The Relationship between the Representation of a IIR Digital Filter in the State Space and the Description by the Topological Matrix

Vladislav Lesnikov, Tatiana Naumovich, Alexander Chastikov
Vyatka State University
Kirov, Russia
Vladislav.Lesnikov.Ru@ieee.org, {ntv_new, alchast}@mail.ru

Abstract—**When analyzing and synthesizing digital filters with different structures, their representation in the state space is most often used. But this is not the only way to describe the filter. A digital filter with an arbitrary structure is most adequately described by a topological matrix, the elements of which are transfer coefficients between nodes of a block diagram. This paper allows us to expand the capabilities of the mathematical model of a digital filter in the form of a topological matrix.**

I. Introduction

The most important problem in the synthesis of IIR digital filters is considering the effects associated with the finite word length (FWL). Many publications used the representation of IIR digital filters in the state space to solve these problems.

In [1], roundoff noise is studied using a state variable formulation. An expression for the variance of the rounding noise at the filter output was obtained. A set of filter invariants are defined. These invariants called second-order modes.

In [2], expressions for the roundoff noise and the dynamic-range constraint equation have been established for the digital filters described by the state equations. The roundoff noise minimization problem was considered. This task was solved subject to the l_2-norm dynamic-range constraint. A lower bound and the global minimum of the noise generated under the assumption that each state equation contains exactly one noise source is obtained.

In [3] two expressions of the output error variance are proposed based on deterministic and statistical approaches for fixed-point state-space digital filters in the time domain. The statistical coefficient sensitivity introduced.

In [4] various possibilities for the individual realization of different structures, depending on coefficient format, state variable format, overflow characteristics are studied.

The results presented in [5] are dedicated to minimizing the upper bound of the rounding noise l_2-norm. Robustness against practical uncertainties is imposed. Finite word length implementation, roundoff errors, and numerical precision included in uncertainty. Convex programming techniques applied to state-space based filter optimization.

Based on the state space representation, a special implicit system description is entered in [6]. Two suitable coefficient sensitivity measures are used.

In [7] the state-space realization utilized to minimize the roundoff noise gain. A genetic algorithm is proposed to implement the optimal structure.

The problem of synthesis the optimal structure of a state-estimate feedback controller with minimum l_2-sensitivity and no overflow is described in [8].

A new measure used for the implementation of filters in state-space form presented in [9]. The classical L2-sensitivity measure is extended with precise consideration on their fixed-point representation in order to make a more valid measure.

In [10] joint optimization of error feedforward, high-order error feedback and state-space realization for minimizing filter output roundoff noise subject to $l2$-scaling constraints for state space digital filters is proposed.

A new measure for evaluating roundoff noise and pole sensitivity is proposed in [11]. This is used to minimize weighted roundoff noise and pole sensitivity subject to l_2-scaling constraints for state-space digital filters.

The algorithm determining the fixed-point formats of all the involved variables (states and outputs) for filters in state-space representation is proposed in [12]. Several quantization modes for coefficients (rounding and truncation) for two's-complement-based fixed-point arithmetic considered.

In [13] a measure for pole and zero sensitivity is proposed. This measure is used to minimize weighted pole and zero sensitivity subject to l_2-scaling constraints for state-space digital filters.

Evaluation of a sufficient condition for the absence of limit cycles of state-space second-order digital filters with minimum l_2-sensitivity subject to l_2-scaling constraints is described in [14].

Thus, the approach described in the reviewed papers allows us to obtain valuable results in analyzing and minimizing the sensitivity of the filter characteristics to their parameters, analyzing and minimizing the noise level of rounding off the results of arithmetic operations, taking into account parasitic oscillations of the limit cycle.

However, despite all the achievements, the problem of the synthesis of IIR digital filters with hard spectral constraints and finite word length remains far from a satisfactory solution.

Traditionally, the implementation of digital signal processing algorithms uses the following approach. First, all calculations are performed in floating point format, and then automatic conversion to fixed point format is performed. Applied to the synthesis of digital filters, it looks like this [15].

At the stage of functional synthesis, the transfer function of the filter or its zeros and poles is computed. All these calculations are in floating point format. Then, at the stage of structural synthesis, the structure of the filter is selected, and the coefficients of the block diagram are calculated in a floating-point format. And only now we are moving to a fixed-point format with a finite word length. However, quantization of the coefficients leads to an unacceptable distortion of the results of functional synthesis. Attempts to correct distortions can lead to either an unacceptable increase in the length of the word or the need to change the structure or return to the stage of functional synthesis.

The authors of this work in the cycle of their publications develop an alternative approach to the synthesis of IIR filters with a FWL [16], [17].

Already at the stage of functional synthesis with a new approach, we get a solution based on the finite word length. The possibility of such a solution is based on the results obtained by the authors of the study of the *z*-plane discretization process due to quantization of the coefficients. These studies have shown that not any *z*-plane point can be a zero or a pole of a FWL filter. Such zeros and poles are elements of the coefficients fractional part length-dependent subset of the set of corresponding degree algebraic numbers. The results of functional synthesis in this case satisfy the requirements of the filter and are final, that is, they are not distorted at the stage of structural synthesis.

At the stage of structural synthesis, a structural scheme is selected from the set of generated structures taking into account the degree of algebraic numbers, which are zeros and poles, the length of the fractional part of the coefficients of equivalent canonical structure, the level of rounding noise of arithmetic operations [18], [19].

