

# On the Time-Frequency Analysis Based Filtering

*Ljubiša Stanković*

*Abstract*— Efficient processing of nonstationary signals requires time-varying approach. An interesting research area within this approach is time-varying filtering. Since there is a certain amount of freedom in the definition of time-varying spectra, several definitions and solutions for the time-varying filtering have been proposed so far. Here we will consider the Wigner distribution based time-varying filtering form defined by using the Weyl correspondence. Its slight modification will be proposed and justified in the processing of noisy frequency modulated signals based on a single signal realization. An algorithm for the efficient determination of the filters' region of support in the time-frequency plane, in the case of noisy signals, will be presented. In the second part of the paper, the theory is applied on the filtering of multicomponent noisy signals. The S-method is used as a tool for the filters' region of support estimation in this case. This method, combined with the presented algorithm, enables very efficient time-varying filtering of the multicomponent noisy signals based on a single realization of the signal and noise. Theory is illustrated by examples.

## I. INTRODUCTION

Analysis and processing of stationary signals is usually performed either in time or in frequency domain. However, when the signals exhibit nonstationary characteristics more efficient processing can be done using joint time-frequency domain tools. They are based on the time-frequency representations of signal. Time-varying filtering is one of the challenging areas where one can benefit from the joint time-frequency representations. This kind of filtering can produce better results, in the nonstationary signal cases, than the processing of signals in either time or frequency domain separately. However, in the definition of time-varying spectra there is a certain amount of freedom [3], [5], [6], [8], [16], [26], what has resulted in several solutions for this ap-

proach. The first, classical solution, has been presented by Zadeh [25]. It can be related to the Richaczek distribution [16]. However, this distribution exhibits some serious drawbacks as a possible tool for the time-frequency representation of signals [6], [8]. This fact was the reason for a redefinition of the time-varying filtering relations into the Wigner distribution framework, using the Weyl correspondence [3], [9], [13], [14], [18]. In this paper we will consider the Wigner distribution approach. Its slight modification will be proposed and justified in the treatment of noisy frequency modulated signals when only one signal realization is available. The algorithm proposed in [23] has been used for determination of the filter's region of support. This algorithm is based on the optimal window length relation in the Wigner distribution. It uses a specific statistics approach of comparing the bias and variance [11], [23]. Theory is extended to the multicomponent noisy signals. The S-method is used as a basic tool for the time-varying filter support estimation [20], [21]. This method combined with the described algorithm, enables very efficient time-varying filtering of the multicomponent noisy signals, *based on a single signal realization*. Examples illustrate the presented theory.

The paper is organized as follows. The basic theory, including the definitions of time-varying filtering, its discrete and pseudo forms, along with the illustrations, are given in Section II. Filtering of a monocomponent signals is considered in the next Section. In Section III the algorithm for the region of support estimation is presented, as well. The algorithm efficiency in the time-varying filtering applications is illustrated on examples. Time-varying filtering of multicomponent signals is studied and illustrated in Section IV.

## II. BASIC THEORY

### A. Definitions

Time-varying filtering of a signal  $x(t)$  has been defined by, [13], [24], [25]:

$$(Hx)(t) = \int_{-\infty}^{\infty} h(t, t - \tau)x(\tau)d\tau, \quad (1)$$

where  $h(t, \tau)$  is the impulse response of the time-varying system  $H$ . If signal  $x(t)$  is a sum of a desired signal  $s(t)$  and noise  $\nu(t)$  then the system  $H$  may be determined by minimizing the mean square error [13], [24]

$$H_{opt} = \arg \min_H E \left\{ |s(t) - (Hx)(t)|^2 \right\}. \quad (2)$$

In the ideal case the system should produce [13]

$$(Hs)(t) = s(t), \quad (H\nu)(t) = 0. \quad (3)$$

Since there is a certain amount of freedom in the definition of a time-varying spectrum, several solutions of this problem have been proposed until now. The classical time-varying filter function proposed by Zadeh [25]

$$Z_H(t, \omega) = \int_{-\infty}^{\infty} h(t, t - \tau)e^{-j\omega\tau} d\tau \quad (4)$$

with the impulse response

$$h(t, t - \tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} Z_H(t, \omega)e^{j\omega\tau} d\omega. \quad (5)$$

It can be related to the Richaczek distribution. Since this distribution exhibits very serious drawbacks as tool for time-frequency analysis, this may be also told for the Zadeh time-varying filter function. This fact was the reason for a redefinition of the time-varying function using the Wigner distribution framework.

Time-varying transfer function, in the Wigner distribution framework, has been defined as the Weyl symbol mapping of the impulse response into the time-frequency plane

[13], [14], [18]

$$L_H(t, \omega) = \int_{-\infty}^{\infty} h\left(t + \frac{\tau}{2}, t - \frac{\tau}{2}\right)e^{-j\omega\tau} d\tau. \quad (6)$$

Using the desired properties of the time-varying system  $(Hs)(t) = s(t)$ , and  $(H\nu)(t) = 0$  and the Moyal's formula relating the inner products of signals and their Wigner distributions,

$$(WD_x(t, \omega), WD_y(t, \omega)) = |(x(t), y(t))|^2,$$

the following properties of the system should ideally hold

$$((Hs)(t), s(t)) = \int_{-\infty}^{\infty} |s(t)|^2 dt =$$

$$(1, WD_s(t, \omega)) = (L_H(t, \omega), WD_s(t, \omega)) \quad (7)$$

$$((H\nu)(t), \nu(t)) = 0 = (L_H(t, \omega), WD_\nu(t, \omega)). \quad (8)$$

Suppose that the Wigner distribution of the signal  $s(t)$  defined by

$$WD_s(t, \omega) = \int_{-\infty}^{\infty} s\left(t + \frac{\tau}{2}\right)s^*\left(t - \frac{\tau}{2}\right)e^{-j\omega\tau} d\tau$$

lies inside a region  $R$ , while the noise lies outside this area, except its small part that can be neglected with respect to the part of noise outside  $R$ . A simple solution satisfying requirements (7),(8) is then given by [13]

$$L_H(t, \omega) = \begin{cases} 1 & \text{for } (t, \omega) \in R \\ 0 & \text{for } (t, \omega) \notin R \end{cases} \quad (9)$$

Another way to get this relation is presented in the Appendix. Note that nothing would qualitatively change if  $L_H(t, \omega)$  assumed any other constant value within  $R$ . The impulse response is obtained as

$$h\left(t + \frac{\tau}{2}, t - \frac{\tau}{2}\right) = \frac{1}{2\pi} \int_{-\infty}^{\infty} L_H(t, \omega)e^{j\omega\tau} d\omega \quad (10)$$

We will use a slightly modified version of filtering relations (1). It turned out that it is

necessary in order to get undistorted frequency modulated signals, when time-varying filtering relation is applied. Time-varying filtering is here defined by

$$(Hx)(t) = \int_{-\infty}^{\infty} h(t + \frac{\tau}{2}, t - \frac{\tau}{2})x(t + \tau)d\tau \quad (11)$$

with the impulse response as in (10). Using the Parseval's theorem, we get

$$(Hx)(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} L_H(t, \omega)X(\omega)e^{j\omega t}d\omega. \quad (12)$$

The realizations and properties of this filtering relation will be studied next. The reason for introducing and using filtering relation (11) will be explained in the sequel. Its analysis is given in the Appendix.

