MINIMIZATION OF L_2-SENSITIVITY FOR A CLASS OF 2-D STATE-SPACE DIGITAL FILTERS SUBJECT TO L_2-SCALING CONSTRAINTS

Takao Hinamoto, Ken-ichi Iwata
Graduate School of Engineering
Hiroshima University
Higashi-Hiroshima 739-8527, Japan
Email: hinamoto@hiroshima-u.ac.jp

Wu-Sheng Lu
Dept. of Elec. and Comp. Engineering
University of Victoria
Victoria, BC, Canada V8W 3P6
Email: wslu@ece.uvic.ca

***Abstract*— This paper is concerned with the minimization of an L_2-sensitivity measure subject to L_2-norm dynamic range scaling constraints for a class of two-dimensional (2-D) state-space digital filters. A novel iterative algorithm is developed to solve the constrained optimization problem directly. The proposed solution method is largely based on the use of a Lagrange function and some matrix theoretic techniques. A numerical example is presented to demonstrate the utility of the proposed technique.**

I. INTRODUCTION

When a transfer function with infinite accuracy coefficients is designed to meet filter specification requirements and realized by a state-space model, coefficients in the state-space model must be truncated or rounded to fit the finite word length (FWL) constraints for the purpose of implementing the filter in a finite binary representation. This coefficient quantization usually alters the characteristics of the filter and may turn a stable filter to an unstable one. This motivates the study of the coefficient sensitivity minimization problem. For 2-D state-space digital filters, the L_1/L_2-sensitivity minimization problem [1]-[5],[8] and L_2-sensitivity minimization problem [6]-[10] have been investigated. On the other hand, it is well known that the use of L_2-scaling constraints can be beneficial for suppressing overflow oscillation [11],[12]. However, not enough research has been reported on the minimization of L_2-sensitivity subject to the L_2-norm dynamic range scaling constraints [13].

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 [14]. The approach taken in the present paper differs significantly from the numerical optimization strategy employed in [13]. First, an expression for evaluating the L_2-sensitivity is introduced and an L_2-sensitivity minimization problem subject to the scaling constraints is formulated. Next, based on a Lagrange function and some matrix-theoretic techniques, a new iterative algorithm is proposed to solve the constrained optimization problem directly. The coordinate transformation matrix is then constructed to satisfy the L_2-scaling constraints. Finally, a numerical example is presented to illustrate the utility of the proposed algorithm.

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 $\boldsymbol{A}$ is denoted by $(\boldsymbol{A})_{ii}$.

II. L_2-SENSITIVITY ANANYSIS

Consider a local state-space model $(\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{b}, \boldsymbol{c}_1, \boldsymbol{c}_2, d)_n$ [14] for a class of 2-D state-space digital filters described by

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

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 2-D filter in (1) is assumed stable, locally controllable and locally observable. The transfer function of the 2-D filter in (1) is given by

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

It is noted that the local state-space model in (1) corresponds to the transposed structure of the Fornasini-Marchesini second model [15].

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 then defined as

$$\boldsymbol{S}_{\boldsymbol{X}} = \frac{\partial f}{\partial \boldsymbol{X}}, \qquad (\boldsymbol{S}_{\boldsymbol{X}})_{ij} = \frac{\partial f}{\partial x_{ij}} \tag{3}$$

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

With these notations, it can easily be shown that

$$\begin{aligned} \frac{\partial H(z_1,z_2)}{\partial \boldsymbol{A}_k} &= z_k^{-1}[\boldsymbol{F}(z_1,z_2)\boldsymbol{G}(z_1,z_2)]^T \\ \frac{\partial H(z_1,z_2)}{\partial \boldsymbol{b}} &= \boldsymbol{G}^T(z_1,z_2) \\ \frac{\partial H(z_1,z_2)}{\partial \boldsymbol{c}_k^T} &= \boldsymbol{F}(z_1,z_2), \qquad k = 1,2 \end{aligned} \tag{4}$$

where

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

The term d in (2) and its sensitivity are independent of the state-space 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 then defined as

$$||\boldsymbol{X}(z_1,z_2)||_2 = \left(\mathrm{tr}\left[\frac{1}{(2\pi j)^2}\oint_{\Gamma_1}\oint_{\Gamma_2}\boldsymbol{X}(z_1,z_2)\boldsymbol{X}^*(z_1,z_2)\frac{dz_1 dz_2}{z_1 z_2}\right]\right)^{\frac{1}{2}} \tag{5}$$

