PAPER

Optimal Synthesis of a Class of 2-D Digital Filters with Minimum L_2-Sensitivity and No Overflow Oscillations

Takao HINAMOTO[†a)], ***Fellow***, **Ken-ichi IWATA**[†b)], ***Nonmember***, **Osemekhian I. OMOIFO**[†c)], ***Student Member***, **Shuichi OHNO**[†d)], ***Member***, ***and*** **Wu-Sheng LU**[††e)], ***Nonmember***

SUMMARY The minimization problem of an L_2-sensitivity measure subject to L_2-norm dynamic-range scaling constraints is formulated for a class of two-dimensional (2-D) state-space digital filters. First, the problem is converted into an unconstrained optimization problem by using linear-algebraic techniques. Next, the unconstrained optimization problem is solved by applying an efficient quasi-Newton algorithm with closed-form formula for gradient evaluation. The coordinate transformation matrix obtained is then used to synthesize the optimal 2-D state-space filter structure that minimizes the L_2-sensitivity measure subject to L_2-norm dynamic-range scaling constraints. Finally, a numerical example is presented to illustrate the utility of the proposed technique.

key words: *L_2-sensitivity minimization, L_2-scaling constraints, no overflow oscillations, optimal synthesis, a class of 2-D state-space digital filters.*

1. INTRODUCTION

This paper is concerned with the optimal realization of a fixed-point 2-D state-space digital filter with finite word length (FWL). The efficiency and performance of the filter are directly influenced by selecting its state-space filter structure. When designing a transfer function with infinite accuracy coefficients so as to meet the filter specification requirements, and implementing it by a state-space model with a finite binary representation, the coefficients in the state-space model must be truncated or rounded to fit the FWL constraints. This coefficient quantization usually alters the characteristics of the filter and may change a stable filter to an unstable one. This motivates the study of the coefficient sensitivity minimization problem. In [1]-[10], two main classes of techniques have been proposed for constructing state-space digital filters that minimize the coefficient sensitivity, that is, L_1/L_2-sensitivity minimization [1]-[5] and L_2-sensitivity minimization [6]-[10]. It has been argued that the sensitivity measure based on the L_2 norm is more natural and reasonable relative to that based on the L_1/L_2-sensitivity minimization [6]-[10]. For 2-D state-space digital filters, the L_1/L_2-mixed sensitivity minimization problem [11]-[15] and L_2-sensitivity minimization problem [10],[16]-[19] have also been investigated. However, to our best knowledge, little has been done for the minimization of L_2-sensitivity subject to the L_2-norm dynamic-range scaling constraints for state-space digital filters [20], although it has been known that the use of scaling constraints can be beneficial for suppressing overflow oscillations [21],[22].

This paper investigates the problem of minimizing an L_2-sensitivity measure subject to L_2-norm dynamic-range scaling constraints for a class of 2-D state-space digital filters [23]. To this end, we introduce an expression for evaluating the L_2-sensitivity and formulate the L_2-sensitivity minimization problem subject to the L_2-norm dynamic-range scaling constraints. Next, the constrained optimization problem is converted into an unconstrained optimization problem by using linear-algebraic techniques. The unconstrained optimization problem is then solved using an efficient quasi-Newton algorithm [24]. A numerical example is presented to demonstrate that the proposed algorithm offers much reduced L_2-sensitivity.

Throughout $\boldsymbol{I}_n$ denotes the identity matrix of dimension $n \times n$. The transpose (conjugate transpose) of a matrix $\boldsymbol{A}$ and trace of a square matrix $\boldsymbol{A}$ are denoted by $\boldsymbol{A}^T$ ($\boldsymbol{A}^*$) and tr[$\boldsymbol{A}$], respectively. The ith diagonal element of a square matrix $\boldsymbol{A}$ are denoted by $(\boldsymbol{A})_{ii}$.

2. L_2-SENSITIVITY ANALYSIS

Consider a local state-space model $(\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{b}, \boldsymbol{c}_1, \boldsymbol{c}_2, d)_n$ for a class of 2-D recursive digital filters which is stable, locally controllable and locally observable [23]

$$\begin{bmatrix} \boldsymbol{x}(i+1, j+1) \\ y(i,j) \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_1 & \boldsymbol{A}_2 \\ \boldsymbol{c}_1 & \boldsymbol{c}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{x}(i, j+1) \\ \boldsymbol{x}(i+1, j) \end{bmatrix} + \begin{bmatrix} \boldsymbol{b} \\ d \end{bmatrix} u(i,j) \qquad (1)$$

Manuscript received June 24, 2005.
[†]The authors are with the Graduate School of Engineering, Hiroshima University, 1-4-1 Higashi-Hiroshima, 739-8527 Japan.
[††]Dept. of Elec. and Comp. Engineering, University of Victoria, Victoria, BC, Canada V8W 3P6
a) E-mail: hinamoto@hiroshima-u.ac.jp
b) E-mail: iwata@hiroshima-u.ac.jp
c) E-mail: osei@hiroshima-u.ac.jp
d) E-mail: ohno@hiroshima-u.ac.jp
e) E-mail: wslu@ece.uvic.ca

where $\boldsymbol{x}(i,j)$ is an $n \times 1$ local state vector, $u(i,j)$ is a scalar input, $y(i,j)$ is a scalar output, and $\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{b}, \boldsymbol{c}_1, \boldsymbol{c}_2$ and d are real constant matrices of appropriate dimensions. The transfer function of (1) is given by

$$H(z_1,z_2) = (z_1^{-1}\boldsymbol{c}_1 + z_2^{-1}\boldsymbol{c}_2) \cdot \left(\boldsymbol{I}_n - z_1^{-1}\boldsymbol{A}_1 - z_2^{-1}\boldsymbol{A}_2\right)^{-1}\boldsymbol{b} + d. \tag{2}$$

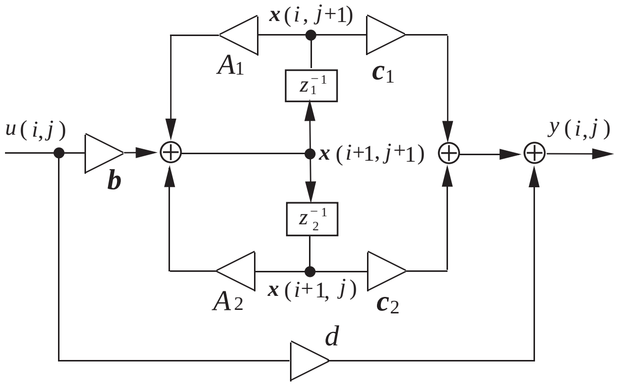

Fig. 1 A LSS model for 2-D filters.

A block diagram of the local state-space (LSS) model in (1) is shown in Fig. 1. It is interesting to note that

$$H^T(z_1,z_2) = \boldsymbol{b}^T\left(\boldsymbol{I}_n - z_1^{-1}\boldsymbol{A}_1^T - z_2^{-1}\boldsymbol{A}_2^T\right)^{-1} \cdot \left(z_1^{-1}\boldsymbol{c}_1^T + z_2^{-1}\boldsymbol{c}_2^T\right) + d \tag{3}$$

can be viewed as a transfer function of the Fornasini-Marchesini second LSS model [25]. Since $H(z_1,z_2) = H^T(z_1,z_2)$, the LSS model in (1) corresponds to a transposed structure of the Fornasini-Marchesini second LSS model.

