The Autogram: An effective approach for selecting the optimal demodulation band in rolling element bearings diagnosis

The Autogram: an effective approach for selecting the optimal demodulation band in rolling element bearings diagnosis

Ali Moshrefzadeh[a,*], Alessandro Fasana[a]

[a] *DIMEAS - Politecnico di Torino*
Corso Duca degli Abruzzi, 24 - 10129 Torino, Italy
Ali.Moshrefzadeh@polito.it
Alessandro.Fasana@polito.it

Abstract

Envelope analysis is one of the most advantageous methods for rolling element bearing diagnostics but finding a suitable frequency band for demodulation has been a substantial challenge for a long time. Introduction of the Spectral Kurtosis (SK) and Kurtogram mostly solved this problem but in situations where signal to noise ratio is very low or in presence of non-Gaussian noise these methods will fail. This major drawback may noticeably decrease their effectiveness and goal of this paper is to overcome this problem. Vibration signals from rolling element bearings exhibit high levels of second-order cyclostationarity, especially in the presence of localized faults. The autocovariance function of a 2nd order cyclostationary signal is periodic and the proposed method, named Autogram, takes advantage of this property to enhance the conventional Kurtogram. The method computes the kurtosis of the unbiased autocorrelation (AC) of the squared envelope of the demodulated signal, rather than the kurtosis of the filtered time signal. Moreover, to take advantage of unique features of the lower and upper portions of the AC, two modified forms of kurtosis are introduced and the resulting colormaps are called Upper and Lower Autogram. In addition, a thresholding method is also proposed to enhance the quality of the frequency spectrum analysis. A new indicator, Combined Squared Envelope Spectrum, is employed to consider all the frequency bands with valuable diag-

nostic information and to improve the fault detectability of the Autogram. The proposed method is tested on experimental data and compared with literature results so to assess its performances in rolling element bearing diagnostics.
Keywords: Autogram, Rolling Element Bearing, Cyclostationary, Autocorrelation, Spectral Kurtosis, Kurtogram, Diagnosis, Defect, Threshold, Undecimated Wavelet Packet Transform

---

Nomenclature

AC unbiased Autocorrelation

$BPFI$ Ballpass Frequency Inner Race

$BPFO$ Ballpass Frequency Outer Race

BSF Ball Spin Frequency

CMS Cyclic Modulation Spectrum

CPU Central Processing Unit

CSA Cyclic Spectral Analysis

$CSES$ Combined Squared Envelope Spectrum

$CWRU$ Case Western Reserve University

EES Enhanced Envelope Spectrum

FF Fundamental Frequency

FK Fast Kurtogram

FP Fundamental Period

GCD Greatest Common Divisor

LCM Least Common Multiple

$MODWPT$ Maximal Overlap Discrete Wavelet Packet Transform

REB Rolling Element Bearing

SC Spectral Correlation

SES Squared Envelope Spectrum

SK Spectral Kurtosis

SNR Signal to Noise Ratio

$STFT$ Short-Time Fourier Transform

WPT Wavelet Packet Transform

1. Introduction

Rolling element bearings (REBs) are one of the most used elements in rotating machinery and their failure is the most important cause of machinery breakdowns [1]. Thus, correctly detecting and diagnosing bearing faults at stages prior to their complete failure is of vital importance. It avoids potential catastrophic damage not only to the apparatus but also to the personnel.

As a localized defect develops, either on an inner race, an outer race or a roller part of a bearing, an impact is generated each time the defect is engaged. Consequently, the bearing and the machine structure are excited, in particular at their resonance frequencies [2]. The corresponding vibration signal will comprise all the harmonics of this impact, which repeats almost periodically at a rate dependent on bearing geometry. Investigation of the generated vibrations is indispensable to detect the faults and many methods have been developed to extract the bearing characteristic frequencies from the measured vibrations. Among them, envelope analysis [2, 3], also called high frequency resonance technique, has been used successfully for a long time: a signal is first bandpass filtered in the excited structural resonance frequency band, and then the spectrum of the envelope signal -which contains the desired diagnostic

frequencies- is formed. The main challenge has always been finding the most suitable frequency band for demodulation.

Spectral Kurtosis (SK) has been a significant step to unravel this problem. It is a method which effectively detects the sequence of impulses in a signal and can be used to determine the proper demodulation frequency band in which a signal has the maximum impulsivity.

Antoni proposed two methods to calculate the SK, one based on short-time Fourier transform (STFT) [4] and another one based on filter banks [5]. In the STFT based SK, the aim is to find the central frequency f and the window length N_w which maximize the value of the SK over all possible choices. A coloured 2D map called Kurtogram displays the values of SK for each couple f and N_w. Antoni [5] also developed the Fast Kurtogram (FK), which is based on the multirate filter-bank structure (MFB) to overcome the rigorous but long computation of full Kurtogram. The method is numerically very efficient and suitable for industrial applications, giving almost the same result as the other one. The improved Kurtogram was further proposed by Lei et. al [6] in which the wavelet packet transform (WPT) is adopted to exploit its good capacity in detection of transients from a noisy signal.

Barszcz and Jablonski [7] argue that, when a signal contains relatively strong non-Gaussian noise such as large impulses, the temporal based kurtosis indicator of Kurtogram would fail as the value of kurtosis, in general, decreases when the transients repetition rate increases. To overcome this drawback, they consequently propose Protrugram, where the kurtosis of the envelope spectrum is displayed, instead of kurtosis of the filtered time signal. However, their proposed method uses a fix demodulation bandwidth and therefore, unlike SK, is not blind and prior knowledge about the defect frequencies is needed. Enhanced Kurtogram was proposed by Wang et. al [8] in which, similar to Protrugram, kurtosis of envelope spectrum of the filtered signal is employed as an indication of faults presence. Nonetheless, both Protrugram and Enhanced Kurtogram may fail when defect frequencies are not dominant in the spectrum in comparison to discrete components generated by other sources such as shaft frequency.

To capture the signature of repetitive transients in both time and frequency domains, Antoni [9] proposed the infogram which combines the concepts of the Kurtogram and Protrugram. In this case, kurtosis is replaced by negentropy, time and frequency domains infograms share the same units and are therefore additive. A review and compendium of more research on the SK theories can be found in Ref. [10].

