Approaches to Time Delay Estimation for Wideband Signals Embedded in Non-Gaussian Noise Received by Two Sensors

Viacheslav Oliinyk[a], Vladimir Lukin[a], Petro Brysin[a]

[a]*National Aerospace University, 17 Chkalova Street, Kharkiv, 61070, Ukraine*

Abstract

A task of time delay estimation for interferometric antennas for wideband signals embedded in Gaussian noise is classical. However, in practice, there are several factors that make its solving problematic including the non-Gaussian nature of the additive noise, low signal-to-noise ratio, limited time of signal registration, source motion, restrictions imposed by the necessity to process data in real-time, and so on. In this paper, we consider possible approaches to providing efficient processing of signal/noise mixtures acquired by two sensors forming a stationary base within a limited time of data registration for signal-noise ratio about unity or smaller in the case of non-Gaussian noise with a priori knowledge on its statistical characteristics. In such conditions, not only quite large normal estimates are possible but also abnormal estimates (outliers) might be observed with high probability. This makes the task of signal source tracking extremely complicated. Thus, we concentrate on considering the approaches to decrease the probability of abnormal error occurrence rather than to reduce the variance of normal estimates. Peculiarities of wideband signals are discussed. Three possible approaches are studied. Their advantages and drawbacks are considered. The final recommendations are given.

Keywords 1

Wideband signal, non-Gaussian noise, interferometric antenna, abnormal estimates, robust processing

1. Introduction

The task of providing an accurate estimation of time delay appears in numerous applications including hydroacoustics (passive sonars) [1, 2], teleconferencing [3, 4], seismic data analysis [5], and others [6, 7]. Originally, the task was formulated as time delay estimation (TDE) or direction of arrival (DOA) estimation for two sensors with fixed and a priori known distance between or for an antenna array with fixed and known geometry for a point-like non-moving signal source where the signal spectrum is known in advance, signal-to-noise ratio (SNR) is considerably larger than unity, observation time is large enough, and noise is additive white Gaussian and independent for all sensors used [2, 8]. For such idealized conditions, methods of optimal signal processing were designed [2] and potential accuracy of TDE that depend on signal and noise power spectra and observation time was determined.

However, it was understood many years ago that a few or several aforementioned assumptions could be violated and can be different depending on an application at hand. For example, for hydroacoustics, SNR can be low (smaller than unity), signal source can move, signal spectrum might change in time and be known only approximately. These obstacles result in necessity to carry out elementary processing for intervals of limited length [9] and non-zero probability of obtaining abnormal estimates of time delay and DOA [1]. This led to necessity to apply more complex methods

IntelITSIS'2023: 4th International Workshop on Intelligent Information Technologies and Systems of Information Security, March 22–24, 2023, Khmelnytskyi, Ukraine
EMAIL: v.o.oliinyk@khai.edu (Viacheslav Oliinyk); lukin@ai.kharkov.com (Vladimir Lukin); p.v.brysin@khai.edu (Petro Brysin)
ORCID: 0000-0002-0914-8664 (Viacheslav Oliinyk); 0000-0002-1443-9685 (Vladimir Lukin)

of joint estimation of TDE and its derivatives [10, 11]. The estimation of derivatives allows to perform tracking better especially if efficient and robust methods and algorithms are employed for this purpose [10]. The use of particle filtering results in more accurate time delay estimation as well [11].

If accuracy of original elementary estimates of time delay is not accurate enough it is worth employing robust post-filtering of elementary time delay estimates [12]. This helps to improve tracking but under condition that probability of abnormal errors of TDE for elementary intervals is not too high. Otherwise, the tracking algorithm might fail and this can lead to undesired consequences [13]. The signal source can be “lost” and the process of its detection should be renewed.

In teleconferencing, TDE used for speaker tracking has other peculiarities. First, a signal under interest (human speech) is non-stationary [14] and changes its spectrum more than for the hydroacoustic case. Recall here that standard Fourier-based methods of spectral analysis can be then replaced by modern methods of temporal-spectral analysis [14]. Second, speech contains pauses that might cause abnormal elementary estimates of time delay if a pause falls on an elementary interval of signal registration. Third, a source of a signal might move more unpredictably compared to the source motion in hydroacoustics.

In other applications, there can be other peculiarities. However, one common specific feature is that noise statistical characteristics are non-Gaussian where the probability density function of the noise can be a priori unknown or change in time [15-17]. Impulsivity of the noise leads to necessity to apply special methods of signal detection [17] and, e.g., processing with fractional lower order moments [15]. Additive noise non-Gaussianity also leads to a radical reduction of TDE accuracy especially if the noise is intensive and impulsive (having heavy tails) [18]. Then, the task is to make signal processing robust to reduce the negative influence of the noise impulsivity.

Thus, the goal of this paper is to consider approaches that provide appropriately fast and efficient data processing in TDE for an antenna consisting of two sensors spaced by a fixed and a priori known distance. Certainly, it is supposed that digital signal processing is performed where amplifiers that can be used as initial cascades have identical amplitude-frequency characteristics. The additional goal of this paper is to point out new directions of research in the field.

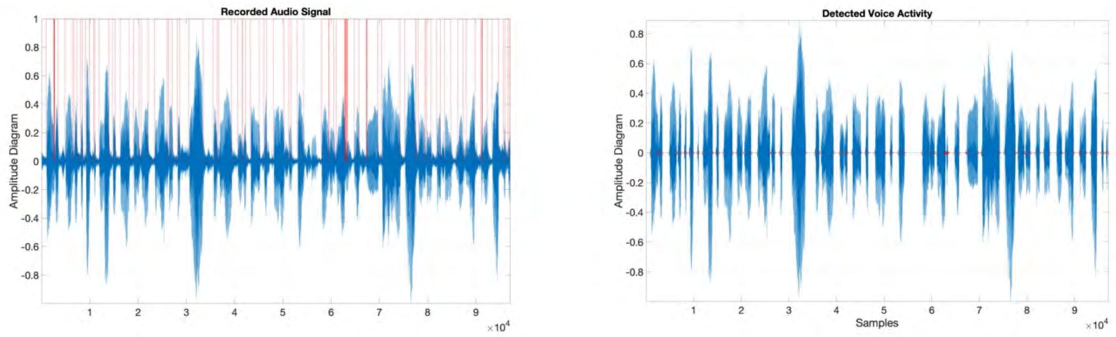
2. Signal/Noise Model and Standard Approach to TDE

We assume that one has two point-like (small dimension) sensors displaced by a distance L (in most applications, it is supposed that the antenna is positioned horizontally [1, 4]). The informative signal is supposed wideband and random. It can be almost stationary or locally stationary as in hydroacoustics or essentially non-stationary as in teleconferencing. In the latter case, some additional information on informative signal properties can be needed for TDE.

We have carried statistical analysis of pause duration in a speech signal. Fig. 1,a shows an example of such a signal of duration 10 s with the total number of signal samples equal to 97000 with marked intervals of speech and pauses. It is possible to see that the signal has intervals of high and low intensity. If an elementary interval of signal processing (mean and variance estimation, spectral analysis, etc.) is less than pause duration, than the following types of intervals are possible: a) intervals of high intensity, b) intervals of low intensity (pauses), c) intervals of transition from high intensity to low one or vice versa.

For signal/noise mixture processing for intervals that corresponds to pauses, obtaining of abnormal estimates of time delay (and DOA) for speech source is practically guaranteed. In the case of interval falling in the transition area, it is also quite possible to have abnormal errors or normal estimates with quite large deviations from a true value of time delay.