Suppose that the LSS model in (1) is implemented by FWL fixed-point arithmetic with a B bit fractional representation, and is realized with coefficient matrices

$$\begin{aligned} \tilde{\boldsymbol{A}}_1 &= \boldsymbol{A}_1 + \Delta\boldsymbol{A}_1, & \tilde{\boldsymbol{A}}_2 &= \boldsymbol{A}_2 + \Delta\boldsymbol{A}_2 \\ \tilde{\boldsymbol{b}} &= \boldsymbol{b} + \Delta\boldsymbol{b}, & \tilde{d} &= d + \Delta d \\ \tilde{\boldsymbol{c}}_1 &= \boldsymbol{c}_1 + \Delta\boldsymbol{c}_1, & \tilde{\boldsymbol{c}}_2 &= \boldsymbol{c}_2 + \Delta\boldsymbol{c}_2 \end{aligned} \tag{4}$$

where $\Delta\boldsymbol{A}_1$, $\Delta\boldsymbol{A}_2$, $\Delta\boldsymbol{b}$, $\Delta\boldsymbol{c}_1$, $\Delta\boldsymbol{c}_2$, and Δd stand for the quantization errors of the coefficient matrices. Then, the transfer function of the FWL realization is expressed as

$$\tilde{H}(z_1,z_2) = (z_1^{-1}\tilde{\boldsymbol{c}}_1 + z_2^{-1}\tilde{\boldsymbol{c}}_2) \cdot \left(\boldsymbol{I}_n - z_1^{-1}\tilde{\boldsymbol{A}}_1 - z_2^{-1}\tilde{\boldsymbol{A}}_2\right)^{-1}\tilde{\boldsymbol{b}} + \tilde{d}. \tag{5}$$

Let $\{p_i\}$ be the set of the ideal parameters of a realization and let $\{\tilde{p}_i\}$ be its FWL version where $\tilde{p}_i = p_i + \Delta p_i$, and Δp_i indicates the corresponding parameter perturbation. If this realization has N parameters, then the first-order approximation of the Taylor series expansion yields

$$\begin{aligned} \Delta H(z_1,z_2) &= \tilde{H}(z_1,z_2) - H(z_1,z_2) \\ &= \sum_{i=1}^{N} \frac{\partial H(z_1,z_2)}{\partial p_i}\Delta p_i. \end{aligned} \tag{6}$$

It is obvious that the smaller $\partial H(z_1,z_2)/\partial p_i$ for $i = 1,2,\cdots,N$ yields the smaller transfer-function error $\Delta H(z_1,z_2)$. For a fixed-point implementation of B bits, the parameter perturbations can be considered to be independent random-variables uniformly distributed within the range $[-2^{-B-1},\ 2^{-B-1}]$. Then a measure of the transfer function error can statistically be defined by

$$\begin{aligned} \sigma^2_{\Delta H} &= \frac{1}{(2\pi j)^2}\oint_{|z_1|=1}\oint_{|z_2|=1} E[|\Delta H(z_1,z_2)|^2]\frac{dz_1 dz_2}{z_1 z_2} \\ &= \frac{1}{(2\pi)^2}\int_0^{2\pi}\int_0^{2\pi} E[|\Delta H(e^{j\omega_1},e^{j\omega_2})|^2]d\omega_1 d\omega_2 \end{aligned} \tag{7}$$

where $E(\cdot)$ denotes the ensemble-average operation. Since $\{\Delta p_i\}$ are independent random variables uniformly distributed, it follows that

$$E[|\Delta H(z_1,z_2)|^2] = \sum_{i=1}^{N}\left|\frac{\partial H(z_1,z_2)}{\partial p_i}\right|^2 \sigma^2 \tag{8}$$

where

$$\sigma^2 = E[(\Delta p_i)^2] = \frac{1}{12}2^{-2B}.$$

Definition 1: Let $\boldsymbol{X}$ be an $m \times n$ real matrix and let $f(\boldsymbol{X})$ be a scalar complex function of $\boldsymbol{X}$, differentiable with respect to all the entries of $\boldsymbol{X}$. The sensitivity function of f with respect to $\boldsymbol{X}$ is defined as

$$\boldsymbol{S}_{\boldsymbol{X}} = \frac{\partial f}{\partial \boldsymbol{X}}, \qquad (\boldsymbol{S}_{\boldsymbol{X}})_{ij} = \frac{\partial f}{\partial x_{ij}} \tag{9}$$

where x_{ij} denotes the (i,j)th entry of matrix $\boldsymbol{X}$.

From (2) and Definition 1, it can easily be shown that

$$\begin{aligned} \frac{\partial H(z_1,z_2)}{\partial \boldsymbol{A}_k} &= z_k^{-1}[\boldsymbol{F}(z_1,z_2)\boldsymbol{G}(z_1,z_2)]^T \\ \frac{\partial H(z_1,z_2)}{\partial \boldsymbol{b}} &= \boldsymbol{G}^T(z_1,z_2) \\ \frac{\partial H(z_1,z_2)}{\partial \boldsymbol{c}_k^T} &= z_k^{-1}\boldsymbol{F}(z_1,z_2), \qquad k = 1,2 \end{aligned} \tag{10}$$

where

$$\begin{aligned} \boldsymbol{F}(z_1,z_2) &= \left(\boldsymbol{I}_n - z_1^{-1}\boldsymbol{A}_1 - z_2^{-1}\boldsymbol{A}_2\right)^{-1}\boldsymbol{b} \\ \boldsymbol{G}(z_1,z_2) &= (z_1^{-1}\boldsymbol{c}_1 + z_2^{-1}\boldsymbol{c}_2) \cdot \left(\boldsymbol{I}_n - z_1^{-1}\boldsymbol{A}_1 - z_2^{-1}\boldsymbol{A}_2\right)^{-1}. \end{aligned}$$

The term d in (2) and its sensitivity are independent of the coordinate and therefore they are neglected here.

Definition 2: Let $\boldsymbol{X}(z_1, z_2)$ be an $m \times n$ complex matrix valued function of the complex variables z_1 and z_2. The L_2 norm of $\boldsymbol{X}(z_1, z_2)$ is defined as

$$\begin{aligned}
&||\boldsymbol{X}(z_1, z_2)||_2 \\
&= \left[\frac{1}{(2\pi)^2} \int_0^{2\pi} \int_0^{2\pi} \sum_{p=1}^{m} \sum_{q=1}^{n} \left| x_{pq}(e^{j\omega_1}, e^{j\omega_2}) \right|^2 d\omega_1 d\omega_2 \right]^{\frac{1}{2}} \\
&= \left(\mathrm{tr} \left[\frac{1}{(2\pi j)^2} \oint_{\Gamma_1} \oint_{\Gamma_2} \boldsymbol{X}(z_1, z_2) \boldsymbol{X}^*(z_1, z_2) \frac{dz_1 dz_2}{z_1 z_2} \right] \right)^{\frac{1}{2}}
\end{aligned}$$

where $\Gamma_i = \{z_i : |z_i| = 1\}$ for $i = 1, 2$.

