Effect of Finite Word-Length on SQNR, Area and Power for Real-Valued Serial FFT

Nanda K. Unnikrishnan
University of Minnesota
Minneapolis, USA
unnik005@umn.edu

Mario Garrido
Linköping University
Linköping, Sweden
mario.garrido.galvez@liu.se

Keshab K. Parhi
University of Minnesota
Minneapolis, USA
parhi@umn.edu

***Abstract*—Modern applications for DSP systems are increasingly constrained by tight area and power requirements. Therefore, it is imperative to analyze effective strategies that work within these requirements. This paper studies the impact of finite word-length arithmetic on the signal to quantization noise ratio (SQNR), power and area for a *real-valued serial* FFT implementation. An experiment is set up using a hardware description language (HDL) to empirically determine the tradeoffs associated with the following parameters: (i) the input word-length, (ii) the word-length of the rotation coefficients, and (iii) length of the FFT on performance (SQNR), power and area. The results of this paper can be used to make design decisions by careful selection of word-length to achieve a reduction in area and power for an acceptable loss in SQNR.**

***Index Terms*—FFT, real-valued FFT, serial commutator, area, power, SQNR**

I. Introduction

The fast Fourier transform (FFT) is one of the fundamental operations in digital signal processing. It is widely used in applications such as OFDM, spectral analysis and in extracting features for classification. Many biomedical signals such as electroencephalogram (EEG), magnetoencephalogram (MEG) and functional magnetic resonance imaging (fMRI) are real. Applications where these signals are used include seizure prediction and detection using EEG [1]–[4], schizophrenia identification from MEG [5] and classification of mental disorders from fMRI [6]. In recent years, there has been significant interest in developing optimized solutions for implementations targeting real-valued signals [7]–[11]. The development of these architectures in more advanced technologies has been driven by tighter constraints on the area and power requirements of the system. This has led to an interest in reducing the number of computations in the system or in improving the utilization of hardware functional units.

Analysis of scaling and roundoff noise is an important consideration in digital signal processing systems [12]. Prior work on understanding SQNR in FFT primarily has focused on analyzing the quantization noise and propagation of the noise through the system [13]–[16]. These quantization noises are introduced due to finite precision during implementation. In finite word-length implementations, the error is introduced due to truncation at the outputs of the adders and multipliers to maintain a constant word-length. Although prior publications compute SQNR for an algorithm, they do not provide accurate SQNR measurements corresponding to the *hardware* implementation. Also, they propose methodologies for improvement of the SQNR of the design without providing the cost of these improvements in terms of area and power. Despite focusing on results that are obtained from actual implementations, the approaches in [17]–[21] do not present a holistic view of the entire system allowing the designer to make an informed choice based on all parameters provided to them. These implementations have only explored a very small region of the design space that is limited to the range of the specific application targeted, i.e., how the word-length would affect the SQNR of the design, without investigating the cost of these improvements. Upto date the only work that has taken this into consideration is [22]. Moreover, the applications considered in these papers are not based on real-valued FFTs.

This research was supported in part by the National Science Foundation under grant number CCF-1814759.

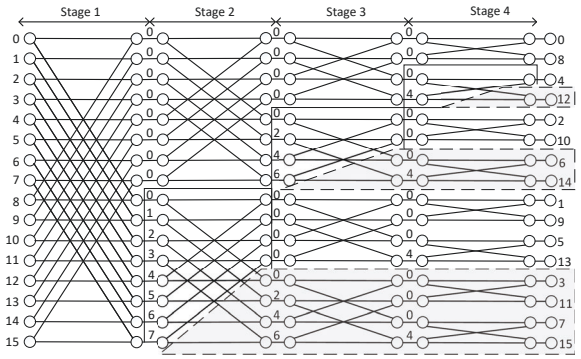


Fig. 1. Flow graph 16-point real-valued FFT from [23].

