An FFT Circuit Using Nested RNS in a Digital Spectrometer for a Radio Telescope

Hiroki Nakahara, Ehime University, Japan
Tsutomu Sasao, Meiji University, Japan
Hiroyuki Nakanishi, Kagoshima University, Japan
Kazumasa Iwai, National Institute of Information and Communications Technology, NICT, Japan
Tohru Nagao, Ehime University, Japan
Naoya Ogawa, Ehime University, Japan

Abstract—A radio telescope analyzes radio frequency (RF) signal received from celestial objects. It consists of an antenna, a receiver, and a spectrometer. The spectrometer converts the time domain signal into the frequency domain signal by an FFT operation. This paper proposes an FFT circuit based on nested residue number system (NRNS). Since the FFT operation is the most computationally intensive part, parallel implementation is necessary to realize a high-speed FFT. We used an FPGA to implement the circuit. The FPGA consists of look-up tables (LUTs) and block RAMs (BRAMs). For direct parallel FFT realization using an existing FPGA library, the number of LUTs for the complex multipliers is the bottleneck. To reduce the number of LUTs in an FPGA, we increase the dynamic range stage by stage. In this case, NRNS2NRNS converters that convert smaller dynamic range to larger dynamic range are necessary. We implemented the proposed NRNS FFT on the Xilinx Corp. Virtex 7 FPGA. Compared with a conventional binary FFT, although the number of block RAMs (BRAMs) was increased by 20.0-156.5%, in the RNS FFT, the number of LUTs was decreased by 42.4-47.8% and the maximum clock frequency was increased by 9.3-41.7%. With this technique, we successfully implemented an FFT that satisfied the required size and speed specifications on an available FPGA, since the excessive number of LUTs was the bottleneck of the binary FFT.

I. INTRODUCTION

A. Digital Spectrometer for a Radio Telescope

A radio telescope analyzes the radio frequency (RF) signal received from celestial objects. Fig. 1 shows the operations in a typical radio telescope. First, the antenna receives the RF signal coming from the celestial objects. Second, the feed horn sends the electromagnetic wave to the receiver through the waveguide. Third, the receiver transforms it to the intermediate frequency (IF) signal. Finally, the digital spectrometer converts the analog signal to the power spectrum. Fig. 2 shows the digital spectrometer. First, the analog-to-digital converter (ADC) converts the IF analog signal into a digital signal. Then, the Fourier transform unit performs the fast Fourier transform (FFT) [3] to obtain the spectrum. Next, the magnitude unit calculates the power spectrum. Finally, to increase the dynamic range of the obtained power spectrum, they are accumulated. Since the FPGAs have a dedicated I/O to connect to the high-speed ADCs and can be reconfigured for a desired specification, they are widely used in the digital spectrometer [1], [2], [6], [8], [9], [12].

We are developing the next generation radio telescope which is called SKA (square kilometers array). It requires wide band (up to 24 GHz) FFT operations [14]. Since the available FPGA operates at only up to a megahertz order, more parallel operations are necessary. The FPGA consists of look-up tables (LUTs) and block RAMs (BRAMs). As for a direct parallel FFT realization using the existing FPGA library [18], the number of LUTs for the complex multipliers becomes the bottleneck [8], [9]. Thus, the reduction of the number of LUTs is essential. To implement an \( n \)-bit parallel multiplier for the FFT circuit, LUTs of \( O(n \cdot 2^n) \) is necessary. In this paper, to reduce the number of LUTs, we decompose large parallel multipliers into smaller ones by the residue number system (RNS).

B. Contributions of the Paper

We use a residue number system (RNS) [15], [19] to decompose \( n \)-input multipliers into smaller ones. With this
technique, we reduce the LUT count to implement the FFT. The RNS decomposes an integer into a set of $L$ integers by residues of moduli set. In a conventional RNS, since no pair of moduli have a common factor with any other, the resulting multipliers have different sizes. This means that RNS-based design requires LUTs with different sizes. To solve this problem, we use the nested RNS (NRNS) [11], which recursively decomposes the numbers in RNS. By using the NRNS, we decompose the multiplier on the FFT circuit into smaller ones so that they can be efficiently implemented by LUT of an FPGA. As a result, the proposed method can realize more FFTs in parallel than the existing FPGA library. Additionally, we compare the NRNS with the conventional RNS for different numbers of inputs $n$ for the arithmetic circuits. Finally, we show the condition that the NRNS realization is smaller than the RNS realization.