From (10) and Definition 2, the overall L_2-sensitivity measure for the LSS model in (1) is evaluated by

$$\begin{aligned}
S &= \sigma^2_{\Delta H} / \sigma^2 \\
&= \sum_{k=1}^{2} \left\| \frac{\partial H(z_1, z_2)}{\partial \boldsymbol{A}_k} \right\|_2^2 + \left\| \frac{\partial H(z_1, z_2)}{\partial \boldsymbol{b}} \right\|_2^2 \\
&\quad + \sum_{k=1}^{2} \left\| \frac{\partial H(z_1, z_2)}{\partial \boldsymbol{c}_k^T} \right\|_2^2 \\
&= 2 \left\| [\boldsymbol{F}(z_1, z_2) \boldsymbol{G}(z_1, z_2)]^T \right\|_2^2 \\
&\quad + \left\| \boldsymbol{G}^T(z_1, z_2) \right\|_2^2 + 2 \left\| \boldsymbol{F}(z_1, z_2) \right\|_2^2 .
\end{aligned} \tag{11}$$

The L_2-sensitivity measure in (11) can be written as

$$S = 2\,\mathrm{tr}[\boldsymbol{M}] + \mathrm{tr}[\boldsymbol{W}_o] + 2\,\mathrm{tr}[\boldsymbol{K}_c] \tag{12}$$

where

$$\begin{aligned}
\boldsymbol{M} &= \frac{1}{(2\pi j)^2} \oint_{\Gamma_1} \oint_{\Gamma_2} [\boldsymbol{F}(z_1, z_2) \boldsymbol{G}(z_1, z_2)]^T \\
&\quad \cdot \boldsymbol{F}(z_1^{-1}, z_2^{-1}) \boldsymbol{G}(z_1^{-1}, z_2^{-1}) \frac{dz_1 dz_2}{z_1 z_2} \\
\boldsymbol{K}_c &= \frac{1}{(2\pi j)^2} \oint_{\Gamma_1} \oint_{\Gamma_2} \boldsymbol{F}(z_1, z_2) \boldsymbol{F}^T(z_1^{-1}, z_2^{-1}) \frac{dz_1 dz_2}{z_1 z_2} \\
\boldsymbol{W}_o &= \frac{1}{(2\pi j)^2} \oint_{\Gamma_1} \oint_{\Gamma_2} \boldsymbol{G}^T(z_1, z_2) \boldsymbol{G}(z_1^{-1}, z_2^{-1}) \frac{dz_1 dz_2}{z_1 z_2}.
\end{aligned}$$

Matrices $\boldsymbol{M}$, $\boldsymbol{K}_c$ and $\boldsymbol{W}_o$ are the 2-D Gramians and can be derived from

$$\begin{aligned}
\boldsymbol{M} &= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \boldsymbol{H}^T(i,j) \boldsymbol{H}(i,j) \\
\boldsymbol{K}_c &= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \boldsymbol{f}(i,j)\, \boldsymbol{f}^T(i,j) \\
\boldsymbol{W}_o &= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \boldsymbol{g}^T(i,j)\, \boldsymbol{g}(i,j)
\end{aligned} \tag{13}$$

where

$$\begin{aligned}
\boldsymbol{f}(i,j) &= \boldsymbol{A}^{(i,j)} \boldsymbol{b} \\
\boldsymbol{g}(i,j) &= \boldsymbol{c}_1 \boldsymbol{A}^{(i-1,j)} + \boldsymbol{c}_2 \boldsymbol{A}^{(i,j-1)} \\
\boldsymbol{A}^{(0,0)} &= \boldsymbol{I}_n, \qquad \boldsymbol{A}^{(i,j)} = \boldsymbol{0}, \;\; i < 0 \text{ or } j < 0 \\
\boldsymbol{A}^{(i,j)} &= \boldsymbol{A}_1 \boldsymbol{A}^{(i-1,j)} + \boldsymbol{A}_2 \boldsymbol{A}^{(i,j-1)} \\
&= \boldsymbol{A}^{(i-1,j)} \boldsymbol{A}_1 + \boldsymbol{A}^{(i,j-1)} \boldsymbol{A}_2, \;\; (i,j) > (0,0) \\
\boldsymbol{H}(i,j) &= \sum_{(0,0) \le (k,r) < (i,j)} \boldsymbol{f}(k,r)\, \boldsymbol{g}(i-k, j-r)
\end{aligned}$$

with the partial ordering for integer pairs (i,j) used in [26, p.2].

We remark that in practice the infinite sums in (13) are approximated with finite sums by truncation. The number of terms that should be used in each of the finite sums depends on how fast the associated series converges which is in turn dependent upon the stability margin of the filter involved. In principle it is advisable that, as long as the available computing resources permit, sufficiently many terms should be utilized in the evaluation so that the error introduced by the truncation becomes negligible.

3. L_2-SENSITIVITY MINIMIZATION

If a coordinate transformation defined by

$$\overline{\boldsymbol{x}}(i,j) = \boldsymbol{T}^{-1} \boldsymbol{x}(i,j) \tag{14}$$

is applied to the LSS model in (1), we obtain a new realization $(\overline{\boldsymbol{A}}_1, \overline{\boldsymbol{A}}_2, \overline{\boldsymbol{b}}, \overline{\boldsymbol{c}}_1, \overline{\boldsymbol{c}}_2, d)_n$ characterized by

$$\begin{aligned}
\overline{\boldsymbol{A}}_1 &= \boldsymbol{T}^{-1} \boldsymbol{A}_1 \boldsymbol{T}, \qquad \overline{\boldsymbol{A}}_2 = \boldsymbol{T}^{-1} \boldsymbol{A}_2 \boldsymbol{T} \\
\overline{\boldsymbol{b}} &= \boldsymbol{T}^{-1} \boldsymbol{b}, \qquad \overline{\boldsymbol{c}}_1 = \boldsymbol{c}_1 \boldsymbol{T}, \qquad \overline{\boldsymbol{c}}_2 = \boldsymbol{c}_2 \boldsymbol{T} \\
\overline{\boldsymbol{K}}_c &= \boldsymbol{T}^{-1} \boldsymbol{K}_c \boldsymbol{T}^{-T}, \qquad \overline{\boldsymbol{W}}_o = \boldsymbol{T}^T \boldsymbol{W}_o \boldsymbol{T}.
\end{aligned} \tag{15}$$

The coordinate transformation in (14) transforms the L_2-sensitivity measure in (12) to

$$S(\boldsymbol{T}) = 2\,\mathrm{tr}[\boldsymbol{M}(\boldsymbol{T})] + \mathrm{tr}[\overline{\boldsymbol{W}}_o] + 2\,\mathrm{tr}[\overline{\boldsymbol{K}}_c] \tag{16}$$

where

$$\boldsymbol{M}(\boldsymbol{T}) = \boldsymbol{T}^T \left[\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \boldsymbol{H}^T(i,j) \boldsymbol{T}^{-T} \boldsymbol{T}^{-1} \boldsymbol{H}(i,j) \right] \boldsymbol{T}.$$

Moreover, if L_2-norm dynamic-range scaling constraints are imposed on the local state vector $\overline{\boldsymbol{x}}(i,j)$, then

$$(\overline{\boldsymbol{K}}_c)_{ii} = (\boldsymbol{T}^{-1} \boldsymbol{K}_c \boldsymbol{T}^{-T})_{ii} = 1 \tag{17}$$

is required for $i = 1, 2, \cdots, n$.