*B. Illustration*

Consider an FM signal

$$s(t) = A(t)e^{j\phi(t)} \quad (13)$$

with a slow-varying amplitude  $A(t)$ , such that its Fourier transform may be obtained using the stationary phase method [3], [15]

$$FT[s(t)] = S(\omega) \cong \int_{-\infty}^{\infty} A(t_0)e^{j\phi(t_0)-j\omega t_0} \times \sqrt{\frac{2\pi j}{\phi''(t_0)}}\delta(t_0 - (\phi'(\omega))^{-1})dt_0$$

where  $(\phi'(\omega))^{-1}$  denotes an inverse function of the instantaneous frequency  $\phi'(t)$ . For asymptotic signals it is equal to the group delay function [3]. Assume that the Wigner distribution of  $s(t) + \nu(t)$  has provided a well localized information about the signal's support in the time-frequency domain so that [3], [6]

$$L_H(t, \omega) = \delta(\omega - \phi'(t)). \quad (14)$$

Note that in the discrete domain, as it is used in numerical realizations, function  $\delta(\omega)$  will be a Kronecker delta function and will satisfy

unity amplitude condition (9) for the time-varying filter. Applying now  $FT[s(t)]$  and  $L_H(t, \omega)$  in (12) we get

$$(Hx)(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(\omega - \phi'(t)) \times A(t_0)e^{j\phi(t_0)-j\omega t_0} \sqrt{\frac{2\pi j}{\phi''(t_0)}}\delta(t_0 - (\phi'(\omega))^{-1}) \times e^{j\omega t} dt_0 d\omega + \nu_{out}(t) = A(t)e^{j\phi(t)} \sqrt{\frac{2\pi j}{\phi''(t)}} + \nu_{out}(t) \quad (15)$$

since

$$\int_{\omega} \int_{t_0} F(t_0)\delta(\omega - \phi'(t)) \times \delta(t_0 - (\phi'(\omega))^{-1})dt_0 d\omega = F(t)$$

for monotone instantaneous frequency function. For the signals with linear instantaneous frequency, or the instantaneous frequency may be considered as a linear function of time within the considered interval, we get

$$(Hx)(t) = cs(t) = cA(t)e^{j\phi(t)}, \quad (16)$$

where  $c$  is a constant, since we assumed that the derivative of the instantaneous frequency is a constant. Other filtering relations than (11) would produce distorted versions of the original signal.

The amount of noise within the support defined by  $\delta(\omega - \phi'(t))$  is neglected, since its energy is reduced by the factor

$$Q = \frac{\iint_D d\omega dt}{\iint_D \delta(\omega - \phi'(t))d\omega dt} \gg 1, \quad (17)$$

where  $D$  is the whole considered time-frequency plane. In the case when the variations of the instantaneous frequency are not small, an amplitude compensation for the IF variations should be done. It could be done by estimating the direction of the instantaneous frequency in the time-frequency plane what, according to the assumed known region of support  $R$ , is not difficult.

**Note 1:** There is an ambiguity in the definition of  $L_H(t, \omega)$  in (14). Instead of the delta

pulse along the instantaneous frequency  $\phi'(t)$ , we may impose the same condition, but with respect to the group delay  $(\phi'(\omega))^{-1}$ , [3], [6], what will give the support function

$$L_H^{(g)}(t, \omega) = \delta(t - (\phi'(\omega))^{-1}). \quad (18)$$

It is interesting to point out that the support function defined as a geometrical mean of (14) and (18)

$$L_{Hg}(t, \omega) = \sqrt{\frac{1}{2\pi j} \delta(\omega - \phi'(t)) \delta(t - (\phi'(\omega))^{-1})} \quad (19)$$

will produce correct both amplitude and phase,

$$(Hx)(t) = A(t)e^{j\phi(t)} \quad (20)$$

for any frequency modulated signal, as far as the stationary phase method (15) holds, since  $\delta(t - (\phi'(\omega))^{-1}) = \delta(\omega - \phi'(t))\phi''(t)$  and  $L_{Hg}(t, \omega) = \delta(\omega - \phi'(t))\sqrt{\phi''(t)/(2\pi j)}$ . The requirement for monotonicity of  $\phi'(t)$  is already included in the stationary phase method definition.

**Note 2:** The signal to noise ratio improvement, that can be gained with the time-varying filter, is defined as a difference of the signal to noise ratio at the output ( $SNR_{out}$ ) and input ( $SNR_{in}$ ) of the system, i.e., as

$$G = SNR_{out} - SNR_{in}.$$

According to (17) and (20) we get

$$G = 10 \log Q. \quad (21)$$

Having in mind (17), this improvement can be significant.

**Note 3:** The Wigner distribution can produce a completely concentrated distribution along the instantaneous frequency (group delay) only for the linear frequency modulated signals. The error analysis and the optimal time-varying data-driven lag window length for the Wigner distribution of nonlinear frequency modulated signals is presented in [11]. If the instantaneous frequency variations are not linear then the complete concentration along the instantaneous frequency, i.e.,  $L_H(\omega, t) \sim \delta(\omega - \phi'(t))$ , may be obtained

using the S-distribution, [22] or the L-Wigner distribution [21]

$$SD(\omega, t) = \int_{-\infty}^{\infty} s^{[L]}(t + \frac{\tau}{2L}) s^{*[L]}(t - \frac{\tau}{2L}) e^{-j\omega\tau} d\tau$$

where for the L-Wigner distribution  $s^{[L]}(t)$  is the  $L$ -th power of the signal, while in the S-distribution only the signal's phase is multiplied by  $L$ . It has been shown that for any FM signal [22], [21]

$$\lim_{L \rightarrow \infty} SD(\omega, t) / A^{[2L]}(t) = \delta(\omega - \phi'(t)).$$

A similar concentration can be achieved by the reassignment method [2], as well. In the case of polynomial phase functions a complete distribution concentration may be produced using the polynomial Wigner-Ville distributions, [3], [17], or the local polynomial distributions [12].

### C. Discrete Form

In the numerical realizations a discrete form of system (11)-(12), given by

$$\begin{aligned} (Hx)(n) &= \sum_{m=-\infty}^{\infty} h(n + \frac{m}{2}, n - \frac{m}{2}) x(n + m) \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} L_H(n, \theta) X(\theta) e^{j\theta n} d\theta \quad (22) \end{aligned}$$

$$\begin{aligned} L_H(n, \theta) &= \sum_{k=-\infty}^{\infty} h(n + \frac{k}{2}, n - \frac{k}{2}) e^{-j\theta k} \\ X(\theta) &= \sum_{m=-\infty}^{\infty} x(m) e^{-j\theta m} \quad (23) \end{aligned}$$

should be used. This form will be used in the analysis starting from the next Section, as well.