The mathematical model of the specific structure of a digital filter with this approach is the square matrix of transfer coefficients between the nodes of the structure, described in [20] and called by the authors the topological matrix. The topological matrix uniquely determines the structure of the filter, whereas there is no one-to-one correspondence between a specific structure and its state-space representation. The topological matrix provides the possibility of generating all possible structures with a given number of nodes. The authors studied the algebraic properties of topological matrices that determine the degree of algebraic numbers (zeros and poles), structural accuracy and complexity of block diagrams. Therefore, it is advisable to express these results in terms of a model based on a topological matrix.

In [21], the authors have already used these results to analyze rounding noise. This paper is devoted to a more detailed study of the correspondence between the descriptions of the structure of the filter topological matrix and representation in the state space.

The remainder of this paper is divided into six sections. Section II describes the use of state space to represent recursive digital filters. The technique based on the application for describing the structure of a digital filter of topological matrices is described in detail in Section III. The ability to generate digital filter structures represented by topological matrices is described in Section IV. This section also shows that the fine structure of the topological matrix can be used for naming the digital filter structure. Section V deals with the process of generating filter structures. For the particular case of structures with four nodes, six possible templates are generated, for which a block diagram and topological matrices are shown. For generated templates, Section VI shows how, based on the topological matrix, all parameters of the state space matrices can be calculated. Conclusions on the article are in Section VII.

II. Representation of Digital Filter Structure in State Space

The state space representation of a filter based on its description by an equation [15], [22]:

$$\begin{bmatrix} \mathbf{s}_{k+1} \\ y_k \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{c} & d \end{bmatrix} \begin{bmatrix} \mathbf{s}_k \\ x_k \end{bmatrix}, \tag{1}$$

where $\mathbf{A} \in \mathrm{R}^{n\times n}$ is the state matrix, $\mathbf{b} \in \mathrm{R}^{n\times 1}$ is the input matrix, $\mathbf{c} \in \mathrm{R}^{1\times n}$ is the output matrix, $d \in \mathrm{R}$ is the feedforward matrix, $x_k \in \mathrm{R}$ are the input samples, $y_k \in \mathrm{R}$ are the output samples, $\mathbf{s}_k \in \mathrm{R}^{n\times 1}$ are the state vectors, $k = 0,1,2,\ldots$. Typically, states of the system believe samples at the outputs of n delay units. The elements of the state matrix **A** are the transfer coefficients from the outputs of the delay blocks to their inputs. The elements of the input matrix **b** are the transfer coefficients from the input of the filter to the inputs of the delay units. The elements of the output matrix **c** are the transfer coefficients from the outputs of the delay units to the output of the filter. The scalar d is the transfer coefficient between input and output of filter.

The block diagram shown in Fig. 1 corresponds to equation (1)

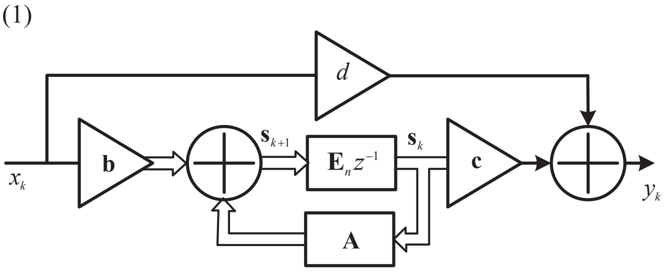

Fig. 1. A state-space representation of any digital filter

III. Description of Digital Filters by Topological Matrix

Any structure of digital filter with *N* nodes can be described by a topological matrix

$$\mathbf{T}(z) = \left[t_{ij}\right], \tag{2}$$

which dimension is equal to $N \times N$ and where t_{ij} is the branch transmittance from a node with number j to a node with number i.

As an example, consider the description of the canonical form of a second-order IIR filter (Fig. 2) with a transfer function

$$\mathrm{H}(z) = \frac{b_0 z^2 + b_1 z + b_2}{z^2 - a_1 z - a_2}. \tag{3}$$

The topological matrix of this structure is equal to

$$\mathbf{T}(z) = \begin{bmatrix} 0 & z^{-1} & 0 & 0 \\ 0 & 0 & z^{-1} & 0 \\ -a_{21} & -a_{11} & 0 & 0 \\ b_{21} & b_{11} & b_{43} & 0 \end{bmatrix}. \tag{4}$$

If the filter structure is physically realizable (computable), then the nodes can be renumbered with natural numbers in such a way that all elements of the topological matrix that are not equal to 0 or 3 will be below the main diagonal.

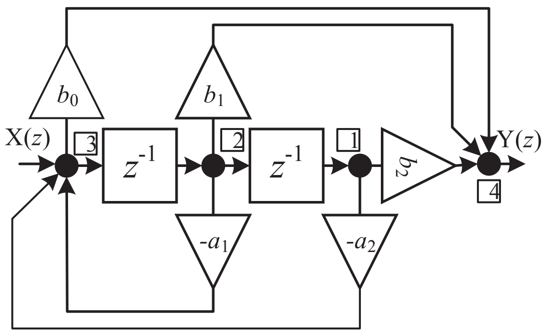

Fig. 2. Canonical form IIR digital filter of second order

Any structure can be described by the equation (20)

$$\mathbf{y}(z) = \mathbf{T}(z)\mathbf{y}(z) + \mathbf{x}(z)\mathrm{X}(z), \tag{5}$$

where

$$\mathbf{y}(z) = \left[\mathrm{Y}_1(z) \quad \mathrm{Y}_2(z) \quad \ldots \quad \mathrm{Y}_N(z)\right]^t, \tag{6}$$

is a vector of z-transforms of sequences of samples, computed in nodes of structure, t is a symbol for the transpose operation,

$$\mathbf{x}(z) = \left[x_1 \quad \ldots \quad x_i \quad \ldots \quad x_N\right]^t, \tag{7}$$

$$x_i = \begin{cases} 1, & \text{if} \quad i = inp, \\ 0, & \text{if} \quad i \neq inp, \end{cases} \tag{8}$$

inp is the input node number, X(z) is the z-transform of input sequence. For the filter shown in Fig. 1.