The problem considered here is as follows: *Given* $\boldsymbol{A}_1$, $\boldsymbol{A}_2$, $\boldsymbol{b}$, $\boldsymbol{c}_1$ *and* $\boldsymbol{c}_2$, *obtain an* $n \times n$ *nonsingular matrix* $\boldsymbol{T}$ *which minimizes* $S(\boldsymbol{T})$ *in (16) subject to the scaling constraints in (17).*

When the LSS model in (1) is assumed to be stable and locally controllable, the local controllability Gramian $\boldsymbol{K}_c$ is symmetric and positive-definite [15]. This implies that $\boldsymbol{K}_c^{1/2}$ satisfying $\boldsymbol{K}_c = \boldsymbol{K}_c^{1/2}\boldsymbol{K}_c^{1/2}$ is also symmetric and positive-definite. Defining

$$\hat{\boldsymbol{T}} = \boldsymbol{T}^T \boldsymbol{K}_c^{-\frac{1}{2}}, \tag{18}$$

the scaling constraints in (17) can be expressed as

$$(\hat{\boldsymbol{T}}^{-T}\hat{\boldsymbol{T}}^{-1})_{ii} = 1, \qquad i = 1, 2, \cdots, n. \tag{19}$$

The constraints in (19) simply state that each column in $\hat{\boldsymbol{T}}^{-1}$ must be a unity vector. If matrix $\hat{\boldsymbol{T}}^{-1}$ is assumed to have the form

$$\hat{\boldsymbol{T}}^{-1} = \left[\frac{\boldsymbol{t}_1}{||\boldsymbol{t}_1||}, \frac{\boldsymbol{t}_2}{||\boldsymbol{t}_2||}, \cdots, \frac{\boldsymbol{t}_n}{||\boldsymbol{t}_n||}\right], \tag{20}$$

then (19) is always satisfied. From (18), it follows that (16) is changed to

$$\begin{aligned} J_o(\hat{\boldsymbol{T}}) = 2\,\mathrm{tr}[\hat{\boldsymbol{T}}\left[\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\hat{\boldsymbol{H}}^T(i,j)\hat{\boldsymbol{T}}^{-1}\hat{\boldsymbol{T}}^{-T}\hat{\boldsymbol{H}}(i,j)\right]\hat{\boldsymbol{T}}^T] \\ +\mathrm{tr}[\hat{\boldsymbol{T}}\hat{\boldsymbol{W}}_o\hat{\boldsymbol{T}}^T] + 2\,\mathrm{tr}[\hat{\boldsymbol{T}}^{-T}\hat{\boldsymbol{T}}^{-1}] \end{aligned} \tag{21}$$

where

$$\hat{\boldsymbol{H}}(i,j) = \boldsymbol{K}_c^{-\frac{1}{2}}\boldsymbol{H}(i,j)\boldsymbol{K}_c^{\frac{1}{2}}, \quad \hat{\boldsymbol{W}}_o = \boldsymbol{K}_c^{\frac{1}{2}}\boldsymbol{W}_o\boldsymbol{K}_c^{\frac{1}{2}}.$$

From the foregoing arguments, the problem of obtaining an $n \times n$ nonsingular matrix $\boldsymbol{T}$ which minimizes $S(\boldsymbol{T})$ in (16) subject to the scaling constraints in (17) can be converted into an unconstrained optimization problem of obtaining an $n \times n$ nonsingular matrix $\hat{\boldsymbol{T}}$ which minimizes $J_o(\hat{\boldsymbol{T}})$ in (21).

Now we apply a quasi-Newton algorithm [24] to minimize $J_o(\hat{\boldsymbol{T}})$ in (21) with respect to matrix $\hat{\boldsymbol{T}}$ given by (20). Let $\boldsymbol{x}$ be the column vector that collects the variables in matrix $\hat{\boldsymbol{T}}$, that is, $\boldsymbol{x} = [\boldsymbol{t}_1^T, \boldsymbol{t}_2^T, \cdots, \boldsymbol{t}_n^T]^T$. Then $J_o(\hat{\boldsymbol{T}})$ is a function of $\boldsymbol{x}$ and denoted by $J(\boldsymbol{x})$. The proposed algorithm starts with an initial point $\boldsymbol{x}_0$ obtained from an initial assignment $\hat{\boldsymbol{T}} = \boldsymbol{I}_n$. Then, in the kth iteration a quasi-Newton algorithm updates the most recent point $\boldsymbol{x}_k$ to point $\boldsymbol{x}_{k+1}$ as

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha_k \boldsymbol{d}_k \tag{22}$$

where

$$\begin{aligned} \boldsymbol{d}_k &= -\boldsymbol{S}_k \nabla J(\boldsymbol{x}_k) \\ \alpha_k &= arg\left[\min_{\alpha}\ J(\boldsymbol{x}_k + \alpha\boldsymbol{d}_k)\right] \\ \boldsymbol{S}_{k+1} &= \boldsymbol{S}_k + \left(1 + \frac{\boldsymbol{\gamma}_k^T\boldsymbol{S}_k\boldsymbol{\gamma}_k}{\boldsymbol{\gamma}_k^T\boldsymbol{\delta}_k}\right)\frac{\boldsymbol{\delta}_k\boldsymbol{\delta}_k^T}{\boldsymbol{\gamma}_k^T\boldsymbol{\delta}_k} \\ &\quad -\frac{\boldsymbol{\delta}_k\boldsymbol{\gamma}_k^T\boldsymbol{S}_k + \boldsymbol{S}_k\boldsymbol{\gamma}_k\boldsymbol{\delta}_k^T}{\boldsymbol{\gamma}_k^T\boldsymbol{\delta}_k} \\ \boldsymbol{S}_0 &= \boldsymbol{I}, \qquad \boldsymbol{\delta}_k = \boldsymbol{x}_{k+1} - \boldsymbol{x}_k \\ \boldsymbol{\gamma}_k &= \nabla J(\boldsymbol{x}_{k+1}) - \nabla J(\boldsymbol{x}_k). \end{aligned}$$

Here, $\nabla J(\boldsymbol{x})$ is the gradient of $J(\boldsymbol{x})$ with respect to $\boldsymbol{x}$, and $\boldsymbol{S}_k$ is a positive-definite approximation of the inverse Hessian matrix of $J(\boldsymbol{x})$. This iteration process continues until

$$|J(\boldsymbol{x}_{k+1}) - J(\boldsymbol{x}_k)| < \varepsilon \tag{23}$$

where $\varepsilon > 0$ is a prescribed tolerance. If the iteration is terminated at step k, then $\boldsymbol{x}_k$ is viewed as a solution point.

The implementation of (22) requires the gradient of $J(\boldsymbol{x})$. Closed-form expressions for $\nabla J(\boldsymbol{x})$ are given below.

