quad (3.11)$$

with

$$W_0 = \sum_{i=0}^{\infty} (\bar{A}^*)^i (\bar{C}^*)^T \bar{C}^* (\bar{A}^*)^i = \begin{bmatrix} w_{11} & w_{12} \\ w_{12}^T & w_{22} \end{bmatrix} \quad (3.12)$$

which is the observability Gramian of the closed loop system.

We note that  $W$  is coordinate dependent. If an initial compensator realization  $(A_r, B_r, C_r, D_r)$  yields  $W^0$ , the realization obtained from  $(A_r, B_r, C_r, D_r)$  by a similarity transformation matrix  $T$  will correspond to  $W = T^T W^0 T$ , while  $\rho$  in (3.10) is some constant value having nothing to do with the choice of coordinate system.

Let  $\{\rho_i\}$  denote the theoretical parameters of the realization of the

compensator. Their FWL version is  $\{p_i^*\}$  with  $p_i = p_i^* + \Delta p_i$  for all  $i$ . Clearly,

$$h_i - h_i^* = \frac{1}{2\pi j} \oint_{|z|=1} [H_c(z) - H_c^*(z)] z^{i-1} dz. \quad (3.13)$$

With a first order approximation, one has

$$\Delta H_c(z) = H_c(z) - H_c^*(z) = \sum_i \frac{\partial H_c}{\partial p_i} \Delta p_i \quad (3.14)$$

In an actual implementation, the coefficient perturbations  $\{\Delta p_i\}$  are determined if the ideal parameters and the number of bits are given. However, following [5], [9] and [10], we shall here adopt a statistical approach where the perturbations of the parameters are considered as random variables. Experiments have shown that the validity of this approach can be guaranteed. Furthermore, this approach can simplify the analysis and yield useful theoretical results. Usually,  $\{\Delta p_i\}$  are modelled as independent zero-mean random variables having a variance  $\sigma_c^2 = (1/12)2^{-2B_c}$  [5] and [10], that is,  $E[\Delta p_i] = 0$ ,  $E[\Delta p_i \Delta p_j] = \sigma_c^2 \delta(i-j)$ . Clearly,  $\Delta y(t)$  depends on the reference signal  $r(t)$ . In a general study,  $r(t)$  is assumed to be white noise with unit variance [12] and. With the assumptions above, one has

$$\sigma_1^2 = \lim_{t \rightarrow \infty} E[\Delta y^*(t)^2] = \sum_{i=1}^N \left\{ \frac{1}{2\pi j} \oint_{|z|=1} \left( \frac{\partial H_c}{\partial p_i} \right) \left( \frac{\partial H_c}{\partial p_i} \right)^H z^{-1} dz \right\} \sigma_c^2 \quad (3.15)$$

where  $N$  is the number of parameters of FWL in the compensator realization  $(A_r, B_1, B_2, C_r, d_1)$ . To compute (3.15), one needs the sensitivity functions  $\{\partial H_c / \partial p_i\}$  of the closed loop transfer function with respect to the parameters concerned. Now, let set up some notations and definitions about sensitivity functions and sensitivity measures.

**Definition 3.1 :** Let  $M \in \mathbb{R}^{n+m}$  be a matrix and let  $f(M) \in \mathbb{C}$  be a scalar complex function of  $M$ , differentiable w.r.t. all the elements of  $M$ . We then define

$$\frac{\partial f}{\partial M} = S \quad \text{with} \quad s_{ij} = \frac{\partial f}{\partial m_{ij}}$$

where  $s_{ij}$  denotes the  $(i,j)$ th element of a matrix  $S$ .

**Definition 3.2 :** Let  $f(z) \in \mathbb{C}^{n+m}$  be any complex matrix valued function of the complex variable  $z$ . We then define the  $l_p$ -norm of  $f(z)$  as

$$\|f\|_p = \left( \frac{1}{2\pi} \int_0^{2\pi} \|f(e^{j\omega})\|_F^p d\omega \right)^{1/p}$$

where  $\|f(e^{j\omega})\|_F$  is the Frobenius norm of the matrix  $f(e^{j\omega})$ .

