Generalized Formulation of 2-D Filter Structures without Global Broadcast for VLSI Implementation

I-Hung Khoo & Hari C. Reddy*
Dept. of Electrical Engineering
California State University Long Beach, U.S.A.
*and College of Computer Science, National Chiao-Tung University, Hsinchu, Taiwan, R.O.C

Lan-Da Van & Chin-Teng Lin
Department of Computer Science
National Chiao-Tung University
Hsinchu, Taiwan, R.O.C

***Abstract*— A generalized formulation is developed that allows the derivation of various new 2-D VLSI filter structures, without global broadcast, using different filter sub-blocks and their interconnections (frameworks). With this formulation, lattice-type and direct-form structures realizing general 2-D IIR and FIR transfer functions, IIR transfer functions with separable denominators, and transfer functions with quadrantal magnitude symmetry are easily obtained. The separable denominator and quadrantal symmetry structures have the advantage of reduced number of multipliers.**

I. INTRODUCTION

Two-dimensional (2-D) digital filters find applications in many digital signal processing areas such as image processing and seismic data processing. Although 2-D digital filters can be simulated on a general purpose computer, for applications involving high data rate, such as real time image processing, dedicated computing structures are needed in order to meet the high throughput demands. Networks using structures such as systolic arrays are popular candidates for VLSI ASIC implementation due to the regularity and modularity of the processing elements involved. Having local data communication (without global broadcast of signals) among the elements is important in such VLSI designs. In [1,2], 2-D systolic digital filter architectures were presented which eliminated the global broadcast of the input and output signals in previous architectures [3,4]. In addition, in [2], new structures realizing transfer functions with separable denominators and having diagonal magnitude symmetry were presented.

In this paper, a generalized formulation is developed that allows the derivation of new 2-D VLSI filter structures, without global broadcast, using different filter sub-blocks and different interconnections (frameworks). A general digital two-pair approach is used to describe the sub-blocks which consist of direct-form or lattice-type FIR filter in one of the frequency variables, as discussed in Section III. (Note that other types can be used too, beside lattice and direct-forms). Then, by applying the sub-blocks in various frameworks, 2-D structures realizing different transfer functions are easily obtained. (The structures presented in [1,2] are among a few of the many possible structures that can be derived using this general formulation). Section IV discusses the filter frameworks for realizing general IIR and FIR transfer functions. Section V presents the frameworks for IIR transfer functions with separable denominators, where the structures exhibit the denominator separability as a filter structural property, and have important symmetry applications. Then, in Section VI, the filter frameworks for realizing transfer functions with quadrantal magnitude symmetry are presented. Following this, the multiplier savings for the separable denominator and quadrantal symmetry structures are discussed. Finally, the roundoff noise and multiplier sensitivity are analyzed for the representative structures.

II. PRELIMINARIES

A general 2-D IIR transfer function can be represented as in (1), where $b_{00}=0$, N_1xN_2 is the order of the filter, and X and Y are respectively the input and output of the filter. The equation can also represent an FIR transfer function if we set $b_{ij}=0$ for all i and j.

$$H(z_1,z_2)=\frac{Y(z_1,z_2)}{X(z_1,z_2)}=\frac{\sum_{i=0}^{N_1}\sum_{j=0}^{N_2}a_{ij}z_1^{-i}z_2^{-j}}{1-\sum_{i=0}^{N_1}\sum_{j=0}^{N_2}b_{ij}z_1^{-i}z_2^{-j}} \tag{1}$$

The transfer function in (1) can also be expressed as:

$$H(z_1,z_2)=\frac{\sum_{i=0}^{N_1}F_i\left(z_2^{-1}\right)\cdot z_1^{-i}}{1-\sum_{i=0}^{N_1}G_i\left(z_2^{-1}\right)\cdot z_1^{-i}} \tag{2}$$

where $F_i\left(z_2^{-1}\right)=\sum_{j=0}^{N_2}a_{ij}z_2^{-j}$ and $G_i\left(z_2^{-1}\right)=\sum_{j=0}^{N_2}b_{ij}z_2^{-j}$ are 1-D FIR functions in z_2 variable only. These 1-D functions can be realized by the two-pair sub-blocks in Section III. These sub-blocks are then used in the filter frameworks in Section IV to realize the overall 2-D transfer function in (2).