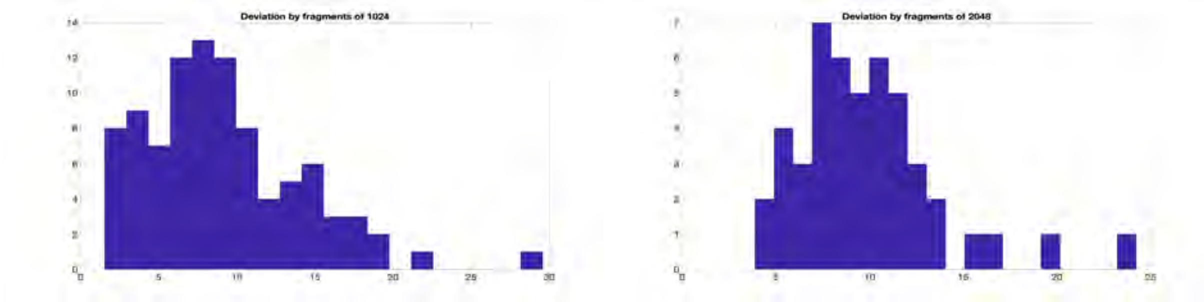
Note that there are pause detectors for speech signals [19, 20]. An example of their operation is shown in Fig. 1,b. It is seen that pauses appear quite often. The pause detector offered at https://www.mathworks.com/help/audio/ref/voiceactivitydetector-system-object.html#d122e45905 analyzes speech in fragments of duration about 30 ms. It has detected 127 blocks of the aforementioned size as pauses whilst 279 have been identified as signals. Thus, for 30 ms blocks, about 30% of them are identified as pauses where one can expect that estimation of time delay can be problematic in the sense of obtaining abnormal estimates.



a b

Figure 1: Speech signal with different fragments (a) and automatically detected speech and pause areas (b)

To analyze the distribution of power in speech intervals of another (fixed) size, we have estimated variances of the considered wideband signal. The histograms are shown in Fig. 2 for two cases: a) observation intervals containing 1024 samples with a duration about 0.1 s; b) observation intervals of duration about 0.2 s that contain 2048 samples. It is seen that the distributions have slightly different properties. The first one corresponds to quite many intervals that have very low intensities (Fig. 2,a) and the distribution also has a heavy right hand tail. The second distribution has larger minimal values and smaller maximal values with the tendency to normalization.



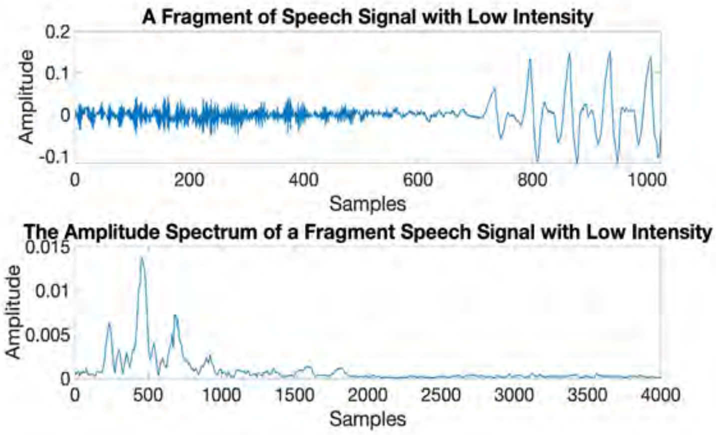
a b

Figure 2: Histograms of variance (signal power) estimates for intervals of the size 1024 samples (a) and 2048 samples (b)

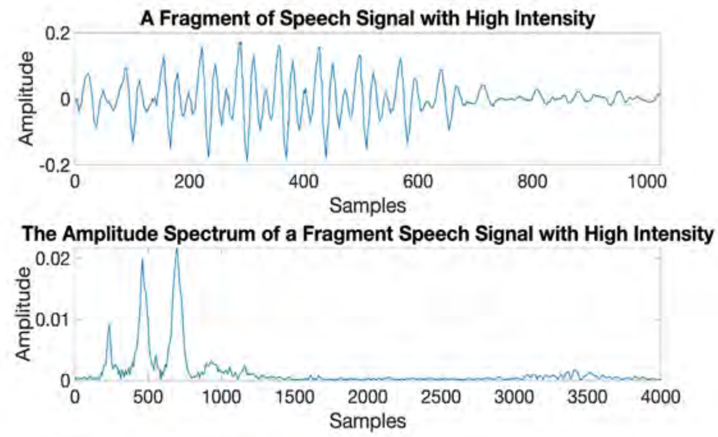
Let us also give some examples of signals for different types of intervals. Examples of speech signal realizations and Fourier spectra for them are given in Fig. 3 for low and high-intensity fragments. Although a few low frequency quasi-harmonic components concentrated in the limits from about 150 to 800 Hz are present in both spectra, the spectra are essentially different.

Thus, there are sufficient differences in signal properties in passive sonar and teleconference applications. In the latter case, the signal component is more non-stationary, which can lead to a higher probability of elementary abnormal errors of TDE for the same mean SNR.

Consider now additive noise properties. For sensors displaced by a rather large distance L (and this is needed to provide appropriately high accuracy of DOA estimation [2]), noise in sensors can be considered independent. Then, we have to be more interested in its statistical characteristics and the corresponding models. In this sense, quite many experiments have been conducted to establish the model and its parameters. Symmetric α-stable (SαS) distribution has been shown to be a good option [15, 16]. This distribution is described by two parameters that can be varied. The first parameter is $\alpha_{S\alpha S}$, which is responsible for tail heaviness. A smaller $\alpha_{S\alpha S}$ corresponds to heavier tails where $\alpha_{S\alpha S}$ relates to standard Gaussian distribution. In practice, it is quite difficult to meet situations where $\alpha_{S\alpha S} < 1.2$. However, the $\alpha_{S\alpha S}$ values of the order 1.4-1.8 are quite typical [15, 16]. The second parameter γ describes the data scale. Larger values of γ correspond to greater noise intensity. If one varies γ in simulations, this is equivalent to SNR variation.



a



b

Figure 3: Examples of speech realizations and spectra for them for low intensity fragment (a) and high intensity fragment (b)

However, the SαS distribution used as the noise model has one drawback. Theoretically, the variance of such a noise is infinite [16]. This does not correspond to physical assumptions on the power of the noise that should be limited. Besides, this makes problematic the simulation of different SNRs that is usually employed in the analysis of TDE method performance. The only thing one can do in simulations based on the SαS distribution of the noise is to get dependences of the accuracy criteria as the variance of normal estimates or probability of abnormal estimates on different values of γ [18-20].

Thus, we consider that two sensors receive the following mixtures of information wideband noise-like (WNL) signal and additive noise:

$$x_1(t) = s(t) + \xi_1(t), \quad x_2(t) = s(t - \tau_0) + \xi_2(t) \tag{1}$$

where $s(t)$, $t = [T_b; T_e]$ is the WNL information signal (irradiated by a considered signal source); $\xi_1(t)$ and $\xi_2(t)$ are the noise realizations for the first and second sensors, respectively, τ_0 is the true delay value. The WNL signal mean is supposed to be equal to zero. Similarly, the means (more correctly, location parameters) of $\xi_1(t)$ and $\xi_2(t)$ are assumed equal to zeroes, too. The observation interval starting and ending time instances are denoted as T_b and T_e where it is assumed that the maximal possible τ_0 (which is determined by the distance between receivers and the speed of wave propagation in a given medium C as $\tau_{max} = L/C$) is sufficiently smaller than $T_e - T_b$. In turn, from the upper side, $T_e - T_b$ is restricted by source motion (angular) speed: $T_e - T_b$ should not be so large that the absolute value of $\tau_0(T_e) - \tau_0(T_b)$ exceeds the main lobe width of autocorrelation function $R_{WNLS}(\tau)$. Besides, in some implementations of data processing algorithms, it is desired to have such

$T_e - T_b$ with taking into account the sampling rate Δt that $(T_e - T_b)/\Delta t$ is equal to power of two (e.g., 1024 or 2048 as in examples above and in some simulations below).

