FPGA Implementation of high speed-low energy RNS based Reconfigurable-FIR Filter for Cognitive Radio Applications

¹C.SRINIVASA MURTHY,²K. SRIDEVI ¹GITAM Institute of Technology, GITAM Deemed to be University, Visakhapatnam, Andhrapradesh, INDIA ²GITAM Institute of Technology, GITAM Deemed to be University, Visakhapatnam, Andhrapradesh, INDIA

Abstract: - The Finite impulse response (FIR) filter is prominently employed in many digital signal processing (DSP) systems for various applications. In this paper, we present a high-performance RNS based FIR filter design for filtration in SDR applications. In general, the residue number system (RNS) gives significant metrics over FIR implementation with its inherent parallelism and data partitioning mechanism. But with increased bit width cause considerable performance trade-off due to both residue computation and reverse conversion. In this paper optimized Residue Number System (RNS) arithmetic is proposed which includes distributed arithmetic based residue computation during RNS multiplication followed by speculative delay optimized reverse computation to mitigate the FIR filter trade-off characteristics with filter length. The proposed RNS design utilizes built-in RAMs block present in the devices of FPGA to accomplish the process of reverse conversion and to store precomputational values. A distinctive feature of the proposed FIR filter implementation with core optimized RNS is to minimize hardware complexity overhead with the improved operating speed. Initially, fetal audio signal detection is carried out to validate the functionality of FIR filter core and FPGA hardware synthesis is carried out for various input word size and FIR length. From the experimental, it is proved that the trade-off exists in conventional RNS FIR over filter length is narrow down along with considerable complexity reduction with our proposed optimized RNS system.

Key-Words: - DA arithmetic, PPA, FIR filter; RNS system; Speculation, Low Energy-power product, area-power product, FPGA

Received: April 30, 2021. Revised: May 21, 2021. Accepted: May 24, 2021. Published: May 26, 2021.

1 Introduction

The most vital element of the residue number system (RNS) that has drawn significant interest is known as the Modulo multiplier and this multiplier is based on the circuit computation. For the first time, the detailed analysis of the architecture of modulo 2ⁿ-1 multipliers is presented in this research work. With help of VLSI modules that are memoryless and lookup-table i.e. through Read-only memory the implementation of modulo multipliers is carried out. On the other hand, the concept is initially used for the word length of low size, and later it can be used for the larger length of the word. To enhance the efficiency of multipliers that are memoryless, the features of residue number systems such as parallel and modular are utilized, and to accelerate multipliers, an algorithm such as Booth-encoding is employed. Additionally, for extremely high range, a highly developed 2ⁿ-1 modulo multiplier is been selected which is dependent on the redundant residue number system (RRNS). Therefore, with the application of the synopsis design compiler tool, the concepts of 2ⁿ-1 modulo multipliers have been successfully analyzed and examined for the latest modulo multipliers advancements. For various applications of signal processing namely image processing, finite impulse response (FIR) filters, communication, cryptography, discrete cosine transform, and for various applications of the digital signal processor (DSP), the 2ⁿ-1 modulo multipliers which are based on Residue number system are widely recognized as a quicker and more convenient arithmetic circuit. When compared to the typical 2's complement system, the Residue number system is more beneficial as it is a number system that is carry-

free and non-weighted. The co-prime integer moduli $\{m_{1,}, m_{2,}, m_{3,...,}, m_k\}$ that is comparative is characterized by RNS, such that whichever arbitrary integer i.e $X = \{x_1, x_2, x_3, \dots, x_k\}$, is referred to as residues of X and $x_{i=X}$ modulo $m_i (x_{i=|X|mi)}$. Forward and reverse converter, inter modulo operations, and arithmetic channels are considered as the four major components of the RNS block diagram and are represented in Figure 1. Here the forward converters and reverse converters are regarded as inter modulo operations. The conversion of residue numbers into weighted numbers is done by reverse converters and vice versa i.e. conversion of weighted numbers into residue numbers are carried out by forward converters. The arithmetic channels carry out operations such as multiplication and addition. However, the operations performed by the arithmetic channels are intra-modulo operations.

Particularly, for modulo multiplication of shorter word-length, modulo multiplier which is based on look-up-table was employed previously [12]. Since the ROM size rises dramatically with the modulus size thus, it is undesirable for relatively large moduli and this technique turns out to be unachievable. In [13] a 2ⁿ-1 modulo multiplier that is based on costeffective hardware loo-up-table is suggested and this technique makes use of cyclic convolution to carry out multiplication operations among two numbers that are decomposed and thus, reducing the size of the ROM. However, for modulus of relatively small size, the technique that is based on look-up-table is quite appropriate. In [14], a 2ⁿ-1 modulo multiplier that is memoryless is presented for the moduli which are significantly larger and medium in size and are carried out through various logic gates, adder, and multiplier. A significantly larger area is required by the Binary multiplier and correspondingly reduces the speed. In [15] a 2^{n} -1 modulo multiplier that is non-encoded and operates at high speed is suggested and is implemented without the application of binary multiplier. When compared to the exiting technique the suggested work in [15] has relatively less complexity in the circuit. Thus, to accelerate the multiplier Booth algorithm is employed and, in the work, presented in [16] a radix-2 Booth encoded multiplier is presented. On the other hand, the bits in the operands are the same as that of the number of partial products (PPs) presented in [15] and [16]. Depending on the number of added PPs, the speed of the multiplier is computed. In [17-20] a $2^{n}-1$ modulo multiplier that is based on the Booth algorithm was proposed to minimize the number of PPs. The partial products that are required for the computation of 2ⁿ-1 modulo multiplication are reduced to $\left[\frac{(n+1)}{2}\right]$ and n/2 for the radix-4 Booth encoded multiplier suggested in [17-18]. But in the case of radix-8 Booth encoded multiplier the total number of partial products are lowered to [(n-2)/3] +1 and is presented in the work proposed in [19-20] and one of the disadvantages of this work is the enhanced delay because of hard multiple existing in the Booth encoding multiplier of radix-8. In [19] the hardmultiple generators were implemented with the application of an improved structure of parallel prefix and ripple carry adders. The intra-carry propagation in the RNS [38] is avoided with the development of a redundant residual number system (RRNS) that is proposed for a high range.

