Realization of 3-D Separable-Denominator Digital Filters with Very Low l_2-Sensitivity

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

Osamu Tanaka
Graduate School of Engineering
Hiroshima University
Higashi-Hiroshima 739-8527, Japan
Email: o–tanaka@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 investigates the problem of reducing the deviation from a desired transfer function caused by the coefficient quantization errors of a three-dimensional (3-D) separable-denominator digital filter. First, a 3-D transfer function with separable denominator is represented with the cascade connection of three one-dimensional (1-D) transfer functions by applying a minimal decomposition technique. Next, the multi-input multi-output (MIMO) 1-D transfer function located in the middle of the cascade connection is realized by a minimal state-space model and then the l_2/l_2-sensitivity of the model is analyzed. Third, a technique for the optimal synthesis of the minimal state-space model is developed so as to minimize the l_2-sensitivity subject to l_2-scaling constraints. Finally, a numerical example is given to demonstrate the validity and effectiveness of the proposed technique.**

I. INTRODUCTION

The problem of minimizing the coefficient sensitivity for two-dimensional (2-D) state-space digital filters has been explored extensively [1]-[7]. Several techniques have been proposed for synthesizing 2-D state-space filter structures with minimum coefficient sensitivity [2]-[7]. Some of them evaluate the sensitivity by using a mixture of l_1/l_2 norms [1]-[4]. The others rely on the use of a pure l_2 norm [5]-[7]. In [6], the weighted-sensitivity minimization of 2-D state-space digital filters has been considered in both cases of a mixture of L_1/L_2 norms and a pure L_2 norm. It should be noted that the l_2-sensitivity minimization is more natural and reasonable than the conventional l_1/l_2 mixed sensitivity minimization, but it is more challenging [7]. A technique has also been presented for synthesizing three-dimensional (3-D) separable-denominator (SD) state-space digital filters with minimum l_2-sensitivity [8]. More recently, the minimization problem of l_2-sensitivity subject to l_2-scaling constraints has been treated for 2-D state-space digital filters [9],[10]. It is known that the use of scaling constraints can be beneficial for suppressing overflow [11]. Alternatively, 3-D digital filters find applications in various image and video signal processing problems, and the coefficient sensitivity of these filters is directly related to their performance on finite word-length devices.

This paper treats the realization of 3-D SD digital filters which reduce l_2-sensitivity subject to l_2-scaling constraints. First, a 3-D transfer function with separable denominator is decomposed into three one-dimensional (1-D) transfer functions with a cascade connection, and the MIMO 1-D transfer function in the middle of the cascade connection is described by minimal state-space realization. Next, an iterative procedure is developed for constructing the optimal state-space model for the MIMO 1-D system so as to minimize the l_2-sensitivity subject to l_2-scaling constraints. Finally, a numerical example is given to demonstrate the validity and effectiveness of the proposed technique.

II. PROBLEM STATEMENT

Consider a stable 3-D SD digital filter described by

$$H(z_1, z_2, z_3) = \frac{N(z_1, z_2, z_3)}{D_1(z_1)D_2(z_2)D_3(z_3)} \tag{1}$$

where

$$\begin{aligned} N(z_1, z_2, z_3) &= \sum_{i=0}^{N_1}\sum_{j=0}^{N_2}\sum_{k=0}^{N_3} a_{ijk} z_1^{-i} z_2^{-j} z_3^{-k} \\ D_1(z_1) &= 1 + b_{11}z_1^{-1} + \cdots + b_{1N_1}z_1^{-N_1} \\ D_2(z_2) &= 1 + b_{21}z_2^{-1} + \cdots + b_{2N_2}z_2^{-N_2} \\ D_3(z_3) &= 1 + b_{31}z_3^{-1} + \cdots + b_{3N_3}z_3^{-N_3}. \end{aligned}$$