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

From (4) and *Definition 2*, the overall L_2-sensitivity measure for the 2-D filter in (1) is evaluated by

$$\begin{aligned} S &= \sum_{k=1}^{2}\left\|\frac{\partial H(z_1,z_2)}{\partial \boldsymbol{A}_k}\right\|_2^2 + \left\|\frac{\partial H(z_1,z_2)}{\partial \boldsymbol{b}}\right\|_2^2 \\ &\quad + \sum_{k=1}^{2}\left\|\frac{\partial H(z_1,z_2)}{\partial \boldsymbol{c}_k^T}\right\|_2^2 \\ &= \left\|[\boldsymbol{F}(z_1,z_2)\boldsymbol{G}(z_1,z_2)]^T\right\|_2^2 + \left\|\boldsymbol{G}^T(z_1,z_2)\right\|_2^2 \\ &\quad + \|\boldsymbol{F}(z_1,z_2)\|_2^2. \end{aligned} \tag{6}$$

It is easy to show that the L_2-sensitivity measure in (6) can be expressed as

$$S = 2\,\mathrm{tr}[\boldsymbol{M}(\boldsymbol{I}_n)] + \mathrm{tr}[\boldsymbol{W}_o] + 2\,\mathrm{tr}[\boldsymbol{K}_c] \tag{7}$$

where

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

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

$$\begin{aligned} \boldsymbol{K}_c &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\boldsymbol{f}(i,j)\,\boldsymbol{f}^T(i,j) \\ \boldsymbol{W}_o &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\boldsymbol{g}^T(i,j)\,\boldsymbol{g}(i,j) \\ \boldsymbol{M}(\boldsymbol{P}) &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\boldsymbol{H}^T(i,j)\boldsymbol{P}^{-1}\boldsymbol{H}(i,j) \end{aligned} \tag{8}$$

where

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

If a coordinate transformation defined by

$$\overline{\boldsymbol{x}}(i,j) = \boldsymbol{T}^{-1}\boldsymbol{x}(i,j) \tag{9}$$

is applied to the 2-D filter 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

$$\begin{aligned} &\overline{\boldsymbol{A}}_k = \boldsymbol{T}^{-1}\boldsymbol{A}_k\boldsymbol{T}, \quad \overline{\boldsymbol{b}} = \boldsymbol{T}^{-1}\boldsymbol{b}, \quad \overline{\boldsymbol{c}}_k = \boldsymbol{c}_k\boldsymbol{T}, \quad k = 1,2 \\ &\overline{\boldsymbol{K}}_c = \boldsymbol{T}^{-1}\boldsymbol{K}_c\boldsymbol{T}^{-T}, \qquad \overline{\boldsymbol{W}}_o = \boldsymbol{T}^T\boldsymbol{W}_o\boldsymbol{T} \end{aligned} \tag{10}$$

Noting that (9) transforms $\boldsymbol{M}(\boldsymbol{I}_n)$ to $\boldsymbol{T}^T\boldsymbol{M}(\boldsymbol{P})\boldsymbol{T}$, it is possible to change the L_2-sensitivity measure in (7) to

$$S(\boldsymbol{P}) = 2\,\mathrm{tr}[\boldsymbol{M}(\boldsymbol{P})\boldsymbol{P}] + \mathrm{tr}[\boldsymbol{W}_o\boldsymbol{P}] + 2\,\mathrm{tr}[\boldsymbol{K}_c\boldsymbol{P}^{-1}] \tag{11}$$

where $\boldsymbol{P} = \boldsymbol{T}\boldsymbol{T}^T$.

