Nonlinear Decomposition Filters with Neural Network Elements

VAIRIS SHTRAUSS
Institute of Polymer Mechanics
University of Latvia
23 Aizkraukles Street, LV 1006 Riga
LATVIA

Abstract: - The paper is devoted to finding ways for improving resolution and accuracy for decomposition of monotonic time- and frequency-domain multi-component signals. For solving the problem, a nonlinear decomposition filter is proposed operating with equally spaced data on a logarithmic time or frequency scale (geometrically spaced on linear scale), which is implemented as a parallel connection of several linear filters, which output signals are transformed by a nonlinear activation function, multiplied by weights and summed. One of the fundamental findings of this study is a square activation function, which ensures physically justified nonnegativity for the recovered distributions of time constants (DTC). It is found that the nonlinear decomposition filter under consideration transforms the Gaussian input noise into the nonnegative output noise with a specific probability distribution having the standard deviation and the mean proportional to the variance of input noise. For most practical cases when the standard deviation of input noise is small, the proposed solution considerably improves the noise immunity of algorithms. Enhancement in the resolution to compare with linear filters is demonstrated for the decomposition of a frequency-domain multi-component signal.

Key-Words: - Decomposition, Monotonic Multi-Component Signals, Distribution of Time Constants, Nonlinear Filters, Square Activation Function, Artificial Neural Networks

1 Introduction
Most materials and various physical objects have responses, which can be considered as the superposition of some monotonic components. These monotonic components often contain very useful information about the object. Despite of huge effort devoted, decomposition of monotonic multi-component signals remains a challenging signal processing problem due to: (i) their exceedingly non-orthogonal behaviour no constituting an orthogonal base, and (ii) the fundamental ill-posedness in the sense that small perturbations in input signal can yield unrealistic high perturbations in the results.

There are two classical approaches for decomposition and analysis of monotonic signals based on the deconvolution in the frequency domain. In the first [1,2], the deconvolution is carried out by: (i) two discrete Mellin transforms of an input signal and the appropriate mono-component (kernel of integral transform), (ii) their division, and (iii) taking the inverse discrete Mellin transform for the division. The second – the Fourier transform based approach [3-6] also performs three similar operations, but the direct discrete Mellin transforms are replaced by the discrete Fourier transforms of the signals, whose independent variables are logarithmically transformed, and the inverse discrete Fourier transform is applied to the division.

Recently, decomposition filters [7-10] have been proposed, which improve the accuracy of the decomposition and simplify its implementation. The decomposition filters invert an integral transform

\[ x(u) = \int_0^\infty f(\tau)K(u,\tau)d\tau / \tau \]

representing the model for a monotonic time- and frequency-domain multi-component signal with the following mono-components

\[ K(u,\tau) = \begin{cases} \exp(-u/\tau) & \text{(1a)} \\ \exp(-u/\tau)/\tau & \text{(1b)} \\ 1 - \exp(-u/\tau) & \text{(1c)} \\ 1/(1+u^2\tau^2) & \text{(1d)} \\ u\tau/(1+u^2\tau^2) & \text{(1e)} \\ u^2\tau^2/(1+u^2\tau^2) & \text{(1f)} \end{cases} \]
where variable \( u \) represents time or frequency, \( \tau \) is time constant, and \( f(\tau) \) is function of distribution of time constants (DTC).

The proposed decomposition filters have three basic features not inherent conventional discrete-time filters [11]. First, they operate on a logarithmic domain, i.e. process equally spaced data on a logarithmic scale, which manifest as geometrically spaced samples on linear scale

\[
u_n = u_0 q^n, \quad n = 0, \pm 1, \pm 2, \ldots, \quad q > 1,
\]
to carry out two general algorithms [7]

\[
f(u_0 q^n) = \sum_{n=-(N-1)/2}^{(N-1)/2} h_n x(u_0 q^{-n}), \quad (2)
\]

for odd number of coefficients \( N \), and

\[
f(u_0 q^n) = \sum_{n=-(N-2)/2}^{(N-2)/2} h_n x(u_0 q^{-n-1/2}), \quad (3)
\]

for even number of coefficients. Progression ratio \( q \) in Eqs. (2) and (3) specifies the sampling rate in the sense that \( \ln q \) plays formally a role of sampling period on a logarithmic scale, whereas its reciprocal represents the appropriate sampling frequency.

