



HAL
open science

Automatic Data-Driven Spectral Analysis Based on a Multi-Estimator Approach

Nadine Martin, Corinne Mailhes

► **To cite this version:**

Nadine Martin, Corinne Mailhes. Automatic Data-Driven Spectral Analysis Based on a Multi-Estimator Approach. *Signal Processing*, 2018, 146, pp.112-125. 10.1016/j.sigpro.2017.12.024 . hal-01677964

HAL Id: hal-01677964

<https://hal.science/hal-01677964>

Submitted on 8 Jan 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Automatic Data-Driven Spectral Analysis Based on a Multi-Estimator Approach

Nadine Martin^{a,b}, Corinne Mailhes^c

^a *Univ. Grenoble Alpes, GIPSA-Lab, F-38000 Grenoble, France*

^b *CNRS, GIPSA-Lab, F-38000 Grenoble, France*

^c *University of Toulouse, IRIT/ENSEEIH/TS, F-31071 Toulouse, France*

Abstract

In signal processing, spectral analysis is widely used but, whereas computing the power spectral density (PSD) by Fourier approaches is relatively easy, its analysis and reading are much more demanding especially for spectrally rich signals. This paper presents an original method which automatically picks out and estimates the relevant spectral structures of an unknown random stationary process, embedded in an unknown non-white Gaussian noise. First, a statistical hypothesis test is applied to each local maximum value of the estimated PSD to detect the potential spectral peaks of interest. Second, an original feature space is proposed for classifying and characterizing the detected structures. Then, one key idea of the proposed strategy is to use not only one spectral estimator but to combine the results of different ones, taking benefits of their good properties. Therefore the detection and classification steps are applied to different spectral estimations. A last fusion step outputs a complete attribute vector, including a confidence index, for each detected structure. Another key idea of this data-driven approach is that all parameters are automatically set up without *a priori* knowledge. This approach is fully adapted to the preventive maintenance of complex systems, as illustrated in the paper.

Keywords: Spectral analysis, Hypothesis test, Peak detection, Data-driven

1. Introduction

Literature on spectral analysis is abundant and lots of different estimators have been proposed (e.g., see [1], [2] and references therein). Depending on the application and on several parameters

Email addresses: nadine.martin@gipsa-lab.grenoble-inp.fr (Nadine Martin),
corinne.mailhes@enseeiht.fr (Corinne Mailhes)

such as signal to noise ratio, sample size or *a priori* spectral shape, a signal processing expert should be able to choose which estimator gives the best results in terms of spectral resolution, frequency estimation or modal detection. However, when no *a priori* information is available about the signal, or when a non-signal processing expert is willing to analyze some specific data, it is interesting to design an automatic spectral analysis tool, providing a complete description of the analyzed signal. Up to now, few methods have been published in this context. A fully automated spectral analysis was investigated in [3]. However, it was restricted to periodic signals focusing on the estimation of the signal period. In a more recent paper, Barbé and Van Moer [4] proposed an automatic detection and estimation procedure for harmonic components. A statistical analysis was considered to discriminate peaks for which amplitudes are higher or lower than a given threshold for any fixed frequency. The threshold was determined according to a minimum inter-group variance and a maximum group distance, taking the whole estimated spectrum into account. Thus, this method can be applied when the noise is white but not for a non-white one where the threshold should be determined individually for each tested peak. When the signal of interest is a sum of sinusoids embedded in white noise, the authors of [5] studied a local least square approach in the frequency domain. Tackling a wider signal class, the authors in [6] set up an automatic selection of the significant spectral components by using two different bandwidth resolutions. However, this method requires two signal measurements, which are not always available.

The objective of this paper is different. We propose a way of automatically analyzing stationary processes without *a priori* information. A new spectral analysis strategy is investigated, allowing specific spectral structures of a signal to be detected and characterized. The proposed strategy is not linked to any artificial intelligence consideration but is only related to the use of signal processing properties. In this context, a spectral structure has to be understood as the frequency signature of a sine wave, a narrow band signal, or a noise peak. The signals of interest are assumed to be composed of a sum of an unknown number of these structures, embedded in an unknown non-white noise. Several applications are demanding spectral analysis of such signals, including vibration, electrical, acoustic, seismic or radar signals. Indeed, these applications lead to process signals with a high number of samples, and having a high number of spectral structures embedded in non-white noise. It is important to highlight that this last assumption is essential in the context of this paper. Indeed, when signals are acquired by rotating machines such as machines in the

nacelle of wind turbines or manufacturing machines, methods restricted to the presence of white noise cannot be applied.

The main idea of the present paper is to combine and take advantage of the different properties of a set of spectral estimators based on the discrete Fourier transform. Combining estimation, detection and classification, the output of the proposed method is a set of spectral structures characterized by a set of appropriate attributes. These attributes include the standard ones as frequency, amplitude, local signal to noise ratio and spectral bandwidth. They are complemented by new ones providing a complete characterization of the spectral structures contained in a given signal. The resulting attribute vectors are referred to as identity cards and allow an easy classification of all the peaks of a power spectral density estimate. These identity cards can be of great interest in many domains and especially for the analysis of acoustic and vibration signals that are generally composed of a high number of spectral structures. Using an automated method combining the results of different spectral analyses of these signals is fruitful and laborsaving, e.g., for the condition monitoring of rotating machines.

A first method has been published in [7, 8]. In this paper we propose a more efficient one in which each step, the detection, the feature space mapping, the classification and the final fusion, has been improved. Furthermore, the method has been implemented on real-time and on-board devices within the frame of two European projects, one devoted to renewable energy in wind turbine farms (<http://www.gipsa-lab.fr/projet/KASTRION/>) and a second one devoted to manufacturing equipment in a coated paper mill (<http://www.supreme-fof.eu>). The paper is organized as follows. Section 2 formulates the problem addressed in this work and provides details about the spectral structures that are considered in this study. Section 3 studies a new spectral peak detection method that is applied on each estimated power spectral density (PSD), through a local hypothesis testing. This hypothesis test is defined in the frequency domain, taking the probability distribution of the estimated noise PSD into account. Section 4 defines the feature space that is used for the classification of the different spectral structures. Section 5 presents a way to choose spectral estimators with a specific example for a sine wave search. The fusion rule adopted to combine the different spectral analysis results is presented in Section 6. This rule leads to the identification and classification of different spectral structures, namely sine waves, narrow band signals and noise peaks. The proposed method is evaluated using simulated and real-world signals in Section 7. In

particular, we show how a real-world signal corresponding to the acoustic noise of a heat pump can be analyzed. Conclusions and perspectives are finally reported in Section 8.

2. Problem statement

2.1. Assumptions

Let us consider a discrete stationary random signal $x[k]$, $k = 0, \dots, K - 1$ defined as the sum of a signal of interest $s[k]$ and a noise $n[k]$,

$$x[k] = s[k] + n[k], k = 0, \dots, K - 1, \quad (1)$$

where K is the sample size. The signal of interest $s[k]$ is a stationary random process, whose PSD is denoted by $\gamma_s(\nu)$, ν being the frequency variable. We assume that the PSD of the signal $s[k]$ can be decomposed as the sum of elementary PSDs referred to as spectral structures. When the observed signal is composed of sine waves or very narrow band signals embedded in noise, a spectral structure typically corresponds to a more or less wide peak in the PSD of this signal. The number of spectral structures in $s[k]$ can be very large. For instance, a vibration signal is frequently composed of more than hundreds of peaks. The additive noise $n[k]$ is assumed to be a zero mean stationary Gaussian noise, independent of the signal of interest, with a PSD denoted by $\gamma_n(\nu)$. The additive noise may be white or not and the number of spectral structures contained in the signal of interest $x[k]$ is unknown.

2.2. Context

The problem addressed in this work consists of analyzing the signal $x[k]$ in order to detect, characterize and classify any spectral structure present in $s[k]$. Motivated by the assumption of a large sample size K , e.g. $K \geq 1000$, Fourier-based estimators will be considered as an appropriate tool for spectral analysis. Fourier estimators are well-known to be robust approaches that are daily used in many domains such as vibration analysis. However, understanding the rich spectral content of vibration signals is not so obvious and needs the help of human interaction, which is precisely the problem addressed in this paper.

In the particular case of a single sine wave embedded in white Gaussian noise, the maximum likelihood estimator of the sine wave frequency is known to be the frequency where the periodogram

reaches its maximum [9]. For multiple sine waves in white Gaussian noise, the maxima of the periodogram provide the maximum likelihood estimators of its sine frequencies as long as the sine waves are well resolved in frequency. However, in the general case of a non-white noise $n[k]$, there is no simple solution for estimating the frequencies contained in $s[k]$ using the maximum likelihood principle.

Therefore, in order to detect and characterize an unknown number of spectral structures embedded in a non-white additive noise, we propose a specific peak detection in the frequency domain. Moreover, to be robust against spectral estimation problems, the continuous PSD is estimated through a set of estimators coping with several compromises. Considering a non-white additive noise is challenging in such an automatic spectral analysis if we want to ensure that the proposed strategy has the same performance for any noise spectral shape.

3. Spectral structure detection

Given $\hat{\gamma}_x(\nu)$ an estimation of the continuous PSD $\gamma_x(\nu)$ of the observed signal $x[k]$, the set of the frequencies at which $\hat{\gamma}_x(\nu)$ is locally maximum is denoted as $\mathcal{M} = \left\{ \max_{\nu} \hat{\gamma}_x(\nu) \right\}$. The detection problem addressed in this Section consists of designing a binary test for detecting the relevant spectral structures in $\hat{\gamma}_x(\nu)$. More precisely, for a given frequency ν belonging to \mathcal{M} , we consider the two following hypotheses

$$H_0 : \gamma_x(\nu) = \gamma_n(\nu), \quad H_1 : \gamma_x(\nu) = \gamma_s(\nu) + \gamma_n(\nu), \quad (2)$$

and a test statistic $T(\nu)$ as a function of $\hat{\gamma}_x(\nu)$, such as

$$T(\nu) \underset{H_1}{\overset{H_0}{\leq}} \mu_{\text{PFA}}, \quad (3)$$

where the threshold μ_{PFA} can be adjusted from the probability of false alarm (PFA) of the test

$$\text{PFA} = \Pr [T(\nu) \geq \mu_{\text{PFA}} \mid H_0], \quad (4)$$

and where $\Pr(A \mid H_0)$ denotes the probability of the event A under hypothesis H_0 . This test will not be applied to all frequencies ν but only to frequencies ν at the local maxima of $\hat{\gamma}_x(\nu)$. The application of test (3) will allow us to keep only maxima in \mathcal{M} under hypothesis H_1 which then correspond to the so-called spectral structures of the analyzed signal.

3.1. Distribution of Fourier-based estimators under H_0

In order to derive the distribution of the test statistic $T(\nu)$ which should be a function of $\hat{\gamma}_x(\nu)$, we propose to study the distribution of the spectral estimators under hypothesis H_0 . In our context, the PSD estimators are Fourier-based with different temporal windows and overlapping. The additive noise is white or not.

Johnson and Long [10] derived the probability density function of the Welch periodogram for overlapping temporal windows of a Gaussian time series. However, this result is difficult to be generalized to other PSD estimators and the distribution of a test statistic resulting from this distribution is computationally expensive. In order to obtain a more general solution, one possibility is to approximate the distribution of the estimated PSD. Different approximations based on the chi-square distribution have been investigated in [1, 10, 11, 12] and more recently in [13]. In this paper, we consider a more general approximation that has to be valid for any Fourier-based PSD estimator. Let us introduce

$$U = a \frac{\hat{\gamma}_x(\nu)}{\gamma_n(\nu)}, \quad (5)$$

where a is a real scalar which will be fixed in what follows. According to the previous results in the literature, the distribution of U is approximated by a chi-square one with a number of degrees of freedom denoted by r . Therefore,

$$E[U] = r \quad \text{and} \quad \text{var}[U] = 2r. \quad (6)$$

Being under hypothesis H_0 and assuming that the noise estimation is unbiased, $E[U] = r$ in (6) gives $a = r$. Using this in the variance expression of (6) leads to

$$r = \frac{2}{\text{varn}[\hat{\gamma}_x(\nu)]} \quad \text{with} \quad \text{varn}[\hat{\gamma}_x(\nu)] = \frac{\text{var}[\hat{\gamma}_x(\nu)]}{\left(E[\hat{\gamma}_x(\nu)^2]\right)}, \quad (7)$$

where $\text{varn}[\hat{\gamma}_x(\nu)]$ is the normalized variance of the spectral estimator. This normalized variance depends only on the chosen PSD estimator and not on the signal $x[k]$ itself [1]. This property is interesting and allows the *a priori* computation of the degree of freedom r only from the variance of the spectrum estimator. Appendix A recalls the equations for the well-known Welch and Blackman-Tuckey estimators, whatever the type and size of the windows and whatever the overlapping. Note that, in the case of a Welch estimator, $r = 2N_{\text{seg}}$ for N_{seg} segments of signal without any overlapping and whatever the window, which gives a lowest value $r = 2$ obtained with one segment.

The proposed approximation of the distribution of $U = r \frac{\hat{\gamma}_x(\nu)}{\gamma_n(\nu)}$ is illustrated Fig. 1 by comparing the probability density function (PDF) of a chi-square variable to the estimated PDF of U for a Blackman-Tukey estimator, a Blackman window and a non-white ARMA Gaussian noise whose parameters are $\{1, -1.3552, 1.228, -0.662, 0.25\}$ for the AR parameters and $\{1, -0.9, 0.81\}$ for the MA ones. This example is a kind of worst-case example since the Blackman-Tukey estimator may lead to an estimated PSD with negative values, which is very penalizing when fitting with a chi-square distribution.