$$\begin{aligned} \frac{\partial J(\hat{\boldsymbol{T}})}{\partial t_{pq}} &= \lim_{\Delta\to 0}\frac{J(\hat{\boldsymbol{T}}_{pq}) - J(\hat{\boldsymbol{T}})}{\Delta} \\ &= 4\beta_1 - 4\beta_2 + 2\beta_3 \end{aligned} \tag{24}$$

where $\hat{\boldsymbol{T}}_{pq}$ is the matrix obtained from $\hat{\boldsymbol{T}}$ with its (p,q)th component perturbed by Δ:

$$\begin{aligned} \hat{\boldsymbol{T}}_{pq} &= \hat{\boldsymbol{T}} + \frac{\Delta\hat{\boldsymbol{T}}\boldsymbol{g}_{pq}\boldsymbol{e}_q^T\hat{\boldsymbol{T}}}{1 - \Delta\boldsymbol{e}_q^T\hat{\boldsymbol{T}}\boldsymbol{g}_{pq}} \\ \beta_1 &= \boldsymbol{e}_q^T\hat{\boldsymbol{T}}\left[\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\hat{\boldsymbol{H}}^T(i,j)\hat{\boldsymbol{T}}^{-1}\hat{\boldsymbol{T}}^{-T}\hat{\boldsymbol{H}}(i,j)\right]\hat{\boldsymbol{T}}^T\hat{\boldsymbol{T}}\boldsymbol{g}_{pq} \\ \beta_2 &= \boldsymbol{e}_q^T\hat{\boldsymbol{T}}^{-T}\left[\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\hat{\boldsymbol{H}}(i,j)\hat{\boldsymbol{T}}^T\hat{\boldsymbol{T}}\hat{\boldsymbol{H}}^T(i,j)\right]\boldsymbol{g}_{pq} \\ \beta_3 &= \boldsymbol{e}_q^T\hat{\boldsymbol{T}}\hat{\boldsymbol{W}}_o\hat{\boldsymbol{T}}^T\hat{\boldsymbol{T}}\boldsymbol{g}_{pq} \\ \boldsymbol{g}_{pq} &= \partial\left\{\frac{\boldsymbol{t}_q}{||\boldsymbol{t}_q||}\right\}/\partial t_{pq} = \frac{1}{||\boldsymbol{t}_q||^3}(t_{pq}\boldsymbol{t}_q - ||\boldsymbol{t}_q||^2\boldsymbol{e}_p) \end{aligned}$$

where $\boldsymbol{e}_p$ denotes an $n\times 1$ unit vector whose pth element equals unity.

4. NUMERICAL EXAMPLE

Let the LSS model $(\boldsymbol{A}_1^o, \boldsymbol{A}_2^o, \boldsymbol{b}^o, \boldsymbol{c}_1^o, \boldsymbol{c}_2^o, d)_4$ in (1) for a class of 2-D digital filters be specified by

$$\boldsymbol{A}_1^o = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -0.00411 & 0.08007 & -0.42458 & 1.04460 \end{bmatrix}$$

$$\boldsymbol{A}_2^o = \begin{bmatrix} -0.22608 & 1.61428 & 0.10054 & -0.00723 \\ -0.40594 & 1.61040 & -0.60615 & 0.24580 \\ -0.30955 & 1.02336 & -0.45322 & 0.38668 \\ -0.14469 & 0.43872 & -0.31019 & 0.56289 \end{bmatrix}$$

$$\boldsymbol{b}^o = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}^T$$

$$\boldsymbol{c}_1^o = \begin{bmatrix} -0.01452 & 0.01234 & 0.02054 & 0.04762 \end{bmatrix}$$

$$\boldsymbol{c}_2^o = \begin{bmatrix} 0.01189 & 0.02351 & -0.00637 & 0.02094 \end{bmatrix}$$

$$d = 0.00943.$$

In this case, it follows from (13) that the Grammians $\boldsymbol{K}_c^o$, $\boldsymbol{W}_o^o$, and $\boldsymbol{M}^o$ are computed as

$$\boldsymbol{K}_c^o = 10^3 \begin{bmatrix} 1.525163 & 0.724615 & 0.352442 & 0.166126 \\ 0.724615 & 0.353199 & 0.176074 & 0.084130 \\ 0.352442 & 0.176074 & 0.092003 & 0.046055 \\ 0.166126 & 0.084130 & 0.046055 & 0.025386 \end{bmatrix}$$

$$\boldsymbol{W}_o^o = \begin{bmatrix} 0.008767 & -0.017767 & 0.005057 & -0.028288 \\ -0.017767 & 0.046357 & -0.023819 & 0.060846 \\ 0.005057 & -0.023819 & 0.230707 & -0.453548 \\ -0.028288 & 0.060846 & -0.453548 & 1.052716 \end{bmatrix}$$

$$\boldsymbol{M}^o = 10^4 \begin{bmatrix} 0.036099 & -0.078957 & 0.088707 & -0.261463 \\ -0.078957 & 0.184748 & -0.237832 & 0.633957 \\ 0.088707 & -0.237832 & 1.316013 & -2.769028 \\ -0.261463 & 0.633957 & -2.769028 & 6.163177 \end{bmatrix}$$

where the infinite sums in (13) were truncated with $(i, j) = (100, 100)$. The L_2-sensitivity measure is then derived from (12) as

$$S = 1.579936 \times 10^5$$

To perform the scaling such that (17) is satisfied, we apply the coordinate transformation matrix given by

$$\boldsymbol{T}_s = \mathrm{diag}\{39.053334,\ 18.793573,\ 9.591835,\ 5.038463\}$$

to the above original realization and obtain a scaled realization characterized by $(\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{b}, \boldsymbol{c}_1, \boldsymbol{c}_2, d)_4$ with

$$\boldsymbol{A}_1 = \begin{bmatrix} 0 & 0.481228 & 0 & 0 \\ 0 & 0 & 0.510378 & 0 \\ 0 & 0 & 0 & 0.525287 \\ -0.031857 & 0.298663 & -0.808282 & 1.044600 \end{bmatrix}$$

$$\boldsymbol{A}_2 = \begin{bmatrix} -0.226080 & 0.776837 & 0.024693 & -0.000933 \\ -0.843550 & 1.610400 & -0.309366 & 0.065898 \\ -1.260339 & 2.005100 & -0.453220 & 0.203118 \\ -1.121498 & 1.636435 & -0.590516 & 0.562890 \end{bmatrix}$$

$$\boldsymbol{b} = \begin{bmatrix} 0 & 0 & 0 & 0.198473 \end{bmatrix}^T$$

$$\boldsymbol{c}_1 = \begin{bmatrix} -0.567054 & 0.231913 & 0.197016 & 0.239932 \end{bmatrix}$$

$$\boldsymbol{c}_2 = \begin{bmatrix} 0.464344 & 0.441837 & -0.061100 & 0.105505 \end{bmatrix}$$

$$d = 0.00943.$$

This scaled realization $(\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{b}, \boldsymbol{c}_1, \boldsymbol{c}_2, d)_4$ will be used as the starting point in the simulations. For the scaled realization, it follows from (13) that the Grammians $\boldsymbol{K}_c$, $\boldsymbol{W}_o$, and $\boldsymbol{M}$ are calculated as

$$\boldsymbol{K}_c = \begin{bmatrix} 1.000000 & 0.987279 & 0.940868 & 0.844274 \\ 0.987279 & 1.000000 & 0.976755 & 0.888478 \\ 0.940868 & 0.976755 & 1.000000 & 0.952963 \\ 0.844274 & 0.888478 & 0.952963 & 1.000000 \end{bmatrix}$$

$$\boldsymbol{W}_o = 10 \begin{bmatrix} 1.337108 & -1.304050 & 0.189462 & -0.556646 \\ -1.304050 & 1.637345 & -0.429399 & 0.576183 \\ 0.189462 & -0.429399 & 2.122604 & -2.191942 \\ -0.556646 & 0.576183 & -2.191942 & 2.672484 \end{bmatrix}$$

$$\boldsymbol{M} = 10^3 \begin{bmatrix} 1.043052 & -1.097577 & 0.637299 & -0.982714 \\ -1.097577 & 1.238937 & -0.830495 & 1.153558 \\ 0.637299 & -0.830495 & 2.324302 & -2.574574 \\ -0.982714 & 1.153558 & -2.574574 & 3.019844 \end{bmatrix}$$

by trancating the infinite sums in (13) with $(i, j) = (100, 100)$. Then the L_2-sensitivity measure in (12) is computed as