In this paper, we bridge the above gap by presenting finite word-length tradeoff analysis associated with SQNR, area and power for a specific *real-valued* FFT architecture. The results we provide are more representative of real-world applications as the architecture is designed, synthesized and simulated in hardware to obtain accurate measurements over a larger design space. The real-valued *serial* FFT architecture is designed based on the architecture in [23], which achieves full hardware utilization and requires lower area. The hardware model is designed to be synthesizable and accounts for the added overheads for the control logic and pipelines to ensure that it meets the timing requirements. We present exhaustive simulation

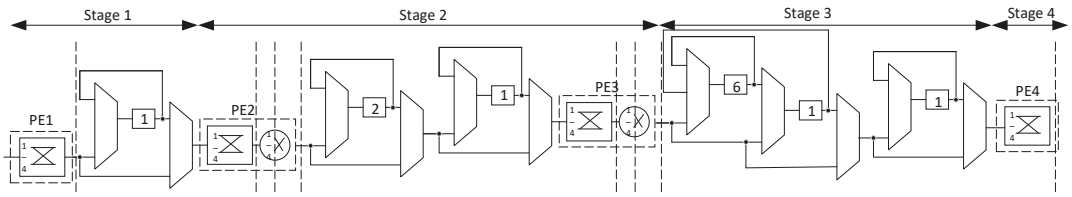


Fig. 2. Architecture of 16-point real-valued FFT from [23]. The numbers inside the squares boxes represent delay elements.

results for all parameters by synthesizing the designs to obtain the area and power consumption values. These results can be used to make informed choices while implementing real-valued serial FFT designs and explore the tradeoffs for a specific application.

This paper is organized as follows. Section II presents the real-valued FFT, its architecture, analysis and implementation. Section III presents the experimental results for SQNR, area and power consumption. In Section IV we summarize the main conclusions.

II. Real-valued FFT architecture

A. Advantages of real-valued FFT

The real-valued FFT eliminates redundant operations by using the property that the spectrum of the signal is conjugate symmetric, i.e., $X[n-k] = X^*[k]$ [24]. This can be observed in Fig. 1 where the redundant calculations are highlighted in a standard decimation in frequency (DIF) data-flow graph. As a consequence of these eliminations in the radix-2 architecture, we can achieve a canonic form of the flow graph, where at each stage the number of signals remains constant.

B. Qualitative analysis

This subsection describes the effects of the following parameters: (i) the word-length of the data, W_D, (ii) the word-length of the coefficient for twiddle factor multiplication, W_C, and (iii) the FFT length, N.

Most FFT implementations use fixed-point precision for the input data and intermediate datapaths. Noise is introduced due to the quantization of the intermediate outputs due to fixed word-length constraints. Therefore, for a proper examination of the SQNR of the system, we must determine all the sources of quantization noise from the architecture. First, the single addition operation in a butterfly module results in the growth of the word-length by one bit; however, to maintain constant word-length we truncate the last bit that introduces noise into the system. Second, the multiplication of the data (W_D) and the coefficient (W_C) in the rotator results in output word-length of $W_D + W_C - 1$. This must be truncated as well to keep the word-length of the datapath fixed. Two consequences arise due to this. First, a change in W_C results in quantization noise being introduced in the rotator. Second, W_D determines the overall SQNR of the system as it affects all modules, and limits the potential benefits obtained by other means.

Based on the implemetation in [23], we see that the number of adders and multipliers increases by a factor $\log_2(N)$; however, the architecture requires $N+9\log_2(N)-19$ memory units. This shows that in longer FFTs the memory units contribute significantly to the area and power of the system. Analyzing the effects of W_C, we see that it is limited to the storage of the coefficients and the width of the multiplier outputs. The multipliers in the design in subsequent datapaths truncate the multiplier output to the width of W_D. Also, an increase in the W_c results in an increase in area to store the coefficients for rotation.