3.2. Noise PSD estimation

In order to determine the distribution of $\hat{\gamma}_x(\nu)$ using (5), we need to know the noise PSD $\gamma_n(\nu)$. We propose to estimate it thanks to a median filtering \mathcal{F} of $\hat{\gamma}_x(\nu)$ denoted $\tilde{\gamma}_n(\nu)$,

$$\tilde{\gamma}_n(\nu) = \mathcal{F}[\hat{\gamma}_x(\nu)]. \quad (8)$$

The median filter is applied to the estimated PSD, i.e., in the frequency domain. The median filter length has to be linked to the equivalent spectral window $W(\nu)$, induced by any Fourier-based spectral estimator and such that

$$E[\hat{\gamma}_x(\nu)] = \gamma_x(\nu) * W(\nu), \quad (9)$$

where $*$ denotes convolution. We propose to adjust the median filter length proportionally to the length of the spectral bandwidth of $W(\nu)$, denoted as B_W and defined as the main lobe width between its first two zeros. On the one hand, the median filter length has to be high enough around frequencies close to sine waves, in order to mitigate the effect of the signal peaks. On the other hand, a too large value of the median filter length fails to track the spectrum variations of the non-white noise. A compromise has to be found between $2B_W$ and $8B_W$; a value of $4B_W$ has been chosen in the examples of Section 7. In practice, the median filter length corresponds to a number of frequency samples which is denoted by M_n in what follows. Note finally that more details about how estimating the noise PSD from the noisy signal PSD can be found in [14], including methods based on median filters or other non-linear filters such as morphological or percentile filters.

3.3. A p -value approach for H_0 rejection

Considering the previous results, we propose to define the test statistic $T(\nu)$ in (3) as

$$T(\nu) = \frac{\hat{\gamma}_x(\nu)}{\tilde{\gamma}_n(\nu)} = r \frac{\hat{\gamma}_x(\nu)}{\gamma_n(\nu)} \left[r \frac{\tilde{\gamma}_n(\nu)}{\gamma_n(\nu)} \right]^{-1}. \quad (10)$$

Under hypothesis H_0 , $U = r \frac{\hat{\gamma}_x(\nu)}{\hat{\gamma}_n(\nu)}$ is considered to be distributed as chi-square with r degrees of freedom as shown in Section 3.1. Given that $\tilde{\gamma}_n(\nu)$ is the median filter of $\hat{\gamma}_x(\nu)$, the distribution of $r \frac{\tilde{\gamma}_n(\nu)}{\hat{\gamma}_n(\nu)}$ can be derived as an order statistics [15, p. 372], however the expression does not lend itself to simple analytic processing. Since noise PSD is estimated thanks to a median filter, we will rather consider $rM_n \frac{\tilde{\gamma}_n(\nu)}{\hat{\gamma}_n(\nu)}$ as the output of a filter with a finite impulse response of length M_n , excited by a chi-square distributed input, U . As a consequence, $T(\nu)$ is viewed as the ratio of two normalized correlated chi-squared distributed variables

$$T(\nu) = \frac{U/r}{rM_n \frac{\tilde{\gamma}_n(\nu)}{\hat{\gamma}_n(\nu)} / rM_n}. \quad (11)$$

The numerator and denominator of this ratio are necessarily correlated. But if we consider that the noise PSD is estimated by a nonlinear filter as described in Section 3.2 and that the test statistic $T(\nu)$ is applied at the peak maxima only, we do the assumption that this correlation can be neglected. Then the distribution of $T(\nu)$ can be approximated by a Fisher-Snedecor one with parameters r and rM_n , denoted by F_{r,rM_n} [15]. This approximation has been validated via several Monte Carlo simulations in the case of white and non-white noise, with different PSD estimators and different windows, with or without time segment overlapping. As an example, Fig. 2 compares the estimated probability density function of the test statistic $T(\nu)$ with its approximated F distribution in the case of a non-white ARMA Gaussian noise, a Blackman-Tukey estimator with a Blackman window and a median filter for the noise PSD estimation.

Under hypothesis H_1 , the distribution of $T(\nu)$ cannot be considered since there is no information about the signal of interest. We propose then to use the p -value of the test (3) [16]. The p -value is defined as the probability of having a value of $T(\nu)$ equal or more extreme than the one observed considering that H_0 is true. It gives a measure of the strength of evidence against H_0 . The lower the p -value, the more unlikely H_0 is, and at some point of low probability, H_0 should be preferably rejected.

Therefore, the p -value denoted as p_0 is computed at each frequency ν_{\max} belonging to \mathcal{M}

$$p_0 = \Pr [T(\nu) \geq T(\nu_{\max}) \mid H_0] = \int_{T(\nu_{\max})}^{+\infty} F_{r,rM_n}(t) dt, \quad (12)$$

where $F_{r,rM_n}(t)$ denotes the PDF of the Fisher-Snedecor distribution with parameters r and rM_n . If this p -value is higher than a given threshold corresponding to a PFA (4), we consider that there is no presumption against hypothesis H_0 and that the spectral structure around the frequency

ν_{\max} is a noise peak. All other spectral structures corresponding to p -values lower than the chosen threshold are detected as potential spectral structures of interest in the observed signal $\hat{\gamma}_x(\nu)$.

This way, the test output is a set of P peaks of amplitude γ_{\max_i} for which the corresponding p -values denoted as p_{0_i} give an idea of how the peaks contradict the hypothesis H_0 suspecting the presence of a signal of interest at frequency ν_{\max_i} . This set of peaks is written as

$$\{(\gamma_{\max_i}, \nu_{\max_i}, p_{0_i}), i = 1, \dots, P\}. \quad (13)$$

3.4. Spectral structure identification

Each triplet $(\gamma_{\max_i}, \nu_{\max_i}, p_{0_i})$ corresponds to a part of the estimated PSD $\hat{\gamma}_x(\nu)$ that we will refer to as a ‘‘spectral structure’’, denoted by γ_i and defined accurately in this Section. This quantity corresponds to one part of interest of the spectrum of the signal.

For each triplet, a local noise variance σ_i^2 is computed as

$$\sigma_i^2 = \frac{1}{\nu_{\min_{i+1}} - \nu_{\min_i}} \int_{\nu_{\min_i}}^{\nu_{\min_{i+1}}} \tilde{\gamma}_n(\nu) d\nu, \quad (14)$$

with $\tilde{\gamma}_n(\nu)$ given by (8) and ν_{\min_i} (resp. $\nu_{\min_{i+1}}$) the argument of the preceding (resp. succeeding) minimum of γ_{\max_i} (see Fig. 3). Therefore, a spectral structure γ_i associated to $(\gamma_{\max_i}, \nu_{\max_i}, p_{0_i})$ is defined by

$$\gamma_i = \{\hat{\gamma}_x(\nu); \nu \in [\nu_i^-, \nu_i^+]\}, \quad (15)$$

with

$$\hat{\gamma}_x(\nu_i^-) = \hat{\gamma}_x(\nu_i^+) = \sigma_i^2 \text{ with } [\nu_i^-, \nu_i^+] \subset [\nu_{\min_i}, \nu_{\min_{i+1}}]. \quad (16)$$

In order to accurately identify each spectral structure, γ_i is compared with the equivalent spectral window $W(\nu)$ defined in (9). The main lobe of $W(\nu)$ defined between the two first zeros and of amplitude one, denoted as $Q(\nu)$, is sampled at a higher rate than the rate of the spectrum $\hat{\gamma}_x(\nu)$. A maximum likelihood approach could have been considered in order to estimate the best position in frequency of $Q(\nu)$ for each γ_i . However, it has been shown to induce a higher computation time without any performance improvement compared to a simple least square minimization between γ_i and shifted versions $Q_S(\nu)$ of $Q(\nu)$. Therefore, this last solution has been preferred. The algorithm is initialized by a first version $Q_0(\nu)$ centered at the maximum frequency

ν_{\max_i} of γ_i and adjusted to its amplitude γ_{\max_i} such that

$$Q_0(\nu) = \gamma_{\max_i} Q(\nu - \nu_{\max_i}). \quad (17)$$

Then this adjusted spectral window will be shifted along the frequency axis, on the increasing or decreasing side, depending on the value of

$$\varepsilon = \text{sign}(\widehat{\gamma}_x(\nu_{\max_i} + \Delta\nu) - \widehat{\gamma}_x(\nu_{\max_i} - \Delta\nu)), \quad (18)$$

with $\Delta\nu$ the sampling period of the spectrum $\widehat{\gamma}_x(\nu)$. If $\varepsilon = 0$, it is a particular and nice case where the peak fits well with the spectral window $Q_0(\nu)$ and will be assigned later on to a sine wave. Otherwise, new versions $Q_S(\nu)$ are computed with a shift step equal to $\varepsilon\Delta\nu/c$, $\Delta\nu/c$ being the sampling period of $Q(\nu)$. Due to the over-sampling rate of $Q_S(\nu)$, c being at least equal to 10, the shift is small compared to the samples of γ_i and the total amplitude of the shift will be limited to one frequency interval $\Delta\nu$. In addition, at each step, the amplitude of $Q_S(\nu)$ is adjusted such that the constraint $Q_S(\nu_{\max_i}) = \gamma_{\max_i}$ is always satisfied.

At each step is computed a normalized quadratic error between γ_i and $Q_S(\nu)$ and only for their common sampling points,

$$e(i) = \frac{1}{\gamma_{\max_i}^2} \sum_{j=j_1}^{j_2} (\widehat{\gamma}_x(\nu_j) - Q_S(\nu_j))^2, \quad (19)$$

with $[\nu_{j_1}, \nu_{j_2}] \subset [\nu_i^-, \nu_i^+]$. The choice of j_1 and j_2 values determines a specific part of the peak and then a specific error. The total error denoted by $e_{\text{tot}}(i)$ takes the whole peak γ_i into account and is used in the least square minimization between γ_i and shifted versions $Q_S(\nu)$ of $Q(\nu)$. This spectral adjustment defined as a matching of the whole peak with the spectral window allows the determination of an adjusted central frequency $\widetilde{\nu}_{\max_i}$ and of an adjusted maximum $\widetilde{\gamma}_{\max_i}$ that will be used in Section 6.

In addition to considering the whole peak which may be distorted by secondary lobes of neighboring peaks, j_1 and j_2 can also be chosen such that (19) computes the -3 dB error denoted $e_{-3\text{dB}}(i)$. This last error will be used in Section 4.2 for defining a structure feature.

4. Spectral structure classification

4.1. Spectral structure class definition

The detection step provides a set of P spectral structures (15) related to the signal of interest. Each of these structures is characterized by a p -value providing a quantitative strength of evidence

against hypothesis H_0 . These detected spectral structures can correspond to sine waves but also to other signal structure with a spectral width broader than the equivalent spectral window. Therefore, in a first approach, three different classes are defined:

- **Sine Wave class (SW)**. This SW class is composed of all spectral structures corresponding to sine waves which are characterized by their frequency ν_s and amplitude A , such that the signal of interest (1) is of the form

$$s[k] = A \cos(2\pi\nu_s k). \quad (20)$$

- **Narrow Band signal class (NB)**. This NB class corresponds to signals having a spectral structure whose bandwidth is between the equivalent spectral window bandwidth B_W of the estimated PSD and the median filter length of the noise PSD estimation. Any wider band signal will be involved in the noise PSD estimation and is out of concern of this paper. This NB class includes slowly non-stationary signals of the form

$$s[k] = A[k] \cos(\phi[k]), \quad (21)$$

where $\phi[k]$ is a slow-varying phase function around a steering frequency and $A[k]$ is a band-limited signal whose maximal frequency is much smaller than this steering frequency [17]. Other stationary signals can be found in this class, such as a random frequency fluctuation over a small frequency band (e.g. a laser noise).

- **Noise peak class (NP)**. This NP class acts as a reject class for peaks departing too much from the equivalent spectral window of the corresponding PSD estimator.

4.2. Spectral structure features

After having defined the three classes to which a spectral structure can belong, we need to define discriminant features to handle the classification of all the detected structures. We propose a list of two features: one dealing with the height of the peak featured by a signal to noise ratio value, the other one with its shape got after a spectral adjustment. The height can be a relevant criterion for noise peak classification while the shape can help to distinguish SW peaks from the others.

Thus the first feature, the height of each spectral structure, is characterized by a local signal to noise ratio denoted as SNR_ℓ and associated to a SW peak,

$$\text{SNR}_\ell = 10 \log \left(\frac{A_i^2}{4\sigma_i^2} \right). \quad (22)$$

The parameter σ_i^2 is defined in (16) and A_i is the sinusoidal amplitude corresponding to the maximum value of the spectral peak γ_{\max_i} . We have chosen this feature SNR_ℓ since it is closely linked to the test statistic (10), as shown in appendix B, by the following relation

$$\text{SNR}_\ell = 10 \log [B_e T_s T(\nu)], \quad (23)$$

with $T_s = 1/F_s$ the sampling period and B_e the noise equivalent bandwidth of the spectral window:

$$B_e = \frac{\int_{-F_s/2}^{+F_s/2} W(\nu) d\nu}{W(0)}. \quad (24)$$

Through (23), the first feature SNR_ℓ is related to the p -value of each detected peak: the higher the SNR_ℓ value, the lower the p -value, the more likely the peak is a signal of interest, either in the SW or NB class. Therefore, by definition SNR_ℓ will be able to separate SW or NB peaks from the other NP peaks.

The second feature, the shape of a spectral structure is deduced from the quadratic spectral error (19) computed at -3dB . If it is low, the spectral structure will belong to SW class, otherwise, it will be assigned to NB or NP class.

Monte Carlo simulations have been performed in order to evaluate the interest of these two features ($\text{SNR}_\ell, e_{-3\text{dB}}$) for classifying the spectral structures into the three classes SW, NB and NP. Two simulated stationary signals of $K = 10,000$ samples have been generated over 50 runs.

- The first signal is a deterministic SW with a unitary amplitude and frequency ν_s .
- The second one is a random NB signal simulated as a sine function with a unitary amplitude and a varying frequency equal to ν_s plus some Gaussian noise with a low variance σ_n^2 .

For simulations, we choose $\sigma_n^2 = 25 \cdot 10^{-6}$ and $\nu_s = 0.11$ in normalized frequencies.

