Instantaneous Frequency Estimation Using Robust Spectrogram with Varying Window Length

Vladimir Katkovnik, Igor Djurović, LJubiša Stanković

Abstract— **Robust M-periodogram is defined for the analysis of signals with heavy-tailed distribution noise. In the form of a robust spectrogram it can be used for the analysis of nonstationary signals. In this paper a robust spectrogram based instantaneous frequency (IF) estimator, with a time-varying window length, is presented. The optimal choice of the window length, based on asymptotic formulae for the variance and bias, can resolve the bias-variance trade-off in the robust spectrogram based IF estimation. However, it depends on the unknown nonlinearity of the IF. The algorithm used in this paper is able to provide the accuracy close to the one that could be achieved if the IF, to be estimated, were known in advance. Simulations show good accuracy ability of the adaptive algorithm and good robustness properties with respect to rare high magnitude noise values.**

I. Introduction

A key-model of the instantaneous frequency (IF) concept is the complex-valued harmonic with a time-varying phase. It is an important model in the general theory of time-frequency distributions. This model has been utilized to study a wide range of signals, including speech, music, acoustic, biological, radar, sonar, and geophysical ones [20]. An overview of the methods for the IF estimation, as well as the interpretation of the IF concept itself, is presented in [2]. One possible approach to the IF estimation is based on time-frequency representations [3], [4], [5], [23], [24]. The spectrogram is a commonly applied distribution within this approach.

In this paper we combine and develop two different ideas: the robust $M-$periodogram and the nonparametric approach [9]-[13] for selection of the time-varying adaptive sliding window length in the corresponding periodogram. The robust $M-$periodogram is developed as a generalization of the standard periodogram for analysis of stationary signals corrupted with heavy tailed distribution noise [14], [16]. Its form applied to the analysis of nonstationary signals will be referred to as the robust spectrogram. Recall that the heavy tailed distribution noise is used as a model of an impulse noise environment (e.g. [18]). The approach which exploits the intersection of confidence interval rule [6] was used in [12] and [17] for the standard periodogram based estimator with varying adaptive window length. It uses only the formula for the variance of the estimate, which does not require information about the IF to be known in advance. Simulations based on the discrete robust spectrogram, with several noisy signal examples, show a good robustness and accuracy ability of the presented adaptive algorithm, as well as an improvement in the spectrogram based time-frequency representation of signals with the nonlinear IF. Finally, the Huber's minimax loss functions are used for design of robust spectrograms with the adaptive window size.

Intern. J. Elect. Comm., AEU, Vol.54, No.4, 2000.

The structure of the paper is as follows. The robust spectrogram as an IF estimator is considered in Section II. The asymptotic bias and variance of the IF estimate, along with the optimal window length, are presented in this section, as well. The adaptive estimate of the IF with a time-varying and data-driven window length is developed in Section III. Numerical examples, along with simulation results, are discussed in Section IV. A generalization of the robust spectrogram is proposed in Section V.

II. Background Theory

A. Robust Spectrogram

Standard spectrogram $I_S(t,\omega)$ definition, of a signal $x(t)$, is based on the standard short-time Fourier transform $\hat{C}_h(t,\omega)$

$$\hat{C}_h(t,\omega) = \frac{1}{\sum_n w_h(nT)} \times$$

$$\sum_n w_h(nT) x(t+nT) \exp(-j\omega nT), \qquad (1)$$

$$I_S(t,\omega) = \left|\hat{C}_h(t,\omega)\right|^2.$$

where the window $w_h(nT) = \frac{T}{h} w(nT/h) \geq 0$ has $h > 0$ as a window length, and $\sum_n w_h(nT) \to 1$ as $h/T \to \infty$. Sampling interval is denoted by T.

The standard short-time Fourier transform $\hat{C}_h(t,\omega)$ may be derived as a solution of the following optimization problem [14]:

$$\hat{C}_h(t,\omega) = \arg\min_C J(\omega, C), \qquad (2)$$

where

$$J(\omega, C) = \sum_n w_h(nT) \times$$

$$|x(t+nT) - C_h(t,\omega)\exp(j\omega nT)|^2. \qquad (3)$$

Here, the weighed square absolute error

$$F(e) = |e(nT)|^2 =$$

$$= |x(t+nT) - C_h(t,\omega)\exp(j\omega nT)|^2 \qquad (4)$$

is used as a loss function and minimized, by determining C. From

$$\frac{\partial J(\omega, C)}{\partial C^*} = 0$$

definition (1) follows.

In [14] it has been shown that the loss functions of other forms than $F(e) = |e|^2$ can be more efficient in the optimization procedure (2). In particular, it has been shown that the loss function of the form $F(e) = |e|$ can produce very good results in the case of a signal corrupted with heavy tailed distribution noise. The periodogram obtained using this loss function is called the robust $M-$periodogram. Its corresponding robust spectrogram is given in the form

$$I_A(t,\omega) = \left|\hat{C}_h(t,\omega)\right|, \qquad (5)$$

$$\hat{C}_h(t,\omega) = \arg\min_C J(\omega, C),$$

$$J(\omega, C) = \sum_n w_h(nT) \times$$

$$|x(t+nT) - C_h(t,\omega)\exp(j\omega nT)|.$$

By minimizing $J(\omega, C)$ we get a solution in the form

$$\hat{C}_h(t,\omega) = \frac{1}{\sum_n d(nT)} \times$$

$$\sum_n d(nT) x(t+nT) \exp(-j\omega nT),$$

$$d(nT) = \gamma(nT) / \sum_n \gamma(nT)$$

$$\gamma(n) = w_h(nT) \times$$

$$|x(t+nT) - C_h(t,\omega)\exp(j\omega nT)|^{-1}.$$

These three equations represent a set of nonlinear equations with unknown $C_h(t,\omega)$. It can be solved using the following iterative procedure [14]:

Step 0. Initialization (standard short time Fourier transform calculation):

$$C_h^{(0)}(t,\omega) = \frac{1}{\sum_n w_h(nT)} \times$$

$$\sum_n w_h(nT) x(t+nT) \exp(-j\omega nT), \qquad (6)$$

$$\gamma^{(0)}(n) = w_h(nT) \times$$

$$\left|x(t+nT) - C_h^{(0)}(t,\omega)\exp(j\omega nT)\right|^{-1}$$

(i) Step k, $k = 1, 2, .., K$:

$$C_h^{(k)}(t,\omega) = \frac{1}{\sum_n \gamma(n)^{(k-1)}} \times$$

$$\sum_n \gamma^{(k-1)}(n) x(t+nT) \exp(-j\omega nT),$$

$$\gamma^{(k)}(n) = w_h(nT) \times$$

$$\left|x(t+nT) - C_h^{(k)}(t,\omega)\exp(j\omega nT)\right|^{-1}$$

with the stopping rule

$$\hat{k} = \min_k \{k : \frac{|C_h^{(k)}(t,\omega) - C_h^{(k-1)}(t,\omega)|}{|C_h^{(k-1)}(t,\omega)|} \leq \eta,$$

$$k \leq K\},$$

where $\eta > 0$ and K are given.

