2020 Moscow Workshop on Electronic and Networking Technologies (MWENT)

Analysis of IIR Filters by Interval Response

Vin´ıcius Borges
Electrical Engineering Department
Federal University of Sa˜o Joa˜o del-Rei
São João del-Rei, Brazil
vnisb@hotmail.com

Erivelton G. Nepomuceno
Electrical Engineering Department
Federal University of Sa˜o Joa˜o del-Rei
São João del-Rei, Brazil
nepomuceno@ufsj.edu.br

Aleksandra V. Tutueva
Computer Aided Design Department
Petersburg Electrotechnical University
Saint Petersburg, Russia
avtutueva@etu.ru

Artur I. Karimov
Youth Research Institute
Petersburg Electrotechnical University
Saint Petesburg, Russia
aikarimov@etu.ru

Carlos Duque
Power Electronics and Automation Group
Federal University of Juiz de Fora
Juiz de Fora, Brazil
carlos.duque@ufjf.edu.br

Timur I. Karimov
Computer Aided Design Department
Petersburg Electrotechnical University
Saint Petersburg, Russia
tikarimov@etu.ru

***Abstract*—The classical theory of signal processing assumes that the designed IIR filters are continuous and have infinitely accurate coefficients. However, when developing filters for real-world digital signal processing tasks, it is necessary to take into account the finite precision of the coefficients representation, especially considering the fixed-point arithmetic. In this paper, we propose a new approach, which we call the interval size approach, making it easy to evaluate the actual digital filter response for various fixed-point arithmetic parameters. We provide Illustrative examples to demonstrate the frequency response bounds evolution as an order, cutoff frequency, and signal frequency are varying. We show that the interval size can be used as a well-suited accuracy factor of a digital filter.**

Keywords—signal processing, digital filters, quantization, interval size

I. Introduction

A digital filter amplifies or reduces certain parameters of a discrete-time signal using mathematical operations [1]. It is widely used in signal processing, and its extraordinary performance is one of the key reasons that DSP has become so popular [2]. Digital filters are categorized as finite duration pulse response (FIR) filters with nonzero pulse response for only a finite number of samples, and infinite duration pulse response (IIR) with an infinite number of non-zero samples in an impulse response [3].

When designing high-performance signal processing systems, a priori it is considered that the filter response is acceptable. Nevertheless, in hardware systems with fixed-point arithmetic, i.e. microcontrollers, digital signal processors and FPGA, or in systems with highly demanding performance specifications, the importance of filter coefficient accuracy increases. With a lack of accuracy, the signal specifications can be completely missed, and the signal – distorted [4–6]. Therefore, a common goal in finite precision analysis is to choose a word length such that the digital system provides a sufficiently accurate realization of the desired frequency response while minimizing the complexity and cost of hardware and software [7].

The reported study was partially supported by CNPq, FAPEMIG, CAPES and by the grant of the Russian Science Foundation (RSF), project № 19-19-00566. By the grant of the RSF, methods to effectively organize the procedure of implementing the signal processing system on the target platform were developed.

In digital signal processing, the effects of finite word length are known as one of the most significant factors. Mullis [8] and Huang [9] demonstrated the influence of quantization errors on digital filter performance depends on the performance chosen for the filter implementation. Rader and Gold [10] showed that small errors in the denominator or numerator coefficients may cause large poles or zeros offsets. Moreover, Goodall et al. [11–12] introduced the concept of coefficients sensitivity to word lengths.

Today there are some common approaches to handling the inaccuracies in parameter values, e.g. to develop design procedures that are inherently insensitive to parameter inaccuracies, or to choose specifications that are consistent with the limited register length [13]. Another possibility is the implementation of interval sets. In [14] it was shown that the frequency response of interval FIR and IIR filters may be guaranteed to lie within specified bounds on corresponding small subsets of extreme filters, each of which has fixed coefficients.

Interval mathematics is well-known in the literature. For example, in previous work, we [15] presented a theorem that identifies the reliability of fixed-point calculations. A technique using two extension intervals to reject a fixed-point design in case a required accuracy is not reached is discussed in [16]. However, the analysis of a digital filter frequency response based on the interval arithmetic still requires a study. This paper uses the proposed concept to evaluate the response of the Butterworth filter.

The paper is organized as follows: In section II, we define the domain and basics of the proposed methodology. In section III, we describe the results. Conclusion and discussion are given in Section IV.