### D. Pseudo forms

Consider now the previous definitions and representations with a limited lag variable  $\tau$ . This case is important for the practical realizations. Introducing a lag window  $w(\tau)$  in the filtering definitions we get

$$(Hx)(t) = \int_{-\infty}^{\infty} h(t + \frac{\tau}{2}, t - \frac{\tau}{2}) w(\tau) x(t + \tau) d\tau$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} L_H(t, \omega) STFT(t, \omega) d\omega \quad (24)$$

where

$$STFT(t, \omega) = \int_{-\infty}^{\infty} w(\tau) x(t + \tau) e^{-j\omega\tau} d\tau \quad (25)$$

is the short time Fourier transform of signal  $x(t)$ , with  $w(0) = 1$ . Applying the same analysis as in (15), and using (19) as the support, we get

$$\begin{aligned} (Hx)(t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(\omega - \phi'(t)) w(\tau_0) \times \\ &A(t + \tau_0) e^{j\phi(t + \tau_0) - j\omega\tau_0} \times \\ &\delta(t + \tau_0 - (\phi'(\omega))^{-1}) d\tau_0 d\omega + \nu_{out}(t) = \\ &= A(t) e^{j\phi(t)} + \nu_{out}(t). \end{aligned} \quad (26)$$

This is a quite interesting conclusion. The lag window does not influence the output signal  $(Hx)(t)$  as far as we are able to determine the region of support  $R$  and  $w(0) = 1$ .

In the realizations, the pseudo Wigner distribution

$$\begin{aligned} PWD_s(t, \omega) &= \\ \int_{-\infty}^{\infty} w_e(\tau) s(t + \frac{\tau}{2}) s^*(t - \frac{\tau}{2}) e^{-j\omega\tau} d\tau \end{aligned} \quad (27)$$

is used for the region of support determination, rather than the Wigner distribution itself. Note that the lag window  $w_e(\tau)$  in the Wigner distribution and the lag window  $w(\tau)$  in (24) are not related, so they may be optimized independently.

The window function  $w_e(\tau)$  will smooth the original Wigner distribution as

$$PWD_s(t, \omega) = WD_s(t, \omega) *_{\omega} W_e(\omega). \quad (28)$$

It will result in a wider region of support  $R_p = R + \Delta R$ , where  $\Delta R$  is determined by the window  $w_e(\tau)$ , i.e., its Fourier transform  $W_e(\omega)$ , length. This will cause that larger amount of the noise passes through the system, than it is theoretically necessary. At the first look it seems that the increase  $\Delta R$  should

not influence the output signal since  $R \subset R_p$ . Initially we made this conclusion, but after numerical simulations we have learned that as a result of  $\Delta R \neq 0$  some signal frequencies, that do not belong to the considered time instant  $t$ , may appear in the output signal and degrade performance. *This is the reason why other smoothed forms of the Wigner distribution, like for example the spectrogram, would not be appropriate for the region  $R$  determination and would not produce good results.* In order to reduce these effects the window  $w_e(\tau)$  should be as wide as possible, i.e., its Fourier transform  $W_e(\omega)$  should be as narrow as possible. This requirement is however contradictory to the reduction of noise and low calculation complexity requirements. Therefore, the optimization of  $w_e(\tau)$  length should be done.

### III. TIME-VARYING FILTERING OF A MONOCOMPONENT SIGNAL

If we have a stochastic process  $\{s(t)\}$ , whose large number of realizations  $s(t)$  is known, and zero-mean white noise  $\nu(t)$  being not correlated with signal, then the region of support  $R$  can be easily obtained. The mean of the Wigner distribution of  $x(t) = s(t) + \nu(t)$ ,  $\overline{WD}_x(t, \omega)$ , is just the Wigner distribution of nonnoisy signal superposed to a pedestal whose weight is equal to the noise variance [1], [7], [19]

$$\begin{aligned} \overline{WD}_x(t, \omega) &= \\ \int_{-\infty}^{\infty} E \left\{ x(t + \frac{\tau}{2}) x^*(t - \frac{\tau}{2}) \right\} e^{-j\omega\tau} d\tau \\ &= \overline{WD}_s(t, \omega) + \sigma_{\nu}^2 \end{aligned} \quad (29)$$

where  $\sigma_{\nu}^2$  is the variance of noise.

The situation significantly complicates if we have to perform time-varying filtering based on a single realization of the signal and noise. In practice this is a very common and important case. Here we can distinguish two steps: 1) Approximate Wigner distribution of the non-noisy signal  $s(t)$ , based on *the single noisy observation*, as accurately as possible, 2) Use this distribution to determine region of support  $R$  and perform time-varying filtering. The first step is crucial, especially in the cases of a very

high noise, i.e., low signal to noise ratio. In the sequel we will describe a method for the efficient implementation of this step, and then demonstrate its accuracy, performing the second (filtering) step. The analysis will be done in the discrete-time domain, since in the analysis of noise in the Wigner distribution in analog domain some forms are not well-defined.

#### A. Region of support $R$ estimation

Consider the case of a monocomponent deterministic discrete-time signal  $s(n)$  with a random white Gaussian complex noise  $\nu(n)$ , with independent real and imaginary parts. The variance of noise is assumed to be  $\sigma_\nu^2$ . This is formally the same situation as the one when the signal  $s(t)$  is random, but we know its single realization only.

The main problem lies in the determination of the region of support  $R$  for  $L_H(t, \omega)$ . It requires the knowledge of the Wigner distribution of signal  $WD_s(t, \omega)$ , see the Appendix. Therefore, the Wigner distribution  $WD_s(t, \omega)$  has to be estimated with the smallest possible error. There are two kind of errors in the pseudo Wigner distribution: one due to the used lag window causing bias and the other due to the noise manifesting itself through variance. Total squared error is defined as a sum of the distribution variance and squared bias. The minimization of this error gives the optimal distribution. It will be then used for the region of support estimation.

The Wigner distribution, in its pseudo form, of a discrete-time noisy signal  $x(n) = s(n) + \nu(n)$  is defined as

$$WD_x(n, \theta; N) = \sum_{k=-\infty}^{\infty} w_e(k)x(n+k)x^*(n-k)e^{-jk2\theta} \quad (30)$$

where  $N$  is the window  $w_e(k) = w(k)w(-k)$  length. Simplified optimization procedure for the Wigner distribution calculation, in the case of a high noise, has been derived in [23]. It results in the adaptive distribution with time-

frequency varying window length

$$WD_x^+(n, \theta) = \begin{cases} WD_x(n, \theta; N_1) & \text{for } \Phi = True \\ WD_x(n, \theta; N_2) & \text{otherwise} \end{cases} \quad (31)$$

where the event  $\Phi = True$  is

$$\Phi = True : |WD_x(n, \theta; N_1) - WD_x(n, \theta; N_2)| \leq (\kappa + \frac{1}{2})[\sigma_{xx}(N_1) + \sigma_{xx}(N_2)]$$

and  $N_1 \ll N_2$ . In order to implement (31) we have to calculate the Wigner distribution with a narrow and a wide lag windows  $N_1$  and  $N_2$ . By using only two distributions, we can expect a significant improvement in the time-frequency representation, since the distribution is usually either very highly concentrated or zero. A theoretical analysis, as in [23], with a large number of window lengths within interval  $(N_1, N_2)$  can prove that we may get the optimal window length within the accuracy of the window length discretization. But, we have concluded that the multi-window approach, although theoretically more accurate, in practice, for the kind of problems treated in this paper, does not produce significant improvement with respect to the very simple two-windows approach, used here.

