

COMPARISON OF TECHNIQUES FOR L_2 -SENSITIVITY MINIMIZATION UNDER L_2 -SCALING CONSTRAINTS IN STATE-SPACE DIGITAL FILTERS

Takao Hinamoto, Kensuke Kai and Wu-Sheng Lu[†]

Hiroshima University, Higashihiroshima 739-8527, Japan. hinamoto@ieee.org

[†]University of Victoria, Victoria, BC, Canada V8W 3P6. wslu@ece.uvic.ca

ABSTRACT

Several techniques which minimize an l_2 -sensitivity measure subject to l_2 -norm dynamic-range scaling constraints are compared for state-space digital filters. This comparison is performed by solving a numerical example from the viewpoint of the convergence rate and the attained minimum value of an l_2 -sensitivity measure.

I. INTRODUCTION

Implementing a fixed-point state-space digital filter with finite word length (FWL), the efficiency and performance of the filter are directly affected by the choice of its state-space filter structure. If a transfer function satisfying specification requirements is designed with infinite accuracy coefficients and realized by a state-space model, the coefficients in the state-space model must be truncated or rounded to fit the FWL constraints. The characteristics of the filter is then altered due to the coefficient quantization, which may turn a stable filter into an unstable one. Therefore, the problem of minimizing the coefficient sensitivity of a digital filter is a significant research topic. Several techniques have been proposed for synthesizing state-space digital filter structures that minimize an l_2 -sensitivity measure [1]-[4]. More recently, a few techniques have developed for minimizing an l_2 -sensitivity measure subject to l_2 -scaling constraints [5]-[7].

We compare the techniques reported in [5]-[7] and a modified version of that in [5] from the viewpoint of the convergence rate and the attained minimum value of an l_2 -sensitivity measure. This comparison is carried out through a numerical example.

II. l_2 -SENSITIVITY ANALYSIS

Consider a state-space digital filter $(\mathbf{A}, \mathbf{b}, \mathbf{c}, d)_n$ which

is stable, controllable and observable

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

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

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

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

$$\mathbf{S}_{\mathbf{X}} = \frac{\partial f(\mathbf{X})}{\partial \mathbf{X}}, \quad (\mathbf{S}_{\mathbf{X}})_{ij} = \frac{\partial f(\mathbf{X})}{\partial x_{ij}} \quad (3)$$

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

Definition 2: Let $\mathbf{X}(z)$ 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 $\mathbf{X}(z)$. The l_2 -norm of $\mathbf{X}(z)$ is then defined as

$$\|\mathbf{X}(z)\|_2 = \left(\text{tr} \left[\frac{1}{2\pi j} \oint_{|z|=1} \mathbf{X}(z)\mathbf{X}^*(z) \frac{dz}{z} \right] \right)^{\frac{1}{2}}. \quad (4)$$

From (2) and *Definitions 1* and *2*, the overall l_2 -sensitivity measure for the state-space digital filter in (1) is defined as

$$\begin{aligned} S &= \left\| \frac{\partial H(z)}{\partial \mathbf{A}} \right\|_2^2 + \left\| \frac{\partial H(z)}{\partial \mathbf{b}} \right\|_2^2 + \left\| \frac{\partial H(z)}{\partial \mathbf{c}^T} \right\|_2^2 \\ &= \left\| [\mathbf{F}(z)\mathbf{G}(z)]^T \right\|_2^2 + \left\| \mathbf{G}^T(z) \right\|_2^2 + \left\| \mathbf{F}(z) \right\|_2^2 \end{aligned} \quad (5)$$

where

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

The term d in (2) and the sensitivity with respect to it are coordinate-independent and therefore they are neglected here.

It is easy to show that the l_2 -sensitivity measure in (5) can be expressed as

$$S = \text{tr}[\mathbf{M}(\mathbf{I}_n)] + \text{tr}[\mathbf{W}_o] + \text{tr}[\mathbf{K}_c] \quad (6)$$

where

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

$$\mathbf{K}_c = \frac{1}{2\pi j} \oint_{|z|=1} \mathbf{F}(z)\mathbf{F}^T(z^{-1}) \frac{dz}{z}$$

$$\mathbf{W}_o = \frac{1}{2\pi j} \oint_{|z|=1} \mathbf{G}^T(z)\mathbf{G}(z^{-1}) \frac{dz}{z}.$$