By some manipulations, one obtains the following sensitivity matrix functions:

$$\frac{\partial H_c}{\partial A_r} = G_2(z) F_2^T(z), \quad \frac{\partial H_c}{\partial B_1} = G_2(z), \quad \frac{\partial H_c}{\partial B_2} = H_c(z) G_2(z),$$

$$\frac{\partial H_c}{\partial C_r^T} = (d + G^T [B^T \quad dB_2^T]^T) F_2, \quad \frac{\partial H_c}{\partial d_1} = \frac{H_0(z)}{1 + H_0(z) C_2(z)}$$

where  $F(z) = [F_1^T(z) \quad F_2^T(z)]^T = (zI - \bar{A})^{-1} \bar{B}$ ,

$$G(z) = [G_1^T(z) \quad G_2^T(z)]^T = (zI - \bar{A}^{-1})^{-1} \bar{C}^T. \quad (3.16)$$

So, with any realization of a given compensator, (3.15) can be computed theoretically. Assuming that the compensator realization concerned is fully parametrized, with the notations in Definition 3-2, (3.15) can be written as follows:

$$\sigma_1^2 = \left\{ \left\| \frac{\partial H_c}{\partial A_r} \right\|_2^2 + \left\| \frac{\partial H_c}{\partial B_1} \right\|_2^2 + \left\| \frac{\partial H_c}{\partial B_2} \right\|_2^2 + \left\| \frac{\partial H_c}{\partial C_r^T} \right\|_2^2 + \left\| \frac{\partial H_c}{\partial d_1} \right\|_2^2 \right\} \sigma_c^2 \\ = \text{tr}[W_{A_r} + W_{B_1} + W_{B_2} + W_{C_r} + W_{d_1}] \sigma_c^2 \quad (3.17a)$$

where

$$W_X = \frac{1}{2\pi j} \oint_{|z|=1} \left( \frac{\partial H_c}{\partial X} \right) \left( \frac{\partial H_c}{\partial X} \right)^H z^{-1} dz \quad \text{with } X = A_r, B_1, B_2, C_r, d_1. \quad (3.17b)$$

Note that  $W_{B_1}$ ,  $W_{B_2}$  and  $W_{C_r}$  can be computed from the corresponding weighted Gramians. Several algorithms are available for the computation of a weighted Gramian [13]. A method for computing  $W_{A_r}$  is available in [14]. As to  $W_{d_1}$ , it can easily be computed with residue theory but it is coordinate independent. Now, with (3.9), (3.10) and (3.17) one can compute the overall output error variance  $\sigma^2$  of the closed loop system. One notes that this variance can be separated into two parts, one is constant and independent of the choice of compensator coordinates, the other totally depends on the coordinate of the compensator. With a given realization of the compensator, one can evaluate the degradation of the closed loop performance due to the FWL implementation of the compensator, but the most interesting problem is to minimize this degradation by choosing optimal compensator realizations. Now, denote  $\Sigma^2(T)$  as the coordinate dependent part of the variance  $\sigma^2$  of (3.9) where "T" shows dependence of this value on the coordinates of the compensator. Then one has, using (3.10) and (3.17a):

$$\Sigma^2(T) = \text{tr}(W) \sigma_n^2 + \text{tr}[W_{A_r} + W_{B_1} + W_{B_2} + W_{C_r}] \sigma_c^2 \quad (3.18)$$

where  $\sigma_n^2$  and  $\sigma_c^2$  are the variances of the roundoff noise and of the statistical coefficient random noise both due to the FWL effects and dependent on the wordlengths of signals and coefficients, respectively. Usually, the two wordlengths are the same [7]. So, one can define a value  $G(T)$  called FWL noise gain as follows:

$$G(T) = \text{tr}[W + W_{A_r} + W_{B_1} + W_{B_2} + W_{C_r}] \quad (3.19)$$

Just like in digital filter design [1], [2], the state  $x_r^*(t)$  in (3.2a) has to be properly scaled in order to maintain its amplitudes within an

acceptable range, and hence to reduce the probability of overflow. If

$$W_c = \begin{bmatrix} W_c(1,1) & W_c(1,2) \\ W_c(2,1) & W_c(2,2) \end{bmatrix} = \sum_{i=0}^{\infty} \bar{A}^i \bar{B} \bar{B}^T (\bar{A})^i \quad (3.20)$$