A traditional approach to TDE in a Gaussian noise environment is to calculate cross-correlation function (CCF) $Y(\tau)$ of a received signal, to find τ that corresponds to its global maximum and accept it as the estimate $\hat{\tau}$ of time delay (that can be further used for estimation of DOA). Conventional expression for CCF (without normalization) is

$$Y(\tau) = \int_{-T/2}^{T/2} x_1(t)x_2(t+\tau)dt \quad (2)$$

where $T = T_e - T_b$. In fact, for a given application, it is enough to calculate $Y(\tau)$ in the limits from $-\tau_{max}$ to τ_{max} that can accelerate calculations. Another conventional way to accelerate processing is to determine cross-spectrum as

$$\dot{S}_{12}(\omega) = \dot{S}_1(\omega)S_2^*(\omega)$$

where $\dot{S}_1(\omega) = FFT\big(x_1(t)\big)$ and $\dot{S}_2(\omega) = FFT\big(x_2(t)\big)$, $\omega = 2\pi f$ is the cyclic frequency. FFT means fast Fourier transform that, under certain conditions as a rather large number of samples N being the power of two, allows calculating the cross-spectrum and the CCF estimate faster than directly where the CCF estimate is obtained as

$$Y(\tau) = FFT^{-1}(\dot{S}_{12}(\omega))$$

where FFT^{-1} denotes inverse fast Fourier transform (FFT). Obviously, $Y(\tau) = R_{WNLS}(\tau - \tau_0) + n(\tau)$ where $n(\tau)$ corresponds to the sum of $\int_{-T/2}^{T/2} S_1(t)\xi_2(t+\tau)dt$, $\int_{-T/2}^{T/2} \xi_1(t)S_2(t+\tau)dt$, and $\int_{-T/2}^{T/2} \xi_1(t)\xi_2(t+\tau)dt$ that are all zero mean random functions containing no useful information on τ_0 and preventing its accurate estimation.

3. Possible approaches to accuracy improvement

In this Section, we consider three approaches to the improvement of TDE accuracy for elementary intervals. It is supposed that if such elementary estimates are accurate, then it will be easier to provide robust tracking of the signal source by robust post-processing of the elementary estimate sequence.

It is clear that if one obtains an estimate of $\dot{S}_1(\omega)$ closer to $\dot{S}_{WNLS}(\omega) = FFT\big(S(t)\big)$ than according to standard methodology, then the influence of $n(\tau)$ can be diminished and, thus, the accuracy of TDE can be improved. One way to do this [18] is based on the so-called robust forms of discrete Fourier transform [21, 22] (RDFT). The main assumption put in the basis of RDFT is that the complex-valued spectrum obtained by RDFT-method in discrete form can be written as

$$\dot{S}_{rob}(p) = R_{rob}(p) + jI_{rob}(p), \quad (3)$$

where R_{rob}(p) and I_{rob}(p) are some robust (with respect to outliers or impulses in a considered signal/noise mixture) estimates of RE and IM components of the spectrum. The index p in (2) relates to a frequency f_p where $f_p=p\Delta f$, $\Delta f=1/T$; $R_{rob}(p)=R_{rob}(f_p)$, $I_{rob}(p)=I_{rob}(f_p)$.

As it is known, the optimal DFT method for signals embedded in Gaussian noise averages $x(n)\cdot exp(-j2\pi f_p nT)=x(n)\cdot exp(-j2\pi pn/N)$ for each frequency:

$$\dot{X}_S(p) = \dot{X}_S(f_p) = \frac{1}{\mathrm{N}}\sum_{n=1}^{N} x(n)\,exp(-j2\pi pn/N) = mean\{x(n)\,exp(-j2\pi pn/N)\} =$$
$$= mean\{Re[x(n)\,exp(-j2\pi pn/N)]\} + jmean\{Im[x(n)\,exp(-j2\pi pn/N)]\}, \quad (4)$$

Here Re[·] and Im[·] are, in general, operators that produce real and imaginary parts of a complex-valued number, respectively. Instead of mean, these can be some robust operators (location estimators) if this provides certain benefits. Generally, the RDFT can be described as:

$$R_{rob}(p) = T\{Re[x(n)\,exp(-j2\pi pn/N)]\},\ I_{rob}(p) = T\{Im[x(n)\,exp(-j2\pi pn/N)]\} \quad (4)$$

where $T\{\cdot\}$ denotes a used robust estimator.

Some robust estimators such as sample median or α-trimmed mean [23] are well known whilst other ones such as sample myriad and meridian [24] are almost unknown. Without going deeply into the theory and practice of robust estimation, let us formulate the requirement for robust estimation for the considered application. Here it is worth stating that there are no fast algorithms for processing based on RDFT in opposite to FFT. Thus, taking into account the necessity of real-time processing in the case of tracking a source of WNL signal, the use of robust estimators that require intensive computations (sample myriad or meridian, Wilcoxon estimate, bootstrap-based estimate) is limited.

Another requirement deals with a priori knowledge of WNL signal spectrum and noise properties. Concerning the WNL signal spectrum, one might have some a priori information about it. For example, lower and upper cut-off frequencies can be known. The main sub-band of signal power concentration can be known as well. Concerning the additive noise, one might know α for the SαS distribution model (at least, approximately) or the range of its possible variation. If these characteristics are known in advance, it is possible to carry out preliminary simulations for the corresponding conditions and to choose the best robust estimator. Otherwise, one has to adapt to the situation at hand to perform some adaptive robust estimation or to provide robustness in the wide sense [23].

To describe some details of this approach, let us give some results taken from [18] and consider the approach's efficiency and its possible modifications in the future. Let us consider four robust estimators, namely standard median (the robust estimate of cross-spectrum obtained using it is denoted as $\dot{S}_{12med}(\omega)$ 0), α-trimmed mean with α_{tr}=0.25 ($\dot{S}_{12\alpha tr}(\omega)$), Hodges-Lehman estimate ($\dot{S}_{12H-L}(\omega)$) and the adaptive estimate based on switching between the Hodges-Lehman and median estimates based on distribution impulsivity estimation ($\dot{S}_{12H-L}(\omega)$, see the details in [18]). The reasons for such analysis and comparisons are the following. First, the mechanism of RDFT action is quite complicated and it is difficult to predict in advance what robust operator would provide the best results [22]. Second, the considered robust estimators possess different robustness where the median has the best ability to remove outliers (impulses) and the adaptive hard-switching operator is able to tune to the situation at hand (distribution of data to be processed).

Here and below, we consider two quantitative criteria of the method efficiency. First, we calculate and analyze the probability of abnormal errors P_{abn} equal to the ratio of abnormal estimates number to the total number of experiments (noise realizations considered). Depending on the method (data processing algorithm) computational efficiency, the total number of experiments is from 1000 to 10000. Second, the variance of normal estimates (σ^2_{norm}) is determined (sometimes we will analyze RMSE of estimates σ_{norm}). An estimate is considered normal if it differs from the true value by no more than the main lobe half-width $\delta\tau_{ml}/2$. Since the simulated WNL signal true spectrum is known in advance or can be estimated, $\delta\tau_{ml}$ can be determined in advance, too. Then, any time delay estimate in experiments can be referred to normal or abnormal.

First, one has to be sure that the approach presuming the use of RDFT is able (at least, in some practical situations) to provide better accuracy than the standard approach to TDE. The results of the preliminary tests are presented in Figure 1. The informative WNL signal has been obtained from white Gaussian noise by linear low-pass filtering in such a manner that its upper cut-off frequency is five times smaller than the data sampling frequency (equal to 20 kHz, T is about 0.05 s)). WNL signal variance is fixed and equal to unity. Equivalent SNR has been varied by changing the parameter γ of the SαS noise.

As one can see (Fig. 4,a), σ_{norm} increases if γ becomes larger, e.g., if SNR reduces. For a certain γ, saturation is achieved, σ_{norm} becomes approximately equal to $\delta\tau_{ml}/3.5$, this happens when P_{abn} exceeds 0.5 (Fig. 4,b), e.g. quite many abnormal errors are observed. The standard method (mean) produces the best accuracy (smaller σ_{norm} and P_{abn} for the same γ), at least until the TDE starts to fail. Among two other variants, the one based on the α-trimmed version of the RDFT produces better performance than the version based on the median form of the RDFT. As one can see, the comparison of the method performance according to P_{abn} (Fig. 4,b) leads to the same conclusions. The values P_{abn} tend to saturation if γ increases although $P_{abn} = 1$ is never attained (see also the dependencies below).

Consider now a more realistic case of α not equal to 2. Fig. 5 presents the dependences like in Fig. 4 but for α = 1.8.