$$\mathbf{y}(z) = \begin{bmatrix} \mathrm{Y}_1(z) \\ \mathrm{Y}_2(z) \\ \mathrm{Y}_3(z) \\ \mathrm{Y}_4(z) \end{bmatrix} = \begin{bmatrix} 0 & z^{-1} & 0 & 0 \\ 0 & 0 & z^{-1} & 0 \\ -a_{21} & -a_{11} & 0 & 0 \\ b_{21} & b_{11} & b_{43} & 0 \end{bmatrix} \begin{bmatrix} \mathrm{Y}_1(z) \\ \mathrm{Y}_2(z) \\ \mathrm{Y}_3(z) \\ \mathrm{Y}_4(z) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \mathrm{X}(z), \tag{9}$$

or

$$\begin{cases} \mathrm{Y}_1(z) = \mathrm{Y}_2(z) z^{-1}, \\ \mathrm{Y}_2(z) = \mathrm{Y}_3(z) z^{-1}, \\ \mathrm{Y}_3(z) = \mathrm{X}(z) - a_2\,\mathrm{Y}_1(z) - a_1\,\mathrm{Y}_2(z), \\ \mathrm{Y}_4(z) = b_2\,\mathrm{Y}_1(z) + b_1\,\mathrm{Y}_2(z) + b_0\,\mathrm{Y}_3(z). \end{cases} \tag{10}$$

From (5) it is easy to get the equations

$$\left(\mathbf{E}_N - \mathbf{T}(z)\right)\mathbf{y}(z) = \mathbf{x}(z)\mathrm{X}(z), \tag{11}$$

where $\mathbf{E}_N$ is identity matrix of size N, and

$$\mathbf{y}(z) / \mathrm{X}(z) = \left(\mathbf{E}_N - \mathbf{T}(z)\right)^{-1} \mathbf{x}(z). \tag{12}$$

From equation (12) it follows that the elements of the matrix

$$\mathbf{H}(z) = \left[\mathrm{H}_{oi}(z)\right] = \left(\mathbf{E}_N - \mathbf{T}(z)\right)^{-1} \tag{13}$$

are the transfer functions of digital filters described by the matrix $\mathbf{T}(z)$, the node i which is the input, and the node o is the output.

IV. Generation and Enumeration of IIR Filters Structures

The structure of the filter Str_N can be divided into two interconnected parts C and Z (Fig. 3). Part C is obtained from the filter structure by removing all delay blocks from it. Part Z consists of delay blocks only.

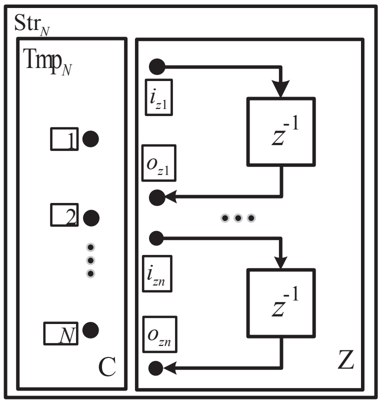

Fig. 3. Decomposition of filter structure

Topology matrix can be represented as

$$\mathbf{T}(z) = \mathbf{T}_\mathrm{C} + z^{-1}\mathbf{T}_\mathrm{Z}, \tag{14}$$

where the matrix $\mathbf{T}_\mathrm{C}$ is obtained by replacing in the matrix $\mathbf{T}(z^{-1})$ elements equal to z^{-1} by zero, the matrix $\mathbf{T}_\mathrm{Z}$ is obtained by replacing in the matrix $\mathbf{T}(z^{-1})$ elements not equal to z^{-1} by zero and replacing elements equal to z^{-1} by 1. The matrix $\mathbf{T}_\mathrm{C}$ is the topological matrix of part C. The matrix $z^{-1}\mathbf{T}_\mathrm{Z}$ is the topological matrix of part Z.

The matrix $\mathbf{T}_\mathrm{C}$ of any scheme with N nodes is obtained from the generalized matrix

$$\mathbf{T}_\mathrm{C} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & \dots & 0 \\ c_{21} & 0 & 0 & \dots & 0 & \dots & 0 \\ c_{31} & c_{32} & 0 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ c_{i1} & c_{i2} & c_{i3} & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ c_{N1} & c_{N2} & c_{N3} & \dots & c_{Ni} & \dots & 0 \end{bmatrix}. \tag{15}$$

The scheme corresponding to the matrix $\mathbf{T}_\mathrm{C}$ can be considered as a template Tmp$_N$ from which all possible concrete schemes with N nodes can be obtained. Obviously, there is a recursion between templates Tmp$_N$ and Tmp$_{N+1}$ (Fig. 4).

The template Tmp$_4$ is described by the matrix

$$\mathbf{T}_\mathrm{C} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ c_{21} & 0 & 0 & 0 \\ c_{31} & c_{32} & 0 & 0 \\ c_{41} & c_{42} & c_{43} & 0 \end{bmatrix} \tag{16}$$

and has the form shown in Fig. 5.

