OPTIMAL FWL DESIGN OF STATE-SPACE DIGITAL SYSTEMS WITH WEIGHTED SENSITIVITY MINIMIZATION AND SPARSENESS CONSIDERATION

B.D.O. Anderson, G. Li and M. Gevers

ABSTRACT

The optimal Finite Word Length (FWL) state-space digital system design problem is investigated. It is argued that it may be desirable to minimize a frequency weighted sensitivity measure over all similarity transformations. The set of optimal realizations minimizing this weighted sensitivity is completely characterized. It is shown that a subset of the optimal realization set consists of sparse Schur realizations. Some nice properties of the Schur realizations are discussed.

Key Words: Finite wordlength, weighted sensitivity measure, Schur realizations.

1. INTRODUCTION

Much attention has recently been paid to the Finite Word Length (FWL) effects in digital system design.[1]-[5]. It is well known that any linear system can be represented by state-space equations and that this state-space model is not unique. In the infinite precision case, all these realizations are equivalent since they yield one and the same transfer function. But the important fact is that different realizations do have different numerical properties such as sensitivity and error propagation. This means that they are no longer equivalent in the finite precision case. The optimal FWL state-space design is to identify those realizations which minimize the degradation of the system performance due to the FWL effects.

In [3] the sensitivity behavior of a transfer function at one frequency point is considered to be as important as at another frequency point. From a practical point of view, we are usually interested in the performance of the transfer function within a specified frequency range, the bandwidth of the transfer function, for example. To achieve this, we define a weighted sensitivity function, and hence a corresponding measure in this paper. The optimal FWL design procedure for a frequency weighted sensitivity measure is given.

From practical considerations, it is desirable that the filter have a nice performance as well as a minimal number of coefficients to be implemented. Noting the fact that the optimal realizations minimizing this sensitivity measure are unique only up to an orthogonal similarity transformation, we further propose the use of Schur realizations within this class of optimal realizations. These Schur realizations have several advantages. They are sparse, and hence require fewer multiplications. Some other useful properties of these realizations are discussed from a practical FWL implementation point of view.

2. WEIGHTED SENSITIVITY MEASURE OF A REALIZATION

In this paper we are concerned with a discrete linear time-invariant single input, single output system H(z) given by a minimal state-space realization:

$$x(k+1) = A\,x(k) + B\,u(k) \tag{2.1a}$$
$$y(k) = C\,x(k) + d\,u(k) \tag{2.1b}$$

with A in $\mathbb{R}^{n\times n}$, B in $\mathbb{R}^n$, C^T in $\mathbb{R}^n$ and d in $\mathbb{R}$. The transfer function can be expressed in terms of state matrices as

$$H(z) = C(zI - A)^{-1}B + d\ . \tag{2.2}$$

We now define a realization set S_H of this system as follows: $S_H = \{ (A, B, C, d)$: (A, B, C, d) satisfies (2.2) $\}$. Clearly, if (A, B, C, d) belongs to S_H, so does ($T^{-1}AT$, $T^{-1}B$, CT, d) for any similarity transformation T. This means that S_H is an infinite set. In practice it is impossible to realize the coefficients in (A, B, C, d) exactly due to Finite Word Length (FWL) constraints. Since different realizations have different sensitivity behaviors as will be shown later, the optimal FWL state-space design is to search for those realizations that minimize the sensitivity in some proper measure.

Definition 2.1: Let $M \in \mathbb{R}^{n\times m}$ be a matrix and let $f(M) \in C$ be a scalar complex function of M, differentiable w.r.t. all the elements of M. We then define

$$\frac{\partial f}{\partial M} = S \ \text{ with } \ s_{ij} \overset{\Delta}{=} \frac{\partial f}{\partial m_{ij}} \tag{2.3}$$

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

With these notations it is easy to show that

$$\frac{\partial H(z)}{\partial A} = G(z)F^T(z)\,,\ \frac{\partial H(z)}{\partial B} = G(z)\ ,\ \frac{\partial H(z)}{\partial C^T} = F(z) \tag{2.4}$$

where

$$F(z) \overset{\Delta}{=} (zI - A)^{-1}B\,,\ G^T(z) \overset{\Delta}{=} C(zI - A)^{-1}. \tag{2.5}$$

Note that the direct feedthrough term d and the sensitivity function w.r.t. it are coordinate independent, so they have nothing to do with the optimal realization problem and hence are not considered in the analysis.