is the controllability Gramian of the closed loop system (2.7), applying the commonly used  $l_2$ -scaling on  $x_r(t)$  can approximately be achieved with the following constraint:

$$(W_c(2,2))_{ii} = 1 \text{ for all } i. \quad (3.21)$$

So, the optimal FWL compensator design can be stated by the following constrained minimization problem:

$$\begin{aligned} \min G(T) &= \text{tr}\{W + W_{A_r} + W_{B_1} + W_{B_2} + W_{C_r}\} \\ (A_r, B_1, B_2, C_r, d_1) &\text{ in } S_c \text{ subject to (3.23)} \end{aligned} \quad (3.22)$$

In the next section, we will discuss this optimal design problem.

#### 4. OPTIMAL FWL COMPENSATOR REALIZATIONS

If an initial realization  $(A_r^0, B_1^0, B_2^0, C_r^0, d_1^0)$  in  $S_c$  corresponds to  $\{W^0, W_{A_r}^0, W_{B_1}^0, W_{B_2}^0, W_{C_r}^0, W_c^0(2,2)\}$ , the realization  $(A_r, B_1, B_2, C_r, d_1)$  obtained from this initial one by a similarity transformation will have  $\{T^T W^0 T, W_{A_r}(T), T^T W_{B_1}^0 T, T^T W_{B_2}^0 T, T^{-1} W_{C_r}^0 T^{-T}, T^{-1} W_c^0(2,2) T^{-T}\}$ . The difficulty in solving (3.22) is due to the fact that  $W_{A_r}(T)$  is a complicated function of  $T$ . To overcome this difficulty, we use an upper bound of  $G(T)$  as is often done in similar numerical analysis problems. It can be shown that

$$\begin{aligned} \text{tr}(W_{A_r}(T)) &= \left\| \frac{\partial H_c}{\partial A_r} \right\|_2^2 = \|G_2(z) F_2^T(z)\|_2^2 \leq \|G_2(z)\|_2^2 \|F_2(z)\|_2^2 \\ &= \text{tr}(W_{B_1}) \text{tr}(W_c(2,2)). \end{aligned}$$

With the constraint (3.21), one has

$$G(T) \leq \bar{G}(T) \triangleq \text{tr}\{T^{-1} W_{C_r}^0 T^{-T}\} + \text{tr}\{T^T [W^0 + (n+1)W_{B_1}^0 + W_{B_2}^0] T\}. \quad (4.1)$$

So, the optimal compensator realization in terms of minimizing the upper bound  $\bar{G}(T)$  of the FWL noise can be formulated as the following general mathematical problem:

$$\begin{aligned} \min \{S(T) &= \text{tr}\{T^{-1} M_1^0 T^{-T}\} + \text{tr}\{T^T M_2^0 T\}\} \\ T: \det T &\neq 0 \text{ and with } (T^{-1} M_2^0 T)_{ii} = 1 \text{ for all } i \\ \text{where } M_i^0 &= i = 1, 2, 3, \text{ are three given positive-definite matrices.} \end{aligned} \quad (4.2)$$

With  $M_1^0 = W_{C_r}^0$ ,  $M_2^0 = W^0 + (n+1)W_{B_1}^0 + W_{B_2}^0$  and  $M_3^0 = W_c^0(2,2)$ , the general problem (4.2) specializes to the minimization of the upper bound (4.1). We now present a solution to the problem defined by (4.2).

**THEOREM 4.1:** Let  $M_i^0$  be symmetric and positive-definite for  $i = 1, 2, 3$ . Then any solution  $T$  for (4.2) can be found from the following equations which have a unique solution  $(P, \lambda)$ :

$$\begin{aligned} \frac{\partial L}{\partial P} &= M_2^0 - P^{-1} M_1^0 P^{-1} - \lambda P^{-1} M_3^0 P^{-1} = 0, \\ \frac{\partial L}{\partial \lambda} &= \text{tr}\{M_3^0 P^{-1}\} - n = 0. \end{aligned} \quad (4.3)$$

where  $P = T T^T$ ,  $R(P) = \text{tr}\{M_1^0 P^{-1}\} + \text{tr}\{M_2^0 P\} = S(T)$  and  $L(P, \lambda) = R(P) + \lambda [\text{tr}\{M_3^0 P^{-1}\} - n]$ .