2. Related Work

Residue Number System (RNS) Applications: Over the past fifty years in computer arithmetic, the RNS is considered as a substitute for the weighted two's complement number system and this system is one of the most confronting and desirable number systems. For the application of rapid computer arithmetic many aspects of residue number systems that are appropriate are considered. Some of the applications of residue number systems are cryptography and digital signal processing [5-8] where subtraction, multiplication, and addition are considered as the most important arithmetic operations and other than these applications the residue number system can detect and correct errors with the help of fault-tolerant and redundant RNS facilities presented in [3,4]. The residue number system can additionally be utilized for the transfer of data securely and without any error in the system of communication and the information is transmitted in the form of residues. As these residues present in the system are not dependent on one another as a result it can minimize the error by itself in their channel of moduli from which the information is transmitted. Therefore, for many wireless sensor networks application, these features of the residue number system are very much appropriate. For the detection and correction of an error in the system of communication, the Redundant RNS (RRNS) can be utilized by the addition of redundant moduli with an existing set of moduli.

RNS in digital signal processing: In the case of processing of digital signal applications, the Finite impulse response (FIR) filter is taken into account as one of the frequently utilized building blocks. Typically, with the application of a binary number system that experiences a significantly high

propagation delay, the implementation of FIR filters is carried out. However, a large delay is introduced in the FIR filters of higher-order due to n number of multiplications and additions. As a result, the speed of the DSP also reduces. Thus, with the help of the Residue Number System, the FIR filter speed can be enhanced as it can execute addition and multiplication without any carry propagation. In this framework, we examine the model of FIR filters that is of low complexity and linear phase. The design suggested in this work are generally employed for the applications of digital communication. One of the important and efficient techniques is dividing the filters into two sub filters namely cosine and comb filters to attain low complexity. The advantages of comb and cosine filters are that it requires fewer hardware resources and the cost required for the computation is also less but one of the disadvantages is that these filters possess reduced magnitude features. Thus, with the help of these two filters, a novel structure has been presented in this research work. When compared to the earlier system, the novel structure developed achieves improved area, power, speed, and also enhanced magnitude features which are particularly employed for sampling rate conversion of low-pass narrowband filtering. The coefficients of filters are defined with no multipliers as these components are costly in terms of power, speed, and the area thus, an appropriate technique for achieving filters with low complexity for constant coefficients filters is realized. Therefore, the work proposed in this thesis mainly concentrates on the constant multiplication's implementation as the network of shifts as well as addition. A new conceptual lower bound has been developed in this work with the factor that every arithmetic operation i.e. subtraction and addition can possess n inputs for a set of pipelined operations which are required in the Multiple Constant Multiplication (MCM) and Single Constant Multiplication (SCM) blocks and the expense of both pipelined operation and single pipelined register is same. The above requirement is vitally useful as it exists in the latest Field Programmable Gate Arrays (FPGAs) families and is considered as the most effective platform for digital signal processing algorithms implementation.

Audio processing, instrumentation or image processing, mobile communications, and many others [1]-[4] are the applications of DSP. Therefore, digital signal processing has been greatly recognized due to its applications in recent years. Approximately 7 billion subscribers to mobile communication were estimated in 2016 alone which signifies 96% of the world's population [5]. Some of the digitally transmitted systems that are used for the communications are Global Navigation Satellite System (GNSS), Digital Subscriber Line (DSL), Cell phones, satellite television, hard drives [6]-[8]. Digital filters are essentially utilized and play an important part in the above-mentioned communication systems. One of the primary functions of the digital filter is to enhance the signal quality, split the components of the signal which are initially combined, or to obtain the signal information. Therefore, because of these functions, digital signal processing is considered an important block [6],[9]-[10]. Nowadays, as the usage of battery-driven mobile devices is rapidly rising so it is necessary to maintain the battery backup for a long time [6],[10]. The number of hardware components that are required is limited due to the increased competition in portable devices for less power consumption. Therefore, to comply with the above features such as less utilization of hardware resources and power [11], the present work mainly concentrates on the advancement of novel digital filter methods. The data rates and bandwidth has been enhanced for succeeding generations in the wireless the communication system. The data rates and bandwidth links achieved by the existing method are 100 M-bit/sec and 20 MHz respectively whereas the upcoming generations are predicted to achieve the bandwidth links of 500MHz and data rates of 1 Gbit/sec [12]. However, the processing of the signal in the digital field is required to be carried out in the future where digital filters are considered as an essential feature of digital signal processing. However, the hardware processing capability is saturated by the process of filtering and the functioning of the system is considered at high rates. Moreover, the lifespan of the batteries is reduced as the digital filters are prohibitively costly (for the necessary arithmetic operations to be executed). Therefore, it is essential to develop algorithms and structures for digital filters with high performance. The digital filters must consume less power and a relatively low number of arithmetic operations, and it should be able to function at higher sampling rates so that these digital filters can be used in the system of communication. When compared to Infinite Impulse Response (IIR) filters, the Finite Impulse Response (FIR) filters possess higher-order for the same conditions of magnitude response. Hence, FIR filters are generally chosen in communications because of their assured stability, linear phase and can execute arithmetic operations in multivariate blocks because of their easy and straightforward decomposition of polyphase.

The data rates and bandwidth has been enhanced for the succeeding generations in the wireless communication system. The data rates and bandwidth links achieved by the existing method are 100 M-bit/sec and 20 MHz respectively whereas the upcoming generations are predicted to achieve the bandwidth links of 500MHz and data rates of 1 Gbit/sec [12]. However, the processing of the signal in the digital field is required to be carried out in the future where digital filters are considered as an essential feature of digital signal processing. However, the hardware processing capability is saturated by the process of filtering and the functioning of the system is considered at high rates. Moreover, the lifespan of the batteries is reduced as the digital filters are prohibitively costly (for the necessary arithmetic operations to be executed). Therefore, it is essential to develop algorithms and structures for digital filters with high performance. The digital filters must consume less power and a relatively low number of arithmetic operations, and it should be able to function at higher sampling rates so that these digital filters can be used in the system of communication. When compared to Infinite Impulse Response (IIR) filters, the Finite Impulse Response (FIR) filters possess higher-order for the same conditions of magnitude response. Hence, FIR filters are generally chosen in communications because of their assured stability, linear phase and can execute arithmetic operations in multivariate blocks because of their easy and straightforward decomposition of polyphase.