The 3-D transfer function in (1) can be decomposed into three 1-D systems as

$$H(z_1, z_2, z_3) = \frac{\boldsymbol{Z}_1^T}{D_1(z_1)}\boldsymbol{H}_2(z_2)\frac{\boldsymbol{Z}_3}{D_3(z_3)} \tag{2}$$

where

$$\begin{aligned} \boldsymbol{Z}_1 &= (1, z_1^{-1}, \cdots, z_1^{-N_1})^T \\ \boldsymbol{Z}_3 &= (1, z_3^{-1}, \cdots, z_3^{-N_3})^T \\ \boldsymbol{H}_2(z_2) &= \frac{\boldsymbol{\Delta}_0 + \boldsymbol{\Delta}_1 z_2^{-1} + \cdots + \boldsymbol{\Delta}_{N_2} z_2^{-N_2}}{D_2(z_2)} \\ \boldsymbol{\Delta}_m &= \begin{bmatrix} a_{0m0} & a_{0m1} & \cdots & a_{0mN_3} \\ a_{1m0} & a_{1m1} & \cdots & a_{1mN_3} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N_1m0} & a_{N_1m1} & \cdots & a_{N_1mN_3} \end{bmatrix} \\ & m = 0, 1, \cdots, N_2. \end{aligned}$$

The above 1-D transfer function $\boldsymbol{H}_2(z_2)$ with (N_3+1) inputs and (N_1+1) outputs can be realized by the following minimal

state-space model $(\boldsymbol{A}_2, \boldsymbol{B}_2, \boldsymbol{C}_2, \boldsymbol{\Delta}_0)_p$:

$$\begin{aligned} \boldsymbol{x}(k+1) &= \boldsymbol{A}_2\boldsymbol{x}(k) + \boldsymbol{B}_2\boldsymbol{u}(k) \\ \boldsymbol{y}(k) &= \boldsymbol{C}_2\boldsymbol{x}(k) + \boldsymbol{\Delta}_0\boldsymbol{u}(k) \end{aligned} \tag{3}$$

where $\boldsymbol{x}(k)$ is a $p \times 1$ state-variable vector, $\boldsymbol{u}(k)$ is an $(N_3+1)\times 1$ input vector, $\boldsymbol{y}(k)$ is an $(N_1+1)\times 1$ output vector, and $\boldsymbol{A}_2, \boldsymbol{B}_2, \boldsymbol{C}_2$ and $\boldsymbol{\Delta}_0$ are real constant matrices of appropriate dimensions. The transfer function of the linear system in (3) is given by

$$\boldsymbol{H}_2(z_2) = \boldsymbol{C}_2(z_2\boldsymbol{I}_p - \boldsymbol{A}_2)^{-1}\boldsymbol{B}_2 + \boldsymbol{\Delta}_0. \tag{4}$$

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 $f(\boldsymbol{X})$ with respect to $\boldsymbol{X}$ is then defined as

$$\boldsymbol{S}_{\boldsymbol{X}} = \frac{\partial f}{\partial \boldsymbol{X}}, \qquad (\boldsymbol{S}_{\boldsymbol{X}})_{ij} = \frac{\partial f}{\partial x_{ij}} \tag{5}$$

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

Definition 2 : Let $\boldsymbol{X}(z_1, z_2, z_3)$ be an $m\times n$ complex matrix-valued function of complex variables z_1, z_2, and z_3. Let $x_{pq}(z_1, z_2, z_3)$ be the (p,q)th entry of $\boldsymbol{X}(z_1, z_2, z_3)$. Then the l_2-norm of $\boldsymbol{X}(z_1, z_2, z_3)$ is defined as

$$\begin{aligned} &\|\boldsymbol{X}(z_1, z_2, z_3)\|_2 \\ &= \left(\mathrm{tr}\left[\frac{1}{(2\pi j)^3} \oint_{|z_1|=1} \oint_{|z_2|=1} \oint_{|z_3|=1} \boldsymbol{X}(z_1, z_2, z_3) \right.\right. \\ &\qquad \left.\left. \cdot \boldsymbol{X}^*(z_1, z_2, z_3) \frac{dz_1}{z_1}\frac{dz_2}{z_2}\frac{dz_3}{z_3} \right] \right)^{\frac{1}{2}}. \end{aligned} \tag{6}$$

