Feedforward FFT Hardware Architectures Based on Rotator Allocation

Mario Garrido Gálvez, Shen-Jui Huang and Sau-Gee Chen

The self-archived postprint version of this journal article is available at Linköping University Institutional Repository (DiVA):
http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-145244

N.B.: When citing this work, cite the original publication.
https://doi.org/10.1109/TCSI.2017.2722690

Original publication available at:
https://doi.org/10.1109/TCSI.2017.2722690
Copyright: Institute of Electrical and Electronics Engineers (IEEE)
http://www.ieee.org/index.html
©2018 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.
Feedforward FFT Hardware Architectures based on Rotator Allocation

Mario Garrido, Member, IEEE, Shen-Jui Huang and Sau-Gee Chen

Abstract—In this paper we present new feedforward FFT hardware architectures based on rotator allocation. The rotator allocation approach consists in distributing the rotations of the FFT in such a way that the number of edges in the FFT that need rotators and the complexity of the rotators are reduced. Radix-2 and radix-2\(^k\) feedforward architectures based on rotator allocation are presented in this paper. Experimental results show that the proposed architectures reduce the hardware cost significantly with respect to previous FFT architectures.

Index Terms—Fast Fourier Transform (FFT), Multi-path delay commutator (MDC), Pipelined architecture, Radix-2, Radix-2\(^k\).

I. INTRODUCTION

The fast Fourier transform (FFT) is one of the most important algorithms in the field of digital signal processing. It is used to calculate the discrete Fourier transform (DFT) efficiently. In order to meet the high performance and real-time requirements of modern applications, hardware designers have always tried to implement efficient architectures for the computation of the FFT. In this context, pipelined hardware architectures [1]–[27] are widely used, because they provide high throughput and low latency suitable for real time, as well as a reasonably low area and power consumption.

There are three main types of pipelined FFT architectures: feedback (FB), feedforward (FF) and serial commutator (SC). First, feedback architectures [1]–[14] are characterized by their feedback loops, i.e., some outputs of the butterflies are fed back to the memories at the same stage. Feedback architectures are divided into single-path delay feedback (SDF) [1]–[5], which process a continuous flow of one sample per clock cycle, and multi-path delay feedback (MDF) or parallel feedback [6]–[14], which process several samples in parallel. Second, feedforward architectures [3], [4], [15]–[21], which mainly consist of multi-path delay commutator (MDC) FFTs [3], do not have feedback loops and each stage passes the processed data to the next stage. MDC architectures process several samples in parallel. A comparison of parallel FFT architectures is provided in [16]. Finally, SC FFT architectures [22] are characterized by the use of circuits for bit-dimension permutation of serial data.

If the purpose is to improve parallel FFT architectures, MDC FFTs already achieve the minimum number of butterflies with 100% usage and the minimum amount of total memory [16]. However, there is still room for improvement regarding rotators: Different radices lead to different amount of rotators, as shown in [16]. For instance, radix-2\(^3\) and radix-2\(^4\) architectures have less general rotators than radix-2 and radix-2\(^2\). However, instead of general rotators, they have a larger number of \(W_8\) and \(W_{16}\) rotators. Therefore, there is a trade-off among rotators.

This paper presents a new idea called rotator allocation. The purpose is to reduce the number and the complexity of the rotators even further. The number of rotators is reduced by only having rotations by 0° in some of the FFT edges. The complexity of the rotators is reduced by distributing the rotations of the FFT so that similar rotations are calculated by the same rotator. For instance, a rotation by 45° and a rotation by 135° can be easily calculated by the same rotator as it only consists of a rotation by 45° plus a trivial rotation by 0° or 90°. New feedforward FFT architectures based on rotator allocation are proposed in this paper. These architectures use radices that lead to a small amount of hardware resources.

The paper is organized as follows. Section II reviews the FFT algorithm. Section III summarizes the types of rotators used in FFT architectures and assigns a cost to them in terms of equivalent adders. Section IV provides general guidelines to design an FFT hardware architecture. Section V explains the design of FFT hardware architectures using rotator allocation. Section VI presents the proposed architectures both for radix-2 and radix-2\(^k\). Section VII compares the proposed architectures to the state-of-the-art in terms of the number of multipliers and equivalent adders. Section VIII provides experimental results and compares them to to previous works in the literature. Finally, Section IX summarizes the main conclusions of the paper.

II. THE FFT ALGORITHM

The \(N\)-point DFT of an input sequence \(x[n]\) is defined as:

\[
X[k] = \sum_{n=0}^{N-1} x[n] W_N^{nk}, \quad k = 0, 1, \ldots, N - 1
\]

(1)

where \(W_N = e^{-j \frac{2\pi}{N}}\).