II. Preliminary Concepts and Methodology

A. IIR filter

IIR digital filters are characterized by infinite impulse response. They have output feedback, which allows achieving a more selective frequency response with a lower number of coefficients. However, IIR digital filters are sensitive to polynomial coefficients varying. IIR digital filters are described by the transfer function [17]:

978-1-7281-2572-5/20/$31.00 ©2020 IEEE

$$H(z) = \frac{\sum_{l=0}^{N} a_l y[n-l]}{\sum_{k=0}^{M} b_k x[n-k]}, \quad (1)$$

where N – zeros coefficients, M – poles coefficients, b_k and a_l – filter coefficients. More general description of (1) with $a_0 = 1$:

$$y(k) = \sum_{k=0}^{M} b_k x[n-k] - \sum_{l=1}^{N} a_1 y[n-l]. \quad (2)$$

B. Quantization error

In the process of converting continuous values to discrete values the difference between the original analog signal and its quantized version appears. This error is called a "quantization error". Thus, quantizing an analog signal corresponds to adding a certain amount of noise. The fewer bits are used in quantization, the more noise is added. The quantization error is given by the formula:

$$\sigma = \frac{A}{2Q}, \quad (3)$$

where $Q = 2^b$, b is the number of bits, A is the full range of analog signal. Using (3), the coefficients of (2) can be rewritten as:

$$\begin{array}{ll} a_k = a_k - \frac{Q}{2} & a_k = a_k + \frac{Q}{2} \\ b_k = b_k - \frac{Q}{2} & b_k = b_k + \frac{Q}{2} \end{array}. \quad (4)$$

C. Simulation

We propose the following simulation algorithm:

Step 1: Generate the poles and zeros of the transfer function (1).

Step 2: Choose a number of bits and calculate quantization error, (3).

Step 3: Insert the quantization error into the poles and zeros. Then, (1) can be rewritten in the following combinations:

$$y_1[k] = \sum_{l=1}^{N} a_1 y[n-l] + \sum_{k=0}^{M} b_x x[n-k], \quad (5)$$

$$y_2[k] = \sum_{l=1}^{N} a_1 y[n-l] + \sum_{k=0}^{M} b_x x[n-k], \quad (6)$$

$$y_3[k] = \sum_{l=1}^{N} a_1 y[n-l] + \sum_{k=0}^{M} b_x x[n-k], \quad (7)$$

$$y_4[k] = \sum_{l=1}^{N} a_1 y[n-l] + \sum_{k=0}^{M} b_x x[n-k]. \quad (8)$$

Step 4: Filter signal using filters (5)-(8).

Step 5: Get the maximum and minimum values in each iteration:

$$y_{\max}[k] = \max(y_1 y_2 y_3 y_4), \quad (9)$$

$$y_{\min}[k] = \min(y_1 y_2 y_3 y_4), \quad (10)$$

Step 6: Calculate the interval size.

$$I = y_{\max}[k] - y_{\min}[k]. \quad (11)$$

The obtained values can be used to analyze any IIR filter. In our study, we focused on Butterworth filter interval errors studying when 16-bit and 32-bit fixed-point arithmetic is used.

III. Tests and Results

To validate the proposed technique, the following tests were performed:

Test 1: Alteration of filter order with fixed cutoff frequency 40 Hz and input sine signal of 40 and 60 Hz frequency. The results obtained for 16 and 32 bits are presented in Tables I and II, respectively. The simulation was stopped when the response became unstable. Thus, we found that the highest possible order for 16 bits is 6th, and for 32 bits is 13th.

Test 2: Alteration of filter cutoff frequency using input sine signal with 40 and 60 Hz frequency and 5th order for 16 bits and 10th order for 32 bits. The results are shown in Fig. 1-4.

Test 3: Alteration of the input sinusoidal signal with filter cutoff frequency at 40 Hz and 5th order for 16 bits and 10th order for 32 bits. The results are shown in Fig. 5-8.

In test 1, the increase in filter order characterizes the increase of the interval size. But in the 32-bit system of 11th order, the size of the range decreased, showing that the level of uncertainty decreased. In test 2, in comparison to the 16-bit high-pass and low-pass filters, the only difference was a larger interval size at the cutoff frequency of 40 Hz. At 32 bits the low-pass filter obtained a smaller interval size. In test 3, the variation of the input signal determines which filter frequencies are least accurate.

Presented results make it possible to determine the best system order, as well as the cutoff frequency and the signal frequency, which provide the lower signal distortion in the fixed-point digital filter.