Multi-rate techniques: Many sampling frequencies are employed in the computation of digital signals and the techniques used to carry out these computations are referred to as Multi-rate systems. Thus, with the application of Multi-rate methods it is shown that the number of multipliers and adders which are necessary for the implementation is reduced [1]-[2]. Therefore, many approaches are available to enhance the multirate filters in DSP. Consider an illustration, an algorithm is introduced in [1] for FIR filters design of Mth band to enhance a polyphase system that is based on two phases for various sampling rate conversion of integer and it is illustrated that the conversion by even factors in a system is not much efficient as the odd factor conversions. In [2]-[3], a novel design technique is introduced to develop the wideband filters and differentiators which subsequently leads to a substantial decrease in the complexity. In the abovementioned novel technique, there exist twofrequency systems that are benefited by the Frequency Response Masking Technique (FRM) to achieve minimized computational complexity and to achieve sharp transition bands. Filter bank systems [4]-[7] are considered as one of the frequently used

applications for multi-rate systems. The Fast Fourier Transform (FFT) and the inverse of Fast Fourier Transform are employed in the technique [4] to accomplish computationally efficient filter banks. While, to develop a prototype filter [5], the method of Interpolated Finite Impulse Response (IFIR) is employed by the technique of trans multiplexers and cosine modulated filter bank (CMFM). In [6]-[7], a technique for enhancing the filter banks coefficients was introduced with the application of natureinspired metaheuristics. The process of interpolation and decimation is divided by the integer D into q stages to achieve computationally efficient techniques, which is the integer D is factorized in q factors. Consider an instance that when the value of q is 2 then $D = M \times R$. Initially, the structure of Cascaded Integrator-Comb (CIC) is employed through down-sampling via M and is significantly more efficient concerning chip area however, these consume high power because of the requirement of integrators at a high rate. Therefore, a decimation system that is based on a multi-stage comb is popularly employed. In the techniques proposed in [8]-[9] the q value is taken as 3 (that is $M = M1 \times M2$), and in [10]-[12] frameworks, the value of q is not less than 3, where D is limited to be 2's or 3's power. The filter of the first stage is implemented with the application of multistage designs in a non-recursive pattern and thereby reducing the power consumption at the cost of an increase in the area of the chip. It has been illustrated that the development of FIR filters was found to be efficient with the application of simple sub-filters. The filters with a smaller number of arithmetic operations and narrow transition band are achieved by the decomposition of the entire filter into simple sub filters when compared to other direct techniques. Therefore, in various applications with minimized computational complexity, these techniques are omnipresent.

The approach of FRM has drawn significant interest in the design of digital filters because of its proficiencies. Masking filters and model filters are considered as the major building blocks in FRM methodology. Due to the various zero-valued coefficients, the model filters are termed as filters with sparse coefficients or it is also referred to as sparse filters. The transition band shape of the entire filter is presented by the sparse filters along with the images of unwanted frequency responses. While these unwanted frequency responses images are eliminated with the help of masking sub-filters. In [23] latest developments carried out for the technique of FRM have been presented. The design of the technique which is based on FRM is introduced in [13] here the model filter is applied in the form of a

C. Srinivasa Murthy, K. Sridevi

hybrid thereby enabling the usage of very fewer hardware resources and minimizing the complexity of the critical path. In [14] a combined structure that is based on the optimization of the convex-concave process is presented. One of the efficient techniques that are based on sub-filters is the Frequency Transformation (FT) which is intended to devise a linear phase Type I FIR filters considering significantly trivial error in the stopbands and passbands as well as narrow transition band. The entire filter has been designed as the interconnection of cascaded identical sub-filters. The design coefficients are present in this interconnection and are seemed to be parallel to the sub-filters. The technique involves mapping of amplitude response of the prototype filter into the bands that can produce structural coefficients with the help of mapping function i.e. sub filter amplitude response. A technique used for developing Hilbert transformers were recently reported [15] where the technique of FT is applied at nesting levels and consist of very fewer multipliers. In [16], a technique for FIR filters in the combined view of frequency transformation is suggested, and the entire filter frequency response is taken into account as a function and is constituted with the help of simpler identical functions. The latest research subject matter especially helpful in comb-based decimation filters (i.e., CIC-based structures) are sharpening techniques that are used to develop filters with enhanced features of magnitude. Other than the above-mentioned sharpening techniques there are also other techniques addressed in [24] namely ACFs and recently in [28]. With the help of optimization, these ACFs can be determined and it cannot be determined with simple formulas. In [28] Saramaki-Ritoniemi introduced a suitable structure for sharpened CIC decimators which is considered as a basis for the remaining sharpened CIC decimators. Several Common Sub-Expression Elimination (CSE) techniques such as [25] describe the constants through Canonical Signed Digit (CSD), Minimal Signed Digit binary. or (MSD) representations. Thus, sub-optimal solutions are one of the disadvantages of these techniques since they are dependent on number representation.

The methodologies which are based on the graph (GB) are presented in [26] are not limited to any specific representations of numbers but it mainly concentrates on determining the intermediate sub-expressions which are required to implement least possible operations. When compared to other CSE techniques, the graph-based methodologies achieve good results by considering a significantly high number of constant realizations. On the other hand, the challenging task of these graph-based techniques

is that is it requires a large area and significantly high computational resources. Therefore, the key purpose of the proposed work is to establish a filter with the least possible number of arithmetic operations to achieve the preferred features of the magnitude. In general, the passband approximately nearer to the tolerable, and an ideal attenuation is the important feature considered for communication by the filters. Therefore, by keeping in mind these critical features, the proposed design is developed in this research work. The results obtained with the application of simple filters in computationally intensive FIR filter is more efficient, and it is observed that the combination of cosine and comb filters enhance their features by introducing a low computational complexity block. However, the comb and cosine filters are functional and consist of poor attenuation and passband droop. The features of the passband are enhanced by the application of compensator filters in cascade. Additionally, the method of sharpening is employed to improve the magnitude features of comb and cosine filters with the help of the tapped cascaded interconnection of these filters. Generally, in the case of sampling rate conversion, the multi-rate techniques are employed to minimize the arithmetic operations to be executed in terms of computational complexity [27].