The matrices \mathbf{K}_c and \mathbf{W}_o are called the controllability and observability Gramians, respectively. The Gramians $\mathbf{M}(\mathbf{P})$, \mathbf{K}_c and \mathbf{W}_o can be obtained by solving the Lyapunov equations [5]

$$\begin{bmatrix} * & * \\ * & \mathbf{M}(\mathbf{P}) \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{bc} \\ \mathbf{0} & \mathbf{A} \end{bmatrix}^T \begin{bmatrix} * & * \\ * & \mathbf{M}(\mathbf{P}) \end{bmatrix} \\ \times \begin{bmatrix} \mathbf{A} & \mathbf{bc} \\ \mathbf{0} & \mathbf{A} \end{bmatrix} + \begin{bmatrix} \mathbf{P}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \quad (7)$$

$$\mathbf{K}_c = \mathbf{A}\mathbf{K}_c\mathbf{A}^T + \mathbf{bb}^T$$

$$\mathbf{W}_o = \mathbf{A}^T\mathbf{W}_o\mathbf{A} + \mathbf{c}^T\mathbf{c}.$$

If a coordinate transformation defined by

$$\bar{\mathbf{x}}(k) = \mathbf{T}^{-1}\mathbf{x}(k) \quad (8)$$

is applied to the state-space digital filter in (1), then the new realization $(\bar{\mathbf{A}}, \bar{\mathbf{b}}, \bar{\mathbf{c}}, d)_n$ can be characterized by

$$\bar{\mathbf{A}} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}, \quad \bar{\mathbf{b}} = \mathbf{T}^{-1}\mathbf{b}, \quad \bar{\mathbf{c}} = \mathbf{c}\mathbf{T}. \quad (9)$$

From (2) and (9), it is clear that the transfer function $H(z)$ is invariant under the coordinate transformation in (8). The coordinate transformation defined by (8) changes (6) to

$$S(\mathbf{P}) = \text{tr}[\mathbf{M}(\mathbf{P})\mathbf{P}] + \text{tr}[\mathbf{W}_o\mathbf{P}] + \text{tr}[\mathbf{K}_c\mathbf{P}^{-1}] \quad (10)$$

where $\mathbf{P} = \mathbf{T}\mathbf{T}^T$. Moreover, if the l_2 -norm dynamic-range scaling constraints are imposed on the new state-variable vector $\bar{\mathbf{x}}(k)$, it is required that for $i = 1, 2, \dots, n$

$$(\bar{\mathbf{K}}_c)_{ii} = (\mathbf{T}^{-1}\mathbf{K}_c\mathbf{T}^{-T})_{ii} = 1. \quad (11)$$

Consequently, the problem of l_2 -sensitivity minimization subject to l_2 -norm dynamic-range scaling constraints is now formulated as follows: For given \mathbf{A} , \mathbf{b} and \mathbf{c} , obtain an $n \times n$ nonsingular matrix \mathbf{T} which minimizes (10) subject to the constraints in (11).

III. MINIMIZATION TECHNIQUES

In order to minimize (10) over an $n \times n$ symmetric positive-definite matrix \mathbf{P} subject to the constraints shown in (11), we define the Lagrange function

$$J(\mathbf{P}, \lambda) = \text{tr}[\mathbf{M}(\mathbf{P})\mathbf{P}] + \text{tr}[\mathbf{W}_o\mathbf{P}] \\ + \text{tr}[\mathbf{K}_c\mathbf{P}^{-1}] + \lambda(\text{tr}[\mathbf{K}_c\mathbf{P}^{-1}] - n) \quad (12)$$

where λ is a Lagrange multiplier. We compute

$$\frac{\partial J(\mathbf{P}, \lambda)}{\partial \mathbf{P}} = \mathbf{M}(\mathbf{P}) - \mathbf{P}^{-1}\mathbf{N}(\mathbf{P})\mathbf{P}^{-1} + \mathbf{W}_o \\ - (\lambda + 1)\mathbf{P}^{-1}\mathbf{K}_c\mathbf{P}^{-1} \quad (13)$$

$$\frac{\partial J(\mathbf{P}, \lambda)}{\partial \lambda} = \text{tr}[\mathbf{K}_c\mathbf{P}^{-1}] - n$$

where $\mathbf{N}(\mathbf{P})$ can be obtained by solving the Lyapunov equation [5]

$$\begin{bmatrix} \mathbf{N}(\mathbf{P}) & * \\ * & * \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{bc} \\ \mathbf{0} & \mathbf{A} \end{bmatrix} \begin{bmatrix} \mathbf{N}(\mathbf{P}) & * \\ * & * \end{bmatrix} \\ \times \begin{bmatrix} \mathbf{A} & \mathbf{bc} \\ \mathbf{0} & \mathbf{A} \end{bmatrix}^T + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{P} & \mathbf{0} \end{bmatrix}.$$