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

$$(\overline{\boldsymbol{K}}_c)_{ii} = (\boldsymbol{T}^{-1}\boldsymbol{K}_c\boldsymbol{T}^{-T})_{ii} = 1 \tag{12}$$

is required for $i = 1, 2, \cdots, n$. As a result, the problem considered here is as follows: *For 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 (11) subject to the constraints in (12).*

III. L_2-SENSITIVITY MINIMIZATION

The problem of minimizing $S(\boldsymbol{P})$ in (11) subject to the constraints in (12) is a constrained nonlinear optimization problem where the variable matrix is $\boldsymbol{P}$. If we sum the n constraints in (12) up, then we have

$$\mathrm{tr}[\boldsymbol{T}^{-1}\boldsymbol{K}_c\boldsymbol{T}^{-T}] = \mathrm{tr}[\boldsymbol{K}_c\boldsymbol{P}^{-1}] = n. \tag{13}$$

Consequently, the above problem can be *relaxed* into the following problem:

$$\begin{aligned} &\text{minimize } S(\boldsymbol{P}) \text{ in (11)} \\ &\text{subject to } \mathrm{tr}[\boldsymbol{K}_c\boldsymbol{P}^{-1}] = n. \end{aligned} \tag{14}$$

Although clearly a solution of problem (14) is not necessarily a solution of the problem of minimizing (11) subject to the constraints in (12), it is important to stress that the ultimate solution we seek for is not matrix $\boldsymbol{P}$ but a nonsingular matrix $\boldsymbol{T}$ that is related to the solution of the problem of minimizing (11) subject to the constraints in (12) as $\boldsymbol{P} = \boldsymbol{T}\boldsymbol{T}^T$. If matrix $\boldsymbol{P}$ is a solution of problem (14) and $\boldsymbol{P}^{1/2}$ denotes a matrix square root of $\boldsymbol{P}$, i.e., $\boldsymbol{P} = \boldsymbol{P}^{1/2}\boldsymbol{P}^{1/2}$, then it is easy to see that any matrix $\boldsymbol{T}$ of the form $\boldsymbol{T} = \boldsymbol{P}^{1/2}\boldsymbol{U}$ where $\boldsymbol{U}$ is an arbitrary orthogonal matrix still holds the relation $\boldsymbol{P} = \boldsymbol{T}\boldsymbol{T}^T$.

As will be shown shortly, under the constraint in (13) there exists an orthogonal matrix $\boldsymbol{U}$ such that matrix $\boldsymbol{T} = \boldsymbol{P}^{1/2}\boldsymbol{U}$ satisfies the constraints in (12), where $\boldsymbol{P}^{1/2}$ is a square root of the solution matrix $\boldsymbol{P}$ for problem (14).

It is for these reasons we now address problem (14) as the first step of our solution strategy. To solve (14), we define the Lagrange function of the problem as

$$J(\boldsymbol{P},\lambda) = 2\,\mathrm{tr}[\boldsymbol{M}(\boldsymbol{P})\boldsymbol{P}] + \mathrm{tr}[\boldsymbol{W}_o\boldsymbol{P}] + 2\,\mathrm{tr}[\boldsymbol{K}_c\boldsymbol{P}^{-1}] + \lambda(\mathrm{tr}[\boldsymbol{K}_c\boldsymbol{P}^{-1}] - n) \tag{15}$$

where λ is a Lagrange multiplier. It is well known that the solution of problem (15) must satisfy the Karush-Kuhn-Tucker (KKT) conditions $\partial J(\boldsymbol{P},\lambda)/\partial\boldsymbol{P} = \boldsymbol{0}$ and $\partial J(\boldsymbol{P},\lambda)/\partial\lambda = 0$ where the gradients are found to be

