Roundoff Noise Analysis for Generalized Direct-Form II Structure of 2-D Separable-Denominator Digital Filters

Takao Hinamoto
Hiroshima University
Higashi-Hiroshima 739-8527, Japan
Email: hinamoto@ieee.org

Akimitsu Doi
Hiroshima Institute of Technology
Hiroshima 731-5193, Japan
Email: doi@cc.it-hiroshima.ac.jp

Wu-Sheng Lu
University of Victoria
Victoria, BC, Canada V8W 3P6
Email: wslu@ece.uvic.ca

Abstract—Based on the concept of polynomial operators, generalized direct-form II structure of two-dimensional (2-D) separable-denominator (SD) digital filters is explored. It is shown that 2-D SD digital filters can be modeled by a generalized SIMO direct-form II and a generalized MISO transposed direct-form II that are connected in cascade. Then an expression for the roundoff noise gain in the resulting structure is derived and investigated. Moreover, the roundoff noise gain is compared with that deduced in a recent study of generalized direct-form II realization of 2-D SD digital filters.

I. INTRODUCTION

In the past decades, delta operator has widely been used in the realization of digital filters to improve finite-word-length (FWL) performance in systems with high sampling rate [1]-[5]. Li and Gevers have studied the roundoff noise gain for the optimal delta-operator direct-form II transposed filter has been presented in terms of the roundoff noise gain and coefficient order delta-operator direct-form II transposed structure has been constructed, and an expression for the roundoff noise gain with respect to free parameters subject to $l_2$-scaling constraints have been examined [13]. In this paper, we present a detailed roundoff noise analysis for generalized direct-form II structure of 2-D SD digital filters using a different approach from [13]. The roundoff noise gain is compared with that deduced in a recent study of generalized direct-form II space-state realization of 2-D SD digital filters in [13].

II. STRUCTURE OF 2-D DIGITAL FILTERS

Consider a 2-D stable SD digital filter of order $(m, n)$ described by

$$H(z_1, z_2) = \sum_{k=0}^{m} \sum_{l=0}^{n} c_{kl} z_1^{-k} z_2^{-l} \left( 1 + \sum_{k=1}^{n} a_k z_1^{-k} \right) \left( 1 + \sum_{l=1}^{n} b_l z_2^{-l} \right)$$  \hspace{1cm}(1)

where the denominator and numerator are assumed to be co-prime. Let $P_1$ and $P_2$ be $(m+1) \times (m+1)$ and $(n+1) \times (n+1)$ nonsingular matrices, respectively, defined by

$$\begin{bmatrix} q_1^0(z_1) & q_1^1(z_1) & \cdots & q_1^m(z_1) \\ q_2^0(z_2) & q_2^1(z_2) & \cdots & q_2^n(z_2) \end{bmatrix}^T = P_1 \begin{bmatrix} z_1^m & \cdots & z_1 \end{bmatrix}^T$$

$$\begin{bmatrix} q_1^0(z_1) & q_1^1(z_1) & \cdots & q_1^m(z_1) \\ q_2^0(z_2) & q_2^1(z_2) & \cdots & q_2^n(z_2) \end{bmatrix}^T = P_2 \begin{bmatrix} z_2^n & \cdots & z_2 \end{bmatrix}^T$$  \hspace{1cm}(2)

Next, scalars $\{\alpha_k | k = 1, 2, \cdots, m\}$, $\{\beta_l | l = 1, 2, \cdots, n\}$ and $\{\tau_{kl} | k = 0, 1, \cdots, m; l = 0, 1, \cdots, n\}$ are defined so that

$$P_1 = \begin{bmatrix} 1 & a_1 & a_2 & \cdots & a_m \end{bmatrix}$$

$$P_2 = \begin{bmatrix} 1 & b_1 & b_2 & \cdots & b_n \end{bmatrix}$$

$$P_{12} = \begin{bmatrix} r_{00} & \cdots & r_{0n} \\ \vdots & \ddots & \vdots \\ r_{m0} & \cdots & r_{mn} \end{bmatrix}$$

$$P = \begin{bmatrix} c_0 & \cdots & c_0 \\ \vdots & \ddots & \vdots \\ c_{00} & \cdots & c_{nn} \end{bmatrix}$$  \hspace{1cm}(3)