Both signals are embedded in an additive white Gaussian noise as in (1). Different signal to noise ratios have been considered from 30 dB to -20 dB. PSD estimation is done using a Welch estimator

[2] on the whole signal (one segment) with a Blackman window. According to Section 3, the noise spectrum is estimated and peaks are detected following the p -value approach detailed in Section 3.3 with $\text{PFA} = 10^{-3}$.

Fig. 4 presents the results of these Monte Carlo simulations for signal to noise ratios equal to 20 dB, 0 dB, -5 dB and -15 dB. It is clear that the three clusters of SW peaks (in red), NB peaks (in green) and NP peaks (in black) are well-separated in the proposed feature space $(\text{SNR}_\ell, e_{-3\text{dB}})$, as long as the signal to noise ratio is equal or above -5 dB. This property is exploited for estimating the class boundaries in the feature space $(\text{SNR}_\ell, e_{-3\text{dB}})$.

4.3. Feature space mapping

To perform peak classification in the proposed feature space $(\text{SNR}_\ell, e_{-3\text{dB}})$, and thanks to the properties of these features highlighted in the previous Section, the class boundaries will be defined from histograms computed from the previous Monte Carlo simulations. Following figures show the results using a Welch PSD estimation with Hanning window. The same reasoning is available for other estimators. Fig. 5 is the histogram of SNR_ℓ cumulating Monte Carlo simulations of SW peaks from 30 dB to -10 dB and Monte Carlo simulations of NB peaks from 30 dB to 5 dB. The threshold T_+ delimits the NP peaks from the others SW or NB without doubt. For lower signal to noise ratio, the boundaries could not be so obvious so we decided to define doubt classes illustrated by the histograms of Fig. 6. The interval between the threshold T_- and the previous one T_+ delimits doubt classes SW/NP or NB/NP between the different classes SW, NB or NP. The threshold T_- is computed such that $\Pr[\text{SNR}_\ell \leq T_- \mid H_1] = P_{\min}$. In our simulations, we have chosen $P_{\min} = 0.01$. Fig. 7 presents the histograms of the second feature $e_{-3\text{dB}}$ showing without doubt that the -3 dB error feature is meaningful to separate the SW peaks from the NB ones when the signal to noise ratio is greater than 0 dB. Both histograms are well-separated by a threshold denoted by T_e . The resulting feature space mapping is shown on Fig. 8.

5. Multi-estimator method

5.1. Combining different spectral estimators

Once spectral peak detection and classification have been set up, the question is: among all existing spectral estimators, which one are we going to choose? The key idea of the proposed

method is not to apply one estimator but to combine the results of several different estimators, taking advantage of their specific properties. Within the frame of spectral analysis, it exists several kinds of Fourier-based estimators [1], [2], [14]. In order to select M different and complementary ones, this Section outlines first a possible list of quality properties to help with the choice and then an example choice among many for a specific view to sine wave search. The number M of selected estimators is not that important, the method works whatever this number. The main claim of this paper is the multi-estimator strategy leading to select and merge different spectral estimations.

5.2. Possible estimator quality properties

In order to select Fourier-based spectral estimators, six estimator quality criteria are pushed forward:

1. **The degree of freedom r of a spectral estimator.** Defined in (7), r is also referred to as a quality ratio of a spectral estimator [11] since it is inversely proportional to the normalized variance (see (7)) and directly linked to the statistical stability of the spectral estimation. The minimum value of r is 2 and the higher the value, the smoother the spectral estimation.
2. **The spectral leakage percentage p .** It measures the power which is transferred from the main lobe of the spectral window $W(\nu)$ to its side lobes. An accurate definition is only given in [18], although this notion is widely discussed in the spectral analysis literature [19], [1],

$$p = 1 - \frac{\int_{-B_W/2}^{+B_W/2} |W(\nu)| d\nu}{\int_{-F_s/2}^{+F_s/2} |W(\nu)| d\nu}, \quad (25)$$

with F_s the signal sampling frequency.

3. **The peak emergence ΔSNR in the spectral estimation.** Related to the emergence of a given peak throughout the noise, ΔSNR is defined from an estimated point of view as the difference between two ratios,

$$\Delta\text{SNR} = \text{SNR}_\ell - \text{SNR}_+. \quad (26)$$

SNR_ℓ is the local signal to noise ratio defined in (22) and SNR_+ is the emerging signal to noise ratio, taking the local level of noise and spectral window effects into account, B_e being defined in (24),

$$\text{SNR}_+ = 10 \log \left(\frac{A_i^2}{4\sigma_i^2 B_e} \right). \quad (27)$$

Thus the peak emergence is defined by

$$\Delta SNR = 10 \log (B_e), \quad (28)$$

and can also be interpreted as the equivalent noise bandwidth of the spectral window [2] given in dB.

4. **The effective statistical bandwidth B_s of the spectral window** [1], [20], [21],

$$B_s = \frac{\left| \int_{-F_s/2}^{+F_s/2} W(\nu) d\nu \right|^2}{\int_{-F_s/2}^{+F_s/2} |W(\nu)|^2 d\nu}. \quad (29)$$

Related to the equivalent spectral window $W(\nu)$, the statistical bandwidth B_s represents the bandwidth of an equivalent rectangular window with the same normalized variance for a white noise as the chosen window. B_s characterizes the quality of the spectral estimation as well as the degree of freedom does but in an opposite way: it is not possible to have a fine spectral resolution (small values of B_s) together with a highly stable spectral estimation (large values of r).

5. **The frequency resolution** directly related to the resolution bandwidth B_{3dB} **of the spectral estimator** [2]. It is defined by

$$B_{\text{3dB}} = \nu_2 - \nu_1, \quad (30)$$

with

$$W(\nu_1) = W(\nu_2) = \frac{W(0)}{2}.$$

Details (e.g., two SW peaks) in a spectrum which are separated by less than this value cannot be resolved. This property has to be associated to the leakage since a compromise has to be found: high frequency resolution (small values of B_{3dB}) is unfortunately associated to high leakage (high values of p).

6. **The side-lobe fall-off SL of the equivalent spectral window** defined as the slope of the envelope [19], [22]. In case of a high signal to noise ratio, it is preferable to consider estimators with high side-lobe fall-off in order that side-lobes do not intervene in the peak detection results.

5.3. One estimator choice example for a sine wave search

Within the case of a sine wave search and based on the previous properties $(p, B_{-3\text{dB}}, B_s, r, \text{SL}, \Delta\text{SNR})$, it would be optimal to have an estimator which simultaneously satisfies a high degree of freedom r , a minimum spectral leakage percentage p , a maximum peak emergence ΔSNR , a small effective statistical bandwidth B_s , a small resolution bandwidth $B_{-3\text{dB}}$ and a high sidelobe fall-off SL . Since there exists no estimator satisfying the six above criteria, we propose a compromise, a right balance among all them by considering five methods ($M = 5$) as following with $m = 1, \dots, M$, [2]

- $m = 1$, a Blackman-Tuckey estimator using a biased estimate of the autocorrelation and a Blackman temporal window ($\text{BT}_b + \text{Blackman}$),
- $m = 2$, a Welch estimator using a Blackman temporal window ($\text{Welch} + \text{Blackman}$),
- $m = 3$, a Welch estimator with a Hanning temporal window ($\text{Welch} + \text{Hanning}$),
- $m = 4$, a Blackman-Tuckey estimator using an unbiased estimate of the autocorrelation with a Blackman temporal window ($\text{BT}_u + \text{Blackman}$),
- $m = 5$, a Blackman-Tuckey estimator using an unbiased estimate of the autocorrelation with a Hanning temporal window ($\text{BT}_u + \text{Hanning}$).

In order to illustrate the pros and cons of these five methods, for each of them, Fig. 9 plots a star chart where each radius represents one of the six previous criteria. As defined in the first chart, each radius scale has been determined such that the radius extremity corresponds to an optimal value of the corresponding criterion.

None of the chosen methods has the best scores for all criteria but each criterion has almost one estimator where the achieved score is the best. One can highlight that,

- for the quality criterion related to the degree of freedom r , the first method outperforms the others,
- for criteria linked to effective statistical bandwidth B_s and to frequency resolution $B_{-3\text{dB}}$, the Blackman-Tuckey estimators are the best,
- regarding leakage p , all methods are quite equivalent except the Blackman-Tuckey estimator with a Hanning window,

- for peak emergence ΔSNR , the three first estimators are of interest.

Lastly, considering the side-lobe fall-off, two different strategies have been developed. When the signal to noise ratio is low, the five chosen methods should be applied regarding their complementary properties ($M = 5$). However, in the case of a high signal to noise ratio (over 10 dB for example), the side-lobe fall-off values of the Blackman-Tuckey methods (methods 1, 4 and 5) could lead to high false-alarm rate within the detection step. Therefore, in this case, only Welch methods are considered ($M = 2$). For this specific search, it is then useful to have the knowledge of the signal to noise ratio. Other choices will not need this knowledge. In this respect, we have implemented a data validation to estimate some prior parameters such as the signal to noise ratio and also a stationary index to qualify the stationarity of the signal. This index is particularly worthwhile for selecting correct vibration data in a wind turbine for example [23], [24]. Moreover for such signals which can be measured under small speed fluctuations, an angular resampling should be applied if a speed signal is available [25]. All this upstream phase is not part of this paper.

As mentioned at the beginning of this Section, the most important point of this paper is the multi-estimator approach which can be used whatever the number and the type of spectral estimators. This Section has proposed one possible choice that has been argued using estimator quality criteria.

6. Fusion

Sections 3 and 4 have led to peak detection and classification on M spectral estimates. The present Section is devoted to the fusion of all these estimations in order to get one final spectral characterization of each detected spectral structure.

After detection and classification, the output of the m th spectral estimator ($m = 1, \dots, M$) is a list of P_m detected peaks or structures denoted as $\gamma_{i,m}, i = 1, \dots, P_m$. At each $\gamma_{i,m}$ is associated an attribute vector composed of the ten following elements, for which the index m has been added relatively to previous Sections:

1. the adjusted central frequency $\tilde{\nu}_{\max_{i,m}}$ obtained after the spectral adjustment of the detected peak with the equivalent spectral window in the classification step, as explained in Section 3.4,
2. the adjusted maximum of the peak $\tilde{\gamma}_{\max_{i,m}}$, corresponding to the adjusted central frequency, as also explained in Section 3.4,

3. the amplitude $A_{i,m} = \sqrt{4\tilde{\gamma}_{\max_{i,m}} B_s T}$, corresponding to the temporal amplitude, B_s being defined in (29) and T being the signal duration,
4. the local noise variance $\sigma_{i,m}^2$ defined in (16),
5. the confidence interval $\mathcal{I}_{i,m}$ corresponding to the frequency resolution of the spectral estimator, centered around $\tilde{\nu}_{\max_{i,m}}$ with a bandwidth B_{3dB} , defined in (30), $\mathcal{I}_{i,m} = [\tilde{\nu}_{\max_{i,m}} - B_{\text{3dB}}/2, \tilde{\nu}_{\max_{i,m}} + B_{\text{3dB}}/2]$,
6. the error couple $(e_{-3\text{dB}_{i,m}}, e_{\text{tot}_{i,m}})$ related to the spectral estimator and to its equivalent spectral window, defined in (19),
7. the local signal to noise ratio, $\text{SNR}_{\ell_{i,m}}$ defined in (22),
8. the emerging signal to noise ratio, $\text{SNR}_{+i,m}$ defined in (27),
9. the p -value, $p_{0_{i,m}}$ introduced in (12),
10. the class label $C_{i,m}$ selected among the set

$$C = \{\text{SW}, \text{NB}, \text{SW/NP}, \text{NB/NP}, \text{NP}\}, \quad (31)$$

resulting from the feature space mapping described in Section 4.3.

Computed for each m th method, these attribute vectors are called intermediate “identity cards”. Consequently, the first step of the fusion is to group or associate the intermediate identity cards corresponding to the same spectral structure among the M spectral estimators. For this, we will use attributes of the cards. The first one, the adjusted central frequency has not been estimated with the same resolution and then has not the same value throughout the different analyses. The association rule will be based on the fifth attribute of the cards, the confidence intervals $\mathcal{I}_{i,m}$. A non-empty interSection will lead to the association of the peaks numbered i and i' if and only if

$$\mathcal{I}_{i,m} \cap \mathcal{I}_{i',m'} \neq \emptyset \text{ for } i \neq i'. \quad (32)$$

Applying this rule leads to the association of maximum M cards which constitutes a sequence describing the same spectral structure viewed through the different spectral estimators. Scanning all the cards through the entire frequency band of the signal provides \mathcal{J} sequences, $j = 1, \dots, \mathcal{J}$ one for each spectral structure of the signal.

At this step, at one sequence j is associated M intermediate identity cards or less if the spectral structure has not been detected by all spectral methods. So the second step of the fusion is to

merge these cards in a final one. This step will be driven by the class label assigned in the feature space $(\text{SNR}_{\ell}, e_{-3\text{dB}})$ of Fig. 8. The fusion computes the centroid position of all the $(\text{SNR}_{\ell}, e_{-3\text{dB}})$ couples of the sequence to provide the final class label C_j .

This centroid position is used to select the nearest couple $(\text{SNR}_{\ell_{i,m}}, e_{-3\text{dB}_{i,m}})$, $m = 1, \dots, M$ and then the corresponding intermediate identity card which is denoted as the best one for keeping the following attributes in the final card of the sequence j , namely the frequency $\tilde{\nu}_{\max_j}$, the corresponding temporal amplitude A_j , the local noise variance σ_j^2 , the SNR_{ℓ_j} , the SNR_{+j} and the p -value p_{0_j} . Rather than keeping the confidence interval, the fusion step computes a relative bandwidth

$$\tilde{B}_j = \frac{B_j}{B_{-3\text{dB}}}, \quad (33)$$

where B_j is the -3 dB spectral bandwidth in the best intermediate card and $B_{-3\text{dB}}$ is the -3 dB spectral bandwidth of the equivalent spectral window for the corresponding spectral estimator of this best card (30).