To calculate the DFT, the FFT based on the Cooley-Tukey algorithm [28] is mostly used. The Cooley-Tukey algorithm reduces the number of operations from \(O(N^2)\) for the DFT to \(O(N \log_2 N)\) for the FFT.
the stages indicates a rotation by:

\[ W_N^\phi = e^{-j \frac{2\pi}{N} \phi} \]  

As a consequence, samples for which \( \phi = 0 \) do not need to be rotated. Likewise, if \( \phi \in [0, N/4, N/2, 3N/4] \) the samples must be rotated by 0°, 270°, 180° or 90°, which correspond to complex multiplications by 1, \(-j\), \(-1\) and \(j\), respectively. These rotations are considered trivial, because they can be carried out by interchanging the real and imaginary components and/or changing the sign of the data.

An index \( I \equiv b_{n-1} \ldots b_0 \) is also added to the graph, where \( b_{n-1} \ldots b_0 \) is the binary representation of \( I \). This index is used in the explanations throughout the paper.

Different radices only differ in the rotations at the different FFT stages [30], whereas the butterflies are the same. Thus, different algorithms lead to different distributions of rotations, which may influence the design of hardware architectures, as shown throughout this paper.

### III. Hardware Rotators for the FFT

This section defines some terms used later in the paper.

A *general rotator* is a rotator that can carry out a rotation by any angle, which is provided as an input. A *single constant rotator* (SCR) is a rotator that carries out a rotation by a single constant angle [31]. A *multiple constant rotator* (MCR) is a rotator that rotates by an angle selected among several constant angles.

A symmetric angle set (SAS) is a set of angles \( n\pi/2 \pm \alpha \), where \( n = 0, \ldots, 3 \) and \( \alpha \in [0, \pi/4] \). Any rotation in a symmetric angle set can be calculated as a rotation by \( \alpha \in [0, \pi/4] \), a trivial rotation and a exchange of the real and imaginary part of the rotation coefficient. Figure 2 shows the angles used in a 16-point FFT (\( \phi = 0, \ldots, 15 \)) and in a 32-point FFT (\( \phi = 0, \ldots, 31 \)) which correspond to dividing the circumference into 16 and 32 equal parts, respectively. These angles form several symmetric angle sets, which are shown in Tables I and II.

A symmetric angle set (SAS) is a set of angles \( n\pi/2 \pm \alpha \), where \( n = 0, \ldots, 3 \) and \( \alpha \in [0, \pi/4] \). Any rotation in a symmetric angle set can be calculated as a rotation by \( \alpha \in [0, \pi/4] \), a trivial rotation and a exchange of the real and imaginary part of the rotation coefficient. Figure 2 shows the angles used in a 16-point FFT (\( \phi = 0, \ldots, 15 \)) and in a 32-point FFT (\( \phi = 0, \ldots, 31 \)) which correspond to dividing the circumference into 16 and 32 equal parts, respectively. These angles form several symmetric angle sets, which are shown in Tables I and II.

An *M-rotator* or *M-rot* is a rotator that can rotate any number of angles in \( M \) different symmetric angle sets. For instance, a rotator that rotates by 0°, 45° and 135° is a 2-rot, as it rotates angles in the symmetric angle sets \( n\pi/2 \) and \( n\pi/2 \pm \pi/4 \). Likewise, the FFT twiddle factor \( W_k \) is a 2-rot, the twiddle factor \( W_{16} \) is a 3-rot (note the 3 SAS in Table I) and the twiddle factor \( W_{32} \) is a 5-rot (note the 5 SAS in Table II). The symbols used in this paper for the different types of rotators are shown in Fig. 3.

Table III shows an estimation of the hardware cost of different types of rotators in terms of adders. The table is divided into rotators for arbitrary scaling and unity scaling. Arbitrary scaling means that any scaling is accepted and unity scaling
TABLE III
ESTIMATION OF EQUIVALENT NUMBER OF ADDERS IN FFT ROTATORS.