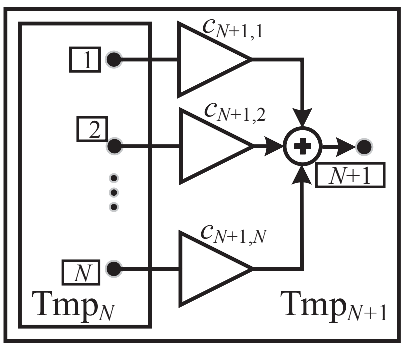

Fig. 4. Recursion at creation of templates

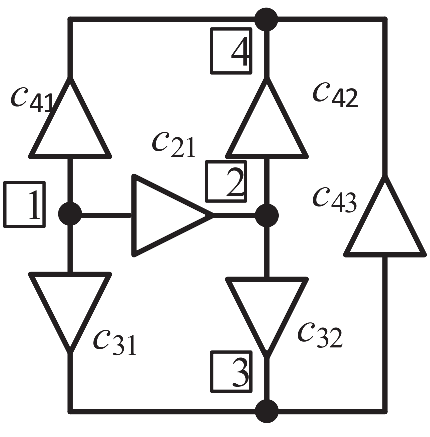

Fig. 5. Template Tmp$_4$ for a block diagram with $N = 4$ nodes

In Fig. 6, template Tmp$_5$ is presented. It is described by the matrix

$$\mathbf{T}_\mathrm{C} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ c_{21} & 0 & 0 & 0 & 0 \\ c_{31} & c_{32} & 0 & 0 & 0 \\ c_{41} & c_{42} & c_{43} & 0 & 0 \\ c_{51} & c_{52} & c_{53} & c_{54} & 0 \end{bmatrix}. \tag{17}$$

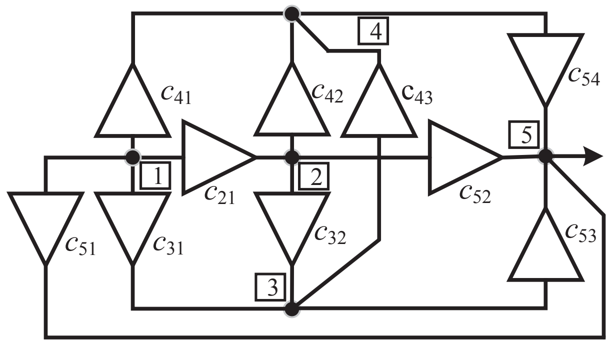

Fig. 6. Template Tmp$_5$ for a block diagram with $N = 5$ nodes

The matrix $\mathbf{T}_\mathrm{Z}$ specifies a subset of structures among structures with the same number of nodes. We confine ourselves to structures in which the number of delay blocks is equal to the order of the filter. In this case, in the $\mathbf{T}_\mathrm{z}$ and $\mathbf{T}(z)$ matrices, the elements z^{-1} are located above the main diagonal. Moreover, in each row and in each column of these matrices there cannot be more than one element z^{-1}.

The position of the elements in the topological matrix allows naming such structures. The authors proposed a set of structures with a given number of nodes and with a given TK matrix to put the name in correspondence as

$$\mathrm{N\{N\}z\{n\}p\{p1\}d\{d1\}\dots p\{pn\}d\{dn\}}.$$

This name describes many structures without specifying the numbers of the input and output nodes. In case of specifying the numbers of the input and output nodes, the name is

$$\mathrm{N}\{N\}\mathrm{z}\{n\}\mathrm{p}\{p_1\}\mathrm{d}\{d_1\}\dots\mathrm{p}\{p_\mathrm{n}\}\mathrm{d}\{d_\mathrm{n}\}\mathrm{i}\{inp\}\mathrm{o}\{out\}.$$

In the names of specific structures instead of curly brackets put specific numeric values: N is the number of nodes in the structure, n is the number of delay blocks, inp is the input node number, out is the output node number, p_i is the number of the line in which the i$^\mathrm{th}$ element z^{-1} is located,

$$d_\mathrm{i} = s_\mathrm{i} - p_\mathrm{i} + 1, \tag{18}$$

where s_i is the column number in which the i$^\mathrm{th}$ element z^{-1} is located.

V. Generating of Structures

The generation of structures with a given number of nodes N consists in the selection of the corresponding matrix $\mathbf{T}_\mathrm{C}$ and the generation of sets of elements z^{-1} that satisfy the above limitations. Below, as an example, we describe the generation of structures with N=4 and n=2. In this case, the matrix $\mathbf{T}_\mathrm{C}$ is described as (16).

A. N4z2p1d2p2d2

The matrix $\mathbf{T}(z)$ for this structure has the form

$$\mathbf{T}(z)=\begin{bmatrix} 0 & z^{-1} & 0 & 0 \\ c_{21} & 0 & z^{-1} & 0 \\ c_{31} & c_{32} & 0 & 0 \\ c_{41} & c_{42} & c_{43} & 0 \end{bmatrix} \qquad (19)$$

and the pattern is shown in Fig. 7.

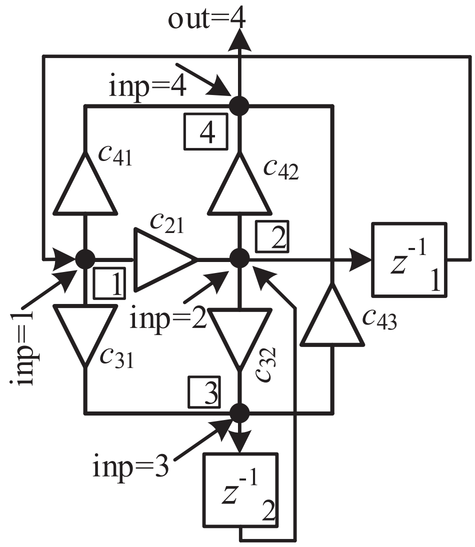

Fig. 7. Pattern of structures N4z2p1d2p2d2i*o4

For this pattern, any node can be input. Since there is no direction from the node number 4 to other nodes, only it can be the output node.

