# Minimization of $L_2$-Sensitivity for 2-D Separable-Denominator State-Space Digital Filters Subject to $L_2$-Scaling Constraints

Takao Hinamoto and Yukihiro Shibata
Graduate School of Engineering
Hiroshima University
Higashi-Hiroshima 739-8527, Japan

Wu-Sheng Lu
Dept. of Elec. and Comp. Engineering
University of Victoria
Victoria, BC, Canada V8W 3P6

***Abstract*— For two-dimensional (2-D) separable-denominator state-space digital filters, we investigate the minimization problem of an $L_2$-sensitivity measure subject to $L_2$-scaling constraints. First, the coefficient sensitivity is analized by using a pure $L_2$ norm. Next, an iterative algorithm is developed for minimizing an $L_2$-sensitivity measure subject to $L_2$-scaling constraints. This approach largely relies on the use of a Lagrange function and some matrix-theoretic techniques.**

***Keywords*—2-D separable-denominator filter, Roesser's local state-space model, finite word length, $L_2$-scaling constraints, $L_2$-sensitivity minimization**

## I. INTRODUCTION

Owing to either truncation or rounding of filter coefficients, the characteristics of an actual transfer function deviate from the original in the fixed-point implementation of recursive digital filters. Several techniques have been proposed for synthesizing 2-D filter structures with low coefficient sensitivity [1]-[7]. Some of these techniques evaluate the sensitivity by using a mixture of $L_1/L_2$ norms [1]-[3], while the others rely on the use of a pure $L_2$ norm [4],[6],[7]. Moreover, minimization of weighted sensitivity for 2-D state-space digital filters has been considered in accordance with both mixed $L_1/L_2$ and pure $L_2$ sensitivity measures [5]. The $L_2$ sensitivity minimization is more natural and reasonable than the convensional $L_1/L_2$ mixed sensitivity minimization, but it is technically more challenging. Alternatively, a state-space digital filter with $L_2$-scaling constraints is beneficial for suppressing overflow oscillations [8],[9]. However, satisfactory solution methods for $L_2$-sensitivity minimization subject to $L_2$-scaling constraints are still needed [10],[11].

This paper formulates an $L_2$-sensitivity minimization problem subject to the scaling constraints for 2-D separable-denominator digital filters. By making use of a Lagrange function and some matrix-theoretic techniques, an iterative algorithm is developed to solve the constraint optimization problem directly. A numerical example is presented to demonstrate the usefulness of the proposed algorithm.

## II. SENSITIVITY ANALYSIS

Without loss of generality, a 2-D digital filter with separable denominator can be described by the Roesser local state-space (LSS) model $\{\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_4, \boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{c}_1, \boldsymbol{c}_2, d\}_{m+n}$

$$\begin{bmatrix} \boldsymbol{x}^h(i+1,j) \\ \boldsymbol{x}^v(i,j+1) \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_1 & \boldsymbol{A}_2 \\ \boldsymbol{0} & \boldsymbol{A}_4 \end{bmatrix} \begin{bmatrix} \boldsymbol{x}^h(i,j) \\ \boldsymbol{x}^v(i,j) \end{bmatrix} + \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \end{bmatrix} u(i,j)$$

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

where $\boldsymbol{x}^h(i,j)$ is an $m\times 1$ horizontal state vector, $\boldsymbol{x}^v(i,j)$ is an $n\times 1$ vertical state vector, $u(i,j)$ is a scalar input, $y(i,j)$ is a scalar output, and $\boldsymbol{A}_1$, $\boldsymbol{A}_2$, $\boldsymbol{A}_4$, $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, $\boldsymbol{c}_1$, $\boldsymbol{c}_2$, and $d$ are real constant matrices of appropriate dimensions. The LSS model in (1) is assumed to be asymptotically stable, separately locally controllable and separately locally observable. The transfer function of the LSS model in (1) is given by

$$\begin{aligned} &H(z_1,z_2) \\ &= \begin{bmatrix} 1 & \boldsymbol{c}_1(z_1\boldsymbol{I}_m - \boldsymbol{A}_1)^{-1} \end{bmatrix} \\ &\qquad \cdot \begin{bmatrix} d & \boldsymbol{c}_2 \\ \boldsymbol{b}_1 & \boldsymbol{A}_2 \end{bmatrix} \begin{bmatrix} 1 \\ (z_2\boldsymbol{I}_n - \boldsymbol{A}_4)^{-1}\boldsymbol{b}_2 \end{bmatrix}. \end{aligned} \tag{2}$$

*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 w.r.t. all the entries of $\boldsymbol{X}$. The sensitivity function of $f$ with respect to $\boldsymbol{X}$ is then defined as

$$\boldsymbol{S}_X = \frac{\partial f}{\partial \boldsymbol{X}} \quad \text{with} \quad (\boldsymbol{S}_X)_{ij} = \frac{\partial f}{\partial x_{ij}} \tag{3}$$