From (2), (4) and Definitions 1 and 2, the l_2-sensitivity of the transfer function $H(z_1, z_2, z_3)$ with respect to the coefficient matrices $\boldsymbol{A}_2$, $\boldsymbol{B}_2$, and $\boldsymbol{C}_2$ is evaluated by

$$\begin{aligned} S &= \left\| \frac{\partial H(z_1, z_2, z_3)}{\partial \boldsymbol{A}_2} \right\|_2^2 + \left\| \frac{\partial H(z_1, z_2, z_3)}{\partial \boldsymbol{B}_2} \right\|_2^2 \\ &\quad + \left\| \frac{\partial H(z_1, z_2, z_3)}{\partial \boldsymbol{C}_2} \right\|_2^2 \\ &= \left\| [\boldsymbol{f}(z_2, z_3)\boldsymbol{g}(z_1, z_2)]^T \right\|_2^2 + \left\| \left[\frac{\boldsymbol{Z}_3}{D_3(z_3)} \boldsymbol{g}(z_1, z_2) \right]^T \right\|_2^2 \\ &\quad + \left\| \left[\boldsymbol{f}(z_2, z_3) \frac{\boldsymbol{Z}_1^T}{D_1(z_1)} \right]^T \right\|_2^2 \end{aligned} \tag{7}$$

where

$$\begin{aligned} \boldsymbol{f}(z_2, z_3) &= (z_2\boldsymbol{I}_p - \boldsymbol{A}_2)^{-1}\boldsymbol{B}_2 \frac{\boldsymbol{Z}_3}{D_3(z_3)} \\ \boldsymbol{g}(z_1, z_2) &= \frac{\boldsymbol{Z}_1^T}{D_1(z_1)} \boldsymbol{C}_2 (z_2\boldsymbol{I}_p - \mathbf{A}_2)^{-1} \\ \frac{\boldsymbol{Z}_1^T}{D_1(z_1)} &= \boldsymbol{c}_1(z_1\boldsymbol{I}_{N_1} - \boldsymbol{A}_1)^{-1}\boldsymbol{B}_1 + \boldsymbol{d}_1 \\ \frac{\boldsymbol{Z}_3}{D_3(z_3)} &= \boldsymbol{C}_3(z_3\boldsymbol{I}_{N_3} - \boldsymbol{A}_3)^{-1}\boldsymbol{b}_3 + \boldsymbol{d}_3. \end{aligned}$$

The l_2-sensitivity measure in (7) can be written as

$$S = \mathrm{tr}[\boldsymbol{M}_A(\boldsymbol{I}_p)] + \mathrm{tr}[\boldsymbol{W}_B] + \mathrm{tr}[\boldsymbol{K}_C] \tag{8}$$

where $\boldsymbol{M}_A(\boldsymbol{I}_p)$, $\boldsymbol{W}_B$, and $\boldsymbol{K}_C$ are obtained by the following general expression:

$$\begin{aligned} \boldsymbol{X} = \frac{1}{(2\pi j)^3} \oint_{|z_1|=1} \oint_{|z_2|=1} \oint_{|z_3|=1} &\boldsymbol{Y}(z_1, z_2, z_3) \\ &\cdot \boldsymbol{Y}^*(z_1, z_2, z_3) \frac{dz_1}{z_1}\frac{dz_2}{z_2}\frac{dz_3}{z_3} \end{aligned}$$

with

$$\begin{aligned} \boldsymbol{Y}(z_1, z_2, z_3) &= [\boldsymbol{f}(z_2, z_3)\boldsymbol{g}(z_1, z_2)]^T \quad \text{for} \quad \boldsymbol{X} = \boldsymbol{M}_A(\boldsymbol{I}_p) \\ \boldsymbol{Y}(z_1, z_2, z_3) &= \left[\frac{\boldsymbol{Z}_3}{D_3(z_3)} \boldsymbol{g}(z_1, z_2) \right]^* \quad \text{for} \quad \boldsymbol{X} = \boldsymbol{W}_B \\ \boldsymbol{Y}(z_1, z_2, z_3) &= \boldsymbol{f}(z_2, z_3) \frac{\boldsymbol{Z}_1^T}{D_1(z_1)} \quad \text{for} \quad \boldsymbol{X} = \boldsymbol{K}_C. \end{aligned}$$