The only parameter that is required in (31) is the Wigner distribution variance  $\sigma_{xx}(N)$ , [1], [19], [23]. There are several ways for its accurate estimation. For high noise cases,  $\sigma_\nu^2 > A^2$ , which are treated in this paper, the estimation is very simple since [1], [19], [23]

$$\begin{aligned} \sigma_{xx}^2(N)/E_w(N) &= \sigma_\nu^2(2A^2(n) + \sigma_\nu^2) \\ &\cong (\sigma_\nu^2 + A^2(n))^2 \cong \left(\frac{1}{N} \sum_{k=-N/2}^{N/2-1} |x(n+k)|^2\right)^2. \end{aligned}$$

Factor  $E_w(N) \sim N$  is a constant for the given window type. The variance  $\sigma_{xx}(N_1)$  can be calculated from the better estimated  $\sigma_{xx}(N_2)$  as  $\sigma_{xx}^2(N_1) = \sigma_{xx}^2(N_2)N_1/N_2$ . Some other approaches for the precise variance estimation, including small noise cases are presented in [11], [23].

### B. Example

Consider a sum of three noisy chirp signals

$$\begin{aligned} x(n) = & A(e^{-25(nT-0.25)^2} e^{j1200(nT)^2} \\ & + e^{-20(nT-0.65)^2} e^{j750(nT+0.75)^2} \\ & + 3.5e^{-22500(nT-0.96875)^2} e^{j1000nT}) + \nu(n). \end{aligned} \quad (32)$$

The sampling interval  $T = 1/2048$ , with  $\mathbf{N} = \{N_1, N_2\} = \{64, 512\}$  samples within the Hanning window is used. The signal amplitude and variance of noise are such that  $10 \log(E_f/\sigma_\nu^2) = -3[dB]$ . The signal may be treated as a monocomponent one since its energy at each time instant  $t$  is mainly concentrated around only one instantaneous frequency.

Step 1. The Wigner distributions with constant window widths and adaptive window width, obtained according to algorithm (31), are presented in Figures 1a),b), and c), respectively. Algorithm (31) has chosen the distribution with  $N_1 = 64$  and very low variance for all regions where the bias is small, including the third signal component. Distribution with  $N_2 = 512$  is chosen by the algorithm only for the small regions where the first two signal components exist, Figure 1. It is exactly what we wanted and expected. Note that the signals do not significantly overlap in time, so the lag-window was sufficient to reduce the "inner interferences" in the Wigner distribution. In order to combine two Wigner distributions, the distribution with  $N_1 = 64$  is interpolated (using zero padding prior the FFT) up to  $N_2 = 512$ .

Step 2. The distribution  $WD_x^+(n, \theta)$ , presented in Figure 1c), is used as an estimate of  $WD_s(n, \theta)$  and its region of support in the time-varying filter definition. According to (10) and the Appendix, the following form of  $L_H(n, \theta)$  is used:  $L_H(n, \theta) = 1$ , for a given time instant  $n$ , on  $\theta$  where the maximum of  $WD_x^+(n, \theta)$  is detected, and zero otherwise. Using this transfer function the signal  $x(n) = s(n) + \nu(n)$  is filtered and

$$y(n) = \frac{1}{2\pi} \int_{-\pi/2}^{\pi/2} L_H(n, \theta) STFT(n, \theta) d\theta \quad (33)$$

is produced. Original signal, the noisy signal and the noisy signal after time-varying filtering, by using  $H(n, \theta)$ , are shown in Figures 2a),b) and c), respectively. Efficiency of the time-varying filter is evident, especially if we have in mind that the signal occupies a wide frequency range and the time-invariant filtering would not produce a significant noise reduction. The signal to noise ratio improvement, that can be approximately expected in this case, according to (17),(21), is  $G = 10 \log(512) = 27[dB]$ . Output signal to noise ratio is then around  $SNR_{out} = 24[dB]$ . In practice this improvement is not so extremely high due to the discretization effects, but it is still evident, Figure 2c).

We may conclude that the presented algorithm may be efficiently used in the time-frequency representation and filtering of monocomponent noisy signals. The theory presented here is quite general and may easily be extended to the other time-frequency distributions. It will be done in the next Section.

## IV. TIME-VARYING FILTERING OF MULTICOMPONENT SIGNALS

The Wigner distribution, when applied on the multicomponent signals exhibits significant drawbacks manifesting themselves as the cross-terms. Consider first the random process  $\{s(t)\}$  whose large number of observations is known. In this case we may apply the time-varying filter relations on the random multicomponent signals, using the expected values of the Wigner distribution  $\overline{WD}_s(t, \omega)$ , called the Wigner spectrum. For the random signal  $s(t) = \sum_{i=1}^M s_i(t)$ , assuming that the components  $s_i(t)$  are not correlated, the cross-terms free Wigner spectrum

$$\overline{WD}_s(t, \omega) = \sum_{i=1}^M \overline{WD}_{s_i}(t, \omega)$$

follows, where

$$\overline{WD}_s(t, \omega) = \int_{-\infty}^{\infty} E \left\{ s\left(t + \frac{\tau}{2}\right) s^*\left(t - \frac{\tau}{2}\right) \right\} e^{-j\omega\tau} d\tau.$$

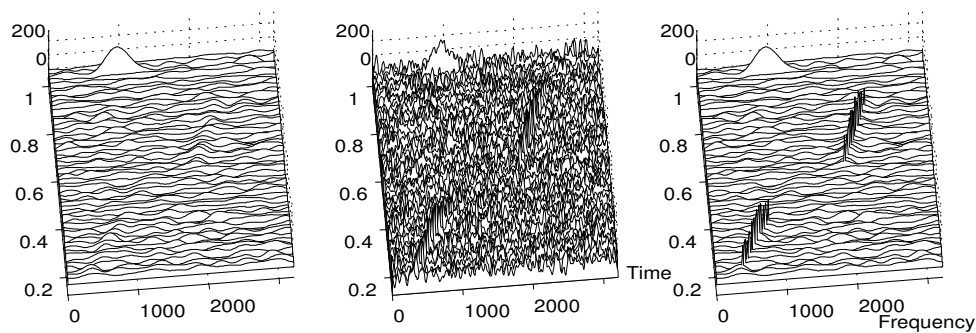


Fig. 1. Wigner distribution of the noisy signal calculated using: a) The constant lag window length  $N_1 = 64$ , b) The constant lag window length  $N_2 = 512$ , c) The adaptive time-frequency varying lag window length.

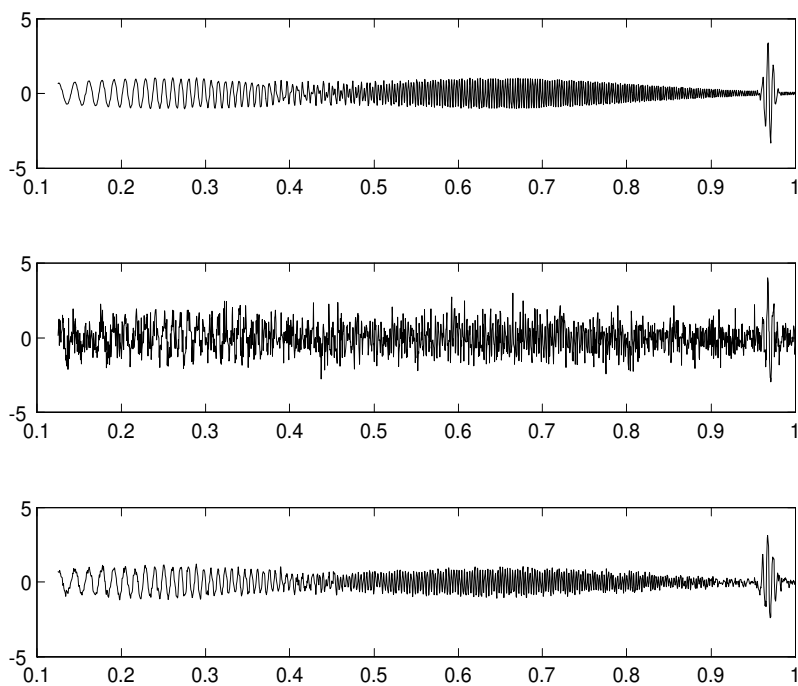


Fig. 2. Time-varying filtering: a) Original signal without noise, b) Noisy signal, c) Signal filtered by using the time-varying filter, with the region of support determination based on the distribution from Figure 1c.



The cross-terms do not exist since we assumed that the components are not correlated

$$E \left\{ s_i \left( t + \frac{\tau}{2} \right) s_j^* \left( t - \frac{\tau}{2} \right) \right\} = 0 \text{ for } i \neq j.$$

These relations mean that if  $R_i$  are the regions of support of  $\overline{WD}_{s_i}(t, \omega)$  then the region of support of  $\overline{WD}_s(t, \omega)$  is

$$R = R_1 \cup R_2 \cup \dots \cup R_M.$$

where  $R_i \cap R_j = \emptyset$ , for  $i \neq j$ .

With these definitions, the application of the time-varying filtering on the multicomponent signals, when a large number of realizations of the same random process  $\{s(t)\}$  is available, is clear and direct.

In practice the time-varying filtering should be performed on a single realization, in the case of multicomponent signals, as well. The classical Wigner distribution is not applicable because of the cross-terms. In the case when a single realization of the random process is considered, the best solution would be a distribution  $P(t, \omega)$  that would produce the sum of the Wigner distributions or the pseudo Wigner distributions of each individual signal component

$$P(t, \omega) = \sum_{i=1}^M PWD_{s_i}(t, \omega). \quad (34)$$

This distribution would have

$$L_H(t, \omega) = \begin{cases} 1 & \text{for } (t, \omega) \in R = \bigcup_{i=1}^M R_i \\ 0 & \text{for } (t, \omega) \notin R \end{cases} \quad (35)$$

As an illustration of the time varying filtering, consider the FM multicomponent signal

$$s(t) = \sum_{i=1}^M A_i(t) e^{j\phi_i(t)}.$$

with the slow-varying amplitudes of each component. Its short-time Fourier transform may be obtained using the stationary phase method as

$$STFT(t, \omega) = \sum_{i=1}^M \int_{-\infty}^{\infty} w(\tau_0) A_i(t + \tau_0) \times$$

$$e^{j\phi_i(t+\tau_0) - j\omega\tau_0} \sqrt{\frac{2\pi j}{\phi''_i(t+\tau_0)}} \times \delta(t + \tau_0 - (\phi'_i(\omega))^{-1}) d\tau_0. \quad (36)$$

Using the filtering function,

$$L_H(t, \omega) = \sum_{i=1}^M \delta(\omega - \phi'_i(t)) \quad (37)$$

or its form given by (19), as it will be used here, and applying  $STFT(t, \omega)$  and  $L_H(t, \omega)$  in (24) we get

$$\begin{aligned} (Hx)(t) &= \frac{1}{2\pi} \sum_{i=1}^M \sum_{k=1}^M \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(\omega - \phi'_k(t)) \times \\ &w(\tau_0) A_i(t + \tau_0) e^{j\phi(t+\tau_0) - j\omega\tau_0} \times \\ &\delta(t + \tau_0 - (\phi'_i(\omega))^{-1}) d\tau_0 d\omega + \nu_{out}(t) \\ &= \sum_{i=1}^M A_i(t) e^{j\phi_i(t)} + \nu_{out}(t). \end{aligned}$$

It has been assumed that the signal components do not simultaneously overlap in time and frequency, what can be understood as: For  $t$  and  $\tau_0$  such that  $\phi'_k(t) = \phi'_i(t + \tau_0)$ , then  $w(\tau_0) A_i(t + \tau_0) = 0$  for all  $\tau_0$  within the lag window  $w(\tau_0)$ .

A distribution that has property (34) is the S-method [20], [21]. It is defined by

$$\begin{aligned} SM(t, \omega) &= \\ &= \frac{1}{\pi} \int_{-\infty}^{\infty} P(\theta) STFT(t, \omega + \theta) STFT^*(t, \omega - \theta) d\theta, \end{aligned} \quad (38)$$

where  $STFT(t, \omega)$  is defined by (25), and  $P(\theta)$  is a finite frequency domain window (we also assume rectangular),  $P(\theta) = 0$ , for  $|\theta| > W_P$ . Numerical realization of the S-method is very simple, according to its discrete form

$$\begin{aligned} SM(n, k) &= SPEC(n, k) + \\ &+ 2 \sum_{i=1}^{L_P} \text{Re}[STFT(n, k + i) STFT^*(n, k - i)], \end{aligned} \quad (39)$$

where  $SPEC(n, k) = |STFT(n, k)|^2$  is the spectrogram and the terms  $\text{Re}[STFT(n, k + i)STFT^*(n, k - i)]$  improve its concentration to the Wigner distribution quality.

In the case when the instantaneous frequencies of the signal components may not be considered as linear functions within the considered lag window, then the L-Wigner distributions or S-distributions realized using the S-method, producing (37) as described in [20], [21], [22], should be used for the regions of support estimation. An alternative approach for these cases is described in [2].

#### A. Example

The S-method and algorithm (31) are applied in order to produce a time-frequency representation, and then the time-varying filtering, of the signal

$$\begin{aligned} x(n) = & A e^{-4(nT-9/16)^2 + j480(nT+1/5)^2} \\ & + A e^{-100(nT-3/8)^2 + j525(nT+3/2)^2} \\ & + A e^{-100(nT-13/16)^2 + j740(nT)^2} + \nu(n) \end{aligned} \quad (40)$$

where the value of  $A$  is such that  $10 \log(E_f/\sigma_v^2) = -1[\text{dB}]$ .

*Step 1:* The same parameters as in the first example, and the algorithm described by (31) in the form

$$SM^+(n, k) = \begin{cases} SM(n, k; N_1, L_P = 0) \\ \quad \text{for } \Phi = \text{True} \\ SM(n, k; N_2, L_P = 2) \\ \quad \text{otherwise} \end{cases} \quad (41)$$

are used. The lag window lengths are  $\{N_1, N_2\} = \{32, 256\}$  with the rectangular window  $P(i)$  that includes only two samples around the central frequency  $\theta = \pi k/N$  on each side ( $L_P = 2$ ). The distribution obtained in this way is presented in Figure 3d. The Wigner distribution of the original *nonnoisy signal* with the constant lag window length  $N = 512$  is given in Figure 3a, while Figure 3b shows the Wigner distribution of noisy signal with the same constant window length. Figure 3c presents the S-method with the constant parameters  $N_2 = 256$ ,  $L_P = 2$ . Note that the Wigner distribution has to be oversampled

with a factor of 2, while the S-method does not require this oversampling. It results in a twice smaller number of samples within the same lag window, i.e., within the basic frequency period in the S-method. This is demonstrated on the frequency axis labels in Figure 3a)-d). Variance for (31), i.e., (41), is estimated by using the approach described in [23].

*Step 2:* Distribution from Figure 3d is used for the regions of support  $L_H(n, k)$  determination. The simplest way to get this regions is by comparing the value of  $SM^+(n, k)$  with a reference level defined, for example, as a fraction of the maximal value of  $SM^+(n, k)$ . Here we used  $M = 0.1 \max_{n, k} \{SM^+(n, k)\}$ , i.e., 1/10th of the maximal distribution value. Then

$$L_H(n, k) = u(SM^+(n, k) - M), \quad (42)$$

where  $u(x)$  is the unit step function. Now we will describe one interesting effect due to the frequency discretization, that may influence the quality of filtering. If the distribution is highly concentrated along the instantaneous frequency and we have an extremely dense grid, then for each component and each time instant  $n$  only one value of  $L_H(n, k)$  will, for that component, assume value 1. However, the true instantaneous frequency value may be, and generally always is, between the grid points. Then it may happen that  $L_H(n, k)$  in two or more points assumes value 1 for one signal component and one time instant. From the numerous experiments we concluded that in those cases it is much better to use only one unity value, where a larger value of  $SM^+(n, k)$  is detected, and then to force other points within that component to take zero value, for that time instant  $n$ . Practically, it may be implemented in a simple way. Detect maximum of  $SM^+(n, k)$  for given  $n$ . If it satisfies the condition that (42) assumes unity value, then assign value of 1 to  $L_H(n, k)$  in that point  $(n, k)$ . Exclude several neighboring points as possible next maximums, find the second maximum check it for (42), and so on until the number of maximums equal to the expected number of components is determined. The region  $L_H(n, k)$  found in this way is shown in Figure 3e.

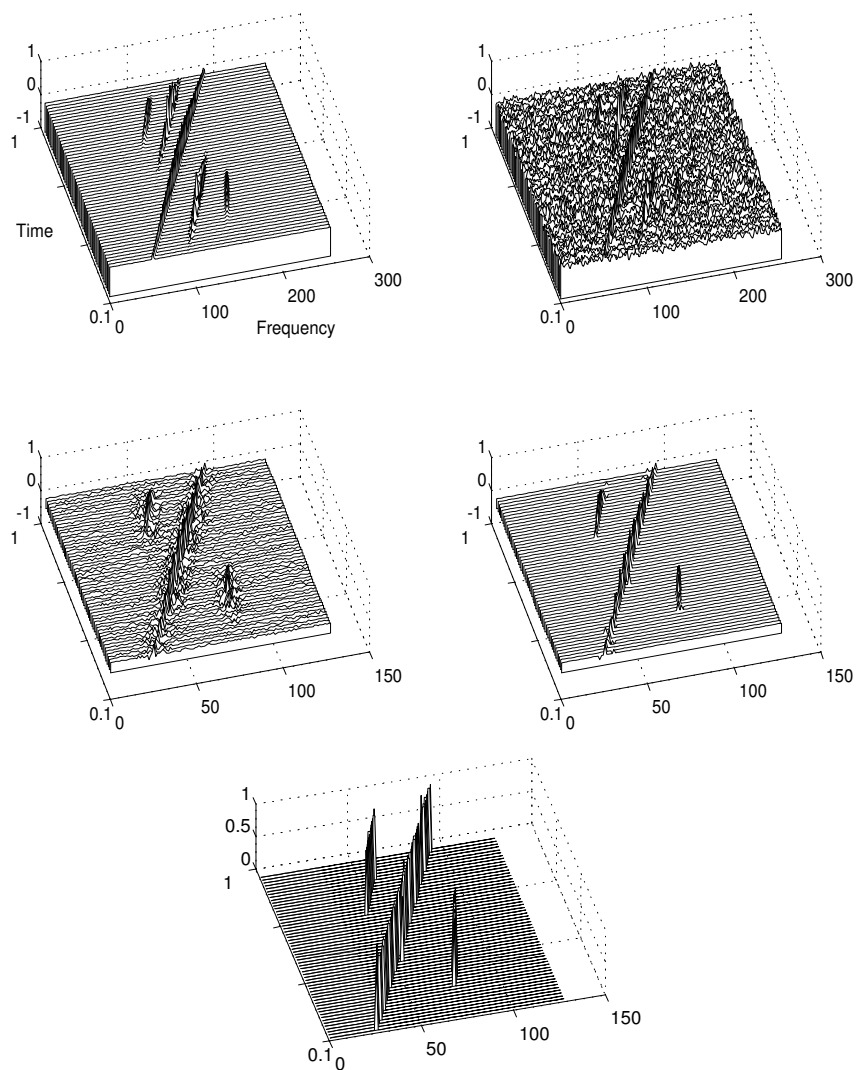


Fig. 3. Time-frequency representation of a multicomponent signal: a) Pseudo Wigner distribution of the original signal without noise, b) Pseudo Wigner distribution of the noisy signal using the constant lag window length  $N_2 = 512$ , c) S-method of the noisy signal with the constant lag window length  $N_2 = 256$ , d) S-method of the noisy signal with the adaptive time-frequency varying window length  $N \in \{32, 256\}$ , e) Time-varying filter region of support. The number of samples is shown on the frequency axis. The Wigner distributions have to be oversampled by factor of 2. Normalized values of distributions are shown.

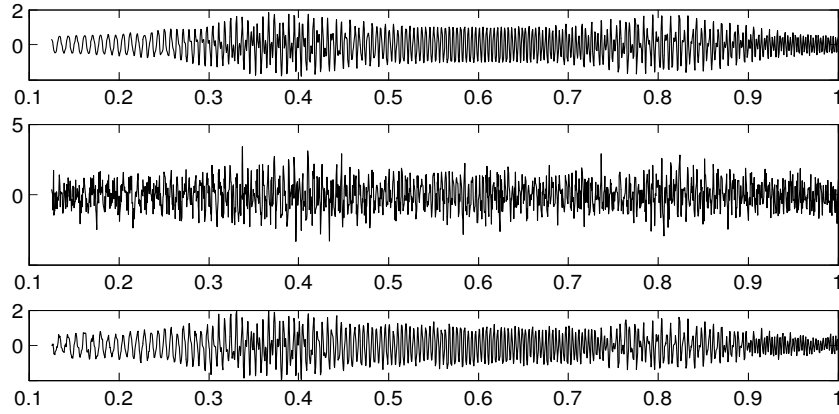


Fig. 4. Time-varying filtering of the multicomponent signal: a) Original signal without noise, b) Noisy signal, c) Signal filtered by using the time-varying filter, with the region of support presented in Figure 3e.

After we have obtained  $L_H(n, k)$  we may directly apply time-varying filtering relation (33). The filtered signal, along with the original one, and the nonfiltered noisy one, is shown in Figure 4. The improvement is evident. Theoretical analysis suggests that, since this signal may be considered as a two-component one (at each time instant there is no more than two instantaneous frequencies where condition for unity value of (42) is satisfied), we may expect the improvement up to  $G = 10 \log(256/2) = 21[\text{db}]$ , i.e., the output signal to noise ratio  $SNR_{out} \approx 20[\text{dB}]$ . Due to the described discretization effects we can not get so high improvement, but it is still rather very high, Figure 4.

## V. CONCLUSION

Time-varying filtering of noisy frequency modulated signals is considered. After a slight modification of the Wigner framework definition, algorithms and methods for the filter's support in the case on noisy multicomponent and monocomponent signals, are presented. They are based on a deterministic or a single realization of random signal. Efficiency of the presented theory is demonstrated on numerical examples.

## VI. APPENDIX: Optimal Filtering

Time-varying filtering is one of the challenging areas where one may benefit from the joint

time-frequency representations. It has been defined by (1), where  $h(t, \tau)$  is the impulse response of the time-varying system  $H$  whose optimal form may be determined by minimizing the mean squared error

$$H_{opt} = \arg \min_H E \left\{ |s(t) - H\{x(t)\}|^2 \right\}. \quad (43)$$

Here we will present a solution based on the modification (11) of the time-varying filtering

$$y(t) = (Hx)(t) = \int_{-\infty}^{\infty} h\left(t + \frac{\tau}{2}, t - \frac{\tau}{2}\right) x(t + \tau) d\tau. \quad (44)$$

The optimal value of  $H$  will be derived by analogy with the Wiener filter in the stationary signal cases. Suppose that the nonstationary stochastic process  $x(t) = s(t) + \nu(t)$ , beside the desired component  $s(t)$  contains noise  $\nu(t)$ . When the mean square error  $e^2 = E\{|s(t) - y(t)|^2\}$  reaches its minimum, the error  $s(t) - y(t)$  is orthogonal to the data  $x^*(t + \alpha)$ , for any  $\alpha$ . From this fact we get

$$E\left\{ \left[ s(t) - \int_{-\infty}^{\infty} h\left(t + \frac{\tau}{2}, t - \frac{\tau}{2}\right) x(t + \tau) d\tau \right] x^*(t + \alpha) \right\} = 0 \quad (45)$$

The expected value of the ambiguity function is defined by

$$\overline{AF}_{xx}(\theta, \tau) = \int_{-\infty}^{\infty} E\{x(t + \frac{\tau}{2})x^*(t - \frac{\tau}{2})\}e^{-j\theta t} dt$$

Taking the Fourier transform over  $t$  of (45), by using appropriate substitutions

$$\int_{-\infty}^{\infty} E\{s(t)x^*(t + \alpha)\}e^{-j\theta t} dt = \overline{AF}_{sx}(\theta, -\alpha)e^{j\alpha\theta/2}$$

and

$$\int_{-\infty}^{\infty} E\{x(t + \tau)x^*(t + \alpha)\}e^{-j\theta t} dt = \overline{AF}_{xx}(\theta, \tau - \alpha)e^{j(\alpha + \tau)\theta/2}$$

we get

$$\overline{AF}_{sx}(\theta, \alpha)e^{-j\alpha\theta/2} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A_H(u, -\tau) \times \overline{AF}_{xx}(\theta - u, \alpha - \tau)e^{-j(\alpha + \tau)(\theta - u)/2} d\tau du \quad (46)$$

where

$$A_H(\theta, \tau) = \int_{-\infty}^{\infty} h(t + \frac{\tau}{2}, t - \frac{\tau}{2})e^{-j\theta t} dt. \quad (47)$$

For

$$|\theta\tau - u\alpha - u\tau|/2 \ll \pi$$

when  $e^{j(\theta\tau - u\alpha - u\tau)/2} \cong 1$ , from (46) follows

$$\overline{AF}_{sx}(\theta, \alpha) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A_H(u, -\tau) \times \overline{AF}_{xx}(\theta - u, \alpha - \tau) d\tau du. \quad (48)$$

Taking two-dimensional Fourier transform of (48) we get

$$\overline{WD}_{sx}(t, \omega) = L_H(t, \omega)\overline{WD}_{xx}(t, \omega)$$

where

$$\begin{aligned} \overline{WD}_{xx}(t, \omega) &= FT_{2d}\{\overline{AF}_{xx}(\theta, \alpha)\} \\ &= \int_{-\infty}^{\infty} E\{x(t + \frac{\tau}{2})x^*(t - \frac{\tau}{2})\}e^{-j\omega\tau} d\tau \end{aligned}$$

is the Wigner spectrum. The Weyl symbol of the filter impulse response is denoted by

$$L_H(t, \omega) = \int_{-\infty}^{\infty} h(t - \frac{\tau}{2}, t + \frac{\tau}{2})e^{-j\omega\tau} d\tau. \quad (49)$$

Therefore, the optimal filter in time-frequency domain is defined by

$$L_H(t, \omega) = \frac{\overline{WD}_{sx}(t, \omega)}{\overline{WD}_{xx}(t, \omega)} \quad (50)$$

what corresponds to the well known Wiener filter in the stationary cases  $H(\omega) = S_{sx}(\omega)/S_{xx}(\omega)$ ; here  $S_{sx}(\omega) = FT\{r_{sx}(\tau)\}$ , since in the stationary cases  $r_{sx}(\alpha) = E\{s(t)x^*(t + \alpha)\}$ ,  $r_{xx}(\tau - \alpha) = E\{x(t + \tau)x^*(t + \alpha)\}$ .

Relation (50) would be obtained if we assumed at the beginning that the random signals are quasistationary when  $r_{sx}(t, t + \alpha)$  could be written as  $r_{sx}(t + \alpha/2, t - \alpha/2)$ , and  $r_{xx}(t + \tau, t + \alpha)$  as  $r_{xx}(t + (\tau - \alpha)/2, t - (\tau - \alpha)/2)$ .

For the signal not correlated with noise follows

$$L_H(t, \omega) = \frac{\overline{WD}_{ss}(t, \omega)}{\overline{WD}_{ss}(t, \omega) + \overline{WD}_{\nu\nu}(t, \omega)} \quad (51)$$

Suppose that the Wigner spectrum of the random signal  $s(t)$  lies inside a region  $R$  in the time-frequency plane, while the noise lies outside this area, except may be its small part that can be neglected with respect to the part of noise outside  $R$ . This is true, for example, for a wide class of frequency modulated (highly concentrated in the time-frequency plane) signals  $s(t)$ , corrupted with a white noise  $\nu(t)$ , widely spread in the time-frequency plane. A simple solution satisfying these requirements is given by

$$L_H(t, \omega) = \begin{cases} 1 & \text{for } (t, \omega) \in R \\ 0 & \text{for } (t, \omega) \notin R \end{cases} \quad (52)$$

Applying this solution with (24), for example, in frequency domain

$$y(t) = (Hx)(t) = \int_{-\infty}^{\infty} L_H(t, \omega)STFT(t, \omega) d\omega$$

we will be able to efficiently filter frequency modulated monocomponent and multicomponent signals corrupted with the white noise.

#### REFERENCES

- [1] M. G. Amin: "Minimum variance time-frequency distribution kernels for signals in additive noise", *IEEE Trans. Signal Processing*, vol. 44, no. 9, pp. 2352-2356, Sep. 1996.
- [2] F. Auger, P. Flandrin: "Improving the readability of time-frequency and time-scale representations by the reassignment method", *IEEE Trans. on Signal Processing*, vol-43, no.5, May 1995, pp.1068-1089.
- [3] B. Boashash: "Estimating and interpreting the instantaneous frequency of a signal-Part 1: Fundamentals", *Proc. of the IEEE*, vol-80, no.4, April 1992, pp.519-538.
- [4] G.F. Boudreaux-Bartels, T.W.Parks: "Time-varying filtering and signal estimation using the Wigner distribution", *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 34, no. 6, pp. 442-451, June 1986.
- [5] T.A.C.M Claasen, W.F.G. Mecklenbauer:"The Wigner distribution: A tool for time-frequency analysis, Parts I,II,II", *Philips J. Res.*, vol.35, no.3, 4/5,6, 1980, pp.217-250, 276-300, 372-389.
- [6] L. Cohen: "Time-frequency analysis", Prentice-Hall, 1995.
- [7] P. Flandrin, W.Martin: "The Wigner-Ville spectrum of nonstationary random signals", in *The Wigner distribution: Theory and applications in signal processing*, Eds. W.Mecklenbrauker, F.Hlawatsch, pp.211-267.
- [8] F. Hlawatsch, G. F. Broudreaux-Bartels: "Linear and quadratic time-frequency signal representation", *IEEE Signal Processing Mag.*, pp.21-67, April 1992.
- [9] A.J.E.M. Janssen: "Wigner weight functions and Weyl symbols of nonnegative definite linear operators", *Philips J. Res.*, vol.44, pp.7-42, 1989.
- [10] J.Jeong,W.J.Williams:"Kernel design for reduced interference distributions", *IEEE Trans.Signal Processing*, vol.40, pp.402-414, Feb.1992.
- [11] V. Katkovnik, L.J. Stanković: "Instantaneous frequency estimation using the Wigner distribution with varying and data-driven window length", *IEEE Trans. on Signal Processing*, vol.46, no.9, Sept.1998, pp.2315-2325.
- [12] V. Katkovnik: "Nonparametric estimation of the instantaneous frequency", *IEEE Trans. on Information Theory*, vol.43, no.1, pp.183-189, Jan1997.
- [13] W.Kozek: "Time-frequency signal processing based on the Wigner-Weyl framework", *Signal Processing*, vol.29, no.1, Oct.1992, pp.77-92.
- [14] G. Matz, F.Hlawatsch, W. Kozek: "Generalized evolutionary spectral analysis and the Weyl spectrum of nonstationary random processes", *IEEE Trans. on Signal Processing*, vol.45, no.6, June 1997, pp.1520-1534.
- [15] A. Papoulis: "Signal Analysis", *McGraw-Hill Book Company*, 1977.
- [16] A.W. Rihaczek: "Signal energy distribution in time and frequency", *IEEE Trans. Information Theory*, vol.14, 1968, pp.369-374.
- [17] B. Ristic, B. Boashash:"Relationship between the polynomial and the higher order Wigner-Ville distribution", *IEEE Signal Processing Letters*, vol-2, no.12, Dec.1995, pp.227-229.
- [18] R.G. Shenoy, T.W. Parks:" The Weyl correspondence and time-frequency analysis", *IEEE Trans. on Signal Processing*, vol.42, no.2, Feb.1994, pp.318-331.
- [19] L.J. Stanković, S.Stanković: "On the Wigner distribution of discrete-time noisy signals with application to the study of quantization effects", *IEEE Trans. on Signal Processing*, vol-42, no.7, July 1994, pp.1863-1867.
- [20] L.J. Stanković: "A method for time-frequency analysis", *IEEE Trans. on Signal Processing*, vol-42, no.1, Jan.1994, pp.225-229.
- [21] L.J. Stanković: "A multitime definition of the Wigner higher order distribution: L-Wigner distribution" *IEEE Signal Processing Letters*, vol.1, no.7, July 1994, pp. 106-109.
- [22] L.J. Stanković: "Highly concentrated time-frequency distribution: Pseudo quantum signal representation", *IEEE Trans. on Signal Processing*, vol-45, no.3, March 1997, pp.534-552.
- [23] L.J. Stanković, V. Katkovnik: "The Wigner distribution of noisy signals with adaptive time-frequency varying window" *IEEE Trans. on Signal Processing*, vol-47, no.4, April 1999, pp.1099-1108.
- [24] H.L. Van Trees, *Detection, Estimation, and Modulation Theory*, New York: Wiley, 1968.
- [25] L.A. Zadeh: "Frequency analysis of variable networks", *Proc. IRE*, vol.67, March 1950, pp.291-299.
- [26] Y.M. Zhu, F. Peyrin, R. Goutte: "Transformation de Wigner-Ville: description d' un nouvel outil de traitement du signal et des images", *Annales des telecom.* vol-42, no 3-4, 1987, pp. 105-117.