In the final card of the sequence is added a specific attribute computed as the detection number percentage of each spectral structure throughout the M spectral analyses and acting as a confidence index denoted CI_j for the sequence j . In the end of the fusion step, the final identity card of each of the \mathcal{J} detected structures is of the form

$$(\tilde{\nu}_{\max_j}, A_j, \sigma_j^2, \text{SNR}_{\ell_j}, \text{SNR}_{+j}, \tilde{B}_j, p_{0_j}, C_j, \text{CI}_j). \quad (34)$$

7. Applications

Figure 10 illustrates the main steps of the method proposed in this paper. The method is fully automatic and no *a priori* knowledge about the signal is needed. Some hyper parameters have to be set *a priori*: the spectral estimators, with their own parameters and thresholds in the feature space as illustrated in Figs. 5, 6 and 7, and the PFA in the detection step. As an illustration, Section 5.2 has proposed a selection of 5 estimators based on complementary quality criteria for a sine wave search. Obviously the method is valuable for any choice and any number of estimators. The PFA acts as a maximum p -value and should be chosen relatively high to be able to detect enough peaks. We use to choose a PFA value equal to 10^{-3} . This Section illustrates some results of the proposed method first for a simulated signal and then for real-world signals. It concludes about a strong interest of this method in the surveillance and maintenance of complex systems.

7.1. A simulated signal

The proposed automatic data-driven multi-estimator approach is applied to the sine-wave signal embedded in an additive non-white noise described in Table 1.

Fig. 11 presents the results of the spectral structure detection detailed in Section 3 for one specific spectral estimator. All local maxima of the PSD associated to a p -value higher than the chosen threshold $PFA = 10^{-3}$ are considered as in favour of hypothesis H_0 and are plotted in black. The remaining maxima corresponds to the P detected peaks and for sake of clarity, their colour corresponds to one p -value interval, rather than giving their associated p -values. The 7-sine wave structures have been well detected with a very low p -value ($\leq 10^{-6}$) and 7 other structures have been also detected but with a higher p -value, from 10^{-6} to 10^{-3} . To characterize further these P structures, the classification step detailed in Section 4 is illustrated in Fig. 12. The 7 sine waves are classified either as SW, or as NB. Among the 7 supplementary spectral structures detected, 5 are classified as NB/NP, 1 as NB and 1 as SW/NP.

The result of this analysis will be enhanced by the multi-estimator approach. Fig. 13 presents the results of detection and classification of each of the M estimators that have detected between 8 and 15 spectral structures. High amplitude structures have been logically detected with a low p -value whereas small amplitude structures close to the noise level have been detected with a higher p -value.

In the end, the fusion step provides 19 sequences that are summarized in Table 2, where it is ordered by decreasing confidence indices and increasing p -values. One can be surprised that the proposed method detects 19 structures whereas the signal has been simulated with 7 ones only. However, this automatic method is able to analyze a signal with no a priori knowledge providing a list of detected spectral structures with several attributes, among which the three more discriminant ones are the confidence, the p -value and the SNR_ℓ .

The 19 structures of Table 2 are plotted in this 3D-space in Fig. 14. In this space, the red “+” correspond to the sequences with a maximum confidence index of 100% and a very low p -value lower than 10^{-6} . Looking at Table 2, these sequences are classified as SW or NB, and their frequencies correspond to the seven peaks of the analysed simulated signal (see Table 1). Even the critical peak at 0.07 Hz is detected but as a NB due to its low signal to noise ratio. Twelve additional sequences have been detected. The green “o” in Fig. 14 correspond to two sequences with high confidence of

80% but with high p -value in the range of 10^{-6} to 10^{-4} and a low local signal to noise ratio SNR_ℓ . Their class labels highlight a doubt between NB and NP. The “x” black in Fig. 14 represent the last ten sequences with a confidence index lower than 50% and then of a lower interest. Note that the chosen simulated signal gathers most of signal peculiarities as a non-white noise, close peaks, and low amplitude ones in different frequency bands. The proposed method succeeds in detecting the 7 signal components with a high confidence (100 %) and a very low p -value.

This example shows the interest of the proposed multi-estimator method which is able to spotlight the true signal spectral structures and to play down the importance of other detected peaks thanks to a confidence attribute got from the results of several spectral estimators. For example, when using a Welch estimator with Hanning or Blackman windows, estimators 2 and 3 of Fig. 13, the peak at frequency 0.42 would be classified as a NB. The fusion strategy helps to adjust this conclusion by classifying it as a SW with a high confidence index (see Table 2). It is important to mention that the proposed method does not need to set *a priori* thresholds and is able to detect all peaks whatever their energy and whatever the color of the noise. The detection performance is that of the set of the chosen methods, and not of one of them. Another example of strategy results on a different simulated signal can be found in [8].

7.2. The acoustic noise of a domestic heat pump

In this Section the proposed method is applied on a real-world signal in order to help for the acoustic study of a domestic heat pump polluting the neighborhood. Fig. 15 presents a zoom on the results obtained by the five selected spectral estimators on one recorded acoustic signal at the middle of the day. For sake of simplicity, the color legend which was given in Fig. 13 is directly detailed for each detected peak. This particular zoom reinforces the idea that choosing a specific spectral analysis would induce conclusions that are for sure to be smoothed by considering several different spectral estimators. For example, the peak at 1,041 Hz is detected by all the five estimators with a very low p -value but three times as a SW and twice as a NB. Moreover, a peak is detected around 1,015 Hz by the three first analysis but not by the two last ones, while another peak is detected around 1,078 Hz by the two last Blackman-Tukey estimators while the three other ones do not detect it. These results have been positively validated by experts in acoustics.

This strategy has been applied to several real-world signals. Fusion results on a hydraulic noise of an oil station can be used to identify vibrations induced by this noise [7]. Once again, as for the

previous simulated signal, this example highlights the interest of the approach.

7.3. Towards a surveillance tool

To go further, this spectral analysis and fusion strategy can be helpful in maintenance for the surveillance of complex machines. Only to illustrate this purpose, the noise of the heat pump, that was accused by neighborhood to produce harmful noise, has been recording twice, one in the middle of the day, the signal analysed in the previous Section, and a second one in the middle of the night. Analytical fusion results can be computed for each signal, which can be directly used for a tracking of the frequencies. For a graphic view of the results, Fig. 16 displays the frequency attributes of the detected sequences in each fusion. For sake of clarity, all detected sequences are represented, whatever their class and p -value, the color scale being linked to their amplitude attribute. This allows to compare immediately the spectral structure of both signals and to highlight the increase in the number of peaks of the signal recorded during the night. Moreover, these adding peaks have also much higher power, within an harmonic structure, at frequencies sensitive to human hearing, all this being the signature of the harmful noise the neighborhood was complaining about. For further insight of a surveillance tool, see [26] where an automatic time-frequency tracking method is proposed.

8. Conclusion

In a context of longtime stationary processes, at least several hundred of samples and more often several thousand, embedded in non-white Gaussian noises of unknown variance and composed of several hundred of spectral structures, this paper presents a new data-driven approach in order to automatically detect and characterize the spectral structures from spectra estimated by Fourier-based estimators. The core idea is to apply a statistical hypothesis testing in the frequency domain. To that end, based on a simple generic model, an approximation of the probability density function of the estimated PSD is derived whatever the window and the overlap and taking into account an estimation of the noise spectrum. Indeed, noise being non-white, noise spectrum estimation is carried through nonlinear filters. Without *a priori* hypothesis on the signal except the Gaussianity of the additive noise, a p -value is used to give a measure of a strength of evidence of a spectral structure against the noise-only hypothesis. Applying this test for all local maxima of the estimated PSD allows the detection of the number of relevant spectral structures with a level of confidence

determined by this p -value. A specific least square approach refines the frequency and amplitude estimates of the so detected structures and provides a characterization of their spectral shape relatively to the oversampled spectral window of the used PSD estimator. These attributes are helpful for mapping the detected structures in an original feature space, defined by a local signal to noise ratio and the quadratic error with the spectral window. A classification step enables then the labelling of each structure according to its signal source (sine wave, narrow band or noise). In order to avoid the choice of a spectral estimator and of an ineluctable compromise, this detection and classification approach is applied on a set of estimators taking advantages of all of their specific properties and allowing the definition of a confidence index. A last fusion step outputs a complete attribute vector referred to as identity card for each detected spectral structure and ready to use in a decision-making process.

A key idea of this data-driven approach is that all parameters are automatically set up according to the input data and the chosen spectral estimators without any other *a priori* knowledge. The output as identity cards can be directly used for physical interpretation or as digital input of a diagnostic system. This approach is free from spectrum display and does not need *a priori* knowledge upon frequency bands as the usual methods in vibration analysis for example. This paper shows the analysis of acoustic measures of a domestic heat pump and the output directly yields a complete digital spectral description of spectra. Users have only to conclude about the links of the detected structures to the physical phenomena helped with a comprehensive list of attributes.

Based on this strategy, additional works have been developed in order to group harmonic families and sidebands [27], to associate detected structures to system kinematic when available [28] and to develop a surveillance tool [26], that has been successfully applied for an early detection and an exact localisation of a mechanical failure in an onshore wind turbine [25].

Appendices

A. Normalized variance for Welch and Blackman estimators

Given that spectral windows and variance for Fourier-based estimators are often mentioned in this paper, this appendix recalls the equations in the general case for the two main estimators, Welch and Blackman-Tuckey. The spectral window denoted $W(\nu)$ is defined as in (9). The time

effective window, denoted $w[m]$, is such that $\text{DFT}(w[m]) = W(\nu)$, where DFT denotes discrete Fourier transform.

For a Welch estimator [29], the signal $x[k]$ is split up into N_{seg} segments of sample size $K_{\text{Welch}} \leq K$, overlapping by D samples with $D < K$, and weighted by any window. In this case, the normalized variance of $\hat{\gamma}_x(\nu)$, denoted $\text{varn}_{\text{Welch}}$, as a function of the time effective window, denoted $g_{\text{Welch}}[m]$, is

$$\text{varn}_{\text{Welch}} = \frac{1}{N_{\text{seg}}} \sum_{n=1-N_{\text{seg}}}^{N_{\text{seg}}-1} \left(1 - \frac{|n|}{N_{\text{seg}}}\right) \frac{g_{\text{Welch}}^2[nD]}{g_{\text{Welch}}^2[0]}.$$

Special cases:

- if $N_{\text{seg}} = 1$, then $\text{varn}_{\text{Welch}} = 1$,
- for N_{seg} without overlapping, $D = 0$ and $\text{varn}_{\text{Welch}} = 1/N_{\text{seg}}$.

For a Blackman-Tukey estimator [30], the autocorrelation function of $x[k]$ is weighted by any window over $2K_{\text{BT}} - 1$ samples, with $2K_{\text{BT}} - 1 \leq K$. In this case, the normalized variance of $\hat{\gamma}_x(\nu)$, denoted varn_{BT} , as a function of the time effective window, denoted $g_{\text{BT}}[m]$, is

$$\text{varn}_{\text{BT}} = \frac{1}{K} \sum_{n=-K_{\text{BT}}+1}^{K_{\text{BT}}-1} \frac{K}{K - |n|} g_{\text{BT}}^2[n].$$

B. Local signal to noise ratio as a function of the statistical test

It is of interest to express the statistical test $T(\nu)$ defined in (10) according to the local signal to noise ratio SNR_ℓ defined in (22). Considering that the statistical test is applied only to frequencies at the local maxima of $\hat{\gamma}_x(\nu)$,

$$\text{for } \nu \in \mathcal{M} = \left\{ \max_{\nu} \hat{\gamma}_x(\nu) \right\}, \quad \hat{\gamma}_x(\nu) = \gamma_{\max_i} = \frac{A_i^2}{4} \frac{1}{B_e T_s},$$

with T_s the sampling period and B_e the noise equivalent bandwidth defined in (24). The noise estimation is assumed to be unbiased (see Section 3.1), therefore

$$\text{for } \nu \in \mathcal{M} = \left\{ \max_{\nu} \hat{\gamma}_x(\nu) \right\}, \quad \tilde{\gamma}_n(\nu) = \sigma_i^2.$$

Hence,

$$T(\nu) = \frac{\hat{\gamma}_x(\nu)}{\tilde{\gamma}_n(\nu)} = \frac{A_i^2}{4\sigma_i^2} \frac{1}{B_e T_s}.$$

Using (22) and writing the last equality in dB scale lead to

$$\text{SNR}_\ell = 10 \log [B_e T_s T(\nu)],$$

where the product $B_e T_s$ is a well-known quantity to characterize a Fourier method.

References

- [1] M. Priestley, Spectral analysis and time series, Academic Press, 1992.
- [2] P. Stoica, R. Moses, Introduction to spectral analysis, Prentice Hall, 1997.
- [3] J. Schoukens, Y. Rolain, G. Simon, R. Pintelon, Fully automated spectral analysis of periodic signals, *IEEE Trans. Instrum. Meas.* 52 (4) (2003) 1021–1024.
- [4] K. Barbé, W. Moer, Automatic detection, estimation and validation of harmonic components in measured power spectra: all-in-one approach, *IEEE Trans. Instrum. Meas.* 60 (3) (2011) 1061–1069.
- [5] L.-M. Zhu, H.-X. Li, H. Ding, Estimation of multi-frequency signal parameters by frequency domain non-linear least squares, *Mechanical Systems and Signal Processing* 19 (2005) 955–973.
- [6] D. Rabijins, G. Vandersteen, W. V. Moer, An automatic detection scheme for periodic signals based on spectrum analyzer measurements, *IEEE Trans. Instrum. Meas.* 53 (3) (2004) 847–853.
- [7] C. Mailhes, N. Martin, K. Sahli, G. Lejeune, Condition monitoring using automatic spectral analysis, in: *Third European Workshop on Structural Health Monitoring (SHM)*, Granada, Spain, 2006, pp. 1316–1323.
- [8] C. Mailhes, N. Martin, K. Sahli, G. Lejeune, A spectral identity card, in: *EUropean Signal Processing Conference, (EUSIPCO)*, Florence, Italy, 2006.
- [9] D. Rife, R. Boorstyn, Single-tone parameter estimation from discrete-time observations, *IEEE Trans. Inf. Theory* 20 (5) (1974) 591–598.
- [10] P. Johnson, D. Long, The probability density of spectral estimates based on modified periodogram averages, *IEEE Trans. Signal Processing* 47 (5) (1999) 1255–1261.