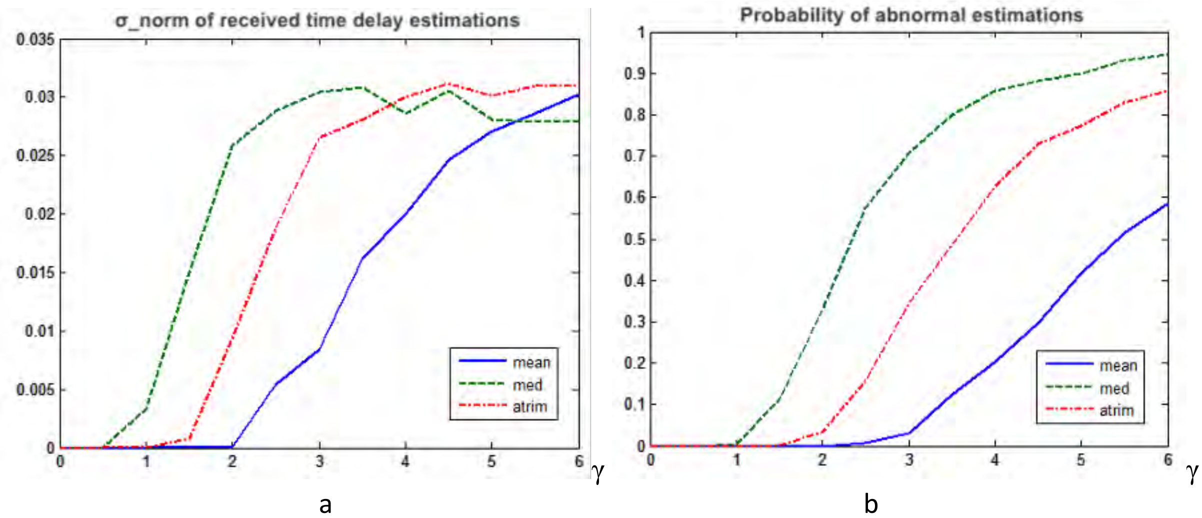


Figure 4: Dependences of σ_{norm} (a) and P_{abn} (b) on γ for the standard approach (mean) and two variants of RDFT-based approach exploiting median (med) and α-trimmed mean (atrim) estimators for α=2 (Gaussian noise)

As it is seen, in this case, the standard method occurs to be the worst. It produces the largest σ_{norm} and P_{abn} for the same small γ and reaches saturation for the smallest γ. Meanwhile, signal processing methods based on robust DFT forms produce sufficiently better results where the α-trimmed form is more efficient than the median form. The presented results allow concluding that it is worth considering different forms of RDFT-based processing for the considered application. It is also clear that more attention has to be paid to the analysis of dependencies of P_{abn} on γ since the reduction of P_{abn} is more important than decreasing normal estimate variance.

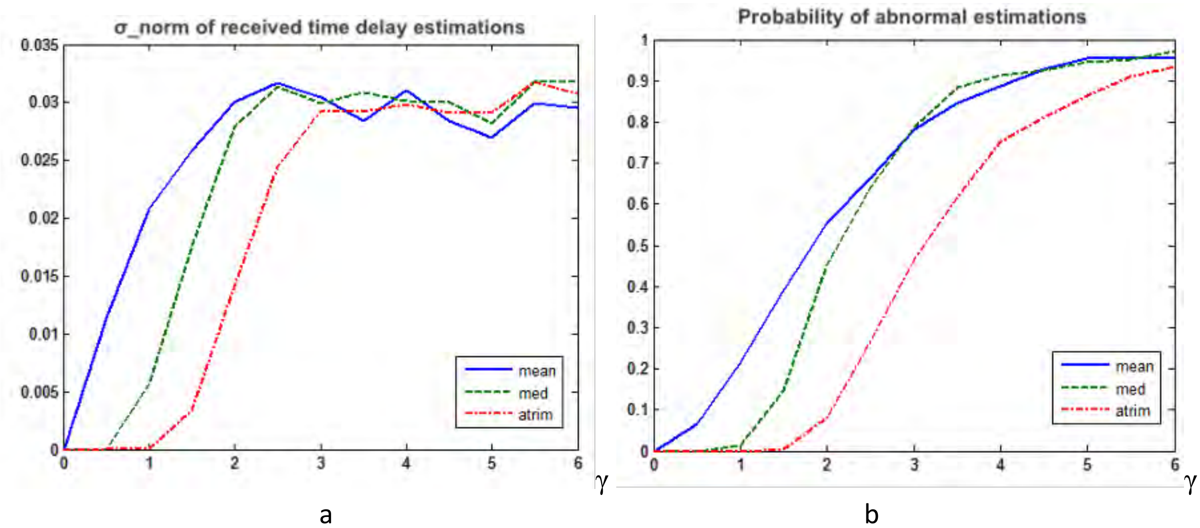


Figure 5: Dependences of σ_{norm} (a) and P_{abn} (b) on γ for the standard approach (mean) and two variants of RDFT-based approach exploiting median (med) and α-trimmed mean (atrim) estimators for α=1.8 (heavy-tailed noise)

So, let us present the results obtained for the same test signal but for the four versions of the RDFT-based processing. The data for α=2 (Gaussian noise) are given in Fig. 6. The results for the standard approach and the RDFT-based version for median and α-trimmed estimators are practically the same as in Fig. 4,b. But we are more interested in two other estimators. They produce practically the same results, which are better than for the median form but slightly worse than for the α-trimmed form of RDFT-based processing. All other tendencies are the same as earlier, i.e. the plots tend to saturation (P_{abn} close to unity) if γ increases.

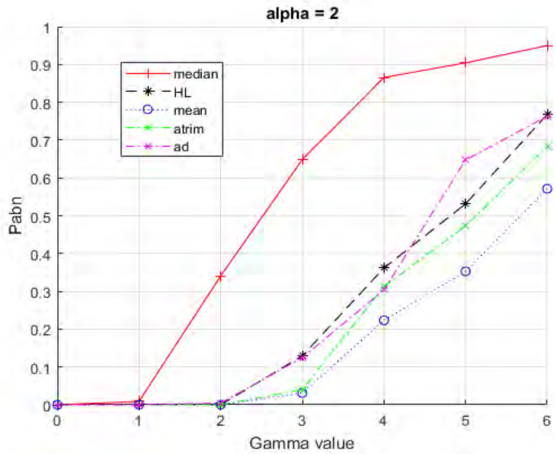
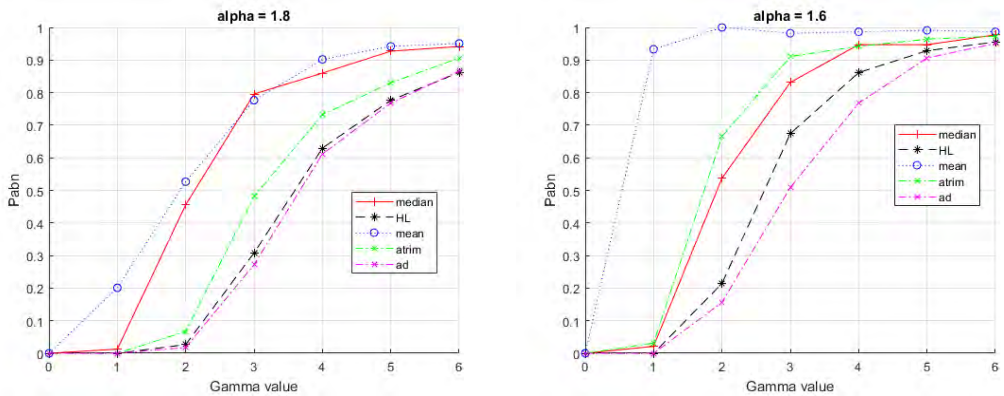


Figure 6: Dependences of P_{abn} on γ for the standard approach (mean) and four variants of RDFT-based approaches exploiting median (med), α-trimmed mean (atrim), Hodges-Lehman (H-L), and adaptive (ad) estimators for α=2 (Gaussian noise)

Let us look what happens if α is smaller than 2, i.e. if noise is impulsive. For α=1.8 (Fig. 7,a), the performance of the standard approach is the worst, and the median form of RDFT performs only slightly better. The processing based on other robust forms of DFT clearly outperforms the latter two methods where the best results are provided by the adaptive form of the RDFT.