The Gramians $\boldsymbol{M}_A(\boldsymbol{P})$, $\boldsymbol{W}_B$, and $\boldsymbol{K}_C$ can be computed using

$$\begin{aligned} \boldsymbol{M}_A(\boldsymbol{P}) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} &\begin{bmatrix} \mathbf{0} & \boldsymbol{I}_p \end{bmatrix} \begin{bmatrix} \boldsymbol{A}_2^T & \mathbf{0} \\ \boldsymbol{C}_2^T\boldsymbol{R}_{ij}^T\boldsymbol{B}_2^T & \boldsymbol{A}_2^T \end{bmatrix}^k \\ &\cdot \begin{bmatrix} \boldsymbol{P}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{A}_2 & \boldsymbol{B}_2\boldsymbol{R}_{ij}\boldsymbol{C}_2 \\ \mathbf{0} & \boldsymbol{A}_2 \end{bmatrix}^k \begin{bmatrix} \mathbf{0} \\ \boldsymbol{I}_p \end{bmatrix} \end{aligned}$$

$$\begin{aligned} \boldsymbol{W}_B &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} (\boldsymbol{A}_2^T)^k \boldsymbol{C}_2^T\boldsymbol{R}_{ij}^T\boldsymbol{R}_{ij}\boldsymbol{C}_2\boldsymbol{A}_2^k \\ \boldsymbol{K}_C &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \boldsymbol{A}_2^k\boldsymbol{B}_2\boldsymbol{R}_{ij}\boldsymbol{R}_{ij}^T\boldsymbol{B}_2^T(\boldsymbol{A}_2^T)^k \end{aligned} \tag{9}$$

where

$$\begin{aligned} \frac{\boldsymbol{Z}_3\boldsymbol{Z}_1^T}{D_3(z_3)D_1(z_1)} &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \boldsymbol{R}_{ij} z_1^{-i} z_3^{-j} \\ \frac{\boldsymbol{Z}_1^T}{D_1(z_1)} &= \sum_{i=1}^{\infty} \boldsymbol{c}_1\boldsymbol{A}_1^{i-1}\boldsymbol{B}_1 z_1^{-i} + \boldsymbol{d}_1 \\ \frac{\boldsymbol{Z}_3}{D_3(z_3)} &= \sum_{j=1}^{\infty} \boldsymbol{C}_3\boldsymbol{A}_3^{i-1}\boldsymbol{b}_3 z_3^{-j} + \boldsymbol{d}_3. \end{aligned}$$

Applying a coordinate transformation defined by $\overline{\boldsymbol{x}}(k) = \boldsymbol{T}^{-1}\boldsymbol{x}(k)$ to the 1-D system $(\boldsymbol{A}_2, \boldsymbol{B}_2, \boldsymbol{C}_2, \boldsymbol{\Delta}_0)_p$ in (3), we obtain a new realization $(\overline{\boldsymbol{A}}_2, \overline{\boldsymbol{B}}_2, \overline{\boldsymbol{C}}_2, \boldsymbol{\Delta}_0)_p$ characterized by

$$\overline{\boldsymbol{A}}_2 = \boldsymbol{T}^{-1}\boldsymbol{A}_2\boldsymbol{T}, \quad \overline{\boldsymbol{B}}_2 = \boldsymbol{T}^{-1}\boldsymbol{B}_2, \quad \overline{\boldsymbol{C}}_2 = \boldsymbol{C}_2\boldsymbol{T}. \tag{10}$$

For the new realization, the l_2-sensitivity measure in (8) is changed to

$$S(\boldsymbol{P}) = \mathrm{tr}[\boldsymbol{M}_A(\boldsymbol{P})\boldsymbol{P}] + \mathrm{tr}[\boldsymbol{W}_B\boldsymbol{P}] + \mathrm{tr}[\boldsymbol{K}_C\boldsymbol{P}^{-1}] \tag{11}$$

where $\boldsymbol{P} = \boldsymbol{T}\boldsymbol{T}^T$. Noting that $\boldsymbol{f}(z_2, z_3)$ is the transfer function from the filter input to the state-variable vector $\boldsymbol{x}(k)$,