(ii). Setting the robust spectrogram $I_A(t,\omega)$ as

$$I_A(t,\omega) = |C_h(t,\omega)|^2,$$

where

$$C_h(t,\omega) = C_h^{(\hat{k})}(t,\omega).$$

Experiments have showed a good convergence of the algorithm. Provided $\eta = 0.1$ a usual number of iteration was about $\hat{k} = 3 \div 5$ and never exceeded 15.

B. Instantaneous Frequency Estimation

Consider now the problem of IF estimation, using the robust spectrogram, from the discrete-time observations

$$x(nT) = m(nT) + \varepsilon(nT),$$

$$\text{with} \quad m(t) = Ae^{j\varphi(t)} \tag{7}$$

where n is an integer, T is a sampling interval and $\varepsilon(nT)$ is a complex-valued circular white noise $E(\varepsilon(nT)) = 0$, $E(|\varepsilon(nT)|^2) = \sigma^2$. By definition, the IF is the first derivative of the phase

$$\Omega(t) = \varphi'(t). \tag{8}$$

Its estimate can be found as

$$\hat{\omega}_h(t) = \arg \max_{\omega \in Q_\omega} I_A(t,\omega) \tag{9}$$

Let us remind that the window $w_h(nT)$ implements the idea of nonparametric estimation of the IF as the time-varying $\Omega(t)$ is fitted by the constant ω within the narrow window around the time-instant t [9], [10], [15].

The asymptotic accuracy analysis of the robust IF estimator (9) has been done in [16]. According to that analysis, with the constraints as stated, the asymptotic formulae for the variance and bias of the IF estimation error

$$\Delta\hat{\omega}_h(t) = \Omega(t) - \hat{\omega}_h(t),$$

are given by

$$var(\Delta\hat{\omega}_h(t)) =$$

$$V(F,G) \cdot \frac{T}{A^2 h^3} W_\omega + o(T/h^3), \tag{10}$$

and

$$E(\Delta\hat{\omega}_h(t)) = B_\omega h^2 \Omega^{(2)}(t) + o(h^2), \tag{11}$$

where $o(x)$ denote a small value such that $o(x)/x \to 0$ as $x \to 0$.

The following notation has been used

$$W_\omega = \frac{\int_{-\infty}^{\infty} w^2(u)u^2 du}{(\int_{-\infty}^{\infty} w(u)u^2 du)^2}, \tag{12}$$

$$B_\omega = \frac{1}{3! \int_{-\infty}^{\infty} w(u)u^2 du} \int_{-\infty}^{\infty} w(u)u^4 du, \tag{13}$$

$$V(F,G) =$$

$$\int (F^{(1)}(v))^2 dG(v) / (\int F^{(2)}(v) dG(v))^2. \tag{14}$$

with $T \to 0$, $h \to 0$, $T/h^4 \to 0$, $\Omega^{(1)}(t) \neq 0$, $\Omega^{(2)}(t) \neq 0$, G is the noise $\varepsilon(nT)$ probability distribution function, and $F^{(1)}$ and $F^{(2)}$ are the derivatives of F.

Comments:

1. Let the noise distribution be Gaussian, $\varepsilon \sim N(0, \sigma^2/2)$, and the loss function F is quadratic $F(e) = e^2$, then $V(F,G) = \sigma^2/2$. Substituting this $V(F,G) = \sigma^2/2$ in (10) gives the known formulae for the variance of the short-time periodogram IF estimates. In particular these formulae can be obtained as a special case from more general results produced in [10] and [13].

In a similar way we obtain for $F(e) = |e|$ that $V(F,G) = \pi\sigma^2/2$.

2. We wish to note that $V(F,G)$ appears as a factor only in the formula for the variance. Thus a choice of the loss function F influences only the variance of estimation but

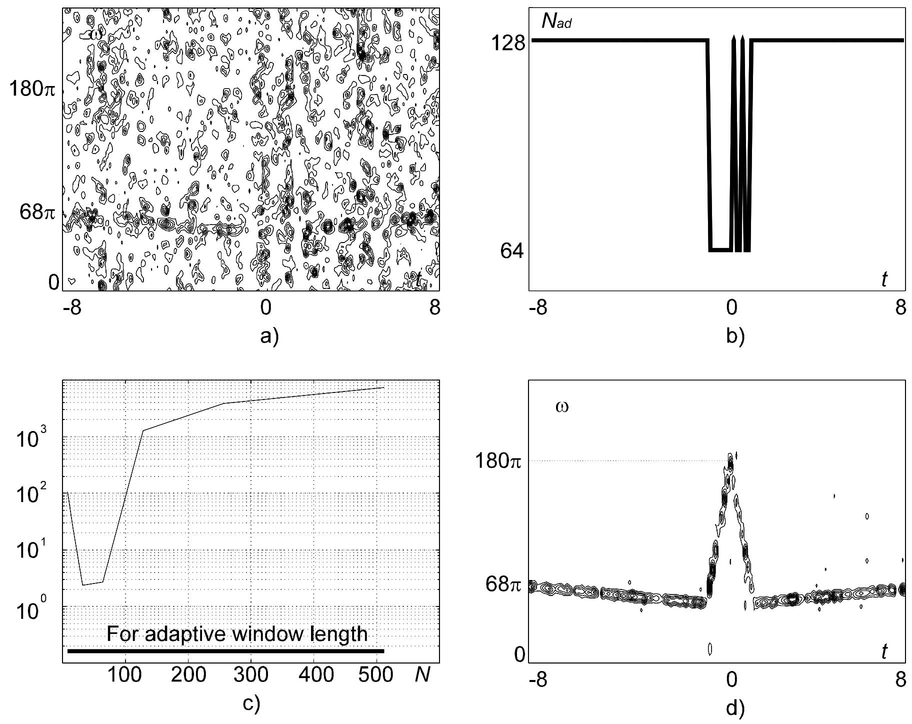

Fig. 1. a) Standard spectrogram of the signal, with $N = 128$, b) Adaptive window length for the robust spectrogram, c) Mean square error versus window lenght, d) Robust spectrogram with adaptive window length.

not the bias. The similar result has a place for the LPA robust regression estimation [25].