C. Organization of the Paper

The rest of the paper is organized as follows: Chapter 2 shows the pipelined binary FFT circuit; Chapter 3 introduces the conventional residue number system (RNS); Chapter 4 shows the FFT circuit based on the RNS (RNS FFT); Chapter 5 proposes the FFT based on the Nested RNS (NRNS FFT); Chapter 6 compares the NRNS FFT with the RNS one; Chapter 7 shows the experimental results; and Chapter 8 concludes the paper.

II. Binary FFT

A. Fast Fourier Transform (FFT)

Let $(x_0, x_1, \ldots, x_{N-1})$ be an input consisting of $N$ complex numbers. The discrete Fourier Transform ($N$ point DFT) is $(c_0, c_1, \ldots, c_{N-1})$, where

$$c_k = \sum_{j=0}^{N-1} a_j w_N^{jk}.$$  

and $w_N^k = \exp(-2\pi i k/N)$ is a twiddle factor. A time complexity of a direct computation for Expr. (1) is $O(N^2)$. By applying decompositions to the $N$ point DFT $s = \lceil \log_s N \rceil$ times recursively, we have a Cooley-Tukey Fast Fourier Transform ($N$ point FFT) [3]. Let $s = \lceil \log_s N \rceil$ be the number of stages, and $r$ be the radix of the FFT. In this paper, we assume that $r = 2$. Fig. 3 shows a signal flow graph obtained by the FFT algorithm, where $N = 8$ and $r = 2$.

B. Pipeline Radix-2 Binary FFT

As shown in Fig. 3, different stages handle points with different distances. By applying an index swap operation replacing indices between adjacent stages, we can adjust the points of the inputs for the butterfly operations. The swap memory performs an index swap operation. Let $w$ be a precision of the FFT. Then, the amount of memory for each swap is $wN$, and the total amount of memory for the swap memory is $wN \lceil \log_s N \rceil$. Fig. 4 shows a radix-2 butterfly operator for $r = 2$, which consists of two complex multipliers. Fig. 5 shows a pipelined radix-2 FFT [4], which allows continuous data processing. Unfortunately, the radix-2 binary FFT requires very large multipliers, since the dynamic range of the latter stages are very large.

III. Residue Number System (RNS)

A residue number system (RNS) [19], [15] is defined by a set of $L$ integer constants as follows:

$$\langle m_1, m_2, \ldots, m_L \rangle,$$

where no pair of moduli have a common factor with any other. An arbitrary integer $Z$ can be uniquely represented by the RNS as a tuple of $L$ integers as follows:

$$(z_1, z_2, \ldots, z_L),$$

where $z_i \equiv Z \pmod{m_i}$.

$$M = \prod_{i=1}^{L} m_i$$ is a dynamic range of the RNS. In the RNS, the addition, the subtraction, and the multiplication are performed in digit-wise. Let $X$ and $Y$ be integers, $x_i$ and $y_i$ be integers in the RNS defined by $m_i$ ($1 \leq i \leq L$). Let $\circ$ denote $+$ (addition), $-$ (subtraction), or $\ast$ (multiplication). Then, $Z = X \circ Y$ can be represented as

$$Z = (z_1, z_2, \ldots, z_L),$$

where $z_i \equiv (x_i \circ y_i) \pmod{m_i}$. Note that, the division is not included in the operations.

Example 3.1: Let $\langle m_1, m_2, m_3 \rangle = \langle 3, 4, 5 \rangle$ be the moduli set. Consider the multiplication $X \times Y$, where $X = 8$ and
IV. RADIX-2 RNS FFT CIRCUIT

Expr. (1) shows that the FFT operation consists of the addition, the subtraction, and the multiplication. By applying the RNS to the FFT, we have the FFT circuit based on the RNS (RNS FFT) shown in Fig. 6. In this circuit, first, the binary input signal is converted into the RNS by read only memories (ROMs). Typically, the input signals from analog-digital converters (ADCs) are 8-14 bits. The binary to RNS converter can be implemented by 18Kb BRAMS on the FPGA. Next, the RNS FFT circuit computes each signal in the digit-wise manner. In this paper, we assume that the conversion from the RNS to the binary is done off-line.