To achieve MP equalizer in the technique proposed in [5], the delay D has been excluded. Therefore, the very first method is to develop an FIR equalizer by employing the same technique as proposed in [5]. Other than the technique [5], there are several other design methodologies proposed for MP FIR filters which are illustrated in [6]-[8]. On the other hand, one of the disadvantages of these techniques is that there is a requirement of multipliers for the filtering results and in digital filters, these multipliers are significantly costlier [1]. Thus, to overcome these challenges, the cascaded expanded CSCF is utilized as a prefilter to execute the complete MP FIR filter with the application of CSCFs without a multiplier. One of the normal operations considered in digital signal processing systems is multiplication with constants and this multiplication is considered a challenging task concerning area and power consumed in hardware. Therefore, with the help of subtractions, shifts, and additions, some of the operations such as Multiple Constant Multiplication (MCM) and Single Constant Multiplication (SCM) operations can be introduced [29]. In [3] the construction of critical path (maximum number of adders connected serially), theoretical lower bounds for the adders and depth levels in single constant multiplication, multiple constant multiplications, and many other constant multiplication blocks are structured with two input adders with the help of shift-and-add technique. In the case of single constant multiplication, the new bounds and tighter lower bounds such as the requirement of additional adders to maintain the depth levels low were introduced in [4]. However, in the case of constant multiplication blocks which consists of several input subtractions, additions and pipelined registers which are included in the arithmetic operations does not include any theoretical lower bounds and in the applications fields such as Field Programmable Gate Array (FPGA), the above-mentioned form of operations is considered during the implementation of pipelined constant multiplication blocks. Since the logic blocks of FPGA consists of memory elements, the computational cost of pipelining is significantly low [5]-[12]. In recent years, adders with three-input have drawn more attention, since the latest FPGAs families consist of larger blocks of logic and enable to fit of significantly more complex adders with the same hardware resources [10]-[12]. In general, several synthesis algorithms of a high level have been successfully implemented for the architecture of constant multiplication blocks without multiplier in the last twenty years. In these algorithms, the reduction in the normal cost function is specified by the number of subtractions and additions that are required to execute the multiplications. On the other hand, the consumption of power and speed is considered as the major drawback of the critical path [13]-[18]. An extensive research study has been performed on FPGAs [5]-[10], [22]-[25], and Application-Specific Integrated Circuits (ASICs) [20]-[21]. The key purpose of the framework is to minimize the number of arithmetic operations which in turn reduces the depth levels.

Related work for Adders and Multipliers: Previously, multipliers of a different form that are efficient have been established. With the application Booth algorithm, Shen and Chen [1] introduced a low power consumption multiplied by minimizing the switching movements of PP (partial product). A low power 2's complement multiplier was proposed by Chen et al. [2]. The PPs switching activities are minimized with the help of the Radix-4 Booth Through the left-right algorithm, a algorithm. multiplier with fixed-width has been proposed by Wang et al. [2]. One of the disadvantages of this fixed-width multiplier is that the dissipation of power is significantly high with the cost and one of the advantages is that it minimizes PP to enhance the multiplication speed. A structure of linear array multiplier that is designed with the integration of various approaches is introduced by Zhijun and Ercegovac [3]. The drawback of this design is that the delay and area required are significantly more but this design is power efficient. A design of the multiplier with relatively low power and speed of operation that is significantly high is been introduced by Chen and Chu [4-5] on compression tree of the multiplier and Booth decoder through the technique of spurious power suppression (SPST). However, the area consumption is more and is considered as one of the drawbacks of this proposed design. For the addition of PPs, ripple carry adder and carry-look-ahead adder is utilized to examine two 32-bit multipliers introduced by Krad and Taie8. BZ-FAD for the shiftadd multipliers proposed by Dastjerdi et al. [6] minimizes the switching performance and thereby reduces power. An 8-bit multiplier with high speed and low power consumption that is based on an algorithm of pair-wise has been proposed by Saha et al. [7]. However, because of the application of wave pipelining the overall efficiency of the proposed model is high. For applications that are based on error-tolerant, an approximate multiplier with a short critical path and power-efficient has been introduced by Liu et al. [8]. For applications such as digital signal processing, a radix-4 Booth multiplier with modified probabilistic estimation bias (PEB) has been introduced by Mohanty and Tiwari [9]. A PEB multiplier for an appropriate adder is also been suggested. A squarer and a 16-bit Booth multiplier with a fixed width that is having efficient area and power is been proposed by Shao and Li [10]. To manage the squarer and multiplier structure, a specific design of the array has been presented and is based on approximate arithmetic Computing (AAAC). A Booth multiplier with fixed width and high speed that is based on the conditional probability of input series (CPIS) has been presented in [11]. One of the advantages of this multiplier is that it minimizes computational cost and enhances the operation speed. With the help of the Dadda algorithm, a multiplier with high speed and low power has been introduced by Shabbir et al.[12]. In the application of multipliers, full adders are recommended. By employing flip-flops in multipliers, errors are minimized in the output.