$$\begin{aligned}\frac{\partial J(\boldsymbol{P},\lambda)}{\partial\boldsymbol{P}} &= 2\,\boldsymbol{M}(\boldsymbol{P}) - 2\,\boldsymbol{P}^{-1}\boldsymbol{N}(\boldsymbol{P})\boldsymbol{P}^{-1} + \boldsymbol{W}_o \\ &\quad -(\lambda+2)\boldsymbol{P}^{-1}\boldsymbol{K}_c\boldsymbol{P}^{-1} \qquad (16)\\ \frac{\partial J(\boldsymbol{P},\lambda)}{\partial\lambda} &= \mathrm{tr}[\boldsymbol{K}_c\boldsymbol{P}^{-1}] - n\end{aligned}$$

where $\boldsymbol{N}(\boldsymbol{P})$ is derived from

$$\boldsymbol{N}(\boldsymbol{P}) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\boldsymbol{H}(i,j)\boldsymbol{P}\boldsymbol{H}^T(i,j).$$

Hence the KKT conditions become

$$\boldsymbol{P}\,\boldsymbol{F}(\boldsymbol{P})\boldsymbol{P} = \boldsymbol{G}(\boldsymbol{P},\lambda), \qquad \mathrm{tr}[\boldsymbol{K}_c\boldsymbol{P}^{-1}] = n \tag{17}$$

where

$$\begin{aligned}\boldsymbol{F}(\boldsymbol{P}) &= 2\,\boldsymbol{M}(\boldsymbol{P}) + \boldsymbol{W}_o\\ \boldsymbol{G}(\boldsymbol{P},\lambda) &= 2\,\boldsymbol{N}(\boldsymbol{P}) + (\lambda+2)\boldsymbol{K}_c.\end{aligned}$$

The first equation in (17) is highly nonlinear with respect to $\boldsymbol{P}$. An effective approach to solving the first equation in (17) is to *relax* it into the following recursive second-order matrix equation:

$$\boldsymbol{P}_{i+1}\boldsymbol{F}(\boldsymbol{P}_i)\boldsymbol{P}_{i+1} = \boldsymbol{G}(\boldsymbol{P}_i,\lambda_i) \tag{18}$$

where $\boldsymbol{P}_i$ is assumed to be known from the previous recursion and the solution $\boldsymbol{P}_{i+1}$ is given by [8]

$$\boldsymbol{P}_{i+1} = \boldsymbol{F}(\boldsymbol{P}_i)^{-\frac{1}{2}}[\boldsymbol{F}(\boldsymbol{P}_i)^{\frac{1}{2}}\boldsymbol{G}(\boldsymbol{P}_i,\lambda_i)\boldsymbol{F}(\boldsymbol{P}_i)^{\frac{1}{2}}]^{\frac{1}{2}}\boldsymbol{F}(\boldsymbol{P}_i)^{-\frac{1}{2}}. \tag{19}$$

To derive a recursive formula for the Lagrange multiplier λ, we use (17) to write

$$\mathrm{tr}[\boldsymbol{P}\boldsymbol{F}(\boldsymbol{P})] = 2\,\mathrm{tr}[\boldsymbol{N}(\boldsymbol{P})\boldsymbol{P}^{-1}] + n(\lambda+2) \tag{20}$$

which naturally suggests the following recursion for λ:

$$\lambda_{i+1} = \frac{\mathrm{tr}[\boldsymbol{P}_i\boldsymbol{F}(\boldsymbol{P}_i)] - 2\,\mathrm{tr}[\boldsymbol{N}(\boldsymbol{P}_i)\boldsymbol{P}_i^{-1}]}{n} - 2. \tag{21}$$

λ_i is the solution of the previous iteration. The initial estimates are given by $\boldsymbol{P}_0 = \boldsymbol{I}_n$ and any value of $\lambda_0 > 0$. This iteration process continues until (17) is satisfied within a prescribed numerical tolerance. Although a mathematical proof of the algorithm's convergence is not yet available, in our computer simulations the proposed algorithm does converge for various types of this class of 2-D state-space digital filters.