**Proof:** see [14].

It seems difficult to solve (4.3) analytically in order to get the optimal  $P_{opt}$ . Here we propose to use the following gradient algorithm to obtain the numerical solutions:

$$\begin{aligned} P(k+1) &= P(k) - \mu_1 \frac{\partial L(P, \lambda)}{\partial P} \Big|_{P=P(k), \lambda=\lambda(k)}, \\ \lambda(k+1) &= \lambda(k) - \mu_2 \frac{\partial L(P, \lambda)}{\partial \lambda} \Big|_{P=P(k), \lambda=\lambda(k)} \end{aligned} \quad (4.4)$$

where  $\mu_1$  and  $\mu_2$  are the (positive) step sizes. Clearly, when these step sizes are chosen to be small, the algorithm (4.4) always converges to the unique solution of (4.3).

**Comments:**

- 1) Once we get the optimal  $P_{opt}$ , using a SVD of  $P_{opt}$  we can obtain the optimal transformation matrices  $T_{opt}$ . An algorithm is available in [14].
- 2) In our optimal procedure, the performance is not the FWL noise gain  $G(T)$  but its upper bound. The question to be asked is how conservative this upper bound is. The optimal realizations obtained by minimizing this upper bound must therefore be validated. We provide some answer to these questions by a numerical example in the next section.

#### 5. NUMERICAL EXAMPLE

We now illustrate our previous theoretical results and give some answers to the questions posed at the end of the previous section with the following example.

The compensator is a 2-degree of freedom controller given by the following transfer functions,

$$\begin{aligned} C_1 &= \frac{0.4372 z^2 + 0.7406 z + 0.3512}{z^2 - 1.8628 z + 0.9316} \times 10^{-2} \text{ and} \\ C_2 &= \frac{-0.9250 z + 0.17700}{z^2 - 1.8628 z + 0.9316} \times 10^{-2}. \end{aligned}$$

One of the corresponding compensators is presented in the observable realization:

$$\begin{aligned} A_r &= \begin{bmatrix} 1.8628 & 1.0000 \\ -0.9316 & 0 \end{bmatrix}, B_r = \begin{bmatrix} 1.3873 & -0.9250 \\ 0.0278 & 1.7700 \end{bmatrix} \times 10^{-2} \\ C_r &= [1.0000 \quad 0], D_r = [0.3472 \quad 0] \times 10^{-2}. \end{aligned} \quad (5.1)$$

An optimal realization which minimizes the upper bound  $\overline{G}(T)$  of the FWL noise gain  $G(T)$  under the  $l_2$ -scaling is computed with the procedure given in the previous section. It yields:

$$A_{opt}^r = \begin{bmatrix} 0.9235 & 0.3096 \\ -0.2073 & 0.9393 \end{bmatrix}, B_{opt}^r = \begin{bmatrix} 0.5524 & 0.3808 \\ -0.1075 & 0.0891 \end{bmatrix}$$

$$C_{opt}^r = [0.0267 \quad -1.1529] \times 10^{-1} \quad (5.2)$$

Here, we also consider a MRH structure [1] and [2] of the compensator without  $l_2$ -scaling on the real compensator states and two  $l_2$ -scaled realizations obtained from the observable realization (5.1) and the MRH structure described above, respectively. The values of the FWL noise gain  $G(T)$  and its upper bound  $\overline{G}(T)$  for these realizations are presented in the following table:

Realization	$G(T)$	$\overline{G}(T)$
(5.1)	$3.0503 \times 10^3$	$3.0503 \times 10^3$
$l_2$ -scaled (5.1)	63.1506	63.1525
MRH structure	7.9193	7.9196
$l_2$ -scaled MRH structure	4.9193	4.9196
optimal (5.2)	4.9075	4.9078

#### Comments:

- 1) One notes that for all the realizations the FWL noise gain  $G(T)$  and its upper bound  $\overline{G}(T)$  are almost the same. This means that the upper bound is not very conservative at least for this example.
- 2) The important point is that the realization having a smaller FWL noise gain yields a smaller upper bound. This signifies that by minimizing the upper bound one can find those realizations which have a low FWL noise gain. So, our optimal procedure is validated, again for this example.
- 3) Clearly, the MRH does not give a nice performance while its  $l_2$ -scaled version does. But our optimal realization (5.2) is still better than the latter and the superiority of our optimal realization over the others always holds.