From $\partial J(\mathbf{P}, \lambda)/\partial \mathbf{P} = \mathbf{0}$ and $\partial J(\mathbf{P}, \lambda)/\partial \lambda = 0$, we get

$$\mathbf{P}\mathbf{F}(\mathbf{P})\mathbf{P} = \mathbf{G}(\mathbf{P}, \lambda), \quad \text{tr}[\mathbf{K}_c\mathbf{P}^{-1}] = n \quad (14)$$

where

$$\mathbf{F}(\mathbf{P}) = \mathbf{M}(\mathbf{P}) + \mathbf{W}_o$$

$$\mathbf{G}(\mathbf{P}, \lambda) = \mathbf{N}(\mathbf{P}) + (\lambda + 1)\mathbf{K}_c.$$

A. Method 1 [5]

Start with initial conditions $\lambda_0 = 0$ and $\mathbf{P}_0 = \mathbf{I}_n$ and repeat

$$\mathbf{P}_{i+1} = \mathbf{F}(\mathbf{P}_i)^{-\frac{1}{2}} [\mathbf{F}(\mathbf{P}_i)^{\frac{1}{2}} \mathbf{G}(\mathbf{P}_i, \lambda_i) \mathbf{F}(\mathbf{P}_i)^{\frac{1}{2}}]^{-\frac{1}{2}} \mathbf{F}(\mathbf{P}_i)^{-\frac{1}{2}} \\ \lambda_{i+1} = \frac{\text{tr}[\mathbf{P}_i \mathbf{F}(\mathbf{P}_i)] - \text{tr}[\mathbf{N}(\mathbf{P}_i) \mathbf{P}_i^{-1}]}{n} - 1 \quad (15)$$

where \mathbf{P}_i and λ_i are solutions of the previous iteration.

B. Method 2 (A modified version of [5])

Start with initial condition $\mathbf{P}_0 = \mathbf{I}_n$ and repeat

$$\lambda_i = \frac{\text{tr}[\mathbf{P}_i \mathbf{F}(\mathbf{P}_i)] - \text{tr}[\mathbf{N}(\mathbf{P}_i) \mathbf{P}_i^{-1}]}{n} - 1 \\ \mathbf{P}_{i+1} = \mathbf{F}(\mathbf{P}_i)^{-\frac{1}{2}} [\mathbf{F}(\mathbf{P}_i)^{\frac{1}{2}} \mathbf{G}(\mathbf{P}_i, \lambda_i) \mathbf{F}(\mathbf{P}_i)^{\frac{1}{2}}]^{-\frac{1}{2}} \mathbf{F}(\mathbf{P}_i)^{-\frac{1}{2}}. \quad (16)$$

C. Method 3 [6]

Start with initial condition $\mathbf{P}_0 = \mathbf{I}_n$ and repeat

$$\mathbf{S}_{i+1} = \mathbf{F}(\mathbf{P}_i)^{-\frac{1}{2}} [\mathbf{F}(\mathbf{P}_i)^{\frac{1}{2}} \mathbf{G}(\mathbf{P}_i) \mathbf{F}(\mathbf{P}_i)^{\frac{1}{2}}]^{-\frac{1}{2}} \mathbf{F}(\mathbf{P}_i)^{-\frac{1}{2}}$$

$$\mathbf{P}_{i+1} = \frac{\text{tr}[\mathbf{K}_c \mathbf{S}_{i+1}^{-1}]}{n} \mathbf{S}_{i+1} \quad (17)$$

where $\mathbf{G}(\mathbf{P}_i) = \mathbf{G}(\mathbf{P}_i, 0)$.

D. Method 4 [7]

Start with initial condition $\mathbf{P}_0 = \mathbf{I}_n$ and repeat

$$\begin{aligned} \mathbf{P}_{i+1} = & \mathbf{F}(\mathbf{P}_i)^{-\frac{1}{2}} [\mathbf{F}(\mathbf{P}_i)^{\frac{1}{2}} \mathbf{G}(\mathbf{P}_i, \lambda_{i+1}) \mathbf{F}(\mathbf{P}_i)^{\frac{1}{2}}]^{\frac{1}{2}} \\ & \times \mathbf{F}(\mathbf{P}_i)^{-\frac{1}{2}}. \end{aligned} \quad (18)$$