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

With these notations, it is easy to show that

$$\begin{aligned} \frac{\partial H(z_1,z_2)}{\partial \boldsymbol{A}_1} &= \boldsymbol{Q}^T(z_1)\boldsymbol{F}^T(z_1,z_2) \\ \frac{\partial H(z_1,z_2)}{\partial \boldsymbol{A}_2} &= \boldsymbol{Q}^T(z_1)\boldsymbol{P}^T(z_2) \\ \frac{\partial H(z_1,z_2)}{\partial \boldsymbol{A}_4} &= \boldsymbol{G}^T(z_1,z_2)\boldsymbol{P}^T(z_2) \\ \frac{\partial H(z_1,z_2)}{\partial \boldsymbol{b}_1} &= \boldsymbol{Q}^T(z_1), \quad \frac{\partial H(z_1,z_2)}{\partial \boldsymbol{b}_2} = \boldsymbol{G}^T(z_1,z_2) \\ \frac{\partial H(z_1,z_2)}{\partial \boldsymbol{c}_1^T} &= \boldsymbol{F}(z_1,z_2), \quad \frac{\partial H(z_1,z_2)}{\partial \boldsymbol{c}_2^T} = \boldsymbol{P}(z_2) \end{aligned}$$

where

$$\boldsymbol{F}(z_1,z_2) = (z_1\boldsymbol{I}_m - \boldsymbol{A}_1)^{-1}[\boldsymbol{b}_1 + \boldsymbol{A}_2\boldsymbol{P}(z_2)]$$
$$\boldsymbol{G}(z_1,z_2) = [\boldsymbol{c}_2 + \boldsymbol{Q}(z_1)\boldsymbol{A}_2](z_2\boldsymbol{I}_n - \boldsymbol{A}_4)^{-1}$$
$$\boldsymbol{P}(z_2) = (z_2\boldsymbol{I}_n - \boldsymbol{A}_4)^{-1}\boldsymbol{b}_2, \;\; \boldsymbol{Q}(z_1) = \boldsymbol{c}_1(z_1\boldsymbol{I}_m - \boldsymbol{A}_1)^{-1}.$$

The term $d$ and the sensitivity with respect to it are coordinate independent, 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 $L_p$ norm of $\boldsymbol{X}(z_1,z_2)$ is then defined as

$$||\boldsymbol{X}||_p = \left[\frac{1}{(2\pi j)^2}\oint\oint_{\Gamma^2} ||\boldsymbol{X}(z_1,z_2)||_F^p \frac{dz_1dz_2}{z_1z_2}\right]^{1/p} \quad (4)$$

where $||\boldsymbol{X}(z_1,z_2)||_F$ is the Frobenius norm of the matrix $\boldsymbol{X}(z_1,z_2)$ defined by

$$||\boldsymbol{X}(z_1,z_2)||_F = \left[\sum_{p=1}^{m}\sum_{q=1}^{n}|x_{pq}(z_1,z_2)|^2\right]^{1/2}.$$

The overall $L_2$-sensitivity measure is now defined by

$$M_2 = \left|\left|\frac{\partial H(z_1,z_2)}{\partial \boldsymbol{A}_1}\right|\right|_2^2 + \left|\left|\frac{\partial H(z_1,z_2)}{\partial \boldsymbol{A}_4}\right|\right|_2^2 + \left|\left|\frac{\partial H(z_1,z_2)}{\partial \boldsymbol{b}_1}\right|\right|_2^2 + \left|\left|\frac{\partial H(z_1,z_2)}{\partial \boldsymbol{b}_2}\right|\right|_2^2 + \left|\left|\frac{\partial H(z_1,z_2)}{\partial \boldsymbol{c}_1^T}\right|\right|_2^2 + \left|\left|\frac{\partial H(z_1,z_2)}{\partial \boldsymbol{c}_2^T}\right|\right|_2^2 + \left|\left|\frac{\partial H(z_1,z_2)}{\partial \boldsymbol{A}_2}\right|\right|_2^2. \quad (5)$$

From (4)-(5), it follows that

$$M_2 = \mathrm{tr}\Big[\boldsymbol{M}_{A_1} + \boldsymbol{M}_{A_4} + \boldsymbol{W}^h + \boldsymbol{W}^v + \boldsymbol{K}^h + \boldsymbol{K}^v\Big] + \mathrm{tr}[\boldsymbol{W}^h]\mathrm{tr}[\boldsymbol{K}^v]$$