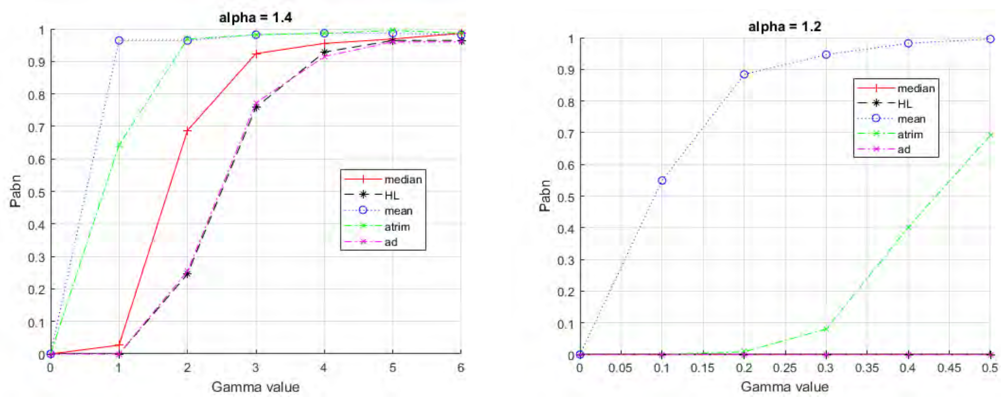
The situation is, in some sense, different for α=1.6 (Fig. 7,b). The standard approach produces the worst results as in the previous case, but the median form of RDFT starts to outperform the α-trimmed form. The reason is, probably, in higher robustness of the median form. As in the previous case, the best results are provided by the adaptive form of RDFT. Recall here that α=1.6 is considered to be typical for hydroacoustic and atmospheric noises.



a b

Figure 7: Dependences of P_{abn} on γ for the standard approach (mean) and four variants of RDFT-based approaches exploiting median (med), α-trimmed mean (atrim), Hodges-Lehman (H-L), and adaptive (ad) estimators for α=1.8 (a) and α=1.6 (b), heavy-tailed noise in both cases

Consider now the cases of heavier tail noises. Fig. 8 presents the plots for α=1.4 and α=1.2. For α=1.4 (Fig. 8,a), the standard approach fails for very small γ. The α-trimmed form of RDFT performs slightly better but its robustness is not high enough. Again, the adaptive and Hodges-Lehman forms of RDFT provide the best results. For α=1.2 (Fig. 8,b), the standard approach produces large P_{abn} for very small γ (note that the considered γ values are in this case smaller than in previous cases). The α-trimmed form of RDFT performs better. However, other forms produce even better results and there are no abnormal errors of TDE for all three forms for the considered range of γ variation. Thus, summarizing the obtained results, it is possible to state that the adaptive robust form of RDFT-base processing proposed in [18] provides the best or almost the best results for a wide range of noise impulsivity variation and, thus, it can be recommended for practical use.



a b

Figure 8: Dependences of P_{abn} on γ for the standard approach (mean) and four variants of RDFT-based approaches exploiting the median (med), α-trimmed mean (atrim), Hodges-Lehman (H-L), and adaptive (ad) estimators for α=1.4 (a) and α=1.2 (b), heavy-tailed noise in both cases

It is worth mentioning the drawbacks of the RDFT-based approach and the directions of its further development. The main drawback is that its computational efficiency is sufficiently worse compared to the standard approach where $\dot{S}_1(\omega) = FFT(x_1(t))$ and $\dot{S}_2(\omega) = FFT(x_2(t))$. RDFT does not allow using FFT, it also requires accomplishing a lot of sorting operations for data samples of a rather large size (if N=1024, then one needs 1024 sortings for data samples of the size N=1024 for real component data and the same for imaginary component data for the median and α-trimmed forms of RDFT; for more complex forms even more operations are needed). Thus, it is desired to find more computationally efficient approaches to signal processing. Concerning directions of further studies, it is possible to try using RDFT instead of FFT at the stage of obtaining the estimate of CCF from cross-spectrum. Besides, it is possible to apply RDFT at both stages of obtaining the estimate of cross-spectrum as well as obtaining the CCF estimate from cross-spectrum.

Keeping in mind the aforementioned drawbacks of the RDFT-based approach, consider two other possible approaches. Let us suppose that the main structure of the signal processing algorithm remains the same, i.e. the cross-spectrum estimate is obtained using to FFTs and the CCF estimate is derived using inverse FFT. Meanwhile, improvements are due to signal/mixture pre-processing (pre-filtering) before applying FFT.

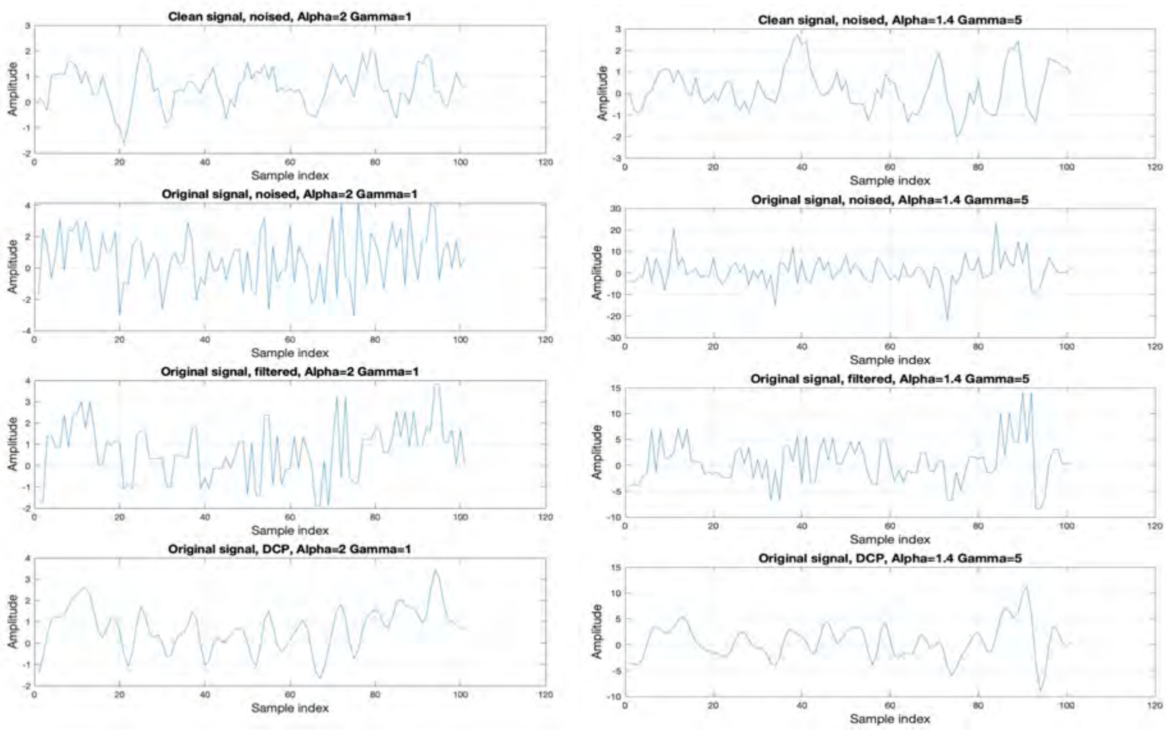
The goal of such pre-filtering is to remove non-Gaussian noise as efficiently as possible with preserving the informative WNL signal. The task of denoising a signal embedded in non-Gaussian noise is a standard one and quite many approaches based on scanning window nonlinear filtering have been already proposed and tested [25, 26]. They have been mainly designed to preserve sharp transitions in signals as, e.g., step or ramp edges. The main advantage of this group of techniques is that they are able to remove impulsive and mixed noise. Later, more advanced methods applicable to electrocardiogram and audio signal denoising have been proposed (let us mention [27, 28] to name a few). They are more based on orthogonal transforms as discrete cosine transform (DCT) and wavelets. Different combined approaches exploiting the impulse removal ability of scanning window nonlinear filters and the Gaussian noise removal ability of transform-based filters have been developed as well.

For the considered application, our desire [19] was not to develop new advanced methods but to exploit the positive features of the already existing filters. Our idea is that some one-dimensional robust scanning window filter removes "obvious" impulses whilst the DCT-based filter (with settings adjusted to residual noise) suppresses the additive noise that is quasi-Gaussian. Then, we have several tasks to be solved: 1) what robust filter and with what parameters to choose? 2) how to choose (set) parameters of the DCT-based filter? 3) what improvement (reduction) in processing efficiency can be gained due to such a pre-filtering?

First of all, let us give some examples of noisy signal pre-processing. Fig. 9,a shows the example of the WNL signal fragment for Gaussian noise (taken from [19]). Comparing the two upper plots, it is seen that additive white Gaussian noise which is quite intensive (γ=1) considerably masks the WNL signal. The center-weighted median filter (scanning window size is equal to 7, the center weight is equal to 3) partly removes noise (see the third plot in the column) but introduces specific distortions. In turn, the DCT-based filter (the fourth plot in the column) has smoothed the peaks but the information signal occurred to be smeared as well. This always happens in signal/image denoising that, alongside with positive effect of noise removal, the negative effect of detail smearing takes place. The main aspect concerns the degree of positive and negative effects and attaining the possible trade-off.

Fig. 9,b demonstrates the case of WNL signal corrupted by SαS noise with α=1.4. It is seen well that the noise absolutely masks the informative signal (pay attention that the scales for two upper plots differ sufficiently). Peak amplitude values in the second plot are by about ten times larger than in the first plot. The center-weighted median filter removes the most obvious impulses but introduces specific distortions. The DCT filter additionally removes noise. The final denoised signal (the lower plot) is still “far away” from the noise-free WNL signal (the upper plot), but there is much less noise than in the original noisy signal.

Recall that we are more interested in the indirect influence of the pre-filtering on TDE accuracy than in the traditional analysis of denoising efficiency. Hence, let us present some results. Two versions of DCT-based denoising have been tested. For the first version, the threshold Thr has been set as $2.7\sigma_{res}$ where σ_{res} is the estimate of residual noise standard deviation (according to traditional recommendation on threshold setting in DCT-based filtering) whilst, for the second version, the threshold has been set as Thr=$2\sigma_{res}$ to improve detail preservation ability.

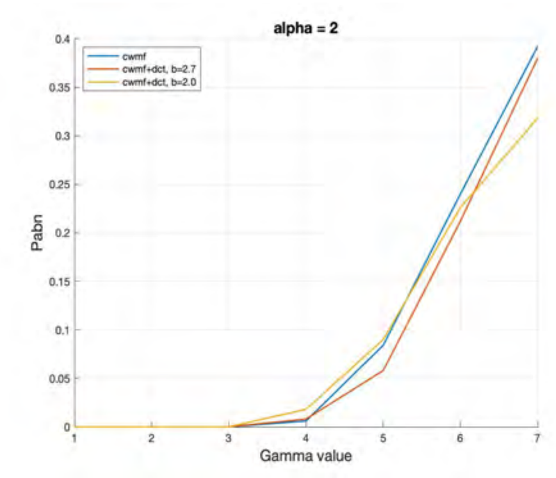


a b

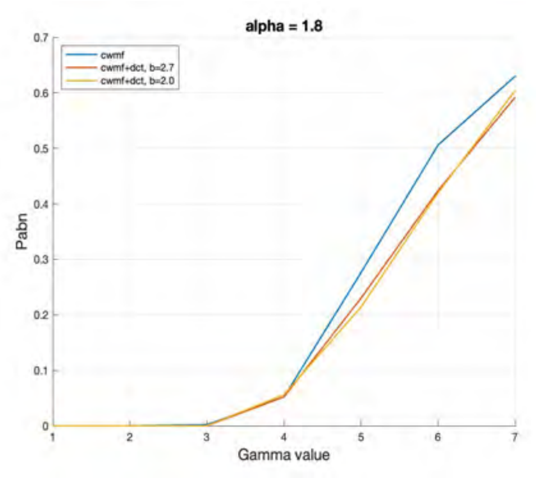
Figure 9: Illustrations of noise pre-filtering for α=2.0 (a) and α=1.4 (b)

Figure 10, a shows the results for the same signal/noise model as earlier for α=2.0. One can compare these plots to the plots in Fig. 4,b and see that pre-filtering leads to reduction of P_{abn}.

Among three considered versions off pre-filtering, the sequential application of the center weighted median filter and DCT with Thr=2.7σ_{res} produces, on average, the best results.



a



b

Figure 10: Dependences of P_{abn} on γ for three variants of data processing with pre-filtering by center-weighted median filter (cwmf) and two-stage denoising (cwmf+dct) for α=2.0 (a) and α=1.8 (b)

The plots for α=1.8 are represented in Fig. 10,b. They can be compared to the plots in Fig. 5,b. As one can see, e.g., for γ=4 or 5, the data processing based on signal pre-filtering provides sufficiently less P_{abn} than the standard procedure and less P_{abn} than data processing by the α-trimmed form of RDFT. Additional removal of residual noise by the DCT-based filter (Fig. 10,b) produces certain benefits compared to the use of only center weighted median filter, especially for large γ.

Fig. 11 gives the results for α=1.6 and α=1.4. They can be compared to the plots in Figures 7,b and 8,a. As one can see, the efficiency for the pre-filtering-based approach for α=1.6 is higher than for the best RDFT-based method (Fig. 7,b, consider, e.g., the data for γ=4). Post-filtering by the DCT-based denoiser produces a certain benefit compared to pre-processing by only the center-weighted median filter but this benefit (reduction of P_{abn}) is not large. Comparison of the plots in Fig. 11,b to the plots in Fig. 8,a show that the pre-filtering-based approach performs sufficiently better than the best RDFT-based approach (analyze the P_{abn} values for, e.g., γ=3). Additional denoising by the DCT-based filter produces certain improvements in accuracy although they are not large.

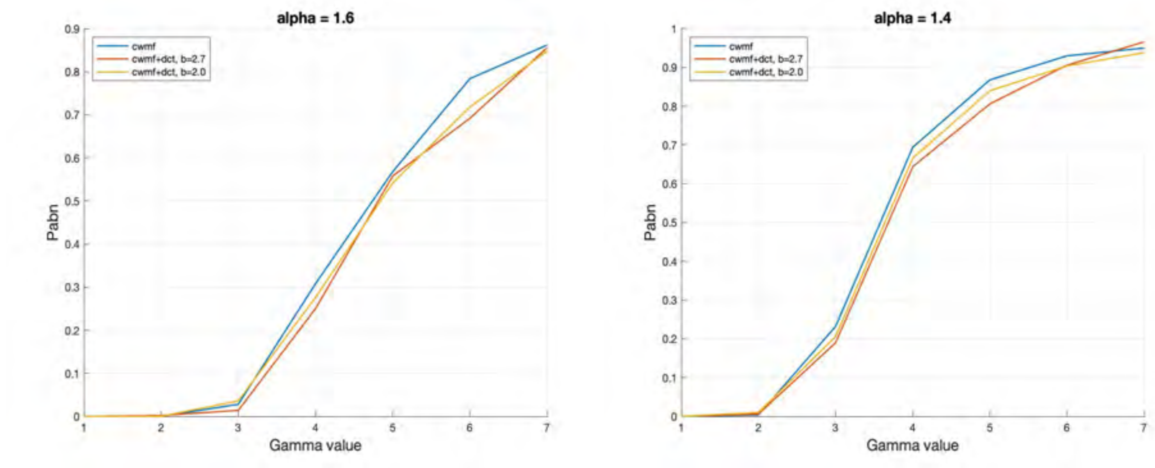
Among the three considered versions of pre-processing the mixture of signal and noise, the simplest version presuming the use of the center-weighted median filter is slightly less efficient in the sense of P_{abn} compared to the two-stage processing, but it requires considerably fewer computations.

Data processing by the center-weighted median filter is very fast while DCT-based denoising requires not only just DCT-based filtering (which is rather fast) but also blind estimation of residual noise standard deviation, which requires efforts comparable to filtering. Keeping this in mind, we can recommend using only the center-weighted median filter for signal pre-processing.

Consider now one more possible approach to TDE. Let us rewrite the expression (2) for CCF as

$$E_1 + E_2 - 2Y(\tau) = \int_{-\frac{T}{2}}^{\frac{T}{2}} ({x_1}^2(t) - 2x_1x_2(t+\tau) + {x_2}^2(t+\tau))dt \qquad (3)$$

where $E_1 = \int_{-T/2}^{T/2} {x_1}^2(t)\, dt$ and $E_2 = \int_{-T/2}^{T/2} {x_2}^2(t+\tau)\, dt$ are the energies of $x_1(t)$ and $x_2(t)$ in the first and second channels, respectively [20]. One can assume that E_1 and E_2 are almost constant if WNL signal and noises are stationary.



a b

Figure 11: Dependences of P_{abn} on γ for three variants of data processing with pre-filtering by center weighted median filter (cwmf) and two-stage denoising (cwmf+dct) for α=1.6 (a) and α=1.4 (b)

Hence, instead of searching for the global maximum of (2), it is possible to find the global minimum of Euclidian distance (3) [29] between the sampled received signals $x_1(t)$ and $x_2(t+\tau)$. In other words, for signals $x_1(i), i=1,\dots,N$ and $x_2(i+j), i=1,\dots,N$, one needs to calculate the similarity measure $S(j), i=-j_{max},\dots,j_{max}$, where $j_{max}\Delta t=\Delta\tau_{max}=L/C$.

Meanwhile, the Euclidian distance $S_E(j)=\Sigma(x_1(i)-x_2(i+j))^2$ (dealing with L^2 norm where the summation is carried out for all available i in a circular manner) is known to be not robust with respect to outliers. If outliers are possible, it is usually replaced by more robust similarity measures [29]. Thus, it was possible to try different other norms for the considered application with an attempt to find more powerful solutions [20].

In a general form, the similarity function for mutually shifted received mixtures can be written as

$$S_\beta(j)=\Sigma|x_1(i)-x_2(i+j)|^\beta, \tag{4}$$

where β denotes the power (that can be any positive value). Let us consider the cases of β equal to 0.5 1.0, and 1.5 under the assumption that some of these three values can be close to a quasi-optimal for given characteristics of a signal and noise. We consider β values smaller than 2 since our focus is on heavy tail noise.

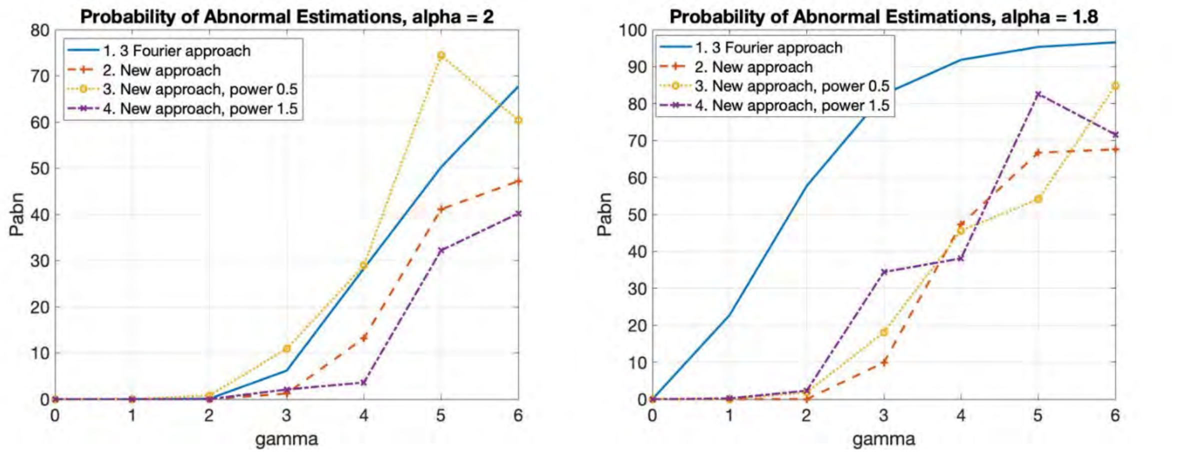
As in the previous cases, the variance of WNL signal has been fixed and equal to 1.0. Meanwhile, its spectral properties were slightly other than in previous cases. To get the WNL signal, AWGN has been passed through the low pass filter in such a way that WNL signal upper frequency occurred to be about three times smaller than the Nyquist frequency (again equal to 20 kHz). This was done to check does the method work well enough for the test signal, which is slightly different from the model data used earlier.

The dependences $P_{abn}(\gamma)$ (given in %) are represented in Fig. 12. Here the following notations are used: the Fourier approach means conventional approach and New approach relates to the use of (4), if β is equal to 0.5 or 1.5, it is marked at plots. It is seen that the approach (4) outperforms the standard approach if β is equal to 1.0 or 1.5 even for Gaussian noise (α=2.0, Fig. 12,a). If α=1.8, the standard method fails even for very small values of γ. The results for the method (4) are considerably better and they are approximately the same for all three considered values of β.

It is impossible to compare the plots in Fig. 12 to the previously analyzed dependences since the signal properties are other. Meanwhile, it is possible to compare the results from [20] to the plots presented above since the signal with the upper frequency 4 kHz was used in [20]. The comparison shows that the approach (4) for β=1 performs slightly worse than the approach based on pre-filtering. The same holds for α=1.8.

Consider now the plots for α=1.6 and α=1.4 given in Fig. 13, a and 13,b, respectively. Obviously, approach (4) outperforms the standard approach that fails even for small γ. It is also interesting that the use of β=0.5 or 1.0 produces better results than the use of β=1.5. This means that for noise with heavier tails one has to use a smaller β in (4).

The “fixed” option in the case of unknown α (which is within the considered range of its variation) could be to use β=1.0, but if one knows or can estimate α, an adaptive setting of β seems possible and reasonable. However, such studies have not been done yet and this is one possible direction for further studies. Besides, other similarity measures can be used and, maybe, some of them can be better (more efficient) than the already considered ones.



a b

Figure 12: Dependences of P_{abn} on γ for four variants of data processing for α=2.0 (a) and α=1.8 (b)

Comparison of data in [20] obtained for the signal with the upper frequency 4 kHz to the pre-filtering-based approach shows that the method (4) is slightly less efficient. Meanwhile, the approach (4) has an obvious advantage that deals with high computational efficiency since it is based on simple arithmetic operations (no sorting is needed compared to two previous approaches).

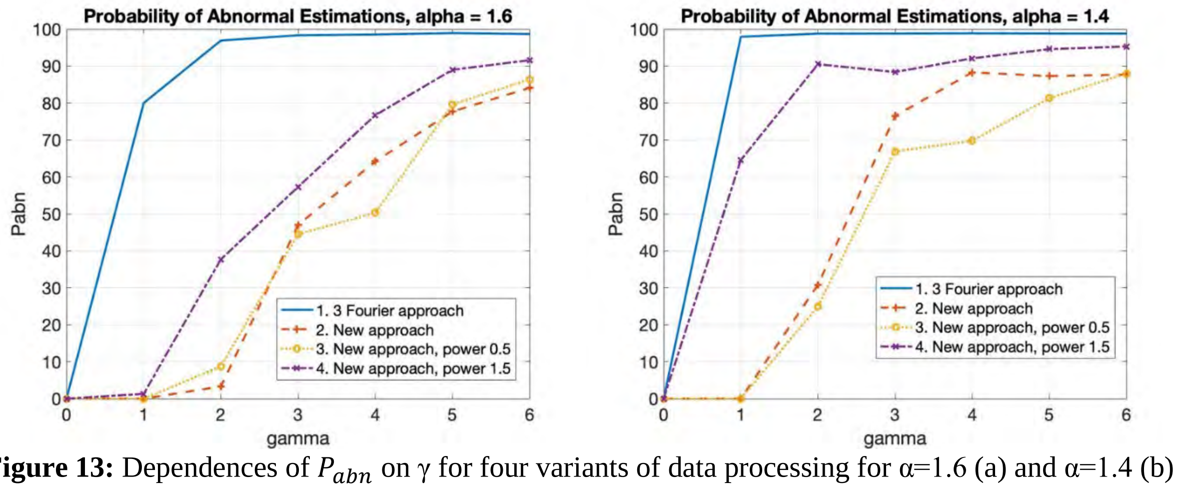


Figure 13: Dependences of P_{abn} on γ for four variants of data processing for α=1.6 (a) and α=1.4 (b)

4. Conclusions

The task of TDE is considered for the case of non-Gaussian environment, which is quite typical for several important applications. The restrictions typical for these applications are discussed. The signal and noise peculiarities are mentioned showing that the task is complicated even for the simplest configuration of receiving antenna having two sensors. If noise is non-Gaussian, performance of the standard approach (optimal for Gaussian noise) becomes worse radically and special means to cope with impulsive nature of the noise is needed.