As the second step of the solution strategy, we now turn our attention to the construction of the optimal coordinate transformation matrix $\boldsymbol{T}$ that solves the problem of minimizing (11) subject to the constraints in (12). As analyzed earlier, the optimal $\boldsymbol{T}$ assumes the form

$$\boldsymbol{T} = \boldsymbol{P}^{\frac{1}{2}}\boldsymbol{U} \tag{22}$$

where $\boldsymbol{P}^{1/2}$ is the square roots of the matrix $\boldsymbol{P}$ obtained above, and $\boldsymbol{U}$ is an $n\times n$ orthogonal matrix to be determined as follows. From (10) and (22) it follows that

$$\begin{aligned}\overline{\boldsymbol{K}}_c &= \boldsymbol{T}^{-1}\boldsymbol{K}_c\boldsymbol{T}^{-T}\\ &= \boldsymbol{U}^T\boldsymbol{P}^{-\frac{1}{2}}\boldsymbol{K}_c\boldsymbol{P}^{-\frac{1}{2}}\boldsymbol{U}.\end{aligned} \tag{23}$$

In order to find an $n\times n$ orthogonal matrix $\boldsymbol{U}$ such that the matrix $\overline{\boldsymbol{K}}_c$ satisfies the scaling constraints in (12), we perform the eigenvalue-eigenvector decomposition for the positive definite matrix $\boldsymbol{P}^{-1/2}\boldsymbol{K}_c\boldsymbol{P}^{-1/2}$ as

$$\boldsymbol{P}^{-\frac{1}{2}}\boldsymbol{K}_c\boldsymbol{P}^{-\frac{1}{2}} = \boldsymbol{R}\boldsymbol{\Theta}\boldsymbol{R}^T \tag{24}$$

where $\boldsymbol{\Theta} = \mathrm{diag}\{\theta_1,\theta_2,\cdots,\theta_n\}$ with $\theta_i > 0$ and $\boldsymbol{R}$ is an orthogonal matrix. Next, an orthogonal matrix $\boldsymbol{S}$ such that

$$\boldsymbol{S}\boldsymbol{\Theta}\boldsymbol{S}^T = \begin{bmatrix} 1 & * & \cdots & * \\ * & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ * & \cdots & * & 1 \end{bmatrix} \tag{25}$$

can be obtained by numerical manipulations [12, p.278]. Using (23), (24) and (25), it can be readily verified that the orthogonal matrix $\boldsymbol{U} = \boldsymbol{R}\boldsymbol{S}^T$ leads to a $\overline{\boldsymbol{K}}_c$ in (23) whose diagonal elements are equal to unity, hence the constraints in (12) are now satisfied. This matrix $\boldsymbol{T}$ together with (22) gives the solution of the problem of minimizing (11) subject to the constraints in (12) as

$$\boldsymbol{T} = \boldsymbol{P}^{\frac{1}{2}}\boldsymbol{R}\boldsymbol{S}^T. \tag{26}$$

IV. NUMERICAL EXAMPLE

Consider a 2-D digital filter, (1), specified by

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

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

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

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

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

$$d = 0.00943.$$

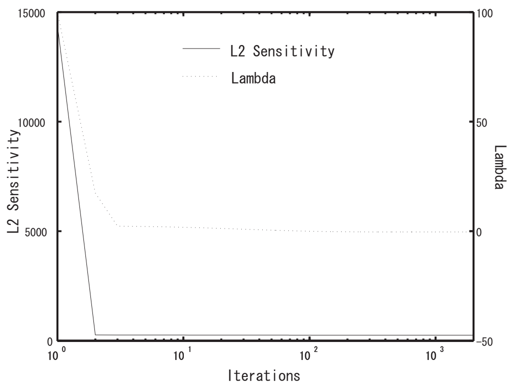

Fig. 1. L_2-Sensitivity and λ Performances

Using (8) over $(0,0) \leq (i,j) \leq (100,100)$, the L_2-sensitivity measure in (7) is computed as $S = 14424.346809$.