The transfer functions for structure N4z2p2d2p3d3i3o4 is

$$\mathrm{H}_{43}(z)=\frac{c_{41}z^{2}+\left(c_{42}-c_{21}c_{43}\right)z+c_{43}}{z^{2}-\left(c_{21}+c_{32}\right)z-c_{31}}.$$

When c_{21}=0, a canonical structure with a transfer function (3) is formed from the pattern. In this case, $c_{41}=b0$, $c_{42}=b_1$, $c_{43}=b_2$, $c_{32}=-a_1$, $c_{31}=-a_2$ (Fig. 2).

B. N4z2p2d2p3d2

The corresponding pattern is shown in Fig. 8. In this case, the topological matrix is expressed as

$$\mathbf{T}(z)=\begin{bmatrix} 0 & 0 & 0 & 0 \\ c_{21} & 0 & z^{-1} & 0 \\ c_{31} & c_{32} & 0 & z^{-1} \\ c_{41} & c_{42} & c_{43} & 0 \end{bmatrix} \qquad (20)$$

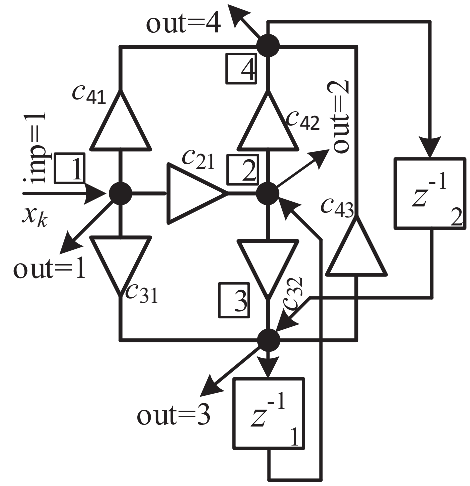

Fig. 8. Pattern of structures N4z2p2d2p3d2i1o*

In this pattern, any node can be an output, and only the node with the number 1 can be an input, since there is no transmission in it. The transfer functions for structure N4z2p2d2p3d3i1o2 is

$$\mathrm{H}_{21}(z)=\frac{c_{41}z^{2}+\left(c_{31}-c_{21}c_{43}\right)z+c_{21}}{z^{2}-\left(c_{43}+c_{32}\right)z-c_{42}}.$$

When c_{43}=0, a transposed canonical structure with a transfer function (3) is formed from the pattern. In this case, $c_{41}=b0$, $c_{31}=b_1$, $c_{21}=b_2$, $c_{32}=-a_1$, $c_{42}=-a_2$.

*C. N4z2p1d2p2d3i*o**

The topological matrix is equal to

$$\mathbf{T}(z)=\begin{bmatrix} 0 & z^{-1} & 0 & 0 \\ c_{21} & 0 & 0 & z^{-1} \\ c_{31} & c_{32} & 0 & 0 \\ c_{41} & c_{42} & c_{43} & 0 \end{bmatrix} \qquad (21)$$

The pattern of the pattern is shown in Fig. 9.

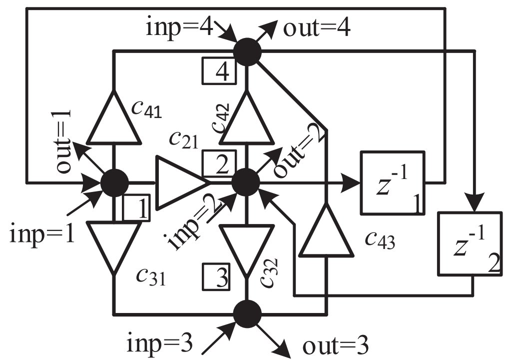

Fig. 9. Pattern of structures N4z2p1d2p2d3i*o*

Any node can be both input and output.

*D. N4z2p1d3p3d2i*o**

The block diagram is shown in Fig. 10.

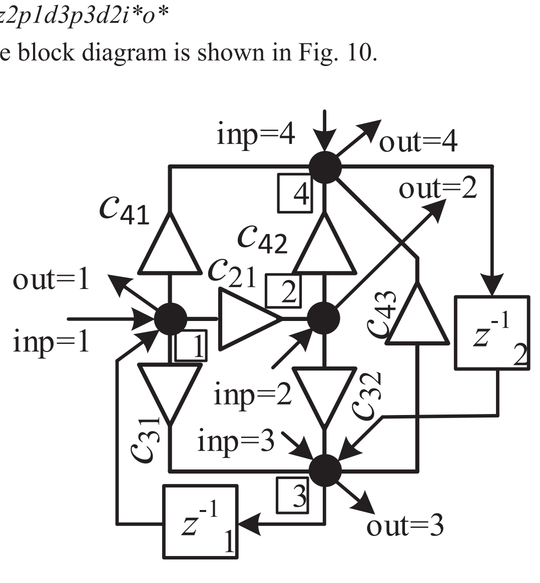

Fig. 10. Pattern of structures N4z2p1d3p3d2i*o*

This scheme corresponds to the topological matrix

$$\mathbf{T}(z)=\begin{bmatrix} 0 & 0 & z^{-1} & 0 \\ c_{21} & 0 & 0 & 0 \\ c_{31} & c_{32} & 0 & z^{-1} \\ c_{41} & c_{42} & c_{43} & 0 \end{bmatrix} \tag{22}$$

The input and output for this pattern can be in any node.