- [11] L. Koopman, The spectral analysis of time series, Academic Press, 1974.
- [12] T. Durrani, J. Nightingale, Probability density function for discrete Fourier spectra, Proc. Inst. Elec. Eng. 120 (2) (1973) 299–311.
- [13] J. Huillery, N. Martin, On the description of spectrograms probabilities with a Chi-squared law, IEEE Trans. Signal Processing 56 (6) (2008) 2249–2258.
- [14] M. Durnerin, A strategy for the interpretation in spectral analysis, detection and characterisation of spectrum components (in French), Ph.D. thesis, INPT, Grenoble, France (1999, available on <https://tel.archives-ouvertes.fr/tel-00789941/>).
- [15] N. Johnson, S. Kotz, N. Balakrishnan, Continuous Univariate Distributions, Vol. 1 (Wiley Series in Probability and Statistics), Wiley-Interscience, 1994.
- [16] E. Lehmann, J. Romano, Testing Statistical Hypothesis, Springer, 3rd Edition, 2005.
- [17] E. Bedrosian, Product theorem for Hilbert transforms, Proc. IEEE 51 (5) (1963) 868–869.
- [18] J. Max, J.-L. Lacoume, Signal processing methods - applications to physical measures (in French), Masson, 1996.
- [19] F. J. Harris, On the use of windows for harmonic analysis with the discrete Fourier transform, Proc. IEEE 6 (1) (1978) 51–83.
- [20] R. Blackman, J.W. Tukey, The measurement of power spectra from the point of view of communication engineering, Bell Syst. Tech. J.
- [21] G. Jenkins, D. Watts, Spectral analysis and its applications, Holden Day, 1968.
- [22] A. H. Nuttall, Some windows with very good sidelobe behavior, IEEE Trans. Acoust., Speech, Signal Processing 29 (1) (1981) 84–91.
- [23] N. Martin, C. Mailhes, About periodicity and signal to noise ratio - The strength of the autocorrelation function, in: Seventh International Conference on Condition Monitoring and Machinery Failure Prevention Technologies (CM 2010 and MFPT 2010), Stratford-upon-Avon, UK, 2010.

- [24] G. Song, Z.-Y. Li, P. Bellemain, N. Martin, C. Mailhes, AStrion data validation of non-stationary wind turbine signals, in: Twelve International Conference on Condition Monitoring and Machinery Failure Prevention Technologies (CM 2015 and MFPT 2015), Oxford, UK, 2015.
- [25] X. Laval, G. Song, Z.-Y. Li, P. Bellemain, M. Lefray, N. Martin, A. Lebranchu, C. Mailhes, AStrion assets for the detection of a main bearing failure in an onshore wind turbine, in: Thirteen International Conference on Condition Monitoring and Machinery Failure Prevention Technologies (CM 2016 and MFPT 2016), Paris, France, 2016.
- [26] T. Gerber, N. Martin, C. Mailhes, Time-frequency tracking of spectral structures estimated by a data-driven method, *IEEE Trans. Industrial Electronics* 52 (10) (2015) 6616 – 6626.
- [27] T. Gerber, N. Martin, C. Mailhes, Identification of harmonics and sidebands in a finite set of spectral components, in: Tenth International Conference on Condition Monitoring and Machinery Failure Prevention Technologies (CM 2013 and MFPT 2013), Cracovie, Poland, 2013.
- [28] M. Firla, Z.-Y. Li, N. Martin, C. Pachaud, T. Barszcz, Automatic characteristic frequency association and all-sideband demodulation for detection of a bearing fault of a test rig, *Mechanical Systems and Signal Processing* 80 (2016) 335–348.
- [29] P. Welch, The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms, *IEEE Trans. Audio and Electroacoustics* 15 (6) (1967) 70–73.
- [30] R. Blackman, J. Tukey, *The Measurement of Power Spectra from the Point of View of Communications Engineering*, Dover Publications, New York, 1958.

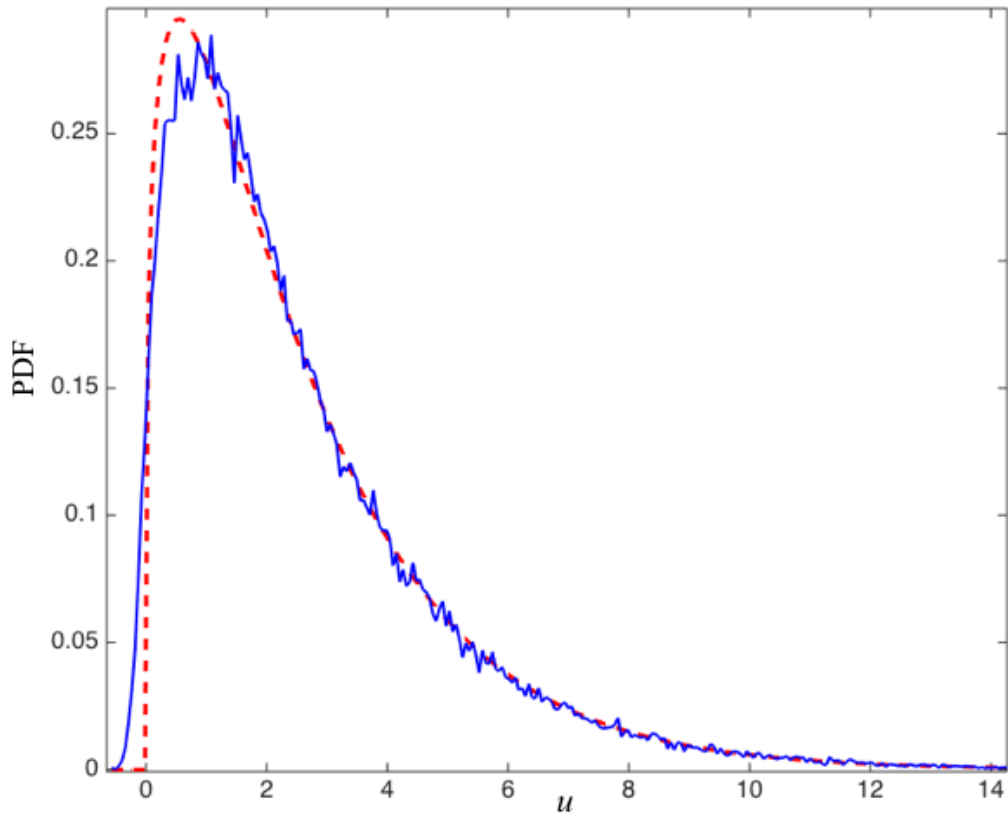


Figure 1: Estimated probability density function (PDF) of a Blackman-Tukey spectral estimator (blue solid line) compared to its corresponding χ_r^2 approximation (red dashed line) with $r = 2.55$

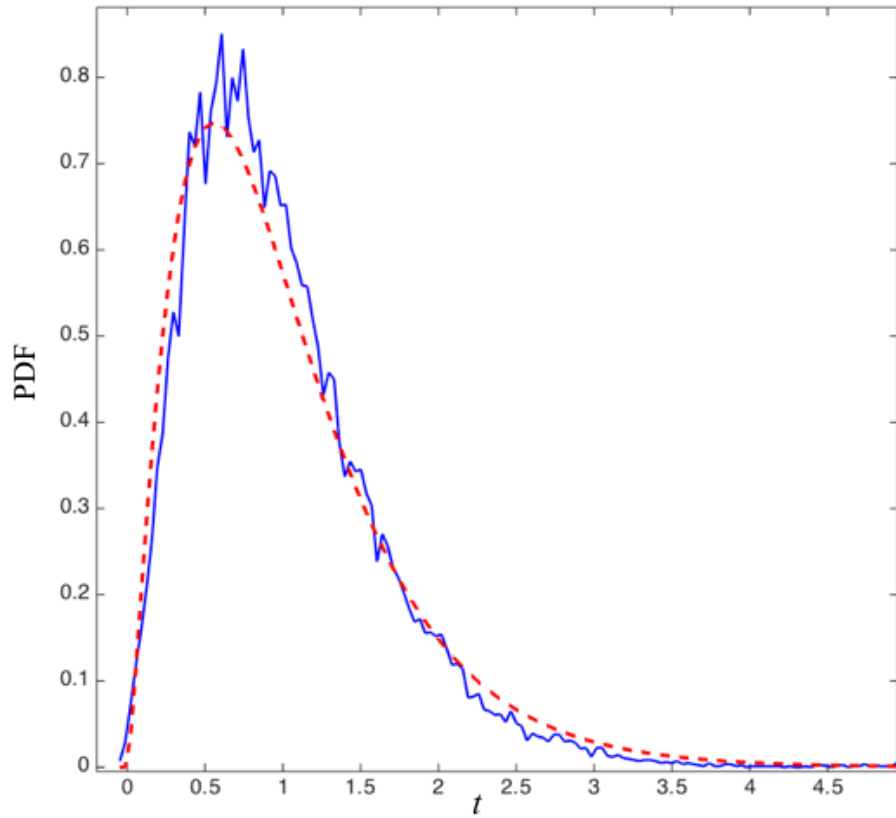


Figure 2: Estimated probability density function (PDF) of the test statistics in the case of a Blackman-Tukey spectral estimator (blue solid line) compared to its corresponding F_{r,rM_n} distribution (red dashed line), with $r = 2.55$ and $M_n = 29$

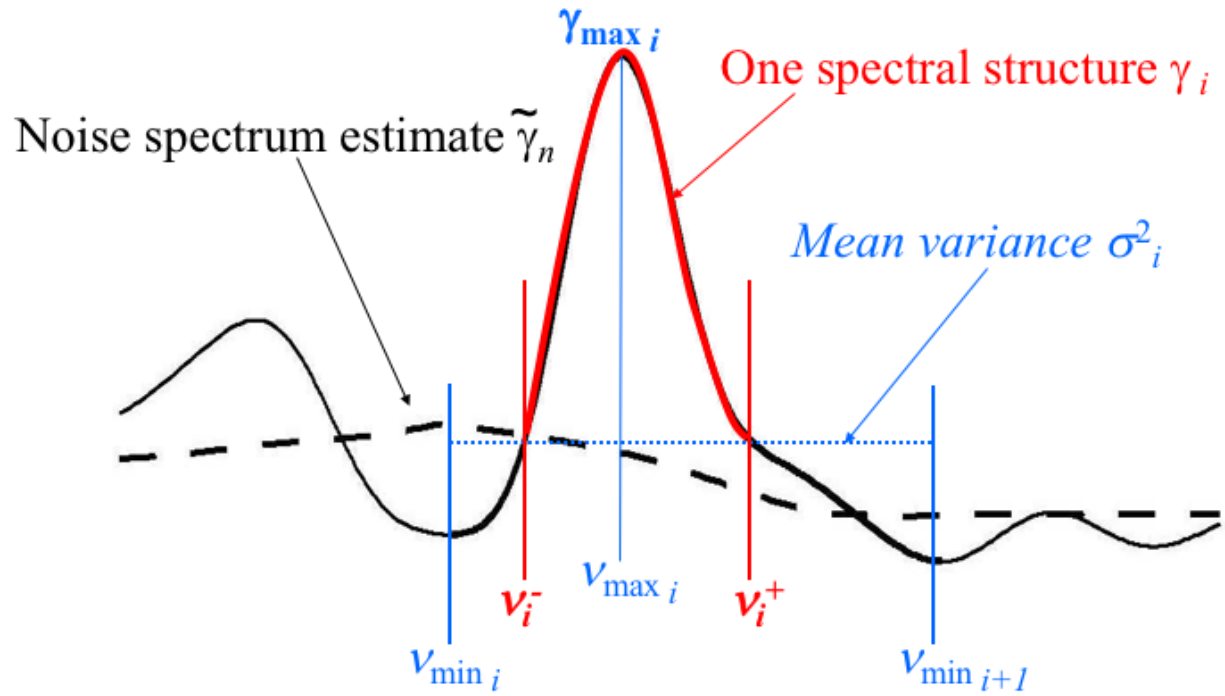


Figure 3: Definition of a spectral structure γ_i (in red) with a maximum γ_{\max_i} at frequency ν_{\max_i} as a portion of the estimated PSD between frequencies ν_i^- and ν_i^+ , such that PSD values at these frequencies equal the local noise variance σ_i^2 (in blue) corresponding to the mean noise PSD between two successive local minima at frequencies ν_{\min_i} and $\nu_{\min_{i+1}}$ around this spectral structure.