The formulae for the bias are the same for the robust and nonrobust estimates [17].

3. Let us consider the mean squared error (MSE) of the estimation. From (10) and (11) follows that for small h the main terms of the MSE can be given in the form

$$E((\Delta\hat{\omega}_h(t)^2)) = V(F,G)\cdot\frac{T}{A^2h^3}W_\omega + \left(B_\omega h^2\Omega^{(2)}(t)\right)^2. \quad (15)$$

Decreasing of the window length h results in decreasing of the bias and in increasing of the variance, and vice versa. The optimal window width is given as

$$h_{opt}(t) = \left(\frac{3V(F,G)\cdot TW_\omega}{4A^2\left(B_\omega\Omega^{(2)}(t)\right)^2}\right)^{1/7}.$$

It gives an optimal bias-variance trade-off, usual for nonparametric estimations. Optimal length depends on the signal-to-noise ratio A/σ_ε, the sampling interval T, noise distribution G, selected loss function F, and the second IF derivative $\Omega^{(2)}(t)$. Thus the optimal, or even reasonable choice of length h, depends on the IF second derivative $\Omega^{(2)}(t)$, which is naturally unknown because the IF itself is to be estimated.

III. Algorithm of Data-Driven Window Length Choice

A. Basic Idea ([12],[17],[22])

The basic idea follows from the IF estimation error analysis. Namely, at least for the asymptotic case, the estimation error can be represented as a sum of the deterministic component (bias) and random component, with the variance given by (10). The estimation error can be written as

$$|\Omega(t) - \hat{\omega}_h(t)| \leq |bias(t,h)| + \kappa\sigma(h), \quad (16)$$

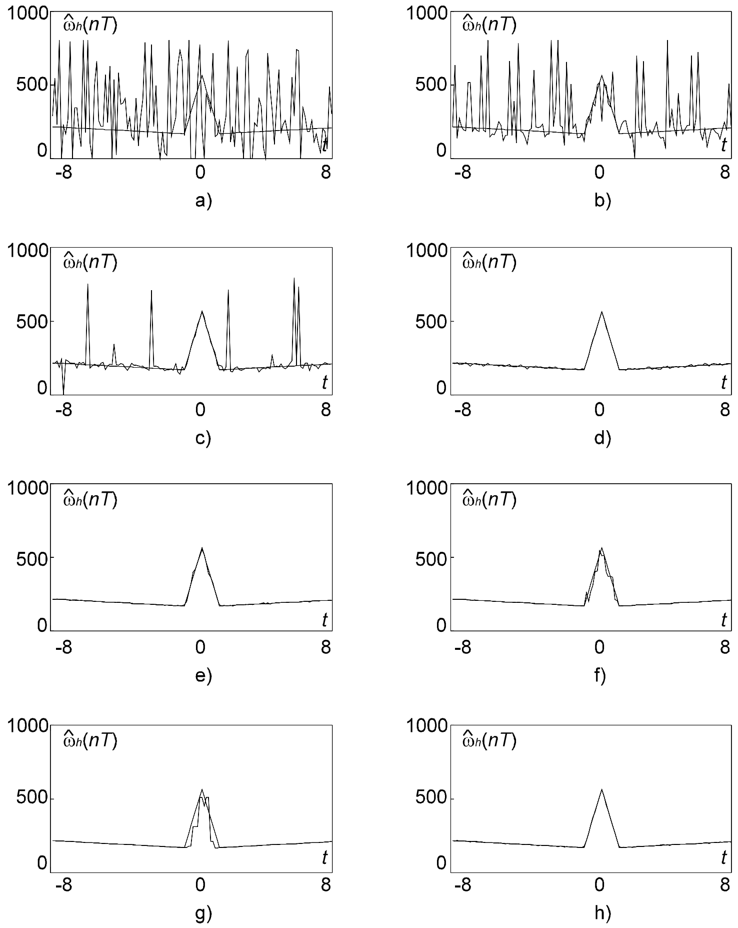

Fig. 2. a)-g) IF estimation of the signal using robust spectrogram with the constant window lengths $N = \{8, 16, 32, 64, 128, 256, 512\}$, respectively; h) IF estimation using robust spectrogram with the adaptive window length.

with $\sigma^2(h) = var(\Delta\hat{\omega}_h(t))$. Inequality (16) holds with probability $P(\kappa)$, where κ is the corresponding quantile of the standard Gaussian distribution $N(0, 1)$. The usual choice $\kappa = 2$ gives $P(\kappa) = 0.95$.

It follows from (11) that $|bias(t, h)| \to 0$ as $h \to 0$. Now, let $h = h_s$ be so small that

$$|bias(t, h_s)| \leq \kappa\sigma(h_s), \tag{17}$$

then

$$|\Omega(t) - \hat{\omega}_{h_s}(t)| \leq 2\kappa\sigma(h_s). \tag{18}$$

It is obvious that, for a set of such small h_s, all of the segments

$$D_s = [\hat{\omega}_{h_s}(t) - 2\kappa\sigma(h_s), \\ \hat{\omega}_{h_s}(t) + 2\kappa\sigma(h_s)], \tag{19}$$

have a point in common, namely $\Omega(t)$.

Consider an increasing sequence of h_s, $h_1 < h_2 < ....$ Let h_{s+} be the largest of those h_s for which the segments D_{s-1} and D_s have a point in common. Let us call this window length h_{s+} 'optimal' and determine the IF estimates with data-driven optimal window length as $\hat{\omega}_{h_{s+}}(t)$.