where

$$\boldsymbol{M}_{A_1} = \frac{1}{(2\pi j)^2}\oint_{|z_1|=1}\oint_{|z_2|=1}[\boldsymbol{F}(z_1^{-1},z_2^{-1})\boldsymbol{Q}(z_1^{-1})] \cdot [\boldsymbol{Q}^T(z_1)\boldsymbol{F}^T(z_1,z_2)]\frac{dz_1dz_2}{z_1z_2}$$
$$\boldsymbol{M}_{A_4} = \frac{1}{(2\pi j)^2}\oint_{|z_1|=1}\oint_{|z_2|=1}[\boldsymbol{G}^T(z_1,z_2)\boldsymbol{P}^T(z_2)] \cdot [\boldsymbol{P}(z_2^{-1})\boldsymbol{G}(z_1^{-1},z_2^{-1})]\frac{dz_1dz_2}{z_1z_2}$$

$$\boldsymbol{K}^h = \frac{1}{(2\pi j)^2}\oint_{|z_1|=1}\oint_{|z_2|=1}\boldsymbol{F}(z_1,z_2)\boldsymbol{F}^*(z_1,z_2)\frac{dz_1dz_2}{z_1z_2}$$
$$\boldsymbol{K}^v = \frac{1}{2\pi j}\oint_{|z_2|=1}\boldsymbol{P}(z_2)\boldsymbol{P}^*(z_2)\frac{dz_2}{z_2}$$
$$\boldsymbol{W}^h = \frac{1}{2\pi j}\oint_{|z_1|=1}\boldsymbol{Q}^*(z_1)\boldsymbol{Q}(z_1)\frac{dz_1}{z_1}$$
$$\boldsymbol{W}^v = \frac{1}{(2\pi j)^2}\oint_{|z_1|=1}\oint_{|z_2|=1}\boldsymbol{G}^*(z_1,z_2)\boldsymbol{G}(z_1,z_2)\frac{dz_1dz_2}{z_1z_2}.$$

The matrices $\boldsymbol{K} = \boldsymbol{K}^h \oplus \boldsymbol{K}^v$ and $\boldsymbol{W} = \boldsymbol{W}^h \oplus \boldsymbol{W}^v$ are called the local controllability Gramian and local obsevability Gramian, respectively, and can be obtained by solving the following Lyapunov equations:

$$\boldsymbol{K}^v = \boldsymbol{A}_4\boldsymbol{K}^v\boldsymbol{A}_4^T + \boldsymbol{b}_2\boldsymbol{b}_2^T, \quad \boldsymbol{W}^h = \boldsymbol{A}_1^T\boldsymbol{W}^h\boldsymbol{A}_1 + \boldsymbol{c}_1^T\boldsymbol{c}_1$$
$$\boldsymbol{K}^h = \boldsymbol{A}_1\boldsymbol{K}^h\boldsymbol{A}_1^T + \boldsymbol{A}_2\boldsymbol{K}^v\boldsymbol{A}_2^T + \boldsymbol{b}_1\boldsymbol{b}_1^T$$
$$\boldsymbol{W}^v = \boldsymbol{A}_4^T\boldsymbol{W}^v\boldsymbol{A}_4 + \boldsymbol{A}_2^T\boldsymbol{W}^h\boldsymbol{A}_2 + \boldsymbol{c}_2^T\boldsymbol{c}_2.$$

Applying the eigenvalue-eigenvector decompositions

$$\boldsymbol{K}^v = \sum_{i=1}^{n}\sigma_i^v\boldsymbol{u}_i\boldsymbol{u}_i^T, \qquad \boldsymbol{W}^h = \sum_{i=1}^{m}\sigma_i^h\boldsymbol{v}_i\boldsymbol{v}_i^T \quad (6)$$

where $\sigma_i^v$ and $\boldsymbol{u}_i$ ($\sigma_i^h$ and $\boldsymbol{v}_i$) are the $i$th eigenvalue and eigenvector of $\boldsymbol{K}^v$ ($\boldsymbol{W}^h$), respectively, we can write (6) as [7]

$$M_2 = \sum_{i=0}^{n}\sigma_i^v\mathrm{tr}[\boldsymbol{W}_i^h(\boldsymbol{I}_m)] + \sum_{i=0}^{m}\sigma_i^h\mathrm{tr}[\boldsymbol{K}_i^v(\boldsymbol{I}_n)] + \mathrm{tr}[\boldsymbol{W}^h + \boldsymbol{W}^v + \boldsymbol{K}^h + \boldsymbol{K}^v] + \mathrm{tr}[\boldsymbol{W}^h]\mathrm{tr}[\boldsymbol{K}^v] \quad (7)$$

where $\sigma_0^v = \sigma_0^h = 1$, $\tilde{\boldsymbol{u}}_0 = \boldsymbol{b}_1$, $\tilde{\boldsymbol{u}}_i = \boldsymbol{A}_2\boldsymbol{u}_i\,(i \geq 1)$, $\tilde{\boldsymbol{v}}_0 = \boldsymbol{c}_2^T$, $\tilde{\boldsymbol{v}}_i = \boldsymbol{A}_2^T\boldsymbol{v}_i\,(i \geq 1)$, and $m\times m$ matrix $\boldsymbol{W}_i^h(\boldsymbol{P}_1)$ and $n\times n$ matrix $\boldsymbol{K}_i^v(\boldsymbol{P}_4)$ are obtained by solving the Lyapunov equations

$$\begin{bmatrix}\boldsymbol{W}_i^h(\boldsymbol{P}_1) & * \\ * & *\end{bmatrix} = \begin{bmatrix}\boldsymbol{A}_1 & \tilde{\boldsymbol{u}}_i\boldsymbol{c}_1 \\ \boldsymbol{0} & \boldsymbol{A}_1\end{bmatrix}\begin{bmatrix}\boldsymbol{W}_i^h(\boldsymbol{P}_1) & * \\ * & *\end{bmatrix} \cdot \begin{bmatrix}\boldsymbol{A}_1 & \tilde{\boldsymbol{u}}_i\boldsymbol{c}_1 \\ \boldsymbol{0} & \boldsymbol{A}_1\end{bmatrix}^T + \begin{bmatrix}\boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{P}_1\end{bmatrix}$$

$$\begin{bmatrix}\boldsymbol{K}_i^v(\boldsymbol{P}_4) & * \\ * & *\end{bmatrix} = \begin{bmatrix}\boldsymbol{A}_4 & \boldsymbol{0} \\ \boldsymbol{b}_2\tilde{\boldsymbol{v}}_i^T & \boldsymbol{A}_4\end{bmatrix}^T\begin{bmatrix}\boldsymbol{K}_i^v(\boldsymbol{P}_4) & * \\ * & *\end{bmatrix} \cdot \begin{bmatrix}\boldsymbol{A}_4 & \boldsymbol{0} \\ \boldsymbol{b}_2\tilde{\boldsymbol{v}}_i^T & \boldsymbol{A}_4\end{bmatrix} + \begin{bmatrix}\boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{P}_4^{-1}\end{bmatrix}.$$

## III. SENSITIVITY MINIMIZATION

The following class of state-space coordinate transformations can be used without affecting the input-output map:

$$\begin{bmatrix}\overline{\boldsymbol{x}}^h(i,j) \\ \overline{\boldsymbol{x}}^v(i,j)\end{bmatrix} = \begin{bmatrix}\boldsymbol{T}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{T}_4\end{bmatrix}^{-1}\begin{bmatrix}\boldsymbol{x}^h(i,j) \\ \boldsymbol{x}^v(i,j)\end{bmatrix} \quad (8)$$

where $\boldsymbol{T}_1$ and $\boldsymbol{T}_4$ are $m\times m$ and $n\times n$ nonsingular constant matrices, respectively. Performing this coordinate transformation to the LSS model in (1) yields a new realization $\{\overline{\boldsymbol{A}}_1, \overline{\boldsymbol{A}}_2, \overline{\boldsymbol{A}}_4, \overline{\boldsymbol{b}}_1, \overline{\boldsymbol{b}}_2, \overline{\boldsymbol{c}}_1, \overline{\boldsymbol{c}}_2, d\}_{m+n}$ characterized by

$$\overline{\boldsymbol{A}}_1 = \boldsymbol{T}_1^{-1}\boldsymbol{A}_1\boldsymbol{T}_1, \qquad \overline{\boldsymbol{A}}_2 = \boldsymbol{T}_1^{-1}\boldsymbol{A}_2\boldsymbol{T}_4$$
$$\overline{\boldsymbol{A}}_4 = \boldsymbol{T}_4^{-1}\boldsymbol{A}_4\boldsymbol{T}_4, \qquad \overline{\boldsymbol{b}}_1 = \boldsymbol{T}_1^{-1}\boldsymbol{b}_1$$
$$\overline{\boldsymbol{b}}_2 = \boldsymbol{T}_4^{-1}\boldsymbol{b}_2, \qquad \overline{\boldsymbol{c}}_1 = \boldsymbol{c}_1\boldsymbol{T}_1, \qquad \overline{\boldsymbol{c}}_2 = \boldsymbol{c}_2\boldsymbol{T}_4$$
$$\overline{\boldsymbol{K}}^h = \boldsymbol{T}_1^{-1}\boldsymbol{K}^h\boldsymbol{T}_1^{-T}, \quad \overline{\boldsymbol{K}}^v = \boldsymbol{T}_4^{-1}\boldsymbol{K}^v\boldsymbol{T}_4^{-T}$$
$$\overline{\boldsymbol{W}}^h = \boldsymbol{T}_1^T\boldsymbol{W}^h\boldsymbol{T}_1, \qquad \overline{\boldsymbol{W}}^v = \boldsymbol{T}_4^T\boldsymbol{W}^v\boldsymbol{T}_4.$$