Definition 2.2: Let $f(z) \in C^{n\times m}$ be any complex matrix valued function of the complex variable z. We then define the l_p-norm of f(z) as

$$\| f \|_P \overset{\Delta}{=} \left(\frac{1}{2\pi} \int_0^{2\pi} \| f(e^{j\omega}) \|_F^P d\omega \right)^{1/P} \tag{2.6a}$$

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

$$\| f(e^{j\omega}) \|_F = \left(\sum_{i=1}^{n} \sum_{j=1}^{m} | f_{ij}(e^{j\omega}) |^2 \right)^{1/2} = \{ tr[f(e^{j\omega}) f^H(e^{j\omega})] \}^{1/2} \tag{2.6b}$$

where $f^H(e^{j\omega}) = f^T(e^{-j\omega})$.

Note that the measure in Definition 2.2 is in fact a frequency independent mean value of a matrix function in the whole frequency range. Therefore, this measure considers the sensitivity behavior of the transfer function at one frequency point to be as important as at another frequency point. It is, however, usually the case that we are only interested in the performance of the transfer function in a specified frequency band or even at some discrete frequency points. This type of requirement can be met by using frequency weighting techniques. This observation leads to the definition of a weighted sensitivity and hence a weighted sensitivity measure.

Let $W_A(z)$, $W_B(z)$ and $W_C(z)$ be three scalar functions of the complex variable z. The weighted sensitivity functions corresponding to those given in (2.4) are defined as

$$\frac{\delta H(z)}{\delta X} \overset{\Delta}{=} W_X(z) \frac{\partial H(z)}{\partial X} \quad \text{with } X = A, B, C. \tag{2.7}$$

The overall frequency weighted sensitivity measure of the transfer function H(z) w.r.t. the parameters in the realization A, B, C is defined as

$$M_a^* \overset{\Delta}{=} \| \frac{\delta Hz)}{\delta A} \|_1^2 + \| \frac{\delta Hz)}{\delta B} \|_2^2 + \| \frac{\delta H(z)}{\delta C^T} \|_2^2 \ . \tag{2.8}$$

The use of different norms in the overall sensitivity measure above is justified by the analytic properties of the first term on the right (2.8), which allow one to perform an analytic minimization procedure [3] and [5]. Now let $W_A(z) = W_1(z)W_2(z)$ be any factorization of $W_A(z)$. One then has the following explicit expressions

$$\frac{\delta H(z)}{\delta A} = [W_1(z)G(z)]\,[W_2(z)F(z)]^T \overset{\Delta}{=} G_1(z)\, F_2(z)^T \tag{2.9a}$$

$$\frac{\delta H(z)}{\delta B} = W_B(z)\, G(z) \overset{\Delta}{=} G_3(z)\, , \quad \frac{\delta H(z)}{\delta C^T} = W_C(z)\, F(z) \overset{\Delta}{=} F_4(z)\, . \tag{2.9b}$$

A similarity transformation x = Tz transforms (A, B, C, $F_2(z)$, $F_4(z)$, $G_1(z)$, $G_3(z)$) into ($T^{-1}AT$, $T^{-1}B$, CT, $T^{-1}F_2(z)$, $T^{-1}F_2(z)$, $T^TG_1(z)$, $T^TG_3(z)$). This means that different realizations yield different sensitivity measures M_a^*. So an interesting problem is to find those realizations that minimize this sensitivity measure. The optimal FWL state-space design can then be formulated as follows:

$$\min_{(A,\, B,\, C) \text{ in } S_H} M_a^* \tag{2.10}$$

3. OPTIMAL FWL REALIZATIONS

The difficulty in solving (2.10) is due to the fact that the first term on the right of (2.8) is a very complicated function of the realization (A, B, C). To overcome this, note that by the Cauchy-Schwartz inequality

$$\left\| \frac{\delta H(z)}{\delta A} \right\|_1^2 = \left\| G_1(z)F_2^T(z) \right\|_1^2 \leq \left\| G_1(z) \right\|_2^2 \left\| F_2(z) \right\|_2^2 \tag{3.1}$$

where equality holds if and only if

$$\rho^2 G_1^H(z)G_1(z) = F_2^H(z)F_2(z) \quad \forall\, z \in \{|z| = 1\}, \text{ for some } \rho \neq 0 \in \mathbb{R}. \tag{3.2}$$

We will study the following upper bound of M_a^*:

$$M_a^* \leq \bar{M}_a^* \overset{\Delta}{=} \left\| G_1(z) \right\|_2^2 \left\| F_2(z) \right\|_2^2 + \left\| \frac{\delta H(z)}{\delta B} \right\|_2^2 + \left\| \frac{\delta H(z)}{\delta C^T} \right\|_2^2 . \tag{3.3}$$

It is easy to show with (2.6) that

$$\bar{M}_a^* = tr(W_{o1})\, tr(W_{c2}) + tr(W_{o3}) + tr(W_{c4}) \tag{3.4}$$

where W_{o1}, W_{c2}, W_{o3} and W_{c4} can be obtained by the following general expression:

$$W = \frac{1}{2\pi j} \oint_{|z|=1} X(z)X^H(z)\, z^{-1}\, dz \tag{3.5}$$

with $X(z) = G_1(z)$, $F_2(z)$, $G_3(z)$ and $F_4(z)$, respectively.

We call these four matrices weighted Gramians. Several algorithms for computing a weighted Gramian are available in [6]. A similarity transformation $x = Tz$ transforms $(A, B, C, W_{c4}, W_{c2}, W_{o3}, W_{o1})$ into $(T^{-1}AT, T^{-1}B, CT, T^{-1}W_{c4}T^{-T}, T^{-1}W_{c2}T^{-T}, T^TW_{o3}T, T^TW_{o1}T)$. So, the optimal FWL design problem of (2.10) is replaced by the following upper bound minimization:

$$\min_{(A,B,C)\in S_H} \{\bar{M}_a^*(T) = tr(W_{o1}P)tr(W_{c2}P^{-1}) + tr(W_{o4}P) + tr(W_{c4}P^{-1}) \overset{\Delta}{=} R(P)\} \tag{3.6b}$$

where $P = TT^T$. Therefore,

$$\min_{T:\, detT\neq 0} \bar{M}_a^*(T) \quad \Longleftrightarrow \quad \min_{P:\, P=TT^T,\, detP\neq 0} R(P) \ . \tag{3.7}$$

The following theorem shows that the minimum of R(P) exists, and that it can be *achieved by non-singular P only. This means that (3.7) has solutions.*

Theorem 3.1: With W_{o3} and W_{c4} non-singular, the minimum of R(P) defined in (3.6b) exists and can be achieved only for non-singular P satisfying the following equation:.

$$P[tr(W_{c2}\, P^{-1})W_{o1} + W_{o3}]P = tr(W_{o1}P)W_{c2} + W_{c4}. \tag{3.8}$$

Proof: see the full version of this paper.

It seems difficult to obtain an explicit expression of the optimal P, that is the solution of equation (3.7), for a general case. However, P can be obtained by an iterative procedure using a gradient algorithm:

$$P(k+1) = P(k) - \mu \left. \frac{dR(P)}{dP} \right|_{P=P(k)} \tag{3.9}$$

where dR(P)/dP can be obtained with (3.6b) and μ is a (positive) step size.

The following theorem shows the uniqueness of the solution of (3.8). Since (3.8) is a necessary condition for (3.7) and since the minimization problem (3.7) has at least one solution, the uniqueness of the solution (3.8) shows that (3.8) is also a sufficient condition for the minimization problem (3.7) and hence guarantees the convergence of this algorithm to the unique solution.

Theorem 3.2: With the four symmetric positive-define matrices W_{o1}, W_{c2}, W_{o3} and W_{c4} given in (3.6), (3.8) has a unique solution, and hence so does (3.7).

Proof: see the full version of this paper.

Note that for any optimal $P = TT^T$, the corresponding optimal transformation matrices can be constructed as

$$T = P^{1/2} V \tag{3.10}$$

for any orthogonal matrix V. All the arguments above can be summarized by the following theorem:

Theorem 3.3: The optimal transformation matrices, that is the solutions of (3.6) are not unique. They can be characterized by (3.10) where P is determined by the system and by the weighting functions (it is the unique solution of (3.8)) while V is an arbitrary orthogonal matrix.

4. SCHUR REALIZATIONS IN DIGITAL FILTER DESIGN

From Theorem 3.3, one can see that the realizations determined by (3.6) form an optimal realization subset. This means that there exist some degrees of freedom in this equivalence subset. In [4], this freedom was used to find a Hessenberg realization in order to reduce the number of components to be implemented. Here, we investigate another realization, called Schur realization, which has some nice properties.

Definition 4.1: Let $(A^s, B^s, C^s) \in S_H$ be a realization of H(z). This realization is called Schur Realization if and only if the matrix A^s is of the following real Schur form:

$$A^s = \begin{bmatrix} A_{11} & x & . & . & x & . & x \\ 0 & A_{22} & . & . & x & . & x \\ . & . & . & . & . & . & . \\ . & . & . & . & A_{ii} & . & . \\ . & . & . & . & . & . & . \\ 0 & 0 & . & . & 0 & . & A_{mm} \end{bmatrix} \tag{4.1}$$

where each A_{ii} is either a 1-by-1 matrix or a 2-by-2 matrix having complex eigenvalues.

Any matrix A can be transformed into the real Schur form (4.1) by an orthogonal similarity transformation. It then follows from Theorem 3.3 that any non Schur realization in the optimal realization subset can be transformed into a Schur realization with an orthogonal similarity matrix. This Schur realization evidently keeps the minimal sensitivity property. From a practical point of view, a general (fully parametrized) optimal realization and its corresponding Schur realization can have different sensitivity behaviors even though they yield the same theoretical minimal sensitivity measure value.

We conclude that equivalent optimal realizations having the same theoretical minimal sensitivity measure could yield different actual sensitivity behaviors.

Theorem 4.1: For any transfer function of McMillan degree n, there exists a Schur realization (A^s, B^s, C^s) of the form (4.1) that belongs to the optimal realization subset and has at least n(n-1)/2 zero parameters.

Proof: see the full version of the paper.

The frequency response characteristics of a filter are determined by its pole-zero positions. This is why in digital filter design, the pole and zero behaviors are also taken as an important design criterion.

An outstanding property of a Schur realization is that its poles are determined only by the main block diagonal elements of A^s. This property allows one to analyze its pole sensitivity behavior easily. By applying an additional orthogonal similarity transformation to any optimal Schur realization, the actual pole sensitivity performance can be improved even though this orthogonal matrix will not change the theoretical pole sensitivity measure of the realization. For a detailed discussion, we refer to the journal version.

5. CONCLUSIONS

From practical considerations, we have defined a weighted sensitivity and sensitivity measure, which are a generalization of those in [3]. Our first contribution in this paper has been to derive a sufficient and necessary condition which characterizes all the optimal similarity transformations and to give an algorithm for solving the minimization problem. The second one is to propose the use of Schur realizations obtained from the optimal realization subset. Some properties of the Schur realizations have been revealed.

ACKNOWLEDGEMENT

The authors would like to thank J.E.Perkins and U. Helmke for many suggestions during the developement of this paper.

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] W.J. Lutz and S. Louis Hakimi (1988), "Design of Multi-Input Multi-Output Systems with Minimum Sensitivity", IEEE. Trans. on Circuits and Systems, Vol.-35, No 9 - Sept., pp 1114-1122.

[5] 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.

[6] G. Li (1990), "Finite Precision Aspects in the Parametrizations of Control, Estimation, and Filtering Problems," Ph.D. dissertation, Louvain University, Belgium.