Choosing $\boldsymbol{P}_0 = \boldsymbol{I}_4$ and $\lambda_0 = 100$ in (19) and (21) as the initial estimates, it took the proposed iterative algorithm 2000 iterations to converge to

$$\boldsymbol{P}^{opt} = \begin{bmatrix} 2.404416 & 2.137212 & 1.871706 & 1.648549 \\ 2.137212 & 1.953194 & 1.743170 & 1.539139 \\ 1.871706 & 1.743170 & 1.617990 & 1.457391 \\ 1.648549 & 1.539139 & 1.457391 & 1.355055 \end{bmatrix}$$

which yields

$$\boldsymbol{T}^{opt} = \begin{bmatrix} 0.787107 & -0.915199 & 0.314478 & 0.921082 \\ 0.608965 & -0.944631 & 0.414720 & 0.719748 \\ 0.330685 & -0.970545 & 0.354388 & 0.664145 \\ 0.146764 & -0.824812 & 0.370182 & 0.718447 \end{bmatrix}.$$

In this case, (11) is minimized subject to the scaling constraints in (12) to $S(\boldsymbol{P}^{opt}) = 255.387433$.

The L_2-sensitivity and λ performances of 2000 iterations are shown in Fig.1, from which it is seen that the iterative algorithm converges with 2000 iterations.

For comparison, only the iterative algorithm in (19) is applied by letting $\lambda_i = 0$ for any i and $\boldsymbol{P}_0 = \boldsymbol{I}_4$ to minimize the L_2-sensivivity measure in (11) (without considering the scaling constraints in (12)) and after 2000 iterations it converges to $\boldsymbol{P} = \boldsymbol{T}\boldsymbol{T}^T$ and $S_2(\boldsymbol{P}) = 255.313680$ where

$$\boldsymbol{T} = \begin{bmatrix} 1.625231 & 0.0 & 0.0 & 0.0 \\ 1.444676 & 0.242716 & 0.0 & 0.0 \\ 1.265260 & 0.359826 & 0.219322 & 0.0 \\ 1.114386 & 0.334013 & 0.326988 & 0.170877 \end{bmatrix}.$$

The above coordinate transformation matrix $\boldsymbol{T}$ is then scaled by an appropriate nonsingular diagonal matrix, so that the scaling constraints in (12) are satisfied. Then the result is

$$S_2(\boldsymbol{P}) = 358.024764$$

where $\boldsymbol{P} = \boldsymbol{T}\boldsymbol{T}^T$ and

$$\boldsymbol{T} = \begin{bmatrix} 1.0 & 0.0 & 0.0 & 0.0 \\ 0.888905 & 0.186972 & 0.0 & 0.0 \\ 0.778511 & 0.277186 & 0.169442 & 0.0 \\ 0.685679 & 0.257302 & 0.252621 & 0.244088 \end{bmatrix}.$$

This shows that the constrained optimization technique offers smaller L_2-sensitivity subject to the scaling constraints relative to the unconstrained counterpart that needs the scaling later.

V. CONCLUSION

This paper has investigated the problem of minimizing an L_2-sensitivity measure subject to the L_2-scaling constraints for a class of 2-D state-space digital filters. An efficient iterative algorithm has been developed by using a Lagrange function and some matrix-theoretic techniques to solve the constrained optimization problem directly. The coordinate transformation matrix has then been constructed to satisfy the L_2-scaling constraints. Computer simulation results have demonstrated the effectiveness of the proposed iterative technique. The extension of the proposed technique to the Roesser model will appear elsewhere.

REFERENCES

[1] 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.

[2] 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.

[3] 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.

[4] 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.

[5] 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.

[6] 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.

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

[8] 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.

[9] 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 Syst. I*, vol.49, pp.1279-1289, Sept. 2002.

[10] 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.

[11] 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.

[12] 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.

[13] T. Hinamoto, H. Ohnishi and W.-S. Lu, "Minimization of L_2-sensitivity for 2-D state-space digital filters subject to L_2-scaling constraints", in *Proc. 2004 IEEE Int. Symp. Circuits Syst.*, vol.3, pp.249-252.

[14] 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.

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