Cyclic Spectral Analysis (CSA) provides a different family of REBs diagnosis techniques which represent a signal in a bi-spectral map, with carrier and modulation frequencies. It has been brought to the field of diagnostics of mechanical systems by Antoni [11, 12] but it did not get the attention it deserves due to, presumably, its advanced theory and high computational cost. Spectral Correlation (SC) [12] and Cyclic Modulation Spectrum (CMS) [13] are two key approaches for CSA. The latter was introduced as a computationally efficient alternative to SC but this advantage is lessened by its inferior statistical properties. The link between SC and CMS, i.e. the estimation of SC by using CMS algorithm, was presented by Ref [14]. Although it is very fast, it still suffers from the same drawback as CMS. The problem was recently solved by Ref. [15] which describes a fast estimator of SC (Fast-SC) by generalization of the CMS algorithm.

The paper is organized as follows. Theoretical background is discussed in section 2. Section 3 establishes a new method for optimal demodulation band selection for non-stationary signals containing repetitive transients, e.g. bearing with localized defects. A comprehensive explanation of the proposed method is presented in section 4. In addition, experimental validation is carried out and the outcome is compared with the FK [16], Fast-SC [15] and literature results [17] to examine the performance of the new proposed method. Finally, the conclusions are drawn in section 5.

2. Theoretical Background

2.1. Spectral Kurtosis and Kurtogram

A brief description of the Spectral Kurtosis and Kurtogram is presented in this subsection, and in the next section the proposed method is introduced.

Kurtosis is a statistical indicator which measures the peakedness of a data set, therefore it can be used to detect faults impulsiveness in signals related to rotary machinery. Kurtosis is defined as

$$\text{Kurtosis} = \frac{\sum_{i=1}^{N} \left(x(t_i) - \mu_x\right)^4}{\left[\sum_{i=1}^{N} \left(x(t_i) - \mu_x\right)^2\right]^2} \tag{1}$$

where $x(t_i)$ is the sample at time $t_i = i/f_s$, f_s is the sampling frequency, μ_x and N are the mean value and length of the data set. The SK is an extension of kurtosis notation to frequency domain where a band is found to demodulate the signal in order to extract its impulsive and non-stationary component. It can be represented by the fourth-order normalized cumulant [4]:

$$K_x(f) = \frac{\left\langle |Y(t_i, f)|^4 \right\rangle}{\left\langle |Y(t_i, f)|^2 \right\rangle^2}, \tag{2}$$

where $\langle\rangle$ is the time-averaging operator. $Y(t_i, f)$ is the short-time Fourier transform (STFT) of a signal $x(t_k)$ obtained at time t_i by moving a constant length window (N_w) along the signal. Therefore, SK is a function of frequency and STFT window length (frequency resolution). Since SK strongly depends on the window length, Full Kurtogram [4] was introduced as the representation of SK for all possible combinations of frequency and N_w. Unfortunately, the procedure is very time consuming and not suitable for on-line industrial applications.

FK [5] was lately proposed to overcome this difficulty. It subdivides the bandwidths into rational ratios that allow the use of fast multirate processing, and then investigates the kurtosis value of the complex filtered signal for each bandwidth. Dyadic and 1/3-binary tree structure are frequently used for designing these band-pass filters. SK is optimal in comparison to FK due to its finer resolution, however two methods generate very similar result.

2.2. Cyclostationarity

Vibration signals are categorized as deterministic, random or a combination of both. Deterministic signals are further considered as periodic and non-periodic, and random signals as stationary and non-stationary. Some processes are not generally periodic functions of time yet are fundamentally generated by a hidden periodic mechanism. Many processes in mechanics emerge form periodic phenomena such as rotation and reciprocation of gears, bearing, belts, chains, shafts, propellers, pistons, and so on. Signals of this kind are called cyclostationary and their statistical properties change cyclically with time. The capacity of traditional signal processing can then be extended to take advantage of these characteristics. A signal is assumed to be cyclostationary of order n when its nth order statistics is periodic, i.e. after passing the signal through any nonlinear transformation including nth power it can be represented by a Fourier series [3, 18].

Randall et al. [19] investigated the bearing signal characteristic in presence of localized faults. The inner race, outer race and rolling elements defects are the most frequent bearing faults. The vibrations produced by the localized faults are a sequence of impulses dominated by the high resonance frequencies of the structure. Slippage of rolling elements and cage introduces some level of randomness [3, 20] in spacing of the impacts and, although it is not larger than a few percent of the rate of bursts repetition, the resulting signals cannot be categorized as a periodic process. Ref. [19] pointed out that a bearing localized fault signal may be modeled as a 2nd order cyclostationary process. A second order cyclostationarity determines processes with a periodic autocovariance function in time.

$$R_{xx}\left(t_i,\tau\right) = \mathbf{E}\{x(t_i-\tau/2)x(t_i+\tau/2)\} \tag{3a}$$

$$R_{xx}\left(t_i,\tau\right) = R_{xx}\left(t_i+T,\tau\right) \tag{3b}$$

where $\mathbf{E}\{\bullet\}$ is the expected value operator. Autocovariance is a function of not only the time lag τ but also the instantaneous time t_i and it should not

be confused with the stationary autocorrelation function calculated as the time average [3].

Ref. [20] argues that signals from a localized fault are not truly cyclostationary and the signals are strictly speaking pseudo-cyclostationary as they seem to be cyclostationary but in fact are not. However, they can be treated as cyclostationary in a first approximation as the departure from cyclostationarity may essentially be really slow.

3. Proposed Method

3.1. Autogram

The usage of FK to determine the most impulsive frequency band, followed by envelope analysis of the bandpass filtered signal, has become the benchmark method for bearing diagnostics for years and has accomplished significant results [17]. FK is numerically very efficient and largely diffused so that also in the present work has been selected as a reference method. Kurtogram is commonly capable of detecting localized hidden non-stationarities, even in presence of strong Gaussian noise, but its performance is limited in several conditions, i.e. low signal to noise ratio (SNR) or strong non-Gaussian noise such as randomly distributed impulses [17]. In these cases, FK was found to be ineffective in seeking the transient signal. These circumstances are common in industrial applications as multiple devices such as gearboxes and bearings work alongside in a complex machine. Also, the acquired signal in harsh environment can be extremely affected by external sources.