Second, the specification (the progression ratio and the number of coefficients) of a decomposition filter is determined in the regularization process [12,13] based on controlling the filter bandwidth by choosing the appropriate sampling rate, which allows ensuring the desired noise immunity for the algorithm.

Third, the decomposition filters are designed by the identification method [14] executing the time-domain optimization where a pair of theoretical functions interrelated with each other by the theoretical deconvolution is used as input and output signals in the filter design and the coefficients are determined by minimizing the error between the exact output function and that obtained by filtering. An advantage of the identification method representing actually a learning algorithm is that it effectively disposes of various secondary effects such as rounding-off, truncation errors, etc.

Practice has shown that the decomposition filters give good results [10] for relatively broad and moderately asymmetric DTC, however, it is problematic to recover discrete, very narrow and essentially asymmetric ones. A limit in resolution of discrete DTCs for the decomposition filters can be estimated as ratio 5 for two time constants [7-9]. Attempts to achieve the better resolution lead to the oscillating solutions producing DTCs with false peaks (lines) and peaks with non-physical negative amplitudes.

The presented contribution is devoted to finding new approaches and algorithms for further enhancing the resolution and accuracy for decomposition of monotonic multi-component signals.

Employment of the strategies of standard artificial neural networks [15] has been studied. However, these studies did not succeed in significant results, and, in most cases, neural net solutions did not give the essential improvement to compare with the results of the decomposition filters. It has been established that commonly used activation functions (sigmoid, radial basis functions, etc.) [15] do not support the specific constant-area normalization principle of DTC curves [16-18] resulting in large amplitudes for narrow and discrete (line) distributions.

**2 Nonlinear Decomposition Filters**

Fig. 1 illustrates the decomposition process for a filter (inside the dashed box) having the impulse response with 3 coefficients. To implement discrete deconvolution, the impulse response \( h_n \) is flipped left-for-right. An output DTC sample is obtained as a weighted sum of the samples of an input function, i.e. is produced by multiplying 3 samples (e.g. 2, 3 and 4) of the input function by the filter coefficients and adding the products. To obtain next output sample the decomposition filter is moved one sample to the left or right.

On the other hand, a decomposition filter can be considered as a single artificial neuron or node [15].
with weights \( h_n \) having linear activation function (not shown in Fig. 1 as no affecting the output). The neuron conception becomes more well-grounded to remember that the decomposition filters are designed by the identification method [14] implementing actually a specific learning algorithm.

Fig. 2. Schematic illustration of the decomposition process by a nonlinear filter or a decomposition network with a hidden layer consisting of two neurons.

Based on this filter-neuron analogy, a nonlinear decomposition filter or a decomposition network has been formed as a parallel connection of several linear filters, which output signals are transformed by a nonlinear activation function, multiplied by weights, and summed (Fig. 2). It must be stressed that the moving deconvolution principle on a logarithmic scale is maintained, i.e. a nonlinear decomposition filter operates with equally spaced data on a logarithmic scale (geometrically spaced on linear scale), and moves to the left or right for obtaining output samples at various time constants.

We find that a square activation function \( g(x) = x^2 \) contrary to conventional activation functions (sigmoid, radial basis functions, etc.) [15] provides some very useful features for DTC recovery, such as:

(i) ensures physically justified nonnegativity for the recovered DTC due squaring the output signals,

(ii) increases radically the noise immunity to compare with the linear decomposition filters due to a small gain for small signals, and

(iii) facilitates supporting the constant-area normalization principle of DTC curves [16-18] by lengthening the peak amplitudes for narrow distributions due to a high gain for large signals.

Simulations have revealed that a hidden layer consisting of some neurons with the square activation function and an output layer with a neuron with the linear activation function (again not shown in Fig. 2) are an optimal architecture for the decomposition problems.