<table>
<thead>
<tr>
<th>Rotator</th>
<th>Hardware Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>------------------</td>
<td>-----------</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>ARBITRARY SCALING</td>
<td></td>
</tr>
<tr>
<td>1-rot CCSSI [31]: SCR arbitrary scaling</td>
<td>6</td>
</tr>
<tr>
<td>2-rot CCSSI [31]: MCR uniform scaling</td>
<td>6</td>
</tr>
<tr>
<td>3-rot CCSSI [31]: MCR uniform scaling</td>
<td>8</td>
</tr>
<tr>
<td>General CORDIC II [32]</td>
<td>16</td>
</tr>
<tr>
<td>General Memoryless CORDIC [33]</td>
<td>21</td>
</tr>
<tr>
<td>General Complex Multiplier: Booth encoding</td>
<td>34</td>
</tr>
<tr>
<td>UNITY SCALING</td>
<td></td>
</tr>
<tr>
<td>1-rot CCSSI [31]: SCR with unity scaling</td>
<td>10</td>
</tr>
<tr>
<td>2-rot CCSSI [31]: MCR with unity scaling</td>
<td>10</td>
</tr>
<tr>
<td>3-rot CCSSI [31]: MCR with unity scaling</td>
<td>10</td>
</tr>
<tr>
<td>General CORDIC II [32]</td>
<td>16</td>
</tr>
<tr>
<td>General Memoryless CORDIC [33]</td>
<td>21</td>
</tr>
<tr>
<td>General Complex Multiplier: Booth encoding</td>
<td>34</td>
</tr>
</tbody>
</table>

Fig. 3. Symbols used in the paper for the different types of rotators.

means that the scaling must be a power of two [31]. M-rots are obtained using the CCSSI approach [31], whereas general rotators are obtained either using the CORDIC algorithm [32], [33] or complex multipliers. The table shows the adders (Add. Rot.) and multiplexers involved in the rotation. Considering that a multiplexer is equivalent to 1/4 adders [34], the column Add. Mux. shows the equivalent number of adders of the multiplexers involved in the rotation. The next column is the number of adders needed for scaling compensation. This is not needed in case of arbitrary scaling. The scaling compensation factor in the CORDIC II is

\[ K = 1/(1.563-1.008) = 0.6347 \approx 0.6348 = 2^{-1} + 2^{-3} + 2^{-7} + 2^{-9} \]

Thus, it needs 6 adders, 3 for the real part and three for the imaginary part. The scaling compensation in the memoryless CORDIC needs 4 adders, 2 for the real part and 2 for the imaginary part [33]. The last column of the table shows the total number of equivalent adders, which is the result of adding the adders involved in the rotation, the equivalent adders of the multiplexers and the adders of the scaling compensation. We use the results in Table III to compute the number of equivalent adders of the proposed FFT architectures.

IV. DESIGNING AN FFT HARDWARE ARCHITECTURE

In order to design an FFT hardware architecture, we have to be aware of the FFT properties introduced in [16]. The first property, which is general for any FFT architecture and any \( N \), is that at any FFT stage, butterflies operate on data whose index \( I \) differ in \( b_{n-s} \), where \( n \) is the number of FFT stages and \( s \) is the specific stage that we are considering. This fact can be observed in the flow graph of Fig. 1. In this flow graph, the index has \( n = 4 \) bits, i.e., \( I \equiv b_3b_2b_1b_0 \). At the first stage, the butterflies operate on samples whose indexes differ in \( b_{n-s} = b_{4-1} = b_3 \). This happens for samples with indexes 0 and 8, 1 and 9, etc. For the second stage, the different bit is \( b_{n-s} = b_{4-2} = b_2 \). Note for instance, that the data with indexes 0 and 4 are operated together in the butterfly at stage 2. For the third and fourth stages, the corresponding bits are \( b_1 \) and \( b_0 \), respectively.

Therefore, if we want to design an FFT hardware architecture, we have to assure that at each stage \( s \), the indexes of the inputs to any butterfly at any time instant differ in \( b_{n-s} \). Note that the term butterfly refers now to a hardware component of the architecture, not to the mathematical operation of the algorithm in the flow graph. If we consider the example of Fig. 3 in [16], we observe that this property is fulfilled at all the stages. In this figure \( b_{n-s} \) is at the lowest parallel dimension, which corresponds to the pair of samples that flow into the butterflies at the same clock cycle.

The property of \( b_{n-s} \) is the only requirement set by the butterflies in FFT hardware architecture. As long as this property is met, we can have any data order at the different FFT stages. This allows for exploring a variety of data orders at the FFT stages. This is what is done in the current paper, as explained later in Section V.

The second FFT property refers to the rotations at the FFT stages. At each stage, any sample with index \( I \) must be rotated according to equation (2) by a value \( \phi \) that depends on the index \( I \) and on the stage, \( s \). We can represent it as \( \phi_s(I) \). The work [30] explains how to calculate \( \phi_s(I) \), which only depends on the FFT algorithm that is used.

By combining these ideas, we lead to the following conclusions: On the one hand, any order at any stage of any FFT hardware architecture is possible as long as the property of \( b_{n-s} \) is met at the input of the butterflies. On the other hand, to any index \( I \) at any FFT stage corresponds a specific rotation \( \phi_s(I) \). As a result, we can play with the data order at different FFT stages to look for patterns that allow for a more optimized distribution of rotations. This is the idea behind the proposed rotator allocation approach.