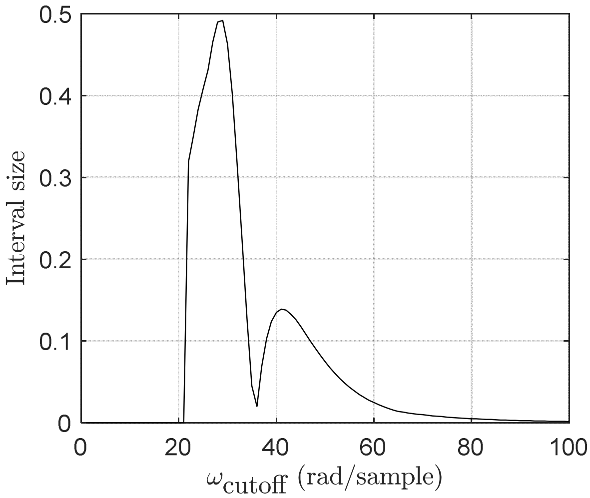

Fig. 1. Test 2: 16-bit low-pass filter of order 5.

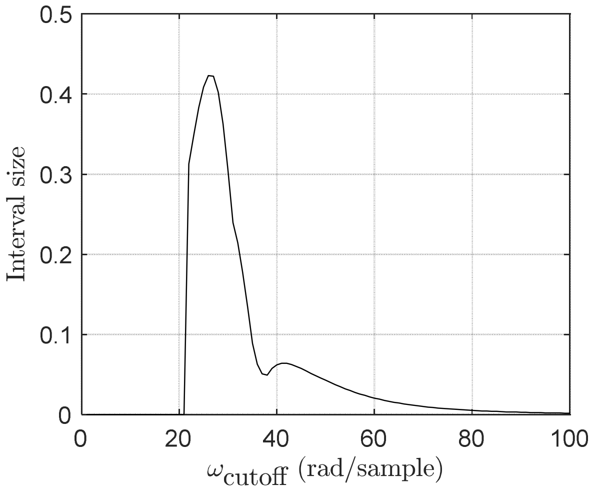

Fig. 2. Test 2: 16-bit high-pass filter of order 5.

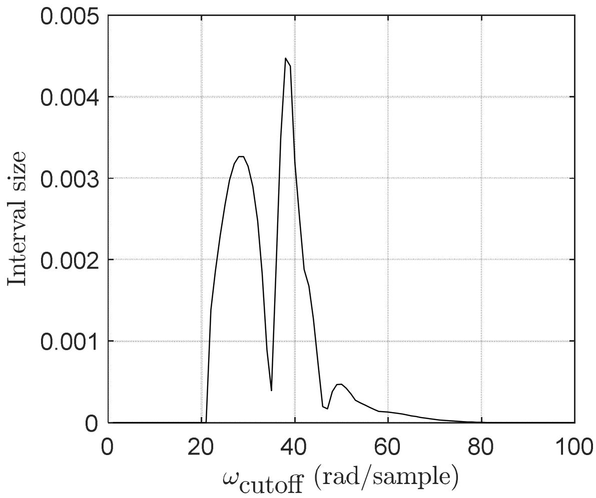

Fig. 3. Test 2: 32-bit low-pass filter of order 10.

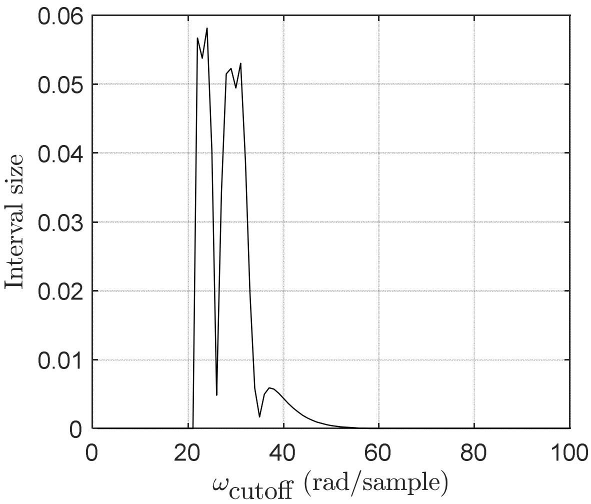

Fig. 4. Test 2: 32-bit high-pass filter of order 10.

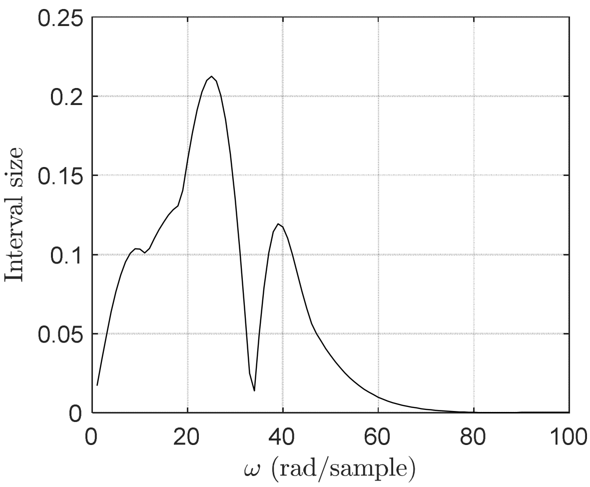

Fig. 5. Test 3: 16-bit low-pass filter of order 5.