*E. N4z2p1d3p2d3i*o**

The structures are shown in Fig. 11.

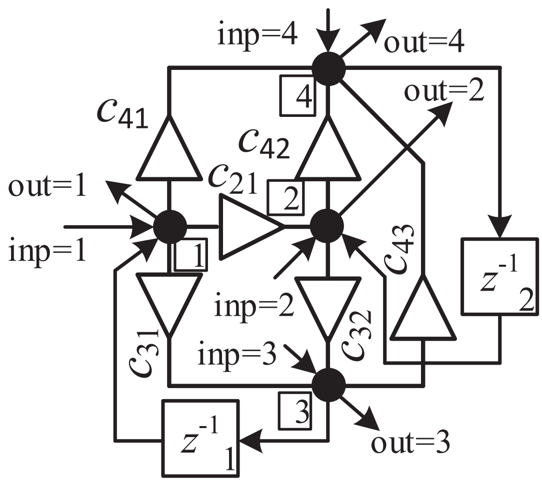

Fig. 11. Pattern of structures N4z2p1d3p2d3i*o*

This scheme is described by a topological matrix

$$\mathbf{T}(z)=\begin{bmatrix} 0 & 0 & z^{-1} & 0 \\ c_{21} & 0 & 0 & z^{-1} \\ c_{31} & c_{32} & 0 & 0 \\ c_{41} & c_{42} & c_{43} & 0 \end{bmatrix} \tag{23}$$

Any node can be both input and output.

*F. N4z2p1d4p2d2i*o**

The matrix $\mathbf{T}(z)$ for this structure has the form

$$\mathbf{T}(z)=\begin{bmatrix} 0 & 0 & z^{-1} & 0 \\ c_{21} & 0 & 0 & 0 \\ c_{31} & c_{32} & 0 & z^{-1} \\ c_{41} & c_{42} & c_{43} & 0 \end{bmatrix} \tag{24}$$

and the pattern is shown in Fig. 12.

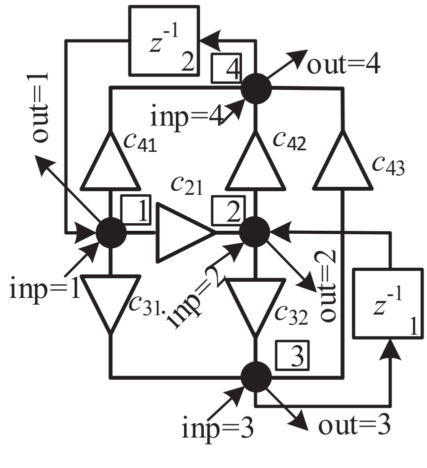

Fig. 12. Pattern of structures N4z2p1d4p2d2i*o*

Any node can be both input and output.

VI. Transition from Topological Matrices to the State Space

The matrix of transfer functions between nodes linear system C is equal to

$$\mathbf{H}_{\mathrm{C}}=\left(\mathbf{E}_N-\mathbf{T}_{\mathrm{C}}\right)^{-1}. \tag{25}$$

All elements of the matrices **A**, **b**, **c**, d describing the state space are elements of the matrix $\mathbf{H}_{\mathrm{C}}$ [21]. Instead of solving a difficult task of symbolic inversion of the matrix, we get analytical expressions for matrix elements.

When N=3

$$\mathbf{H}_{\mathrm{C}}=\begin{bmatrix} 1 & 0 & 0 \\ c_{21} & 1 & 0 \\ c_{21}c_{32}+c_{31} & c_{32} & 1 \end{bmatrix}. \tag{26}$$

At N=4

$$\mathbf{H}_{\mathrm{C}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ c_{21} & 1 & 0 & 0 \\ c_{21}c_{32} + c_{31} & c_{32} & 1 & 0 \\ c_{41} + c_{21}c_{42} + c_{31}c_{43} + c_{21}c_{32}c_{43} & c_{32}c_{43} + c_{42} & c_{43} & 1 \end{bmatrix}. \quad (27)$$

Considering the peculiarities of the inversion of triangular matrices we write recurrent expressions for calculating the elements $H_{\mathrm{C,ij}}$ of the matrix $\mathbf{H}_{\mathrm{C}}$:

$$H_{\mathrm{C},ij} = \begin{cases} 0, & i < j, \\ 1, & i = j, \\ \sum_{m=1}^{i-j} H_{\mathrm{C},i-m,j} c_{i,j-m}, & i > j. \end{cases} \quad (28)$$

For the convenience of calculating the elements of the state space matrices for N=4 and n=2, TABLE 1 lists the numbers of input and output nodes of the delay blocks.

TABLE I. INPUT AND OUTPUT NODES OF DELAY UNITS

		Delay unit 1		Delay unit 2	
		input	output	input	output
1	N4z2p1d2p2d2i*o4	2	1	3	2
2	N4z2p2d2p3d2i1o*	3	2	4	3
3	N4z2p1d2p2d3i*o*	2	1	4	2
4	N4z2p1d3p3d2i*o*	3	1	4	3
5	N4z2p1d3p2d3i*o*	3	1	4	2
6	N4z2p1d4p2d2i*o*	3	2	4	1

The position of the elements of the state matrices **A** for the six considered patterns is shown in Fig. 13.