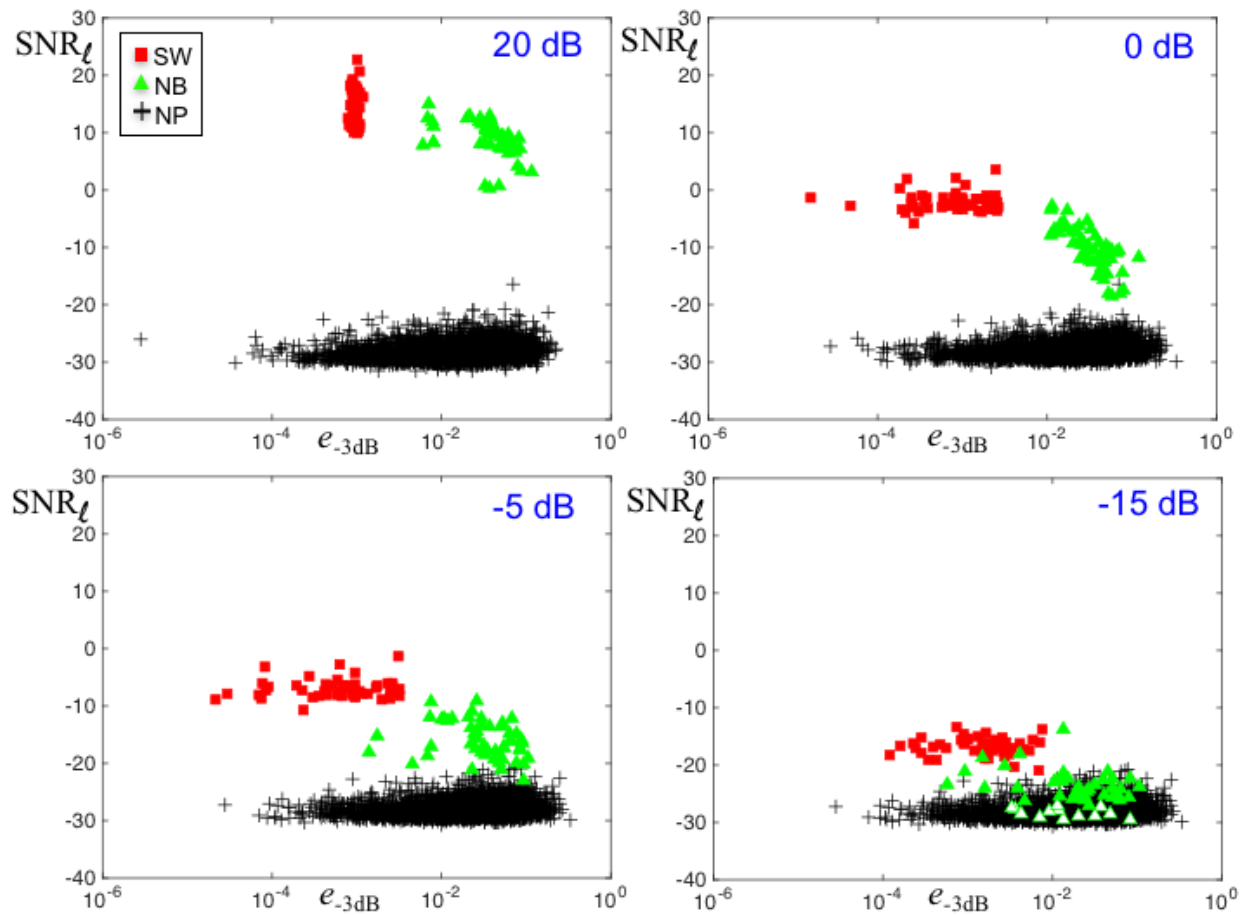


Figure 4: Monte Carlo simulations (50 runs) for a SW signal and a NB signal embedded in an additive Gaussian noise (signal to noise ratio at the upper right of each plot): each symbol corresponds to a detected peak, using a Welch estimator on the whole signal (one segment of $K = 10000$ samples) with a Blackman window, $r = 2$ and $M_n = 81$, the red squares are the SW ones, the green triangles, the NB ones, the rest the NP ones.

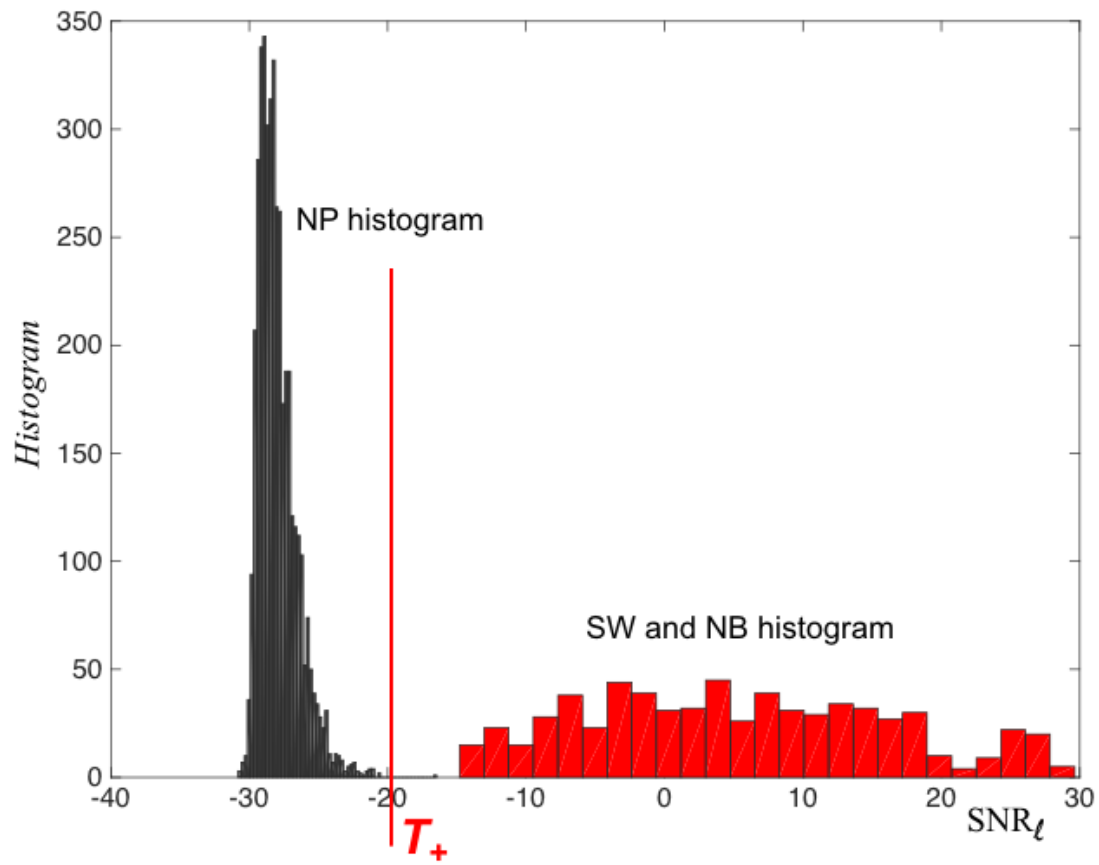


Figure 5: SNR_ℓ histograms on NP (left black) and signal peaks (SW and NB, right red) of simulated signals embedded in an additive Gaussian noise using Welch PSD estimation with Hanning window, SNR from 30 dB to -10 dB for SW and from 30 dB to 5 dB for NB, 50 signal realizations of 10000 samples per SNR value.

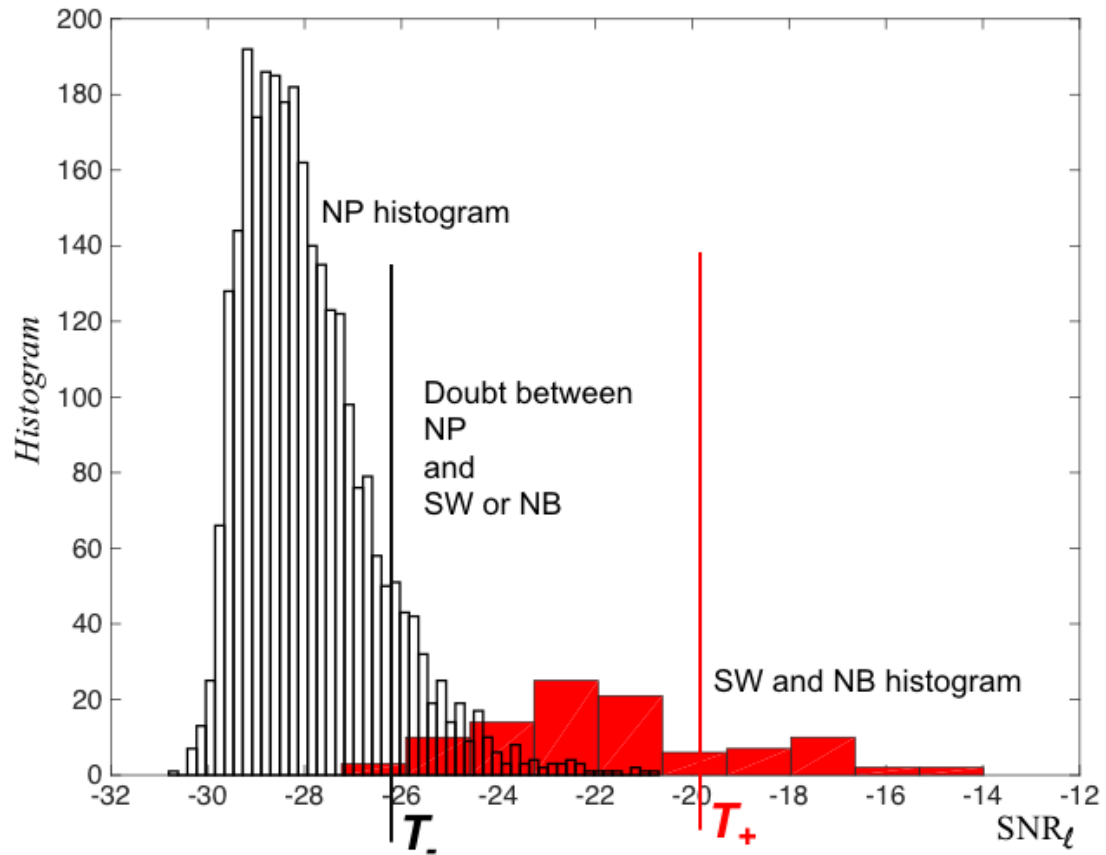


Figure 6: SNR_l histograms on NP (left black) and signal peaks (SW and NB, right red) of simulated signals embedded in an additive Gaussian noise using Welch PSD estimation with Hanning window, at very low SNR, -20 dB for SW and -10 dB for NB, 50 signal realizations of 10000 samples per SNR value.

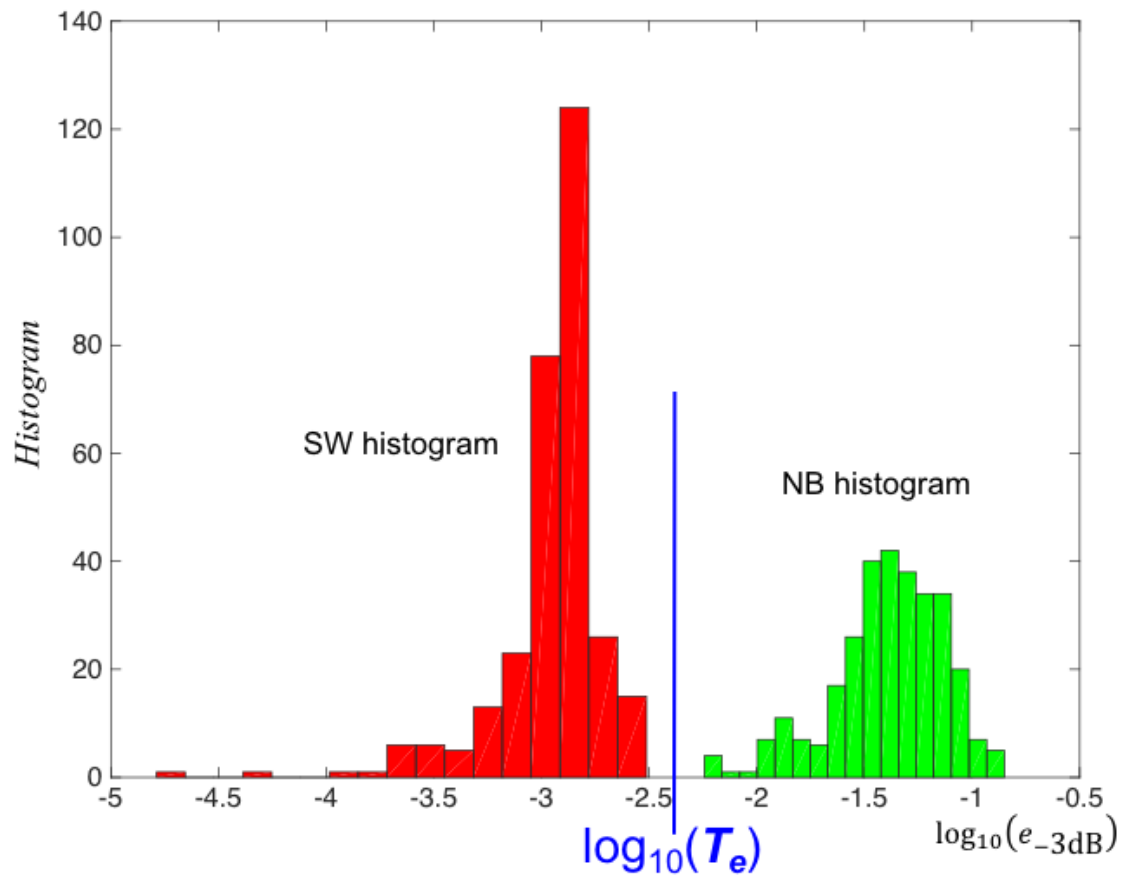


Figure 7: -3dB error histograms on SW (left red) and NB (right green) simulated signals embedded in an additive Gaussian noise using Welch PSD estimation with Hanning window, SNR from 30 dB to 0 dB, 50 signal realizations of 10000 samples per SNR value.