Fig. 7 shows the modulo $m_i$ RNS butterfly, which is derived from the binary butterfly operator shown in Fig. 4. In Fig. 7, $A = (A_R, A_I)$ and $B = (B_R, B_I)$ denote the complex inputs, $W = (W_R, W_I)$ denotes the complex twiddle factor, $R$ denotes the real part, and $I$ denotes the imaginary part. The $m_i$ RNS butterfly module can be realized by small LUTs [16]. Let $m_i$ be the modulo in the RNS. Then, the total amount of memory for the modulo $m_i$ butterfly is $10 \times (m_i)^2 \cdot \lceil \log_2 m_i \rceil$ bits, since the butterfly operator uses 10 arithmetic circuits. Fig. 7 shows that the $N$ points RNS FFT requires memories with $\sum_{i=1}^{L} 10 \times (m_i)^2 \cdot \lceil \log_2 m_i \rceil \cdot \lceil \log_2 N \rceil$ bits. In other words, the number of LUTs is $O((m_i)^2 \cdot (\log m_i) \cdot (\log N))$. Therefore, we can decrease the number of LUTs by decreasing $m_i$.

Fig. 8 shows the swap memory on the radix-2 RNS FFT. As shown in Fig. 6, the RNS FFT consists of $L$ moduli FFTs. Swap values for $L$ butterfly operators are stored in the swap memory. When $r = 2$, the RNS butterfly operator swaps $\frac{2r}{2}$ signals. Also, each RNS butterfly operator produces $\lceil \log_2 m_i \rceil$ bits. Thus, the amount of swap memory $\text{Mem}$ for each stage is

$$\text{Mem} = \frac{N}{2} \times \left( \sum_{i=1}^{L} \lceil \log_2 m_i \rceil \right).$$

Since the number of stages is $\lceil \log_2 N \rceil$, the total amount of swap memories is $\text{Mem} \cdot \lceil \log_2 N \rceil$ bits.

V. FFT CIRCUIT BASED ON NESTED RNS

A. RNS FFT USING RNS2RNS CONVERTERS

A direct realization of the RNS FFT shown in Fig. 6 requires an excessive number of LUTs, since the dynamic range is too large for the first half stages of the RNS butterflies. In this paper, we increase the dynamic range stage by stage. We insert RNS2RNS converters [10] that converts a small dynamic range to a large dynamic range. Fig. 9 shows the RNS FFT inserted RNS2RNS converters. In this circuit, since large moduli are removed, the number of LUTs is reduced. However, it requires additional RNS2RNS converters. We
decompose the RNS2RNS converter by applying the functional decomposition.

Fig. 10 shows the function table for the RNS2RNS converter which converts \( (m_1, m_2) = (2, 3) \) to \( (m_1, m_2, m_3) = (2, 3, 5) \). Although we can use an arbitrary moduli set in the RNS2RNS converter, to reduce the amount of hardware, we use \( g(m_1, m_2, \ldots, m_L) = (m_1, m_2, \ldots, m_L, m_{L+1}) \) as the RNS2RNS converter. In this case, as shown in Fig. 11, the RNS2RNS converter requires only the circuit for \( g' (m_1, m_2, \ldots, m_L) = m_{L+1} \). Let \( M = \prod_{i=1}^{L} m_i \) be the dynamic range. A single-memory realization of the RNS2RNS converter requires a memory with \( M \lceil \log_2 m_{L+1} \rceil \) bits. In our implementation, as shown in Fig. 12, we decompose the RNS2RNS converter into the RNS2Bin converter and the Bin2RNS converter. Let \( m_{L+1} \) be the modulo in the Bin2RNS converter, then the column multiplicity of the decomposition is at most \( m_{L+1} \). Thus, we can reduce the necessary number of LUTs by applying the functional decomposition [13]. In a similar manner, the RNS2Bin converter representing the mod-EVMD: Modulo Edge-Valued MDD, which is a combination of the multi-valued decision diagram (MDD) [5] and the edge-valued decision diagram (EVDD) [7], can be reduced by the functional decomposition [10].