V. FFT DESIGN USING ROTATOR ALLOCATION

The purpose of the FFT design using rotator allocation is to distribute the rotations of the FFT in such a way that the number and complexity of the required rotators is reduced. Fig. 4(a) shows an example of a layout for the first three stages of a 16-point 4-parallel FFT. The indexes in the figure
show how data flows at the different stages. Each element in the matrices of indexes is the index value according to Fig. 1. Data flows from left to right. Thus, values in the same column are data that flow in parallel and values in the same row flow through the same path in consecutive clock cycles. The matrices of rotations show the value $\phi_n(I)$ that corresponds to each of the indexes according to the flow graph in Fig. 1. For instance, the rotation corresponding to the index 10 at stage 1 is $\phi_n(I) = \phi_1(10) = 2$ according to Fig. 1. This case is highlighted in Fig. 4(a).

As for the indexes, each column in the matrices of rotations are rotations that are calculated in parallel at the same clock cycle, whereas rotations in the same row are calculated in consecutive clock cycles by the same rotator. According to this, each row of the matrices of rotations that must be calculated by a single rotator in consecutive clock cycles. For instance, stage 1 needs a rotator by $\phi = \{0, 1, 2, 3\}$ and a rotator by $\phi = \{4, 5, 6, 7\}$, which are 3-rots according to Table I. Stage 2 includes 2 2-rots and stage 3 includes 2 trivial rotators (T).

The layout in Fig. 4(a) translates into the FFT architecture in Fig. 4(b), which shows the rotators at each stage, as well as the content of the rotation memories.

By applying rotator allocation we aim to reduce the number of rotators and their complexity. Rotator allocation simply consists of reorganizing the matrices of indexes and, therefore, the matrices of rotations, in such a way that the matrices of rotations have less rotators (if possible) and the complexity of the rotators is smaller. On the one hand, fewer rotators are achieved when there are more rows in the matrices of rotations whose elements are $\phi = 0$. On the other hand, the complexity of the rotators is reduced when the rotations in the same row are in less SAS.

The procedure of rotator allocation consists in distributing the bits $b_{n-1} \ldots b_0$ of the index $I$ into serial and parallel dimensions. Serial dimensions correspond to data arriving at the same input terminal in series and parallel dimensions refers to data arriving at parallel terminals. Depending on the FFT size $N = 2^n$ and the number of parallel data in the FFT $P = 2^p$, the number of serial bits is $n - p = \log_2(N) - \log_2(P)$ and the number of parallel ones is $p = \log_2(P)$. For the example in Fig. 4, $N = 16$ and $P = 4$, so there are $n - p = \log_2(16) - \log_2(4) = 2$ serial dimensions and $p = \log_2(4) = 2$ parallel ones. The alternatives to allocate the bits correspond to all the possible permutations of the bits, i.e.,

$$\begin{align*}
\text{Serial} & | \text{Parallel} \\
| & | \\
| & |
\end{align*}$$

However, neither the order of the serial bits nor the order of the parallel bits affect the complexity of the rotators: A different order of the serial bits changes the order of the rotations, but the same rotations are calculated by the rotators. A different order of the parallel bits changes the edges in which the rotators are placed, but the rotators are the same. Therefore, the alternatives in each row of equation (3) have
the same complexity, so only one alternative per row needs to be evaluated. According to this, the number of alternatives that need to be evaluated at each FFT stage is

$$\binom{n}{p} = \frac{n!}{p!(n-p)!}.$$  

(4)

Note that we can choose any order in equation (3) and design the data management of the FFT to achieve the desired order. However, it is in general a good idea to place the rotators close to the butterflies of the same state. This requires that the data order at the rotator respects the property of close to the butterflies of the same state. This requires that the additional rotators are simpler than a general rotator. Therefore, the FFT hardware does not duplicate when duplicating the number of samples in parallel. Only the number of butterflies is duplicated. The memory is reduced and the rotators are almost duplicated in number, but simplified in complexity.

Figure 7 shows the proposed 4-parallel radix-2 FFTs for decimation in time (DIT) [29]. It can be noted that this architecture is symmetric with respect to the DIF one: Whereas for the DIF case the length of the buffers and the complexity of the rotators decrease with the number of the stage, for the DIT case it is the opposite. As a result, the amount of hardware resources for the DIF and DIT versions is the same. The 8-parallel radix-2 DIT FFT is obtained in a similar way.