a controllability Gramian $\boldsymbol{K}$ can be derived from

$$\begin{aligned}\boldsymbol{K} &= \frac{1}{(2\pi j)^2}\oint_{|z_2|=1}\oint_{|z_3|=1}\boldsymbol{f}(z_2,z_3)\boldsymbol{f}^*(z_2,z_3)\frac{dz_2}{z_2}\frac{dz_3}{z_3}\\ &= \sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\boldsymbol{A}_2^k\boldsymbol{B}_2\boldsymbol{r}_j\boldsymbol{r}_j^T\boldsymbol{B}_2^T(\boldsymbol{A}_2^T)^k\end{aligned} \tag{12}$$

where

$$\frac{\boldsymbol{Z}_3}{D_3(z_3)} = \sum_{j=0}^{\infty}\boldsymbol{r}_j z_3^{-j}.$$

In this case, l_2-scaling constraints are given by

$$(\overline{\boldsymbol{K}})_{ii} = (\boldsymbol{T}^{-1}\boldsymbol{K}\boldsymbol{T}^{-T})_{ii} = 1 \quad \text{for} \quad i = 1, 2, \cdots, p. \tag{13}$$

The problem at hand can now be formulated as to obtain the coordinate transformation matrix $\boldsymbol{T}$ that minimizes $S(\boldsymbol{P})$ in (11) subject to the l_2-scaling constraints in (13).

III. L_2-SENSITIVITY MINIMIZATION

Let the l_2-scaling constraints in (13) be relaxed as

$$\text{tr}[\boldsymbol{T}^{-1}\boldsymbol{K}\boldsymbol{T}^{-T}] = \text{tr}[\boldsymbol{K}\boldsymbol{P}^{-1}] = p. \tag{14}$$

If $\text{tr}[\boldsymbol{K}\boldsymbol{P}^{-1}] = p$ is satisfied, then a $p \times p$ orthogonal matrix $\boldsymbol{U}$ can always be constructed so that $\boldsymbol{T} = \boldsymbol{P}^{1/2}\boldsymbol{U}$ satisfies l_2-scaling constraints in (13) [9]. This justifies the relaxation made in (14) and in this way, we now focus on the problem

$$\begin{aligned}&\text{minimize } S(\boldsymbol{P}) \text{ in (11)}\\ &\text{subject to } \text{tr}[\boldsymbol{K}\boldsymbol{P}^{-1}] = p.\end{aligned} \tag{15}$$

To solve problem (15), we define the following Lagrange function of the problem:

$$\begin{aligned}J(\boldsymbol{P},\lambda) = &\ \text{tr}[\boldsymbol{M}_A(\boldsymbol{P})\boldsymbol{P}] + \text{tr}[\boldsymbol{W}_B\boldsymbol{P}]\\ &+\text{tr}[\boldsymbol{K}_C\boldsymbol{P}^{-1}] + \lambda(\text{tr}[\boldsymbol{K}\boldsymbol{P}^{-1}] - p)\end{aligned} \tag{16}$$

where λ is the Lagrange multiplier. Setting $\partial J(\boldsymbol{P},\lambda)/\partial\boldsymbol{P} = \boldsymbol{0}$, it follows that

$$\boldsymbol{P}\boldsymbol{F}(\boldsymbol{P})\boldsymbol{P} = \boldsymbol{G}(\boldsymbol{P},\lambda) \tag{17}$$

where

$$\begin{aligned}\boldsymbol{F}(\boldsymbol{P}) &= \boldsymbol{M}_A(\boldsymbol{P}) + \boldsymbol{W}_B\\ \boldsymbol{G}(\boldsymbol{P},\lambda) &= \boldsymbol{N}_A(\boldsymbol{P}) + \boldsymbol{K}_C + \lambda\boldsymbol{K}\end{aligned}$$

with

$$\begin{aligned}\boldsymbol{N}_A(\boldsymbol{P}) = &\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\begin{bmatrix}\boldsymbol{I}_p & \boldsymbol{0}\end{bmatrix}\begin{bmatrix}\boldsymbol{A}_2 & \boldsymbol{B}_2\boldsymbol{R}_{ij}\boldsymbol{C}_2\\ \boldsymbol{0} & \boldsymbol{A}_2\end{bmatrix}^k\\ &\cdot\begin{bmatrix}\boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{P}\end{bmatrix}\begin{bmatrix}\boldsymbol{A}_2^T & \boldsymbol{0}\\ \boldsymbol{C}_2^T\boldsymbol{R}_{ij}^T\boldsymbol{B}_2^T & \boldsymbol{A}_2^T\end{bmatrix}^k\begin{bmatrix}\boldsymbol{I}_p\\ \boldsymbol{0}\end{bmatrix}.\end{aligned}$$