B. Nested RNS (NRNS)

Since moduli set of RNS consists of pairwise prime numbers, we need LUTs with different sizes. Assume that an integer \( Z \) is represented by \( Z = (z_1, z_2, \ldots, z_L) \) in the RNS, and \( z_i \) is represented by the RNS recursively. Then, we have the nested RNS (NRNS) [11] which is defined by a set of \( j \) moduli \( \langle m_1, m_2, \ldots, m_{ij} \rangle \), where no pair of moduli have a common factor with any other. In the NRNS, \( z_i \) can be uniquely represented as a tuple of \( j \) integers as follows:

\[
Z = (z_1, z_2, \ldots, (z_{i1}, z_{i2}, \ldots, z_{ij}), \ldots, z_L).
\]

In the NRNS, we describe the original moduli \( m_i \) in the subscript of the right parenthesis. Since the NRNS can be applied to each element in the RNS recursively, moduli can be decomposed into smaller sizes. In this case, we must consider the dynamic range. Let \( X \) and \( Y \) be integers represented by \( X = (x_1, x_2, \ldots, x_L) \) and \( Y = (y_1, y_2, \ldots, y_L) \) in the RNS, respectively. By applying the NRNS to \( x_i \) and \( y_i \), we have \( X = (x_1, x_2, \ldots, (x_{i1}, x_{i2}, \ldots, x_{ij}), \ldots, x_L) \) and \( Y = (y_1, y_2, \ldots, (y_{i1}, y_{i2}, \ldots, y_{ij}), \ldots, y_L) \). Since the dynamic range for the NRNS is \( M_m = \prod_{i=1}^{L} m_i \), the relation \( M_m \geq x_i \circ y_i \) must be satisfied. By rewriting the output of the Bin2RNS converter \( z_i \), shown in Fig. 12 to \( (z_{i1}, z_{i2}, \ldots, z_{ij}) \), we have the Bin2NRNS converter.

Example 5.2: Consider that the RNS \( \langle 3, 4, 5 \rangle \) is represented by the moduli set \( \langle 3, 4, 3, 4 \rangle \) by the NRNS. Let \( X = 17 \) and \( Y = 3 \). In the NRNS, \( X \) and \( Y \) are represented by \( X \equiv (2, 1, 2, 2) \) and \( Y \equiv (0, 3, 0, 3) \), respectively. Thus, \( X + Y \) in the NRNS is computed by

\[
X + Y \equiv (2 \ mod \ 3, 4 \ mod \ 4, 2 \ mod \ 3, 5 \ mod \ 4) \equiv (2, 0, (2, 1)).
\]

By converting \( (2, 1) \) in the RNS \( \langle 3, 4 \rangle \) to the binary number, we have 5. Note that, \( 5 \ (mod \ 5) = 0 \). Then, we have

\[
(2, 0, (2, 1)) = (2, 0, 5) \equiv (2, 0, 0).
\]

By converting \( (2, 0, 0) \) in the RNS \( \langle 3, 4, 5 \rangle \) into the binary number, we have 20, which is equal to the result of \( 17 + 3 \).

VI. COMPARISON OF NRNS-BASED ARITHMETIC CIRCUITS WITH RNS-BASED ONES

Let \( \langle m_1, m_2, \ldots, m_L \rangle \) be the moduli set for the RNS. Suppose that the moduli set \( \langle m_1, m_2, \ldots, m_{L+1} \rangle \) where the RNS2RNS converter adds a new modulo \( m_{L+1} \). By applying the NRNS to \( m_{L+1} \), we have the moduli set for the NRNS as \( \langle m_1, m_2, \ldots, m_L, m_{L+1, 1, m_{L+1, 2}, \ldots, m_{L+1, j}} \rangle \). In this section, we analyze the number of LUTs with respect to \( m_{L+1} \). First, we estimate the number of LUTs. Let \( L_{RNS} \) be the necessary number of LUTs to realize an arithmetic circuit based on the RNS, and \( L_{NRNS} \) be that based on the NRNS. Note that, in this paper, since we assume that the RNS2RNS converter is realized by BRAMs instead of LUTs, we do not count the number of LUTs for each converter. Also, we assume a moduli set consisting of prime numbers. The number of LUTs for the RNS is estimated by