For the considered above nonlinear filters, algorithms (2) and (3) modify into

\[
 f(u_o q^n) = \sum_{k=1}^{K} w_k \left( \sum_{m_{(N-k)/2}}^{(N-1)/2} h_n x(u_o q^{m-n}) \right)^2 ; \quad (4)
\]

\[
 f(u_o q^n) = \sum_{k=1}^{K} w_k \left( \sum_{m_{(N-2)/2}}^{(N-1)/2} h_n x(u_o q^{m-0.5-n}) \right)^2 \quad (5)
\]

respectively, where \( K \) is number of neurons in the hidden layer, and \( w_k \) is a weight for the output of \( k \)-th neuron.

3 Noise Transformation

3.2 Linear Filters

The noise behaviour of a linear decomposition filter, as a linear system (on the logarithmic domain) is rather simple. Random input signal or noise with a Gaussian probability distribution produces an output signal that also have a Gaussian probability distribution [11] and a linear decomposition filter amplifies linearly (proportionally) input noise variance \( \sigma^2_x \) to yield the output noise variance \( \sigma^2_y \):

\[
 \sigma^2_y = \sigma^2_x \sum h_n = \sigma^2_x S ,
\]

where proportional coefficient \( S = \sum h_n \) is named ‘noise coefficient’. Summation index \( n \) runs in accordance with Eqs. (2) or (3).

As shown in [7-10], the noise coefficient depends on the filter bandwidth and, consequently, on progression ratio \( q \) and the large values of \( S \) is a reason for the ill-posedness in the decomposition. A
regularization method developed in [12,13] controls
the filter bandwidth by choosing the appropriate
sampling rate, to ensure in such way the desired noise
coefficient.

3.2 Nonlinear Filters
A nonlinear decomposition filter exhibits the more
complex noise behaviour. It can be shown that the
squaring the Gaussian input noise with zero mean
\( \mu = 0 \) produces a nonnegative output noise having
specific probability distribution function
\[
p(x) = \begin{cases} 0 & \text{for } x < 0 \\ \frac{3}{\sigma_x^2} \exp\left(-\frac{6x}{\sigma_x^2}\right) & \text{for } x \geq 0 \end{cases}
\]
depending on standard deviation \( \sigma_x \) of the input
noise. Distribution (6) has nonzero mean
\[
\mu_{\text{square}} = \sigma_x^2
\]
and standard deviation
\[
\sigma_{\text{square}} = 1.26 \mu = 1.26 \sigma_x^2
\]
proportional to mean (7).

Therefore, the output noise of a neuron in the
hidden layer has probability distribution (6), and its
mean \( \mu_k \) depending on weights \( h_n^k \) is equal to
\[
\mu_k = \sum_n (h_n^k)^2 \sigma_x^2
\]
In its turn, mean of the output noise of an entire
nonlinear decomposition filter can be calculated as
\[
\mu_y = \sum_{k=1}^{K} w_k \sum_n (h_n^k)^2 \sigma_x^2
\]
The probability distribution of the output noise of an
entire nonlinear filter shapes from the distributions of
individual neurons, and for small number of
neurons, the function is similar to that of (6).

Therefore, both standard deviation and mean of
the output noise of a nonlinear decomposition filter
are proportional to the variance of the input noise.
For most practical cases, when \( \sigma_x < 1 \), it results in
smaller the standard deviation of the output noise to
compare with that of the input noise. The nonzero
mean of output noise would be considered as a
shortcoming for nonlinear filters, fortunately, its
influence usually is not so essential.

In Fig. 3, a histogram of the values of output
noise is shown for a nonlinear filter with two
neurons \( K = 2 \) in the hidden layer having eight
inputs \( N = 8 \) for the Gaussian input noise with
amplitude 0.05 and zero mean. The histogram is in
good agreement with probability distribution function (6). As seen, the values output noise practically does not exceed input noise amplitude
0.05, and main part of the values is concentrated
within interval [0, 0.01]. In the table bellow,
parameters of the input and output noise are
compared obtained from the histogram and
theoretical evaluations (8) and (9).