For the new realization, the $M_2$ in (7) is changed to

$$M_2(\boldsymbol{P}) = \sum_{i=0}^{n} \sigma_i^v \mathrm{tr}[\boldsymbol{W}_i^h(\boldsymbol{P}_1)\boldsymbol{P}_1^{-1}] + \sum_{i=0}^{m} \sigma_i^h \mathrm{tr}[\boldsymbol{K}_i^v(\boldsymbol{P}_4)\boldsymbol{P}_4] + \mathrm{tr}[\boldsymbol{W}^h\boldsymbol{P}_1 + \boldsymbol{W}^v\boldsymbol{P}_4 + \boldsymbol{K}^h\boldsymbol{P}_1^{-1} + \boldsymbol{K}^v\boldsymbol{P}_4^{-1}] + \mathrm{tr}[\boldsymbol{W}^h\boldsymbol{P}_1]\mathrm{tr}[\boldsymbol{K}^v\boldsymbol{P}_4^{-1}]$$

where $\boldsymbol{P} = \boldsymbol{P}_1 \oplus \boldsymbol{P}_4$ and $\boldsymbol{P}_i = \boldsymbol{T}_i\boldsymbol{T}_i^T$ for $i = 1, 4$.

If $l_2$-norm dynamic-range scaling constraints are imposed on the local state vector $[\overline{\boldsymbol{x}}^h(i,j)^T, \overline{\boldsymbol{x}}^v(i,j)^T]^T$, then

$$(\overline{\boldsymbol{K}}^h)_{ii} = (\boldsymbol{T}_1^{-1}\boldsymbol{K}^h\boldsymbol{T}_1^{-T})_{ii} = 1$$

$$(\overline{\boldsymbol{K}}^v)_{jj} = (\boldsymbol{T}_4^{-1}\boldsymbol{K}^v\boldsymbol{T}_4^{-T})_{jj} = 1$$

are required for $i = 1, 2, \cdots, m$ and $j = 1, 2, \cdots, n$.

The problem of minimizing $M_2(\boldsymbol{P})$ in (9) subject to the constraints in (9) is a constrained nonlinear optimization problem where the variable is matrix $\boldsymbol{P}$. If we sum up the $m$ or $n$ constraints in (9), then we have

$$\mathrm{tr}[\boldsymbol{K}^h\boldsymbol{P}_1^{-1}] = m, \qquad \mathrm{tr}[\boldsymbol{K}^v\boldsymbol{P}_4^{-1}] = n. \tag{9}$$

Consequently, the problem of minimizing (9) subject to the constraints in (9) can be *relaxed* into the problem

$$\begin{aligned} &\text{minimize } M_2(\boldsymbol{P}) \text{ in (9)} \\ &\text{subject to } \mathrm{tr}[\boldsymbol{K}^h\boldsymbol{P}_1^{-1}] = m \text{ and } \mathrm{tr}[\boldsymbol{K}^v\boldsymbol{P}_4^{-1}] = n. \end{aligned} \tag{10}$$

To solve (10), define a Lagrange function of the problem as

$$\begin{aligned} J(\boldsymbol{P}, \lambda_1, \lambda_4) &= M_2(\boldsymbol{P}) + \lambda_1(\mathrm{tr}[\boldsymbol{K}^h\boldsymbol{P}_1^{-1}] - m) \\ &\quad + \lambda_4(\mathrm{tr}[\boldsymbol{K}^v\boldsymbol{P}_4^{-1}] - n) \end{aligned} \tag{11}$$

where $\lambda_1$ and $\lambda_4$ are Lagrange multipliers. It is well known that the solution of problem (10) must satisfy the Karush-Kuhn-Tucker (KKT) conditions $\partial J(\boldsymbol{P}, \lambda_1, \lambda_4)/\partial \boldsymbol{P}_i = \boldsymbol{0}$ and $\partial J(\boldsymbol{P}, \lambda_1, \lambda_4)/\partial \lambda_i = 0$ for $i = 1, 4$ where the gradients are found to be

$$\begin{aligned} \frac{\partial J(\boldsymbol{P}, \lambda_1, \lambda_4)}{\partial \boldsymbol{P}_1} &= \boldsymbol{F}_1(\boldsymbol{P}) - \boldsymbol{P}_1^{-1}\boldsymbol{F}_2(\boldsymbol{P}_1, \lambda_1)\boldsymbol{P}_1^{-1} \\ \frac{\partial J(\boldsymbol{P}, \lambda_1, \lambda_4)}{\partial \boldsymbol{P}_4} &= \boldsymbol{F}_3(\boldsymbol{P}_4) - \boldsymbol{P}_4^{-1}\boldsymbol{F}_4(\boldsymbol{P}, \lambda_4)\boldsymbol{P}_4^{-1} \\ \frac{\partial J(\boldsymbol{P}, \lambda_1, \lambda_4)}{\partial \lambda_1} &= \mathrm{tr}[\boldsymbol{K}^h\boldsymbol{P}_1^{-1}] - m \\ \frac{\partial J(\boldsymbol{P}, \lambda_1, \lambda_4)}{\partial \lambda_4} &= \mathrm{tr}[\boldsymbol{K}^v\boldsymbol{P}_4^{-1}] - n \end{aligned}$$