Here, λ_{i+1} is obtained using a bisection method so that

$$f(\lambda_{i+1}) = n - \text{tr}[\hat{\mathbf{K}}_c(\mathbf{P}_i) \hat{\mathbf{G}}(\mathbf{P}_i, \lambda_{i+1})] = 0 \quad (19)$$

where

$$\begin{aligned} \hat{\mathbf{K}}_c(\mathbf{P}_i) &= \mathbf{F}(\mathbf{P}_i)^{\frac{1}{2}} \mathbf{K}_c \mathbf{F}(\mathbf{P}_i)^{\frac{1}{2}} \\ \hat{\mathbf{G}}(\mathbf{P}_i, \lambda_{i+1}) &= [\mathbf{F}(\mathbf{P}_i)^{\frac{1}{2}} \mathbf{G}(\mathbf{P}_i, \lambda_{i+1}) \mathbf{F}(\mathbf{P}_i)^{\frac{1}{2}}]^{-\frac{1}{2}}. \end{aligned}$$

The above iterative procedures continue until

$$|J(\mathbf{P}_{i+1}, \lambda_{i+1}) - J(\mathbf{P}_i, \lambda_i)| < \varepsilon \quad (20)$$

is satisfied where ε is a prescribed tolerance.

IV. A NUMERICAL EXAMPLE

Let a state-space digital filter in (1) be specified by

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0.453770 & -1.556160 & 1.974860 \end{bmatrix}$$

$$\mathbf{b} = [0 \quad 0 \quad 1]^T$$

$$\mathbf{c} = [0.023170 \quad 0.023020 \quad 0.79300]$$

$$d = 0.015940.$$

Performing the computation of (7) and the l_2 -scaling, the Grammians \mathbf{K}_c , \mathbf{W}_o and $\mathbf{M}(\mathbf{I}_3)$ were calculated as

$$\mathbf{K}_c = \begin{bmatrix} 1.000000 & 0.872501 & 0.562821 \\ 0.872501 & 1.000000 & 0.872501 \\ 0.562821 & 0.872501 & 1.000000 \end{bmatrix}$$

$$\mathbf{W}_o = \begin{bmatrix} 0.820741 & -2.035330 & 1.628162 \\ -2.035330 & 5.307276 & -4.264906 \\ 1.628162 & -4.264906 & 3.941493 \end{bmatrix}$$

$$\mathbf{M}(\mathbf{I}_3) = \begin{bmatrix} 8.921384 & -22.046468 & 17.916293 \\ -22.046468 & 55.671739 & -46.052035 \\ 17.916293 & -46.052035 & 42.522180 \end{bmatrix}.$$

The l_2 -sensitivity measure in (6) was computed as

$$S = 120.184738.$$

Applying Method 1 [5] and Method 2 (A modified version of [5]) with $\varepsilon = 10^{-7}$, it took the iterative algorithms 246 and 231 iterations, respectively, to converge to

$$\mathbf{P}^{opt} = \begin{bmatrix} 2.307530 & 1.375668 & 0.514401 \\ 1.375668 & 1.103115 & 0.678193 \\ 0.514401 & 0.678193 & 0.666912 \end{bmatrix}$$

which yields

$$\mathbf{T}^{opt} = \begin{bmatrix} -0.816495 & 0.220555 & 1.261833 \\ -0.093245 & -0.076884 & 1.043316 \\ 0.437414 & 0.006321 & 0.689595 \end{bmatrix}.$$

In both cases, the l_2 -sensitivity measure in (10) was minimized subject to the scaling constraints in (11) to

$$S(\mathbf{P}^{opt}) = 8.672129.$$

The l_2 -sensitivity performances of 500 iterations for Method 1 [5] and Method 2 (A modified version of [5]) are shown in Fig. 1.

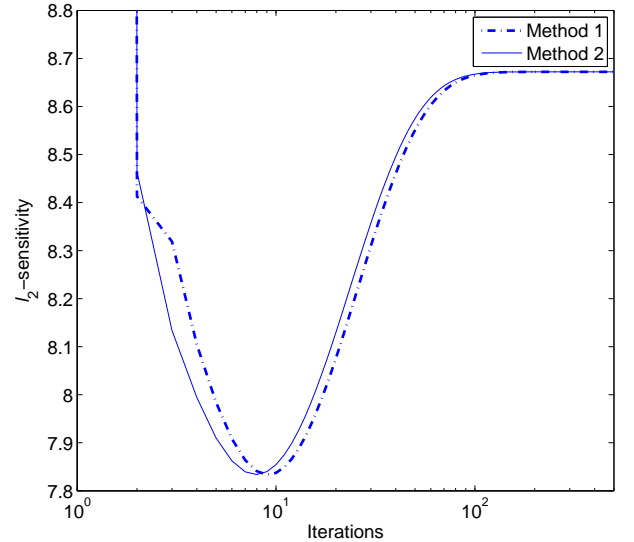


Figure 1: Performances of Methods 1 and 2.