For larger $N$, any $N$-point $P$-parallel radix-2 DIF FFT architectures based on rotator allocation is obtained by adding one extra stage at the beginning of the $N/2$-point $P$-parallel radix-2 DIF architecture, as can be observed when comparing the 32-point 4-parallel radix-2 FFT in Fig. 5 and the 16-point 4-parallel radix-2 FFT in Fig 4(d). Analogously, for the DIT case any $N$-point $P$-parallel DIT FFT is obtained by adding one extra stage at the end of the $N/2$-point $P$-parallel radix-2 DIT architecture. In both DIF and DIT cases, further stages added to the 32-point FFT require general rotators. Thus, the 64-point radix-2 FFT requires the same rotators as the 32-point radix-2 FFT plus two general rotators. For any radix-2 4-parallel (DIF or DIT) FFT based on rotator allocation, the type and number of rotators are shown in Table V. For the DIF case the length of the buffers and the complexity of the rotators decrease with the number of the stage, for the DIT case it is the opposite. As a result, the amount of hardware resources for the DIF and DIT versions is the same. The 8-parallel radix-2 DIT FFT is obtained in a similar way.

For larger $N$, any $N$-point $P$-parallel radix-2 DIF FFT architectures based on rotator allocation is obtained by adding one extra stage at the beginning of the $N/2$-point $P$-parallel radix-2 DIF architecture, as can be observed when comparing the 32-point 4-parallel radix-2 FFT in Fig. 5 and the 16-point 4-parallel radix-2 FFT in Fig 4(d). Analogously, for the DIT case any $N$-point $P$-parallel DIT FFT is obtained by adding one extra stage at the end of the $N/2$-point $P$-parallel radix-2 DIT architecture. In both DIF and DIT cases, further stages added to the 32-point FFT require general rotators. Thus, the 64-point radix-2 FFT requires the same rotators as the 32-point radix-2 FFT plus two general rotators. For any radix-2 4-parallel (DIF or DIT) FFT based on rotator allocation, the type and number of rotators are shown in Table V. For the DIF case the length of the buffers and the complexity of the rotators decrease with the number of the stage, for the DIT case it is the opposite. As a result, the amount of hardware resources for the DIF and DIT versions is the same. The 8-parallel radix-2 DIT FFT is obtained in a similar way.

Table IV lists the twiddle factors values ($\phi$) for the 32-point 4-parallel radix-2 architecture in Fig. 5.

<table>
<thead>
<tr>
<th>Stage</th>
<th>Alternative</th>
<th>1-rot</th>
<th>2-rot</th>
<th>3-rot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 1</td>
<td>0 0 0 0 0 0 0 0</td>
<td>0 1 3 5 7 9 11 13</td>
<td>0 0 0 0 0 0 0 0</td>
<td></td>
</tr>
<tr>
<td>Stage 2</td>
<td>0 0 0 0 0 0 0 0</td>
<td>0 4 8 12 0 4 8 12</td>
<td>0 0 0 0 0 0 0 0</td>
<td></td>
</tr>
<tr>
<td>Stage 3</td>
<td>0 8 0 8 0 8 0 8</td>
<td>2 6 10 14 2 6 10 14</td>
<td>0 0 0 0 0 0 0 0</td>
<td></td>
</tr>
<tr>
<td>Stage 4</td>
<td>0 0 0 0 0 0 0 0</td>
<td>0 0 0 0 0 0 0 0</td>
<td>0 0 0 0 0 0 0 0</td>
<td></td>
</tr>
</tbody>
</table>

Table V lists the number of butterflies and rotators for the 32-point 4-parallel and 8-parallel radix-2 DIF architecture. The architectures are shown in Fig. 4(c) and Fig. 4(d), respectively.

Table VI lists the number of butterflies and rotators for the 32-point 4-parallel and 8-parallel radix-2 DIF architecture. The architectures are shown in Fig. 4(c) and Fig. 4(d), respectively.

### Table IV

<table>
<thead>
<tr>
<th>Stage</th>
<th>Alternative</th>
<th>1-rot</th>
<th>2-rot</th>
<th>3-rot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 1</td>
<td>0 0 0 0 0 0 0 0</td>
<td>0 1 3 5 7 9 11 13</td>
<td>0 0 0 0 0 0 0 0</td>
<td></td>
</tr>
<tr>
<td>Stage 2</td>
<td>0 0 0 0 0 0 0 0</td>
<td>0 4 8 12 0 4 8 12</td>
<td>0 0 0 0 0 0 0 0</td>
<td></td>
</tr>
<tr>
<td>Stage 3</td>
<td>0 8 0 8 0 8 0 8</td>
<td>2 6 10 14 2 6 10 14</td>
<td>0 0 0 0 0 0 0 0</td>
<td></td>
</tr>
<tr>
<td>Stage 4</td>
<td>0 0 0 0 0 0 0 0</td>
<td>0 0 0 0 0 0 0 0</td>
<td>0 0 0 0 0 0 0 0</td>
<td></td>
</tr>
</tbody>
</table>
difference in the length of the buffers of the first 16-point FFT. The connection stage has 4 general rotators. This is the same for all radix-\(2^k\) FFTs presented next. As a result, the radix-\(2^k\) feedforward FFTs have \(P/2 \cdot \log_2 N\) butterflies, a total memory size \(N - P\), and the rotators of the radix-2 sub-FFTs plus \(P\) general rotators per connection stage.