This paper proposes a new procedure based on unbiased AutoCorrelation (AC) to overcome the restrictions imposed by heavy Gaussian and also non-Gaussian background noise. The proposed method is thought to be sufficiently general for detection of rotary machinery faults with impulsive signal, e.g. REBs and gears. Similar to FK, the new method is blind and no prior knowledge of signals is required. The flowchart of the proposed method is shown in Figure 1 and the details of each step are described as follows.

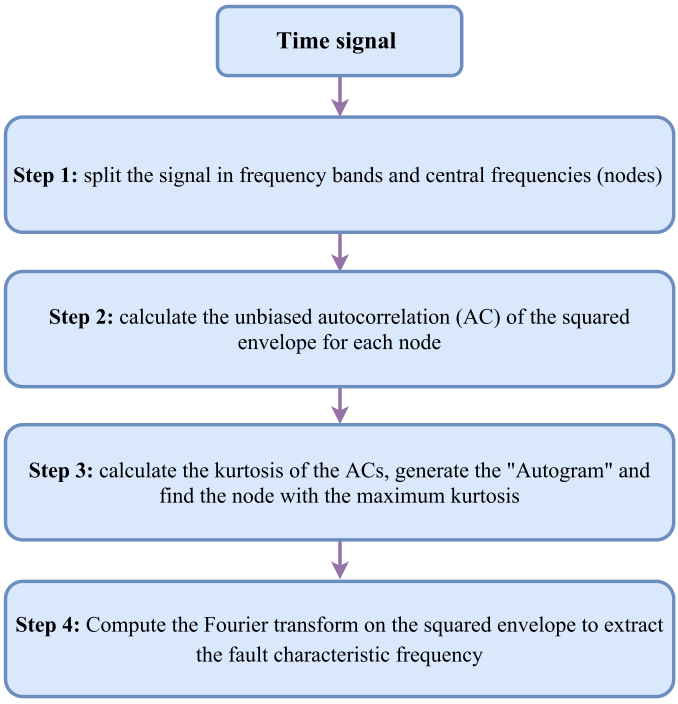

Figure 1: Flowchart of the proposed method

Step 1: In this first step, time domain data are divided in frequency bands, according to a dyadic tree structure, by means of the wavelet transform (WT). Wavelets have very good local properties in both time and frequency domains, and WT can be used as an effective filter to split a signal in different frequency bands and central frequencies [6]. However, due to downsampling operations, the length of the time history halves at each level of decomposition, which can limit the ability to investigate the WT coefficients. Furthermore, the transform can be sensitive to the selection of starting point of the time series, i.e. a change in the starting point can produce rather different outcomes [21]. These drawbacks can be overcome by using the maximal overlap (undecimated) discrete wavelet packet transform (MODWPT) which removes the downsampling step in discrete wavelet packet transform (DWPT). Details can be found in Ref. [21].

Basically, the MODWPT is applied as a filter to the investigated time history and a series of signals is consequently produced at each level of decomposition. The filtered signals, each corresponding to a frequency band and central frequency (node), are the inputs for the following Step 2.

This division in frequency bands is similar to the decimated filterbanks used by the FK [5], although MODWPT is computationally more expensive as it does not make use of multirate processing. Unquestionably, the procedure is not as fast as the FK and a comparison of the required CPU (Central Processing Unit) times is presented in subsection 4.7. MODWPT has been chosen because, besides overcoming the DWPT drawbacks, it preserves full-time resolution [21] which is essential for the proposed method. But other filters may be used, provided that the length of the resulting time history remains unchanged. It should be noted the comparison of the MODWPT filter with other filters, e.g. STFT or the filternabks used in Ref. [5], is out of the scope of this paper. A comparison of the transfer functions of FIR filters and WPT may be found in Ref. [5].

Step 2: The fundamental concept which motivates this work is to take advantage of periodicity of the autocovariance function which characterizes the 2nd order cyclostationarity of bearing vibration signals (subsection 2.2). Therefore,

the unbiased AC of the (periodic) instantaneous autocovariance of the signal $R_{xx}(t_i, 0)$ is calculated, where x is the signal filtered by MODWPT at Step 1. It can be seen from Equation 3a that the instantaneous autocovariance is computed by the ensemble average operator. Unfortunately, in many situations the expected value cannot be obtained since only a single record of data rather than a set of records is available. Once cyclostationarity is assumed, other features of the signal, e.g. the envelope function, provide similar information on periodicity as the instantaneous autocovariance [18]. A simple example is provided by the white noise with periodical amplitude modulation, shown in Figure 2a. Figure 2b and Figure 2c (left columns) illustrate the instantaneous autocovariance $R_{xx}(t_i, 0)$ and the envelope of the signal. The frequency representation of the signal (Figure 2a, right) does not provide any useful information. On the other hand, the modulation frequency can be clearly detected in the very similar Figure 2b and Figure 2c (right). Note that the ensemble average in Figure 2b is produced by using 200 realizations.

In this step, unbiased AC is computed on the squared envelope of the signal as follows:

$$\hat{R}_{XX}(\tau) = \frac{1}{N-q} \sum_{i=1}^{N-q} X(t_i) X(t_i + \tau) \tag{4}$$

where X is the squared envelope of the filtered signal, $\tau = q/f_s$ is the delay factor and $q = 0, ..., N-1$.

The AC has the benefit of removing the uncorrelated components of the signal, i.e. noise and random impulsive contents, both unrelated to any specific bearing fault. Furthermore, the periodic part of the signal (directly related to the defects) is enhanced, showing an additional virtue of this process. This is even more advantageous since it is done for each node separately rather than on the complete raw signal, so that SNR for each demodulated band signal is increased.

With τ increasing, the number of data samples for computing the AC will decrease (Equation 4), and therefore the resulting part will not have an adequate estimation variance. As a result, only a part of the computed AC is chosen for

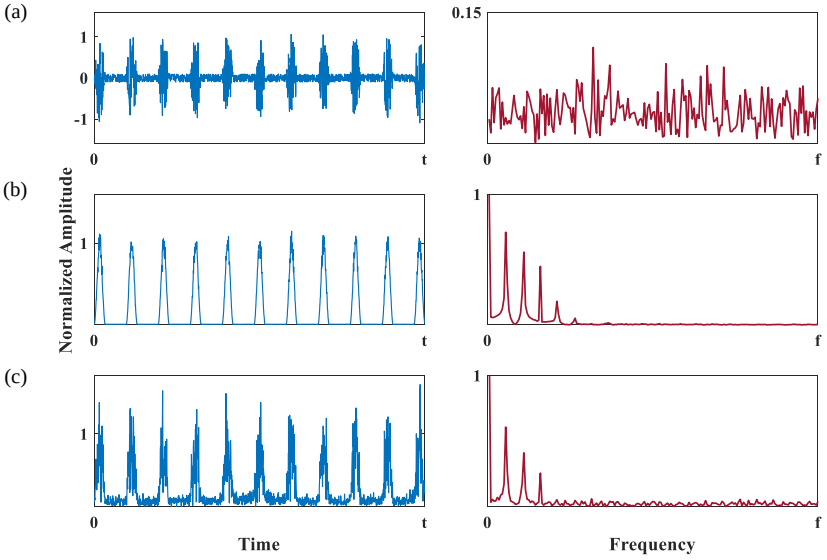

Figure 2: Example of (a) amplitude modulated white noise and its spectrum (b) instantaneous autocovariance $R_{xx}(t, 0)$ and its spectrum (c) envelope signal and its spectrum. Generic engineering units may be associated to the data

further investigation: unless stated otherwise, throughout this paper the first half of the AC is selected. Moreover, the first coefficients of filtered signal, that are affected by the transients of filters, should not be included in calculation of the AC [5].

The output of this step is then the AC which leads to a more accurate diagnostic process than possible with the original outputs of MODWPT. In fact, impulsive noise, which ineffectively assigns very high kurtosis to a signal, can largely be removed.

Step 3: The objective of this step is to find the most suitable frequency band for demodulation. This is substantial to have a successful diagnosis of bearings faults, since fault information cannot be extracted from the demodulated signal if the appropriate frequency band and central frequency are not selected. In this regard, FK and Protrugram are two widely known methods which compute kurtosis of the filtered time signal and spectral lines of the enve-

lope spectrum respectively. In this paper, an alternative approach is introduced to achieve an optimal frequency band of demodulation. The proposed method differs from both mentioned techniques because the kurtosis is evaluated for the signals resulting from step 2, i.e. the unbiased AC of the squared envelope, for each level and frequency band (nodes). Subsequently, the kurtosis values of all nodes, similar to FK, are presented in a colormap, whose color scale is proportional to kurtosis value and the vertical and horizontal axis represent level of the MODWPT decomposition and frequency respectively. Since the concept is analogous to Kurtogram, and this proposal is based on autocorrelation, the authors suggest the name of "Autogram" for this newly developed approach.

To quantify the impulsivity of the AC for each node, three equations which are modified versions of the kurtosis (Equation 1) are proposed as follows:

$$\text{Kurtosis}(X) = \frac{\sum_{i=1}^{\frac{N}{2}} \left[\hat{R}_{XX}(i) - \min(\hat{R}_{XX}(\tau)\right]^4}{\left[\sum_{i=1}^{\frac{N}{2}} \left[\hat{R}_{XX}(i) - \min(\hat{R}_{XX}(\tau))\right]^2\right]^2}, \tag{5a}$$

$$\text{Kurtosis}_u(X) = \frac{\sum_{i=1}^{\frac{N}{2}} |\hat{R}_{XX}(i) - \overline{X}_T(i)|_+^{\ 4}}{\left[\sum_{i=1}^{\frac{N}{2}} |\hat{R}_{XX}(i) - \overline{X}_T(i)|_+^{\ 2}\right]^2}, \tag{5b}$$

$$\text{Kurtosis}_l(X) = \frac{\sum_{i=1}^{\frac{N}{2}} |\hat{R}_{XX}(i) - \overline{X}_T(i)|_-^{\ 4}}{\left[\sum_{i=1}^{\frac{N}{2}} |\hat{R}_{XX}(i) - \overline{X}_T(i)|_-^{\ 2}\right]^2}, \tag{5c}$$

where N is length of the original signal and operators $|\blacksquare|_+$ and $|\blacksquare|_-$ mean that only positive or negative values are accepted and values of other data points are set to zero. Also, $\overline{X}_T$ is the threshold level and it is defined as the moving mean value of the AC:

$$\overline{X}_T(i) = \frac{1}{k} \sum_{j=i}^{i+k-1} \hat{R}_{XX}(i) \tag{6}$$

k is length of the windowed signal to be averaged, typically a very small fraction of the total number of samples N.

Equation 5a is analogous to the standard definition, except that the minimum is used instead of the mean value. The AC of the squared envelope is in fact a positive function, whose minimum value is arguably different from zero.

Colormap presentation of the results based on the Equation 5a-c will be called Autogram, Upper Autogram and Lower Autogram correspondingly. The advantages and characteristics of each proposed Autogram and the condition in which using Upper/Lower Autogram is beneficial will be discussed in detail within the first example of section 4.

Ultimately the signal associated with the node with the highest kurtosis is considered for further investigation.

Step 4: The Fourier transform is finally applied to the squared envelope of the signal associated to the node selected in Step 3. The fault characteristic frequencies are extracted and a diagnosis of the bearing is performed.

3.2. Lower/Upper threshold and Squared Envelope Spectrum

As a consequence of step 2 of the proposed method, the level of uncorrelated signals such as noise and random impulses is reduced and the distinction between two parts of the signal (noise and defect impulses) has become more clear. Periodic impulses - corresponding to defective bearing - are in fact enhanced at step 2 and can be separated from noise more effectively. This gives the opportunity to separate the two parts without losing any useful information for diagnosis of bearings, as the main interest is in repetition frequency of the peaks corresponding to bearings defects, which is conserved after this process. To achieve this objective, in this step a thresholding procedure is introduced and performed on the resulting coefficient of step 2 (AC of the envelope signal), for the node with the highest kurtosis selected in step 3.

In this paper, the non-constant $\overline{X}_T$, defined in Equation 6, is used as the threshold level. The Lower/Upper thresholding process extracts the significant information by setting to threshold level the AC values with higher/lower than the threshold level.

This process directly affects the quality of the frequency analysis as it controls which coefficients will be retained and which will be discarded. The benefit of this step is not merely limited to suppression of noise, since both lower and upper part of the squared envelope signal after performing AC has some unique

and virtuous features which will be discussed with the help of the first example in the section 4.

3.3. Combined Squared Envelope Spectrum (CSES)

A major problem related to all methods which try to find the best frequency band and center frequency for demodulation, such as Autogram or Kurtogram, is that other nodes which may contain useful information are neglected. It is more problematic especially in cases with multiple defects on different bearings as they may have different carrier frequencies, or when a defect on a bearing excites different resonance frequencies. To overcome this difficulty, we propose an approach to combine the outcomes of the all nodes with valuable diagnostic information. The steps of this procedure can be explained as follow:

1. For each level of decomposition, the nodes with valuable diagnostic information are selected. It is achieved by selecting the node with highest kurtosis in each level, in addition to the nodes with kurtosis larger than a certain threshold. In this paper, the threshold for each level is selected equivalent to half of the maximum kurtosis of that level.
2. For each level of decomposition, the squared envelope spectrum (SES) is computed for all nodes selected in previous step. These spectrums are normalized within 0 and 1, then they are combined. Hereafter, this spectral quantity is called "Combined Squared Envelope Spectrum" (CSES):

$$\mathrm{CSES}(level) = \sum_{i=1}^{n_{level}} \mathrm{SES}(i, level) \tag{7}$$

where n_{level} is the number of the nodes selected in the previous step for each level.

4. Results and Discussion

The data sets provided by the Case Western Reserve University (CWRU) bearing data center [22] has become a standard reference for diagnosis of bearings. For example, a detailed benchmark study has been provided by Smith

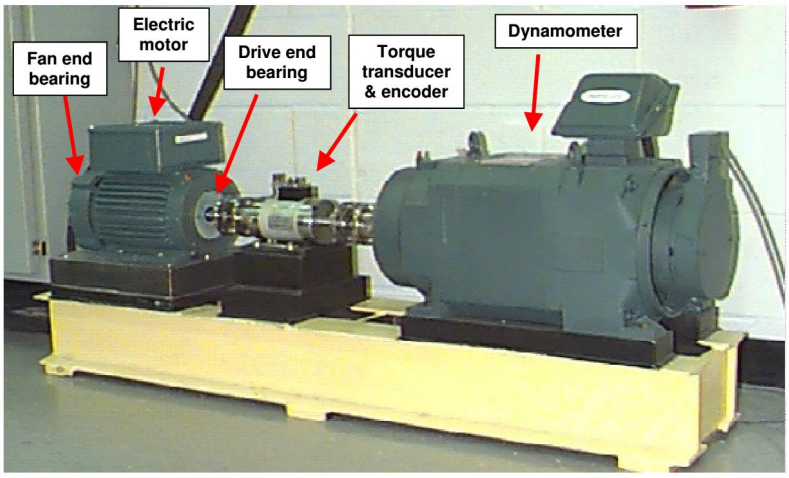

Figure 3: CWRU bearing test apparatus [17, 22]

and Randall [17] in which three diagnostic methods such as envelope analysis of the raw signal, cepstrum pre-whitening and discrete/random separation (DRS) followed by SK and bandpass filtering were applied to all the data sets. Therefore, these data will be used to examine the performance of the proposed method. The results will be compared with the benchmark study, FK and Fast-SC, whose codes have been provided by [16, 15]. The comparison makes it possible to properly evaluate the proposed method.

The benchmark study characterized the result of its diagnosis in the same six categories employed throughout the following section - Table 1.

The basic experimental setup is shown in Figure 3. The test rig includes a 2 hp electric motor, a dynamometer which provides the load and a torque transducer/encoder. Two test bearings, namely drive-end and fan-end bearings, support the motor shaft and a single localized fault is introduced on inner race, outer race or rolling element. The vibration data were acquired for about 10 seconds with sampling frequencies of 12 and 48 kHz for each case. More details on bearings and faults specification, can be found in Refs. [22, 17].

Note that signal pre-whitening and/or separation of deterministic and ran-

Table 1: Categorisation of diagnosis outcomes Diagnosis [17]

Diagnosis category	**Diagnosis success**	**Explanation**
Y1	Yes	Data clearly diagnosable and showing classic characteristics for the given bearing fault in both the time and frequency domains
Y2	Yes	Data clearly diagnosable but showing non-classic characteristics in either or both of the time and frequency domains
P1	Partial	Data probably diagnosable; e.g., envelope spectrum shows discrete components at the expected fault frequencies but they are not dominant in the spectrum
P2	Partial	Data potentially diagnosable; e.g., envelope spectrum shows smeared components that appear to coincide with the expected fault frequencies
N1	No	Data not diagnosable for the specified bearing fault, but with other identifiable problems (e.g., looseness)
N2	No	Data not diagnosable and virtually indistinguishable from noise, with the possible exception of shaft harmonics in the envelope spectrum

dom signals is usually conducted on the raw signal before the diagnosis. The aim is to reduce the effects of the deterministic components (discrete frequency "noise") such as gears signals and therefore increase the bearings SNR. To this purpose Refs. [17, 8] recommend discrete/random separation (DRS) and autoregressive (AR) model methods respectively, while some other techniques are discussed in Ref. [23]. Nevertheless, in this paper neither pre-whitening nor separation of deterministic and random parts is performed prior to analysis. If the parameters of the filters (e.g. DRS and AR), which should be set in advance, are not correctly selected, their performances would be improper.

Daubechies wavelets (db12) are employed to decompose the signal in Step 1. Moreover, to have proper minimum frequency band of demodulation for data with sampling frequencies of 12 and 48 kHz, 5 and 6 levels of decomposition are performed respectively, for both FK and Autogram.

Note that in the envelope spectra throughout this section, green dash-dot lines cursors are depicted at the nominal shaft frequency, red dashed lines at harmonics of the expected fault frequency, and red dotted lines show the first order modulation sidebands around the fault frequency and its harmonics. These sidebands are spaced at shaft and cage frequencies for defects on the inner race and rolling element respectively. For outer race faults there will be no sidebands because the outer race is fix and as a result the transmission path between the defect and the signal acquisition point does not vary.

This section is divided in two parts. The purpose of the first part is to explain in deeper details the steps described in subsection 3.1. Additionally, the performance of the Upper, Lower and original Autogram and the effects of different proposed thresholding will be illustrated. In the second part the CSES results will be presented and a comparison with Fast-SC will be shown.

4.1. Case 1: investigation of different Autograms and thresholding process

4.1.1. Autogram

In this case, record 176 FE with an inner race defect is examined. The vibration signal of the record is plotted in Figure 4a. Defect repetitive transients

are visible in the time waveform and the fault is categorized as P1, Y2 and P2 by methods 1, 2 and 3 of the benchmark study [17] respectively.

FK is depicted in Figure 4b and two frequency bands, which are both related to the defect, can be detected in the plot. The center frequency 9187.5 Hz with bandwidth 375 Hz has the highest kurtosis. The SES of the filtered signal, selected by the FK, is displayed in Figure 4d. Although it is categorized as P2 in the benchmark, the strongest component in the spectrum is indeed the ballpass frequency inner race (BPFI). This misclassification might be a result of the DRS filter or insufficient maximum level of decomposition, since the latter is not mentioned in the paper.

The proposed method is also applied to the same signal. The Autogram is shown in Figure 4c and the maximum value is found with center frequency 6750 Hz and bandwidth 1500 Hz, at node (4,5). The SES of the signal related to node (4, 5) is presented in Figure 4e. An extra defect frequency band, in addition to the couple in the FK, is present in the Autogram with very high frequency.

Figure 5a illustrates the squared envelope of the filtered signal associated to the mentioned node (4, 5), while its AC is depicted in Figure 5b (to have a closer look, only short portions are presented). The filtered signal has high SNR and the uncorrelated part, Gaussian noise in this case, is canceled in the AC. As a result, almost all the defect pulses are evidently distinguishable. One of the main reason behind proposition of the Autogram is to utilize this very beneficial feature. This is the reason why the Kurtogram is not able to detect the high frequency band as clearly as the Autogram does (see Figure 4b and c). The defect pulses are mixed and masked by high frequency noise, and therefore the filtered signal has low kurtosis.

Bearing signals are usually modulated by other sources such as shaft frequency, and in some cases more than one defect exists in a bearing. Since the carrier frequencies, e.g. bearing defect frequencies, are not usually integer multiple of the modulation frequency, e.g. shaft frequency (f_s), the signal (and mainly its envelope) will also have, in addition to shaft and defect periodicities, a fundamental period (FP) which is different from the carrier and modulation

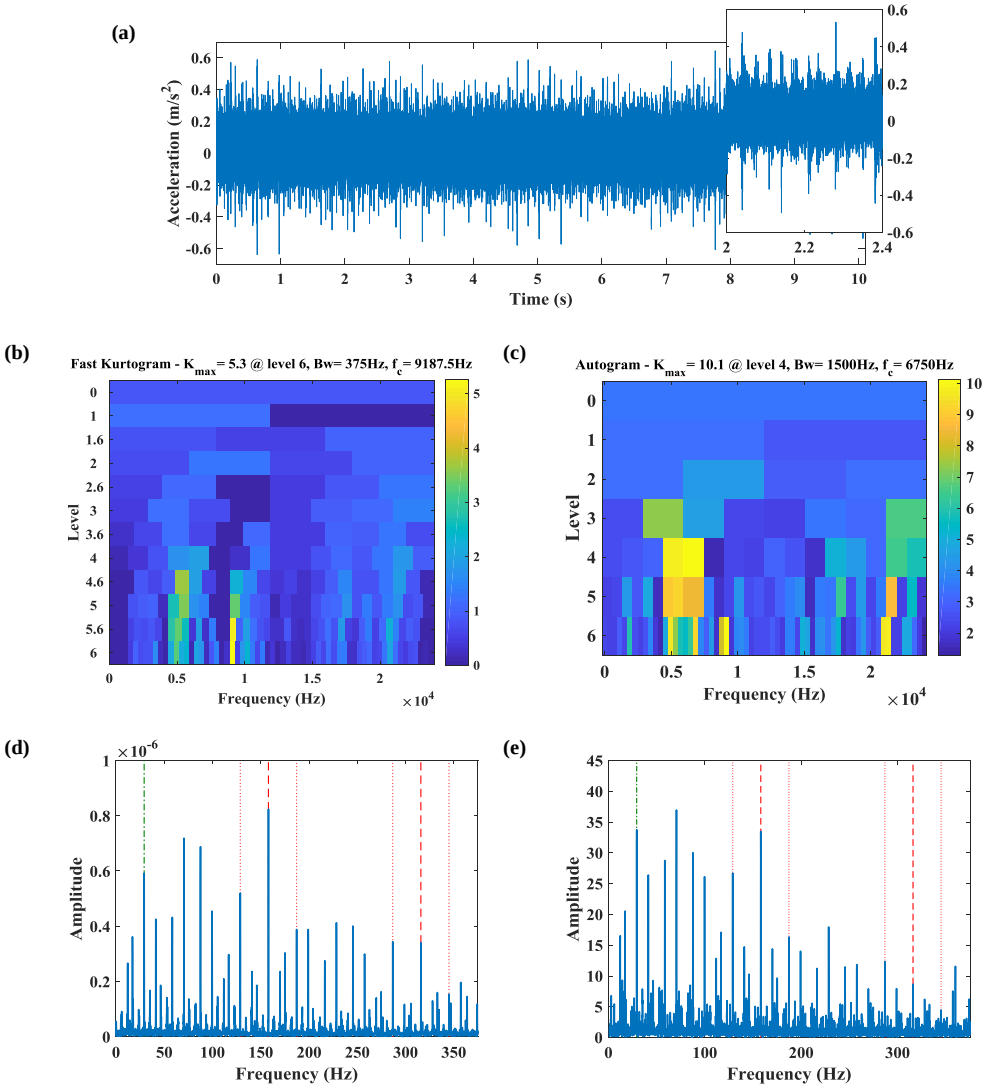

Figure 4: Case 1 (a) Time domain signal: 176 FE (b) FK (c) Autogram, squared envelope spectrum (SES) of the signal related to node with highest kurtosis in (d) FK (e) Autogram (Green dash-dot line: nominal shaft frequency, red dashed lines: first two harmonics of the BPFI, red dotted lines: first order modulation sidebands at shaft speed around the BPFI and its harmonics)