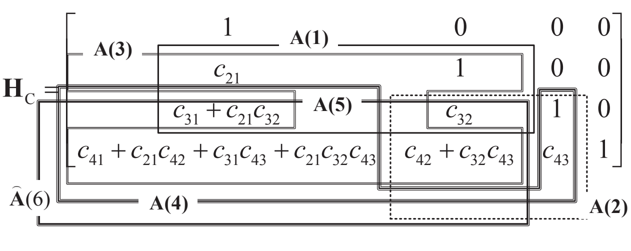

Fig. 13. Elements of matrices **A(*)** in matrix $\mathbf{H}_{\mathrm{C}}$

For the first five patterns for the first delay block, the numbers of both the input and the output nodes are less than for the second delay block. For the sixth pattern for the first delay block, the number of the input node is less than for the second, and the number of the output node is greater. Therefore, in Fig. 13 shows the position of the elements of the matrix

$$\widehat{\mathbf{A}}(6) = \mathbf{A}(6)\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \quad (29)$$

The position of the elements of the matrices A does not depend on the numbers of the input and output nodes of the filter.

In fig. 14 shows the arrangement of matrix elements **b** for the first pattern for different numbers of the input filter node. Also the state matrix **A**(1) is shown.

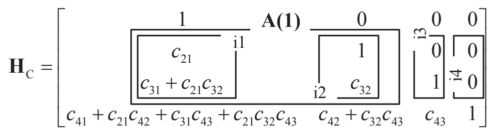

Fig. 14. Elements of **b** in matrix $\mathbf{H}_{\mathrm{C}}$ for the first pattern

For the second, fourth, fifth and sixth patterns, the numbers of the input nodes of the delay blocks are the same, so the matrices **b** have the same arrangement for these patterns (Fig. 15).

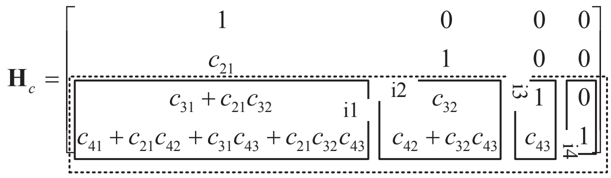

Fig. 15. Elements of **b** in matrix $\mathbf{H}_{\mathrm{C}}$ for the patterns 2, 4, 5, and 6

In Fig. 16 positions of matrices **b** for third pattern is shown.

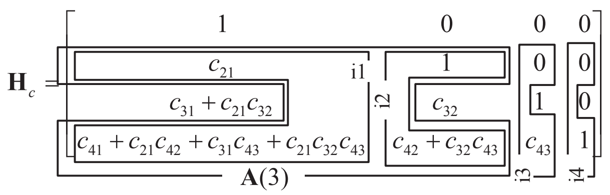

Fig. 16. Elements of **b** in matrix $\mathbf{H}_{\mathrm{C}}$ for the third pattern

In Fig. 17 positions of matrices **c** for the patterns 1, 3, and 5 for different numbers of output nodes is shown.

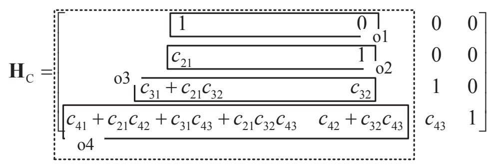

Fig. 17. Elements of **c** in matrix $\mathbf{H}_{\mathrm{C}}$ for the first, third and fifth patterns

Fig. 18 and Fig. 19 depict the position of the matrices c for the second and fourth patterns, respectively.

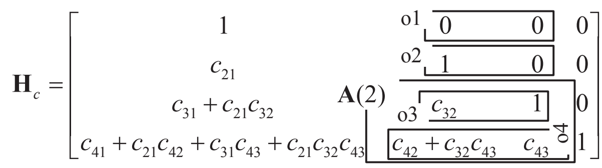

Fig. 18. Elements of **c** in matrix $\mathbf{H}_{\mathrm{C}}$ for the second pattern

$$
\mathbf{H}_c = \begin{array}{c} \mathrm{o1} \\ \mathrm{o2} \\ \mathrm{o3} \\ \mathrm{o4} \end{array} \begin{bmatrix} 1 & 0 & 0 & 0 \\ c_{21} & 1 & 0 & 0 \\ c_{31} + c_{21}c_{32} & c_{32} & 1 & 0 \\ c_{41} + c_{21}c_{42} + c_{31}c_{43} + c_{21}c_{32}c_{43} & c_{42} + c_{32}c_{43} & c_{43} & 1 \end{bmatrix}
$$

Fig. 19. Elements of **c** in matrix $\mathbf{H}_C$ for the fourth pattern

For the scalar d, the ratio

$$d = H_{C,oi} \tag{30}$$

is true, where o and i are the output and input nodes, respectively.

VII. Conclusion

The results described in this paper are supposed to be used as follows. At the stage of functional synthesis, the values of zeros and poles that are unchanged at subsequent stages are calculated. Further, taking into account the characteristics of structural accuracy and structural complexity, the degree of algebraic numbers, which are zeros and poles, the appropriate filter structures are generated, which are described by topological matrices. To select the generated structures, the noise level due to quantization coefficients at the filter output, sensitivity measures, the possibility of parasitic oscillations of the limiting cycle are analyzed. The results obtained for the state-space digital filters are used for this. The results described here are used to calculate the state-space.

Acknowledgment

The work was supported by a grant from the Russian Foundation for Basic Research 18-07-00986.