<table>
<thead>
<tr>
<th>Table. Noise parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma )</td>
</tr>
<tr>
<td>Input</td>
</tr>
<tr>
<td>Output (histogram)</td>
</tr>
<tr>
<td>Output (Eqs. (8), (9))</td>
</tr>
</tbody>
</table>

Thus, the proposed structure for the nonlinear
filters under consideration provides considerable
noise reduction. This allows decreasing progression
ratio \( q \), which leads to increase of decomposition
accuracy [7-10]. Analysis has shown that the
progression ratio from \( q = 3 \ldots 4 \), necessary to use
for linear filters to ensure acceptable low noise
coefficient of order 10, can be reduced to \( q = 2 \) or
even to smaller value. (The histogram in Fig. 3 is
obtained for a filter operating at \( q = 2 \)). It must
remember [10] that progression ratio \( q = 2 \) for linear
decomposition filters produces the noise coefficients
of an order of some thousands.
4 Simulation Results

In this Section, some simulation results are considered obtained by both linear and nonlinear filters for decomposition of frequency-domain multi-component signals with mono-components (1d).

For the linear case, a filter, described in [10], operating at $q = 3.3$ with 6 coefficients having noise coefficient $S = 2.28$ is used. For the nonlinear case, a filter operating at $q = 2$ containing two neurons with 8 inputs in the hidden layer has been designed. The histogram of the noise values for this filter is shown in Fig. 3.

In Figs. 4 and 5, the decomposition results are compared obtained from noiseless input data.

\[ x_{\text{noisy}}(u_m) = x_{\text{exact}}(u_m) + e \cdot n(m), \]  

where $n(m)$ is the pseudorandom sequence within interval [-1,1] with zero mean having the Gaussian probability distribution, and $e$ is a factor specifying amplitude of the input noise. The recovered noisy DTC is smoothed by simple 5-point averaging

\[ f(\tau_m) = \frac{1}{5} \sum_{n=-2}^{2} f(\tau_{m+n}). \]  

As it is seen from Fig. 4, both linear and nonlinear filter gives DTCs without non-physical oscillations, however the peaks obtained by the nonlinear filter are higher. If the limit in the resolution of linear decomposition filters can be estimated as $\tau_1/\tau_2 \geq 5$ ($\tau_2/\tau_1 \leq 0.2$), then the limit in the resolution for the nonlinear filter under consideration is $\tau_1/\tau_2 \geq 4$ ($\tau_2/\tau_1 \leq 0.25$). In Fig. 5, the limit case with $\tau_1/\tau_2 = 4$ is shown. If two separate distributions are seen still in their merging for the nonlinear filter, than the distributions are is fully melted together for DTCs recovered by the linear filter.

In Fig. 5, decomposition results are shown obtained from the noisy input data corrupted by additive random noise

\[ f(\tau) = \frac{1}{5} \sum_{n=-2}^{2} f(\tau_{m+n}). \]  

The results show that the noise can be effectively reduced by smoothing. The smoothed DTCs (dotted curves) are in rather good agreement with these
obtained from the noiseless input data. Thus, smoothing of recovered DTCs is strongly recommended for both linear and nonlinear decomposition filters.

5 Conclusions
To enhance resolution and accuracy for decomposition of monotonic multi-component signals, nonlinear filters are proposed implemented as a parallel connection of several linear decomposition filters, which output signals are transformed by the square activation function, multiplied by weights and summed. It is found that the square activation function has several notable features for decomposition of monotonic signals including physically justified nonnegativity of the recovered distributions of time constants (DTC) and increased noise immunity of the algorithms. In addition, the square activation function supports the constant-area normalization principle for the recovered DTC. It is found that a nonlinear decomposition filter transforms the Gaussian input noise into the nonnegative output noise with specific probability distribution having the standard deviation and the mean proportional to the variance of input noise. Simulation results of decomposition of a frequency-domain multi-component signal are compared obtained by linear and nonlinear filters. Enhancement in resolution of DTC is demonstrated for the proposed nonlinear filters.

References: