Analysis and Minimization of L_2-Sensitivity for Linear Systems and Two-Dimensional State-Space Filters Using General Controllability and Observability Gramians

Takao Hinamoto, *Fellow, IEEE*, Shuichi Yokoyama, Takuya Inoue, Wu Zeng, and Wu-Sheng Lu, *Fellow, IEEE*

***Abstract*—A novel expression for the evaluation of L_2-sensitivity is developed for the cases of linear discrete-time systems, linear continuous-time systems, and two-dimensional (2-D) state-space digital filters. This is accomplished by introducing the concept of general controlability and observability Gramians in each case. Moreover, the L_2-sensitivity measures obtained here contain the conventional L_1/L_2-sensitivity measures as a special case. An iterative procedure for constructing the optimal coordinate transformation matrix that minimizes the L_2-sensitivity measure is then presented in each case. This procedure is advantageous since the initial estimate and the estimate at each iteration can be calculated analytically. Finally, three numerical examples are given to illustrate the utility of the proposed techniques.**

***Index Terms*—L_2-sensitivity, linear continuous-time systems, linear discrete-time systems, optimal realization, sensitivity minimization, two-dimensional state-space digital filters.**

I. Introduction

ONE OF THE primary finite-length register effects in fixed point digital filters is changes in the input–output description of the filter due to approximating real-number parameters with a finite binary representation. Such an effect is loosely called coefficient sensitivity and has been an important research topic since coefficient truncation or rounding may cause an originally stable filter to be an unstable one. It is well-known that the undesirable finite-word-length (FWL) effects can be reduced considerably by the appropriate selection of the filter structure. Several techniques for synthesizing linear discrete-time systems that minimize the coefficient sensitivity have been reported in [1]–[9]. These techniques can be divided into two main classes: L_1/L_2-sensitivity minimization [1]–[5] and L_2-sensitivity minimization [6]–[9]. It has been pointed out [6]–[9] that the sensitivity measure based on the L_2 norm only is natural and reasonable. The L_1/L_2-sensitivity minimization has also been considered in linear continuous-time systems [10]. The problem of minimizing the coefficient sensitivity measure evaluated by using a mixture of L_1/L_2 norms has been studied for two-dimensional (2-D) state-space digital filters [11]–[15]. Recently, the L_2-sensitivity minimization problem has also been treated in the 2-D case [16]–[18].

This paper addresses the L_2-sensitivity minimization problem for the cases of linear discrete-time systems, linear continuous-time systems, and 2-D state-space digital filters. To this end, we introduce the concept of general controllability and observability Gramians, and derive a novel expression for evaluating the L_2-sensitivity for each case. This expression includes the conventional L_1/L_2-sensitivity measure as a special case. Next, an iterative procedure for the minimization of an L_2-sensitivity measure with respect to a positive-definite and symmetric matrix is presented in each case. This iteration procedure can be performed analytically. Three numerical examples are presented to demonstrate that the proposed algorithms offer faster convergence as well as improved solutions.

Throughout this paper, $\boldsymbol{I}$ denotes the identity matrix of appropriate dimension. The transpose (conjugate transpose) of a matrix $\boldsymbol{A}$ is indicated by $\boldsymbol{A}^T$ ($\boldsymbol{A}^*$), and $\mathrm{tr}[\boldsymbol{A}]$ and $\oplus$ are used to denote the trace of a square matrix $\boldsymbol{A}$ and the direct sum of matrices, respectively.

II. L_2-Sensitivity Minimization of Linear Systems

A. Linear Discrete-Time Systems

Let $(\boldsymbol{A}, \boldsymbol{b}, \boldsymbol{c}, d)_n$ be a state-space description of a stable, controllable, and observable linear discrete-time system, i.e.,

$$\begin{aligned} \boldsymbol{x}(k+1) &= \boldsymbol{A}\boldsymbol{x}(k) + \boldsymbol{b}u(k) \\ y(k) &= \boldsymbol{c}\boldsymbol{x}(k) + du(k) \end{aligned} \tag{1}$$

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

$$H(z) = \boldsymbol{c}(z\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{b} + d. \tag{2}$$

Suppose that (1) is implemented by finite-word-length (FWL) fixed-point arithmetic with a B-bit fractional representation and is realized by $(\tilde{\boldsymbol{A}}, \tilde{\boldsymbol{b}}, \tilde{\boldsymbol{c}}, \tilde{d})_n$ with

$$\begin{aligned} \tilde{\boldsymbol{A}} &= \boldsymbol{A} + \Delta\boldsymbol{A} \quad \tilde{\boldsymbol{b}} = \boldsymbol{b} + \Delta\boldsymbol{b} \\ \tilde{\boldsymbol{c}} &= \boldsymbol{c} + \Delta\boldsymbol{c} \quad \tilde{d} = d + \Delta d \end{aligned} \tag{3}$$

Manuscript received May 21, 2001; revised April 15, 2002. This paper was recommended by Associate Editor R. W. Newcomb.

T. Hinamoto, S. Yokoyama, T. Inoue, and W. Zeng are with the Graduate School of Engineering, Hiroshima University, Higashi-Hiroshima 739-8527, Japan.

W.-S. Lu is with the Department of Electrical and Computer Engineering, University of Victoria, Victoria, BC V8W 3P6, Canada.

Publisher Item Identifier 10.1109/TCSI.2002.802362.

where $\Delta\boldsymbol{A}$, $\Delta\boldsymbol{b}$, $\Delta\boldsymbol{c}$, and Δd stand for the quantization errors of the coefficient matrices. The transfer function of the FWL realization is then expressed in the form

$$\tilde{H}(z) = \tilde{\boldsymbol{c}}(z\boldsymbol{I} - \tilde{\boldsymbol{A}})^{-1}\tilde{\boldsymbol{b}} + \tilde{d}. \tag{4}$$

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. Assuming that this realization has N parameters, the first-order approximation of Taylor's series expansion yields

$$\begin{aligned}\Delta H(z) =& \tilde{H}(z) - H(z)\\ =& \sum_{i=1}^{N} \frac{\partial H(z)}{\partial p_i}\Delta p_i.\end{aligned} \tag{5}$$

It is obvious that the smaller $\partial H(z)/\partial p_i$, $i = 1, 2, \ldots, N$ yield the smaller transfer-function error $\Delta H(z)$. For a fixed-point implementation of B bits, the parameter perturbations are considered 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

$$\begin{aligned}\sigma_{\Delta H}^2 =& \frac{1}{2\pi j}\oint_{|z|=1} E[|\Delta H(z)|^2]\frac{dz}{z}\\ =& \frac{1}{2\pi}\int_0^{2\pi} E[|\Delta H(e^{j\omega})|^2]d\omega\end{aligned} \tag{6}$$

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

$$E[|\Delta H(z)|^2] = \sum_{i=1}^{N}\left|\frac{\partial H(z)}{\partial p_i}\right|^2 \sigma^2 \tag{7}$$

where

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

By carrying out a coordinate transformation

$$\overline{\boldsymbol{x}}(k) = \boldsymbol{T}^{-1}\boldsymbol{x}(k) \tag{8}$$

to (1), a new realization $(\overline{\boldsymbol{A}}, \overline{\boldsymbol{b}}, \overline{\boldsymbol{c}}, d)_n$ characterized by

$$\overline{\boldsymbol{A}} = \boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}, \quad \overline{\boldsymbol{b}} = \boldsymbol{T}^{-1}\boldsymbol{b}, \quad \overline{\boldsymbol{c}} = \boldsymbol{c}\boldsymbol{T} \tag{9}$$

can be derived. From (2) and (9), it is clear that the transfer function $H(z)$ is invariant under the coordinate transformation of (8).

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

$$\boldsymbol{S}_{\boldsymbol{X}} = \frac{\partial f}{\partial \boldsymbol{X}}, \quad (\boldsymbol{S}_{\boldsymbol{X}})_{ij} = \frac{\partial f}{\partial x_{ij}} \tag{10}$$

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

Definition 2: Let $\boldsymbol{X}(z)$ be an $m \times n$ complex matrix-valued function of a complex variable z and let $x_{pq}(z)$ be the (p, q)th entry of $\boldsymbol{X}(z)$. The L_2-norm of $\boldsymbol{X}(z)$ is then defined as [19, p. 48]

$$\begin{aligned}\|\boldsymbol{X}(z)\|_2 =& \left[\frac{1}{2\pi}\int_0^{2\pi}\left(\sum_{p=1}^{m}\sum_{q=1}^{n}\left|x_{pq}(e^{j\omega})\right|^2\right)d\omega\right]^{1/2}\\ =& \left(\mathrm{tr}\left[\frac{1}{2\pi j}\oint_{|z|=1}\boldsymbol{X}(z)\boldsymbol{X}^*(z)\frac{dz}{z}\right]\right)^{1/2}.\end{aligned} \tag{11}$$

From the foregoing arguments, the overall L_2-sensitivity measure is defined by

$$\begin{aligned}S_d =& \frac{\sigma_{\Delta H}^2}{\sigma^2}\\ =& \left\|\frac{\partial H(z)}{\partial \boldsymbol{A}}\right\|_2^2 + \left\|\frac{\partial H(z)}{\partial \boldsymbol{b}}\right\|_2^2 + \left\|\frac{\partial H(z)}{\partial \boldsymbol{c}^T}\right\|_2^2 + \left\|\frac{\partial H(z)}{\partial d}\right\|_2^2\\ =& \left\|\boldsymbol{G}^T(z)\boldsymbol{F}^T(z)\right\|_2^2 + \left\|\boldsymbol{G}^T(z)\right\|_2^2 + \|\boldsymbol{F}(z)\|_2^2 + 1\end{aligned} \tag{12}$$

where

$$\boldsymbol{F}(z) = (z\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{b}, \quad \boldsymbol{G}(z) = \boldsymbol{c}(z\boldsymbol{I} - \boldsymbol{A})^{-1}.$$

The terminology "L_2-sensitivity" used here reflects the fact that the terms involved in (12) are all L_2-norms, and the definition differs itself from the mixed L_1/L_2 sensitivity defined by

$$\tilde{S}_d = \left\|\frac{\partial H(z)}{\partial \boldsymbol{A}}\right\|_1^2 + \left\|\frac{\partial H(z)}{\partial \boldsymbol{b}}\right\|_2^2 + \left\|\frac{\partial H(z)}{\partial \boldsymbol{c}^T}\right\|_2^2 + \left\|\frac{\partial H(z)}{\partial d}\right\|_2^2 \tag{13a}$$

whose upper bound, i.e.,

$$\hat{S}_d = \left\|\boldsymbol{G}^T(z)\right\|_2^2\|\boldsymbol{F}(z)\|_2^2 + \left\|\boldsymbol{G}^T(z)\right\|_2^2 + \|\boldsymbol{F}(z)\|_2^2 + 1 \tag{13b}$$

was the sensitivity measure employed by the authors of [1]–[5]. From a technical point of view, the main motivation of investigating the L_1/L_2 sensitivity was to overcome the difficulties introduced if the L_1-norm term in $\tilde{S}_d$ would have been replaced by a L_2-norm term. See [6]–[9] for some detailed accounts of this and other issues concerning the relationship between these two sensitivity measures.

Note that the L_2-sensitivity measure, (12), can also be expressed as

$$S_d = \mathrm{tr}[\boldsymbol{M}] + \mathrm{tr}[\boldsymbol{W}_0] + \mathrm{tr}[\boldsymbol{K}_0] + 1 \tag{14a}$$

where

$$\boldsymbol{M} = \frac{1}{2\pi j}\oint_{|z|=1}\boldsymbol{F}(z^{-1})\boldsymbol{G}(z^{-1})[\boldsymbol{F}(z)\boldsymbol{G}(z)]^T\frac{dz}{z} \tag{14b}$$

$$\boldsymbol{K}_0 = \frac{1}{2\pi j}\oint_{|z|=1}\boldsymbol{F}(z)\boldsymbol{F}^T(z^{-1})\frac{dz}{z} \tag{14c}$$

$$\boldsymbol{W}_0 = \frac{1}{2\pi j}\oint_{|z|=1}\boldsymbol{G}^T(z)\boldsymbol{G}(z^{-1})\frac{dz}{z}. \tag{14d}$$

The matrices $\boldsymbol{K}_0$ and $\boldsymbol{W}_0$ are called the controllability and observability Gramians, respectively, and can be obtained by solving the Lyapunov equations [20]

$$\boldsymbol{K}_0 = \boldsymbol{A}\boldsymbol{K}_0\boldsymbol{A}^T + \boldsymbol{b}\boldsymbol{b}^T \tag{15a}$$

$$\boldsymbol{W}_0 = \boldsymbol{A}^T\boldsymbol{W}_0\boldsymbol{A} + \boldsymbol{c}^T\boldsymbol{c}. \tag{15b}$$

Definition 3: The general controllability and observability Gramians for a linear discrete-time system (1) are defined by

$$\boldsymbol{K}_i = \boldsymbol{A}\boldsymbol{K}_i\boldsymbol{A}^T + \frac{1}{2}\left[\boldsymbol{A}^i\boldsymbol{b}\boldsymbol{b}^T + \boldsymbol{b}\boldsymbol{b}^T(\boldsymbol{A}^T)^i\right] \tag{16a}$$

$$\boldsymbol{W}_i = \boldsymbol{A}^T\boldsymbol{W}_i\boldsymbol{A} + \frac{1}{2}\left[\boldsymbol{c}^T\boldsymbol{c}\boldsymbol{A}^i + (\boldsymbol{A}^T)^i\boldsymbol{c}^T\boldsymbol{c}\right] \tag{16b}$$

respectively, where $i = 0, 1, 2, \ldots$.

Notice that, with $i = 0$, (16a) and (16b) become (15a) and (15b), respectively. In other words, the conventional controllabilty and observability Gramians, $\boldsymbol{K}_0$ and $\boldsymbol{W}_0$, can be viewed as the special cases of the general controllability and observability Gramians, $\boldsymbol{K}_i$ and $\boldsymbol{W}_i$.

The theorem below relates the evaluation of the matrix $\boldsymbol{M}$ to the general controllability and observability Gramians.

Theorem 1: The matrix $\boldsymbol{M}$ defined in (14b) is symmetric and positive-definite and can be expressed in the form

$$\boldsymbol{M} = \mathrm{tr}[\boldsymbol{W}_0]\boldsymbol{K}_0 + 2\sum_{i=1}^{\infty}\mathrm{tr}[\boldsymbol{W}_i]\boldsymbol{K}_i. \tag{17}$$

Proof: The symmetry of matrix $\boldsymbol{M}$ can be verified by evaluating $\boldsymbol{M}^T$ using (14b) as

$$\begin{aligned}\boldsymbol{M}^T =& \frac{1}{2\pi j}\oint_{|z|=1}\boldsymbol{F}(z)\boldsymbol{G}(z)[\boldsymbol{F}(z^{-1})\boldsymbol{G}(z^{-1})]^T\frac{dz}{z}\\ =& \frac{1}{2\pi}\int_0^{2\pi}\boldsymbol{F}(e^{j\omega})\boldsymbol{G}(e^{j\omega})[\boldsymbol{F}(e^{-j\omega})\boldsymbol{G}(e^{-j\omega})]^T d\omega\\ =& \frac{1}{2\pi}\int_0^{2\pi}\boldsymbol{F}(e^{-j\omega})\boldsymbol{G}(e^{-j\omega})[\boldsymbol{F}(e^{j\omega})\boldsymbol{G}(e^{j\omega})]^T d\omega\\ =& \frac{1}{2\pi j}\oint_{|z|=1}\boldsymbol{F}(z^{-1})\boldsymbol{G}(z^{-1})[\boldsymbol{F}(z)\boldsymbol{G}(z)]^T\frac{dz}{z}\\ =& \boldsymbol{M}.\end{aligned} \tag{18}$$

To show the positive definiteness of $\boldsymbol{M}$, let $\boldsymbol{x}$ be an arbitrary nonzero column vector of dimension n and use (14b) to compute

$$\begin{aligned}\boldsymbol{x}^T\boldsymbol{M}\boldsymbol{x} =& \frac{1}{2\pi j}\oint_{|z|=1}\boldsymbol{x}^T\boldsymbol{F}(z^{-1})\boldsymbol{G}(z^{-1})[\boldsymbol{F}(z)\boldsymbol{G}(z)]^T\boldsymbol{x}\frac{dz}{z}\\ =& \frac{1}{2\pi j}\oint_{|z|=1}\left|[\boldsymbol{F}(z)\boldsymbol{G}(z)]^T\boldsymbol{x}\right|^2\frac{dz}{z}\\ =& \frac{1}{2\pi}\int_0^{2\pi}\left|[\boldsymbol{F}(e^{j\omega})\boldsymbol{G}(e^{j\omega})]^T\boldsymbol{x}\right|^2 d\omega > 0.\end{aligned} \tag{19}$$

Next, noting that

$$\begin{aligned}(z\boldsymbol{I}-\boldsymbol{A})^{-1} =& z^{-1}\boldsymbol{I} + z^{-2}\boldsymbol{A} + z^{-3}\boldsymbol{A}^2 + \cdots\\ =& \sum_{k=0}^{\infty}z^{-(k+1)}\boldsymbol{A}^k\end{aligned} \tag{20}$$

and utilizing the *Cauchy integral theorem*

$$\frac{1}{2\pi j}\oint_C z^k\frac{dz}{z} = \begin{cases}1, & k=0\\ 0, & k\neq 0\end{cases} \tag{21}$$

where C is a counterclockwise contour that encircles the origin, it is possible to write the $\boldsymbol{M}$ matrix as

$$\begin{aligned}\boldsymbol{M} =& \frac{1}{2\pi j}\oint_{|z|=1}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{i=0}^{\infty}\sum_{p=0}^{\infty}\boldsymbol{A}^k\boldsymbol{b}\boldsymbol{c}\boldsymbol{A}^l(\boldsymbol{A}^T)^i\\ &\cdot\boldsymbol{c}^T\boldsymbol{b}^T(\boldsymbol{A}^T)^p z^{k+l-i-p}\frac{dz}{z}\\ =& \sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{i=0}^{\infty}\boldsymbol{A}^k\boldsymbol{b}\boldsymbol{c}\boldsymbol{A}^l(\boldsymbol{A}^T)^i\boldsymbol{c}^T\boldsymbol{b}^T(\boldsymbol{A}^T)^{k+l-i}\end{aligned} \tag{22}$$

where $p = k + l - i \geq 0$. Since $\boldsymbol{c}\boldsymbol{A}^l(\boldsymbol{A}^T)^i\boldsymbol{c}^T$ is a scalar, it follows that

$$\begin{aligned}\boldsymbol{M} =& \sum_{l=0}^{\infty}\sum_{i=0}^{\infty}\boldsymbol{c}\boldsymbol{A}^l(\boldsymbol{A}^T)^i\boldsymbol{c}^T\sum_{k=0}^{\infty}\boldsymbol{A}^k\boldsymbol{b}\boldsymbol{b}^T(\boldsymbol{A}^T)^{k+l-i}\\ =& \sum_{l=0}^{\infty}\boldsymbol{c}\boldsymbol{A}^l(\boldsymbol{A}^T)^l\boldsymbol{c}^T\sum_{k=0}^{\infty}\boldsymbol{A}^k\boldsymbol{b}\boldsymbol{b}^T(\boldsymbol{A}^T)^k\\ &+\sum_{l=1}^{\infty}\boldsymbol{c}\boldsymbol{A}^l(\boldsymbol{A}^T)^{l-1}\boldsymbol{c}^T\sum_{k=0}^{\infty}\boldsymbol{A}^k\boldsymbol{b}\boldsymbol{b}^T(\boldsymbol{A}^T)^{k+1}\\ &+\sum_{l=2}^{\infty}\boldsymbol{c}\boldsymbol{A}^l(\boldsymbol{A}^T)^{l-2}\boldsymbol{c}^T\sum_{k=0}^{\infty}\boldsymbol{A}^k\boldsymbol{b}\boldsymbol{b}^T(\boldsymbol{A}^T)^{k+2}\\ &+\cdots\\ &+\sum_{i=1}^{\infty}\boldsymbol{c}\boldsymbol{A}^{i-1}(\boldsymbol{A}^T)^i\boldsymbol{c}^T\sum_{k=1}^{\infty}\boldsymbol{A}^k\boldsymbol{b}\boldsymbol{b}^T(\boldsymbol{A}^T)^{k-1}\\ &+\sum_{i=2}^{\infty}\boldsymbol{c}\boldsymbol{A}^{i-2}(\boldsymbol{A}^T)^i\boldsymbol{c}^T\sum_{k=2}^{\infty}\boldsymbol{A}^k\boldsymbol{b}\boldsymbol{b}^T(\boldsymbol{A}^T)^{k-2}\\ &+\cdots.\end{aligned} \tag{23}$$

Hence

$$\begin{aligned}\boldsymbol{M} =& \sum_{l=0}^{\infty}\boldsymbol{c}\boldsymbol{A}^l(\boldsymbol{A}^T)^l\boldsymbol{c}^T\sum_{k=0}^{\infty}\boldsymbol{A}^k\boldsymbol{b}\boldsymbol{b}^T(\boldsymbol{A}^T)^k\\ &+\frac{1}{2}\sum_{l=0}^{\infty}\left[\boldsymbol{c}\boldsymbol{A}\boldsymbol{A}^l(\boldsymbol{A}^T)^l\boldsymbol{c}^T + \boldsymbol{c}\boldsymbol{A}^l(\boldsymbol{A}^T)^l\boldsymbol{A}^T\boldsymbol{c}^T\right]\\ &\cdot\sum_{k=0}^{\infty}\left[\boldsymbol{A}^k\boldsymbol{b}\boldsymbol{b}^T\boldsymbol{A}^T(\boldsymbol{A}^T)^k + \boldsymbol{A}^k\boldsymbol{A}\boldsymbol{b}\boldsymbol{b}^T(\boldsymbol{A}^T)^k\right]\\ &+\frac{1}{2}\sum_{l=0}^{\infty}\left[\boldsymbol{c}\boldsymbol{A}^2\boldsymbol{A}^l(\boldsymbol{A}^T)^l\boldsymbol{c}^T + \boldsymbol{c}\boldsymbol{A}^l(\boldsymbol{A}^T)^l(\boldsymbol{A}^T)^2\boldsymbol{c}^T\right]\\ &\cdot\sum_{k=0}^{\infty}\left[\boldsymbol{A}^k\boldsymbol{b}\boldsymbol{b}^T(\boldsymbol{A}^T)^2(\boldsymbol{A}^T)^k + \boldsymbol{A}^k\boldsymbol{A}^2\boldsymbol{b}\boldsymbol{b}^T(\boldsymbol{A}^T)^k\right]\\ &+\cdots\\ =& \mathrm{tr}\left[\sum_{l=0}^{\infty}(\boldsymbol{A}^T)^l\boldsymbol{c}^T\boldsymbol{c}\boldsymbol{A}^l\right]\sum_{k=0}^{\infty}\boldsymbol{A}^k\boldsymbol{b}\boldsymbol{b}^T(\boldsymbol{A}^T)^k\\ &+\frac{1}{2}\mathrm{tr}\left[\sum_{l=0}^{\infty}(\boldsymbol{A}^T)^l(\boldsymbol{c}^T\boldsymbol{c}\boldsymbol{A} + \boldsymbol{A}^T\boldsymbol{c}^T\boldsymbol{c})\boldsymbol{A}^l\right]\\ &\cdot\sum_{k=0}^{\infty}\boldsymbol{A}^k\left(\boldsymbol{b}\boldsymbol{b}^T\boldsymbol{A}^T + \boldsymbol{A}\boldsymbol{b}\boldsymbol{b}^T\right)(\boldsymbol{A}^T)^k\\ &+\frac{1}{2}\mathrm{tr}\left[\sum_{l=0}^{\infty}(\boldsymbol{A}^T)^l(\boldsymbol{c}^T\boldsymbol{c}\boldsymbol{A}^2 + (\boldsymbol{A}^T)^2\boldsymbol{c}^T\boldsymbol{c})\boldsymbol{A}^l\right]\\ &\cdot\sum_{k=0}^{\infty}\boldsymbol{A}^k\left(\boldsymbol{b}\boldsymbol{b}^T(\boldsymbol{A}^T)^2 + \boldsymbol{A}^2\boldsymbol{b}\boldsymbol{b}^T\right)(\boldsymbol{A}^T)^k\\ &+\cdots\end{aligned} \tag{24}$$

which coincides with (17). This completes the proof of Theorem 1.

Making use of (17), we can write (14a) as

$$S_d = (\mathrm{tr}[\boldsymbol{W}_0]+1)\,(\mathrm{tr}[\boldsymbol{K}_0]+1) + 2\sum_{i=1}^{\infty}\mathrm{tr}[\boldsymbol{W}_i]\mathrm{tr}[\boldsymbol{K}_i]. \quad (25)$$

This is a new formula for expressing the L_2-sensitivity of (12) in terms of the general controllability and observability Gramians. Once $\boldsymbol{K}_i$ and $\boldsymbol{W}_i$ are obtained, the computation quantity of each term in the right-hand side of (25) is constant independent of i. Also, the infinite sum can be approximated with the finite sum by truncation provided that the system is stable. It should be noted that the first term in the right-hand side of (25) is identical to the upper bound of the L_1/L_2-mixed sensitivity measure [1]–[5]. Therefore, the new formula for the L_2-sensitivity reveals its connection to the L_1/L_2-mixed sensitivity measure in a natural way.

Next, we consider the problem of synthesizing the optimal state-space model with minimum L_2-sensitivity.

From (16a), (16b), and (9), the general controllability and observability Gramians $\overline{\boldsymbol{K}}_i$ and $\overline{\boldsymbol{W}}_i$ in the new realization $(\overline{\boldsymbol{A}},\overline{\boldsymbol{b}},\overline{\boldsymbol{c}},d)_n$ are related to the original ones by

$$\overline{\boldsymbol{K}}_i = \boldsymbol{T}^{-1}\boldsymbol{K}_i\boldsymbol{T}^{-T}, \quad \overline{\boldsymbol{W}}_i = \boldsymbol{T}^{T}\boldsymbol{W}_i\boldsymbol{T} \quad (26)$$

where $i = 0,1,2,\ldots$. For the new realization $(\overline{\boldsymbol{A}},\overline{\boldsymbol{b}},\overline{\boldsymbol{c}},d)_n$, (25) is changed to

$$S_d(\boldsymbol{T}\boldsymbol{T}^T) = \left(\mathrm{tr}[\overline{\boldsymbol{W}}_0]+1\right)\left(\mathrm{tr}[\overline{\boldsymbol{K}}_0]+1\right) + 2\sum_{i=1}^{\infty}\mathrm{tr}[\overline{\boldsymbol{W}}_i]\mathrm{tr}[\overline{\boldsymbol{K}}_i] \quad (27)$$

or equivalently

$$S_d(\boldsymbol{P}) = (\mathrm{tr}[\boldsymbol{W}_0\boldsymbol{P}]+1)\left(\mathrm{tr}[\boldsymbol{K}_0\boldsymbol{P}^{-1}]+1\right) + 2\sum_{i=1}^{\infty}\mathrm{tr}[\boldsymbol{W}_i\boldsymbol{P}]\mathrm{tr}[\boldsymbol{K}_i\boldsymbol{P}^{-1}] \quad (28)$$

where $\boldsymbol{P} = \boldsymbol{T}\boldsymbol{T}^T$ is an $n\times n$ positive-definite symmetric matrix.

Note that in (27) and (28) the dependence of the L_2-sensitivity measure on the coordinate transformation matrix $\boldsymbol{T}$ (equivalently $\boldsymbol{P}$) has been explicitly indicated. In this regard, it can be readily verified that the L_2-sensitivity measure of the original realization as defined in (12) [and evaluated in (14a) and (25)] can be referred to as $S_d(\boldsymbol{I})$ where $\boldsymbol{I}$ is the identity matrix of dimension n.

Using the formula for evaluating matrix gradient [21, p. 275]

$$\frac{\partial[\mathrm{tr}(\boldsymbol{M}\boldsymbol{X})]}{\partial\boldsymbol{X}} = \boldsymbol{M}^T$$
$$\frac{\partial[\mathrm{tr}(\boldsymbol{M}\boldsymbol{X}^{-1})]}{\partial\boldsymbol{X}} = -(\boldsymbol{X}^{-1}\boldsymbol{M}\boldsymbol{X}^{-1})^T \quad (29)$$

it follows from (28) that

$$\frac{\partial S_d(\boldsymbol{P})}{\partial\boldsymbol{P}} = (\mathrm{tr}[\boldsymbol{K}_0\boldsymbol{P}^{-1}]+1)\boldsymbol{W}_0 - (\mathrm{tr}[\boldsymbol{W}_0\boldsymbol{P}]+1)\boldsymbol{P}^{-1}\boldsymbol{K}_0\boldsymbol{P}^{-1} + 2\sum_{i=1}^{\infty}\left(\mathrm{tr}[\boldsymbol{K}_i\boldsymbol{P}^{-1}]\boldsymbol{W}_i - \mathrm{tr}[\boldsymbol{W}_i\boldsymbol{P}]\boldsymbol{P}^{-1}\boldsymbol{K}_i\boldsymbol{P}^{-1}\right). \quad (30)$$

Letting $\partial S_d(\boldsymbol{P})/\partial\boldsymbol{P} = \boldsymbol{0}$ provides

$$\boldsymbol{P}\boldsymbol{G}(\boldsymbol{P})\boldsymbol{P} = \boldsymbol{H}(\boldsymbol{P}) \quad (31a)$$

where

$$\boldsymbol{H}(\boldsymbol{P}) = (\mathrm{tr}[\boldsymbol{W}_0\boldsymbol{P}]+1)\boldsymbol{K}_0 + 2\sum_{i=1}^{\infty}\mathrm{tr}[\boldsymbol{W}_i\boldsymbol{P}]\boldsymbol{K}_i \quad (31b)$$

$$\boldsymbol{G}(\boldsymbol{P}) = (\mathrm{tr}[\boldsymbol{K}_0\boldsymbol{P}^{-1}]+1)\boldsymbol{W}_0 + 2\sum_{i=1}^{\infty}\mathrm{tr}[\boldsymbol{K}_i\boldsymbol{P}^{-1}]\boldsymbol{W}_i. \quad (31c)$$

The matrix $\boldsymbol{G}(\boldsymbol{P})$ in (31c) can be expressed as

$$\boldsymbol{G}(\boldsymbol{P}) = \mathrm{tr}[\boldsymbol{K}_0\boldsymbol{P}^{-1}]\boldsymbol{W}_0 + 2\sum_{i=1}^{\infty}\mathrm{tr}[\boldsymbol{K}_i\boldsymbol{P}^{-1}]\boldsymbol{W}_i + \boldsymbol{W}_0 = \mathrm{tr}[\overline{\boldsymbol{K}}_0]\boldsymbol{W}_0 + 2\sum_{i=1}^{\infty}\mathrm{tr}[\overline{\boldsymbol{K}}_i]\boldsymbol{W}_i + \boldsymbol{W}_0. \quad (32)$$

Using an argument similar to that of the proof for Theorem 1, it can be shown that the sum of the first two terms in the above expression gives a symmetric and positive-definite matrix. Since the third term, matrix $\boldsymbol{W}_0$ [see (14d)], is also symmetric and positive-definite, matrix $\boldsymbol{G}(\boldsymbol{P})$ is positive-definite.

It is noted that the sensitivity measure $S_d(\boldsymbol{P})$ here is essentially equivalent to a function $R(\boldsymbol{P})$ defined in [8, p. 104] or $f(\boldsymbol{P})$ defined in [17, p. 811], which has been shown to be unimodal whose stationary point is unique, symmetric, and positive-definite and is the global minimum of the function. Consequently, in what follows, our attention will be focused on solving (31a).

Note that $\boldsymbol{P}\boldsymbol{W}\boldsymbol{P} = \boldsymbol{M}$ has the unique solution [5] as

$$\boldsymbol{P} = \boldsymbol{W}^{-(1/2)}[\boldsymbol{W}^{1/2}\boldsymbol{M}\boldsymbol{W}^{1/2}]^{1/2}\boldsymbol{W}^{-(1/2)} \quad (33)$$

in the case when $\boldsymbol{W} > 0$ and $\boldsymbol{M} \geq 0$ are symmetric. A natural iteration scheme suggested by (31a) is

$$\boldsymbol{P}_{i+1}\boldsymbol{G}(\boldsymbol{P}_i)\boldsymbol{P}_{i+1} = \boldsymbol{H}(\boldsymbol{P}_i). \quad (34)$$

By (33), the matrix $\boldsymbol{P}_{i+1}$ satisying (34) can be obtained as

$$\boldsymbol{P}_{i+1} = \boldsymbol{G}^{-(1/2)}(\boldsymbol{P}_i)\left[\boldsymbol{G}^{1/2}(\boldsymbol{P}_i)\,\boldsymbol{H}(\boldsymbol{P}_i)\,\boldsymbol{G}^{1/2}(\boldsymbol{P}_i)\right]^{1/2}\boldsymbol{G}^{-(1/2)}(\boldsymbol{P}_i). \quad (35)$$

This iteration process continues until

$$|S_d(\boldsymbol{P}_{i+1}) - S_d(\boldsymbol{P}_i)| < \varepsilon \quad (36)$$

where $\varepsilon > 0$ is a prescribed tolerance.

To obtain a reasonable initial estimate $\boldsymbol{P}_0$ analytically, we denote the principal term of $S_d(\boldsymbol{P})$ in (28) as

$$S_d^{(0)}(\boldsymbol{P}) = (\mathrm{tr}[\boldsymbol{W}_0\boldsymbol{P}]+1)\left(\mathrm{tr}[\boldsymbol{K}_0\boldsymbol{P}^{-1}]+1\right). \quad (37)$$

Note that $S_d^{(0)}(\boldsymbol{P})$ in (37) corresponds to the upper bound of L_1/L_2-mixed sensitivity measure [1]–[5]. The positive-definite symmetric matrix $\boldsymbol{P}$, say $\boldsymbol{P}_0$, minimizing (37) is then derived uniquely as

$$\boldsymbol{P}_0 = \boldsymbol{W}_0^{-(1/2)}\left[\boldsymbol{W}_0^{1/2}\boldsymbol{K}_0\boldsymbol{W}_0^{1/2}\right]^{1/2}\boldsymbol{W}_0^{-(1/2)} \quad (38)$$

which serves as the initial estimate in the iteration process (35).

Once the optimal positive-definite symmetric matrix $\boldsymbol{P}$ that minimizes (28) is obtained, the optimal coordinate transformation matrix $\boldsymbol{T}$ is constructed as

$$\boldsymbol{T} = \boldsymbol{P}^{1/2}\boldsymbol{U} \quad (39)$$

where $\boldsymbol{U}$ is any $n \times n$ orthogonal matrix. The optimal state-space model with minimum L_2-sensitivity can then be realized by substituting (39) into (9).

B. Linear Continuous-Time Systems

In a linear continuous-time system, the transfer function is given by

$$H(s) = \boldsymbol{c}(s\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{b} + d \tag{40}$$

instead of (2). Equation (14a) is then written in the form

$$S_c = \mathrm{tr}[\boldsymbol{M}] + \mathrm{tr}[\boldsymbol{W}_0] + \mathrm{tr}[\boldsymbol{K}_0] + 1 \tag{41}$$

with

$$\begin{aligned}
\boldsymbol{M} =& \frac{1}{2\pi j}\int_{-j\infty}^{j\infty} \boldsymbol{F}(-s)\boldsymbol{G}(-s)[\boldsymbol{F}(s)\boldsymbol{G}(s)]^T ds \\
\boldsymbol{K}_0 =& \int_0^\infty e^{\boldsymbol{A}t}\boldsymbol{b}\boldsymbol{b}^T e^{\boldsymbol{A}^T t} dt \\
\boldsymbol{W}_0 =& \int_0^\infty e^{\boldsymbol{A}^T t}\boldsymbol{c}^T\boldsymbol{c}e^{\boldsymbol{A}t} dt \\
\boldsymbol{F}(s) =& (s\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{b} \\
\boldsymbol{G}(s) =& \boldsymbol{c}(s\boldsymbol{I} - \boldsymbol{A})^{-1}
\end{aligned}$$

where the controllability and observability Gramians $\boldsymbol{K}_0$ and $\boldsymbol{W}_0$ can be obtained by solving the Lyapunov equations [20]

$$\boldsymbol{A}\boldsymbol{K}_0 + \boldsymbol{K}_0\boldsymbol{A}^T = -\boldsymbol{b}\boldsymbol{b}^T \tag{42a}$$

$$\boldsymbol{A}^T\boldsymbol{W}_0 + \boldsymbol{W}_0\boldsymbol{A} = -\boldsymbol{c}^T\boldsymbol{c} \tag{42b}$$

in place of (15a) and (15b).

Definition 4: The general controllability and observability Gramians for a linear continuous-time system are defined by

$$\boldsymbol{A}\boldsymbol{K}_i + \boldsymbol{K}_i\boldsymbol{A}^T = -\frac{1}{2}\left(\boldsymbol{b}\boldsymbol{b}^T e^{\boldsymbol{A}^T i\Delta} + e^{\boldsymbol{A}i\Delta}\boldsymbol{b}\boldsymbol{b}^T\right) \tag{43a}$$

$$\boldsymbol{A}^T\boldsymbol{W}_i + \boldsymbol{W}_i\boldsymbol{A} = -\frac{1}{2}\left(\boldsymbol{c}^T\boldsymbol{c}e^{\boldsymbol{A}i\Delta} + e^{\boldsymbol{A}^T i\Delta}\boldsymbol{c}^T\boldsymbol{c}\right) \tag{43b}$$

where $i = 0, 1, 2, \ldots$, and Δ stands for a sampling interval.

Theorem 2: The $\boldsymbol{M}$ matrix defined in (41) can be approximated by

$$\boldsymbol{M} \simeq \mathrm{tr}[\boldsymbol{W}_0]\boldsymbol{K}_0 + \sqrt{2}\Delta\sum_{i=1}^{\infty}\mathrm{tr}[\boldsymbol{W}_i]\boldsymbol{K}_i. \tag{44}$$

Proof: Applying the inverse Laplace transform yields

$$\mathcal{L}^{-1}[\boldsymbol{F}(s)\boldsymbol{G}(s)] = \int_0^\infty e^{\boldsymbol{A}(t-\tau)}\boldsymbol{b}\boldsymbol{c}e^{\boldsymbol{A}\tau} d\tau \tag{45a}$$

$$\mathcal{L}^{-1}[\boldsymbol{F}(s)\boldsymbol{G}(s)]^T = \int_0^\infty e^{\boldsymbol{A}^T\tau}\boldsymbol{c}^T\boldsymbol{b}^T e^{\boldsymbol{A}^T(t-\tau)} d\tau \tag{45b}$$

where [20, p. 67]

$$e^{\boldsymbol{A}t} = \begin{cases} \sum\limits_{k=0}^{\infty}\frac{1}{k!}\boldsymbol{A}^k t^k, & t \geq 0 \\ \boldsymbol{0}, & t < 0. \end{cases}$$

Using Parseval's theorem, we can write the matrix $\boldsymbol{M}$ in (41) as

$$\begin{aligned}
\boldsymbol{M} =& \int_0^\infty\int_0^\infty\int_0^\infty e^{\boldsymbol{A}(t-\lambda)}\boldsymbol{b}\boldsymbol{c}e^{\boldsymbol{A}\lambda} \\
& \cdot e^{\boldsymbol{A}^T\tau}\boldsymbol{c}^T\boldsymbol{b}^T e^{\boldsymbol{A}^T(t-\tau)} d\lambda d\tau dt \\
=& \int_0^\infty\int_0^\infty \boldsymbol{c}e^{\boldsymbol{A}\lambda}e^{\boldsymbol{A}^T\tau}\boldsymbol{c}^T \\
& \cdot \left[\int_0^\infty e^{\boldsymbol{A}(t-\lambda)}\boldsymbol{b}\boldsymbol{b}^T e^{\boldsymbol{A}^T(t-\tau)} dt\right] d\lambda d\tau
\end{aligned} \tag{46}$$

which can be approximated by

$$\begin{aligned}
\boldsymbol{M} \simeq& \int_0^\infty \boldsymbol{c}e^{\boldsymbol{A}\lambda}e^{\boldsymbol{A}^T\lambda}\boldsymbol{c}^T \\
& \cdot \left[\int_0^\infty e^{\boldsymbol{A}(t-\lambda)}\boldsymbol{b}\boldsymbol{b}^T e^{\boldsymbol{A}^T(t-\lambda)} dt\right] d\lambda \\
& + \frac{\Delta}{\sqrt{2}}\sum_{i=1}^{\infty}\int_0^\infty \boldsymbol{c}e^{\boldsymbol{A}\lambda}e^{\boldsymbol{A}^T(\lambda+i\Delta)}\boldsymbol{c}^T \\
& \cdot \left[\int_0^\infty e^{\boldsymbol{A}(t-\lambda)}\boldsymbol{b}\boldsymbol{b}^T e^{\boldsymbol{A}^T(t-\lambda-i\Delta)} dt\right] d\lambda \\
& + \frac{\Delta}{\sqrt{2}}\sum_{i=1}^{\infty}\int_0^\infty \boldsymbol{c}e^{\boldsymbol{A}(\tau+i\Delta)}e^{\boldsymbol{A}^T\tau}\boldsymbol{c}^T \\
& \cdot \left[\int_0^\infty e^{\boldsymbol{A}(t-\tau-i\Delta)}\boldsymbol{b}\boldsymbol{b}^T e^{\boldsymbol{A}^T(t-\tau)} dt\right] d\tau \\
=& \int_0^\infty \boldsymbol{c}e^{\boldsymbol{A}\lambda}e^{\boldsymbol{A}^T\lambda}\boldsymbol{c}^T \\
& \cdot \left[\int_0^\infty e^{\boldsymbol{A}(t-\lambda)}\boldsymbol{b}\boldsymbol{b}^T e^{\boldsymbol{A}^T(t-\lambda)} dt\right] d\lambda \\
& + \frac{\Delta}{\sqrt{2}}\sum_{i=1}^{\infty}\int_0^\infty \boldsymbol{c}e^{\boldsymbol{A}\lambda}e^{\boldsymbol{A}^T(\lambda+i\Delta)}\boldsymbol{c}^T \\
& \cdot \left[\int_0^\infty e^{\boldsymbol{A}(t+i\Delta)}\boldsymbol{b}\boldsymbol{b}^T e^{\boldsymbol{A}^T t} dt\right] d\lambda \\
& + \frac{\Delta}{\sqrt{2}}\sum_{i=1}^{\infty}\int_0^\infty \boldsymbol{c}e^{\boldsymbol{A}(\tau+i\Delta)}e^{\boldsymbol{A}^T\tau}\boldsymbol{c}^T \\
& \cdot \left[\int_0^\infty e^{\boldsymbol{A}t}\boldsymbol{b}\boldsymbol{b}^T e^{\boldsymbol{A}^T(t+i\Delta)} dt\right] d\tau
\end{aligned} \tag{47}$$

where Δ denotes a sampling interval. From (47) it follows that

$$\begin{aligned}
\boldsymbol{M} \simeq& \int_0^\infty \boldsymbol{c}e^{\boldsymbol{A}\lambda}e^{\boldsymbol{A}^T\lambda}\boldsymbol{c}^T d\lambda \int_0^\infty e^{\boldsymbol{A}t}\boldsymbol{b}\boldsymbol{b}^T e^{\boldsymbol{A}^T t} dt \\
& + \frac{\Delta}{\sqrt{2}}\sum_{i=1}^{\infty}\int_0^\infty \boldsymbol{c}e^{\boldsymbol{A}\lambda}e^{\boldsymbol{A}^T\lambda}e^{\boldsymbol{A}^T i\Delta}\boldsymbol{c}^T d\lambda \\
& \cdot \int_0^\infty e^{\boldsymbol{A}t}e^{\boldsymbol{A}i\Delta}\boldsymbol{b}\boldsymbol{b}^T e^{\boldsymbol{A}^T t} dt \\
& + \frac{\Delta}{\sqrt{2}}\sum_{i=1}^{\infty}\int_0^\infty \boldsymbol{c}e^{\boldsymbol{A}i\Delta}e^{\boldsymbol{A}\tau}e^{\boldsymbol{A}^T\tau}\boldsymbol{c}^T d\tau \\
& \cdot \int_0^\infty e^{\boldsymbol{A}t}\boldsymbol{b}\boldsymbol{b}^T e^{\boldsymbol{A}^T i\Delta}e^{\boldsymbol{A}^T t} dt \\
=& \mathrm{tr}\left[\int_0^\infty e^{\boldsymbol{A}^T\lambda}\boldsymbol{c}^T\boldsymbol{c}e^{\boldsymbol{A}\lambda} d\lambda\right]\int_0^\infty e^{\boldsymbol{A}t}\boldsymbol{b}\boldsymbol{b}^T e^{\boldsymbol{A}^T t} dt + \frac{\Delta}{2\sqrt{2}} \\
& \cdot \sum_{i=1}^{\infty}\mathrm{tr}\left[\int_0^\infty e^{\boldsymbol{A}^T\lambda}\left(e^{\boldsymbol{A}^T i\Delta}\boldsymbol{c}^T\boldsymbol{c} + \boldsymbol{c}^T\boldsymbol{c}e^{\boldsymbol{A}i\Delta}\right)e^{\boldsymbol{A}\lambda} d\lambda\right] \\
& \cdot \int_0^\infty e^{\boldsymbol{A}t}\left(e^{\boldsymbol{A}i\Delta}\boldsymbol{b}\boldsymbol{b}^T + \boldsymbol{b}\boldsymbol{b}^T e^{\boldsymbol{A}^T i\Delta}\right)e^{\boldsymbol{A}^T t} dt
\end{aligned} \tag{48}$$

which is identical to (44). This completes the proof of Theorem 2.

Making use of (44), one can write (41) as

$$\begin{aligned}
S_c \simeq& (\mathrm{tr}[\boldsymbol{W}_0] + 1)(\mathrm{tr}[\boldsymbol{K}_0] + 1) \\
& + \sqrt{2}\Delta\sum_{i=1}^{\infty}\mathrm{tr}[\boldsymbol{W}_i]\mathrm{tr}[\boldsymbol{K}_i].
\end{aligned} \tag{49}$$

This can be viewed as the continuous-time version of (25), but in the present case (49) is an approximate expression for evaluating the L_2-sensitivity of a linear continuous-time system in the state-space description. The infinite sum in (49) can be approximated with a finite sum by truncation provided that the system is stable. Note that the first term in the right-hand side of (49) is the upper bound of the L_1/L_2-mixed sensitivity measure [10], thus the L_2-sensitivity expression in (49) offers an analytical relation to the L_1/L_2-mixed sensitivity measure.

In the realization of the optimal linear continuous-time system with minimum L_2-sensitivity, (49) is changed to

$$S_c(\boldsymbol{TT}^T) \simeq (\mathrm{tr}[\boldsymbol{T}^T\boldsymbol{W}_0\boldsymbol{T}]+1)(\mathrm{tr}[\boldsymbol{T}^{-1}\boldsymbol{K}_0\boldsymbol{T}^{-T}]+1) + \sqrt{2}\Delta\sum_{i=1}^{\infty}\mathrm{tr}[\boldsymbol{T}^T\boldsymbol{W}_i\boldsymbol{T}]\mathrm{tr}[\boldsymbol{T}^{-1}\boldsymbol{K}_i\boldsymbol{T}^{-T}] \quad (50)$$

or equivalently

$$S_c(\boldsymbol{P}) \simeq (\mathrm{tr}[\boldsymbol{W}_0\boldsymbol{P}]+1)(\mathrm{tr}[\boldsymbol{K}_0\boldsymbol{P}^{-1}]+1) + \sqrt{2}\Delta\sum_{i=1}^{\infty}\mathrm{tr}[\boldsymbol{W}_i\boldsymbol{P}]\mathrm{tr}[\boldsymbol{K}_i\boldsymbol{P}^{-1}] \quad (51)$$

where $\boldsymbol{P}=\boldsymbol{TT}^T$. Similar to the discrete-time case, the L_2-sensitivity measure evaluated in (41) and (49) may be referred to as $S_c(\boldsymbol{I})$. It follows that the formulas in (31a), (34), and (35) also hold for the present case provided that $\boldsymbol{H}(\boldsymbol{P})$ and $\boldsymbol{G}(\boldsymbol{P})$ are modified to

$$\boldsymbol{H}(\boldsymbol{P}) = (\mathrm{tr}[\boldsymbol{W}_0\boldsymbol{P}]+1)\boldsymbol{K}_0 + \sqrt{2}\Delta\sum_{i=1}^{\infty}\mathrm{tr}[\boldsymbol{W}_i\boldsymbol{P}]\boldsymbol{K}_i \quad (52a)$$

$$\boldsymbol{G}(\boldsymbol{P}) = (\mathrm{tr}[\boldsymbol{K}_0\boldsymbol{P}^{-1}]+1)\boldsymbol{W}_0 + \sqrt{2}\Delta\sum_{i=1}^{\infty}\mathrm{tr}[\boldsymbol{K}_i\boldsymbol{P}^{-1}]\boldsymbol{W}_i. \quad (52b)$$

Consequently, the iteration process described by (34) and (35) and the initial estimate given by (38) can also apply to the linear continuous-time systems.

III. L_2-Sensitivity Minimization of 2-D Digital Filters

A 2-D digital filter can be represented by the following local state-space (LSS) model $(\boldsymbol{A},\boldsymbol{b},\boldsymbol{c},d)_{m,n}$ which is stable, separately locally controllable, and separately locally observable [22]:

$$\begin{aligned}\begin{bmatrix}\boldsymbol{x}^h(i+1,j)\\ \boldsymbol{x}^v(i,j+1)\end{bmatrix} &= \begin{bmatrix}\boldsymbol{A}_1 & \boldsymbol{A}_2\\ \boldsymbol{A}_3 & \boldsymbol{A}_4\end{bmatrix}\begin{bmatrix}\boldsymbol{x}^h(i,j)\\ \boldsymbol{x}^v(i,j)\end{bmatrix} + \begin{bmatrix}\boldsymbol{b}_1\\ \boldsymbol{b}_2\end{bmatrix}u(i,j)\\ &\triangleq \mathbf{A}\boldsymbol{x}(i,j)+\boldsymbol{b}u(i,j)\\ y(i,j) &= [\boldsymbol{c}_1 \quad \boldsymbol{c}_2]\begin{bmatrix}\boldsymbol{x}^h(i,j)\\ \boldsymbol{x}^v(i,j)\end{bmatrix} + du(i,j)\\ &\triangleq \mathbf{c}\boldsymbol{x}(i,j)+du(i,j)\end{aligned} \quad (53)$$

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 $\boldsymbol{A}_1$, $\boldsymbol{A}_2$, $\boldsymbol{A}_3$, $\boldsymbol{A}_4$, $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, $\boldsymbol{c}_1$, $\boldsymbol{c}_2$, and d are real constant matrices of appropriate dimensions. The transfer function of (53) is given by

$$H(z_1,z_2) = \boldsymbol{c}(z_1\boldsymbol{I}_m \oplus z_2\boldsymbol{I}_n - \boldsymbol{A})^{-1}\boldsymbol{b} + d. \quad (54)$$

Let the coordinate transformation be specified by

$$\overline{\boldsymbol{x}}(i,j) = \boldsymbol{T}^{-1}\boldsymbol{x}(i,j) \quad (55)$$

where $\boldsymbol{T}=\boldsymbol{T}_1\oplus\boldsymbol{T}_4$ with $\boldsymbol{T}_1$ and $\boldsymbol{T}_4$ nonsingular. Then the LSS model (53) is equivalent to a new realization $(\overline{\boldsymbol{A}},\overline{\boldsymbol{b}},\overline{\boldsymbol{c}},d)_{m,n}$ characterized by

$$\overline{\boldsymbol{A}} = \boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}, \quad \overline{\boldsymbol{b}} = \boldsymbol{T}^{-1}\boldsymbol{b}, \quad \overline{\boldsymbol{c}} = \boldsymbol{c}\boldsymbol{T} \quad (56)$$

in the sense that the transfer function $H(z_1,z_2)$ remains invariant under such a transformation.

Definition 5: 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

$$\|\boldsymbol{X}\|_p = \left[\frac{1}{(2\pi j)^2}\oint\oint_{\Gamma^2}\|\boldsymbol{X}(z_1,z_2)\|_F^p\,\frac{dz_1dz_2}{z_1z_2}\right]^{1/p} \quad (57)$$

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

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

and $\Gamma^2=\{(z_1,z_2):|z_1|=1,|z_2|=1\}$. An L_2-sensitivity is then evaluated by

$$\begin{aligned}S_{2d} &= \left\|\frac{\partial H(z_1,z_2)}{\partial \boldsymbol{A}}\right\|_2^2 + \left\|\frac{\partial H(z_1,z_2)}{\partial \boldsymbol{b}}\right\|_2^2 + \left\|\frac{\partial H(z_1,z_2)}{\partial \boldsymbol{c}^T}\right\|_2^2 + \left\|\frac{\partial H(z_1,z_2)}{\partial d}\right\|_2^2\\ &= \left\|\boldsymbol{G}^T(z_1,z_2)\boldsymbol{F}^T(z_1,z_2)\right\|_2^2 + \left\|\boldsymbol{G}^T(z_1,z_2)\right\|_2^2 + \|\boldsymbol{F}(z_1,z_2)\|_2^2 + 1\end{aligned} \quad (58)$$

where

$$\begin{aligned}\boldsymbol{F}(z_1,z_2) &= (z_1\boldsymbol{I}_m \oplus z_2\boldsymbol{I}_n - \boldsymbol{A})^{-1}\boldsymbol{b}\\ \boldsymbol{G}(z_1,z_2) &= \boldsymbol{c}(z_1\boldsymbol{I}_m \oplus z_2\boldsymbol{I}_n - \boldsymbol{A})^{-1}.\end{aligned}$$

By Definition 5, we can write (58) as

$$S_{2d} = \mathrm{tr}[\boldsymbol{M}] + \mathrm{tr}[\boldsymbol{W}_{00}] + \mathrm{tr}[\boldsymbol{K}_{00}] + 1 \quad (59)$$

where

$$\begin{aligned}\boldsymbol{M} &= \frac{1}{(2\pi j)^2}\oint\oint_{\Gamma^2}\boldsymbol{F}\left(z_1^{-1},z_2^{-1}\right)\boldsymbol{G}\left(z_1^{-1},z_2^{-1}\right) \cdot [\boldsymbol{F}(z_1,z_2)\boldsymbol{G}(z_1,z_2)]^T\frac{dz_1dz_2}{z_1z_2}\\ \boldsymbol{K}_{00} &= \frac{1}{(2\pi j)^2}\oint\oint_{\Gamma^2}\boldsymbol{F}(z_1,z_2)\boldsymbol{F}^T\left(z_1^{-1},z_2^{-1}\right)\frac{dz_1dz_2}{z_1z_2}\\ \boldsymbol{W}_{00} &= \frac{1}{(2\pi j)^2}\oint\oint_{\Gamma^2}\boldsymbol{G}^T(z_1,z_2)\boldsymbol{G}\left(z_1^{-1},z_2^{-1}\right)\frac{dz_1dz_2}{z_1z_2}.\end{aligned}$$

The matrices $\boldsymbol{K}_{00}$ and $\boldsymbol{W}_{00}$ here are called 2-D local controllability and observability Gramians, respectively, and can be derived from

$$\boldsymbol{F}(z_1,z_2) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\boldsymbol{f}(i,j)z_1^{-i}z_2^{-j} \quad (60a)$$

$$\boldsymbol{G}(z_1,z_2) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\boldsymbol{g}^T(i,j)z_1^{-i}z_2^{-j} \quad (60b)$$

where

$$\begin{aligned}\boldsymbol{f}(i,j) =& \boldsymbol{A}^{(i-1,j)}\begin{bmatrix}\boldsymbol{I}_m & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}\begin{bmatrix}\boldsymbol{b}_1\\ \boldsymbol{0}\end{bmatrix}\\ &+\boldsymbol{A}^{(i,j-1)}\begin{bmatrix}\boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{I}_n\end{bmatrix}\begin{bmatrix}\boldsymbol{0}\\ \boldsymbol{b}_2\end{bmatrix}\\ \boldsymbol{g}^T(i,j) =& \boldsymbol{c}\boldsymbol{A}^{(i-1,j)}\begin{bmatrix}\boldsymbol{I}_m & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}+\boldsymbol{c}\boldsymbol{A}^{(i,j-1)}\begin{bmatrix}\boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{I}_n\end{bmatrix}\end{aligned}$$

and the state transition matrix $\boldsymbol{A}^{(i,j)}$ is derived from

$$\begin{aligned}\boldsymbol{A}^{(1,0)} =& \begin{bmatrix}\boldsymbol{I}_m & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}\boldsymbol{A}\quad \boldsymbol{A}^{(0,1)}=\begin{bmatrix}\boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{I}_n\end{bmatrix}\boldsymbol{A}\\ \boldsymbol{A}^{(0,0)} =& \boldsymbol{I}_{m+n},\qquad \boldsymbol{A}^{(-i,j)}=\boldsymbol{0}\ (i\ge 1)\\ \boldsymbol{A}^{(i,-j)} =& \boldsymbol{0}\quad (j\ge 1)\\ \boldsymbol{A}^{(i,j)} =& \boldsymbol{A}^{(1,0)}\boldsymbol{A}^{(i-1,j)}+\boldsymbol{A}^{(0,1)}\boldsymbol{A}^{(i,j-1)}\\ =& \boldsymbol{A}^{(i-1,j)}\boldsymbol{A}^{(1,0)}+\boldsymbol{A}^{(i,j-1)}\boldsymbol{A}^{(0,1)},\\ & (i,j)>(0,0).\end{aligned}$$

By utilizing the following 2-D Cauchy integral theorem:

$$\frac{1}{(2\pi j)^2}\oint\oint_{C^2} z_1^m z_2^n\frac{dz_1dz_2}{z_1z_2}=\begin{cases}1, & (m,n)=(0,0)\\ 0, & (m,n)\ne(0,0)\end{cases}\tag{61}$$

where C is a counterclockwise contour that encircles the origin, it is possible to write the $\boldsymbol{K}_{00}$ and $\boldsymbol{W}_{00}$ in the form

$$\boldsymbol{K}_{00}=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\boldsymbol{f}(i,j)\boldsymbol{f}^T(i,j)\tag{62a}$$

$$\boldsymbol{W}_{00}=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\boldsymbol{g}(i,j)\boldsymbol{g}^T(i,j).\tag{62b}$$

Definition 6: The general 2-D local controllability and observability Gramians for a 2-D state-space digital filter, (53), are defined by

$$\begin{aligned}\boldsymbol{K}_{\xi\zeta} =& \frac{1}{2}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\Big[\boldsymbol{f}(i,j)\boldsymbol{f}^T(i+\xi,j+\zeta)\\ &+\boldsymbol{f}(i+\xi,j+\zeta)\boldsymbol{f}^T(i,j)\Big]\end{aligned}\tag{63a}$$

$$\begin{aligned}\boldsymbol{W}_{\xi\zeta} =& \frac{1}{2}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\big[\boldsymbol{g}(i,j)\boldsymbol{g}^T(i+\xi,j+\zeta)\\ &+\boldsymbol{g}(i+\xi,j+\zeta)\boldsymbol{g}^T(i,j)\big]\end{aligned}\tag{63b}$$

where $(0,0)\le(\xi,\zeta)<(\infty,\infty)$.

Theorem 3: The matrix $\boldsymbol{M}$ defined in (59) can be evaluated as

$$\boldsymbol{M}=\mathrm{tr}[\boldsymbol{W}_{00}]\boldsymbol{K}_{00}+2\sum_{(0,0)<(\xi,\zeta)<(\infty,\infty)}\sum\mathrm{tr}[\boldsymbol{W}_{\xi\zeta}]\boldsymbol{K}_{\xi\zeta}.\tag{64}$$

Proof: This can be proved by substituting (60a) and (60b) into the $\boldsymbol{M}$ matrix defined in (59) and utilizing the 2-D Cauchy integral theorem stated in (61).

Substituting (64) into (59) yields

$$\begin{aligned}S_{2d}=&(\mathrm{tr}[\boldsymbol{W}_{00}]+1)(\mathrm{tr}[\boldsymbol{K}_{00}]+1)\\ &+2\sum_{(0,0)<(\xi,\zeta)<(\infty,\infty)}\sum\mathrm{tr}[\boldsymbol{W}_{\xi\zeta}]\mathrm{tr}[\boldsymbol{K}_{\xi\zeta}].\end{aligned}\tag{65}$$

For the new realization $(\overline{\boldsymbol{A}},\overline{\boldsymbol{b}},\overline{\boldsymbol{c}},d)_{m,n}$ characterized by (56), it is easy to show that (65) is changed to

$$\begin{aligned}S_{2d}(\boldsymbol{P})=&(\mathrm{tr}[\boldsymbol{W}_{00}\boldsymbol{P}]+1)(\mathrm{tr}[\boldsymbol{K}_{00}\boldsymbol{P}^{-1}]+1)\\ &+2\sum_{(0,0)<(\xi,\zeta)<(\infty,\infty)}\sum\mathrm{tr}[\boldsymbol{W}_{\xi\zeta}\boldsymbol{P}]\mathrm{tr}[\boldsymbol{K}_{\xi\zeta}\boldsymbol{P}^{-1}]\end{aligned}\tag{66}$$

where

$$\boldsymbol{P}=\boldsymbol{T}\boldsymbol{T}^T,\quad \boldsymbol{P}=\begin{bmatrix}\boldsymbol{P}_1 & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{P}_4\end{bmatrix}.$$

According to the partition of $\boldsymbol{P}$, the Gramians $\boldsymbol{W}_{\xi\zeta}$ and $\boldsymbol{K}_{\xi\zeta}$ become

$$\boldsymbol{W}_{\xi\zeta}=\begin{bmatrix}\boldsymbol{W}_{\xi\zeta}^{(1)} & \boldsymbol{W}_{\xi\zeta}^{(2)}\\ \boldsymbol{W}_{\xi\zeta}^{(3)} & \boldsymbol{W}_{\xi\zeta}^{(4)}\end{bmatrix},\quad \boldsymbol{K}_{\xi\zeta}=\begin{bmatrix}\boldsymbol{K}_{\xi\zeta}^{(1)} & \boldsymbol{K}_{\xi\zeta}^{(2)}\\ \boldsymbol{K}_{\xi\zeta}^{(3)} & \boldsymbol{K}_{\xi\zeta}^{(4)}\end{bmatrix}.\tag{67}$$

By introducing

$$\hat{\boldsymbol{W}}_{\xi\zeta}=\begin{bmatrix}\boldsymbol{W}_{\xi\zeta}^{(1)} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{W}_{\xi\zeta}^{(4)}\end{bmatrix},\quad \hat{\boldsymbol{K}}_{\xi\zeta}=\begin{bmatrix}\boldsymbol{K}_{\xi\zeta}^{(1)} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{K}_{\xi\zeta}^{(4)}\end{bmatrix}\tag{68}$$

we can express $S_{2d}(\boldsymbol{P})$ in (66) as

$$\begin{aligned}S_{2d}(\boldsymbol{P})=&(\mathrm{tr}[\hat{\boldsymbol{W}}_{00}\boldsymbol{P}]+1)(\mathrm{tr}[\hat{\boldsymbol{K}}_{00}\boldsymbol{P}^{-1}]+1)\\ &+2\sum_{(0,0)<(\xi,\zeta)<(\infty,\infty)}\sum\mathrm{tr}[\hat{\boldsymbol{W}}_{\xi\zeta}\boldsymbol{P}]\mathrm{tr}[\hat{\boldsymbol{K}}_{\xi\zeta}\boldsymbol{P}^{1}].\end{aligned}\tag{69}$$

Hence, it suffices to deal with the matrix $\hat{\boldsymbol{K}}$ instead of $\boldsymbol{K}$. To make the subsequent derivation simple, in the following we omit the hat and write $\boldsymbol{K}$ for $\hat{\boldsymbol{K}}$.

By letting $\partial S_{2d}(\boldsymbol{P})/\partial\boldsymbol{P}_1=\boldsymbol{0}$ and $\partial S_{2d}(\boldsymbol{P})/\partial\boldsymbol{P}_4=\boldsymbol{0}$, we obtain

$$\boldsymbol{P}_1\boldsymbol{F}_1(\boldsymbol{P})\boldsymbol{P}_1=\boldsymbol{F}_2(\boldsymbol{P})\tag{70a}$$

$$\boldsymbol{P}_4\boldsymbol{F}_3(\boldsymbol{P})\boldsymbol{P}_4=\boldsymbol{F}_4(\boldsymbol{P})\tag{70b}$$

respectively, where

$$\begin{aligned}\boldsymbol{F}_1(\boldsymbol{P}) =& \mathrm{tr}[\boldsymbol{K}_{00}\boldsymbol{P}^{-1}]\boldsymbol{W}_{00}^{(1)}\\ &+2\sum_{(0,0)<(\xi,\zeta)<(\infty,\infty)}\sum\mathrm{tr}[\boldsymbol{K}_{\xi\zeta}\boldsymbol{P}^{-1}]\boldsymbol{W}_{\xi\zeta}^{(1)}\\ \boldsymbol{F}_2(\boldsymbol{P}) =& \mathrm{tr}[\boldsymbol{W}_{00}\boldsymbol{P}]\boldsymbol{K}_{00}^{(1)}\\ &+2\sum_{(0,0)<(\xi,\zeta)<(\infty,\infty)}\sum\mathrm{tr}[\boldsymbol{W}_{\xi\zeta}\boldsymbol{P}]\boldsymbol{K}_{\xi\zeta}^{(1)}\\ \boldsymbol{F}_3(\boldsymbol{P}) =& \mathrm{tr}[\boldsymbol{K}_{00}\boldsymbol{P}^{-1}]\boldsymbol{W}_{00}^{(4)}\\ &+2\sum_{(0,0)<(\xi,\zeta)<(\infty,\infty)}\sum\mathrm{tr}[\boldsymbol{K}_{\xi\zeta}\boldsymbol{P}^{-1}]\boldsymbol{W}_{\xi\zeta}^{(4)}\\ \boldsymbol{F}_4(\boldsymbol{P}) =& \mathrm{tr}[\boldsymbol{W}_{00}\boldsymbol{P}]\boldsymbol{K}_{00}^{(4)}\\ &+2\sum_{(0,0)<(\xi,\zeta)<(\infty,\infty)}\sum\mathrm{tr}[\boldsymbol{W}_{\xi\zeta}\boldsymbol{P}]\boldsymbol{K}_{\xi\zeta}^{(4)}.\end{aligned}$$

The iterative procedure reported in [18] can be used to solve (70a) and (70b) for $\boldsymbol{P}_1$ and $\boldsymbol{P}_4$, that is,

$$\begin{aligned}\boldsymbol{P}_1^{(i+1)} =& \boldsymbol{F}_1^{-(1/2)}(\boldsymbol{P}^{(i)})[\boldsymbol{F}_1^{1/2}(\boldsymbol{P}^{(i)})\boldsymbol{F}_2(\boldsymbol{P}^{(i)})\\ &\cdot\boldsymbol{F}_1^{1/2}(\boldsymbol{P}^{(i)})]^{1/2}\boldsymbol{F}_1^{-(1/2)}(\boldsymbol{P}^{(i)})\end{aligned}\tag{71a}$$

$$\begin{aligned}\boldsymbol{P}_4^{(i+1)} =& \boldsymbol{F}_3^{-(1/2)}(\boldsymbol{P}^{(i)})[\boldsymbol{F}_3^{1/2}(\boldsymbol{P}^{(i)})\boldsymbol{F}_4(\boldsymbol{P}^{(i)})\\ &\cdot\boldsymbol{F}_3^{1/2}(\boldsymbol{P}^{(i)})]^{1/2}\boldsymbol{F}_1^{-(1/2)}(\boldsymbol{P}^{(i)})\end{aligned}\tag{71b}$$

where $\boldsymbol{P}^{(i)}$ is derived from the previous iteration. The iteration continues until

$$|S_{2d}(\boldsymbol{P}^{(i+1)}) - S_{2d}(\boldsymbol{P}^{(i)})| < \varepsilon \tag{72}$$

where $\varepsilon > 0$ is a prescribed tolerance.

The initial estimate $\boldsymbol{P}^{(0)} = \boldsymbol{P}_1^{(0)} \oplus \boldsymbol{P}_4^{(0)}$ can be obtained by minimizing the upper bound of L_1/L_2-norms mixed sensitivity measure [11], [14]

$$S_{2d}^{(0)}(\boldsymbol{P}) = (\mathrm{tr}[\boldsymbol{W}_{00}\boldsymbol{P}] + 1)(\mathrm{tr}[\boldsymbol{K}_{00}\boldsymbol{P}^{-1}] + 1) \tag{73}$$

which yields a closed-form solution as

$$\boldsymbol{P}^{(0)} = \boldsymbol{W}_{00}^{-(1/2)} \left[\boldsymbol{W}_{00}^{1/2}\boldsymbol{K}_{00}\boldsymbol{W}_{00}^{1/2}\right]^{1/2} \boldsymbol{W}_{00}^{-(1/2)}. \tag{74}$$

Once the positive-definite L_2-optimal symmetric matrix $\boldsymbol{P} = \boldsymbol{P}_1 \oplus \boldsymbol{P}_4$ is obtained, the L_2-optimal coordinate transformation matrix can be constructed as

$$\boldsymbol{T} = \left[\boldsymbol{P}_1^{1/2} \oplus \boldsymbol{P}_4^{1/2}\right] [\boldsymbol{U}_1 \oplus \boldsymbol{U}_4] \tag{75}$$

where $\boldsymbol{U}_1$ and $\boldsymbol{U}_4$ are arbitrary $m \times m$ and $n \times n$ orthogonal matrics, respectively. The L_2-optimal filter structures that minimize (69) can be synthesized by substituting (75) into (56).

IV. Numerical Examples

Example 1: Consider a linear discrete-time system, (1), specified by

$$\begin{aligned}
\boldsymbol{A} &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0.453\,77 & -1.556\,16 & 1.974\,86 \end{bmatrix} \\
\boldsymbol{b} &= [0 \quad 0 \quad 1]^T \\
\boldsymbol{c} &= [0.023\,17 \quad 0.023\,02 \quad 0.079\,30] \\
d &= 0.015\,94
\end{aligned}$$

whose poles are at $z = 0.657\,881\,7$ and $z = 0.658\,489\,2 \pm j0.506\,098\,9$.

Using (15a) and (15b), its controllability and observability Gramians are calculated as

$$\boldsymbol{K}_0 = \begin{bmatrix} 17.061\,835\,37 & 14.886\,464\,08 & 9.602\,767\,53 \\ 14.886\,464\,08 & 17.061\,835\,37 & 14.886\,464\,08 \\ 9.602\,767\,53 & 14.886\,464\,08 & 17.061\,835\,37 \end{bmatrix}$$

$$\boldsymbol{W}_0 = \begin{bmatrix} 0.048\,103\,939 & -0.119\,291\,362 & 0.095\,427\,125 \\ -0.119\,291\,362 & 0.311\,061\,239 & -0.249\,967\,573 \\ 0.095\,427\,125 & -0.249\,967\,573 & 0.231\,012\,260 \end{bmatrix}.$$

By using (16a), (16b), (25), and (37), the L_2-sensitivity measure and the first term in it are computed as

$$\begin{aligned}
S_d &= 159.890\,910\,94 \\
S_d^{(0)}(\boldsymbol{I}) &= 82.984\,214\,45
\end{aligned}$$

where the infinite sum in (25) was truncated with $i = 100$ (the same truncation will be applied hereafter in this example).

The initial estimate $\boldsymbol{P}_0$ is derived from (38) as

$$\boldsymbol{P}_0 = \begin{bmatrix} 68.335\,599\,95 & 39.198\,457\,69 & 13.722\,889\,21 \\ 39.198\,457\,69 & 33.217\,505\,17 & 21.483\,618\,75 \\ 13.722\,889\,21 & 21.483\,618\,75 & 22.737\,263\,42 \end{bmatrix}$$

TABLE I
L_2-Sensitivity Profile of the First Ten Iterations

i	L_2-Sensitivity	i	L_2-Sensitivity
	159.8909109417472	5	8.832683342813343
0	8.943594607334033	6	8.832683342812404
1	8.833335506929652	7	8.832683342812398
2	8.832687360817378	8	8.832683342812402
3	8.832683367564366	9	8.832683342812402
4	8.832683342964870	10	8.832683342812393

or equivalently

$$\boldsymbol{T}_0 = \begin{bmatrix} 1.538\,415\,708 & 4.703\,727\,215 & 6.621\,467\,159 \\ 3.039\,406\,990 & 4.443\,971\,631 & 2.056\,848\,672 \\ 4.246\,846\,977 & 2.126\,308\,653 & -0.424\,694\,816 \end{bmatrix}.$$

Using (28) and (37), we evaluate

$$\begin{aligned}
S_d(\boldsymbol{P}_0) &= 8.943\,594\,61 \\
S_d^{(0)}(\boldsymbol{P}_0) &= 5.755\,476\,18.
\end{aligned}$$

Applying (35) to minimize (28), after ten iterations we obtain

$$\boldsymbol{P}_{10} = \begin{bmatrix} 81.469\,138\,71 & 48.384\,231\,45 & 17.980\,091\,71 \\ 48.384\,231\,45 & 39.032\,444\,13 & 24.143\,338\,09 \\ 17.980\,091\,71 & 24.143\,338\,09 & 23.951\,552\,07 \end{bmatrix}$$

or equivalently

$$\boldsymbol{T}_{10} = \begin{bmatrix} 1.283\,114\,259 & 5.356\,474\,502 & 7.150\,590\,005 \\ 3.056\,235\,553 & 4.759\,869\,047 & 2.652\,454\,530 \\ 4.338\,392\,561 & 2.264\,582\,260 & 0.039\,613\,686 \end{bmatrix}.$$

Using (28) and (37), we evaluate

$$\begin{aligned}
S_d(\boldsymbol{P}_{10}) &= 8.832\,683\,34 \\
S_d^{(0)}(\boldsymbol{P}_{10}) &= 5.814\,855\,21.
\end{aligned}$$

The L_2-sensitivity profile of the first ten iterations is given in Table I, from which we see that with a tolerance $\varepsilon = 10^{-7}$ the algorithm converges after four iterations. By substituting $\boldsymbol{T} = \boldsymbol{T}_{10}$ into (9), the optimal state-space model with minimum L_2-sensitivity is then obtained as

$$\begin{aligned}
\overline{\boldsymbol{A}} &= \begin{bmatrix} 0.796\,804\,001 & -0.414\,267\,323 & -0.025\,895\,398 \\ 0.414\,267\,323 & 0.565\,420\,531 & -0.316\,443\,915 \\ -0.025\,895\,398 & 0.316\,443\,915 & 0.612\,635\,467 \end{bmatrix} \\
\overline{\boldsymbol{b}} &= \begin{bmatrix} 0.444\,118\,830 \\ -0.413\,263\,073 \\ 0.229\,880\,039 \end{bmatrix} \\
\overline{\boldsymbol{c}} &= [0.444\,118\,830 \quad 0.413\,263\,073 \quad 0.229\,880\,039] \\
d &= 0.015\,940\,000.
\end{aligned}$$

For comparison purposes, the gradient-flow-based algorithm [8] in which $\boldsymbol{P}_{i+1}$ is generated by

$$\boldsymbol{P}_{i+1} = \boldsymbol{P}_i - \mu \frac{\partial S_d(\boldsymbol{P}_i)}{\partial \boldsymbol{P}_i} \tag{76}$$

was applied to the above example with the same initial estimate. In the simulation, the step-size μ in (76) is set to 0.28 (which seems the best), and the algorithm after 1000 iterations yields

$$S_d(\boldsymbol{P}_{1000}) = 8.832\,746\,57.$$

This is slightly larger than

$$S_d(\boldsymbol{P}_{10}) = 8.832\,683\,34$$

obtained by our method.

Example 2: Now we consider a linear continuous-time system, (40), characterized by

$$\boldsymbol{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -0.162\,484\,210 & -0.526\,273\,961 & -0.628\,831\,301 \end{bmatrix}$$
$$\boldsymbol{b} = [0 \quad 0 \quad 1]^T$$
$$\boldsymbol{c} = [0.250\,120\,406 \quad 0 \quad 0.191\,751\,308]$$
$$d = -0.539\,352\,1$$

whose poles are given by $s = -0.376\,731\,3$ and $s = -0.126\,050 \pm j0.644\,524\,2$.

The controllabilty and observability Gramians $\boldsymbol{K}_0$ and $\boldsymbol{W}_0$ are computed from (42a) and (42b) as

$$\boldsymbol{W}_0 = \begin{bmatrix} 0.148\,531\,757 & 0.152\,935\,083 & 0.192\,511\,683 \\ 0.152\,935\,083 & 0.248\,817\,754 & 0.290\,599\,751 \\ 0.192\,511\,683 & 0.290\,599\,752 & 0.491\,362\,363 \end{bmatrix}$$
$$\boldsymbol{K}_0 = \begin{bmatrix} 11.487\,179\,114 & 0.000\,000\,000 & -2.968\,181\,165 \\ 0.000\,000\,000 & 2.968\,181\,165 & 0.000\,000\,000 \\ -2.968\,181\,165 & 0.000\,000\,000 & 1.562\,076\,459 \end{bmatrix}.$$

By letting $\Delta = 0.01$ and using (43a) and (43b), the formula (49), where the infinite sum was truncated to its first 8000 terms, gives

$$S_c \simeq 67.383\,280$$
$$S_c^{(0)}(\boldsymbol{I}) = 32.141\,035$$

where

$$S_c^{(0)}(\boldsymbol{P}) = (\mathrm{tr}[\boldsymbol{W}_0\boldsymbol{P}] + 1)(\mathrm{tr}[\boldsymbol{K}_0\boldsymbol{P}^{-1}] + 1).$$

The same truncation will be applied hereafter in Example 2. The initial estimate $\boldsymbol{P}_0$ was obtained using (38) as

$$\boldsymbol{P}_0 = \begin{bmatrix} 13.645\,073\,461 & -2.672\,229\,834 & -3.167\,232\,201 \\ -2.672\,229\,834 & 6.099\,162\,847 & -1.239\,205\,730 \\ -3.167\,232\,201 & -1.239\,205\,730 & 2.880\,278\,340 \end{bmatrix}$$

or equivalently

$$\boldsymbol{T}_0 = \begin{bmatrix} 2.058\,386\,621 & 2.912\,443\,826 & 0.962\,179\,266 \\ 0.993\,607\,653 & -0.940\,504\,298 & -2.056\,054\,071 \\ -0.223\,906\,291 & -1.287\,114\,675 & 1.083\,272\,877 \end{bmatrix}.$$

In this case, (51) yields

$$S_c(\boldsymbol{P}_0) \simeq 20.704\,544$$
$$S_c^{(0)}(\boldsymbol{P}_0) = 10.256\,183.$$

In order to minimize $S_c(\boldsymbol{P})$ in (51), we apply (35) where $\boldsymbol{H}(\boldsymbol{P})$ and $\boldsymbol{G}(\boldsymbol{P})$ are given by (52a) and (52b). Ten iterations yield

$$\boldsymbol{P}_{10} = \begin{bmatrix} 14.932\,872\,358 & -3.417\,647\,079 & -3.136\,150\,798 \\ -3.417\,647\,079 & 6.452\,249\,132 & -1.114\,438\,061 \\ -3.136\,150\,798 & -1.114\,438\,061 & 2.621\,026\,248 \end{bmatrix}$$

or equivalently

$$\boldsymbol{T}_{10} = \begin{bmatrix} 2.001\,302\,974 & 3.020\,964\,828 & 1.342\,173\,712 \\ 1.185\,944\,380 & -1.030\,092\,946 & -1.996\,169\,728 \\ -0.360\,253\,310 & -1.235\,733\,847 & 0.981\,939\,744 \end{bmatrix}.$$

Now (51) gives

$$S_c(\boldsymbol{P}_{10}) \simeq 20.576\,702$$
$$S_c^{(0)}(\boldsymbol{P}_{10}) = 10.308\,948.$$

The L_2-sensitivity profile of the first ten iterations is given in Table II. It is observed that, with a tolerance $\varepsilon = 10^{-8}$, the algorithm converges after four iterations.

TABLE II
L_2 SENSITIVITY PROFILE OF THE FIRST TEN ITERATIONS

i	L_2-Sensitivity	i	L_2-Sensitivity
	67.38327994750175	5	20.57670199135822
0	20.70454380351688	6	20.57670199135815
1	20.57708100856075	7	20.57670199135823
2	20.57670318054942	8	20.57670199135817
3	20.57670199508960	9	20.57670199135819
4	20.57670199136992	10	20.57670199135813

By substituting $\boldsymbol{T} = \boldsymbol{T}_{10}$ into (9), the optimal state-space model with minimum L_2-sensitivity is then obtained as

$$\overline{\boldsymbol{A}} = \begin{bmatrix} -0.083\,598\,330 & -0.505\,801\,407 & -0.130\,205\,421 \\ 0.505\,801\,407 & -0.191\,285\,835 & -0.417\,260\,962 \\ -0.130\,205\,421 & 0.417\,260\,962 & -0.353\,947\,136 \end{bmatrix}$$
$$\overline{\boldsymbol{b}} = [0.431\,487\,669 \quad -0.518\,651\,368 \quad 0.523\,993\,264]^T$$
$$\overline{\boldsymbol{c}} = [0.431\,487\,669 \quad 0.518\,651\,368 \quad 0.523\,993\,264]$$
$$d = -0.539\,352\,1.$$

Example 3: Consider a 2-D stable state-space digital filter of order (2,2) modeled by (53) where

$$\boldsymbol{A} = \begin{bmatrix} 1.888\,99 & -0.912\,19 & -1.0 & 0.0 \\ 1.0 & 0.0 & 0.0 & 0.0 \\ 0.027\,71 & -0.025\,80 & 1.888\,99 & 1.0 \\ -0.025\,80 & 0.024\,31 & -0.912\,19 & 0.0 \end{bmatrix}$$
$$\boldsymbol{b} = [0.219\,089 \quad 0.0 \quad -0.028\,889 \quad 0.091\,219]^T$$
$$\boldsymbol{c} = [0.288\,89 \quad -0.091\,219 \quad -0.219\,089 \quad 0.0]$$
$$d = 0.870\,21.$$

First, (60a), (60b), (62a), and (62b) were used with the truncation $(0,0) \le (i,j) \le (100,100)$ to evaluate $\boldsymbol{K}_{00}$ and $\boldsymbol{W}_{00}$ as shown at the bottom of the next page. Next, $\boldsymbol{K}_{\xi\zeta}$ and $\boldsymbol{W}_{\xi\zeta}$ were calculated using (63a) and (63b) with truncations $(0,0) \le (\xi,\zeta) \le (100,100)$ and $(0,0) \le (i,j) \le (100,100)$. The L_2-sensitivity measure was then computed using (65) as

$$S_{2d} = 1.049\,619\,37 \times 10^7$$
$$S_{2d}^{(0)}(\boldsymbol{I}) = 2.304\,669\,43 \times 10^5$$

where $S_{2d}^{(0)}(\boldsymbol{I})$ is defined in (73).