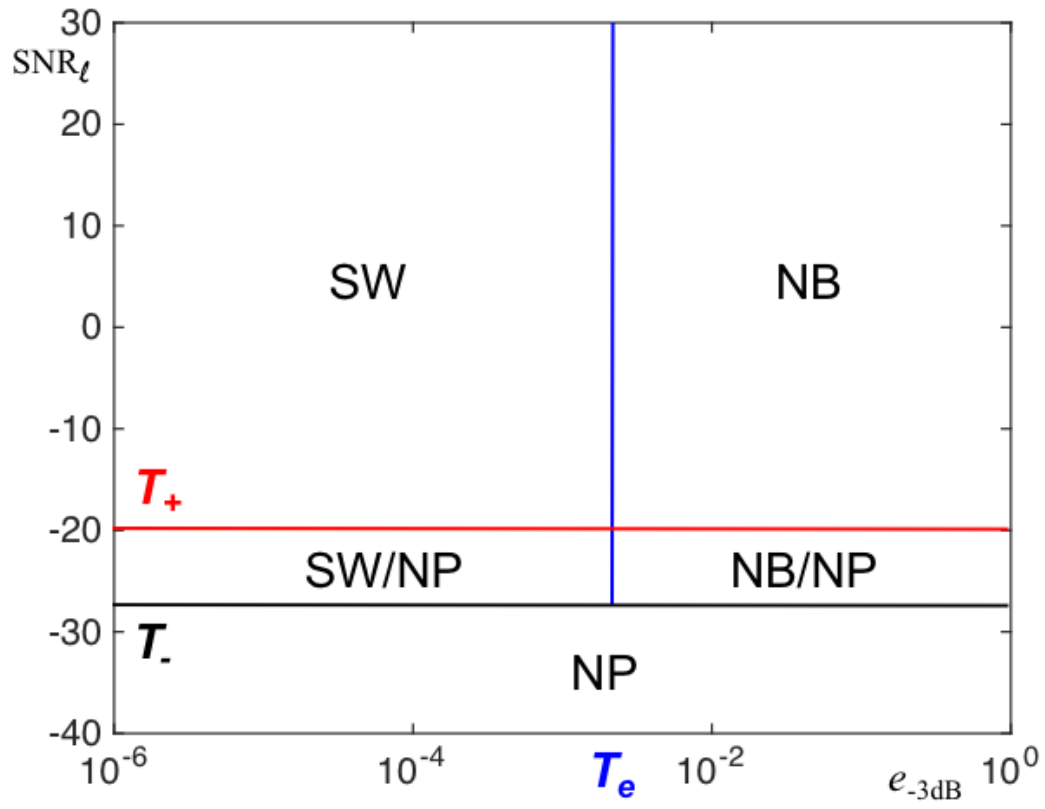


Figure 8: Feature space mapping of the $(\text{SNR}_\ell, e_{-3\text{dB}})$ plane into 5 classes.

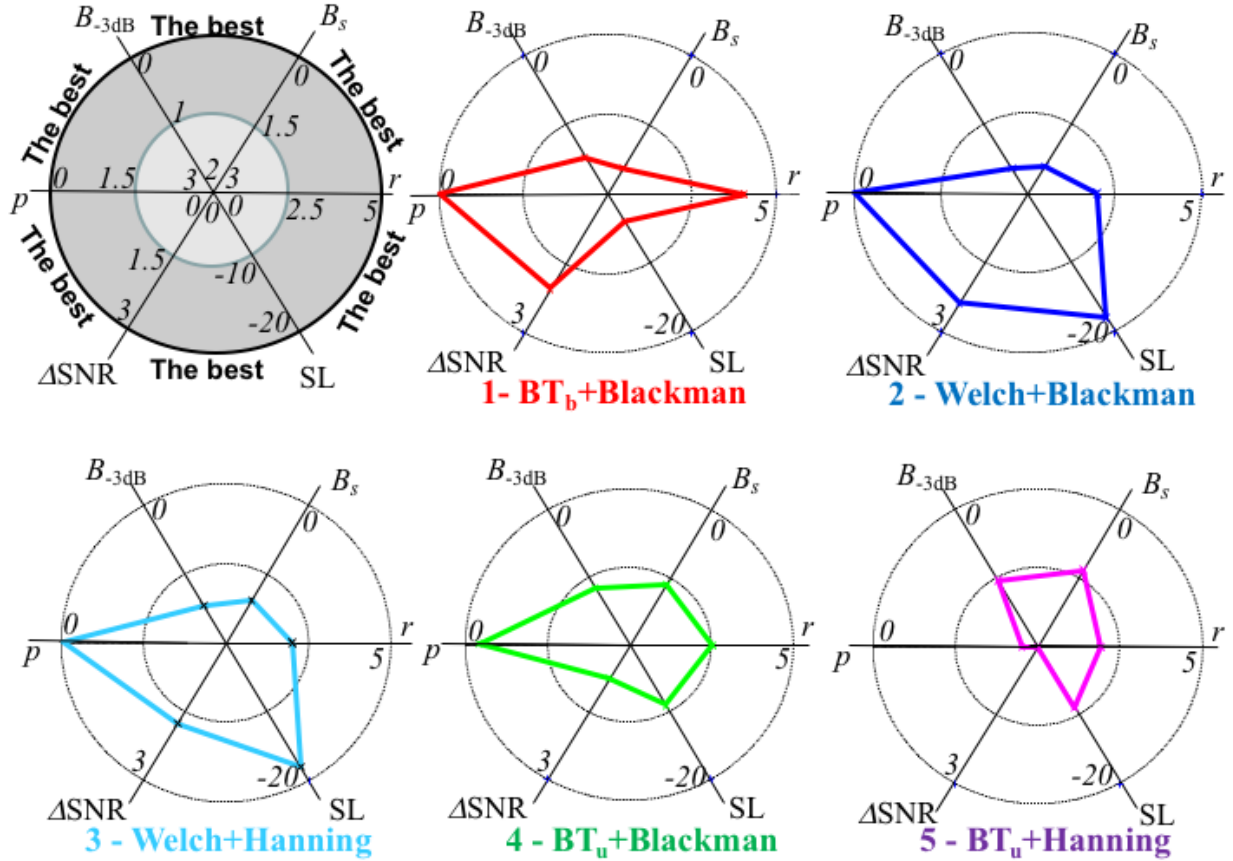


Figure 9: estimator quality properties ($p, B_{-3dB}, B_s, r, SL, \Delta SNR$) of different estimators: a Blackman-Tukey estimator using a biased estimate of the autocorrelation and a Blackman temporal window ($BT_b + \text{Blackman}$), a Welch estimator using a Blackman temporal window (Welch + Blackman), a Welch estimator with a Hanning temporal window (Welch + Hanning), a Blackman-Tukey estimator using an unbiased estimate of the autocorrelation with a Blackman temporal window ($BT_u + \text{Blackman}$), a Blackman-Tukey estimator using an unbiased estimate of the autocorrelation with a Hanning temporal window ($BT_u + \text{Hanning}$). The first grey diagram defines the different axis scales.

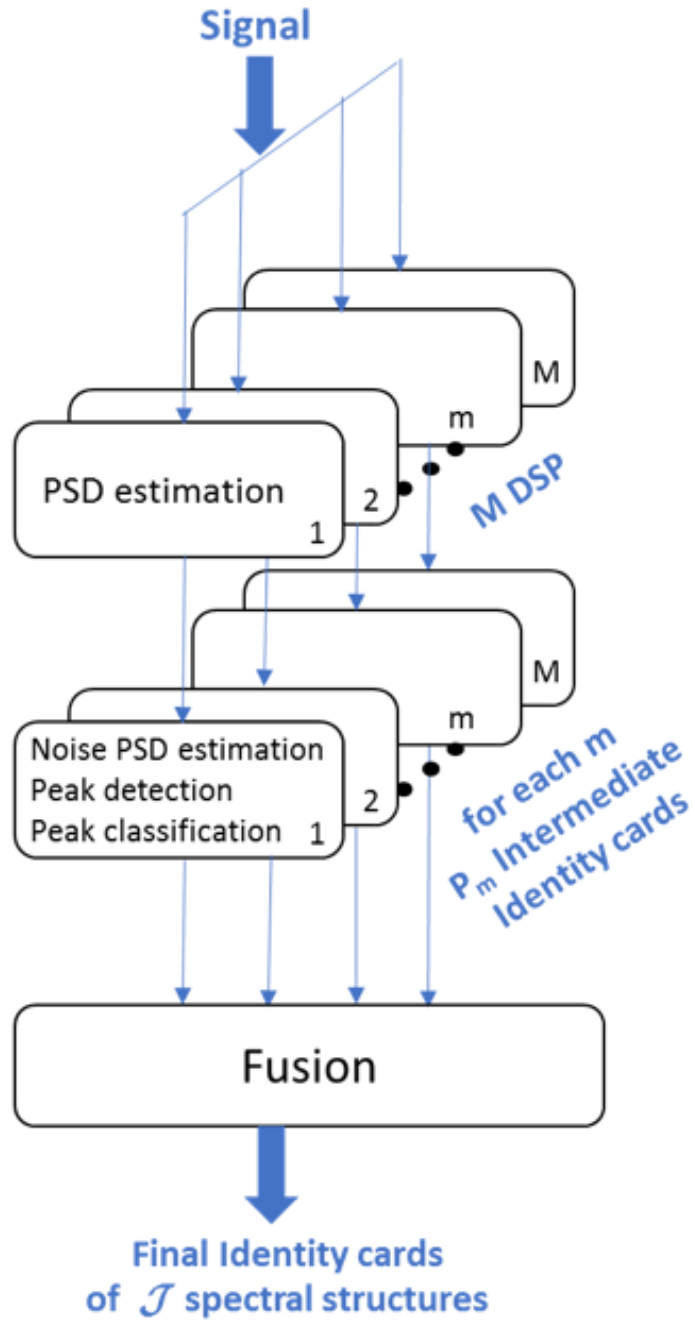


Figure 10: Global scheme of the proposed multi-estimator approach.

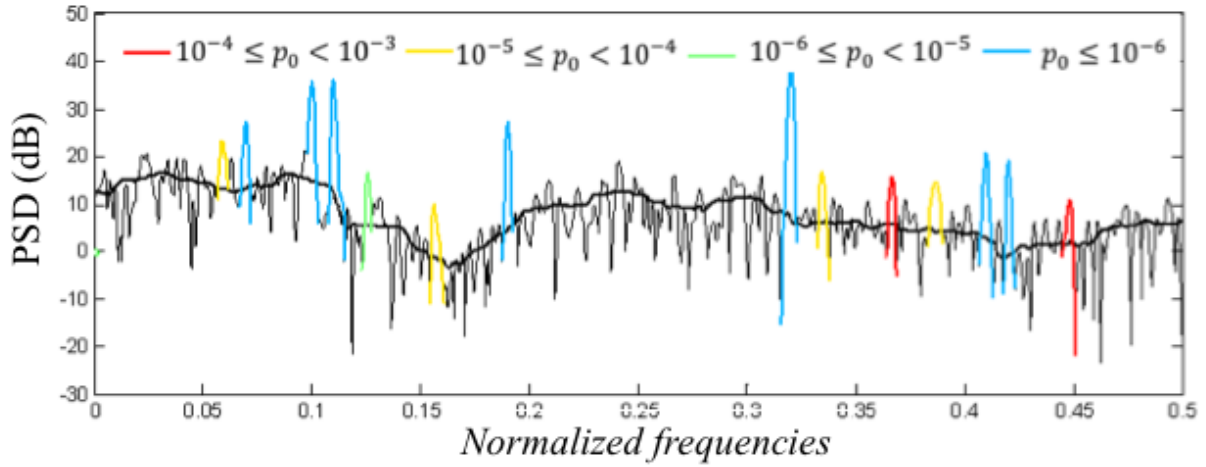


Figure 11: Detection result on the simulated signal described in Table 1 for Welch PSD estimation with Hanning window ($r = 2$ and $M_n = 37$): detected peaks are colored in red, yellow, green or cyan depending on their associated p -value, while the continuous black line along the PSD represents the noise PSD estimation.

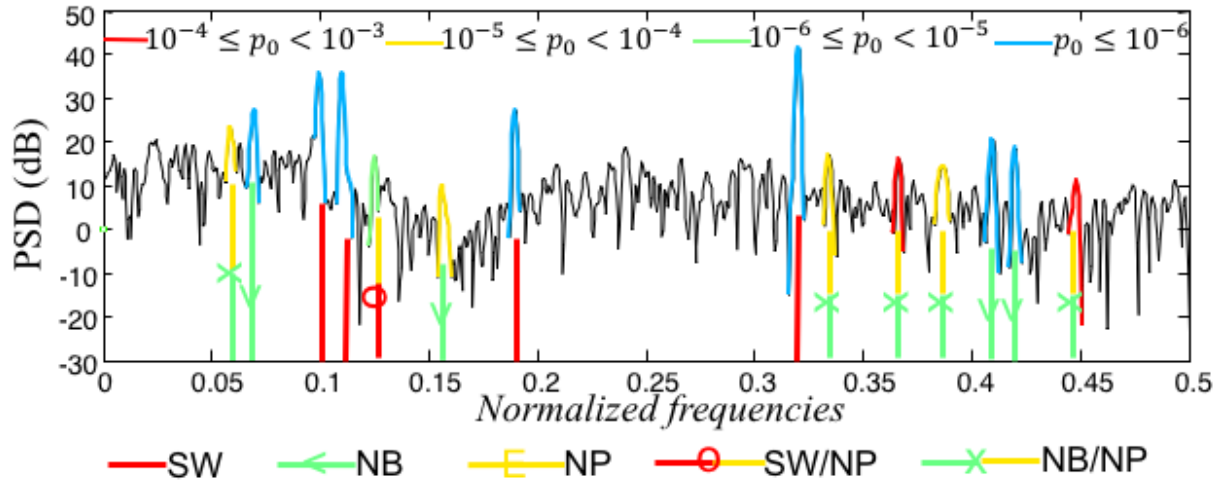


Figure 12: Classification result on the noisy signal described in Table 1 for Welch PSD estimation with Hanning window ($r = 2$ and $M_n = 37$): classified peaks have a colored bar corresponding to different classes, as indicated.