Table VII presents the proposed radix-\(2^k\) feedforward FFTs based on rotator allocation for 4-parallel samples. The table is divided into Selected architectures and Other alternatives with higher hardware cost. The selected architectures are the ones that achieve smallest hardware cost. The first two columns of the table show the number of points, \(N\), and the radix. The third column indicates the FFT algorithms used in these architectures, which are represented by the binary trees in Fig. 9. Note that the binary tree provides the values of the rotation angles at each FFT stage [35]. The fourth to the seventh columns show the number of rotators that the architecture requires. The general rotators are separated into two groups: The general rotators in the radix-2 sub-FFTs and the general rotators used to interconnect sub-FFTs. Finally, the last column indicates the total number of equivalent real adders of the entire architecture, including the adders of the butterflies and rotators, and the equivalent number of real adders of the multiplexers:

\[
\text{Equiv. Adders} = 2P \cdot \log_2 N + \text{AddersRot} + 2P/4 \cdot \log_2 (N/P),
\]

where the adders for the rotators (AddersRot) are computed according to the estimations in Table III.
Fig. 7. Proposed 32-point 4-parallel radix-2 MDC DIT FFT architecture.

Fig. 8. 256-point 4-parallel radix-2⁴ MDC DIF FFT architecture.

<table>
<thead>
<tr>
<th>FFT Algorithm</th>
<th>Rotators</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>Radix</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>16</td>
<td>2</td>
</tr>
<tr>
<td>32</td>
<td>2</td>
</tr>
<tr>
<td>64</td>
<td>2</td>
</tr>
<tr>
<td>128</td>
<td>2</td>
</tr>
<tr>
<td>256</td>
<td>2</td>
</tr>
<tr>
<td>512</td>
<td>2</td>
</tr>
<tr>
<td>1024</td>
<td>2</td>
</tr>
<tr>
<td>2048</td>
<td>2</td>
</tr>
<tr>
<td>4096</td>
<td>2</td>
</tr>
</tbody>
</table>

In Table VII, radix-$2^k$ starts to get better results than radix-2 for $N = 64$, whereas radix-2 is preferable for sizes smaller than or equal to 32.

Analogously, Table VIII shows the proposed radix-$2^k$ feed-forward FFTs based on rotator allocation for 8-parallel samples. In this case, radix-$2^k$ starts to obtain better results than radix-2 for $N = 128$.

VII. COMPARISON

Table IX compares the proposed architectures to previous ones in terms of complex adders and rotators. The table includes the rotators used in previous architectures as a function of $n = \log_2(N)$. For the proposed architectures there is no generalized expression. Instead, the number of rotators is provided by the Selected architectures in Tables VII and VIII. For a better comparison of the number of rotators, Table IX includes the examples of $N = 1024$ and $N = 4096$. The architectures are sorted out by the exponent $k = 0, \ldots, 4$ of the radix-$2^k$. It can be observed that radix-2 architectures lead to the largest number of complex rotators, and this number decreases as we increase $k$ until radix-$2^4$. For $N = 1024$ the proposed architectures reduce this number even further, as they require half the general rotators of previous radix-$2^4$ FFT architectures, both for 4-parallel and 8-parallel designs. Note also that when not using rotator allocation, a 1024-point FFT that uses radix-$2^5$ requires 8 general rotators ($4W_{1024}$ and $4W_{32}$), 4 3-rot and 4 2-rot, and a 1024-point FFT that
uses radix-2^4 [16] requires 8 general rotators (4 W_{1024} and 4 W_{32}) and 6 3-rot, as shown in Table IX. Thus, radix-2^4 and a radix-2^5 FFTs without rotator allocation are comparable in area. However, when using rotator allocation, these numbers are improved noticeably.

For the case of N = 4096 in Table IX, the number of general rotators of the proposed architectures is the same as those in previous radix-2^4 FFTs. However, the number of non-general non-trivial rotators and their complexity is reduced significantly by using the proposed approach.