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

$$\boldsymbol{P}^{(k+1)}\boldsymbol{F}(\boldsymbol{P}^{(k)})\boldsymbol{P}^{(k+1)} = \boldsymbol{G}(\boldsymbol{P}^{(k)},\lambda^{(k+1)}) \tag{18}$$

where $\boldsymbol{P}^{(k)}$ is assumed to be known from the previous recursion and the solution $\boldsymbol{P}^{(k+1)}$ is given by [9]

$$\begin{aligned}\boldsymbol{P}^{(k+1)} = &\ \boldsymbol{F}(\boldsymbol{P}^{(k)})^{-\frac{1}{2}}[\boldsymbol{F}(\boldsymbol{P}^{(k)})^{\frac{1}{2}}\\ &\cdot\boldsymbol{G}(\boldsymbol{P}^{(k)},\lambda^{(k+1)})\boldsymbol{F}(\boldsymbol{P}^{(k)})^{\frac{1}{2}}]^{\frac{1}{2}}\boldsymbol{F}(\boldsymbol{P}^{(k)})^{-\frac{1}{2}}.\end{aligned} \tag{19}$$

Here, the Lagrange multiplier $\lambda^{(k+1)}$ can be efficiently obtained using a bisection method so that

$$f(\lambda^{(k+1)}) = p - \text{tr}[\tilde{\boldsymbol{K}}^{(k)}\tilde{\boldsymbol{G}}^{(k)}(\lambda^{(k+1)})] = 0 \tag{20}$$

where

$$\begin{aligned}\tilde{\boldsymbol{K}}^{(k)} &= \boldsymbol{F}(\boldsymbol{P}^{(k)})^{\frac{1}{2}}\boldsymbol{K}\boldsymbol{F}(\boldsymbol{P}^{(k)})^{\frac{1}{2}}\\ \tilde{\boldsymbol{G}}^{(k)}(\lambda^{(k+1)}) &= [\boldsymbol{F}(\boldsymbol{P}^{(k)})^{\frac{1}{2}}\mathbf{G}(\boldsymbol{P}^{(k)},\lambda^{(k+1)})\boldsymbol{F}(\boldsymbol{P}^{(k)})^{\frac{1}{2}}]^{-\frac{1}{2}}.\end{aligned}$$

This iteration process continues until

$$|J(\boldsymbol{P}^{(k)},\lambda^{(k+1)}) - J(\boldsymbol{P}^{(k-1)},\lambda^{(k)})| < \varepsilon \tag{21}$$

is satisfied for a prescribed tolerance $\varepsilon > 0$. If the iteration is terminated at step k, then $\boldsymbol{P}^{(k)}$ is claimed to be a solution point.

IV. NUMERICAL EXAMPLE

Consider a stable 3-D SD digital filter specified by

$$\boldsymbol{\Delta}_0 = 10^{-2}\begin{bmatrix}0.00730 & 0.34297 & -0.09594 & 0.20541\\ 3.33408 & -5.73707 & 3.94939 & -1.61598\\ -1.46081 & 2.66051 & -1.68094 & 0.68022\\ 1.12651 & -1.62192 & 1.24735 & -0.55781\end{bmatrix}$$

$$\boldsymbol{\Delta}_1 = 10^{-2}\begin{bmatrix}2.81318 & -5.00467 & 3.46926 & -0.84798\\ -5.29980 & 9.24831 & -6.29206 & 2.80791\\ 4.95232 & -8.39641 & 5.73329 & -1.62170\\ 0.72029 & -1.34272 & 0.95941 & 0.54827\end{bmatrix}$$

$$\boldsymbol{\Delta}_2 = 10^{-2}\begin{bmatrix}-0.69409 & 1.54874 & -0.94779 & 0.39116\\ 3.93785 & -6.79910 & 4.66564 & -1.96344\\ -2.37995 & 4.20737 & -2.75482 & 0.95329\\ 0.70545 & -0.90615 & 0.73168 & -0.55633\end{bmatrix}$$

$$\boldsymbol{\Delta}_3 = 10^{-2}\begin{bmatrix}1.67681 & -2.69078 & 1.98218 & -0.33567\\ -0.59937 & 1.11289 & -0.71981 & 0.43504\\ 1.82472 & -2.93685 & 2.11591 & -0.43417\\ 1.28875 & -2.01749 & 1.51782 & -0.09016\end{bmatrix}$$

$$\begin{aligned}\begin{bmatrix}b_{11} & b_{12} & b_{13}\end{bmatrix} &= \begin{bmatrix}b_{31} & b_{32} & b_{33}\end{bmatrix}\\ &= \begin{bmatrix}-1.81600 & 1.23756 & -0.31382\end{bmatrix}\\ \begin{bmatrix}b_{21} & b_{22} & b_{23}\end{bmatrix} &= \begin{bmatrix}-1.81611 & 1.23775 & -0.31391\end{bmatrix}.\end{aligned}$$

The above data can be realized by a minimal state-space model in (3) as

$$\boldsymbol{A}_2 = \begin{bmatrix}0.00000 & -0.19089 & 0.29060\\ 0.74393 & -86.40470 & 133.71075\\ -0.27211 & -57.01643 & 88.22081\end{bmatrix}$$

$$\boldsymbol{B}_2 = 10^3 \begin{bmatrix} 0.00602 & -0.00921 & 0.00699 & -0.00095 \\ -1.10247 & 1.68622 & -1.27902 & 0.17267 \\ -0.71455 & 1.09291 & -0.82977 & 0.11192 \end{bmatrix}$$

$$\boldsymbol{C}_2 = \begin{bmatrix} 0.07236 & 0.06711 & -0.10298 \\ 0.01930 & 0.01789 & -0.02745 \\ 0.05887 & 0.05460 & -0.08378 \\ 0.07079 & 0.06565 & -0.10073 \end{bmatrix}.$$

Using (9) and (12) with truncation $(0,0,0) \leq (i,j,k) \leq (100,100,100)$ to evaluate the Gramians $\boldsymbol{M}_A$, $\boldsymbol{W}_B$, $\boldsymbol{K}_C$ and $\boldsymbol{K}$, we obtained the coordinate transformation matrix $\boldsymbol{T}$ as

$$\boldsymbol{T} = 10^3 \mathrm{diag}\{0.01077, 2.58588, 1.68384\}$$

which yields

$$\boldsymbol{K} = \begin{bmatrix} 1.00000 & -0.84067 & -0.83915 \\ -0.84067 & 1.00000 & 0.99999 \\ -0.83915 & 0.99999 & 1.00000 \end{bmatrix}$$

$$\boldsymbol{K}_C = 10 \begin{bmatrix} 5.70413 & -4.79529 & -4.78664 \\ -4.79529 & 5.70413 & 5.70410 \\ -4.78664 & 5.70410 & 5.70413 \end{bmatrix}$$

$$\boldsymbol{W}_B = 10^8 \begin{bmatrix} 0.00014 & 0.02404 & -0.02399 \\ 0.02404 & 4.36159 & -4.35410 \\ -0.02399 & -4.35410 & 4.34663 \end{bmatrix}$$

$$\boldsymbol{M}_A = 10^7 \begin{bmatrix} 0.00021 & 0.03478 & -0.03471 \\ 0.03478 & 5.83540 & -5.82400 \\ -0.03479 & -5.82400 & 5.81261 \end{bmatrix}.$$