In our discussion, we assume that the filter is used to process a square image of size MxM and the pixel values in the image are fed to the filter in raster-scan mode, i.e. the input sequence is $x(0,0)$, $x(0,1)$, …, $x(0,M-1)$, $x(1,0)$, $x(1,1)$, … etc. We can then replace z_2^{-1} by a single delay register, z^{-1}, and z_1^{-1} by a shift register of length M, z^{-M}, provided $M>N_2$. Without loss of generality, we will assume $N_1=N_2=N$ in discussing the filters.

III. FILTER SUB-BLOCKS

The filter sub-blocks are formulated as general digital two-pair networks and realize 1-D FIR functions in z_2. Here we assume $z^{-1}=z_2^{-1}$.

Sub-block #1, shown in Fig. 1, has 2 inputs and 1 output. It is direct form, i.e. the multiplier values are the same as the polynomial coefficients. It realizes the following two FIR functions.

$$F_i\left(z^{-1}\right)=\left.\frac{Y_i}{X_i}\right|_{W_i=0}=\sum_{j=0}^{N}a_{ij}z^{-j}\quad,\quad G_i\left(z^{-1}\right)=\left.\frac{Y_i}{W_i}\right|_{X_i=0}=\sum_{j=0}^{N}b_{ij}z^{-j} \tag{3}$$

Note that the special arrangement of the delays is to eliminate global broadcast of the signals, X_i and W_i, and also

978-1-4244-7773-9/10/$26.00

to control the critical period. The critical period is the time required for the signal through the slowest (critical) path of the structure and determines the highest possible clock speed of the structure.

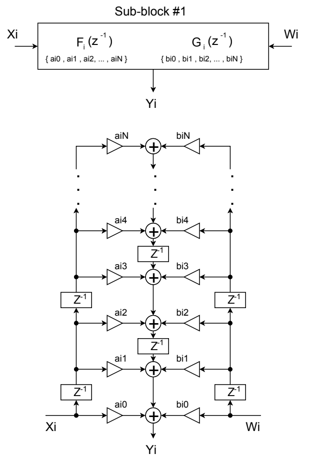

Figure 1. Sub-block #1 (2-inputs-1-outputs, direct-form)

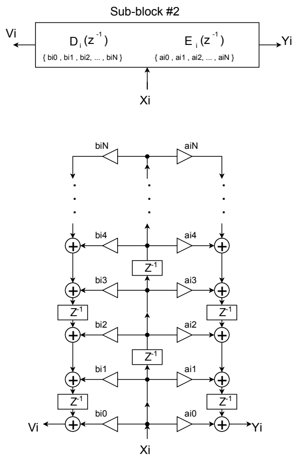

Figure 2. Sub-block #2 (1-input-2-outputs, direct-form)

Sub-block #2 is shown in Fig.2. It has 1 input and 2 outputs and realizes the following two FIR functions. It is also direct form.

$$D_i\left(z^{-1}\right)=\frac{V_i}{X_i}=\sum_{j=0}^{N} b_{ij} z^{-j} \quad , \quad E_i\left(z^{-1}\right)=\frac{Y_i}{X_i}=\sum_{j=0}^{N} a_{ij} z^{-j} \tag{4}$$

The alternate lattice form for sub-block#2 is shown in Fig.3. Unlike a regular lattice, this structure has different multipliers in the top and bottom branches. A special extraction algorithm has been developed to determine the multiplier values, ka_{ij} and kb_{ij}, from the function coefficients a_{ij} and b_{ij}. Because of lack of space, it is not stated here. Like a regular lattice, the structure cannot realize functions where the constant term is zero, i.e. a_{i0}=0 or b_{i0}=0. Also note the extra delays (circled in Fig. 3) added to control the critical period. This will result in latency in the form of extra z^{-N} factors in (4). However, the factors can be cancelled in the final reconfiguration to be discussed in Section IV so that the overall 2-D transfer function will not have any latency. The lattice-form in Fig.3 has the same number of multipliers and adders as the direct-form in Fig.2. It does, however, require more delay elements.