C. Implementation

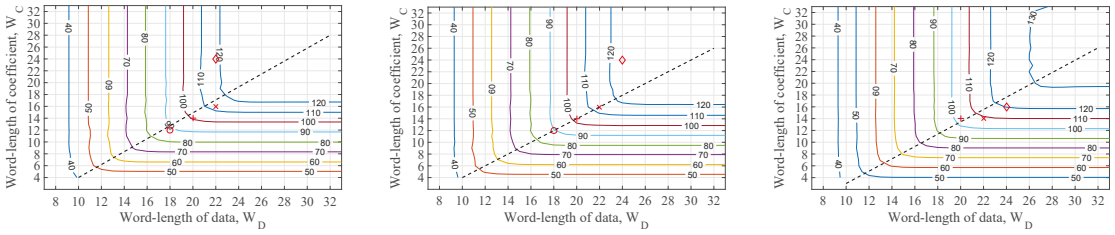
In this paper, we implemented the design based on the architecture presented in [23], making use of the butterfly module, data permutation circuits, and the rotator module as shown in Fig. 2. The architecture allows for the insertion of pipeline stages, as shown in Fig. 2, to support the target frequency of 500 MHz using a 45 nm technology node. Each stage's control logic is implemented in a parametrized form and in higher stages form the critical path of the design. Therefore, these are also carefully pipelined along with the adders and multipliers to ensure that the design meets the required specification. The work in [15] suggests word-length growth as a method to increase the SQNR of the system; however, in this paper, we focus on a fixed word-length throughout the entire design. We are required to scale the output of the design as the truncation in the butterfly module results in the output being scaled up by a factor of 2 at each stage. Thus an input which was defined in the fixed-point notation, $-1 \leq x < 1$, can represent any number in the range of $-N \leq X < N$. However, since the maximum value that an FFT can take is N we require to scale down the input by a factor of 2 after each stage to prevent overflow.

III. Experimental analysis

This section describes the experimental methodology as well as how the three main parameters under consideration: word-length of the data, W_D, the word-length of the coefficient, W_C, and FFT length, N, impact the characteristics of the design. Several designs are analyzed with respect to area, power and SQNR, by varying these individually and in combinations.

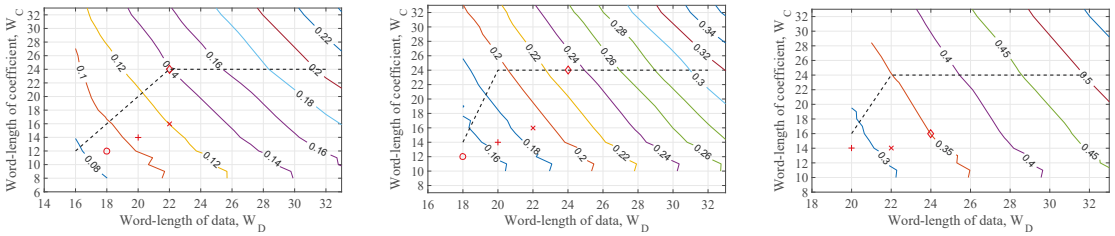
A. Methodology

The experiments were conducted on a SystemVerilog based synthesizable RTL, parametrized by the word-length of the data, the word-length of the coefficient and the length of the FFT. The design was synthesized using a 45 nm technology node at 1.1 V and 500 MHz. These were chosen to give a good reference setup to compare the different parameters



(a) (b) (c)

Fig. 3. Variation of SQNR (dB) with the word-length of coefficient W_C and the word-length of the data,W_D, over different FFT lengths. (a) FFT length, $N = 256$. (b) FFT length, $N = 512$. (c) FFT length, $N = 1024$. The dashed lines represent desired operating points for specified SQNR that minimize area and power consumption.



(a) (b) (c)

Fig. 4. Variation of area (mm^2) with word-length of coefficient, W_C and word-length of the data,W_D, over different FFT lengths. (a) FFT length, $N = 256$. (b) FFT length, $N = 512$. (c) FFT length, $N = 1024$. The dashed lines represent desired operating points for an area constraint that maximize SQNR.

while ensuring that the timing is met at this frequency. For the evaluation of SQNR, we use a sum of sinusoids as the input and calculate the SQNR. W_D and W_C are varied by steps of one bit until 32 bits. The size of the FFT, N, is chosen as 256, 512 or 1024.