Three approaches based on RDFT, signal pre-filtering and the use of robust similarity measures are presented. Simulation results that allow analyzing and comparing their performance are given and discussed. The advantages and drawbacks of these approaches are considered.

It is shown that the approach based on RDFT is less efficient than others in the sense of accuracy and it requires sufficiently more computations since Fast Fourier Transform algorithms cannot be employed in RDFT calculation. Two other approaches can be implemented easily and are very fast. The pre-filtering approach is able to produce a little bit better accuracy whilst the approach based on robust similarity measures is faster.

5. References

[1] G. P. Kousiopoulos, G. N. Papastavrou, D. Kampelopoulos, N. Karagiorgos, S. Nikolaidis, Comparison of Time Delay Estimation Methods Used for Fast Pipeline Leak Localization in High-Noise Environment, Technologies, 8(2): 27 (2020). doi:10.3390/technologies8020027.

[2] T. Padois, O. Doutres, F. Sgard, On the use of modified phase transform weighting functions for acoustic imaging with the generalized cross correlation, J. Acoust. Soc. Am., 145 (2019) 1546–1555. doi:0.1121/1.5094419.

[3] J. Benesty, J. Chen, Study and Design of Differential Microphone Arrays, Springer-Verlag, Berlin, 2013.

[4] H. He, J. Lu, L. Wu, X. Qiu, Time delay estimation via non-mutual information among multiple microphones, Applied Acoustics, 74(8) (2013) 1033–1036. doi:10.1016/j.apacoust.2013.02.001.

[5] A. Carrier, J.-L. Got, A maximum a posteriori probability time-delay estimation for seismic signals, Geophysical Journal International, 198(3) (2014) 1543–1553. doi:10.1093/gji/ggu218.

[6] M. M. Saad, C. J. Bleakley, T. Ballal, S. Dobson, High-Accuracy Reference-Free Ultrasonic Location Estimation, IEEE Transactions on Instrumentation and Measurement, 61(6) (2012) 1561–1570. doi:10.1109/TIM.2011.2181911.

[7] Yi Zhu, L. Liu, J. Zhang, Joint Angle and Delay Estimation for 2D Active Broadband MIMO-OFDM Systems, in: Proceedings of the IEEE Global Communications Conference, GLOBECOM ’13, IEEE, Atlanta, GA, 2013, pp. 3322–3327. doi:10.1109/GLOCOM.2013.6831581.

[8] H. Peyvandi1, M. Farrokhrooz, H. Roufarshbaf, S.-J. Park, SONAR Systems and Underwater Signal Processing: Classic and Modern Approaches, in: N. Kolev (Ed.), Sonar Systems, IntechOpen, London, 2011, pp. 9–35. doi:10.5772/17505.

[9] D. Dash, V. Jayaraman, Time delay estimation issues for target detection and transmitter identification in multistatic radars, Engineering Reports, 2(10) (2020). doi:10.1002/eng2.12236.

[10] H. Li, E. D. Gedikli, R. Lubbad, Exploring time-delay-based numerical differentiation using principal component analysis, Physica A: Statistical Mechanics and its Applications, 556 (2020) 1–20. doi:10.1016/j.physa.2020.124839.

[11] P. Marmaroli, X. Falourd, H. Lissek, A comparative study of time delay estimation techniques for road vehicle tracking, in: Proceedings of the Acoustics 2012 Nantes Conference 11th Congrès Français d'Acoustique ’2012, pp. 4135–4140. URL:http://www.conforg.fr/acoustics2012/cdrom/data/articles/000462.pdf.

[12] A. Mehrjouyan, A. Alfi, Robust adaptive unscented Kalman filter for bearings-only tracking in three dimensional case, Applied Ocean Research, 87 (2019), 223–232. doi: 10.1016/j.apor.2019.01.034.

[13] Ali Massoud, Direction of arrival estimation in passive sonar systems, Ph.D. thesis, Queen’s University Kingston, Ontario, Canada, 2012.

[14] A. Meynard, B. Torr´esani, Spectral analysis for nonstationary audio, 2018. URL: https://arxiv.org/pdf/1712.10252.pdf.

[15] M. Shao, C. L. Nikias, Signal processing with fractional lower order moments: stable processes and their applications, Proc. of IEEE, 81(7) (1993) 986–1010. doi:10.1109/5.231338.

[16] J. P. Nolan, Univariate Stable Distributions, Springer, Cham, 2020. doi:10.1007/978-3-030-52915-4.

[17] Z. Luo, P. Lu, G. Zhang, Locally optimal detector design in impulsive noise with unknown distribution, EURASIP Journal on Advances in Signal Processing, 34 (2018). doi:10.1186/s13634-018-0560-x.
[18] V. Oliinyk, V. Lukin, I. Djurovic, Time Delay Estimation for Noise-Like Signals Embedded in Non-Gaussian Noise Using Adaptive Robust DFT, in: Proceedings of the IEEE 7th Mediterranean Conference on Embedded Computing, MECO ’18, IEEE, Podgorica, Montenegro, 2018, pp. 267–270. doi:10.1109/MECO.2018.8406054.
[19] V. Oliinyk, V. Lukin, Time Delay Estimation for Noise-Like Signals Embedded in Non-Gaussian Noise Using Pre-filtering in Channels, in: Proceedings of the IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering. TCSET ’20, IEEE, Lviv, Ukraine, pp. 638–643. doi:10.1109/TCSET49122.2020.235510.
[20] I. Djurovic, V. Oliinyk, V. Lukin, A Fast and Efficient Method for Time Delay Estimation for the Wideband Signals in Non-gaussian Environment, in: M. Nechyporuk, V. Pavlikov, D. Kritskiy (Eds.), Integrated Computer Technologies in Mechanical Engineering, volume 188 of Lecture Notes in Networks and Systems, Springer, Cham, 2021, pp. 30–41. doi:10.1007/978-3-030-66717-7_3.
[21] I. Djurovic, A WD-RANSAC Instantaneous Frequency Estimator, IEEE Signal Processing Letters, 23(5) (2016) 757–761. doi:10.1109/LSP.2016.2551732.
[22] O. Roienko, V. Lukin, V. Oliinyk, I. Djurović, M. Simeunović, An Overview of the Adaptive Robust DFT and Its Applications, in: Technological Innovation in Engineering Research, volume 4, Book Publisher International, Hooghly, India, 2022, pp. 68–89. doi:10.9734/bpi/tier/v4/6314F.
[23] International Encyclopedia of Statistical Science, M. Lovric (Ed.), Springer, Berlin, 2011. doi:10.1007/978-3-642-04898-2_594.
[24] S. M. Vovk, General approach to building the methods of filtering based on the minimum duration principle, Radioelectronics and Communications Systems, 59 (2016) 281–292. doi:10.3103/S0735272716070013.
[25] J. Astola, P. Kuposmanen, Fundamentals of nonlinear signal processing, 2nd. ed., CRC Press, Boca Raton, USA, 2019. doi:10.1201/9781003067832.
[26] M. Mafi, H. Martin, M. Cabrerizo, J. Andrian, A. Barreto, M. Adjouadi, A comprehensive survey on impulse and Gaussian denoising filters for digital images Signal Processing, 157 (2019) 236–260. doi:10.1016/j.sigpro.2018.12.006.
[27] S. Chatterjee, R. S. Thakur, R. N. Yadav, L. Gupta, D. K. Raghuvanshi, Review of noise removal techniques in ECG signals, IET Signal Processing, 14(9) (2020) 569–590.
[28] A. Awad, Impulse noise reduction in audio signal through multi-stage technique, Engineering Science and Technology, 22(2) (2019) 629–636. doi:10.1016/j.jestch.2018.10.008.
[29] S. Bandyopadhyay, S. Saha, Unsupervised Classification: Similarity Measures, Classical and Metaheuristic Approaches, and Applications, Springer-Verlag, Berlin, 2013. doi:10.1007/978-3-642-32.