3. Proposed Methodology

RRNS system for FIR filter design: Some of the applications of residue number systems are cryptography and digital signal processing where subtraction, multiplication, and addition are considered as the most important arithmetic operations, and other than these applications the residue number system can detect and correct errors

with the help of fault-tolerant and redundant RNS facilities. The residue number system can additionally be utilized for the transfer of data securely and without any error in the system of communication and the messages are transmitted in the form of residues. As these residues present in the system are not dependent on one another as a result it can minimize the error by itself in their channel of moduli from which the information is transmitted. Therefore, for many wireless sensor networks application, these features of the residue number system are very much appropriate. For the detection and correction of an error in the system of communication, the Redundant RNS (RRNS) can be utilized by the addition of redundant moduli with an existing set of moduli. In the case of digital signal processing applications, the Finite impulse response (FIR) filter is considered as one of the most frequently utilized building blocks. Typically, with the application of a binary number system that experiences a significantly high propagation delay, the implementation of FIR filters is carried out. However, a large delay is introduced in the FIR filters of higher-order due to n number of multiplications and additions. As a result, the speed of the DSP also reduces. Thus, with the help of the Residue Number System, the FIR filter speed can be enhanced as it can execute addition and multiplication without any carry propagation. Intra-modulo is the operation performed by modulo multiplication. Therefore, propagation of carry does not take place between modulo channels. A redundant encoding has been developed to avoid intra-carry propagation from arithmetic operations in residue number systems as it is suitable for significantly high dynamic range. Thus, to achieve an effective modulo multiplier in the application of digital signal processing, the redundant residue number system (RRNS) can be studied.

In this paper, we propose DA based speculative RNS MAC unit for FIR filer design which has metrics as follows:

- DA based residue computation unit will offer considerable complexity reduction over a wide range of moduli set.
- It can be used for higher-order FIR design to solve tradeoff constraints related to the conventional RNS system.
- Speculative reverse conversion at the final stage can enhance the accumulation speed with direct RAM-based data accessibility.
- RNS-Ternary PPA is mainly to optimize area as shown in Fig.5.

3.1 RNS System

the Residue number system, In arithmetic computation is carried out using a predefined moduli set, which comprises prime integers as the moduli. The input operands range accommodated by the RNS number system without causing any truncation in the results is statistically formulated based on moduli set values and the total elements used as moduli's. RNS system for given dynamic range M can accommodate the results in the range of [0, M - 1] irrespective of the arithmetic used. In addition to this, during RNS computation each moduli and associated computations are carried out as isolated channels in L parallel paths which reduce the path propagation delays considerably. Moreover, this path delay is further optimized using modified parallel prefix adder topology-based accumulation within the RNS system in the proposed method in existing designs and the detailed information's and its design part are discussed in [14] as shown in Figure.1 [13,14]. Moduli conversion: Considered modulis $\{m_1, \dots, m_n\}$ m_2, \dots, m_p and its associated residue $\{r_1, r_2, \dots, r_p\}$ are related to the input operands X as given below equations:

$$r_i = |X|m_i$$
 and $|X|M = |\sum_{i=1}^n r_i|M_{i-1}|_{m_i} M_i|M$ -(1)

Where M is the product of all modulis, and $M_i=M/m_i$. This can be rewritten as:

 $X_i = r_i * \{0, 0, ..., 1, ..., 0, 0\} = r_i * X_i ..., ..., (3)$ Where X_i is computed in such a way that $|X_i| m_i = 1$. The Equation that relates r_i and its inverse r_i -1 is as follows:

According to the dynamic range evaluated before the computation both input operands and moduli's are selected within the required range and finally, modulo reduction is performed on both sides of the Eqn. (7).

Where **k** is length of filter, in this design, the k is 0 to 63. Equation (1) proposes FIR filter operation with direct structure which requires less number of registers and its internal architecture is shown in Fig.2. The Fig.2. is a sub part in the RNS based RFIR which is shown in Fig.1. In FIR filter is most commonly used in DSP applications, the traditional FIR filter design uses binary number system for design of adders and multipliers which leads to larger propagation & net delays and it limits the speed of operations. To address these limitations, the proposed RNS based FIR filter given in equation (1) which uses faster modified Parallel Prefix Adder (PPA) where carry bit propagation is avoided, the results of existing PPA and modified PPA are shown in Table.1.

 $M_{i} \text{ is defined as } \frac{M}{M_{i}}, \text{ where } M$ $= \prod_{i=1}^{k} P_{i}, \text{ then } ||M_{1-i}||m_{i}M_{i}|m_{i}$ = 1.....(5)All m_{i} 's are relatively prime, the inverses exist: $X_{i} = |\overline{M_{i-1}}|m_{i}\overline{M_{i}}, \dots \dots \dots \dots (6)$ X $= \sum_{i=1}^{n} X_{i}$ $= \sum_{i=1}^{n} r_{i}|\overline{M_{i-1}}|m_{i}\overline{M_{i}} \dots \dots \dots \dots (7)$ $\vdots \quad \text{to the dynamic range evaluated before th$

According to the dynamic range evaluated before the computation both input operands and moduli's are selected within the required range and finally, modulo reduction is performed on both sides of the Eqn. (7).

The basic FIR filter function is given by $y(z) = \sum_{j=0}^{M-1} x(j)h^{-j} = \sum_{j=0}^{M-1} x(n)h(n-k) - - - (1)$

Where **k** is length of filter, in this design, the k is 0 to 63. Equation (1) proposes FIR filter operation with direct structure which requires less number of registers and its internal architecture is shown in Fig.2. The Fig.2. is a sub part in the RNS based RFIR which is shown in Fig.1. In FIR filter is most commonly used in DSP applications, the traditional FIR filter design uses binary number system for design of adders and multipliers which leads to larger

propagation & net delays and it limits the speed of operations. To address these limitations, the proposed RNS based FIR filter given in equation (1) which uses faster modified Parallel Prefix Adder (PPA) where carry bit propagation is avoided, the results of existing PPA and modified PPA are shown in Table.1.

3.2 Memory efficient post computation

After residue computation which follows conventional arithmetic operation reverse conversion is performed as post computation to convert the residue number to an integer value. This process involves numbers with a maximum range of values depends on the size of each moduli. Here all possible results are pre-computed and stored in memory as readily available blocks to perform the reverse conversion process. It will reduce the hardware complexity of the reverse conversion unit since all these memory units are transformed into dedicated on board block RAMs during hardware synthesis. As compared to other computations this on-chip memory access not only minimize the computation cost the access time is also minimized with the least critical path as shown in Fig.6.