The total area occupied by the rotators is compared in Figs. 10 and 11 for 4-parallel and 8-parallel FFTs, respectively. The area is measured in terms of equivalent adders, considering the adder cost of each type of rotator according to Table III. Figs. 10 and 11 show that the rotator area of the proposed architectures is smaller than that in previous approaches for all FFT sizes. For 4-parallel architectures, the rotator area reduction ranges from 17% to 36%. For 8-parallel architectures, it ranges from 23% for to 50%. This highlights the significant improvements of using the rotator allocation approach.
For the area of the entire FFT architecture, Figs. 12 and 13 compare the proposed feedforward FFTs based on rotator allocation with previous feedforward FFT architectures for 4-parallel and 8-parallel, respectively. In all the architectures the total memory is the same, $N - P$. Therefore, the comparison is done in terms of equivalent adders of rotators, butterflies and shuffling circuits. For the proposed architectures, the number of equivalent adders comes from Tables VII and VIII, and equation (5). For the previous architectures, we calculate the number of equivalent adders in the same way as for the proposed architectures, i.e., considering the equivalent adders in butterflies and shuffling circuits, and an optimized implementation of the rotators according to Table III.

Figs. 12 and 13 show that the proposed architectures reduce
the number of equivalent adders with respect to previous approaches. For 4-parallel and $N = 4096$ points, the savings of using rotator allocation are around 15% with respect to radix-$2^4$ [16], and 40% with respect to radix-$2^2$ [16], [17]. For 8-parallel and $N = 4096$, the savings are around 22% with respect to radix-$2^4$ architectures [16], and 48% with respect to radix-$2^2$ [16], [17].

VIII. IMPLEMENTATION

In order to verify the advantages of the proposed architectures, we have implemented the proposed 1024-point 4-parallel radix-$2^5$ DIF FFT based on rotator allocation, shown in Table VII. The input data word length is chosen to be 16 bits. Similar to the case of the 256-point 4-parallel radix-$2^4$ MDC DIF FFT in Fig. 8, the 1024-point 4-parallel radix-$2^5$ DIF FFT can be derived by cascading two 32-point MDC DIF FFTs as that in Fig. 5, changing the length of the buffers of one of them, and including rotators and shuffling circuits between both of them.

A. FPGA results

Table X compares 4-parallel 1024-point FFT architectures on FPGAs. The references to previous works include the first author and the year of the work. For instance, ‘Wang16’ means ‘Wang, 2016’. The table includes the area in terms of slices, BRAM and DSP slices, and the performance in terms of clock frequency (Clk) and throughput (Th). The proposed results are obtained for a Virtex-6 XC6VSX475T-1-FF1156 FPGA. The comparison shows that the proposed architecture reduces by 64% or more the number of DSP slices with respect to previous approaches. This is due to the fact that the proposed architecture has only 4 general rotators, which are implemented in the DSP slices. The rest of rotators are implemented as distributed logic (slices). This, however, does not increase significantly the amount of distributed logic: The proposed architecture requires approximately the same distributed logic as [16], [17], which also require 12 BRAM.

Figure 14 shows the area comparison as a bar diagram. It can be observed that our approach saves a significant amount of DSP slices used for rotators with respect to previous approaches, while keeping a small number of slices and BRAM.
The difference between these two versions is that the high performance approach includes additional pipelining. The experimental results show that the low power version only requires 8.88 mW at 180 MHz. The high performance version calculates the 1024-point FFT in 0.83 µs at 1.28 GS/s. In both versions, the SQNR [38], [39] is 40.3 dB, which meets the SQNR needed in applications such as UWB [18], Wi-Fi [18] and gigabit WPAN [40]. Finally, the area of the proposed FFTs is small, around 0.2 mm². This meets the main purpose of the paper, i.e., achieving a small area for the FFT.

IX. Conclusions

In this paper, we have presented the rotation allocation approach. It consists in finding an efficient distribution of FFT rotations that reduces the number of rotators and their complexity. This leads to new radix-2 and radix-2⁵ MDC FFT architectures. These architectures require the same memory and butterflies as previous MDC FFTs, but fewer and/or simpler rotators. For 4-parallel FFTs, the area of rotators is reduced by a factor from 17% to 36% with respect to previous approaches. For 8-parallel FFTs, the rotator area reduction ranges from 23% to 50%. These savings are confirmed with the implementation of a 4-parallel 1024-point radix-2⁵ MDC FFT based on rotation allocation. Experimental results on FPGAs show significant reduction in area with respect to previous 4-parallel 1024-point FFTs, with 64% less DSP slices than previous FFT architectures. This leads to 33% less total FFT area. Experimental results on ASICs lead to a very small area, together with low power consumption or high performance.

X. Acknowledgment

The authors would like to thank Dr. Martin Kumm for his help with the calculation of the adder cost of the rotators.