The basic idea behind this choice is as follows: If the segments D_{s-1} and D_s do not have a point in common it means that at least one of the inequalities (18) does not hold, i.e. the bias is too large as compared with the standard deviation in (17). Thus, the statistical hypotheses to be tested for the bias is given in the

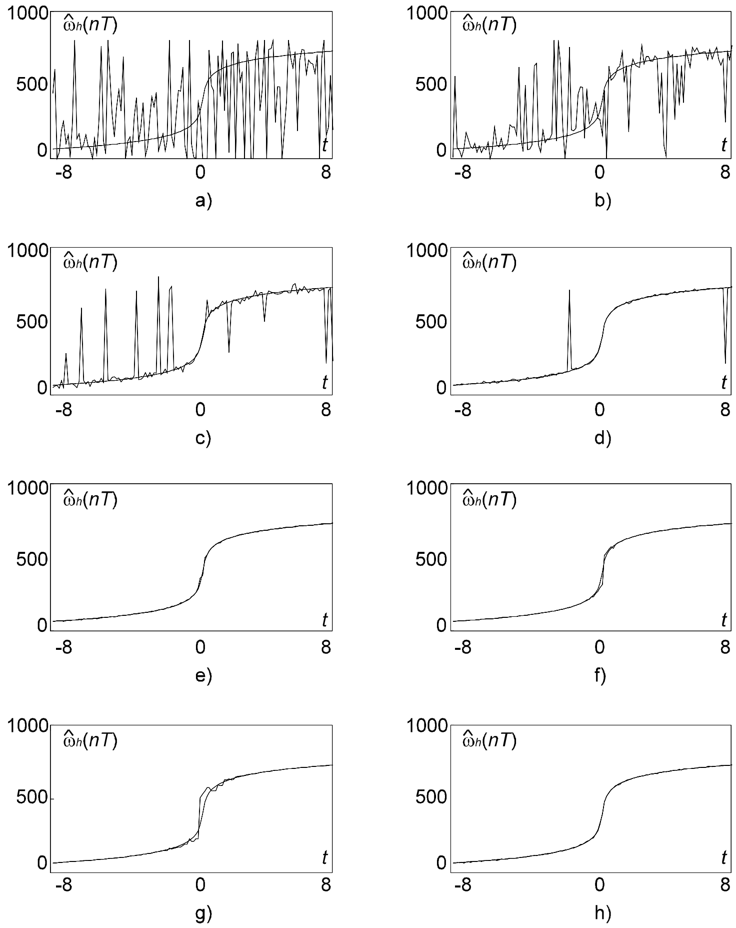

Fig. 3. a)-g) IF estimation of the signal using robust spectrogram with the constant window lengths $N = \{8, 16, 32, 64, 128, 256, 512\}$, respectively; h) IF estimation using robust spectrogram with the adaptive window length.

form of the sequence of inequalities (18) and the largest length h_s for which these inequalities have a point in common is considered as a bias-variance compromise, when the bias and variance are of the same order. Details on this two-segments intersection approach may be found in [21], [22].

B. Algorithm

Let us initially assume that the amplitude A and the standard deviation σ of the noise are known. Let H be an increasing sequence of the window length values

$$H = \{h_s \mid h_1 < h_2 < h_3 < ... < h_J\}. \quad (20)$$

In general, any reasonable choice of H is acceptable. In particular, the lengths with dyadic numbers $N_s = 2N_{s-1}$ of observations within the window length, until the largest h_J is reached, will be assumed. This scheme corresponds to the radix-2 FFT algorithms. Note that the relation between the window length and the number of observation within that length is $h_s = N_s T$. However, we want to emphasize that the minimum window size h_1 should not be too small (say $h/T > 20 \div 40$) in order to preserve the property of algorithm to be robust with respect to the heavy-tailed distribution noise.

The following steps are generated for each t.

1. The robust spectrogram is calculated for

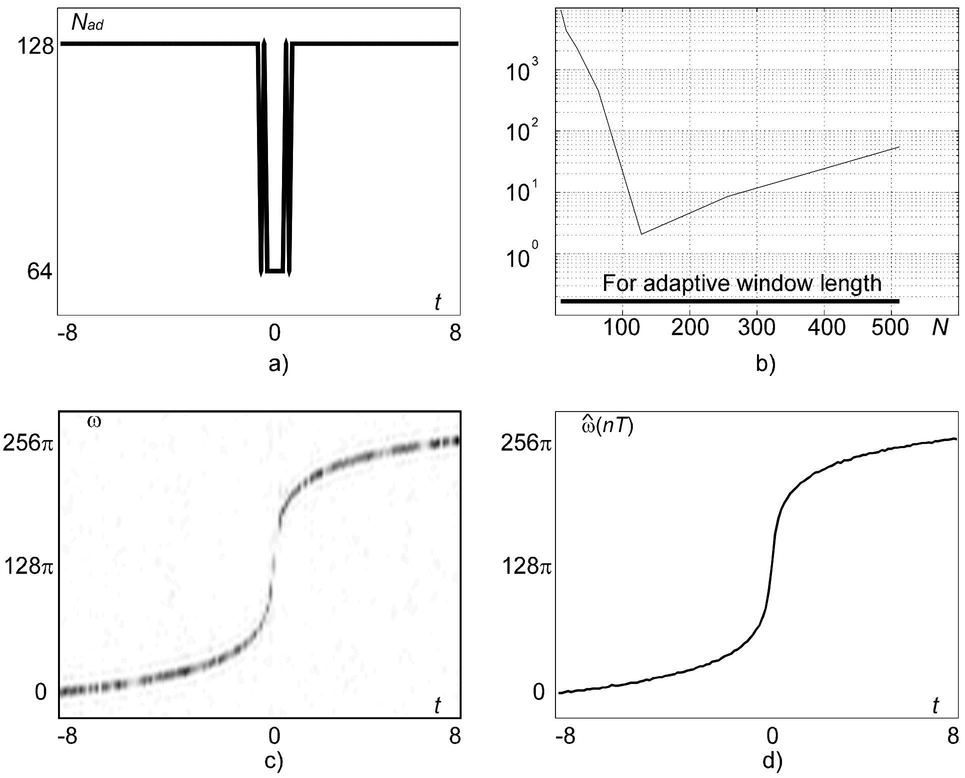

Fig. 4. a) Adaptive window length for the robust spectrogram, b) Mean square error versus window lenght, c) Robust spectrogram with adaptive window length, c) IF estimation using proposed adaptive algorithm and robust spectrogram.

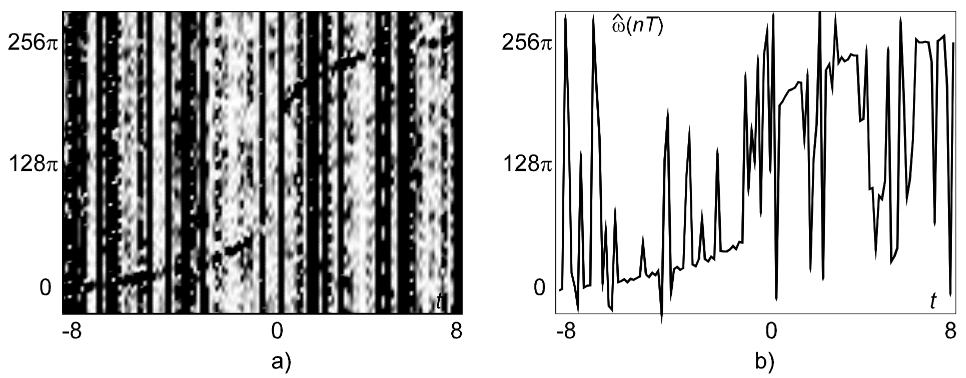

Fig. 5. a) The standard spectrogram with constant window length $N = 128$. The spectrogram values are limited to the expected maximal value equal to 1. b) IF estimation using standard spectrogram.

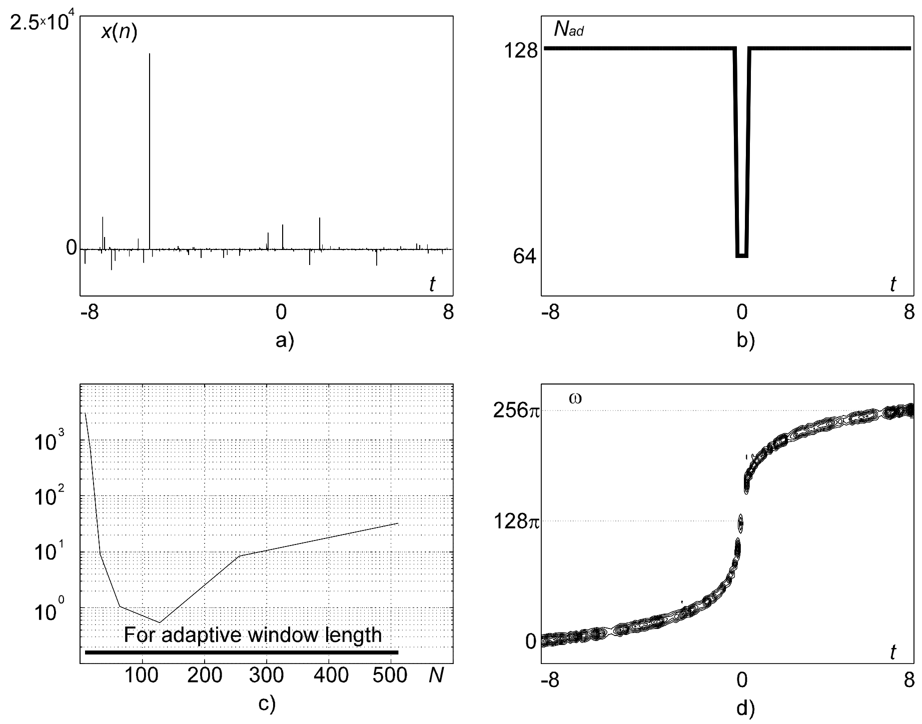

Fig. 6. a) Signal corrupted with the Cauchy noise, b) Adaptive window length for the robust spectrogram, c) Mean square error versus window lenght, d) Robust spectrogram with adaptive window length.

all of $h_s \in H$. Thus, we obtain a set of robust spectrograms for a fixed time instant t

$$\{I_A(\omega, t; h_s)\},\ h_s \in H.$$

The IF estimates are found as

$$\hat{\omega}_{h_s}(t) = \arg[\max_{\omega \in Q_\omega} I_A(\omega, t; h_s)]. \tag{21}$$

2. The upper and lower bounds of the confidence intervals D_s in (19) are built as follows

$$U_s(t) = \hat{\omega}_{h_s}(t) + 2\kappa\sigma(h_s),$$

$$L_s(t) = \hat{\omega}_{h_s}(t) - 2\kappa\sigma(h_s). \tag{22}$$

The variance $\sigma^2(h_s)$ is estimated by $\hat{\sigma}^2(h_s) = \hat{\sigma}^2(h_J)h_J^3/h_s^3$, where $\hat{\sigma}^2(h_J)$ is the variance estimation obtained by using the widest window h_J, according to

$$\hat{\sigma}^2(h_J) = \frac{1}{N}\sum_{i=1}^{N} |x(t+iT)|^2 - \hat{A}^2,$$

while $\hat{A}$ is the estimated amplitude of signal. It can be obtained applying the methods described in [21] on signal $x(t+nT)/e(t+nT)$, where $e(t+nT)$ is the error (4).

3. The optimal window length h_{s^+} is determined as the largest $s = s^+$ $(s = 1, 2, ..., J)$ when

$$\left|\hat{\omega}_{h_{s-1}}(t) - \hat{\omega}_{h_s}(t)\right| \le 2\kappa(\sigma(h_{s-1}) + \sigma(h_s))$$

is still satisfied,

$$\hat{h}(t) = h_{s+}(t), \tag{23}$$

and $\hat{\omega}_{\hat{h}(t)}(t)$ is the adaptive IF estimator with the data driven window for a given instant t.

4. The robust spectrogram with the optimal window length is

$$I^+(\omega, t) = I_A(\omega, t; \hat{h}(t)). \tag{24}$$

Steps 1 through 4 are repeated for each considered time-instant t.

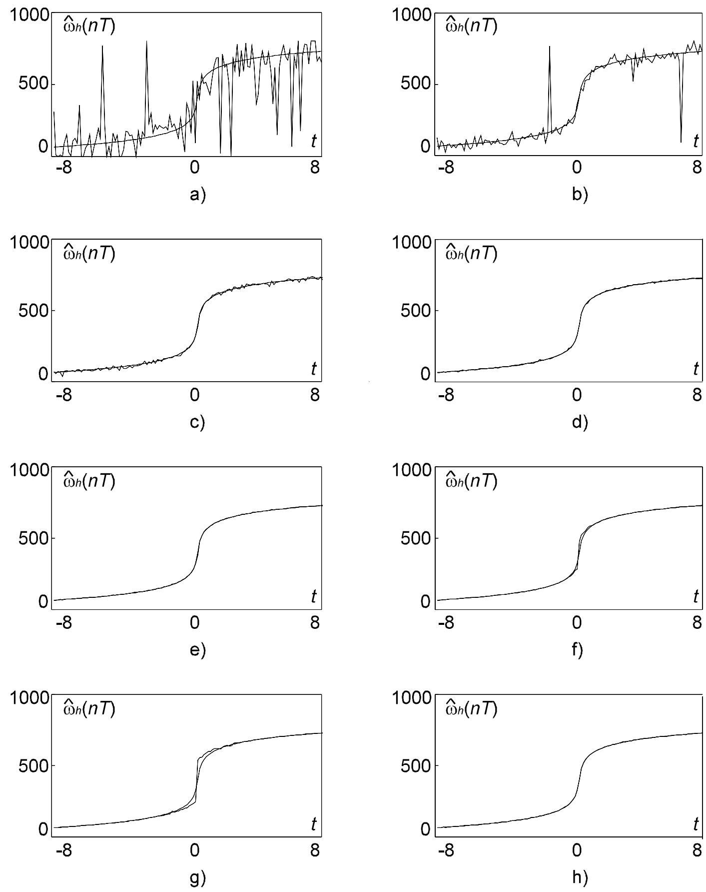

Fig. 7. a)-g) IF estimation of the signal corrupted with the Cauchy noise using robust spectrogram with the constant window lengths $N = \{8, 16, 32, 64, 128, 256, 512\}$, respectively; h) IF estimation using robust spectrogram with the adaptive window length.

IV. Examples

Example 1: Signal with the IF given as

$$\Omega(t) = \begin{cases} 180\pi - 126\pi|t| & |t| < 1 \\ 52\pi + 2\pi|t| & |t| \geq 1 \end{cases} \quad (25)$$

is considered. The signal is embedded with a high amount of heavy tailed noise

$$\varepsilon_H(nT) = 1.5(\varepsilon_R^3(nT) + j\varepsilon_I^3(nT))/\sqrt{2}, \quad (26)$$

where $\varepsilon_R(nT)$ and $\varepsilon_I(nT)$ are mutually independent white Gaussian noises $N(0, 1)$. In this case standard spectrogram is useless for IF estimation, Fig.1. Application of the robust spectrogram, Section II, along with the algorithm from Section III resulted in the adaptive window length that is shown in Fig.1b. Robust spectrogram calculated using the adaptive window length is shown in Fig.1d. It is easy to conclude that, in contrast to the standard spectrogram, the robust adaptive one is almost not influenced by the heavy tailed noise. Mean square error of the IF estimation, using the robust spectrogram, versus window length is shown in Fig.1c. The straight line shows mean square error for the IF estimation using adaptive robust spectrogram. We may conclude that the adaptive estimation produces smaller mean square error than *the best constant window length, which is also a priori unknown.* The IF estimates using different window lengths in the robust spectrogram, along with the adaptive one, are shown in

Fig.2. Obviously, for slow IF changes adaptive algorithm takes wider window length, while for faster changes it takes narrower window length, as expected.

Example 2: Consider now a signal with highly nonlinear IF

$$\Omega(t) = 20\pi \text{asinh}(12.5t) + 128\pi \tag{27}$$

The signal is embedded with same kind of noise as in the previous example. The IF estimates are shown in Fig.3. The adaptive window length, mean square error, the robust spectrogram with adaptive window length, and the IF estimate using the proposed algorithm and the robust spectrogram are shown in Fig.4.

Next, consider signal (27) corrupted with noise $\varepsilon_C(nT) = \varepsilon_R(nT) + j\varepsilon_I(nT)$, where $\varepsilon_R(nT)$ and $\varepsilon_I(nT)$ are independent Cauchy noises. The probability density function of the Cauchy noise is

$$g(x) = \frac{a}{\pi(1 + (ax)^2)}$$

Note, that the Cauchy noise can be simulated as $\varepsilon_R(nT) = \varepsilon_1(nT)/\varepsilon_2(nT)$, where $\varepsilon_i(nT)$, $i = 1, 2$, is the Gaussian white noise with variance σ_i, and $a = \sigma_2/\sigma_1$. Signal (27) corrupted with the Cauchy noise with $a = 0.5$ is shown in Fig.6a. Standard spectrogram could not produce any reasonable result for this noise. The standard spectrogram with constant window length $N = 128$ and the IF estimate based on it are shown in Fig.5. The adaptive window length is shown in Fig.6b. The mean square error is shown in Fig.6c, while the robust spectrogram with adaptive window length is shown in Fig.6d. The IF estimators are shown in Fig.7. Obviously the robust spectrogram produced very accurate results, in this case, as well.

Finally, consider signal (27) with mixed Gaussian and heavy tailed distribution noise (26). Two cases will be considered:

(A) One being a sum of these two noises

$$\varepsilon(nT) = (1 - \beta)\varepsilon_G(nT) + \beta\varepsilon_H(nT)$$

where $\varepsilon_G(nT)$ is a white complex Gaussian noise with variance σ_ε^2, and independent real and imaginary parts. Real and imaginary parts of the heavy tailed noise are given as $\text{Re}\{\varepsilon_H(nT)\} = (\text{Re}\{\varepsilon_G(nT)\})^3$ and $\text{Im}\{\varepsilon_H(nT)\} = (\text{Im}\{\varepsilon_G(nT)\})^3$.

(B) The other case when the probability density function is of the form

$$p(\varepsilon) = (1 - \beta)p_G(\varepsilon) + \beta p_H(\varepsilon), \tag{28}$$

where $p_G(\varepsilon)$ is the Gaussian probability density $N(0, \sigma)$, $p_H(\varepsilon)$ is the probability density of an heavy tailed distributed noise, while $1 \geq \beta \geq 0$ determines a proportion of these random noises in the mixed noise with the probability density $p(\varepsilon)$.

Different proportions $\beta = \{0, 0.2, 0.4, 0.6, 0.8, 1\}$ are considered. The IF estimators with standard and robust spectrogram are compared for different β in cases (A) and (B), Tables I,II,III,IV. In both cases similar conclusions hold. For $\beta = 0$ (pure Gaussian noise)

TABLE I
Mean square error for the standard spectrogram with constant and adaptive window lengths for the noise in case (a)

β	$N = 64$	*Adaptive*
0	0.6856	0.1455
0.2	1.0392	0.5467
0.4	5.1502	4.6032
0.6	1060.9	253.22
0.8	3964.3	1434.3
1	5503.6	2819.2

TABLE II
Mean square error for the robust spectrogram with constant and adaptive window lengths for the noise in case (a)

β	$N = 64$	*Adaptive*
0	0.7124	0.1968
0.2	0.8974	0.1930
0.4	1.0478	0.2363
0.6	1.0602	0.2356
0.8	1.0843	0.2721
1	1.4888	0.2851

TABLE III

Mean square error for the standard spectrogram with constant and adaptive window lengths for the noise in case (b)

β	0	0.2	0.4	0.6	0.8	1
N=8	1807.4	6440.6	7865.7	9231.7	9551.5	12581
N=16	52.897	4098.3	4965.5	5690.4	6452.0	8856.3
N=32	5.8204	2316.8	2223.4	3706.6	4715.8	5485.9
N=64	0.6187	1287.4	1699.1	1865.1	2608.8	3625.6
N=128	1.0585	8.4956	234.84	750.23	1261.4	1386.5
N=256	10.011	395.22	520.47	966.21	1182.8	269.60
N=512	32.244	1777.0	2748.3	2725.9	2700.1	222.92
Adaptive	0.1973	227.18	498.03	665.69	1221.7	1442.0

TABLE IV

Mean square error for the robust spectrogram with constant and adaptive window lengths for the noise in case (b)

β	0	0.2	0.4	0.6	0.8	1
N=8	3019.8	4763.9	4095.0	4914.0	5473.8	6869.1
N=16	130.89	255.20	141.88	364.21	779.45	726.77
N=32	61.585	15.397	13.625	10.599	9.3672	8.8703
N=64	0.7643	1.0887	1.0104	0.9121	0.8122	1.3339
N=128	0.7484	2.0492	1.3116	1.1158	1.4329	1.5898
N=256	9.6263	9.2244	8.9314	8.9170	8.9780	9.4466
N=512	31.821	33.203	32.380	33.558	33.117	32.471
Adaptive	0.2152	0.4152	0.5998	0.5502	0.1435	0.1914

both of the spectrograms show similar results, first rows in Tables I and II. The standard spectrogram is just slightly better than the robust spectrogram. By increasing of β the IF estimation error for the standard spectrogram significantly increases, while the robust spectrogram based IF estimation error remains low and close to the case of the Gaussian noise only. The mean square IF estimation errors for the standard spectrogram with $N = 64$ (this spectrogram have the smallest IF with constant window width) and adaptive window length, for various β, are given in Tables I and II. Note that the quantization error for all examples is 0.08333, defining the lower accuracy limit. It can be seen that as the amount of the heavy tailed noise, i.e., β, increases the error increases significantly. The similar results are presented for the standard and robust spectrogram in Tables III and IV. The standard spectrogram becomes useless for the estimation when $\beta \geq 0.4$ in case (A) and for $\beta \geq 0.2$ in case (B). From the last column of Table II or the last row of Table IV we can see that the mean square error remains low and of the same order for all proportions between Gaussian and heavy tailed noise, when the robust adaptive spectrogram is used for the IF estimation. However, when the standard spectrogram is used that is not the case, the last column of Table I or the last row of Table III.

V. Further development

The L_2-norm in (3) results in the estimate $\hat{C}_h(t,\omega)$ linear on observations and the conventional power spectrogram (1). It is well known that linear estimates give good results and good filtering of observation errors provided they are subjected to the Gaussian distribution. However, the heavy-tailed distribution errors are able to destroy any linear estimate and, in particular, can result in a complete degradation of the beamformer power function.

If a noise distribution is known the maximum likelihood (ML) method is a powerful universal tool in order to design the best estimates for the given distribution. However, for many cases these ML estimates are quite sensitive with respect to a deviation from the parametric model and the hypothetical distribution. Even a slight deviation from the hypothesis can result in a strong degradation of the ML estimate. In particular, this high-sensitivity with respect to the distribution hypotheses has a place for the Gaussian noise. The minimax robust approach has been developed in statistics as an alternative to the conventional ML in order to decrease the mentioned above sensitivity of the ML estimates to the hypothetical distribution and to improve their efficiency in an environment of errors with the heavy-tailed distributions (see. [1], [7], [8], [19]). Loss function $F(e) = |e|$ is just an example of a function that enables robust estimation in some classes of heavy-tailed distributions. More classes of the robust distributions, as well as more details about the minimax robustness concept can be found in [1], [8], [19].

Here we wish to present a loss function for which the IF estimation algorithm is slightly different from the considered one. Based on the previous results (Tables I-IV) we can conclude that a combination of loss functions $F(e) = e^2$ for small errors and $F(e) = |e|$ for large error would further improved results. This loss function can be defined as

$$F(e) = \begin{cases} e^2/2 & |e| \leq \Delta \\ \Delta|e| - \Delta^2/2 & |e| > \Delta \end{cases} \qquad (29)$$

It is optimal in the minimax sense over the class of distributions (28). The adjustable parameter Δ in (29) depends on the noise parameters [8]. It has been used, for example, in the robust multiuser detection in nonGaussian channels in [26].

For the loss function (29) the complex valued amplitude $\hat{C}_h(t,\omega)$ can be calculated as a solution of the equation:

$$\hat{C}_h(t,\omega) = \frac{P_h(t,\omega)}{w_m} \qquad (30)$$

where

$$P_h(t,\omega) = \sum_{n,|e|\leq\Delta} w_h(nT)x(t+nT)\exp(-j\omega nT) + \sum_{n,|e|>\Delta} w_h(nT)e^{-1}(nT)x(t+nT)\exp(-j\omega nT)$$

$$w_m = \sum_{n,|e|\leq\Delta} w_h(nT) + \sum_{n,|e|>\Delta} w_h(nT)e^{-1}(nT)$$

$$|e| = |x(t+nT) - C_h(t,\omega)\exp(j\omega nT)|$$

From (29) it is clear that when error values $e(nT)$ are greater than Δ values of signal $x(t+nT)$ are decreased by $e^{-1}(nT)$. Calculation of this modified robust spectrogram can be done according to the previously described recursive procedure for the robust spectrogram, Section II. In our simulations we use $\Delta = 1$. The experiments show fast convergence of the algorithm.

As an example, consider the IF estimation of the signal (27) with noise (28) using the modified robust spectrogram. Table V gives the IF estimation errors for the modified robust spectrogram with constant ($N = 64$) and for the adaptive window length. The modified robust spectrograms have smaller IF estimation errors than corresponding robust spectrograms, Table IV. For this kind of noise standard spectrogram [17] is not able to produce any acceptable result.

VI. Conclusion

The robust spectrogram, being time-varying form of the robust $M-$periodogram, with the varying adaptive window size is developed. The intersection of confidence intervals rule

TABLE V
MEAN SQUARE ERROR FOR THE MODIFIED ROBUST SPECTROGRAM WITH CONSTANT AND ADAPTIVE WINDOW LENGTHS

β	$N = 64$	*Adaptive*
0	0.6714	0.1632
0.2	1.0021	0.3752
0.4	0.9866	0.5751
0.6	0.9235	0.5493
0.8	0.7726	0.1344
1	1.2568	0.1846

is applied for varying window size selection. Simulation demonstrates that the new spectrogram gives the estimates of the varying IF which are strongly robust with respect to the noise having a heavy-tailed distribution.

REFERENCES

[1] D.F. Andrews, P.J. Bickel, F.R.Hampel, P.J. Huber, W.H. Rogers, and J.W. Tukey: *Robust estimation of location: Survey and advances.* Prinston Univ. Press, Prinston, N.J., 1972.

[2] B. Boashash: "Estimating and interpreting the instantaneous frequency of a signal-Part 1: Fundamentals", *Proc. IEEE*, vol. 80, no.4, pp.519-538, April 1992.

[3] L. Cohen and C. Lee: "Instantaneous bandwidth", in *Time-frequency signal analysis*, B. Boashash ed., Longman Cheshire, 1992.

[4] L. Cohen: "Distributions concentrated along the instantaneous frequency" *SPIE*, Advanced Signal Proc. Alg., Architect., and Implementations, vol. 1348, pp.149-157, 1990.

[5] L. Cohen: *Time-frequency analysis*, Prentice Hall, Englewood Cliffs, N.J., 1995.

[6] A. Goldenshluger and A. Nemirovski: "On spatial adaptive estimation of nonparametric regression". *Mathematical methods of Statistics,* vol. *6, 2, pp. 135-170*, 1997.

[7] P.J. Huber: "Robust regression: asymptotics, conjectures and Monte Carlo", *Ann. Math. Statist.*, vol. 1, N 5, pp. 799-821, 1973 .

[8] P.J. Huber: *Robust statistics*, John Wiley &Sons Inc., 1981.

[9] V. Katkovnik: "Local polynomial periodogram for time-varying frequency estimation," *South African Statistical Journal,* vol. 29, pp. 169-198, 1995.

[10] V. Katkovnik: "Nonparametric estimation of instantaneous frequency", *IEEE Trans. on Information Theory*, vol. 43, no.1, pp.183-189, Jan.1997.

[11] V. Katkovnik: "Local polynomial periodogram for signals with the time-varying frequency and amplitude", *Proc. IEEE Int Conf. ASSP*, Atlanta, Georgia, May 7-10, 1996, USA , pp. 1399-1402.

[12] V. Katkovnik: "Adaptive local polynomial periodogram for time-varying frequency estimation", in *Proc. IEEE-SP on TFTS Analysis*, Paris, pp.329-332, June 1996.

[13] V. Katkovnik: "Discrete-time local polynomial approximation of instantaneous frequency", *IEEE Trans. Sig. Proc.*, vol. 46, N 10, 1998, pp. 2626-2637.

[14] V. Katkovnik: "Robust $M-$periodogram", *IEEE Trans. Sig. Proc.*, vol.46, No. 11, 1998, pp. 3104-3109.

[15] V. Katkovnik: "On multiple window local polynomial approximation with varying adaptive bandwidths," in *Proceedings in Computational Statistics: 13th symposium held in Bristol,*1998; *COMPSTAT. Ed.by R.Payne and P. Green,* Heidelberg. Physica-Verlag, pp. 353-*358,* 1998.

[16] V. Katkovnik: "Robust $M-$estimates of the frequency and amplitude of a complex-valued harmonic", *Sig. Proc.*, to appear.

[17] V. Katkovnik and LJ. Stanković, "Periodogram with varying and data-driven window length", *Signal Processing*, vol. 67, N3, 1998, pp. 345-358.

[18] C.L. Nikias and M. Shao: *Signal Processing with Alpha-Stable Distributions and Applications.* New York: Wiley, 1995.

[19] B.T. Poljak and J.Z. Tsypkin: "Robust identification", *Automatika*, vol. 16, pp. 53-63, 1980.

[20] Proceedings of the IEEE, *Special issue on Time-Frequency Analysis*, vol. , no.9, Sept. 1996.

[21] LJ. 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.

[22] LJ. Stankovic, V. Katkovnik: "Algorithm for the instantaneous frequency estimation using time-frequency distributions with variable window width", *IEEE Sign. Proc. Letters*, vol. 5, No. 9, Sept. 1998, pp. 224-227.

[23] LJ. Stanković, S.Stanković: "On the Wigner distribution of discrete-time noisy signals with application to the study of quantization effects", *IEEE Trans. Sig. Proc.*, vol. 42, no.7, pp.1863-1867, July 1994, .

[24] LJ. Stanković: "A method for improved distribution concentration in the time-frequency analysis of multicomponent signals the L-Wigner distribution" *IEEE Trans. Sig. Proc.*, vol. 43, no.5, May 1995.

[25] A. B. Tsybakov: "Robust reconstruction of functions by the local-approximation methods," *Problems of Information Transmission*, vol. 22, pp. 133–146, 1986.

[26] X. Wang, H. V. Poor: "Robust multiuser detection in non-Gaussian channels", *IEEE Trans. Sig. Proc.*, Vol. 47, No.2, pp.289-305, Feb.1999.