

JOINT OPTIMIZATION OF ERROR FEEDBACK AND COORDINATE TRANSFORMATION FOR ROUND-OFF NOISE MINIMIZATION IN STATE-SPACE DIGITAL FILTERS

Takao Hinamoto, Hiroaki Ohnishi and Wu-Sheng Lu[†]

Graduate School of Engineering, Hiroshima University, Japan
hinamoto@hiroshima-u.ac.jp

[†]Dept. Elect. Comput. Eng., University of Victoria, Canada
wslu@ece.uvic.ca

ABSTRACT

An iterative approach for joint optimization of a scalar error-feedback matrix and a coordinate transformation matrix is developed to minimize the roundoff noise subject to the l_2 -norm dynamic-range scaling constraints. When the iterative algorithm converges and the optimal coordinate transformation matrix is obtained, the diagonal error-feedback matrix is derived to minimize the noise gain in the optimal state-space realization. This diagonal error-feedback matrix enables one to produce more reduction of the noise gain. Finally, a numerical example is given to illustrate the utility of the proposed technique.

1. INTRODUCTION

As is well known, error feedback is an effective technique for the reduction of the roundoff noise due to signal quantizations [1]-[4]. Alternatively, the problem of synthesizing state-space digital filter structures which minimize roundoff noise under l_2 -norm scaling constraints has been investigated [5],[6]. Moreover, it has been shown that the output quantization noise can be reduced more effectively by choosing the filter structure and applying the concept of error feedback [7]-[9].

This paper investigates the problem of minimizing roundoff noise under l_2 -norm dynamic-range scaling constraints in state-space digital filters by means of joint optimization of error feedback and coordinate transformation. An iterative procedure is proposed for the joint optimization. When the iterative algorithm converges, the optimal coordinate transformation matrix is obtained, which is used to construct the optimal state-space realization. Then, the diagonal error-feedback matrix is derived to minimize the noise gain in the optimal realization. Our simulation results demonstrate the validity of the proposed technique.

2. ERROR FEEDBACK IN STATE-SPACE DIGITAL FILTERS

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

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

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 filter (1) is assumed controllable and observable.

Taking into account the quantizations performed before matrix-vector multiplication, one can express an finite-word-length (FWL) implementation of (1) with error feedback as

$$\begin{aligned}\tilde{\mathbf{x}}(k+1) &= \mathbf{A}\mathbf{Q}[\tilde{\mathbf{x}}(k)] + \mathbf{b}u(k) + \mathbf{D}\mathbf{e}(k) \\ \tilde{y}(k) &= \mathbf{c}\mathbf{Q}[\tilde{\mathbf{x}}(k)] + du(k)\end{aligned}\quad (2)$$

where $\mathbf{e}(k) = \tilde{\mathbf{x}}(k) - \mathbf{Q}[\tilde{\mathbf{x}}(k)]$ and \mathbf{D} is referred to as an *error-feedback matrix* of dimension $n \times n$.

All coefficient matrices $\mathbf{A}, \mathbf{b}, \mathbf{c}$, and d are assumed to have an exact fractional B_c bit representation. The FWL state-variable vector $\tilde{\mathbf{x}}(k)$ and the output $\tilde{y}(k)$ all have a B bit fractional representation, while the input $u(k)$ is a $(B - B_c)$ bit fraction. The quantizer $\mathbf{Q}[\cdot]$ in (2) rounds the B bit fraction $\tilde{\mathbf{x}}(k)$ to $(B - B_c)$ bits after completing the multiplications and additions, where the sign bit is not counted. It is assumed that the roundoff error $\mathbf{e}(k)$ can be modeled as a zero-mean noise process with covariance $\sigma^2 \mathbf{I}_n$. Subtracting (2) from (1) yields

$$\begin{aligned}\Delta\mathbf{x}(k+1) &= \mathbf{A}\Delta\mathbf{x}(k) + (\mathbf{A} - \mathbf{D})\mathbf{e}(k) \\ \Delta y(k) &= \mathbf{c}\Delta\mathbf{x}(k) + \mathbf{c}\mathbf{e}(k)\end{aligned}\quad (3)$$

where

$$\Delta \mathbf{x}(k) = \mathbf{x}(k) - \tilde{\mathbf{x}}(k), \quad \Delta y(k) = y(k) - \tilde{y}(k).$$

Taking the z-transform on both sides of (3) and setting $\Delta \mathbf{x}(0) = \mathbf{0}$ gives

$$\begin{aligned} \Delta Y(z) &= \mathbf{G}_D(z) \mathbf{E}(z) \\ \mathbf{G}_D(z) &= \mathbf{c}(z\mathbf{I}_n - \mathbf{A})^{-1}(\mathbf{A} - \mathbf{D}) + \mathbf{c} \end{aligned} \quad (4)$$

where $\Delta Y(z)$ and $\mathbf{E}(z)$ represent the z-transforms of $\Delta y(k)$ and $\mathbf{e}(k)$, respectively.

The normalized noise gain $I(\mathbf{D}) = \sigma_{out}^2/\sigma^2$ is then defined by

$$I(\mathbf{D}) = \text{tr}[\mathbf{W}_D] \quad (5)$$

where

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

By utilizing the Cauchy integral theorem, matrix \mathbf{W}_D defined in (5) can be expressed in closed form as

$$\mathbf{W}_D = (\mathbf{A} - \mathbf{D})^T \mathbf{W}_o (\mathbf{A} - \mathbf{D}) + \mathbf{c}^T \mathbf{c} \quad (6)$$

where \mathbf{W}_o is the observability Gramian of the filter that can be obtained by solving the Lyapunov equation

$$\mathbf{W}_o = \mathbf{A}^T \mathbf{W}_o \mathbf{A} + \mathbf{c}^T \mathbf{c}. \quad (7)$$

Alternatively, the controllability Gramian \mathbf{K}_c can be obtained by solving the Lyapunov equation

$$\mathbf{K}_c = \mathbf{A} \mathbf{K}_c \mathbf{A}^T + \mathbf{b} \mathbf{b}^T. \quad (8)$$

The filter (1) is changed via coordinate transformation $\bar{\mathbf{x}}(k) = \mathbf{T}^{-1} \mathbf{x}(k)$ to a new realization $(\bar{\mathbf{A}}, \bar{\mathbf{b}}, \bar{\mathbf{c}}, d)_n$ with

$$\begin{aligned} \bar{\mathbf{A}} &= \mathbf{T}^{-1} \mathbf{A} \mathbf{T}, & \bar{\mathbf{b}} &= \mathbf{T}^{-1} \mathbf{b}, & \bar{\mathbf{c}} &= \mathbf{c} \mathbf{T} \\ \bar{\mathbf{W}}_o &= \mathbf{T}^T \mathbf{W}_o \mathbf{T}, & \bar{\mathbf{K}}_c &= \mathbf{T}^{-1} \mathbf{K}_c \mathbf{T}^{-T}. \end{aligned} \quad (9)$$

The problem considered here is to jointly optimize a scalar error-feedback matrix $\alpha \mathbf{I}_n$ and a coordinate transformation matrix \mathbf{T} for roundoff noise minimization under l_2 -norm dynamic-range scaling constraints:

$$(\bar{\mathbf{K}}_c)_{ii} = (\mathbf{T}^{-1} \mathbf{K}_c \mathbf{T}^{-T})_{ii} = 1, \quad i = 1, 2, \dots, n. \quad (10)$$

If the unit noise matrix \mathbf{W}_D , (6), with $\mathbf{D} = \alpha \mathbf{I}_n$ is denoted by \mathbf{W}_α , then the noise gain $I(\mathbf{D})$ defined in (5) is written as $\text{tr}[\mathbf{T}^T \mathbf{W}_\alpha \mathbf{T}]$ under joint optimization of scalar error-feedback and coordinate transformation.

The proposed joint optimization will be carried out in an iterative manner.

3. AN ITERATIVE PROCEDURE FOR JOINT OPTIMIZATION

In order to minimize $\text{tr}[\mathbf{T}^T \mathbf{W}_\alpha \mathbf{T}]$ (with α fixed) over an $n \times n$ nonsingular matrix \mathbf{T} subject to the constraints shown in (10), we define the Lagrange function

$$J(\alpha, \mathbf{P}, \lambda) = \text{tr}[\mathbf{W}_\alpha \mathbf{P}] + \lambda(\text{tr}[\mathbf{K}_c \mathbf{P}^{-1}] - n) \quad (11)$$

where $\mathbf{P} = \mathbf{T} \mathbf{T}^T$ and λ is a Lagrange multiplier. We compute

$$\begin{aligned} \frac{\partial J(\alpha, \mathbf{P}, \lambda)}{\partial \alpha} &= 2(\alpha \text{tr}[\mathbf{W}_o \mathbf{P}] - \text{tr}[\mathbf{W}_o \mathbf{A} \mathbf{P}]) \\ \frac{\partial J(\alpha, \mathbf{P}, \lambda)}{\partial \mathbf{P}} &= \mathbf{W}_\alpha - \lambda \mathbf{P}^{-1} \mathbf{K}_c \mathbf{P}^{-1} \end{aligned} \quad (12)$$

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

Letting $\partial J(\alpha, \mathbf{P}, \lambda)/\partial \alpha = 0$ yields

$$\alpha = \frac{\text{tr}[\mathbf{W}_o \mathbf{A} \mathbf{P}]}{\text{tr}[\mathbf{W}_o \mathbf{P}]} \quad (13)$$

Letting $\partial J(\alpha, \mathbf{P}, \lambda)/\partial \mathbf{P} = \mathbf{0}$ and $\partial J(\alpha, \mathbf{P}, \lambda)/\partial \lambda = 0$,

$$\mathbf{P} \mathbf{W}_\alpha \mathbf{P} = \lambda \mathbf{K}_c, \quad \text{tr}[\mathbf{K}_c \mathbf{P}^{-1}] = n. \quad (14)$$

It follows from (14) that

$$\begin{aligned} \mathbf{P} &= \sqrt{\lambda} \mathbf{W}_\alpha^{-\frac{1}{2}} [\mathbf{W}_\alpha^{\frac{1}{2}} \mathbf{K}_c \mathbf{W}_\alpha^{\frac{1}{2}}]^{\frac{1}{2}} \mathbf{W}_\alpha^{-\frac{1}{2}} \\ \frac{1}{\sqrt{\lambda}} \text{tr}[\mathbf{K}_c \mathbf{W}_\alpha]^{\frac{1}{2}} &= \frac{1}{\sqrt{\lambda}} \left(\sum_{i=1}^n \theta_i \right) = n \end{aligned} \quad (15)$$

where θ_i^2 for $i = 1, 2, \dots, n$ are the eigenvalues of $\mathbf{K}_c \mathbf{W}_\alpha$. This can be used to obtain

$$\mathbf{P} = \frac{1}{n} \left(\sum_{i=1}^n \theta_i \right) \mathbf{W}_\alpha^{-\frac{1}{2}} [\mathbf{W}_\alpha^{\frac{1}{2}} \mathbf{K}_c \mathbf{W}_\alpha^{\frac{1}{2}}]^{\frac{1}{2}} \mathbf{W}_\alpha^{-\frac{1}{2}}. \quad (16)$$

Substituting (16) into (11) yields the minimum value of $J(\alpha, \mathbf{P}, \lambda)$ for a given scalar α as

$$\min_{\mathbf{P}, \lambda} J(\alpha, \mathbf{P}, \lambda) = \frac{1}{n} \left(\sum_{i=1}^n \theta_i \right)^2. \quad (17)$$

This completes the first round of iteration and this process may continue until both \mathbf{P} and α converge. Having obtained an $n \times n$ symmetric positive-definite matrix \mathbf{P} , an improved value of scalar α can be obtained using (13).

This iterative procedure for minimizing (11) with respect to a scalar parameter α as well as matrix \mathbf{P} can be summarized as follows:

1) Set $i = 1$ and

$$\mathbf{P}(0) = \text{diag}\{(\mathbf{K}_c)_{11}^{-1}, (\mathbf{K}_c)_{22}^{-1}, \dots, (\mathbf{K}_c)_{nn}^{-1}\}.$$

2) Compute a scalar $\alpha(i)$ using

$$\alpha(i) = \frac{\text{tr}[\mathbf{W}_o \mathbf{A} \mathbf{P}(i-1)]}{\text{tr}[\mathbf{W}_o \mathbf{P}(i-1)]}.$$

3) Compute

$$I_{\min}(\alpha(i) \mathbf{I}_n) = (1 - \alpha(i)^2) \text{tr}[\mathbf{W}_o \mathbf{P}(i-1)].$$

4) Replace \mathbf{W}_α by $\mathbf{W}_{\alpha(i)}$ computed using

$$\mathbf{W}_{\alpha(i)} = (1 + \alpha(i)^2) \mathbf{W}_o - \alpha(i) (\mathbf{A}^T \mathbf{W}_o + \mathbf{W}_o \mathbf{A}).$$

5) Derive matrix \mathbf{P} from (17), and take the resulting \mathbf{P} as $\mathbf{P}(i)$.

6) Compute $\text{tr}[\mathbf{W}_{\alpha(i)} \mathbf{P}(i)]$.

7) Update i to $i + 1$.

8) Repeat from Step 2) until the change in either $I_{\min}[\alpha(i) \mathbf{I}_n]$ or $\text{tr}[\mathbf{W}_{\alpha(i)} \mathbf{P}(i)]$ becomes negligible.

Next, the coordinate transformation matrix \mathbf{T} will be constructed so that (10) is satisfied. From (16), the optimal coordinate transformation matrix \mathbf{T} that minimizes (11) can be obtained in closed form as

$$\mathbf{T} = \frac{1}{\sqrt{n}} \left(\sum_{i=1}^n \theta_i \right)^{\frac{1}{2}} \mathbf{W}_\alpha^{-\frac{1}{2}} [\mathbf{W}_\alpha^{\frac{1}{2}} \mathbf{K}_c \mathbf{W}_\alpha^{\frac{1}{2}}]^{\frac{1}{4}} \mathbf{U} \quad (18)$$

where \mathbf{U} is an arbitrary $n \times n$ orthogonal matrix. From (18) it follows that

$$\begin{aligned} \overline{\mathbf{K}}_c &= \mathbf{T}^{-1} \mathbf{K}_c \mathbf{T}^{-T} \\ &= n \left(\sum_{i=1}^n \theta_i \right)^{-1} \mathbf{U}^T [\mathbf{W}_\alpha^{\frac{1}{2}} \mathbf{K}_c \mathbf{W}_\alpha^{\frac{1}{2}}]^{\frac{1}{2}} \mathbf{U}. \end{aligned} \quad (19)$$

Let us choose the $n \times n$ orthogonal matrix \mathbf{U} such that the matrix $\overline{\mathbf{K}}_c$ in (9) satisfies the l_2 -norm dynamic-range scaling constraints, (10), on the state-variables. To this end, we perform the eigenvalue-eigenvector decomposition

$$[\mathbf{W}_\alpha^{\frac{1}{2}} \mathbf{K}_c \mathbf{W}_\alpha^{\frac{1}{2}}]^{\frac{1}{2}} = \mathbf{R} \mathbf{\Theta} \mathbf{R}^T \quad (20)$$

where $\mathbf{\Theta} = \text{diag}\{\theta_1, \theta_2, \dots, \theta_n\}$ and $\mathbf{R} \mathbf{R}^T = \mathbf{I}_n$. This yields

$$n \left(\sum_{i=1}^n \theta_i \right)^{-1} [\mathbf{W}_\alpha^{\frac{1}{2}} \mathbf{K}_c \mathbf{W}_\alpha^{\frac{1}{2}}]^{\frac{1}{2}} = \mathbf{R} \mathbf{\Lambda}^{-2} \mathbf{R}^T \quad (21)$$

where $\mathbf{\Lambda} = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ and for $i = 1, 2, \dots, n$, $\lambda_i = ((\theta_1 + \theta_2 + \dots + \theta_n)/n\theta_i)^{\frac{1}{2}}$. Now an $n \times n$ orthogonal matrix \mathbf{S} such that

$$\mathbf{S} \mathbf{\Lambda}^{-2} \mathbf{S}^T = \begin{bmatrix} 1 & * & \cdots & * \\ * & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ * & \cdots & * & 1 \end{bmatrix} \quad (22)$$

can be obtained by numerical manipulations [6, p.278]. By choosing $\mathbf{U} = \mathbf{R} \mathbf{S}^T$ in (18), the optimal coordinate transformation matrix \mathbf{T} both satisfying (10) and minimizing (11) can now be constructed as

$$\mathbf{T} = \frac{1}{\sqrt{n}} \left(\sum_{i=1}^n \theta_i \right)^{\frac{1}{2}} \mathbf{W}_\alpha^{-\frac{1}{2}} [\mathbf{W}_\alpha^{\frac{1}{2}} \mathbf{K}_c \mathbf{W}_\alpha^{\frac{1}{2}}]^{\frac{1}{4}} \mathbf{R} \mathbf{S}^T. \quad (23)$$

Suppose the iterative algorithm converges after N iterations and the optimal coordinate transformation matrix $\mathbf{T}(N)$ has been computed from (20)-(23). Then, the diagonal error-feedback matrix $\mathbf{D} = \text{diag}\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ that minimizes

$$\begin{aligned} I(\mathbf{D}) &= \text{tr}[\mathbf{T}^T(N) \mathbf{W}_o \mathbf{T}(N)] + \text{tr}[\mathbf{T}^T(N) \mathbf{W}_o \mathbf{T}(N) \mathbf{D}^2] \\ &\quad - 2 \text{tr}[\mathbf{T}^T(N) \mathbf{A}^T \mathbf{W}_o \mathbf{T}(N) \mathbf{D}] \end{aligned} \quad (24)$$

is given by

$$\alpha_i = \frac{(\mathbf{T}^T(N) \mathbf{W}_o \mathbf{A} \mathbf{T}(N))_{ii}}{(\mathbf{T}^T(N) \mathbf{W}_o \mathbf{T}(N))_{ii}}, \quad i = 1, 2, \dots, n. \quad (25)$$

This diagonal error-feedback matrix \mathbf{D} makes it possible to produce more reduction of the noise gain, i.e.,

$$I_{\min}(\mathbf{D}) < I_{\min}[\alpha(N) \mathbf{I}_n]. \quad (26)$$

4. A NUMERICAL EXAMPLE

Let a state-space digital filter $(\mathbf{A}, \mathbf{b}, \mathbf{c}, d)_3$ be described in a controllable canonical form as

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0.339377 & -1.152652 & 1.520167 \end{bmatrix}$$

$$\mathbf{b} = [0 \ 0 \ 0.437881]^T$$

$$\mathbf{c} = [0.212964 \ 0.293733 \ 0.718718]$$

$$d = 6.59592 \times 10^{-2}$$

which satisfies l_2 -norm dynamic-range scaling constraints, and yields $I(\mathbf{0}) = \text{tr}[\mathbf{W}_o] = 11.133150$.

We apply the iterative optimization procedure in Section 3 to this filter. The convergent profile of first 10 iterations is given in Table I, from which we see that the algorithm converges after six iterations to a scalar $\alpha = 0.647686$ and $I_{min}(\alpha \mathbf{I}_n) = \text{tr}[\mathbf{T}^T \mathbf{W}_\alpha \mathbf{T}] = 1.450048$. In this case, the coordinate transformation matrix \mathbf{T}^o is given by

$$\mathbf{T}^o = \begin{bmatrix} -1.973853 & -0.153371 & -2.328357 \\ -0.063334 & -1.398294 & -1.260527 \\ 1.402772 & -0.676604 & -0.969851 \end{bmatrix}.$$

TABLE I
CONVERGENT PROFILE
OF FIRST 10 ITERATIONS

| i | $\alpha(i)$ | $I_{min}[\alpha(i)\mathbf{I}_n]$ | $\text{tr}[\mathbf{W}_{\alpha(i)}\mathbf{P}(i)]$ |
|-----|-------------|----------------------------------|--|
| 1 | 0.764400 | 4.627965 | 1.482085 |
| 2 | 0.655454 | 1.451873 | 1.450188 |
| 3 | 0.648286 | 1.450059 | 1.450049 |
| 4 | 0.647733 | 1.450048 | 1.450048 |
| 5 | 0.647689 | 1.450048 | 1.450048 |
| 6 | 0.647686 | 1.450048 | 1.450048 |
| 7 | 0.647686 | 1.450048 | 1.450048 |
| 8 | 0.647686 | 1.450048 | 1.450048 |
| 9 | 0.647686 | 1.450048 | 1.450048 |
| 10 | 0.647686 | 1.450048 | 1.450048 |

If $\alpha = 0.647686$ is rounded to power-of-two representation with 3 bits after binary point, then the noise gain is founded to be $I(\alpha \mathbf{I}_n) = 1.451335$ where $\alpha = 0.625$.

Next, a refined solution which offers further reduced noise gain is deduced by applying an optimal diagonal error-feedback matrix to the optimized realization, i.e., $(\mathbf{T}^{o-1} \mathbf{A} \mathbf{T}^o, \mathbf{T}^{o-1} \mathbf{b}, \mathbf{c} \mathbf{T}^o, d)_3$. The optimal diagonal error-feedback matrix obtained using (25) is given by

$$\mathbf{D} = \text{diag}\{0.705402, 0.510713, 0.683277\}$$

which yields $I_{min}(\mathbf{D}) = 1.433755$.

The above diagonal error-feedback matrix after 3-bit quantization (power-of-two representation with 3 bits after binary point) gives $I_{min}(\mathbf{D}) = 1.438801$, which is less than $I_{min}(\mathbf{D}) = 1.450049$ in the optimal scalar error-feedback.

To compare the proposed method with those reported in [6]-[8], we choose $\mathbf{D} = \mathbf{0}$, i.e., $\mathbf{W}_D = \mathbf{W}_o$ as in [6] or $\mathbf{D} = \mathbf{I}_n$ as in [7],[8], and minimize $\text{tr}[\mathbf{T}^T \mathbf{W}_D \mathbf{T}]$ with respect to matrix \mathbf{T} under the constraints of (10):

$$\min_{\mathbf{T}} \text{tr}[\mathbf{T}^T \mathbf{W}_o \mathbf{T}] = 2.355360$$

$$\min_{\mathbf{T}} \text{tr}[\mathbf{T}^T \mathbf{W}_D \mathbf{T}] = 1.752546$$

which are considerably larger than our results described above.

5. CONCLUSION

The roundoff noise minimization in state-space digital filters has been considered. The noise minimization problem has been addressed in scenario where a scalar error-feedback matrix and a coordinate transformation matrix are jointly optimized subject to usual l_2 -norm dynamic-range scaling constraints. Simulation results have been presented to illustrate and support our theoretical analysis and proposed algorithm.

The extension of the results obtained in this paper to multidimensional case will appear elsewhere.

REFERENCES

- [1] W. E. Higgins and D. C. Munson, "Noise reduction strategies for digital filters: Error spectrum shaping versus the optimal linear state-space formulation," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 30, pp. 963-973, Dec. 1982.
- [2] W. E. Higgins and D. C. Munson, "Optimal and suboptimal error-spectrum shaping for cascade-form digital filters," *IEEE Trans. Circuits Syst.*, vol.CAS-31, pp.429-437, May 1984.
- [3] T. I. Laakso and I. O. Hartimo, "Noise reduction in recursive digital filters using high-order error feedback," *IEEE Trans. Signal Processing*, vol.40, pp.1096-1107, May 1992.
- [4] P. P. Vaidyanathan, "On error-spectrum shaping in state-space digital filters," *IEEE Trans. Circuits Syst.*, vol.CAS-32, pp.88-92, Jan. 1985.
- [5] C. T. Mullis and R. A. Roberts, "Synthesis of minimum roundoff noise fixed point digital filters," *IEEE Trans. Circuits Syst.*, vol.CAS-23, pp.551-562, Sept. 1976.
- [6] S. Y. Hwang, "Minimum uncorrelated unit noise in state-space digital filtering," *IEEE Trans. Acoust. Speech, Signal Processing*, vol.ASSP-25, pp.273-281, Aug. 1977.
- [7] D. Williamson, "Roundoff noise minimization and pole-zero sensitivity in fixed-point digital filters using residue feedback," *IEEE Trans. Acoust. Speech, Signal Processing*, vol.ASSP-34, pp.1210-1220, Oct. 1986.
- [8] G. Li and M. Gevers, "Roundoff noise minimization using delta-operator realizations," *IEEE Trans. Signal Processing*, vol.41, pp.629-637, Feb. 1993.
- [9] T. Hinamoto and S. Kanemori, "Error spectrum shaping in state-space digital filters with l_2 -scaling constraints," in Proc. *2000 IEEE Int. Symp. Circuits Syst.*, Geneva, Switzerland, vol.2, pp.329-332, May 2000.