Figure 15 compares the throughput versus the FPGA utilization for 1024-point FFTs on FPGAs. The FPGA utilization is calculated as the percentage of FPGA hardware resources used of a Virtex-6 XC6VSX475T-1-FF1156 FPGA, according to the definition in [17].

Fig. 15 includes both serial FFTs (1P) and parallel FFTs (2P, 4P, 8P). The architectures improve as they increase in throughput and/or decrease in FPGA utilization, towards the upper left corner. Both the throughput and the FPGA utilization increase with the parallelization of the FFT. Serial FFTs [2], [27] appear in the lower left corner and 8-parallel FFTs appear towards the upper right one. The proposed FFT provides a good trade-off between throughput and resources. Its FPGA utilization is 1.03% and the throughput is 1012 MS/s. This is 33% less area and 11% more throughput than the closest FFT [16], [17]. Furthermore, it has significantly smaller FPGA utilization than previous parallel FFTs, while keeping a high throughput.

B. ASIC results

Table XI shows the ASIC results of the proposed FFT in two different versions; low power and high performance.

Figure 15. Throughput vs. FPGA utilization of 1024-point FFT hardware architectures.

<table>
<thead>
<tr>
<th>FFT Parameters</th>
<th>Proposed (low power)</th>
<th>Proposed (high performance)</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Radix</td>
<td>2⁵</td>
<td>2⁵</td>
</tr>
<tr>
<td>Word length</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Technology (nm)</td>
<td>55</td>
<td>55</td>
</tr>
<tr>
<td>Voltage (V)</td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td>Ck (MHz)</td>
<td>180</td>
<td>320</td>
</tr>
<tr>
<td>Th (MS/s)</td>
<td>1290</td>
<td>1290</td>
</tr>
<tr>
<td>Latency (cycles)</td>
<td>263</td>
<td>265</td>
</tr>
<tr>
<td>Latency (µs)</td>
<td>1.46</td>
<td>0.83</td>
</tr>
<tr>
<td>Area (mm²)</td>
<td>0.198</td>
<td>0.212</td>
</tr>
<tr>
<td>SQNR (dB)</td>
<td>40.3</td>
<td>40.3</td>
</tr>
<tr>
<td>Power (mW)</td>
<td>8.88</td>
<td>17.02</td>
</tr>
</tbody>
</table>

@ 180 MHz @ 320 MHz

The authors would like to thank Dr. Martin Kumm for his help with the calculation of the adder cost of the rotators.

REFERENCES

IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS PART I: REGULAR PAPERS 12


Mario Garrido (M’07) received the M.S. degree in electrical engineering and the Ph.D. degree from the Technical University of Madrid (UPM), Madrid, Spain, in 2004 and 2009, respectively. In 2010 he moved to Sweden to work as a postdoctoral researcher at the Department of Electrical Engineering at Linköping University. Since 2012 he is Associate Professor at the same department. His research focuses on optimized hardware design for signal processing applications. This includes the design of hardware architectures for the calculation of transforms, such as the fast Fourier transform (FFT), circuits for data management, the CORDIC algorithms, and circuits to calculate statistical and mathematical operations. His research covers high-performance circuits for real-time computation, as well as designs for small area and low power consumption.

Shen-Jui Huang received his M.S. degree from the Department of Electrical Engineering, National Taiwan University in 1997, and his Ph.D. degree in Electronics from National Chiao Tung University in 2012. His research interests include baseband signal processing and circuit implementation. Currently he is a deputy manager in Novatek Corp., Hsinchu, Taiwan.

Sau-Gee Chen received his B.S. degree from National Tsing Hua University, Taiwan, in 1978, M.S. degree and Ph.D. degree in electrical engineering, from the State University of New York at Buffalo, NY, in 1984 and 1988, respectively. Currently, he is a Professor at the Department of Electronics Engineering, National Chiao Tung University (NCTU), Taiwan, and a member of Board of Governor, IEEE Taipei Section. He was the department Chair of the Department of Electronics Engineering, NCTU, during 2012-2015. He has been appointed as the Coordinator of IEEE VTS Asia-Pacific Chapters, since 2014. He was the Chair, IEEE Vehicular Technology Society, Taipei Chapter, during 2012-2013. He was Director of Honors Program, College of Electrical & Computer Engineering/College of Computer Science from 2011 to 2012 at NCTU. He also was the Associate Dean, Office of International Affairs, during March-July, 2011, as well as the director of Institute of Electronics from 2003 to 2006, all at the same organization. During 2004-2006, he served as an associate editor of IEEE Transactions on Circuits and Systems I. His research interests include digital communication, multi-media computing, digital signal processing, and VLSI signal processing. He has published more than 100 conference and journal papers, and holds 19 US and Taiwan patents.