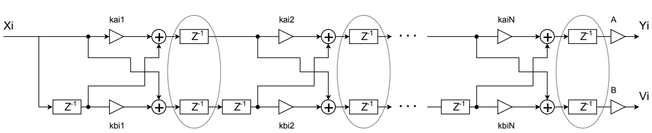

Figure 3. Sub-block #2 (1-input-2-outputs, lattice-form)

The direct-form version of sub-block #3 is shown in Fig. 4. The alternating delay arrangement is to eliminate global broadcast of the signal X_i and to control the critical period. The sub-block has 1 input and 1 output and realizes the FIR function in (5). Note that ρ_{ij} can represent either the numerator or denominator coefficient a_{ij} or b_{ij}.

$$C_{\rho i}\left(z^{-1}\right)=\frac{Y_i}{X_i}=\sum_{j=0}^{N} \rho_{ij} z^{-j} \tag{5}$$

The lattice version of sub-block #3 is shown in Fig.5. This is based on the one-multiplier lattice in [6]. The multiplier values, k_{ij}, can be determined from the function coefficients a_{ij}, using the regular lattice extraction plus appropriate scaling of the results [6]. Note that extra delays (in circle) are added to control the critical period, which will result in latency in the form of z^{-N} factor in (5). The latency can be removed in the filter framework to be discussed in Section IV. One limitation of the structure is that it cannot realize functions where ρ_{i0}=0. Also, compared to direct-form, the lattice version requires more delays and adders, but the number of multipliers is the same.

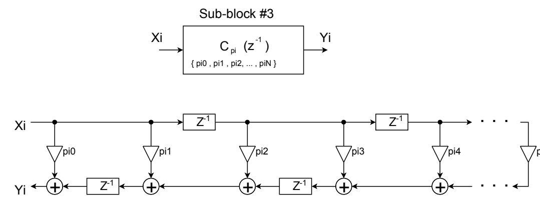

Figure 4. Sub-block #3 (1-input-1-output, direct-form)

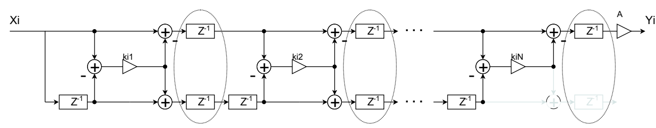

Figure 5. Sub-block #3 (1-input-1-output, lattice-form)

IV. FILTER FRAMEWORKS FOR REALIZING GENERAL TRANSFER FUNCTION

The sub-blocks discussed in the previous section are used in the filter frameworks in this section to realize the general 2-D transfer function in (2).

Filter framework A is shown in Fig. 6. It uses only filter sub-block #1. Notice that the shift registers (SR) are of length *M-1* due to the additional delays added at the input and output branches to eliminate the global broadcast. It can be verified using Mason's gain formula that the structure, with $z^{-1}=z_2^{-1}$ and $\mathrm{SR}=z_1^{-1}z_2$, possesses the transfer function in (2).

Filter framework B is shown in Fig. 7. It utilizes sub-block#2 and realizes the 2-D transfer function in (2) with the notation change from F_i and G_i to E_i and D_i respectively, which highlights the difference in sub-blocks. 2-D filter framework B is the transpose of 2-D filter framework A.

As discussed in Section III, there are two versions of sub-block #2 – direct-form (Fig. 2) and lattice-form (Fig.3). They can be used in any combination in the framework. The only restriction is that the bottom sub-block has to be direct-form. The reason is that the lattice-form cannot realize a function where the constant term is zero, as is needed for function D_0. Also, if the lattice-forms are used (which introduce a latency of z_2^{-N}), the length of the SR can be adjusted to compensate for the latency so that the overall 2-D transfer function will not have any latency. For instance, if the bottom sub-block is direct-form while the rest are lattice, then the bottom SR will need to be of length M-N-1 (realizing $z_1^{-1}z_2^{N+1}$) rather than M-1 (realizing $z_1^{-1}z_2$).

Filter framework C is shown in Fig. 8. It uses only sub-block #3, either the direct-form of Fig. 4 or the lattice-form of

Fig.5. It realizes the 2-D transfer function in (2) with the notation change from F_i to C_{ai} and from G_i to C_{bi} . Note that the direct and lattice-form sub-blocks can be used in any combination in the framework. The only restriction is that the C_{b0} sub-block needs to be direct-form as the lattice-type cannot again realize a function with a zero constant term. Once again, if the lattice-forms are used, appropriate adjustments can be made to the SR to avoid the latency.