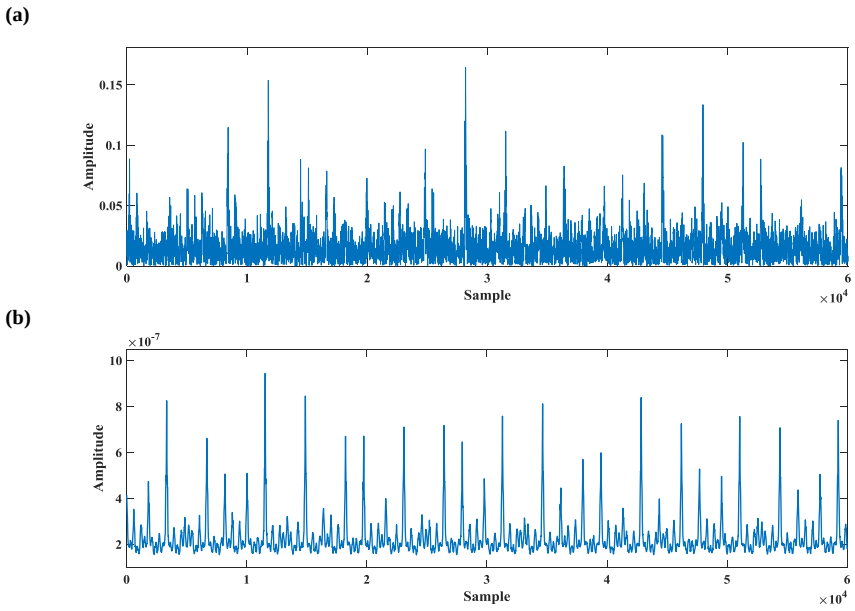

Figure 5: Case 1 (a) squared envelope of the filtered signal associated to the node selected by Autogram and (b) its AC

periodicities.

FP is the *least common multiple* (LCM) of all individual periods of a signal, i.e. the fundamental frequency (FF) is the *greatest common divisor* (GCD) of all frequency components of a signal. When maxima for both modulation and carrier signals occur at the same time, the corresponding peaks, multiplication of these two values, have the largest amplitudes. Their periodicity will be equal to the FP, i.e. the peaks with the same amplitude will have the period equal to the FP. From now on, to distinguish between theses peaks and peaks spaced by the defect period we will refer to them as the "second kind peaks". Similarly, peaks amplitude in the signal's envelope and its AC will manifest the same pattern.

As the distinction between peaks is more clear in the AC than in the envelope signal, it is possible to effectively alleviate the effect of the shaft frequency modulation. This is achieved by using the threshold introduced with equation

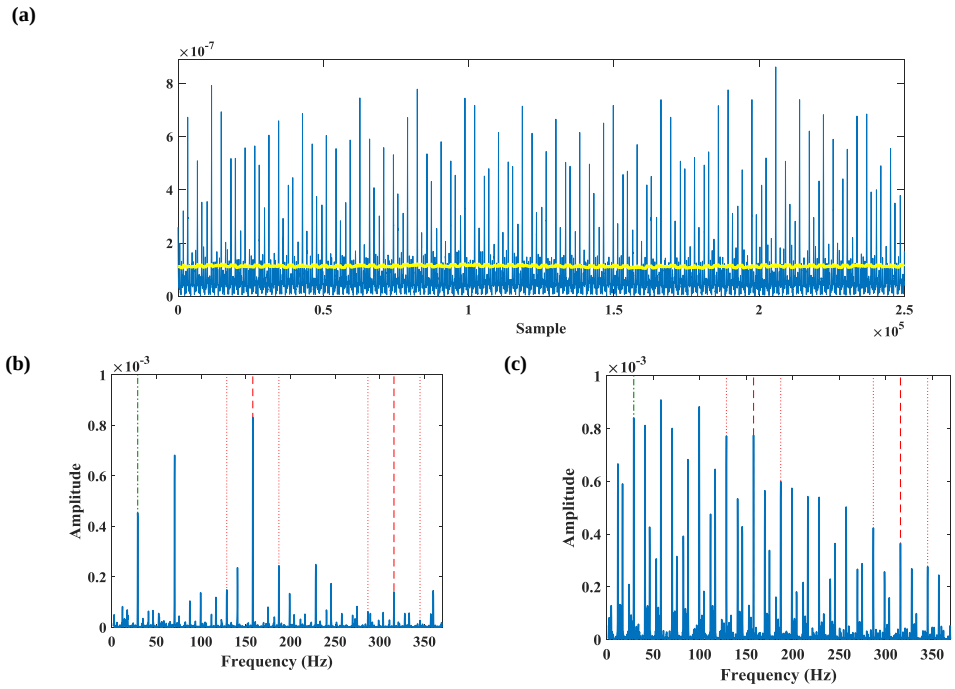

Figure 6: Case 1 (a) AC of the envelope signal (blue line) and the threshold level (yellow line), spectrum of the AC signal after performing (b) lower threshold (c) upper threshold

Equation 6. The AC and the threshold level (yellow line) are depicted in Figure 6a. Figure 6b shows the spectrum of AC signal after implementing "lower threshold" (only samples below the yellow line are selected) where the amplitude of shaft frequency and its harmonic are decreased. On the contrary, the amplitude of shaft frequency and its harmonic as well as the corresponding sidebands around defect frequency are increased after the "upper threshold" (Figure 6c) as the resultant signal mostly contains the "second kind peaks". These characteristics of the lower and upper parts of the AC are one of the main reasons behind the proposed Lower and Upper Autogram which are discussed in the following.

4.1.2. Lower Autogram

Another interesting observation is that for each maximum in the AC there is a corresponding minimum and, in contrast to maxima, minima are not affected by the modulation sources. To exploit this property, Lower Autogram is proposed in which a lower threshold is applied on the AC of the envelope signal and then Autogram is computed for the modified signal. It is advantageous to employ Lower Autogram when large non-periodic impulses in the signal cause high kurtosis in both FK and Autogram, even after performing AC. However, their effect on kurtosis is less prominent when just the lower part is selected. In the present case, the signal does not contain impulsive noise and the Lower Autogram, which is not depicted here, is almost comparable to the originally proposed Autogram (Figure 4c).