$$S = 1.533797 \times 10^4.$$

Choosing $\hat{\boldsymbol{T}} = \boldsymbol{I}_n$ as the initial assignment, $J_o(\boldsymbol{I}_n) = 5.601499 \times 10^2$ was computed from (21). Applying the quasi-Newton algorithm in (22) for the minimization of $J_o(\hat{\boldsymbol{T}})$ in (21), it took 30 iterations to converge to the solution

$$\hat{\boldsymbol{T}} = \begin{bmatrix} 1.236271 & -0.693521 & -0.153525 & -0.306992 \\ 0.026339 & 1.274328 & -0.447779 & -0.256231 \\ 0.359964 & -0.434998 & 1.126999 & -0.142857 \\ 1.255885 & 0.180785 & 0.005688 & 0.573318 \end{bmatrix}$$

which yields

$$\boldsymbol{T} = \begin{bmatrix} 0.282749 & 0.429995 & 0.428701 & 1.113378 \\ 0.070152 & 0.442238 & 0.441342 & 0.989134 \\ -0.065510 & 0.248649 & 0.532508 & 0.893925 \\ -0.152303 & 0.070079 & 0.400554 & 0.889667 \end{bmatrix}.$$

In this case, the new realization $(\overline{\boldsymbol{A}}_1, \overline{\boldsymbol{A}}_2, \overline{\boldsymbol{b}}, \overline{\boldsymbol{c}}_1, \overline{\boldsymbol{c}}_2, d)_4$ in (15) is constructed as

$$\overline{\boldsymbol{A}}_1 = \begin{bmatrix} 0.318221 & 0.363292 & -0.214913 & -0.146352 \\ -0.014377 & 0.137344 & 0.565138 & -0.073837 \\ -0.081823 & -0.077103 & 0.167820 & 0.186979 \\ -0.013435 & 0.075530 & -0.037542 & 0.421215 \end{bmatrix}$$

$$\overline{\boldsymbol{A}}_2 = \begin{bmatrix} 0.537741 & -0.043781 & 0.150393 & 0.212192 \\ 0.088487 & 0.384327 & 0.015711 & 0.009448 \\ -0.224833 & 0.361068 & 0.119172 & -0.070477 \\ -0.093959 & -0.049634 & 0.142214 & 0.452750 \end{bmatrix}$$

$$\overline{\boldsymbol{b}} = \begin{bmatrix} -0.381083 & -0.363714 & -0.778565 & 0.537030 \end{bmatrix}^T$$

$$\overline{\boldsymbol{c}}_1 = \begin{bmatrix} -0.193514 & -0.075468 & 0.060275 & -0.012374 \end{bmatrix}$$

$$\overline{\boldsymbol{c}}_2 = \begin{bmatrix} 0.150222 & 0.387265 & 0.403791 & 0.993274 \end{bmatrix}$$

$$d = 0.00943$$

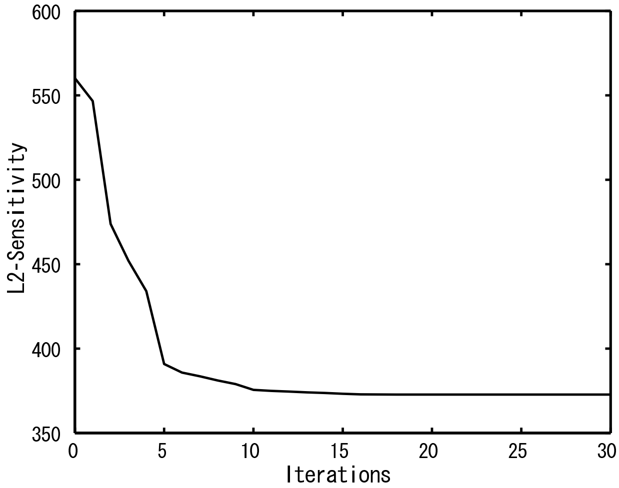

Fig. 2 L_2-sensitivity performance.

which yields

$$\overline{\boldsymbol{K}}_c = \begin{bmatrix} 1.000000 & 0.412457 & -0.000234 & -0.669182 \\ 0.412457 & 1.000000 & 0.546814 & -0.408864 \\ -0.000234 & 0.546814 & 1.000000 & -0.185690 \\ -0.669182 & -0.408864 & -0.185690 & 1.000000 \end{bmatrix}$$

$$\overline{\boldsymbol{W}}_o = \begin{bmatrix} 1.231416 & 0.433200 & -0.206292 & -0.537717 \\ 0.433200 & 0.877046 & 0.499273 & 0.360230 \\ -0.206292 & 0.499273 & 0.640216 & 0.820198 \\ -0.537717 & 0.360230 & 0.820198 & 2.412811 \end{bmatrix}$$

$$\boldsymbol{M}(\boldsymbol{T}) = 10\begin{bmatrix} 6.141820 & 2.122331 & -1.144342 & -3.890787 \\ 2.122331 & 3.507444 & 1.791768 & -0.662469 \\ -1.144342 & 1.791768 & 2.314498 & 1.729572 \\ -3.890787 & -0.662469 & 1.729572 & 6.016978 \end{bmatrix}.$$

The L_2-sensitivity measure in (16) is then minimized subject to the L_2-scaling constraints in (17) to

$$S(\boldsymbol{T}) = 372.776303997204.$$

The L_2-sensitivity performance of 30 iterations in (21) is shown in Fig. 2, from which it is observed that the iterative algorithm converges before 30 iterations where for $\varepsilon = 10^{-11}$, (23) was satisfied at $k = 26$.

The simulation results of applying the technique in [27] to this example can be found in Appendix. It is noted that the most significant 10 digits of the minimum L_2-sensitivity value $S(\boldsymbol{T}) = 372.776303997204$ shown in this section coincides with those of the minimum L_2-sensitivity value $S(\boldsymbol{P}) = 372.776303987459$ shown in Appendix. In addition, the optimal coordinate transformation matrix $\boldsymbol{T}$ in this section, denoted as $\boldsymbol{T}_{here}$, which minimizes the L_2-sensitivity measure subject to the L_2-scaling constraints is related to the corresponding optimal matrix $\boldsymbol{T}$ in Appendix, denoted as $\boldsymbol{T}_{[27]}$, by

$$\boldsymbol{T}_{[27]} = \boldsymbol{T}_{here}\boldsymbol{U}$$

where

$$\boldsymbol{U} = \begin{bmatrix} 0.212855 & 0.485601 & 0.833142 & 0.157347 \\ -0.268700 & -0.210930 & 0.014111 & 0.939739 \\ -0.292144 & -0.736880 & 0.552713 & -0.257229 \\ -0.892830 & 0.420366 & 0.013524 & -0.161137 \end{bmatrix}$$

and $\boldsymbol{U}$ is an orthogonal matrix satisfying $\boldsymbol{U}\boldsymbol{U}^T = \boldsymbol{U}^T\boldsymbol{U} = \boldsymbol{I}_4$. This means that

$$\boldsymbol{P} = \boldsymbol{T}_{[27]}\boldsymbol{T}_{[27]}^T = \boldsymbol{T}_{here}\boldsymbol{T}_{here}^T$$

is valid. For these reasons, we can conclude that the minimum L2-sensitivity value obtained by the proposed algorithm is practically identical to that of [27].

Concerning the computational complexity of the two algorithms, the algorithm in [27] took 2762 iterations and 311.528 seconds of CPU time on the AthlonXP 2500+ with clock 1.83 GHz and memory 480 MB to converge to the solution, while the identical solution was obtained by the proposed algorithm with 26 iterations and 94.736 seconds of CPU time on the same computer.

5. CONCLUSION

We have investigated the problem of minimizing the L_2-sensitivity measure subject to L_2-norm dynamic-range scaling constraints for a class of 2-D state-space digital filters. It has been shown that the L_2-sensitivity minimization problem subject to L_2-norm dynamic-range scaling constraints can be converted into an unconstrained optimization problem by using linear algebraic techniques. An efficient quasi-Newton algorithm has then been applied to solve the unconstrained optimization problem. The coordinate transformation matrix obtained has allowed us to construct the optimal 2-D state-space filter structure with minimum L_2-sensitivity and no overflow oscillations. Computer simulation results have demonstrated the effectiveness of the proposed technique.

It shoud be pointed out that the same problem was solved recently by relying on a Lagrange function [27]. The technique proposed in this paper can be viewed as an alternative mehod for soving the L_2-sensitivity minimization problem subject to L_2-scaling constraints.

References

[1] L. Thiele, "Design of sensitivity and round-off noise optimal state-space discrete systems," *Int. J. Circuit Theory A pl.*, vol. 12, pp.39-46, Jan. 1984.

[2] __________, "On the sensitivity of linear state-space systems," *IEEE Trans. Circuits Syst.*, vol.CAS-33, pp.502-510, May 1986.

[3] M. Iwatsuki, M. Kawamata and T. Higuchi, "Statistical sensitivity and minimum sensitivity structures with fewer coefficients in discrete time linear systems," *IEEE Trans. Circuits Syst.*, vol.37, pp.72-80, Jan. 1989.

[4] G. Li and M. Gevers, "Optimal finite precision implementation of a state-estimate feedback controller," *IEEE Trans. Circuits Syst.*, vol.37, pp.1487-1498, Dec. 1990.

[5] G. Li, B. D. O. Anderson, M. Gevers and J. E. Perkins, "Optimal FWL design of state-space digital systems with weighted sensitivity minimization and sparseness consideration," *IEEE Trans. Circuits Syst. I*, vol.39, pp.365-377, May 1992.

[6] W.-Y. Yan and J. B. Moore, "On L^2-sensitivity minimization of linear state-space systems," *IEEE Trans. Circuits Syst. I*, vol.39, pp.641-648, Aug. 1992.

[7] G. Li and M. Gevers, "Optimal synthetic FWL design of state-space digital filters," in *Proc. 1992 IEEE Int. Conf. Acoust., Speech, Signal Processing*, vol.4, pp.429-432.

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

[9] U. Helmke and J. B. Moore, *Optimization and Dynamical Systems*, Springer-Verlag, London, 1994.

[10] T. Hinamoto, S. Yokoyama, T. Inoue, W. Zeng and W.-S. Lu, "Analysis and minimization of L_2-sensitivity for linear systems and two-dimensional state-space filters using general controllability and observability Gramians," *IEEE Trans. Circuits S yst. I*, vol.49, pp.1279-1289, Sept. 2002.

[11] M. Kawamata, T. Lin and T. Higuchi, "Minimization of sensitivity of 2-D state-space digital filters and its relation to 2-D balanced realizations," in *Proc. 1987 IEEE Int. Symp. Circuits Syst.*, pp.710-713.

[12] T. Hinamoto, T. Hamanaka and S. Maekawa, "Synthesis of 2-D state-space digital filters with low sensitivity based on the Fornasini-Marchesini model," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.ASSP-38, pp.1587-1594, Sept. 1990.

[13] T. Hinamoto, T. Takao and M. Muneyasu, "Synthesis of 2-D separable-denominator digital filters with low sensitivity," *J. Franklin Institute*, vol.329, pp.1063-1080, 1992.

[14] T. Hinamoto and T. Takao, "Synthesis of 2-D state-space filter structures with low frequency-weighted sensitivity," *IEEE Trans. Circuits Syst. II*, vol.39, pp.646-651, Sept. 1992.

[15] T. Hinamoto and T. Takao, "Minimization of frequency-weighting sensitivity in 2-D systems based on the Fornasini-Marchesini second model, " in *1992 IEEE Int. Conf. Acoust., Speech, Signal Processing*, pp.401-404.

[16] G. Li, "On frequency weighted minimal L_2 sensitivity of 2-D systems using Fornasini-Marchesini LSS model", *IEEE Trans. Circuits Syst. I*, vol.44, pp.642-646, July 1997.

[17] G. Li, "Two-dimensional system optimal realizations with L_2-sensitivity minimization," *IEEE Trans. Signal Processing*, vol.46, pp.809-813, Mar. 1998.

[18] T. Hinamoto, Y. Zempo, Y. Nishino and W.-S. Lu, "An analytical approach for the synthesis of two-dimensional state-space filter structures with minimum weighted sensitivity," *IEEE Trans. Circuits Syst. I*, vol.46, pp.1172-1183, Oct. 1999.

[19] T. Hinamoto and Y. Sugie, "L_2-sensitivity analysis and minimization of 2-D separable-denominator state-space digital filters," *IEEE Trans. Signal Processing*, vol.50, pp.3107-3114, Dec. 2002.

[20] T. Hinamoto, H. Ohnishi and W.-S. Lu, "Minimization of L_2-sensitivity for state-space digital filters subject to L_2-scaling constraints," in *Proc. 2004 IEEE Int. Symp. Circuits Syst.*, vol.III, pp.137-140.

[21] C. T. Mullis and R. A. Roberts, "Synthesis of minimum roundoff noise fixed-point digital filters," *IEEE Trans. Circuits Syst.*, vol. 23, pp. 551-562, Sept. 1976.

[22] S. Y. Hwang, "Minimum uncorrelated unit noise in state-space digital filtering," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 25, pp. 273-281, Aug. 1977.

[23] T. Hinamoto, "A novel local state-space model for 2-D digital filters and its properties," in *Proc. 2001 IEEE Int. Symp. Circuits Syst.*, vol.2, pp.545-548.

[24] R. Fletcher, Practical Methods of Optimization, 2nd ed. Wiley, New York, 1987.

[25] E. Fornasini and G. Marchesini, "Doubly-indexed dynamical systems: State-space models and structural properties," *Math Syst. Theory*, vol.12, pp.59-72, 1978.

[26] R. P. Roessor, "A discrete state-space model for linear image processing," *IEEE Trans. Automat. Contr.*, vol.AC-20, pp.1-10, Feb. 1975.

[27] T. Hinamoto, K. Iwata and W.-S. Lu "Minimization of L_2-sensitivity for a class of 2-D state-space digital filters subject to L_2-scaling constraints," in *Proc. 2005 IEEE Int. Symp. Circuits Syst.*, pp.2401-2404.

Appendix

In the simulations of applying the technique reported in [27], the same scaled realization $(\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{b}, \boldsymbol{c}_1, \boldsymbol{c}_2, d)_4$ as in Section 4 was used as the starting point. In the same manner as in Section 4, the L_2-sensitivity measure of the scaled realization was computed as

$$S = 15337.965477.$$

It is noted that the previous L_2-sensitivity value $S = 14424.346809$ in [27] was calculated not using $\boldsymbol{H}(i,j)$ defined in (13), but using

$$\boldsymbol{H}(i,j) = \sum_{(0,0)\le(k,r)\le(i-1,j-1)}\sum \boldsymbol{f}(k,r)\,\boldsymbol{g}(i-k,j-r).$$

When choosing $\boldsymbol{P}_0 = \boldsymbol{I}_4$ and $\lambda_0 = 100$ as the initial estimates in [27], it took the iterative algorithm in [27] 2000 iterations to converge to

$$\boldsymbol{P} = \begin{bmatrix} 1.688238 & 1.500480 & 1.311958 & 1.149324 \\ 1.500480 & 1.373665 & 1.224597 & 1.077089 \\ 1.311958 & 1.224597 & 1.148785 & 1.035997 \\ 1.149324 & 1.077089 & 1.035997 & 0.980059 \end{bmatrix}$$

which yields

$$\boldsymbol{T} = \begin{bmatrix} -1.174654 & 0.198729 & 0.493644 & 0.158893 \\ -1.115961 & 0.031366 & 0.322000 & 0.153714 \\ -1.034448 & -0.100879 & 0.255344 & -0.057664 \\ -0.962590 & -0.009915 & 0.107523 & -0.204501 \end{bmatrix}$$

and the L_2-sensitivity measure is minimized subject to the L_2-scaling constraints $(\overline{\boldsymbol{K}}_c)_{ii} = (\boldsymbol{T}^{-1}\boldsymbol{K}_c\boldsymbol{T}^{-T})_{ii} = 1$ for $i = 1, 2, \cdots, n$ to

$$S(\boldsymbol{P}) = 372.776303987459.$$

The profiles of the L_2-sensitivity, parameter λ, as well as $\mathrm{tr}[\boldsymbol{K}_c\boldsymbol{P}^{-1}]$ during the first 2000 iterations of the algorithm are shown in Figs. 3 and 4, respectively, from which it is observed that the iterative algorithm converges before 2000 iterations. However, for a prescribed tolerance $\varepsilon = 10^{-11}$,

$$|J(\boldsymbol{P}_{i+1}, \lambda_{i+1}) - J(\boldsymbol{P}_i, \lambda_i)| < \varepsilon$$

was satisfied at $i = 2762$.

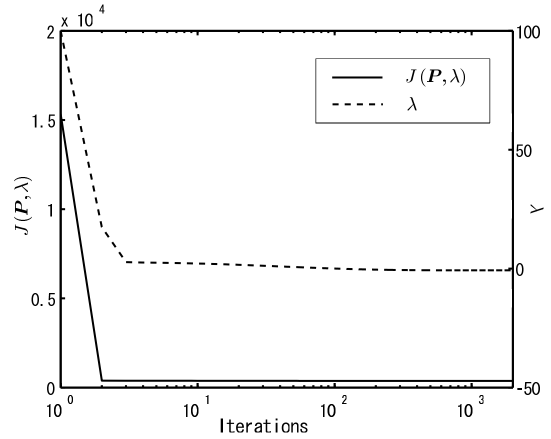

Fig. 3 L_2-sensitivity and λ performances.

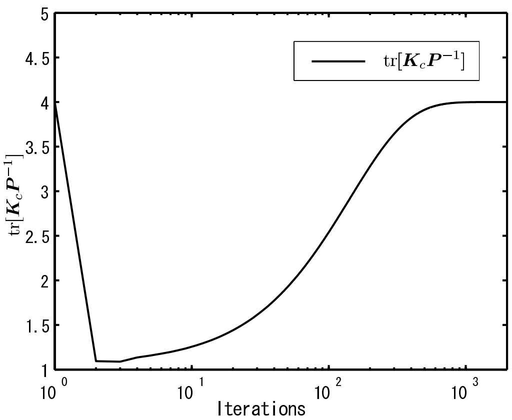

Fig. 4 $\mathrm{tr}[\boldsymbol{K}_c\boldsymbol{P}^{-1}]$ performance.

Takao HINAMOTO received the B.E. degree from Okayama University, Okayama, Japan, in 1969, the M.E. degree from Kobe University, Kobe, Japan, in 1971, and the Dr. Eng. degree from Osaka University, Osaka, Japan, in 1977, all in electrical engineering. From 1972 to 1988, he was with the Faculty of Engineering Kobe University, Kobe, Japan. From September 1979 to March 1981, he was on leave from Kobe University as a Visiting Member of Staff in the Department of Electrical Engineering, Queen's University, Kingston, Ontario, Canada. During 1988-1991, he was a Professor of electrical circuits in the Faculty of Engineering, Tottori University, Tottori, Japan. Since January 1992, he has been a Professor of electronic control in the Department of Electrical Engineering, Hiroshima University, Hiroshima, Japan. His teaching and research interests include digital signal processing, system theory, and control systems engineering. Dr. Hinamoto was an Associate Editor of the IEEE Transactions on CIRCUITS AND SYSTEMS II during 1993-1995, and an Associate Editor of the IEEE Transactions on CIRCUITS AND SYSTEMS I during 2002-2003.

Ken-ichi IWATA received the B.E. and M.E. degrees in electrical engineering from Hiroshima University, Hiroshima, Japan, in 2003 and 2005, respectively. He was engaged in research on digital signal processing during his graduate studies. Since April 2005, he has been with Mitsubishi Electric Corporation, Himeji, Japan.

Osemekhian I. OMOIFO received the B.E. degree in electrical/ electronics engineering from the University of Benin, Edo state, Nigeria, in 1998, the M.E. degree from Hiroshima University, Hiroshima, Japan, in 2004. His fields of interest are in signal processing (both analog and digital), filter design and stability analysis. He is currently a second year PhD student in engineering at Hiroshima University.

Shuichi OHNO received the B.E., M.E., and Dr. Eng. degrees in applied mathematics and physics from Kyoto University, Kyoto, Japan, in 1990, 1992, and 1995, respectively. From 1995 to 1999, he was a research associate with the Department of Mathematics and Computer Science, Shimane University, Shimane, Japan, where he became an Assistant Professor. He spent 14 months in 2000 and 2001 at the University of Minnesota, Minneapolis, as a visiting researcher. Since 2002, he has been an Associate Professor with the Department of Artificial Complex Systems Engineering, Hiroshima University, Hiroshima, Japan. His current interests are in the areas of signal processing in communication, wireless communications, adaptive signal processing, and multirate signal processing. Dr. Ohno is a member of IEICE. He has been serving as an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS since 2001.

Wu-Sheng Lu received the B.Sc. degree in Mathematics from Fudan University, Shanghai, China, in 1964, and the M.S. degree in Electrical Engineering and the Ph.D. degree in Control Science from the University of Minnesota, Minneapolis, USA, in 1983 and 1984, respectively. He was a post-doctoral fellow at the University of Victoria, Victoria, B.C., Canada, in 1985 and a visiting assistant professor with the University of Minnesota in 1986. Since 1987, he has been with the University of Victoria where he is a professor. His current teaching and research interests are in the general areas of digital signal processing and application of optimization methods. He is the co-author with A. Antoniou of Two-Dimensional Digital Filters (Marcel Dekker, 1992). Dr. Lu served as an associate editor of the Canadian Journal of Electrical and Computer Engineering in 1989, and editor of the same journal from 1990 to 1992. He served as an associate editor for the IEEE Transactions on Circuits and Systems, Part II, from 1993 to 1995 and for Part I of the same journal from 1999 to 2001. Presently he is serving as associate editor for the IEEE Transaction on Circuits and Systems, Part I, and the International Journal of Multidimensional Systems and Signal Processing.