The points marked by symbols in the graphs and tables correspond to designs that satisfy the SQNR requirement of: $\circ = 90$ dB, $+ = 100$ dB, $* = 110$ dB and $\diamond = 120$ dB.

B. Analysis of SQNR

This experiment evaluates the effect of varying the parameters W_D, W_C and N on the SQNR of the system. The SQNR is calculated as

$$\text{SQNR(dB)} = 10\log_{10}\left(\frac{||X_I||^2}{||X_I - X_q||^2}\right) \quad (1)$$

Where X_I and X_q, respectively, represent the output of an FFT *without* quantization and the output of the current design *with* quantization. In Fig. 3 we compare the effects of W_D and W_C on SQNR across different length FFTs. We observe that W_D is more of a limiting factor when it comes to the overall SQNR of the system. W_C can be set 4 to 6 bits less than W_D depending on N to achieve a good SQNR with less area and power consumption. The best operating points for a given SQNR are shown in dashed lines in Fig. 3. The designs at these operating points minimize the area and power consumption for a specified SQNR. Note that similar observations on operating points for specified SQNR were made for complex FFTs [22]. The choice of N for a design is ideally application-specific and this choice has minimal impact on the SQNR of the system.

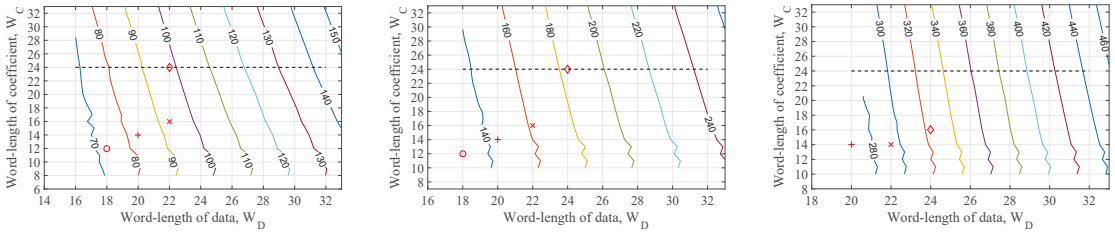
TABLE I
EFFECTS OF THE WORD-LENGTH OF THE DATA W_D AND THE WORD-LENGTH OF THE COEFFICIENT W_C ON SQNR (DB) FOR SELECTED CONFIGURATIONS

W_D	W_C					
	10	**12**	**14**	**16**	**24**	**32**
$N = 256$						
16	76.56	80.36	80.53	80.88	80.82	80.82
18	79.86	$\circ$ 88.69	91.96	92.19	92.46	92.46
20	80.15	91.55	$+$ 100.76	104.75	105.78	105.78
22	80.16	91.82	103.66	$*$ 113.56	$\diamond$ 119.12	119.07
24	80.15	91.83	103.73	115.48	125.58	125.70
32	80.15	91.82	103.73	115.65	127.42	127.42
$N = 512$						
18	82.51	$\circ$ 89.67	91.79	91.88	91.91	92.01
20	83.15	94.30	$+$ 102.22	104.59	106.01	105.80
22	83.18	94.79	106.10	$*$ 114.52	117.13	117.12
24	83.19	94.82	106.67	117.46	$\diamond$ 123.80	123.78
32	83.19	94.82	106.73	118.73	125.78	125.78
$N = 1024$						
20	86.05	97.05	$+$ 103.06	105.10	105.42	105.43
22	86.13	98.11	$*$ 108.64	115.18	117.09	117.13
24	86.14	98.22	109.75	$\diamond$ 120.28	125.96	126.06
32	86.14	98.24	109.91	121.75	130.37	130.38

Table I presents the SQNR values for different values of W_D, W_C and N.

C. Analysis of Area

This experiment evaluates the effect of varying the parameters W_D, W_C and N on the area of the system. The design was synthesized in a single technology node and can be used as a reference to study these parameters.



(a) (b) (c)

Fig. 5. Variation of power (mW) with word-length of coefficient W_C and word-length of the data,W_D, over different FFT lengths. (a) FFT length, $N = 256$. (b) FFT length, $N = 512$. (c) FFT length, $N = 1024$. The dashed lines represent desired operating points for a power constraint that maximize SQNR.

TABLE II
EFFECTS OF THE WORD-LENGTH OF THE DATA W_D AND THE WORD-LENGTH OF THE COEFFICIENT W_C ON AREA (mm^2), FOR SELECTED CONFIGURATIONS

W_D	W_C					
	10	**12**	**14**	**16**	**24**	**32**
$N = 256$						
16	0.073	0.076	0.080	0.084	0.096	0.108
18	0.086	○ 0.090	0.093	0.097	0.114	0.128
20	0.096	0.100	+ 0.104	0.109	0.126	0.143
22	0.107	0.111	0.115	∗ 0.120	◇ 0.139	0.156
24	0.116	0.121	0.126	0.130	0.150	0.170
32	0.156	0.161	0.169	0.174	0.204	0.230
$N = 512$						
18	0.140	○ 0.146	0.151	0.157	0.176	0.197
20	0.156	0.163	+ 0.166	0.172	0.194	0.215
22	0.172	0.178	0.183	∗ 0.190	0.212	0.235
24	0.188	0.194	0.201	0.207	◇ 0.230	0.255
32	0.254	0.261	0.270	0.277	0.310	0.352
$N = 1024$						
20	0.269	0.276	+ 0.283	0.290	0.321	0.347
22	0.296	0.304	∗ 0.312	0.319	0.349	0.376
24	0.322	0.331	0.341	◇ 0.350	0.376	0.408
32	0.433	0.442	0.453	0.462	0.504	0.549

TABLE III
EFFECTS OF THE WORD-LENGTH OF THE DATA W_D AND THE WORD-LENGTH OF THE COEFFICIENT W_C ON POWER CONSUMPTION (mW), FOR SELECTED CONFIGURATIONS

W_D	W_C					
	10	**12**	**14**	**16**	**24**	**32**
$N = 256$						
16	63.11	63.77	64.54	65.60	68.17	71.79
18	72.65	○ 73.60	74.33	75.78	79.32	84.04
20	81.09	82.08	+ 82.98	84.45	88.98	93.43
22	89.63	90.52	91.44	∗ 93.08	◇ 97.99	102.68
24	97.93	99.11	100.19	102.05	107.46	112.57
32	131.82	133.27	134.64	136.77	143.76	149.77
$N = 512$						
18	128.12	○ 129.22	130.21	131.91	136.14	141.76
20	142.82	144.23	+ 145.04	146.72	152.06	157.36
22	157.65	159.00	160.06	∗ 161.92	167.57	173.23
24	172.46	173.99	175.32	177.34	◇ 183.68	189.96
32	232.59	234.45	236.00	238.42	246.20	254.46
$N = 1024$						
20	263.22	264.53	+ 265.89	267.93	274.23	280.98
22	290.74	292.12	∗ 293.59	295.70	302.28	308.85
24	318.10	319.85	321.53	◇ 323.94	331.17	338.24
32	428.45	430.76	432.62	435.31	444.66	453.55

Fig. 4 compares the effects of W_D and W_C on area. As seen from the figure, W_D and W_C have a similar impact on the area of the system. The choice of N matters here, as different values of N vary the ratio of contribution between W_D and W_C to the area of the design. As we increase N, the contribution of W_D increases as we require larger circuits for data reordering. Also, an increase in N results in an increase in contribution from W_C as more coefficients need to be stored. For a specified *area constraint*, a point corresponding to a lower W_D and higher W_C is a desirable choice for achieving higher SQNR. The best operating points for an *area-constrained* design are shown in dashed lines in Fig. 4. Note that, for higher values of W_D, W_C is fixed at 24 bits for *area-constrained* designs. Table II presents the area in mm^2 for different values of W_D, W_C and N.