From (1)-(3), the transfer function in (1) can be expressed as

$$H(z_1, z_2) = \sum_{k=0}^{m} \sum_{l=0}^{n} r_{kl} q_1^k(z_1) q_2^l(z_2)$$

$$+ \sum_{k=0}^{m} \sum_{l=0}^{n} \alpha_k q_1^k(z_1) \sum_{l=0}^{n} \beta_l q_2^l(z_2)$$  \hspace{1cm}(4)
where scaling factors $\kappa_1$ and $\kappa_2$ are determined by requiring $\alpha_0 = 1$ and $\beta_0 = 1$. We now define

$$\rho_k^h(z_1) = \frac{z_1 - \gamma_k}{\Delta_k} \quad \text{for} \quad k = 1, 2, \cdots, m$$

$$\rho_l^v(z_2) = \frac{z_2 - \gamma_l}{\Delta_l} \quad \text{for} \quad l = 1, 2, \cdots, n$$

(5)

where \{\gamma_k\}, \{\gamma_l\}, \{\Delta_k > 0\} and \{\Delta_l > 0\} are four sets of constants [13] and polynomial operators are chosen as

$$q^h_k(z_1) = \rho_{k+1}^h(z_1)\rho_{k+2}^h(z_1)\cdots\rho_m^h(z_1), \quad k = 0, 1, \cdots, m-1$$

$$q^v_l(z_2) = \rho_{l+1}^v(z_2)\rho_{l+2}^v(z_2)\cdots\rho_n^v(z_2), \quad l = 0, 1, \cdots, n-1$$

and $q^h_m(z_1) = q^v_m(z_2) = 1$. Using (5), we can specify the corresponding transformation matrices $P_1$, $P_2$ and scalars $\kappa_1 = \Delta_1\Delta_2\cdots\Delta_m$, $\kappa_2 = \Delta_1\Delta_2\cdots\Delta_n$. Making use of (6), the transfer function in (4) can be written as [13]

$$H(z_1, z_2) = \frac{\sum_{k=0}^{m} \sum_{l=0}^{n} r_{kl}}{\sum_{k=0}^{m} \rho_k^h(z_1)^{-1} \sum_{l=0}^{n} \rho_l^v(z_2)^{-1}}$$

(7)

where $\alpha_0 = \beta_0 = 1$ and $\rho_0^h(z_1)^{-1} = \rho_0^v(z_2)^{-1} = 1$.

The implementations of $\rho_k^h(z_1)^{-1}$ and $\rho_l^v(z_2)^{-1}$ are depicted in Fig. 1. As an illustrative example, the structure of (7) for a 2-D filter with $(m, n) = (3, 3)$ is depicted in Fig. 2 where $u(i, j)$ is a scalar input and $y(i, j)$ is a scalar output.

From Figures 1 and 2, we deduce

$$y(i, j) = w_0(i, j) + r_{00}[u(i, j) - \sum_{l=1}^{n} \beta_l x_l^v(i, j)]$$

$$+ r_{01}[u(i, j) - \sum_{l=1}^{n} \beta_l x_l^v(i, j)]$$

(8)

$$w_k(i, j) = \rho_k^h(z_1)^{-1}w_{k+1}(i, j) + r_{kl}x_l^v(i, j)$$

$$+ r_{k0}[u(i, j) - \sum_{l=1}^{n} \beta_l x_l^v(i, j)] - \alpha_k y(i, j)$$

for $k = 1, 2, \cdots, m$ where $w_{m+1}(i, j) = 0$. In addition,

$$x_1(v, i, j) = \delta v_0(i, j) + \Delta v_1(i, j) u(i, j) - \sum_{l=1}^{n} \beta_l x_l^v(i, j)$$

(9)

$$l = 2, 3, \cdots, n.$$

We note that the model in (7) contains $2(m + n) + (m + 1)(n + 1)$ nontrivial parameters \{\kappa_i\}, \{\Delta_i\}, \{\gamma_i\} and \{r_{kl}\} plus $m + n$ free parameters \{\gamma_k\} and \{\gamma_l\}.

III. ROUNDOFF NOISE ANALYSIS

We begin by examining the roundoff noise caused by the term $\alpha_k y(i, j)$ for $1 \leq k \leq m$ at the output. Due to the product quantization, for the actual filter implemented by a FWL system, (8) can be written as

$$\hat{y}(i, j) = \hat{w}_0(i, j) + r_{00}[u(i, j) - \sum_{l=1}^{n} \hat{\beta}_l x_l^v(i, j)]$$

$$+ \sum_{l=1}^{n} r_{0l}x_l^v(i, j)$$

$$\hat{w}_s(i, j) = \rho_k^h(z_1)^{-1}(\hat{w}_{s+1}(i, j) + \sum_{l=1}^{n} r_{sl}x_l^v(i, j))$$

$$+ r_{s0}[u(i, j) - \sum_{l=1}^{n} \hat{\beta}_l x_l^v(i, j)] - \alpha_s \hat{y}(i, j) + \tau_s[i, j]$$

(10)

for $1 \leq s \leq m$ where $\hat{w}_{m+1}(i, j) = 0, \tau_s(i, j) = 0$ unless $s = k$, $\hat{y}(i, j)$ is the actual output, $\hat{w}_s(i, j)$ is the actual signal of $w_s(i, j)$ and $\tau_s(i, j) = Q[\alpha_k \hat{y}(i, j)] - \alpha_k \hat{y}(i, j)$ is the roundoff noise due to quantizer $Q[\cdot]$.

Subtracting (8) from (10) yields

$$\delta y(i, j) = \delta w_0(i, j)$$

$$\delta w_s(i, j) = \rho_k^h(z_1)^{-1}(\delta w_{s+1}(i, j) - \alpha_s \delta y(i, j) - \tau_s(i, j))$$

(11a)

where

$$\delta y(i, j) = \hat{y}(i, j) - y(i, j)$$

$$\delta w_s(i, j) = \hat{w}_s(i, j) - w_s(i, j)$$

(11b)

If a 1-D state-space model $(A_1, \tau_k, c_1)_m$ is realized using (5) from (11a), the transfer function from $-\tau_k(i, j)$ to $\delta y(i, j)$ is given by

$$H_{1k}(z_1) = c_1(z_1 I_m - A_1)^{-1}\tau_k$$

(12)
where \( \xi_k \) is the \( k \)th column of an identity matrix \( I_m \) and
\[
c_1 = \begin{bmatrix} \Delta_1 & 0 & \cdots & 0 \\ 0 & \Delta_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\alpha_m \Delta_1 & 0 & \cdots & 0 \end{bmatrix}
\]
\[
A_1 = \begin{bmatrix} -\alpha_1 \Delta_1 & \Delta_2 & \cdots & 0 \\ -\alpha_2 \Delta_1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\alpha_m \Delta_1 & 0 & \cdots & 0 \end{bmatrix} + \begin{bmatrix} \tau_1 & 0 & \cdots & 0 \\ 0 & \tau_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \tau_m \end{bmatrix}
\]
Based on above analysis, it is natural to define the roundoff noise gain in terms of \( H_{1k}(z_1) \) as
\[
J_1(\alpha_k) = \frac{E[\delta y(i,j)^2]}{E[\xi_k(i,j)^2]} = \frac{1}{2\pi j} \int_{|z_1|=1} H_{1k}^H(z_1) H_{1k}(z_1) \frac{dz_1}{z_1}
\]
where \( H_{1k}^H(z_1) \) denotes the conjugate transpose of \( H_{1k}(z_1) \). Substituting (12) into (13), it follows that for \( k = 1, 2, \cdots, m \)
\[
J_1(\alpha_k) = \mathbf{e}_k^T \left[ \sum_{i=0}^{\infty} (c_1 A_1)^T c_1 A_1 \right] \mathbf{e}_k = \mathbf{e}_k^T W^h \mathbf{e}_k
\]
where \( W^h \) is the horizontal observability Grammian which can be obtained by solving the Lyapunov equation \([9],[13]\)
\[
W^h = A_1^T W^h A_1 + c_1^T c_1
\]
Similarly, the roundoff noise gain produced by the coefficient \( r_{kl} \) for \( l = 0, 1, \cdots, n \) in the second equation of (8) can be expressed as
\[
J_2(r_{kl}) = \mathbf{e}_k^T W^h \mathbf{e}_k \quad \text{for} \quad k = 1, 2, \cdots, m
\]
With \( w_{k+1}(i,j) \) replaced by \( \Delta_{k+1} x_{k+1}^h(i,j) \) in the second equation of (8), the roundoff noise gain due to \( \Delta_{k+1} \) can be viewed as a function of \( r_{kl} \), i.e.,
\[
J_2(\Delta_{k+1}) = J_2(r_{kl}) = \mathbf{e}_k^T W^h \mathbf{e}_k \quad \text{for} \quad k = 1, 2, \cdots, m
\]
As shown in Fig. 1(a), parameter \( \tau_k \) induces a multiplication \( \tau_k x_k^h(i,j) \) which produces no roundoff noise if \( \tau_k = 0, \pm 1 \). Let \( \psi(\tau_k) \) denote the roundoff noise due to \( \tau_k \) where \( \psi(\tau_k) = 1 \) for all \( \tau_k \) except \( \tau_k = 0, \pm 1 \) for which \( \psi(\tau_k) = 0 \), and \( \delta y(i,j) \) is the corresponding output deviation. Then the transfer function from \( \psi(\tau_k) e_k^T(i,j) \) to \( \delta y(i,j) \) becomes \( H_{1k}(z_1) \) in (12). Actually, this roundoff noise can be viewed as that generated by the term \( r_{kl} x_{k+1}^h(i,j) \). Hence
\[
J_3(\tau_k) = \psi(\tau_k) J_2(r_{kl}) = \psi(\tau_k) \mathbf{e}_k^T W^h \mathbf{e}_k
\]
for \( k = 1, 2, \cdots, m \) where
\[
\psi(\gamma) = \begin{cases} 1 \text{ for } \gamma \neq 0, \pm 1 \\ 0 \text{ for } \gamma = 0, \pm 1 \end{cases}
\]
Concerning the roundoff noise due to coefficient \( r_{0l} \) for \( l = 0, 1, \cdots, n \) in the first equation of (8), the first equation in (11a) needs to be changed to
\[
\delta y(i,j) = \delta w_1(i,j) + \varepsilon_0(i,j)
\]
When a 1-D state-space model \((A_1, \alpha, -c_1, 1)_m\) is realized from (11a) whose first equation was replaced by (18) and \( \varepsilon_0(i,j) = 0 \) in the second equation, the transfer function \( H_{10}(z_1) \) from \( \varepsilon_0(i,j) \) to \( \delta y(i,j) \) is given by
\[
H_{10}(z_1) = -c_1(z_1 I_m - A_1)^{-1} \alpha + 1
\]
where \( \alpha = [\alpha_1, \alpha_2, \cdots, \alpha_m]^T \). Hence the roundoff noise gain caused by the coefficient \( r_{0l} \) is found to be
\[
J_4(r_{0l}) = \alpha^T W^h \alpha + 1 \quad \text{for} \quad l = 0, 1, \cdots, n
\]
Supposing that \( w_1(i,j) \) in the first equation of (8) is replaced by \( \Delta_1 x_{1}^h(i,j) \), the roundoff noise gain produced by \( \Delta_1 \) is identical to that by \( r_{0l} \), which leads to
\[
J_4(\Delta_1) = J_4(r_{0l}) = \alpha^T W^h \alpha + 1
\]
We now examine the roundoff noise caused by the term \( \beta p x_{p}^h(i,j) \) for \( 1 \leq p \leq n \) at the output. Due to the product quantization, for the actual filter implemented by a FWL system, (8) and (9) can be written as
\[
\tilde{y}(i,j) = \tilde{w}_1(i,j) + \sum_{l=1}^{n} r_{0l} \tilde{x}_{p}^h(i,j)
\]
\[
\tilde{w}_k(i,j) = \rho_k^l(z_1) - \delta \tilde{x}_{p}^h(i,j) - \varepsilon_p(i,j)
\]
\[
\tilde{x}_{p}^h(i,j) = \sum_{l=1}^{n} (\beta_l \tilde{x}_{p}^h(i,j) - \varepsilon_p(i,j)) - \alpha_k \tilde{y}(i,j)
\]
For \( k = 1, 2, \cdots, m \) where \( \tilde{w}_{m+1}(i,j) = 0 \), and
\[
\tilde{x}_{p}^h(i,j) = \tilde{x}_{p}^h(i,j) + \Delta_1 [u(i,j) - \sum_{l=1}^{n} \beta_l \tilde{x}_{p}^h(i,j) - \varepsilon_p(i,j)]
\]
\[
\tilde{x}_{p}^h(i,j) = \Delta_1 [u(i,j) - \sum_{l=1}^{n} \beta_l \tilde{x}_{1}^h(i,j) - \varepsilon_p(i,j)]
\]
for \( l = 2, 3, \cdots, n \) respectively, where \( \tilde{x}_{p}^h(i,j) \) denotes the actual signal of \( x_{p}^h(i,j) \) and \( \varepsilon_p(i,j) = Q[\beta_p \tilde{x}_{p}^h(i,j)] - \beta_p x_{p}^h(i,j) \) is the roundoff noise due to quantizer \( Q[\cdot] \). Subtracting (8) from (22) yields
\[
\delta y(i,j) = \delta w_1(i,j) + \sum_{l=1}^{n} r_{0l} \delta x_{p}^h(i,j)
\]
\[
\delta w_k(i,j) = \rho_k^l(z_1) - \delta \tilde{x}_{p}^h(i,j) - \varepsilon_p(i,j)
\]
\[
\delta \tilde{w}_k(i,j) = \sum_{l=1}^{n} (\beta_l \delta \tilde{x}_{p}^h(i,j) - \delta \varepsilon_p(i,j)) - \alpha_k \delta \tilde{y}(i,j)
\]
for \( k = 1, 2, \cdots, m \) where \( \delta w_{m+1}(i,j) = 0 \) and \( \delta \tilde{x}_{p}^h(i,j) - \varepsilon_p(i,j)] = \tilde{x}_{p}^h(i,j) - \varepsilon_p(i,j)] \) for \( l = 1, 2, \cdots, n \). Using (5) and (11b),
we can write (24) as
\[
\delta y(i, j) = \Delta_1 \delta x_1^j(i, j) + \sum_{l=1}^{n} (r_{0l} - r_{00} \beta_l) \delta x_l^j(i, j)
+ r_{00} - \hat{\delta}_p(i, j)
\]
\[
\delta x_k^b(i, j) = -\alpha_k \Delta_1 \delta x_1^j(i, j) + \sum_{l=1}^{n} \delta x_{l+1}^b(i, j)
+ \sum_{l=1}^{n} (r_{kl} - r_{k0} \beta_l + r_{00} \alpha_k \beta_l) \delta x_l^j(i, j)
+ \gamma_k \hat{\delta}_k(i, j) + (r_{k0} - r_{00} \alpha_k) \cdot \hat{\delta}_p(i, j)
\]
for \(k = 1, 2, \ldots, m\) where \(x_{p+1}^m(i, j) = 0\). By subtracting (9) from (23), we obtain
\[
\delta x_1^j(i, j + 1) = \hat{\gamma}_1 \delta x_1^j(i, j) + \Delta_1 \left( -\hat{\delta}_p(i, j) - \sum_{l=1}^{n} \beta_l \delta x_l^j(i, j) \right)
\]
\[
\delta x_k^j(i, j + 1) = \Delta_1 \delta x_{k-1}^j(i, j) + \gamma_k \delta x_1^j(i, j)
\]
(26)
At this point, we consider a 2-D local state-space realization of (25) and (26). The transfer function from \(-\hat{\delta}_p(i, j)\) to \(\delta y(i, j)\) is then found to be [13]
\[
H(z_1, z_2) = d + c_1(z_1 \mathbf{I}_m - A_1)^{-1} b_1
+ [c_1(z_1 \mathbf{I}_m - A_1)^{-1} A_2 + c_2] (z_2 \mathbf{I}_n - A_4)^{-1} b_2
\]
(27)
where
\[
A_2 = \begin{bmatrix}
    r_{11} & r_{12} & \cdots & r_{1n} \\
    r_{21} & r_{22} & \cdots & r_{2n} \\
    \vdots & \vdots & \ddots & \vdots \\
    r_{m1} & r_{m2} & \cdots & r_{mn}
\end{bmatrix}, \quad
A_4 = \begin{bmatrix}
    \alpha_1 & \alpha_2 & \cdots & \alpha_n \\
    \alpha_1 & \alpha_2 & \cdots & \alpha_n \\
    \vdots & \vdots & \ddots & \vdots \\
    \alpha_1 & \alpha_2 & \cdots & \alpha_n
\end{bmatrix}
\]
\[
b_1 = \begin{bmatrix}
    r_{11} & r_{12} & \cdots & r_{1n} \\
    r_{21} & r_{22} & \cdots & r_{2n} \\
    \vdots & \vdots & \ddots & \vdots \\
    r_{m1} & r_{m2} & \cdots & r_{mn}
\end{bmatrix}^T - r_{00} \begin{bmatrix}
    \alpha_1 & \alpha_2 & \cdots & \alpha_n
\end{bmatrix}^T
\]
\[
b_2 = \begin{bmatrix}
    \Delta_1 & 0 & \cdots & 0 \\
    \Delta_2 & 0 & \cdots & 0 \\
    \vdots & \vdots & \ddots & \vdots \\
    \Delta_n & 0 & \cdots & 0
\end{bmatrix}^T
\]
\[
c_2 = \begin{bmatrix}
    r_{11} & r_{12} & \cdots & r_{1n} \\
    r_{21} & r_{22} & \cdots & r_{2n} \\
    \vdots & \vdots & \ddots & \vdots \\
    r_{m1} & r_{m2} & \cdots & r_{mn}
\end{bmatrix} - r_{00} \begin{bmatrix}
    \beta_1 & \beta_2 & \cdots & \beta_n
\end{bmatrix}
\]
Based on this, the roundoff noise gain defined by \(J_5(\beta_p) = E[\delta y(i, j)^2] / E[\hat{\delta}_p(i, j)^2]\) can be expressed as
\[
J_5(\beta_p) = \frac{1}{(2\pi)^2} \int_{|z_1|=1} \int_{|z_2|=1} |H(z_1, z_2)|^2 |dz_1| dz_2
\]
(28)
Substituting (27) into (28) yields
\[
J_5(\beta_p) = \hat{\theta}_1^T \mathbf{W}^v b_1 + \hat{\theta}_2^T \mathbf{W}^v b_2 + d^2
= b_1^T \mathbf{W}^v b_1 + \hat{\Delta}_1^2 \hat{\delta}_1^T \mathbf{W}^v \hat{\delta}_1 + r_{00}^2
\]
(29)
where \(\mathbf{W}^v\) is the vertical observability Gramian which can be obtained by solving the Lyapunov equation [9], [13]
\[
\mathbf{W}^v = A_4^T \mathbf{W}^v A_4 + A_2^T \mathbf{W}^h A_2 + \epsilon_1^T \epsilon_2
\]
Due to the product quantization caused by \(\Delta_1\) in the first equation of (9), the actual filter implemented by a FWL system can be written as
\[
\hat{y}(i, j) = \hat{w}_1(i, j) + \sum_{l=1}^{n} r_{0l} \hat{x}_l^j(i, j)
+ r_{00} \left[ u(i, j) - \sum_{l=1}^{n} \beta_l \hat{x}_l^j(i, j) \right]
\]
(30)
\[
\hat{w}_k(i, j) = \rho_k^0 (z_1)^{-1} \left[ \hat{w}_{k+1}(i, j) + \sum_{l=1}^{n} r_{kl} \hat{x}_l^j(i, j) \right]
+ r_{k0} \left[ u(i, j) - \sum_{l=1}^{n} \beta_l \hat{x}_l^j(i, j) \right] - \alpha_k \hat{y}(i, j)
\]
(31)
for \(k = 1, 2, \ldots, m\) where \(\hat{w}_{m+1}(i, j) = 0\), and
\[
\hat{x}_1^j(i, j + 1) = \gamma_1 \hat{x}_1^j(i, j) + \Delta_1 \left[ u(i, j) - \sum_{l=1}^{n} \beta_l \hat{x}_l^j(i, j) \right]
+ \epsilon_1(i, j)
\]
\[
\hat{x}_k^j(i, j + 1) = \Delta_1 \hat{x}_{k-1}^j(i, j) + \gamma_k \hat{x}_1^j(i, j)
\]
(32)
for \(k = 1, 2, \ldots, m\) where \(\hat{w}_{m+1}(i, j) = 0\). Subtracting (8) from (30) yields
\[
\delta y(i, j) = \delta w_1(i, j) + \sum_{l=1}^{n} r_{0l} \delta x_l^j(i, j)
- r_{00} \sum_{l=1}^{n} \beta_l \delta x_l^j(i, j)
\]
\[
\delta w_k(i, j) = \rho_k^0 (z_1)^{-1} \left[ \delta w_{k+1}(i, j) + \sum_{l=1}^{n} r_{kl} \delta x_l^j(i, j) \right]
- r_{k0} \sum_{l=1}^{n} \beta_l \delta x_l^j(i, j) - \alpha_k \delta y(i, j)
\]
(32)
for \(k = 1, 2, \ldots, m\) where \(\delta w_{m+1}(i, j) = 0\). Subtracting (9) from (31), we obtain
\[
\delta x_1^j(i, j + 1) = \hat{\gamma}_1 \delta x_1^j(i, j) - \Delta_1 \sum_{l=1}^{n} \beta_l \delta x_l^j(i, j) + \epsilon_1(i, j)
\]
\[
\delta x_k^j(i, j + 1) = \Delta_1 \delta x_{k-1}^j(i, j) + \gamma_k \delta x_1^j(i, j)
\]
(33)
for \(l = 2, 3, \ldots, n\), respectively. Using (5) and (11b), we consider a 2-D local state-space realization of (32) and (33). The transfer function from \(\epsilon_1(i, j)\) to \(\delta y(i, j)\) is then found to be
\[
H_2(z_1, z_2) = \left[ c_1(z_1 \mathbf{I}_m - A_1)^{-1} A_2 + c_2 \right] (z_2 \mathbf{I}_n - A_4)^{-1} \hat{\delta}_1
\]
(34)
where \(\hat{\delta}_1\) denotes the \(l\)th column of an identity matrix \(\mathbf{I}_n\).
Hence the roundoff noise gain due to \(\Delta_1\) becomes
\[
J_6(\Delta_1) = \hat{\theta}_1^T \mathbf{W}^v \hat{\delta}_1
\]
(35)
Similarly, the roundoff noise gain due to \(\Delta_l\) for \(l = 2, 3, \ldots, n\) is given by
\[
J_6(\Delta_l) = \hat{\theta}_1^T \mathbf{W}^v \hat{\delta}_l \quad \text{for} \quad l = 2, 3, \ldots, n
\]
(36)
and the roundoff noise gain due to $\gamma_l$ for $l = 1, 2, \cdots, n$ can be written as

$$J_0(\gamma_l) = \psi(\gamma_l)\hat{e}_l^T W^r \hat{e}_l$$

for $l = 1, 2, \cdots, n$ \hspace{1cm} (37)

Based on the above analysis, the total roundoff noise gain of the filter structure in Fig. 2 can be defined as

$$J_p = \sum_{k=1}^{m-1} J_1(\alpha_k) + \sum_{l=0}^{n-1} J_2(\delta_0^x) + J_3(\gamma_k)$$

$$+ \sum_{k=1}^{m-1} J_2(\Delta_k + 1) + J_4(\Delta_1) + J_4(\Delta_0)$$

$$+ \sum_{l=1}^{n} \left[ J_4(\delta_0^x) + J_5(\beta_l) + J_6(\Delta_l) + J_6(\gamma_l) \right]$$

which can be written as

$$J_p = (n+3)\text{tr}[W^h] - \text{tr}[W^r] + \left[ \Psi W^h \right]$$

$$+ (n+2)(\alpha^T W^h \alpha + 1) + \left[ \Psi W^v \right] + \left[ \Psi W^v \right]$$

$$+ n\left[ b_i^T W^h b_i + (\Delta^2 - 1) \right]$$

where

$$\Psi = \text{diag}\{\psi(\gamma_1), \psi(\gamma_2), \cdots, \psi(\gamma_m)\}$$

$$\Psi = \text{diag}\{\psi(\gamma_1), \psi(\gamma_2), \cdots, \psi(\gamma_n)\}$$

Remark 1: At this point, it is of interest to note that the roundoff noise gain for state-space realization of the filter structure in (7) can be evaluated as [13]

$$J_{SP} = (n+3)\text{tr}[W^h] - \text{tr}[W^r]$$

$$+ \sum_{k=2}^{m} \psi(\gamma_k)\hat{e}_l^T W^h \hat{e}_l + n\hat{e}_l^T W^v \hat{e}_l$$

$$+ \text{tr}[W^v] + \sum_{l=2}^{n} \psi(\gamma_l)\hat{e}_l^T W^v \hat{e}_l + n + 2$$

From (39) and (40), it follows that

$$J_p - J_{SP} = \psi(\gamma_1)\hat{e}_l^T W^h \hat{e}_l + \psi(\gamma_1)\hat{e}_l^T W^v \hat{e}_l + (n+2)\alpha^T W^h \alpha$$

$$+ n\left[ b_i^T W^h b_i + (\Delta^2 - 1)\right] \hat{e}_l^T W^v \hat{e}_l + n + 2$$

(41)

It is noted that the difference $J_p - J_{SP}$ evaluated in (41) is due to the different number of parameters (coefficients) between the filter structure in (7) and its state-space realization.

IV. A NUMERICAL EXAMPLE

Consider a 2-D stable SD digital filter of order $(m, n) = (3, 3)$ in (1) with

$$[a_1 \ a_2 \ a_3] = [-2.173645 \ 1.386929 \ -0.599655]$$

$$[b_1 \ b_2 \ b_3] = [-2.280029 \ 1.887939 \ -0.564961]$$

$$[c_{kl}] = \begin{bmatrix}
0.019421 & -0.027724 & 0.011468 & 0.000087 \\
0.048393 & 0.017545 & -0.050267 & 0.033061 \\
-0.004328 & -0.008847 & 0.096260 & -0.083801 \\
0.000138 & 0.007979 & -0.052927 & 0.062007
\end{bmatrix}$$

The numerical results obtained by applying the technique in [13] were summarized in application with two cases of

$$\gamma_z = [0, 0, \cdots, 0]^T$$

and

$$\gamma_z = [1, 1, \cdots, 1]^T$$

in Table I, where

$$\gamma(J_{p_{\text{opt}}}) = [1.000 \ 0.625 \ 0.750 \ 0.000 \ 1.000 \ 0.625]$$

$$\gamma(J_{SP_{\text{opt}}}) = [0.250 \ 0.625 \ 0.750 \ -0.750 \ 0.750 \ 0.750]$$

V. CONCLUSION

An expression of the roundoff noise gain for the resulting structure has been derived and investigated. Moreover, the roundoff noise gain has been compared with that deduced in a recent study of generalized direct-form II state-space realization of 2-D SD digital filters. In a numerical example, the roundoff noise gains have been minimized with respect to the free parameters subject to $l_2$-scaling constraints through exhaustive search in a finite element space [13].

REFERENCES