\[
L_{RNS} = \lceil k2^{2k}/L_{FPGA} \rceil,
\]

where the arithmetic circuit has \( 2k \) inputs and \( k \) outputs, where \( k = \lceil \log_2 m_{L+1} \rceil \).
where $k = \lceil \log_2 m_{L+1} \rceil$ and $L_{FPGA}$ is the number of bits for the LUT in an FPGA$^3$. We estimate the number of LUTs for the NRNS whose moduli set is \( \langle m_1, m_2, \ldots, m_L, \{m_{L+1,1}, m_{L+1,2}, \ldots, m_{L+1,j}\} \rangle \). The dynamic range for the NRNS must satisfy the relation \( (m_{L+1})^2 \leq \prod_{i=1}^{L} r_{L+1,i} \). Let $s_i \in \{0, 1\}$ be the select variable that satisfies the following:

$$r_{L+1,i} = \begin{cases} m_{L+1,i} & (s_i = 1) \\ 1 & (s_i = 0) \end{cases}$$

Then, the number of LUTs for the NRNS is estimated by

$$L_{NRNS} = \sum_{i=1}^{j} \lfloor s_i k_i 2^{k_i}/L_{FPGA} \rfloor,$$

where $k_i = \lceil \log_2 m_{L+1,i} \rceil$.

Let $G$ be the gain for the number of LUTs. Then, we have

$$G = \frac{L_{NRNS}}{L_{FPGA}}. \quad (3)$$

Expr. (3) shows that when $G > 1.00$, the NRNS-based realization requires fewer LUTs than the RNS-based one.

Fig. 13 shows the gain $G$ for $L_{FPGA} = 64$, for different value of the moduli $m_{L+1}$ in the RNS. Note that, in our implementation, we used $s_j$ that minimizes $L_{NRNS}$. Fig. 13 shows that for $m_{L+1} \geq 8$ except for $m_{L+1} = 15$, $G > 1.00$. Thus, in that case, by applying the NRNS, we can reduce the number of LUTs. Note that for $m_{L+1} = 15$, $G < 1.00$, and the NRNS-based design requires more LUTs.

$^3$In the experiment, we used an FPGA that has six-input LUTs. Thus, $L_{FPGA} = 64$

### Table I. Synthesis Options for the Xilinx Inc. Binary FFT.

<table>
<thead>
<tr>
<th>Option</th>
<th>Parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Implementation</td>
<td>Pipelined, Streaming IO</td>
</tr>
<tr>
<td>Data Format</td>
<td>Fixed Point</td>
</tr>
<tr>
<td>Input Data Width</td>
<td>8bit</td>
</tr>
<tr>
<td>Phase Factor Width</td>
<td>18bit</td>
</tr>
<tr>
<td>Scaling Options</td>
<td>Unscaled</td>
</tr>
<tr>
<td>Output Ordering</td>
<td>Bit/Digit Reversed</td>
</tr>
<tr>
<td>Complex Multipliers</td>
<td>Use CLB Logic</td>
</tr>
<tr>
<td>Butterfly Arithmetic</td>
<td>Use CLB Logic</td>
</tr>
</tbody>
</table>

### Table II. Moduli Set Using in the Implementation.

<table>
<thead>
<tr>
<th>FFT # of points, $N$</th>
<th>Moduli set</th>
</tr>
</thead>
<tbody>
<tr>
<td>1024</td>
<td>( \langle 3, 7, 11, 13, 17, 19 \rangle )</td>
</tr>
<tr>
<td>2048</td>
<td>( \langle 7, 9, 11, 13, 17 \rangle )</td>
</tr>
<tr>
<td>4096</td>
<td>( \langle 7, 11, 13, 15, 17, 19 \rangle )</td>
</tr>
<tr>
<td>8192</td>
<td>( \langle 7, 11, 13, 15, 17, 19 \rangle )</td>
</tr>
<tr>
<td>16384</td>
<td>( \langle 11, 13, 14, 15, 17, 19 \rangle )</td>
</tr>
</tbody>
</table>

Fig. 14. Comparison of the numbers of 6-input LUTs.