D. Analysis of Power

This experiment evaluates the effect of varying the parameters W_D, W_C and N on the power consumption of the system.

Fig. 5 and Table III illustrate the tradeoff of power consumption in mW with respect to W_D, W_C and N. For a *power-constrained* design, a point corresponding to a lower W_D and higher W_C is a desirable choice for achieving higher SQNR. The best operating points are shown by the dashed lines in Fig. 5. Note that for *power-constrained* designs, W_C is fixed at 24 bits for the range of W_D considered in this paper.

IV. CONCLUSION

This paper has quantified the impact of variation in word-lengths and FFT length on the SQNR, power, and area. This study is important especially in advanced complex designs that are constrained by low power consumption and low area requirements. Therefore we provide design choices on the selection of word-length for a given specification (area, power, and SQNR).

Future work on trade-off analysis of SQNR, area and power can be directed towards use of fixed-width multipliers based on approximations using statistical corrections [25], [26]. These designs can reduce area and power consumption without degrading SQNR. Similar empirical analysis can also be carried out for parallel real FFTs [7]–[9]. Analysis of effect of finite word-length on Welch power spectral density [27] for real-valued signals is also of interest.

REFERENCES

[1] Z. Zhang and K. K. Parhi, "Low-complexity seizure prediction from iEEG/sEEG using spectral power and ratios of spectral power," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 10, no. 3, pp. 693–706, June 2016.

[2] ——, "Seizure prediction using polynomial SVM classification," in *Proc. Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug 2015, pp. 5748–5751.

[3] ——, "Seizure detection using regression tree based feature selection and polynomial SVM classification," in *Proc. Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug 2015, pp. 6578–6581.

[4] Y. Park, L. Luo, K. K. Parhi, and T. Netoff, "Seizure prediction with spectral power of EEG using cost-sensitive support vector machines," *Epilepsia*, vol. 52, no. 10, pp. 1761–1770, 2011.

[5] T. Xu, M. Stephane, and K. K. Parhi, "Abnormal neural oscillations in schizophrenia assessed by spectral power ratio of MEG during word processing," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 11, pp. 1148–1158, Nov 2016.

[6] T. Xu, K. R. Cullen, B. Mueller, M. W. Schreiner, K. O. Lim, S. C. Schulz, and K. K. Parhi, "Network analysis of functional brain connectivity in borderline personality disorder using resting-state fMRI," *NeuroImage: Clinical*, vol. 11, pp. 302 – 315, 2016.

[7] M. Garrido, K. K. Parhi, and J. Grajal, "A pipelined FFT architecture for real-valued signals," *IEEE Trans. on Circuits and Systems I: Regular Papers*, vol. 56, no. 12, pp. 2634–2643, Dec 2009.

[8] S. A. Salehi, R. Amirfattahi, and K. K. Parhi, "Pipelined architectures for real-valued FFT and hermitian-symmetric IFFT with real datapaths," *IEEE Trans. on Circuits and Systems II: Express Briefs*, vol. 60, no. 8, pp. 507–511, Aug 2013.

[9] M. Ayinala and K. K. Parhi, "FFT architectures for real-valued signals based on radix-$2^3$and radix-2^4algorithms," *IEEE Trans. on Circuits and Systems I: Regular Papers*, vol. 60, no. 9, pp. 2422–2430, Sep. 2013.

[10] A. Chinnapalanichamy and K. K. Parhi, "Serial and interleaved architectures for computing real FFT," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, April 2015, pp. 1066–1070.

[11] P. K. Meher, B. K. Mohanty, S. K. Patel, S. Ganguly, and T. Srikanthan, "Efficient VLSI architecture for decimation-in-time fast Fourier transform of real-valued data," *IEEE Trans. on Circuits and Systems I: Regular Papers*, vol. 62, no. 12, pp. 2836–2845, Dec 2015.