## 6. CONCLUSIONS

In this paper, we have considered the optimal parametrization problem for the FWL implementation of a compensator. The contribution of this paper has been twofold. The first was to derive a computable expression for a newly defined FWL noise gain of a closed-loop transfer function. The second contribution was to compute the set of optimal realizations, i.e. the set of realizations of the compensator that minimize the closed loop system performance degradation due to the FWL implementation of the compensator under dynamic range constraint. We have illustrated with a numerical example the nice performance that can be achieved by the optimal realization as compared to the observable realizations and the MRH structures with or without dynamic range constraint. The validity of our optimal design technique is confirmed by this example.

## REFERENCES

- [1] C.T. Mullis and R.A. Roberts (1976), "Filter structures which minimize roundoff noise in fixed-point digital filters", Proc. IEEE Int. Conf. Acoust. Speech and Signal Processing, pp. 505-508.
- [2] S.Y. Hwang (1977), "Minimum Uncorrelated Unit Noise in State-Space Digital Filtering" IEEE Trans. on Acoust. Speech and Signal Processing, Vol. ASSP-25, No 4 - Aug., pp.273-281.
- [3] V. Tavsanoglu and L. Thiele (1984), "Optimal Design of State-Space Digital Filters by Simultaneous Minimization of Sensitivity and Roundoff Noise", IEEE Trans. on Circuits and Systems, Vol. CAS-31, No 10 - Oct., pp. 884-888.
- [4] L. Thiele (1986), "On the Sensitivity of Linear State-Space Systems," IEEE Trans. on Circuits and Systems, Vol. CAS-33, No.5, May, pp.502-510.
- [5] P. Moroney, A.S. Willsky, and P.K. Houpt (1980), "The Digital Implementation of Control Compensators: The Coefficient Wordlength Issue," IEEE Trans. on Automatic Contr. Vol. AC-25, No.4, Aug. pp.621-630.
- [6] G. Li and M. Gevers (1990), "Optimal Finite Precision Implementation of a State-Estimate Feedback Controller," IEEE Trans. on Circuits and Systems, Vol. CAS-37, No.12, Dec., pp.1487-1498.
- [7] D. Williamson and K. Kadiman (1990), "Optimal Finite Wordlength Linear Quadratic Regulation," IEEE Trans. on Auto. Control, Vol. AC-34, No.12, Dec., pp.453-460.
- [8] W.S. Gray (1989), "A Geometric Approach to The Parametric Sensitivity of Dynamical Systems," Ph.D. dissertation, Georgia Institute of Technology, Georgia, U.S.A.
- [9] J.B. Knowles and E.M. Olcayto (1968), "Coefficient Accuracy and Digital Filter response," IEEE Trans. on Circuits Syst., Vol. CAS-15, Mar. pp.31-41.
- [10] R.E. Crochiere (1974), "Digital Network Theory and Its Application to The Analysis and Design of Digital Filters," Ph.D. dissertation, Dept. of Elec. Eng., M.I.T. Apr.
- [11] K. Liu and R. Skelton (1990), "Optimal Controllers for Finite Wordlength Implementation," Proceeding of American Control Conference, San Diego.
- [12] Darrell Williamson (1986), "Roundoff Noise Minimization and Pole-Zero Sensitivity in Fixed-Point Digital Filters Using Residue Feedback", IEEE Trans. on Acoustics, Speech and Signal Processing, Vol-ASSP-34, No 5 - Oct., pp. 1210-1220.
- [13] G. Li (1990), "Finite Precision Aspects in the Parametrizations of Control, Estimation, and Filtering Problems," Ph.D. dissertation, Louvain University, Belgium.
- [14] G.Li and M.Gevers (1990), "Optimal FWL State-Space Compensator Design," Technical Report, Unité AUTO, Louvain University, Belgium.