Applying Method 3 [6] with $\varepsilon = 10^{-7}$, it took the iterative algorithm 8 iterations to converge to

$$\mathbf{P}^{opt} = \begin{bmatrix} 2.323990 & 1.368404 & 0.507227 \\ 1.368404 & 1.088538 & 0.665792 \\ 0.507227 & 0.665792 & 0.653786 \end{bmatrix}$$

which yields

$$\mathbf{T}^{opt} = \begin{bmatrix} -0.811257 & 0.218876 & 1.271985 \\ -0.077911 & -0.069596 & 1.038087 \\ 0.442333 & 0.026430 & 0.676335 \end{bmatrix}.$$

Table 1: Iterations to become less than $\varepsilon = 10^{-7}$

Method	1	2	3	4
Iterations	246	231	8	8

Table 2: Iterations to become less than $\varepsilon = 10^{-8}$

Method	1	2	3	4
Iterations	286	269	10	9

In this case, the l_2 -sensitivity measure in (10) was minimized subject to the scaling constraints in (11) to

$$S(\mathbf{P}^{opt}) = 8.677268.$$

Applying Method 4 [7] with $\varepsilon = 10^{-7}$, it took the iterative algorithm 8 iterations to converge to the minimum value identical to those obtained by Methods 1 and 2.

The l_2 -sensitivity performances of 500 iterations for Methods 3 [6] and 4 [7] are shown in Fig. 2.

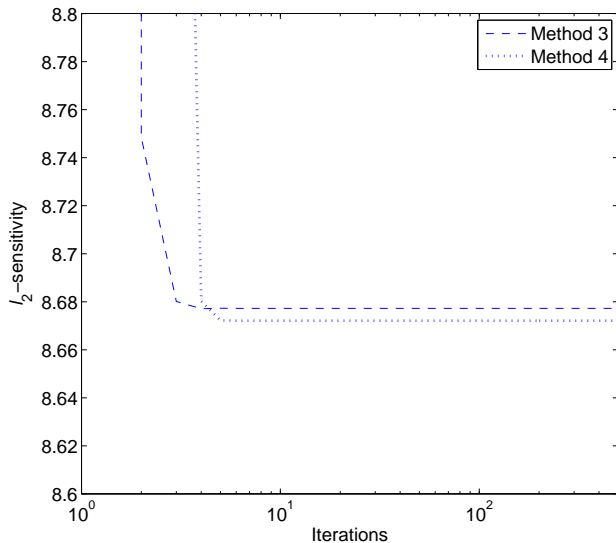


Figure 2: Performances of Methods 3 and 4.

V. CONCLUSION

This paper has compared the techniques reported in [5]-[7] and a modified version of that in [5] from the viewpoint of the convergence rate and the attained minimum value of an l_2 -sensitivity measure. This compar-

Table 3: Iterations to attain $J(\mathbf{P}, \lambda) = 8.672132$

Method	1	2	3	4
Iterations	274	256	Not Applicable	8

ison has been accomplished by solving a numerical example. As a result, it has been observed that the modified version of the technique in [5] yields faster convergence than that in [5]. Moreover, the minimum value attained by the technique reported in [6] is higher than those obtained by the other techniques. It is noted that the technique reported in [6] minimizes an l_2 -sensitivity measure without l_2 -scaling constraints, and then the l_2 -scaling is carried out to the resulting solution (filter).

REFERENCES

- [1] 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.
- [2] G. Li and M. Gevers, "Optimal synthetic FWL design of state-space digital filters," *Proc. 1992 IEEE Int. Conf. Acoust., Speech, Signal Processing*, vol.4, pp.429-432.
- [3] M. Gevers and G. Li, *Parameterizations in Control, Estimation and Filtering Problems: Accuracy Aspects*, Springer-Verlag, 1993.
- [4] 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.
- [5] T. Hinamoto, K. Iwata and W.-S. Lu, "L2-sensitivity minimization of one- and two-dimensional state-space digital filters subject to L2-scaling constraints," *IEEE Trans. Signal Process.*, vol. 54, no. 5, pp. 1804-1812, May 2006.
- [6] S. Yamaki, M. Abe and M. Kawamata, "A novel approach to L_2 -sensitivity minimization of digital filters subject to L_2 -scaling constraints," *Proc. 2006 IEEE Int. Symp. Circuits Syst.*, pp. 5219-5222.
- [7] T. Hinamoto and T. Oumi, "Minimization of L_2 -sensitivity for linear discrete-time systems subject to L_2 -scaling constraints using a Lagrange function and a bisection method," *Proc. 2006 SICE-ICASE Int. Joint Conf.*, pp. 1963-1967.