Identification of non-informative noise component in time-frequency representations. Application to vibration-based local damage detection.

Katarzyna SKOWRONEK[1], Tomasz BARSZCZ[2], Jérôme ANTONI[3]
Radosław ZIMROZ[1], Agnieszka WYŁOMAŃSKA[1]*

[1]Wroclaw University of Science and Technology, Wybrzeze Wyspianskiego 27
50-370 Wroclaw, Poland
[2]AGH University of Science and Technology, Al. Mickiewicza 30, 30-059 Krakow, Poland
[3]Univ Lyon, INSA-Lyon, LVA, EA677, F-69621 Villeurbanne, France
agnieszka.wylomanska@pwr.edu.pl

Abstract

In this research, we highlight the importance of the background noise properties in the context of the vibration-based local damage detection. In the case, when the background noise has Gaussian characteristics, the classical methods for local damage detection can be applied. However, for many real signals the assumption of Gaussian distribution of the noise is not satisfied and one may expect the large impulses that influence the noise characteristics. In that case the impulsiveness criteria fail. Since, most of the methods for local damage detection are based on the autocovariance function properly defined for signals with finite second moment, we indicate here the important role of the finite variance of the signal and present a new approach for the assessment of the probabilistic properties of the noise. The methodology is applied for the TF representation of the signals. The problem is illustrated for the simulated signals from exemplary non-Gaussian distribution.

1 Introduction

In many technical systems, the measurement of any physical variable is performed to acquire some important information about an object or a process (called SOI). Although measurement systems are very advanced, the informative components may be noisy because of the presence of other stronger sources. Even though the background noise is considered as a non-informative component, its properties can have a significant impact for further analysis of the signal. We share a novel perspective to the problem of local damage detection and propose to analyze the probabilistic properties of the noise component as a pre-processing step. The main attention is paid on the existence of finite variance of the noise distribution. We note that for signals with finite-variance distribution one can apply methods based on classical measures of dependence (like autocovatiance function) for SOI detection. On the other hand, if the distribution of the background noise has infinite variance, the classical techniques may be not efficient. It is worth highlighting that infinite variance of a given distribution implies the higher moments are also infinite. Thus, the mixture of the SOI and impulsive (non-Gaussian infinite-variance) noise excludes also impulsive criteria.

We highlight, the considered problem is much more general than the classical goodness-of-fit testing, i.e. testing if the underlying signal has a given theoretical distribution. In many cases, the identification of the noise distribution is not possible (as other components may disturb this information).

In our research, to assess the background noise properties we propose a simple technique that is applied for the time-frequency representation (here spectrogram) of the signal. The approach is based on the statistic called empirical cumulative fourth moment (ECFM) and the observation that it exhibits irregular chaotic behaviour for data from infinite-variance distribution while it stabilizes for data with finite variance. This specific behavior

*The work of KS, RZ and AW is supported by National Center of Science under Sheng2 project No. UMO-2021/40/Q/ST8/00024.

of ECFM is a starting point for the algorithm proposed in this research. In this extended abstract we present the results for the simulated signals from selected non-Gaussian distribution (t location-scale), however, the extended analysis 9also for real signals) one can find in [1, 2].

2 Methodology

The ECFM statistic is defined as follows

$$C(k) = \frac{1}{k}\sum_{m=1}^{k}(x_m - me(x))^4, \qquad k = 1,2,\cdots,n, \tag{1}$$

where $x_1, x_2, \cdots, x_n$ is the considered signal of independent identically distributed (i.i.d.) observations and $me(x)$ the corresponding sample median. As it was mentioned, the methods for local damage detection are mostly based on the analysis of the signals in TF domain. Thus, we propose to apply the presented procedure to the time-frequency representation (spectrogram) of a given signal. However, this algorithm can be also applied for identification of infinite-variance behavior for signals in any other domains, see [2]. We remind, the spectrogram $S(\cdot,\cdot)$ is a square of the absolute value of STFT

$$S(t,f) = |STFT(t,f)|^2, \;\; STFT(t,f) = \sum_{m=1}^{n} x_m w(t-m) e^{-i2\pi f \frac{m}{n}}, \tag{2}$$

where $w(\cdot)$ is a given window, $t \in T$ is time point and $f \in \mathscr{F}$ is frequency. The procedure for finite-variance assessment consists of the following steps. 1) First, we transform the signal $x_1, x_2, \cdots, x_n$ to TF domain (spectrogram). 2) Next, we select the frequencies corresponding to the noise component. We denote this set as $\tilde{\mathscr{F}}$. 3) For each $f \in \tilde{\mathscr{F}}$ we normalize the vector $S(\cdot, f)$. 4) Next, for each vector $S(\cdot, f)$ we calculate the ECFM statistic. 5) In the next step, for each $f \in \tilde{\mathscr{F}}$ we identify the segments of the ECFM statistic between the jumps. 6) For each $f \in \tilde{\mathscr{F}}$ we select the last ”long” segment of the ECFM statistic. The corresponding vector (segment) we denote as $D(\cdot, f)$. 7) For each $f \in \tilde{\mathscr{F}}$ we fit the straight line to the vector $D(\cdot, f)$ using the least squares method. The estimated value of the slope parameter for frequency f we denote as a_f. 8) Finally, we analyze the distribution of the estimated slopes along the selected frequencies to asses the background noise properties.