with

$$\begin{aligned} \boldsymbol{F}_1(\boldsymbol{P}) &= \sum_{i=0}^{n} \sigma_i^v \boldsymbol{K}_i^h(\boldsymbol{P}_1) + (1 + \mathrm{tr}[\boldsymbol{K}^v\boldsymbol{P}_4^{-1}])\boldsymbol{W}^h \\ \boldsymbol{F}_2(\boldsymbol{P}_1, \lambda_1) &= \sum_{i=0}^{n} \sigma_i^v \boldsymbol{W}_i^h(\boldsymbol{P}_1) + (\lambda_1 + 1)\boldsymbol{K}^h \\ \boldsymbol{F}_3(\boldsymbol{P}_4) &= \sum_{i=0}^{m} \sigma_i^h \boldsymbol{K}_i^v(\boldsymbol{P}_4) + \boldsymbol{W}^v \\ \boldsymbol{F}_4(\boldsymbol{P}, \lambda_4) &= \sum_{i=0}^{m} \sigma_i^h \boldsymbol{W}_i^v(\boldsymbol{P}_4) + (\lambda_4 + 1 + \mathrm{tr}[\boldsymbol{W}^h\boldsymbol{P}_1])\boldsymbol{K}^v \end{aligned}$$

$$\begin{bmatrix} \boldsymbol{K}_i^h(\boldsymbol{P}_1) & * \\ * & * \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_1 & \boldsymbol{0} \\ \tilde{\boldsymbol{u}}_i\boldsymbol{c}_1 & \boldsymbol{A}_1 \end{bmatrix}^T \begin{bmatrix} \boldsymbol{K}_i^h(\boldsymbol{P}_1) & * \\ * & * \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{A}_1 & \boldsymbol{0} \\ \tilde{\boldsymbol{u}}_i\boldsymbol{c}_1 & \boldsymbol{A}_1 \end{bmatrix} + \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{P}_1^{-1} \end{bmatrix}$$

$$\begin{bmatrix} \boldsymbol{W}_i^v(\boldsymbol{P}_4) & * \\ * & * \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_4 & \boldsymbol{b}_2\tilde{\boldsymbol{v}}_i^T \\ \boldsymbol{0} & \boldsymbol{A}_4 \end{bmatrix} \begin{bmatrix} \boldsymbol{W}_i^v(\boldsymbol{P}_4) & * \\ * & * \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{A}_4 & \boldsymbol{b}_2\tilde{\boldsymbol{v}}_i^T \\ \boldsymbol{0} & \boldsymbol{A}_4 \end{bmatrix}^T + \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{P}_4 \end{bmatrix}.$$

Hence the KKT conditions in (12) become

$$\begin{aligned} \boldsymbol{P}_1\boldsymbol{F}_1(\boldsymbol{P})\boldsymbol{P}_1 &= \boldsymbol{F}_2(\boldsymbol{P}_1, \lambda_1), \quad \mathrm{tr}[\boldsymbol{K}^h\boldsymbol{P}_1^{-1}] = m \\ \boldsymbol{P}_4\boldsymbol{F}_3(\boldsymbol{P}_4)\boldsymbol{P}_4 &= \boldsymbol{F}_4(\boldsymbol{P}, \lambda_4), \quad \mathrm{tr}[\boldsymbol{K}^v\boldsymbol{P}_4^{-1}] = n\,. \end{aligned} \tag{12}$$

The two equations on the left-hand side in (12) are highly nonlinear with respect to $\boldsymbol{P}_1$ and $\boldsymbol{P}_4$. An effective approach to solving the first two equations in (12) is to *relax* them into the following recursive second-order matrix equations:

$$\begin{aligned} \boldsymbol{P}_1^{(i+1)}\boldsymbol{F}_1(\boldsymbol{P}^{(i)})\boldsymbol{P}_1^{(i+1)} &= \boldsymbol{F}_2(\boldsymbol{P}_1^{(i)}, \lambda_1^{(i)}) \\ \boldsymbol{P}_4^{(i+1)}\boldsymbol{F}_3(\boldsymbol{P}_4^{(i)})\boldsymbol{P}_4^{(i+1)} &= \boldsymbol{F}_4(\boldsymbol{P}^{(i)}, \lambda_4^{(i)}) \end{aligned} \tag{13}$$

where $\boldsymbol{P}_1^{(i)}$, $\boldsymbol{P}_4^{(i)}$, $\lambda_1^{(i)}$ and $\lambda_4^{(i)}$ are known from the previous recursion. The solutions $\boldsymbol{P}_1^{(i+1)}$ and $\boldsymbol{P}_4^{(i+1)}$ of (13) are given by