[12] K. K. Parhi, *VLSI Digital Signal Processing Systems: Design and Implementation*. Hoboken, NJ, USA: Wiley, 1999.

[13] W. Chang and T. Q. Nguyen, "On the fixed-point accuracy analysis of FFT algorithms," *IEEE Trans. on Signal Processing*, vol. 56, no. 10, pp. 4673–4682, Oct 2008.

[14] N. Sulaiman and T. Arslan, "A multi-objective genetic algorithm for on-chip real-time optimisation of word length and power consumption in a pipelined FFT processor targeting a MC-CDMA receiver," in *Proc. NASA/DoD Conference on Evolvable Hardware*, June 2005, pp. 154–159.

[15] R. B. Perlow and T. C. Denk, "Finite wordlength design for VLSI FFT processors," in *Conf. Rec. Asilomar Conference on Signals, Systems and Computers*, vol. 2, Nov 2001, pp. 1227–1231 vol.2.

[16] P. Welch, "A fixed-point fast Fourier transform error analysis," *IEEE Trans. on Audio and Electroacoustics*, vol. 17, no. 2, pp. 151–157, June 1969.

[17] C. Yang, Y. Xie, H. Chen, and Y. Deng, "New quantization error assessment methodology for fixed-point pipeline FFT processor design," in *Proc. IEEE International System-on-Chip Conference (SOCC)*, Sep. 2014, pp. 299–305.

[18] X. Shih, H. Chou, and Y. Liu, "Design and implementation of flexible and reconfigurable SDF-based FFT chip architecture with changeable-radix processing elements," *IEEE Trans. on Circuits and Systems I: Regular Papers*, vol. 65, no. 11, pp. 3942–3955, Nov 2018.

[19] C. Yang, Y. Xie, H. Chen, and Y. Deng, "New quantization error assessment methodology for fixed-point pipeline FFT processor design," in *Proc. IEEE International System-on-Chip Conference (SOCC)*, Sep. 2014, pp. 299–305.

[20] J.-C. Kuo, C.-H. Wen, C.-H. Lin, and A.-Y. A. Wu, "VLSI design of a variable-length FFT/IFFT processor for OFDM-based communication systems," *EURASIP Journal on Advances in Signal Processing*, vol. 2003, no. 13, p. 439360, Dec 2003.

[21] S. Johansson, S. He, and P. Nilsson, "Wordlength optimization of a pipelined FFT processor," in *Conf. Rec. 42nd Midwest Symposium on Circuits and Systems*, vol. 1, Aug 1999, pp. 501–503.

[22] D. Guinart, "Deterministic analysis of the accuracy in FFT hardware architectures," Master's thesis, Linköping University, Jun. 2012.

[23] M. Garrido, N. K. Unnikrishnan, and K. K. Parhi, "A serial commutator fast Fourier transform architecture for real-valued signals," *IEEE Trans. on Circuits and Systems II: Express Briefs*, vol. 65, no. 11, pp. 1693–1697, Nov 2018.

[24] H. Sorensen, D. Jones, M. Heideman, and C. Burrus, "Real-valued fast Fourier transform algorithms," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. 35, no. 6, pp. 849–863, June 1987.

[25] K.-J. Cho, K.-C. Lee, J.-G. Chung, and K. K. Parhi, "Design of low-error fixed-width modified booth multiplier," *IEEE Trans. on Very Large Scale Integration (VLSI) Systems*, vol. 12, no. 5, pp. 522–531, May 2004.

[26] S.-M. Kim, J.-G. Chung, and K. K. Parhi, "Low error fixed-width CSD multiplier with efficient sign extension," *IEEE Trans. on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 50, no. 12, pp. 984–993, Dec 2003.

[27] K. K. Parhi and M. Ayinala, "Low-complexity Welch power spectral density computation," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 61, no. 1, pp. 172–182, Jan 2014.