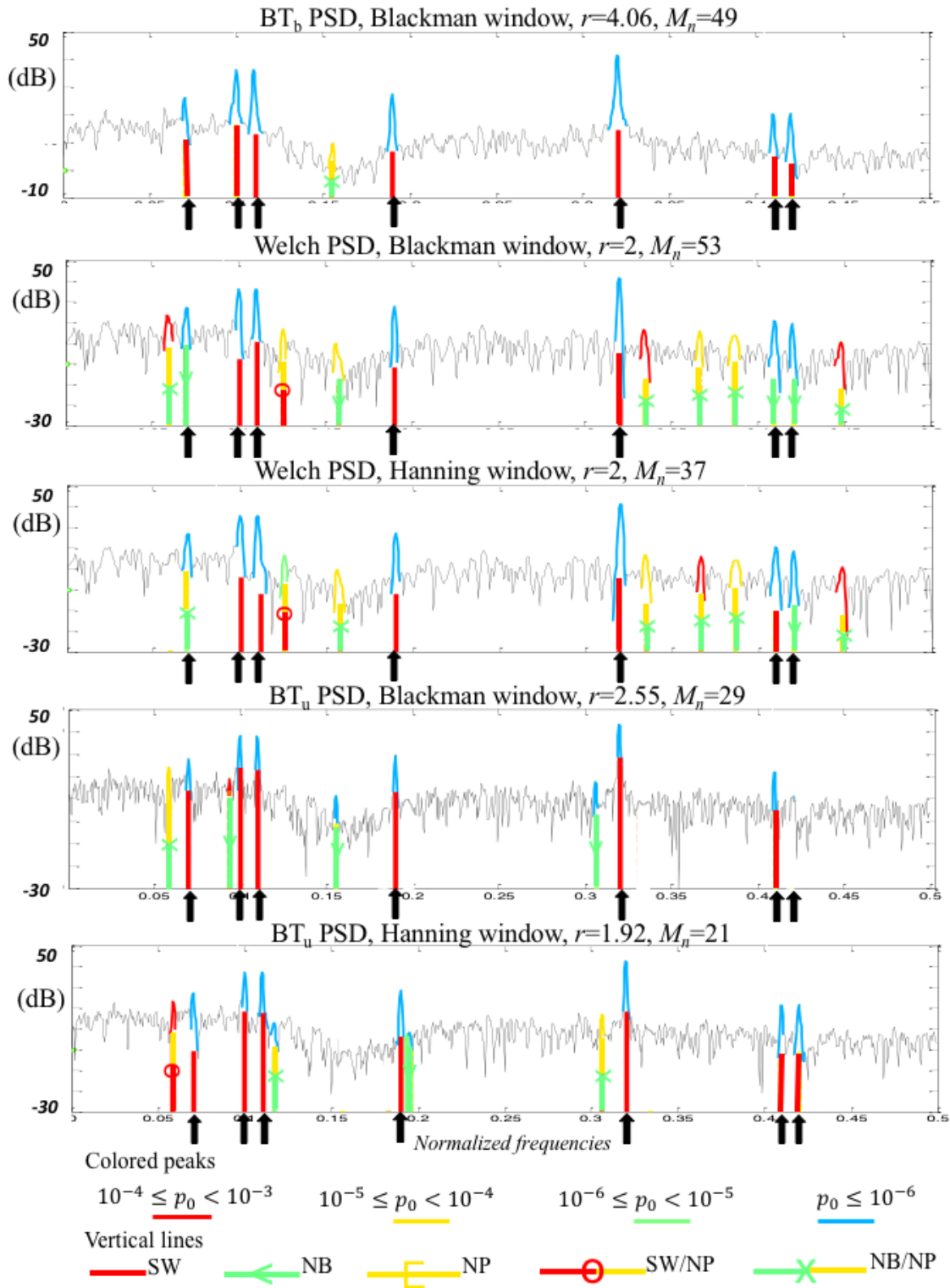


Figure 13: Results of detection and classification on signal described in Table 1 for the five selected spectral estimators: the true frequency values of the seven sine waves are spotted with black arrows.

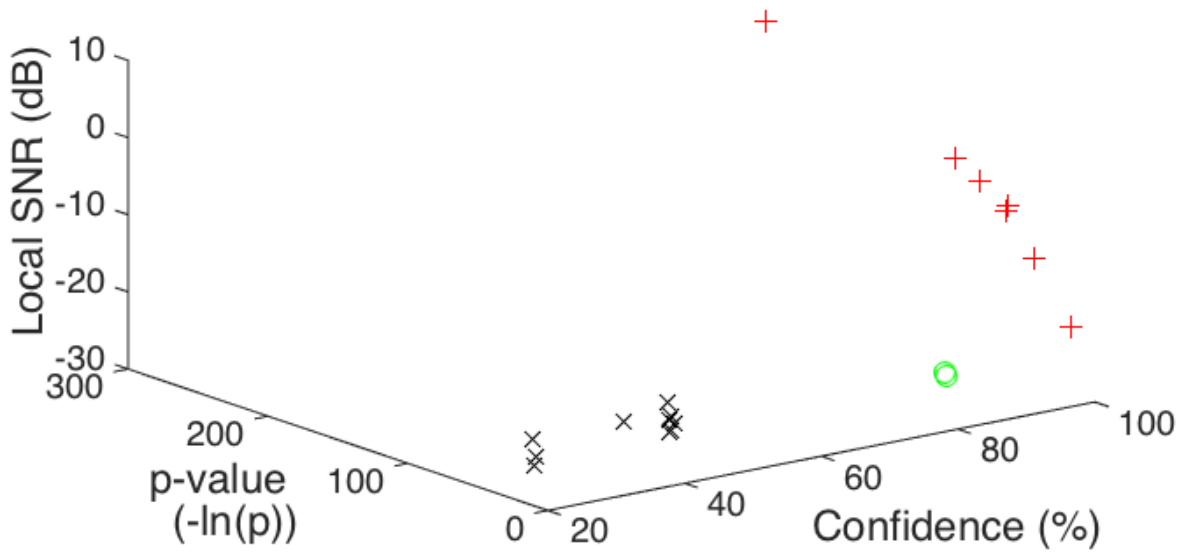


Figure 14: Graphic view of the fusion results on simulated signal described in Table 1: the 19 sequences are plotted as a function of confidence index, p -value and SNR_{ℓ} : red “+” for high confidence and low p -value, green “o” for high confidence but higher p -value, black “x” for low confidence.

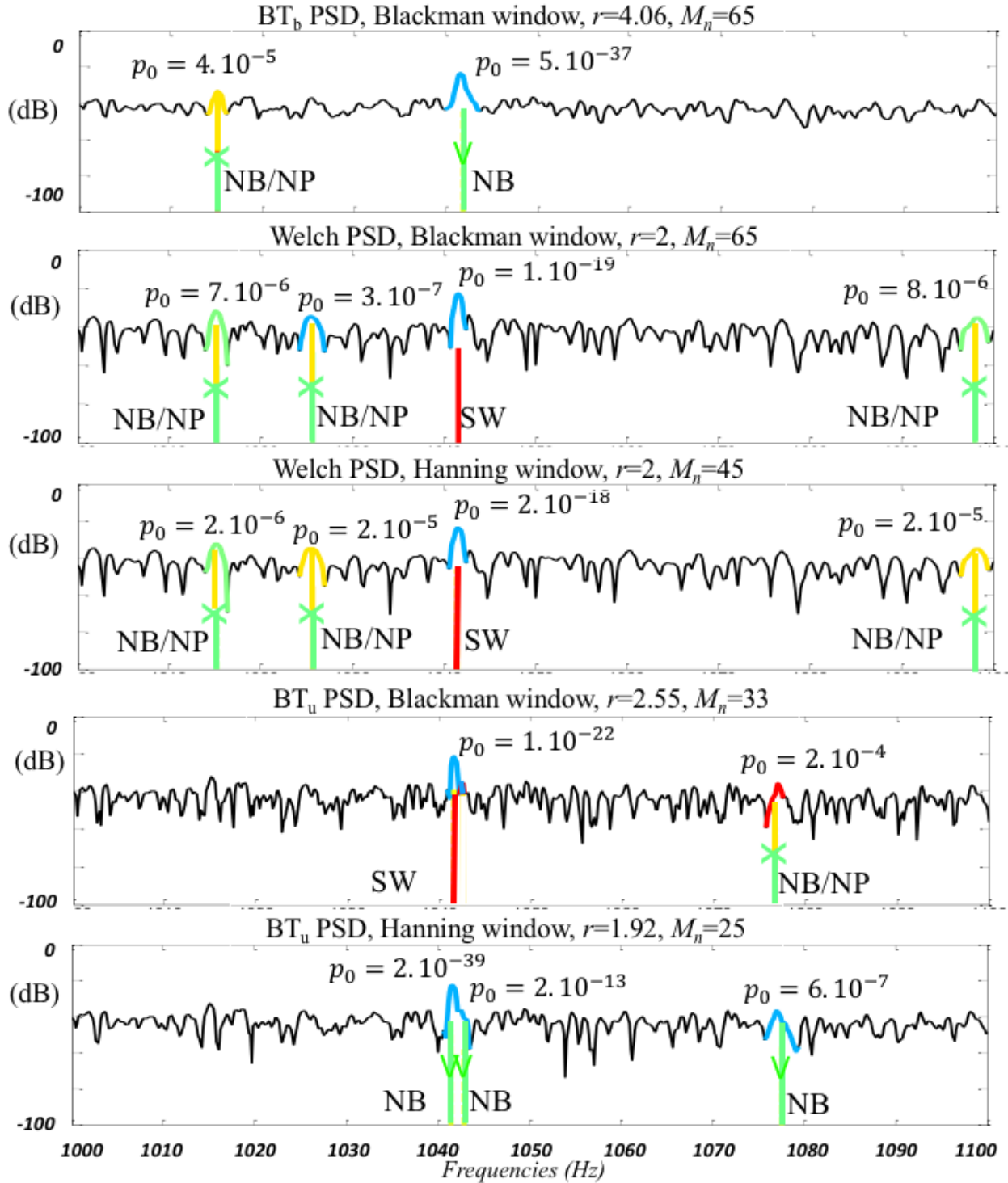


Figure 15: Results of detection and classification on a domestic heat pump acoustic signal.

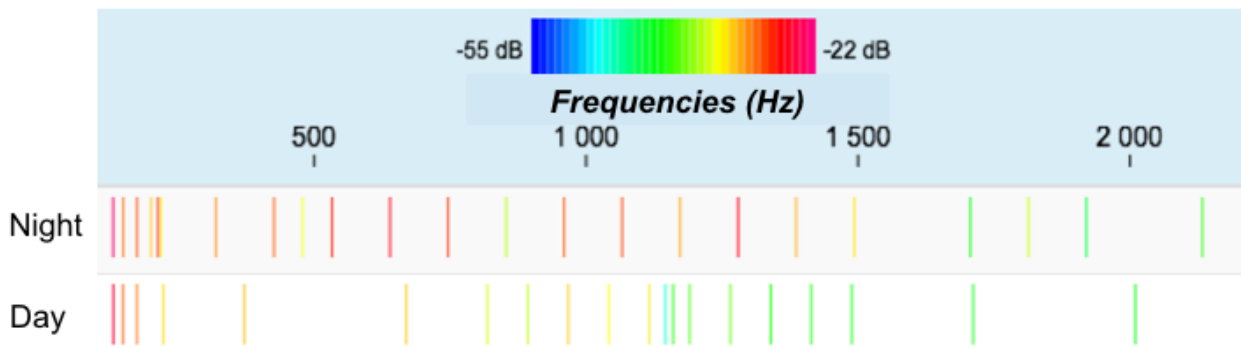


Figure 16: Proposed method in a comparative objective for the monitoring of a domestic heat pump acoustic signal recorded in the middle of the day (below) and in the middle of the night (above). Plot of the frequency attribute of the detected sequences, color scale being the amplitude attribute.

Table 1: Parameters of the sine signal with additive non-white ARMA noise.

Normalized frequencies	Amplitudes	Local signal to noise ratios (dB)
0.07	1.2	-20.9
0.10	5	-7.2
0.11	5	-6.1
0.19	2	-6.6
0.32	10	4.
0.41	0.8	-14.5
0.42	0.9	-13.2

<p>ARMA noise parameters AR parameters: $\{1, -1.3552, 1.228, -0.662, 0.25\}$ MA parameters: $\{1, -0.9, 0.81\}$</p>
--

Table 2: Fusion results on simulated signal described in Table 1: display of 8 attributes of the final identity cards.

Frequency	Amplitude	SNR $_{\ell_j}$	SNR $_{+j}$	Rel. Bandwidth	p -value	Class	Confidence
$\tilde{\nu}_{\max_j}$	A_j	(dB)	(dB)	\tilde{B}_j	p_{0_j}	C_j	CI $_j$
		(22)	(27)	(33)	(12)	(31)	
0.32	9.78	4.93	32.66	1.29	4.28E-103	SW	100%
0.11	5.17	-4.43	24.9	1.04	3.51E-44	SW	
0.19	1.92	-6.34	22.78	0.89	2.48E-36	SW	
0.42	0.86	-9.04	19.29	1.29	4.42E-28	SW	
0.1	5.11	-8.35	20.99	1.04	7.50E-28	SW	
0.41	0.84	-13.95	14.09	0.86	1.13E-19	SW	
0.07	1.63	-21.42	7.92	0.86	3.98E-08	NB	
0.0587	1.00	-23.29	4.75	0.86	7.91E-05	NB/NP	80%
0.1557	0.19	-23.64	5.70	0.86	3.48E-04	NB/NP	
0.1829	0.21	-24.81	4.53	1.5	8.56E-20	NB	40%
0.1257	0.56	-20.22	7.50	1.19	5.01E-06	SW/NP	
0.3863	0.45	-22.33	5.39	2.38	2.88E-05	NB/NP	
0.3343	0.56	-24.06	5.27	1.19	3.32E-05	NB/NP	
0.3063	0.44	-23.68	5.65	0.86	6.36E-05	NB/NP	
0.3666	0.49	-21.98	6.35	1.19	1.09E-04	NB/NP	
0.448	0.27	-22.69	5.64	1.19	9.99E-04	NB/NP	
0.094	0.53	-21.47	6.25	0.86	1.92E-05	NB/NP	20%
0.4159	0.23	-24.66	4.67	0.86	9.44E-05	NB/NP	
0.3556	0.31	-23.64	5.70	0.86	2.13E-04	NB/NP	