There are two additional 2-D filter configurations. They are not shown due to the lack of space. Filter framework D is the transpose of framework C. Filter framework E is the same as framework C but without the C_{bi} sub-blocks, so it realizes 2-D FIR transfer functions.

V. FILTER FRAMEWORKS FOR REALIZING TRANSFER FUNCTIONS WITH SEPARABLE DENOMINATOR

By mixing the sub-blocks in specific ways, filter frameworks realizing transfer functions with separable denominator of the form in (6) can be obtained. The idea is to form two non-touching loops in different variables as per Mason's gain formula.

$$H(z_1,z_2)=\frac{Y(z_1,z_2)}{X(z_1,z_2)}=\frac{\sum_{i=0}^{N_1}\sum_{j=0}^{N_2}a_{ij}z_1^{-i}z_2^{-j}}{\left(1-\sum_{i=1}^{N_1}b_{i0}z_1^{-i}\right)\cdot\left(1-\sum_{j=1}^{N_2}b_{0j}z_2^{-j}\right)} \tag{6}$$

The significant feature of these structures is that they exhibit the denominator separability as a filter structural property, independent of the choice of multiplier values. The separable denominator transfer function has several advantages over the general one in (1). Firstly, the stability can be checked by simply solving for the poles of the two 1-D polynomials, and any unstable pole is easy to stabilize. Secondly, the separable denominator requires fewer multipliers to realize. Thirdly, the separable denominator is required in realizing stable magnitude responses possessing various symmetries (except for the diagonal symmetry) [5].

Filter framework F is shown in Fig. 9. It uses sub-block #2 at the bottom while the rest are sub-block #1. (Note that the lattice-form cannot be used here). It realizes the transfer function in (7), with G_i's being constants.

$$\frac{Y}{X}=\frac{E_0\left(z_2^{-1}\right)+\sum_{i=1}^{N}F_i\left(z_2^{-1}\right)\cdot z_1^{-i}}{\left[1-D_0\left(z_2^{-1}\right)\right]\cdot\left[1-\sum_{i=1}^{N}G_i\cdot z_1^{-i}\right]} \tag{7}$$

By taking the transpose of filter framework F, framework G can be obtained. (It is again not shown due to space limitation.) It uses sub-block #1 at the bottom of the framework while the rest are sub-block#2. It realizes the transfer function in (7) but with *E* swapping with *F,* and *D* swapping with *G*.

Filter framework H is shown in Fig 10. It uses only sub-block#3. It realizes the transfer function in (7) but with C_a replacing *E* and *F*, and C_b replacing *D* and *G*. Note that the bottom two sub-blocks have reversed input-output compared to the rest of the sub-blocks. Finally, filter framework I is the transpose of framework H.

VI. QUADRANTAL SYMMETRY FILTER FRAMEWORKS

The presence of symmetry in the 2-D frequency response induces certain relationship among the filter coefficients which can result in fewer multipliers in the implementation. There are many types of symmetries [5]. Here, we will focus on one of the symmetries, namely, quadrantal symmetry.

A 2-D magnitude response possesses quadrantal symmetry if $\left|H(z_1,z_2)\right|=\left|H(z_1^{-1},z_2)\right|$ with $z_1=e^{j\theta_1}$ and $z_2=e^{j\theta_2}$, $\forall\ (\theta_1,\theta_2)$ [5]. Assuming the separable denominator transfer function in (6) is adopted, it will have quadrantal symmetry if $a_{ij}=a_{(N-i)j}$ for all *i, j*. Applying this constraint to the separable denominator framework F and H, the new quadrantal symmetry configurations can be obtained as shown in Fig 11 and 12. Note that the changes are highlighted in red, and although the structures shown are 2x2, they can easily be generalized to higher orders. In a similar manner, the symmetry constraint can be applied to filter frameworks G and I, as well as FIR framework E, to yield new structures with the symmetry. They are omitted here due to lack of space.