Algorithm: Proposed Ternary-RNS based FIR filter

Input: Sampled Audio Signal, Moduli sets $(m_1, m_2, m_3, \dots, m_n)$

Output: Filtered sampled audio signal y(n-k) **Start:**

Process 1: RNS Computation

////Define the moduli sets Module set (7,8,9)

Input operands: A=14; and B=20; /////for RNS multiplier

Input to FIR filter: Sampled Audio Signal (x)

Process 2: Forward Conversion for RNS computation

Let Ar_1 is modulo operation between one of the operand and moduli set

 $Ar_1=(A) \mod (m_1); Ar_2=(A) \mod (m_2); Ar_3=(A) \mod (m_3); Br_1=(B) \mod (m_1); Br_2=(B) \mod (m_2); Br_3=(B) \mod (m_3);$

Compute above Ar_1 Ar_3 to get Input Residual 1= (0, 6, 5) and Br_1 Br_3 to get Input Residual 2 = (6, 4, 2).

Process 3: Residue Computation for Multiplication, where * is RNS based multiplier

Compute $r_1 = (Ar_1 * Br_1)mod(m_1)=0$, Compute $r_2 = (Ar_2 * Br_2)mod(m_2)=0$

Compute $r_3 = (Ar_3 * Br_3)mod(m_3)=1$, Output of the process 3 i.e Residue Components are (0,0,1)

Process 4: Residue Computation for Adder (+) where + is Ternary-RNS based adder

Compute $r_1 = (Ar_1 + Br_1)mod(m_1)=6$, Compute $r_2 = (Ar_2 + Br_2)mod(m_2)=2$,

Compute $r_3 = (Ar_3 + Br_3)mod(m_3)=7$

Output of the process 4 i.e Residue Components are (6,2,7)

Process 5: Reverse Conversion for RNS computation

Inputs: Moduli set and $invM_{1},\ invM_{2,}$ $invM_{3}$

Outputs: ROM1, ROM2 and ROM3 Where ROM1 stores computational of $(M_1 * invM_1)mod(m_1)=1$, Where ROM2 stores computational of $(M_2 * invM_2)mod(m_2)=1$, Where ROM3 stores computational of $(M_3 * invM_3)mod(m_3)=1$ end: $\begin{array}{l} \text{Compute } M_1 = m_2 \ast m_3, \text{ Compute } \\ M_2 = m_1 \ast m_3, \text{Compute } M_3 = m_1 \ast m_2 \\ \text{Compute } do \text{ iterative process till to get } \\ (M_n \ast invM_n)mod \ (m_n) = 1 \\ |M_1 \ast invM_1|mod \ (m_1) = 1, \\ M_2 \ast invM_2|mod \ (m_2) = 1, \\ |M_3 \ast invM_3|mod \ (m_3) = 1 \\ \text{At } M_1 = 72, \\ M_2 = 63 \text{ and } M_3 = 56 \\ \text{we will get above three equations output is } 1 \end{array}$

Process 6: Final Reverse Conversion for RNS computation

RNS output = $(|M_1 * invM_1 * r_1| + |M_2 * invM_2 * r_2| + |M_3 * invM_3 * r_3|)mod(m_1 * m_2 * m_3)$ RNS output=|280|mod 504=280 i.e 14*20=280 hence it is proved









3.3 DA based residue computation using approximated speculation

The inclusion of speculation during accumulation allows overall path delay propagation with appropriate carry approximation. Here it is incorporated to perform multiplication as a sequential of addition as given in [10]. Here multiplier less DA arithmetic includes with most appropriate preprocessing units to narrow down the carry propagation path and this will keep the critical path delay as constant using prior computations. Without using any inner stage pipelining units accumulation speed is increased with inherent metrics of low complexity. Here MAC unit designed with speculative delay optimized accumulation unit helps to meet the demand requirements of FIR filter design, and reduce the performance penalty gap that arises with FIR tap extension. Moreover, speculative units compute accumulation in identical blocks as parallel compared to conventional prefix computation methods which lead to significant path delay optimization is shwon in Fig.3. Implementation of FIR filter using Residue Number System (RNS) with DA based arithmetic has the following advantages: improved data rate with inherent parallelism, modularity, and path optimized speculation operation. Here the entire MAC operation is performed using RNS in each FIR tap as shown in Fig.2. MAC is the core operation used in FIR filters in which speculation features of prefix topology with the RNS system beneficial for high implementation [9].



Fig.3. Internal Architecture of proposed sub-filter RFIR filter for module n_i^1

3.4 Delay optimization

The major limitations come out with improved FIR impulse response is mitigated with carry approximated accumulation and speculation driven reverse conversion in RNS multiplication. Moreover, error propagation in DA based sequentially stagewise shifting operation during FIR computation is solved with an error correction unit.



Fig.4. Proposed Internal micro-architecture of one coefficient and audio signal of RNS-Ternary PPA based RFIR filter for high speed and low energy consumption in SDR application.

Both error correction and reverse computation in the RNS system play a prominent role in overall performance in terms of complexity reduction as well as critical path reduction as shown in Fig.4. And the performance metrics of DA based residue computation both in terms of complexity reduction and performance retention is increased with the FIR order as shown in Table.3.

4. Results and Discussion

Experimental results are carried out over different sets of moduli's to validate the performance metrics of the proposed RNS FIR design. Here digital design is carried out using Verilog HDL and its functionality is verified using ModelSim simulation and metrics are evaluated using Artix-7 Development FPGA hardware for synthesis and place & route, the obtained synthesis report and device utilizations are shown in Table 1 and Table 2. From the experimental results, it is well proved speculative DA model exploits gives considerable path delay reduction with significant resource optimization level. Hardware complexity is evaluated as logical element utilization during hardware synthesis which showed the superiority of speculative DA based RNS system in various dynamics of FIR filter characteristics.

Table 1. Performance trade off comparison over input word length

Input word lengt h size	Moduli set(2n+ 1,2n,2n -1)	RN specul conve re comp	S with ation and entional sidue outation	RNS with speculation and residue computation		
		Area (LEs)	Fmax	Area(LEs)	Fmax	
8 bit	(7,8,9)	1396	57.3MH z	396	75.48 MHz	
16 bit	(31,32, 33)	1823	24.96M Hz	1263	29.78 MHz	

Table 2. Performance analyzes of speculativeDA based RNS FIR design.

FIR lengt h	RN spec conv re com	IS with culation and rentional esidue putation	RNS with speculation and DA based residue computation		
	Are a (LE s)	Fmax	Area(LE s)	Fmax	
8 tap	2096	63.46M Hz	2009	209.600M Hz	
16 tap	8623	57.3MH z	8193	356.125M Hz	



(a) Hardware complexity overhead



(b) Performance area gap

Fig.5. Improved performance trade off comparison of DA based arithmetic in RNS over FIR length.

Table 3. Performance analyzes of speculative DA based RNS FIR design in terms of energy, power and product of area, delay and power

Multiplier	Slices (area)	LUT	Delay (ns)	Power (mW)	Area* delay	Time* power	Area*time *power
Two speed Radix-4 Serial-Parallel multiplier for 32 bit [19]	1590		26.820	86.6	7186x10 ⁻⁶	143.27 x10 ⁻⁹	22.852 x10 ⁻ 9
Booth Serial- Parallel Multiplier for 16 bits[19]	1200		27.2	0.85	19.06 x10 ⁻⁶	23.13 x10 ⁻⁹	16.19 x10 ⁻⁶
Modified Shift-Add Multiplier for 16 its [19,23]	2107		20.51	0.1	21.07	25.1	52.8
Bawooley1 multiplier	10475		10.25	22.62	1.07368x10 ⁻ 04	1.691x10 ⁻¹⁰	2.42x10 ⁻⁰⁶
Wallace tree multiplier	111		8.51ns	16.5	9.4461x10 ⁻⁷	1.40x10 ⁻¹⁰	1.55x10 ⁻⁰⁸
Proposed Ternary-RNS based FIR design for SDR	174	249	4.771	0.088	8.30 x10 ⁻⁶	41.9 x10 ⁻⁹	7.3 x10 ⁻⁶



Fig.6. Overall Device utilization comparison between proposed ternary-RNS based FIR filter design for SDR applications



Fig.7. Simulated results for Audio samples and after filtering reconstructed output

Fig.7. shows that audio signal for SDR applications using methodology is presented to validate the efficiency FIR filter design with extended length and its functionality. The filter process is carried out using high-precision filter coefficients with improved frequency response and accurate filtering operations.

5. Conclusions

An extensive analysis of various types of 2^{n} -1 modulo multipliers without memory is presented in this research work and this paper is evaluated by implementing the Booth algorithm so that the efficiency of 2^{n} -1 modulo multipliers is enhanced by limiting the number of PPs and this number is limited to one third for radix-8 Booth-encoded multiplier. However, this multiplier is not much use for a smaller dynamic range because of the existence of hard-multiple, but for a higher dynamic range, it is more efficient in terms of power and area. Apart from the

above-mentioned modulo multiplier employed in the residue number system, a redundant encoding has been developed to avoid intra-carry propagation from arithmetic operations in residue number systems as it is suitable for significantly high dynamic range. However, still there exists a scope for advanced study in the area of 2^{n} -1 modulo multipliers. In future research work, we consider the structure of modified Booth-encoded multiplier with higher-radix by utilizing hard-multiple generation. Further, to achieve an effective modulo multiplier in the application of digital signal processing, the redundant

residue number system (RRNS) can be studied. This paper focused on the implementation of high-end FIR filters using optimized RNS units for the fetal ECG signal detection process. The hardware synthesis results presented in this work proved that each level of optimizations carryout in RNS computation has a direct impact on hardware rate and performance retention of the FIR filter design. Here both the RAM-based speculative reverse conversion and DAbased residue computation are used for path delay reduction in the RNS system which can able to reduce the performance penalty gap in FIR filter design. This work restores consistent performance metrics with the extension of the FIR filter tap by incorporating an optimized RNS MAC and memory-efficient reverse converter unit. Based on Table.3, the obtained results in terms of area, delay, power, Area & time product & power product and Area & delay product are compared with existing works and it is found that there are 35% optimization in power consumption, 15% improvement in Area & time product, 23% improvement in Area & delay.

References

- N. P. Smart, F. Vercauteren, "Fully homomorphic encryption with relatively small key and ciphertext sizes" In Public Key Cryptography - PKC'10, Vol. 6056 of LNCS, pp. 420-443, Springer, 2010.
- [2]. Oscal T C. Chen, Sandy Wang and Yi-Wen Wu, Minimization of switching activities of partial products for designing low-power multipliers, IEEE Trans. Very Large Scale Integration (VLSI) Systems. 11 (2003) 418–43
- [3]. Jinn-Shyan Wang, Chien-Nan Kuo, and Tsung-Han Yang. Low-power fixed-width array multipliers, in Proc. IEEE Int. Symp. Low Power Electronics and Design (ISLPED04), 2004, pp. 307–312.
- [4]. Huang Zhijun and Milos D. Ercegovac, Highperformance low-power left-to-right array multiplier design, IEEE Trans. Computers. 54 (2005) 272–283
- [5]. Kuan Hung Chen and Yuan Sun Chu, A low-power multiplier with the spurious power suppression technique, IEEE Trans. Very Large Scale Integration (VLSI) Systems. 15 (2007) 846–850
- [6]. Hasan Krad and Aws Yousif Al Taie, Performance analysis of a 32-bit multiplier with a carry look ahead adder and a 32-bit multiplier with a ripple adder using VHDL, J. Computer Science. 4(4) (2008) 305–30
- [7]. M. Mottaghi-Dastjerdi, A. Afzali-Kusha, and M. Pedram, BZ-FAD: A low-power low area multiplier based on shift-and-add architecture, IEEE Trans. Very Large Scale Integration (VLSI) Systems. 17 (2009) 302–306

- [8]. Aloke Saha, Dipankar Pal and Mahesh Chandra, Low-power 6-GHz wave-pipelined 8b × 8b multiplier, IET Circuits, Devices & Systems. 7(3) (2013) 124–140.
- [9]. Cong Liu, Jie Han and Fabrizio Lombardi, A lowlower, high-performance approximate multiplier with configurable partial error recovery, in Proc. IEEE Design, Automation and Test in Europe Conf. and Exhibition (DATE), (2014), pp. 1–4.
- [10]. Basant K. Mohanty and Vikas Tiwari, Modified PEB formulation for hardware efficient fixed-width Booth multiplier, J. Circuits Syst. Signal Process. 33 (2014) 3981–3994.
- [11]. Botang Shao and Peng Li, Array-based approximate arithmetic computing: A general model and applications to multiplier and squarer design, IEEE Trans. Circuits and Systems-I: Regular Papers. 62 (2015) 1081–10
- [12]. When-Quan He, Yuan-Ho Chen and Shyh-Jye Jou, Dynamic error compensated fixed width Booth multiplier based on conditional-probability of input series, J. Circuits Syst. Signal Process. 35 (2016) 2972–2991
- [13]. Zain Shabbir, Anas Razzaq Ghumman and Shabbir Majeed Chaudhry, A reduced-sp D3Lsum adder based high frequency 4 × 4 bit multiplier using Dadda algorithm, J. Circuits Syst. Signal Process. 35 (2016) 3113–313
- [14]. Ahmad Hiasat.et.al, "A Scaler Design for the RNS Three-Moduli Set {2ⁿ⁺¹-1, 2ⁿ, 2ⁿ-1}Based on Mixed-Radix Conversion", Journal of Circuits, Systems, and Computers, Vol. 29, No. 3 (2020) 2050041 (12 pages), World Scientific Publishing Company, DOI: 10.1142/S021812662050041
- [15]. Raj Kumar.et.al, "Perspective and Opportunities of Modulo 2n-1 Multipliers in Residue Number System: A Review", Journal of Circuits, Systems, and Computers Vol. 29, No. 11 (2020) 2030008 (24 pages), World Scientific Publishing Company, DOI: 10.1142/S021812662030008
- [16]. Grande Naga Jyothi.et.al, "ASIC implementation of distributed arithmetic based FIR filter using RNS for high speed DSP systems", International Journal of Speech Technology, Spinger, 2020, <u>https://doi.org/10.1007/s10772-020-09683-1</u>
- [17]. Rami Akeela.at.al, "Software-defined Radios: Architecture, State-of-the-art, and Challenges", INTERNET OF THINGS RESEARCH LAB, DEPARTMENT OF COMPUTER ENGINEERING, SANTA CLARA UNIVERSITY, USA — MARCH 2018
- [18]. Lamjed Touil.et.al, "Design of Low-Power Structural FIR Filter Using Data Driven Clock Gating and Multibit Flip-Flops", Hindawi, Journal of Electrical and Computer Engineering Volume 2020, Article ID 8108591, 9 pages, https://doi.org/10.1155/2020/8108591
- [19]. Lamjed Touil,et.al, "Design of Low-Power Structural FIR Filter Using Data-Driven Clock

Gating and Multibit Flip-Flops", Hindawi Journal of Electrical and Computer Engineering Volume 2020, Article ID 8108591, 9 pages, https://doi.org/10.1155/2020/8108591

- [20]. C. Efstathiou.et.al, Modified Booth modulo 2ⁿ -¹ multipliers, IEEE Trans. Comput. 53 (2004) 370– 374.
- [21]. R. Muralidharan.et.al, Radix-8 booth encoded modulo 2ⁿ -¹ multipliers with adaptive delay for high dynamic range residue number system, IEEE Transactions Circuit Systems. International Regular. Pap. 58 (2011) 982–993
- [22]. R. Muralidharan.et.al, Area-power efficient modulo 2ⁿ-¹ and modulo 2n þ 1 multipliers for 2ⁿ-¹, 2ⁿ, 2ⁿ⁺¹ based RNS, IEEE Trans. Circuits Syst. I Regul.Pap. 59 (2012) 2263–2274
- [23]. R. Muralidharan.et.al., Radix-4 and Radix-8 booth encoded multi-modulus multipliers, IEEE Trans. Circuits Syst. I Regul. Pap. 60 (2013) 2940–2952.
- [24]. H. Pettenghi.et.al, Efficient method for designing modulo {2n+k} multipliers, J. Circ. Syst. Comp. 23 (2014) 1450001
- [25]. Romero, D.E.T. "High velocity multiplierless Frequency Response Masking (FRM) FIR channels with diminished use of equipment assets," IEEE International Midwest Symposium on Circuits and Systems (MWSCAS), pp. 1-4, 2015.
- [26]. Lu, W.- S. what's more, Takao H., "A bound together way to deal with the plan of added and recurrence reaction concealing FIR channels," IEEE Transactions on Circuits and Systems I – Reg. Papers, 2016.
- [27]. Demirtas, S. what's more, Oppenheim A. V., "A useful sythesis way to deal with channel honing and secluded channel plan," IEEE Transactions on Signal Processing, 2016.
- [28]. Candan, C. "Ideal Sharpening of CIC channels and a proficient usage through Saramaki-Ritoniemi obliteration channel structure," 2011. http://www.eee.metu.edu.tr/~ccandan/bar dir/pick honed CIC filt broadened new.pdf. (keep going access on February 2017)
- [29]. Molnar G., Dudarin A. what's more, Vucic M. "Minimax plan of multiplier less honed CIC channels dependent on span examination," IEEE Internat. Show on Information and Communication Technology, Electronics and Microelectronics (MIPRO), May 2016.
- [30]. Aksoy, L., Costa, E., Flores, P. and Monteiro, J. *Multiplierless design of linear DSP transforms*, in VLSI-SoC: Advanced Research for Systems on Chip, Springer, Chap. 5, pp. 73–93, 2012.
- [31]. Meyer-Baese, U. Digital Signal Processing with Field Programmable Gate Arrays, Springer, 2014.

Creative Commons Attribution License 4.0 (Attribution 4.0 International, CC BY 4.0)

This article is published under the terms of the Creative Commons Attribution License 4.0 https://creativecommons.org/licenses/by/4.0/deed.en_US