3 Results for simulated signals

We analyze the vector of independent observations from t location-scale $\mathscr{T}(\nu,\delta)$ distribution defined through the probability density function

$$f(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\delta\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)}\left[\frac{\nu+\frac{x^2}{\delta^2}}{\nu}\right]^{-\left(\frac{\nu+1}{2}\right)}, \qquad x \in R, \tag{3}$$

where $\Gamma(\cdot)$ is the gamma function, $\nu > 0$ is a shape parameter and $\delta > 0$ is a scale parameter. The variance for t-location scale distribution is finite for $\nu > 2$. Otherwise, it is infinite. However, when we analyze the data from $\mathscr{T}(\nu,\delta)$ distribution in TF representation (spectrogram), the finite variance we expect for $\nu > 4$ and for other cases the variance is infinite, see [1, 2] for more details.

In Fig. 1 (top panel) we demonstrate the exemplary simulated signals (of length 10000 observations) from $\mathscr{T}(\nu,\delta)$ distribution. We consider three vales of parameters responsible for heavy-tailed behavior, namely $\nu = 6$ (Fig. 1, left top panel), $\nu = 3$ (Fig. 1, middle top panel) and $\nu = 1.5$ (Fig. 1, right top panel). In each case we assume $\delta = 1$. Let us emphasize that $\nu = 6$ corresponds to the finite-variance case while for $\nu = 1.5$ we have infinite-variance distribution (in time and TF domain). In the middle top panel of Fig. 1 we present the intermediate case, i.e. for $\nu = 3$ the t location-scale distribution has finite variance in time domain while in TF domain it transfers to infinite-variance case. In the bottom panel of Fig. 1 we demonstrate the spectrograms of the simulated signals.

For each considered case, we simulated signals of length 10000 from $\mathscr{T}(\nu,\delta)$ distribution and applied the presented above procedure. Finally, we calculated the median of the obtained a_f values (estimated for all frequencies from spectrogram). In Fig. 2 we present the distribution of the medians of the slopes a_f calculated

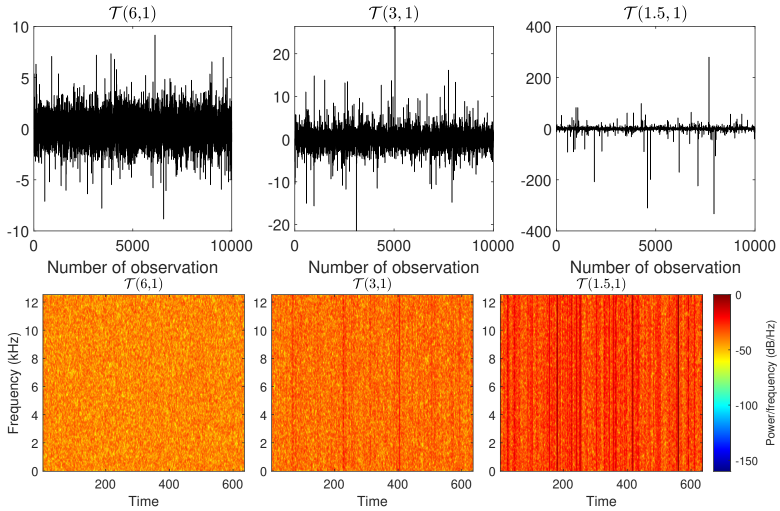

Figure 1: Examples of simulated signals from $\mathscr{T}(\nu,\delta)$ distribution (top panels) and the corresponding spectrograms (bottom panels).

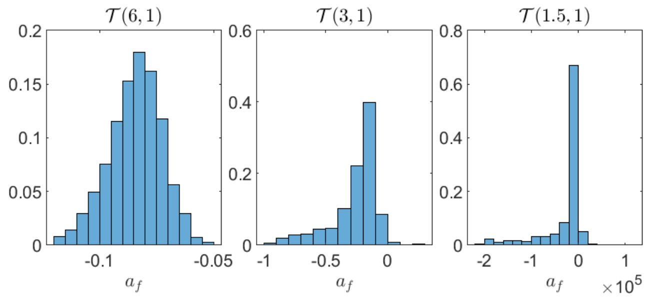

Figure 2: The distribution of the slopes a_f for simulated signals from $\mathscr{T}(\nu,\delta)$ distribution.

for 1000 simulated signals from the considered cases. One can conclude that there are clearly visible significant differences between the values of a_f estimated for distribution with finite- and infinite- variances. To underline the differences between the considered cases we also calculated the medians of the values presented in Fig. 2. For finite variance case (i.e. when $\nu = 6$) the median of a_f values is equal to e^{-2} while for the infinite-variance case (i.e. when $\nu = 1.5$) it is equal e^{10}. The differences we also see in the IQR statistic, considered as the dispersion measure. For the finite-variance case the IQR of the values presented in Fig. 2 is equal to e^{-4} while for the most extreme case (i.e. when $\nu = 1.5$) - it is equal to e^{15}.

4 Conclusions

In our research, we present another perspective for local damage detection and indicated that the properties of the background noise are significant for the selection of the appropriate tools. We propose a simple approach and demonstrate its efficiency for simulated signals from t location-scale distribution which may have finite- and infinite-variance. More deeper analysis, also for real signals, are presented in our ongoing paper [1].

References

[1] K. Skowronek, T. Barszcz, J. Antoni, R. Zimroz, A. Wyłomańska, *Assessment of background noise properties in time and time-frequency domains in the context of vibration-based local damage detection in real environment*, accepted to MSSP, 2023.

[2] K. Skowronek, R. Zimroz, A. Wyłomańska, *Testing for finite variance - analysis in time and time-frequency representations of the data. Applications to vibration signals from rotating machines*, in preparation, 2023.