The quadrantal symmetry structure has the lowest number of multipliers compared to all the structures discussed so far. Filter frameworks A through D realizing regular 2-D IIR transfer function require $2(N+1)^2-1$ multipliers. The separable denominator frameworks, F through I, require fewer multipliers: $(N+1)^2+2N$. The quadrantal symmetry structures require the least number of multipliers: only $(N+1)^2/2+2N$ and $(N/2+1)\cdot(N+1)+2N$ when N is odd and even respectively. For 2-D FIR structures, the number of required multiplier is reduced from $(N+1)^2$ to $(N+1)^2/2$ (for N odd) or $(N/2+1)\cdot(N+1)$ (for N even).

VII. ROUNDOFF NOISE AND MULTIPLIER SENSITIVITY

Because of the numerous possible structures, the noise and sensitivity are studied only for the representative structures.

The signal roundoff noise is studied for filter frameworks A through D using the direct-form sub-blocks. Assuming a 2-D lowpass filter, the roundoff noise is plotted against the filter cutoff frequency as shown in Fig 13 for different filter orders. It can be seen that filter frameworks B and D have the lowest roundoff noise and their advantage increases with the filter order.

The multiplier sensitivity is studied next. Comparison is made between direct and lattice version of sub-block #3 as applied to either filter framework C or D. (For the lattice version, the bottom-row sub-blocks are direct-form in order to realize the zero constant term). The result is shown in Fig. 14 assuming a lowpass 2-D filter. It can be seen that, for most of the passband frequencies, the lattice-type possesses better multiplier sensitivity.

VIII. CONCLUSION

A generalized formulation is developed that allows the derivation of several new 2-D VLSI filter structures, without global broadcast, using different filter sub-blocks and their interconnection frameworks. Using this formulation, structures realizing general 2-D IIR and FIR transfer functions, IIR transfer functions with separable denominators, and transfer functions with quadrantal magnitude symmetry were obtained. The separable denominator and quadrantal symmetry structures have the advantage of reduced number of multipliers. The roundoff noise and multiplier sensitivity were studied for some of the representative structures.

References

[1] L. D. Van, "A new 2-D systolic digital filter architecture without global broadcast", *IEEE Trans. VLSI syst*., v10, no. 4, pp477-486, Aug. 2002.

[2] I. H. Khoo, H. C. Reddy, L .D. Van, C. T. Lin, "2-D Digital Filter Architectures without Global Broadcast and Some Symmetry Applications", *Proceedings of the 2009 ISCAS*, pp 952-959, May 2009.

[3] M. A. Sid-Ahmed, "A systolic realization for 2-D digital filters," *IEEE Trans. Acoust., Speech, Signal Processing*, v37, pp560-565, Apr. 1989.

[4] S. Sunder, F. El-Guibaly, A. Antoniou, "Systolic implementations of two-dimensional recursive digital filters", *Proc. IEEE Int. Symp. Circuits Syst.*, pp1034-1037, May 1990.

[5] H. Reddy, I. Khoo, P. Rajan, "2-D symmetry: theory and filter design applications", *IEEE Circuits and Systems Magazine*, v3, pp4-33, 2003.

[6] J. Makhoul, "A class of all-zero lattice digital filters: properties and applications", *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 26, pp304-314, Aug. 1978.