The initial estimate $\boldsymbol{P}^{(0)} = \boldsymbol{P}_1^{(0)} \oplus \boldsymbol{P}_4^{(0)}$ was computed using (74) as

$$\boldsymbol{P}_1^{(0)} = \begin{bmatrix} 18.230\,312\,72 & 18.793\,858\,26 \\ 18.793\,858\,26 & 20.020\,376\,79 \end{bmatrix}$$
$$\boldsymbol{P}_4^{(0)} = \begin{bmatrix} 0.186\,683\,52 & -0.175\,208\,54 \\ -0.175\,208\,54 & 0.171\,174\,80 \end{bmatrix}.$$

Then, using (69) and (73), we obtain

$$S_{2d}(\boldsymbol{P}^{(0)}) = 7.900\,097\,87 \times 10^3$$
$$S_{2d}^{(0)}(\boldsymbol{P}^{(0)}) = 501.395\,756\,76.$$

Finally, by applying (71a) and (71b), with two iterations we obtain $\boldsymbol{P}^{(2)} = \boldsymbol{P}_1^{(2)} \oplus \boldsymbol{P}_4^{(2)}$ where

$$\boldsymbol{P}_1^{(2)} = \begin{bmatrix} 39.187\,326\,93 & 40.669\,774\,97 \\ 40.669\,774\,97 & 43.444\,552\,78 \end{bmatrix}$$
$$\boldsymbol{P}_4^{(2)} = \begin{bmatrix} 0.088\,082\,60 & -0.082\,082\,80 \\ -0.082\,082\,80 & 0.079\,469\,89 \end{bmatrix}$$

or equivalently

$$\boldsymbol{T}_1^{(2)} = \begin{bmatrix} 4.696\,426\,05 & 4.138\,950\,27 \\ 4.138\,950\,27 & 5.129\,682\,59 \end{bmatrix}$$
$$\boldsymbol{T}_4^{(2)} = \begin{bmatrix} 0.233\,207\,00 & -0.183\,567\,69 \\ -0.183\,567\,69 & 0.213\,945\,78 \end{bmatrix}.$$

The L_2-sensitivity profile of the first ten iterations is given in Table III. Then, (69) and (73) were calculated as

$$S_{2d}(\boldsymbol{P}^{(2)}) = 4.779\,951\,63 \times 10^3$$
$$S_{2d}^{(0)}(\boldsymbol{P}^{(2)}) = 807.469\,530\,86.$$

TABLE III
L_2 SENSITIVITY PROFILE OF THE FIRST TEN ITERATIONS

i	L_2-sensitivity	i	L_2-sensitivity
	1.04961937×10^7	5	4.77998761×10^3
0	7.90009787×10^3	6	4.77998760×10^3
1	4.78462354×10^3	7	4.77998760×10^3
2	4.77995163×10^3	8	4.77998760×10^3
3	4.77998982×10^3	9	4.77998760×10^3
4	4.77998749×10^3	10	4.77998760×10^3

By substituting $\boldsymbol{T}^{(2)} = \boldsymbol{T}_1^{(2)} \oplus \boldsymbol{T}_4^{(2)}$ into (56), the optimal filter structure with minimum L_2-sensitivity was obtained as shown at the bottom of the page.

It is noted that the L_2-sensitivity measure $S_{2d}(\boldsymbol{P}^{(2)}) = 4.779\,951\,63 \times 10^3$ in the optimal filter structure obtained here is considerably smaller than $S_{2d} = 6.222\,578\,56 \times 10^3$ in the optimal filter structure obtained by Li [17].

V. CONCLUSION

Novel expressions for the L_2-sensitivity of linear discrete-time systems, linear continuous-time systems, and 2-D state-space digital filters have been developed. Each expression is derived in terms of the general controllability and observability Gramians. It has been shown that, for each class of systems, the new formula offers a natural connection between the L_2-sensitivity and the L_1/L_2-mixed sensitivity. An iterative procedure has also been proposed to facilitate the minimization of the new expressions of the L_2-sensitivity measure. Our computer simulation results have demonstrated the effectiveness of the proposed techniques compared with several existing methods.

$$\boldsymbol{K}_{00} = \begin{bmatrix} 87.096\,159\,05 & 85.181\,964\,73 & 1.636\,490\,20 & -1.536\,011\,92 \\ 85.181\,964\,73 & 87.092\,707\,06 & 1.317\,925\,48 & -1.230\,067\,10 \\ 1.636\,490\,20 & 1.317\,925\,48 & 1.133\,755\,98 & -1.034\,846\,50 \\ -1.536\,011\,92 & -1.230\,067\,10 & -1.034\,846\,50 & 0.972\,803\,28 \end{bmatrix}$$
$$\boldsymbol{W}_{00} = \begin{bmatrix} 11.337\,679\,13 & -10.352\,076\,67 & 15.856\,787\,62 & 12.350\,254\,75 \\ -10.352\,076\,67 & 9.661\,973\,14 & -15.007\,594\,46 & -11.566\,136\,34 \\ 15.856\,787\,62 & -15.007\,594\,46 & 638.959\,288\,02 & 622.907\,317\,71 \\ 12.350\,254\,75 & -11.566\,136\,34 & 622.907\,317\,71 & 638.944\,553\,68 \end{bmatrix}$$

$$\overline{\boldsymbol{A}} = \begin{bmatrix} 0.962\,973\,55 & -0.147\,675\,17 & -0.171\,872\,45 & 0.135\,288\,51 \\ 0.138\,551\,74 & 0.926\,016\,45 & 0.138\,677\,49 & -0.109\,159\,27 \\ 0.075\,570\,78 & -0.030\,143\,12 & 0.983\,227\,00 & 0.143\,465\,12 \\ -0.031\,211\,45 & 0.057\,885\,53 & -0.150\,694\,21 & 0.905\,763\,00 \end{bmatrix}$$
$$\overline{\boldsymbol{b}} = [0.161\,467\,55 \quad -0.130\,282\,17 \quad 0.652\,248\,55 \quad 0.986\,001\,03]^T$$
$$\overline{\boldsymbol{c}} = [0.979\,199\,62 \quad 0.727\,776\,83 \quad -0.051\,093\,09 \quad 0.040\,217\,66]$$

REFERENCES

[1] L. Thiele, "Design of sensitivity and round-off noise optimal state-space discrete systems," *Int. J. Circuit Theory Appl.*, 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. Acoustics, Speech, and Signal Processing*, vol. 4, pp. 429–432.

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

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

[10] W. J. Lutz and S. L. Hakimi, "Design of multi-input multi-output systems with minimum sensitivity," *IEEE Trans. Circuits Syst.*, vol. 35, pp. 1114–1122, Sept. 1988.

[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 and Systems*, 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. 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 Inst.*, 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] ——, "Minimization of frequency-weighting sensitivity in 2-D systems based on the Fornasini–Marchesini second model," in *1992 IEEE Int. Conf. Acoustics, Speech, and 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] ——, "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] K. Zhou and J. C. Doyle, *Essentials of Robust Control*. Upper Saddle River, NJ: Prentice-Hall, 1998.

[20] T. Kailath, *Linear System*. Englewood Cliffs, N.J.: Prentice-Hall, 1980.

[21] L. L. Scharf, *Statistical Signal Processing*. Reading, MA: Addison-Wesley, 1991.

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

Takao Hinamoto (M'77–SM'84–F'01) received the B.E. degree from Okayama University in 1969, the M.E. degree from Kobe University in 1971, and the Dr. Eng. degree from Osaka University in 1977, all in electrical engineering.

From 1972 to 1988, he was with the Faculty of Engineering, Kobe University. From 1979 to 1981, he was on leave from Kobe University as a visiting member of staff in the Department of Electrical Engineering, Queen's University, Kingston, ON, Canada. During 1988-1991, he was a Professor of electronic circuits in the Faculty of Engineering, Tottori University, Japan. Since January 1992, he has been a Professor of electronic control in the Department of Electrical Engineering, Hiroshima University, Higashi-Hiroshima, Japan. His research interests include digital signal processing, system theory, and control engineering. He has published more than 250 papers in these areas and is the co-editor and coauthor of *Two-Dimensional Signal and Image Processing* (Tokyo, Japan: SICE, 1996).

Dr. Hinamoto served as an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II from 1993 to 1995 and presently serves as an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS I. He also served as Chair of the 12th Digital Signal Processing (DSP) Symposium held in Hiroshima in November 1997, sponsored by the DSP Technical Committee of IEICE. He was the Guest Editor of the special section of DSP in the August 1998 issue of the *IEICE Transactions on Fundamentals*. Since 1995, he has been a member of the steering committee of the IEEE Midwest Symposium on Circuits and Systems and, since 1998, a member of the Digital Signal Processing Technical Committee in the IEEE Circuits and Systems Society. He senied as a member of the Technical Program Committee for ISCAS'99. From 1993 to 2000, he served as a senator or member of the Board of Directors in the Society of Instrument and Control Engineers (SICE), and from 1999 to 2001, he was Chair of the Chugoku Chapter of SICE. He played a leading role in establishing the Hiroshima Section of IEEE, and served as the Interim Chair of the section. He is a recipient of the IEEE Third Millennium Medal.

Shuichi Yokoyama received the B.E. and M.E. degrees in electrical engineering from Hiroshima University, Higashi-Hiroshima, Japan, in 1996 and 1998, respectively.

He was engaged in research on digital signal processing during his graduate studies. Since 1998, he has been with NTT DoCoMo Chugoku Inc., Hiroshima, Japan.

Takuya Inoue received the B.E. and M.E. degrees in electrical engineering from Hiroshima University, Higashi-Hiroshjma, Japan, in 1999 and 2001, respectively.

He was engaged in research on digital signal processing during his graduate studies. Since 2001, he has been with Fujitsu System Solutions Limited, Tokyo, Japan.

Wu Zeng received the B.E. degree in automatic control engineering from Beijing Institute of Mechanical Industry, Beijing, China, and the M.E. degree in electrical engineering from Hiroshima University, Higashi-Hiroshjma, Japan, in 1996 and 2000, respectively. He was engaged in research on digital signal processing during his graduate studies. Since 2000, he has been with Precision Digital Corporation, Osaka, Japan.

Wu-Sheng Lu (S'81–M'85–SM'90–F'99) received his undergraduate education in mathematics from Fudan University, Shanghai, China, from 1959 to 1964, and the M.S. degree in electrical engineering and the Ph.D. degree in control science from the University of Minnesota, Minneapolis, in 1983, and 1984, respectively.

He was a Post-Doctoral Fellow at University of Victoria, Victoria, BC, Canada, in 1985 and a Visiting Assistant Professor at the University of Minnesota in 1986. Since 1987, he has been with University of Victoria where he is currently a Professor. His teaching and research interests are in the areas of digital signal processing and application of optimization methods. He is the coauthor, with A. Antoniou, of *Two-Dimensional Digital Filters* (New York: Marcel Dekker, 1992).

Dr. Lu was an Associate Editor of the *Canadian Journal of Electrical and Computer Engineering* in 1989, and the Editor of the same journal from 1990 to 1992. He served as an Associate Editor for IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II from 1993 to 1995 and for IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS I from 1999 to 2001. He is presently an Associate Editor for the *International Journal of Multidimensional Systems and Signal Processing*. He is a Fellow of the Engineering Institute of Canada.