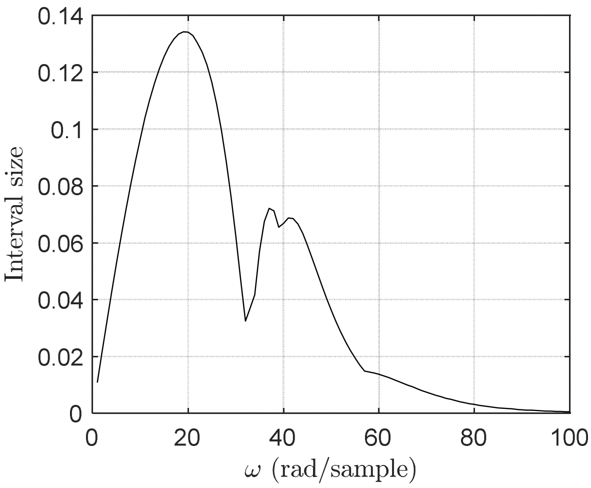

Fig. 6. Test 3: 16-bit high-pass filter of order 5.

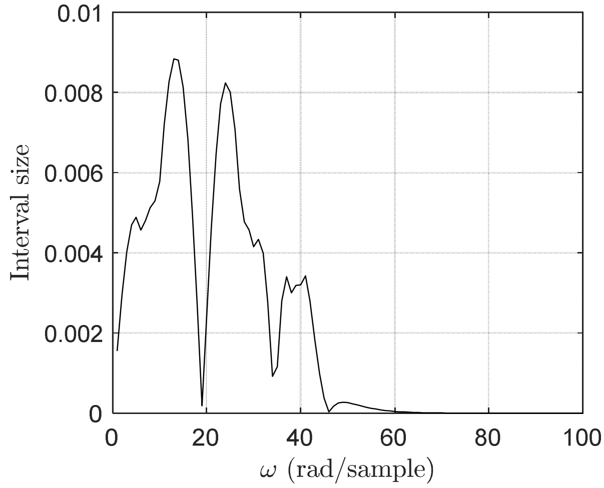

Fig. 7. Test 3: 32-bit low-pass filter of order 10.

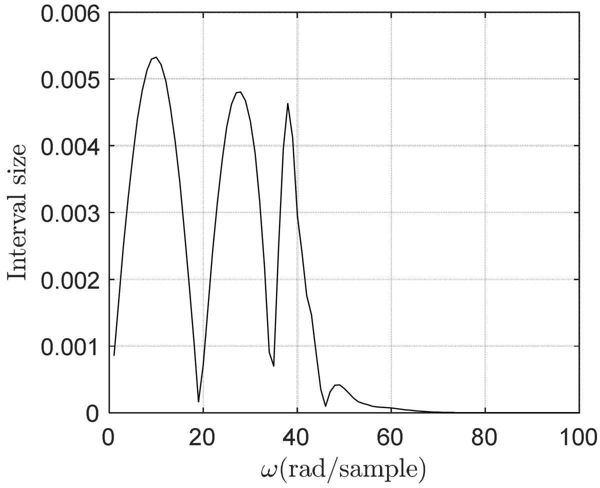

Fig. 8. Test 3: 32-bit high-pass filter of order 10.

TABLE I. INTERVAL SIZE WITH 16 BITS

n^{th} order	Low-pass	High-pass
1	0.00013157	0.00013783
2	0.00112199	0.00112199
3	0.00615655	0.00510382
4	0.02887566	0.01370685
5	0.10441428	0.02955294
6	0.16903701	0.11799196

TABLE II. INTERVAL SIZE WITH 32 BITS

n^{th} order	low-pass	high-pass
1	0.0000000020077	0.0000000021030
2	0.0000000171213	0.0000000171213
3	0.0000000939418	0.0000000778782
4	0.0000004406045	0.0000002061205
5	0.0000015834518	0.0000004456805
6	0.0000020584737	0.0000016033810
7	0.0000575346152	0.0000181295321
8	0.0003223693824	0.0001619635586
9	0.0009850016594	0.0010024738655
10	0.0043719068664	0.0043719234919
11	0.0175192104908	0.0136181514784
12	0.0080165777634	0.0080168544070
13	0.2815356824937	0.1973608462015

IV. CONCLUSION

An effective method to avoid such unwanted effects of filter coefficients quantization as quantization noise, limit cycles, and overflow oscillations is to simulate the system with various quantization settings and discard solutions that do not meet the criteria of stability and accuracy. To avoid extensive calculations, we propose an interval-based technique, according to which it is possible to determine the boundary values of the filter response accuracy. Therefore, the frequency response of the IIR filters may have characteristics within strictly specified limits, and the interval size can be used as an accuracy factor of a digital filter.

The limitation of this work is that we investigated only one type of filter zeros and poles placement, and also only two different machine word lengths were considered. In the future, the results will be extended to a wide class of filters.

REFERENCES

[1] E. Lai, "Practical digital signal processing," Elsevier, 2003.

[2] S. Smith, "Digital signal processing: a practical guide for engineers and scientists," Elsevier, 2013.

[3] I. Grout, "Digital systems design with FPGAs and CPLDs," Elsevier, 2011.

[4] M. Antosova, V. Davidek and B. Psenicka, "Finite word-length effects in digital state-space filters," Radioengineering, vol. 8(4), 1999. URL: http://hdl.handle.net/11012/58278

[5] B. Liu, "Effect of finite word length on the accuracy of digital filters–a review," IEEE Transactions on Circuit Theory, vol. 18(6), pp. 670–677, 1971. DOI: 10.1109/TCT.1971.1083365

[6] T.I. Karimov, D.N. Butusov, V.S. Andreev, V.G. Rybin, D.I. Kaplun, "Compact Fixed-Point Filter Implementation," in Proceedings of the 22nd Conference of FRUCT Association, Jyvaskyla, Finland, 15–18 May 2018; Vol. 2, pp. 73–78. DOI: 10.23919/FRUCT.2018.8468289.

[7] H.M. Cheng, G.C. Chiu, "Finite precision controller implementation-limitation on sample rate," in Advanced Intelligent Mechatronics, 2003. AIM 2003. Proceedings. 2003 IEEE/ASME International Conference on, vol. 1. IEEE, pp. 634–639, 2003. DOI:10.1109/AIM.2003.1225168.

[8] C. Mullis, R. Roberts, "Synthesis of minimum roundoff noise fixed point digital filters," IEEE Transactions on Circuits and Systems, vol. 23(9), pp. 551–562, 1976. DOI: 10.1109/TCS.1976.1084254

[9] S. Hwang, "Minimum uncorrelated unit noise in state-space digital filtering," IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 25(4), pp. 273–281, 1977. DOI: 10.1109/TASSP.1977.1162971

[10] C.M. Rade, B. Gold, "Digital filter design techniques in the frequency domain," Proceedings of the IEEE, vol. 55(2), pp. 149– 171, 1967. DOI: 10.1109/TEMC.1968.302949

[11] R.M. Goodall, B. Donoghue, "Very high sample rate digital filters using the δ operator," IEE Proceedings G (Circuits, Devices and Systems), vol. 140(3), pp. 199–206, 1993. DOI: 10.1049/ip-g-2.1993.0032

[12] S. Jones, R. Goodall, M. Gooch, "Targeted processor architectures for high-performance controller implementation," Control Engineering Practice, vol. 6, no. 7, pp. 867–878, 1998. DOI: 10.1016/S1474-6670(17)44459-4

[13] A.V. Oppenheim, C.J. Weinstein, "Effects of finite register length in digital filtering and the fast fourier transform," Proceedings of the IEEE, vol. 60(8), pp. 957–976, 1972. DOI: 10.1109/PROC.1972.8820.

[14] N.K. Bose, K.D. Kim, "Boundary implications for frequency response of interval fir and iir filters," IEEE Transactions on Signal Processing, vol. 39(10), pp. 2167–2173, 1991. DOI: 10.1109/78.91173

[15] E.G. Nepomuceno, "Convergence of recursive functions on computers," The Journal of Engineering, Institution of Engineering and Technology, pp. 1–3, 2014. DOI: 10.1049/joe.2014.0228.

[16] E.G. Nepomuceno, S.A Martins, "A lower bound error for free-run simulation of the polynomial NARMAX," Systems Science & Control Engineering, vol. 4(1), pp. 50–58, 2016. DOI: 10.1080/21642583.2016.1163296.

[17] A. Petraglia, "Fundamental frequency response bounds of direct-form recursive switched-capacitor filters with capacitance mismatch," IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing, vol. 48(4), pp. 340–350, 2001. DOI: 10.1109/82.933792.