$$\begin{aligned} &\boldsymbol{P}_1^{(i+1)} = \\ &\boldsymbol{F}_1^{-\frac{1}{2}}(\boldsymbol{P}^{(i)})[\boldsymbol{F}_1^{\frac{1}{2}}(\boldsymbol{P}^{(i)})\boldsymbol{F}_2(\boldsymbol{P}_1^{(i)}, \lambda_1^{(i)})\,\boldsymbol{F}_1^{\frac{1}{2}}(\boldsymbol{P}^{(i)})]^{\frac{1}{2}}\boldsymbol{F}_1^{-\frac{1}{2}}(\boldsymbol{P}^{(i)}) \\ &\boldsymbol{P}_4^{(i+1)} = \\ &\boldsymbol{F}_3^{-\frac{1}{2}}(\boldsymbol{P}_4^{(i)})[\boldsymbol{F}_3^{\frac{1}{2}}(\boldsymbol{P}_4^{(i)})\boldsymbol{F}_4(\boldsymbol{P}^{(i)}, \lambda_4^{(i)})\,\boldsymbol{F}_3^{\frac{1}{2}}(\boldsymbol{P}_4^{(i)})]^{\frac{1}{2}}\boldsymbol{F}_3^{-\frac{1}{2}}(\boldsymbol{P}_4^{(i)}), \end{aligned} \tag{14}$$

respectively. To derive recursive formulas for the Lagrange multipliers $\lambda_1$ and $\lambda_4$, we employ (12) to write

$$\begin{aligned} \mathrm{tr}[\boldsymbol{P}_1F_1(\boldsymbol{P})]) &= \sum_{i=0}^{n} \sigma_i^v \mathrm{tr}[\boldsymbol{W}_i^h(\boldsymbol{P}_1)\boldsymbol{P}_1^{-1}] + m(\lambda_1 + 1) \\ \mathrm{tr}[\boldsymbol{P}_4F_3(\boldsymbol{P}_4)] &= \sum_{i=0}^{m} \sigma_i^h \mathrm{tr}[\boldsymbol{W}_i^v(\boldsymbol{P}_4)\boldsymbol{P}_4^{-1}] \\ &\quad + n(\lambda_4 + 1 + \mathrm{tr}[\boldsymbol{W}^h\boldsymbol{P}_1]) \end{aligned} \tag{15}$$

which naturally suggest the recursions for $\lambda_1$ and $\lambda_4$:

$$\begin{aligned} \lambda_1^{(i+1)} &= \frac{\mathrm{tr}[\boldsymbol{P}_1^{(i)}\boldsymbol{F}_1(\boldsymbol{P}^{(i)})] - \sum_{i=0}^{n} \sigma_i^v \mathrm{tr}[\boldsymbol{W}_i^h(\boldsymbol{P}_1^{(i)})\boldsymbol{P}_1^{(i)-1}]}{m} - 1 \\ \lambda_4^{(i+1)} &= \frac{\mathrm{tr}[\boldsymbol{P}_4^{(i)}\boldsymbol{F}_3(\boldsymbol{P}_4^{(i)})] - \sum_{i=0}^{m} \sigma_i^h \mathrm{tr}[\boldsymbol{W}_i^v(\boldsymbol{P}_4^{(i)})\boldsymbol{P}_4^{(i)-1}]}{n} - 1 \\ &\quad - \mathrm{tr}[\boldsymbol{W}^h\boldsymbol{P}_1^{(i)}]. \end{aligned} \tag{16}$$

The iteration process starts with $\boldsymbol{P}^{(0)} = \boldsymbol{I}_{m+n}$ and any values of $\lambda_1^{(0)} > 0$ and $\lambda_4^{(0)} > 0$, and continues until (12) is satisfied within a prescribed numerical tolerance.

Having obtained the optimal $\boldsymbol{P} = \boldsymbol{P}_1 \oplus \boldsymbol{P}_4$ and noticing $\boldsymbol{P} = \boldsymbol{T}\boldsymbol{T}^T$, the optimal coordinate-transformation matrix $\boldsymbol{T} = \boldsymbol{T}_1 \oplus \boldsymbol{T}_4$ satisfying the constraints in (9) can now be readily determined using the technique described in [11].

## IV. ILLUSTRATIVE EXAMPLE

As an example, consider a 2-D separable-denominator state-space digital filter in (1) specified by

$$\boldsymbol{A}_1 = \begin{bmatrix} 0.0 & 1.0 & 0.0 \\ 0.0 & 0.0 & 1.0 \\ 0.599655 & -1.836929 & 2.173645 \end{bmatrix}$$

$$\boldsymbol{A}_2 = \begin{bmatrix} 0.064564 & 0.033034 & 0.012881 \\ 0.091213 & 0.110512 & 0.102759 \\ 0.097256 & 0.151864 & 0.172460 \end{bmatrix}$$

$$\boldsymbol{A}_4 = \begin{bmatrix} 0.0 & 0.0 & 0.564961 \\ 1.0 & 0.0 & -1.887939 \\ 0.0 & 1.0 & 2.280029 \end{bmatrix}, \quad \boldsymbol{b}_1 = \begin{bmatrix} 0.047053 \\ 0.062274 \\ 0.060436 \end{bmatrix}$$

$$\boldsymbol{b}_2 = \begin{bmatrix} 1.0 & 0.0 & 0.0 \end{bmatrix}^T, \quad \boldsymbol{c}_1 = \begin{bmatrix} 1.0 & 0.0 & 0.0 \end{bmatrix}$$

$$\boldsymbol{c}_2 = \begin{bmatrix} 0.016556 & 0.012550 & 0.008243 \end{bmatrix}, \; d = 0.019421.$$

After carrying out the $L_2$-scaling for the above LSS model with a diagonal coordinate-transformation matrix, the $L_2$-sensitivity of the scaled LSS model was found to be $M_2 = 4526.0790$. Profiles of the $L_2$-sensitivity, parameters $\lambda_1$ and $\lambda_4$ as well as tr$[\boldsymbol{K}^h\boldsymbol{P}_1^{-1}]$ and tr$[\boldsymbol{K}^v\boldsymbol{P}_4^{-1}]$ during the first 500 iterations of the proposed algorithm are shown in Figs. 1 and 2, respectively. Together these figures clearly reveal a two-stage convergence behavior of the algorithm in that the first stage (which consists of just one iteration) of the algorithm reduces the $L_2$-sensitivity drastically without maintaining the constraint tr$[\boldsymbol{K}^h\boldsymbol{P}_1^{-1}] = 3$ and tr$[\boldsymbol{K}^v\boldsymbol{P}_4^{-1}] = 3$, and the second stage of the algorithm is able to restore the constraints tr$[\boldsymbol{K}^h\boldsymbol{P}_1^{-1}] = 3$ and tr$[\boldsymbol{K}^v\boldsymbol{P}_4^{-1}] = 3$ while further reducing the $L_2$-sensitivity slightly. The $L_2$-sensitivity became $M_2(\boldsymbol{P}) = 101.0064$ after 500 iterations of the algorithm.

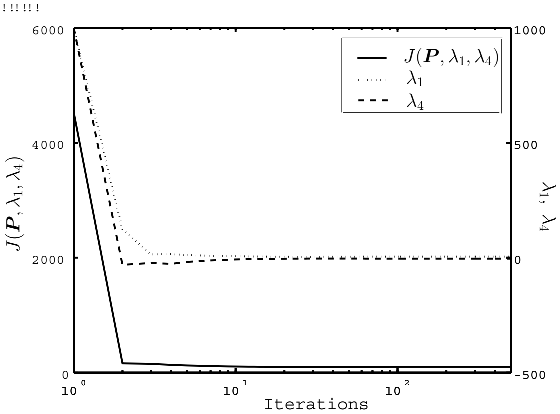


Fig. 1. $L_2$-Sensitivity, $\lambda_1$ and $\lambda_4$ Performances.

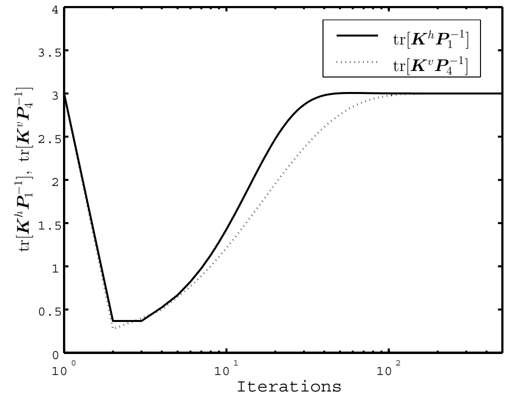


Fig. 2. tr$[\boldsymbol{K}^h\boldsymbol{P}_1^{-1}]$ and tr$[\boldsymbol{K}^v\boldsymbol{P}_4^{-1}]$ Performances.

## V. CONCLUSION

This paper has developed an iterative algorithm for minimizing the $L_2$ sensitivity measure subject to the $L_2$-scaling constraints for 2-D state-space digital filters with separable denominator. This relies on the use of a Lagrange function and some matrix-theoretic techniques. The results of a numerical example have demonstrated the effectiveness of the proposed technique.

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

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

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] T. Hinamoto, K. Iwata and W.-S. Lu, "State-space digital filters with minimum $L_2$-sensitivity subject to $L_2$-scaling constraints," in *Proc. 2005 IEEE Int. Conf. Acoust., Speech, Signal Processing*, vol.4, pp.265-268.