In this case, the l_2-sensitivity measure was computed as

$$S = 9.87318749 \times 10^8.$$

Choosing $\boldsymbol{P}^{(0)} = \boldsymbol{I}_3$ in (19) as an intial estimate and a tolerance $\epsilon = 10^{-8}$ in (21) as well as in the bisection method, it took the Lagrange-based algorithm 15 iterations to converge to the solution

$$\boldsymbol{P}^{opt} = \begin{bmatrix} 2.267890 & -2.297027 & -2.289396 \\ -2.297027 & 3.274871 & 3.268974 \\ -2.289396 & 3.268974 & 3.263110 \end{bmatrix}$$

or equivalently,

$$\boldsymbol{T}^{opt} = \begin{bmatrix} 0.180276 & 1.360946 & -0.619043 \\ -1.018340 & -1.083999 & 1.030932 \\ -1.020420 & -1.079492 & 1.027884 \end{bmatrix}$$

and then the l_2-sensitivity measure was computed as

$$J(\boldsymbol{P}^{opt}, \lambda) = 3.243563304 \times 10^3$$

where $\lambda = 8.42923 \times 10^2$. The profiles of the l_2-sensitivity measure $J(\mathbf{P}, \lambda)$ and the Lagrange multiplier λ during the first 15 iterations are shown in Fig. 1, from which it is observed that with a tolerance $\epsilon = 10^{-8}$ the algorithm converges within 15 iterations.

V. CONCLUSION

The realization problem of a 3-D SD digital filter has been investigated so as to reduce the l_2-sensitivity under the l_2-scaling constraints. An efficient iterative technique has been developed to solve the problem. To this end, we utilize a Lagrange function as well as an efficient bisection method, and solve the constrained optimization problem directly. Our computer simulation results have demonstrated the validity and effectiveness of the proposed technique.

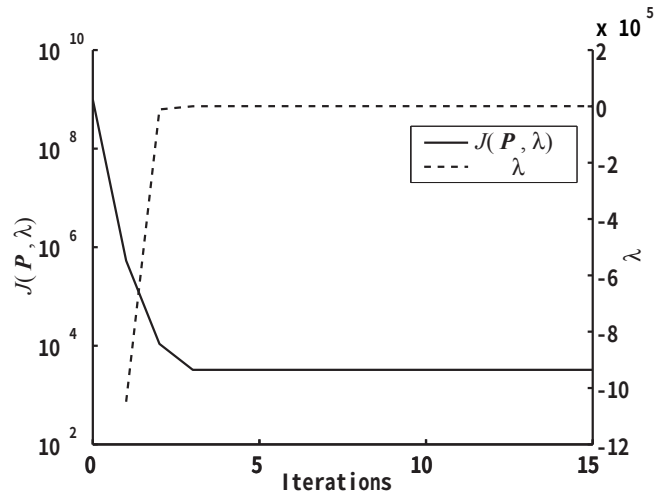

Fig. 1. Profiles of l_2-sensitivity and λ performances during the first 15 iterations.

REFERENCES

[1] A. Zilouchian and R. L. Carroll, "A coefficient sensitivity bound in 2-D state-space digital filtering," *IEEE Trans. Circuits Syst.*, vol.CAS-33, pp.665-667, June 1986.

[2] 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," *Proc. 1987 IEEE Int. Symp. Circuits Syst.*, pp.710-713.

[3] 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.

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

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

[6] 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.

[7] 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.

[8] T. Hinamoto, Y. Sugie, A. Doi and M. Muneyasu, "Synthesis of 3-D separable-denominator state-space digital filters with minimum L_2-sensitivity," *Multidimensional Systems and Signal Processing*, vol.15, pp.147-167, Apr. 2004.

[9] T. Hinamoto, K. Iwata and W.-S. Lu, "L_2-sensitivity Minimization of one- and two-dimensional state-space digital filters subject to L_2-scaling constraints," *IEEE Trans. Signal Processing*, vol. 54, pp. 1804-1812, May 2006.

[10] T. Hinamoto, T. Oumi, O. I. Omoifo and W.-S. Lu, "Minimization of frequency-weighted l_2-sensitivity subject to l_2-scaling constraints for two-dimensional state-space digital filters, " IEEE Trans. on Signal Processing, vol.56, no.10, pp.5157-5168, Oct. 2008.

[11] W.-S. Lu and A. Antoniou, "Synthesis of 2-D state-space fixed-point digital filter structures with minimum roundoff noise," *IEEE Trans. Circuits Syst.*, vol. CAS-33, pp. 965-973, Oct. 1986.