References

[1] C. T. Mullis and R. A. Roberts, “Synthesis of minimum roundoff noise fixed point digital filters,” *IEEE Trans. Circuits Syst.*, vol. CAS-23, no. 9, pp. 551–562, Sept. 1976.

[2] S. Y. Hwang, “Minimum uncorrelated unit noise in state-space digital filtering”, *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-25, no. 4, pp. 273–281, Aug. 1977.

[3] M. Kawamata and T. Higuchi, “A unified approach to the optimal synthesis of fixed-point state-space digital filters”, *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, no. 4, pp. 911–920, Aug. 1985.

[4] G. F. Dehner, "Noise optimized IIR digital filter design–tutorial and some new aspects", *Signal Processing*, vol. 83, no. 8, 2003, pp. 1565-1582

[5] J. C. Geromel, R. A. Borges, "Joint optimal design of digital filters and state-space realizations", *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 53, no. 12, pp. 1353-1357, Dec. 2006.

[6] T. Hilaire, P. Chevrel, J. F. Whidborne, "A unifying framework for finite wordlength realizations", *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 54, no. 8, pp. 1765-1774, 2007

[7] Z. Zhao, “An efficient state-space realization with minimum roundoff noise gain”, *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 54, issue 2, Feb. 2007, pp. 432-440.

[8] T. Hinamoto, T. Kawagoe, "Optimal synthesis of state-estimate feedback controllers with minimum l_2-sensitivity and no overflow oscillations", *2007 IEEE International Symposium on Circuits and Systems (ISCAS 2007)*, New Orleans, LA, USA, 27-30 May 2007, pp. 3996-3999, 2007.

[9] T. Hilaire, "On the Transfer Function Error of State-Space Filters in Fixed-Point Context", *Circuits and Systems II: Express Briefs IEEE Transactions on*, vol. 56, no. 12, pp. 936-940, 2009.

[10] T. Hinamoto, A. Doi, W.-S. Lu, "Jointly optimal error feedforward high-order error feedback and realization for roundoff noise minimization in IIR digital filters", *2014 IEEE International Symposium on Circuits and Systems (ISCAS)*, Melbourne VIC, Australia, 1-5 June 2014, pp. 982-985.

[11] Y. Hinamoto and A. Doi, "Minimized roundoff noise and pole sensitivity subject to L_2-scaling constraints for IIR filters", *2015 23rd European Signal Processing Conference (EUSIPCO)*, Nice, France, 31 Aug.-4 Sept. 2015, pp. 1123 - 1127

[12] A. Volkova, T. Hilaire, C. Lauter, "Determining fixed-point formats for a digital filter implementation using the worst-case peak gain measure", *2015 49th Asilomar Conference on Signals Systems and Computers*, pp. 737-741, Nov 2015.

[13] T. Hinamoto, A. Doi, W.-S. Lu, "Realization with minimal weighted pole and zero sensitivity subject to ℓ_2-scaling constraints for recursive digital filters", *2016 IEEE International Symposium on Circuits and Systems (ISCAS)*, Montreal, QC, Canada, 22-25 May 2016, pp. 2178-2181

[14] S. Yamaki, M. Abe, M. Kawamata, "Conjecture on the absence of limit cycles in second-order digital filters with minimum L_2-sensitivity subject to L_2-scaling constraints", *Journal of Signal Processing*, vol. 21, issue 2, pp. 47-51, 2017.

[15] P. S. R. Diniz, E. A. B. da Silva and S. L. Netto, *Digital Signal Processing: System Analysis and Design*: 2nd edition, New York: Cambridge University Press, 2010.

[16] V. Lesnikov, T. Naumovich, A. Chastikov, S. Armishev, “A new paradigm in design of IIR digital filters”, *2010 IEEE East-West Design & Test Symposium (EWDTS 2010)*, Saint Petersburg, Russia, 17-20 Sept. 2010, pp. 282 – 285.

[17] V. Lesnikov, T. Naumovich, A. Chastikov, S. Armishev, “Implementation of a new paradigm in design of IIR digital filters”, *2010 IEEE East-West Design & Test Symposium (EWDTS 2010), Saint Petersburg, Russia, 17-20 Sept. 2010*, pp. 156 –159.

[18] V. Lesnikov, T. Naumovich, “Generation and enumeration of structures of IIR digital filters”, *GSPx-2005, Pervasive Signal Processing (The Embedded Signal Processing Conference)*, Santa Clara, Ca, USA, October 24 - 27, 2005 (ISBN 0-9728718-2-9, paper number: 1837).

[19] V. Lesnikov, T. Naumovich, and A. Chastikov, “Generation and decomposition of digital filter topology”, *2017 IEEE East-West Design & Test Int. Symposium (EWDTS 2017)*, Novi Sad, Serbia, 29 Sept.-2 Oct. 2017, pp. 1 – 4.

[20] R. E. Crochier, and A. V. Oppenheim, “The analysis of linear digital circuits”, *Proceedings of IEEE*, 1975, 63 (4), pp. 581 – 595.

[21] V. Lesnikov, T. Naumovich, A. Chastikov, and S.Armishev, “Universal technique of the analysis of round-off noise in digital filters with arbitrary structure described by topological matrixes”, *2013 IEEE East-West Design & Test Symposium (EWDTS 2013)*, Rostov-on-Don, Russia, 27-30 Sept. 2013, pp. 1 – 4.

[22] T. Hinamoto, W.-S. Lu, *Digital Filter Design and Realization*, Delft: River Publishers, 2017.