For $m_{L+1} = 15$, since $k = \lceil \log_2 m_{L+1} \rceil = 4$, $L_{NRNS} = \lceil k \cdot 2^{k+2}/L_{FPGA} \rceil = 16$. We used an optimal moduli set \( \langle 3, 7, 11 \rangle \) that minimizes the number of LUTs for the NRNS. In this case, $L_{NRNS} = \lceil 2 \cdot 2^{3}/L_{FPGA} \rceil + \lceil 3 \cdot 2^{3}/L_{FPGA} \rceil + \lceil 4 \cdot 2^{4}/L_{FPGA} \rceil = 21$. We use the NRNS instead of the RNS for $m_{L+1} \geq 8$, except for $m_{L+1} = 15$.

### VII. Experimental Results

We implemented the NRNS FFT (NRNS2NRNS) on the Xilinx Inc. Virtex 7 FPGA. Then, we compared it with the Xilinx Inc. binary FFT library (Binary) [18], the conventional RNS FFT (RNS) [16], and the RNS FFT with the RNS2RNS converter (RNS2RNS) [10]. Table I shows the synthesis options for the Xilinx binary FFT library. As for the RNS FFT, we chose moduli set in Table II. Note that, the input signal is represented by 8 bits, and the twiddle factor is represented by 18 bits. Using the experimental result shown in Fig. 13, we applied the NRNS to moduli set that is greater than or equal to 8 except for 15. In Table II, bold numbers show moduli set applied to NRNS.

Fig. 14 compares the numbers of 6-input LUTs, while Fig. 15 compares the numbers of 18Kb BRAMs for different number of FFT points. Since the NRNS decomposed the FFT circuit, it reduced the number of LUTs by 42.4-47.8%. Compared with the RNS FFT, the NRNS reduced the LUT count by 9.4-20.5%. Since the dynamic range is often greater than the bit range of the binary FFT, the amount of swap memory for the RNS FFT are larger than that for the binary FFTs. Fig. 15 shows that compared with the binary FFT, the NRNS-based design required 20.0-156.5% more BRAMs. Compared with the RNS FFT, the NRNS-based design required 34.1% more BRAMs. However, for large $N$, the number of LUTs is the bottleneck [8], [9], while that of the BRAMs is not. Fig. 16 compares the maximum clock frequencies. Since the proposed NRNS FFT has a smaller realization, it has a shorter critical path. Thus, the proposed one operates 9.3-41.7% faster than conventional ones.
VIII. CONCLUSION

In this paper, we proposed the FFT circuit based on the NRNS to reduce the number of LUTs. For large modulo $m$, the NRNS reduces the number of LUTs for arithmetic circuits, while for small $m$, it does not. This paper compared the NRNS with the RNS for arithmetic circuits. The experimental analysis showed that for $m \geq 8$, the NRNS-based design requires fewer LUTs except for $m = 15$. We realized the NRNS FFT on the Xilinx Inc. Virtex 7 FPGA. Since the NRNS decomposes the FFT circuit into smaller one so that the circuit is implemented efficiently by LUTs, compared with the Xilinx binary FFT, it required 42.4-47.8% fewer LUTs. Compared with the conventional RNS FFT, it required 9.4-20.5% fewer LUTs. Although the NRNS FFT requires more BRAMs, for large inputs $N$, it is not the bottleneck. As for the maximum clock frequency, since the proposed RNS FFT has a smaller realization, it operated in 9.3-41.7% higher frequency.

IX. ACKNOWLEDGMENTS

This research is supported in part by the Grants in Aid for Scientific Research of JSPS.

REFERENCES