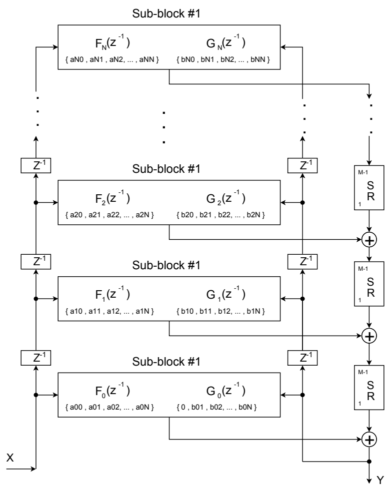

Figure 6. Filter framework A

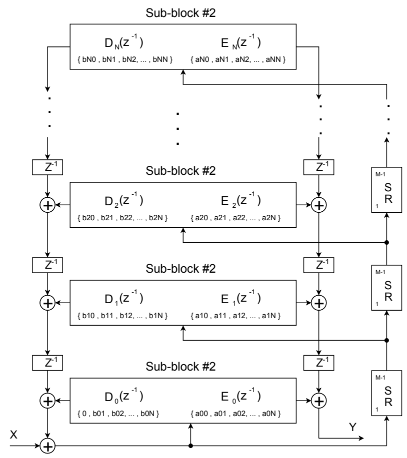

Figure 7. Filter framework B

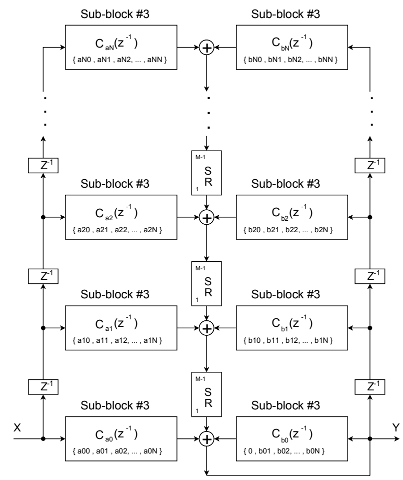

Figure 8. Filter framework C

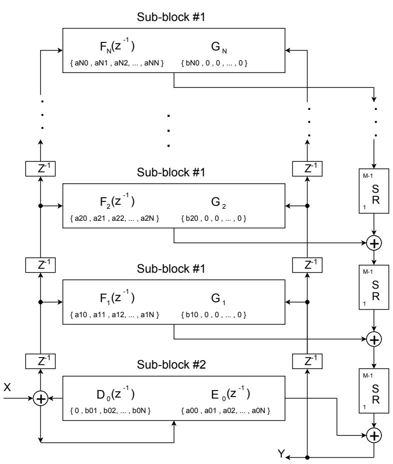

Figure 9. Filter framework F (separable denominator)

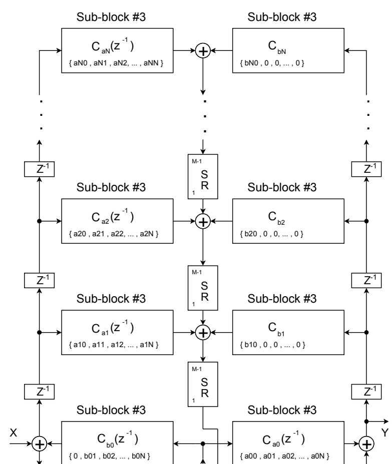

Figure 10. Filter framework H (separable denominator)

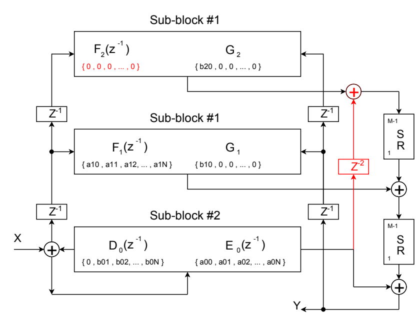

Figure 11. Quadrantal symmetry framework based on framework F

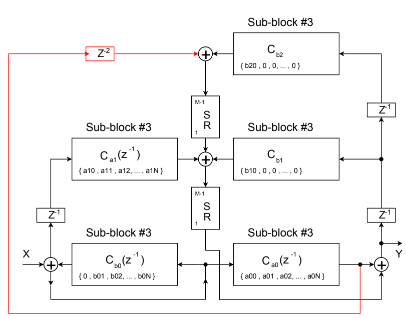

Figure 12. Quadrantal symmetry framework based on framework H

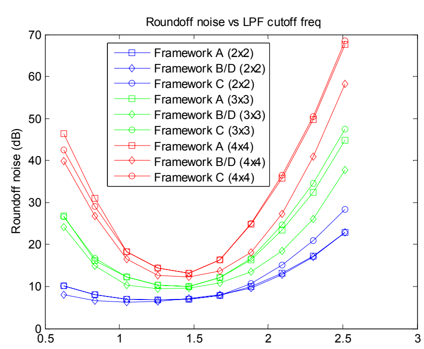

Figure 13. Roundoff noise

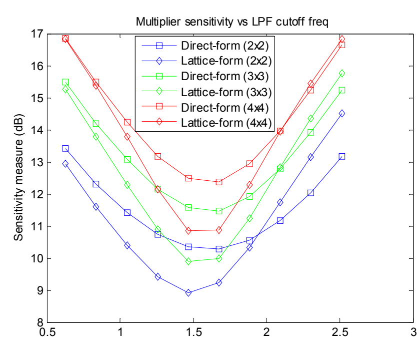

Figure 14. Multiplier sensitivity