4.1.3. Upper Autogram

As it was discussed earlier, the largest peaks of the AC are spaced by the FP (second kind peaks). Therefore, when the SNR is low for a filtered signal, many defect peaks are masked by noise and mainly the second kind peaks can be extracted. However, since the level of noise is high and only a few peaks are present, the kurtosis value of the signal is very low and a proper frequency band cannot be detected for demodulation. Upper Autogram is introduced in subsection 3.1 to overcome this difficulty, in which an upper threshold is applied to the AC of the envelope signal to remove the noisy part and then Autogram, which is called Upper Autogram, is constructed for the modified signal. Consequently, the value of kurtosis for each frequency band which mainly includes second kind peaks will rise drastically as the lower part is removed.

The Upper Autogram for record 176 FE is shown in Figure 7a and the maximum value is assigned to the node (5, 24), with center frequency 17625 Hz and bandwidth 750 Hz. Even though this frequency band comprises valuable diagnostic information it cannot be detected by neither FK nor Autogram.

The squared envelope of the filtered signal is displayed in Figure 7b and its AC and the threshold level (yellow line) are depicted in Figure 7c. After

performing AC, the level of noise is reduced, the defect pulses are enhanced and now these peaks are more noticeable. These large second kind peaks are spaced at the envelope's period and other peaks related to defects with considerably lower amplitude can also be spotted. The spectrum of the envelope signal is shown in Figure 7d while Figure 7e presents the AC's spectrum after performing upper threshold, where it can be realized that the level of noise in the spectrum is reduced. Moreover, sidebands are more dominant in this spectrum because the lower part, which mainly contains the unmodulated part of the signal, is removed. Furthermore, since mainly second kind peaks are present after upper thresholding, the spectrum will mainly encompass the fundamental frequency and its harmonics. This phenomenon is greatly advantageous in finding bearing defect frequencies when the SNR is low because even in presence of a few peaks, the existence of the defect frequency could be revealed in the spectrum.

The spectrum of the lower part is plotted in the Figure 7f where the defect frequency dominates the spectrum and the shaft frequency and sidebands are almost negligible.

It must be noted Lower and Upper Autograms should be considered as complementary approaches to the original Autogram, taking into account that they highlight some specific features of the signals.

4.2. Case 2: impulsive noise

In this case, record 275 DE - Figure 8a - with a defect on the inner race is examined. A number of transients are visible in the time waveform and the signal shows some level of non-stationarity. The bearing fault is diagnoseable (P1) by methods 1 and 2 of the benchmark study but method 3 does not determine the presence of the defect (N1) [17]. The reason is that FK is vulnerabile to impulsive noise, i.e. it tends to highlight the presence of individual impulses rather than sequences of transients. Figure 8b illustrates the FK in which the highest kurtosis, for the node with center frequency 3625 Hz and bandwidth 250 Hz, is related to impulsive noise. The SES is plotted in Figure 8d but does not provide valuable diagnostic information.

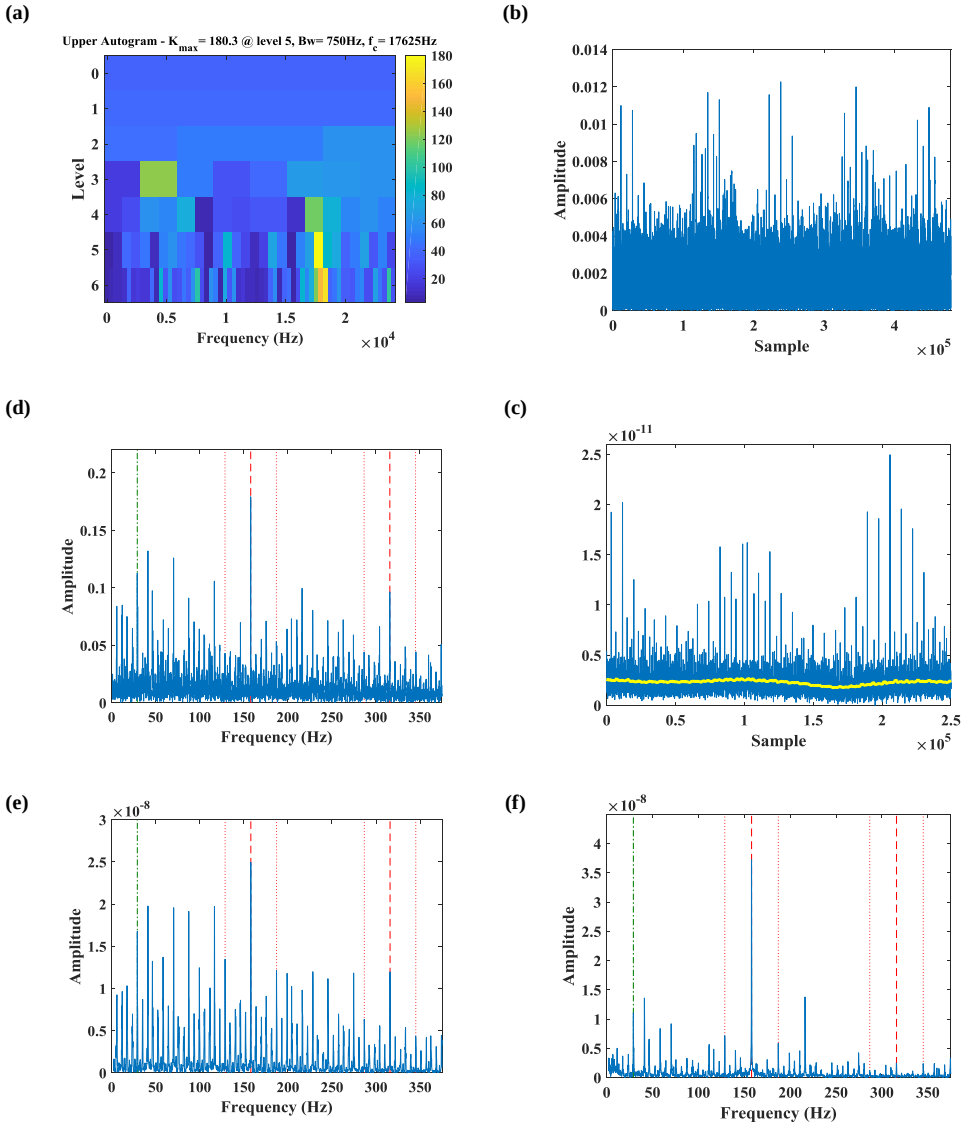

Figure 7: Case 1 (a) Upper Autogram (b) envelope of the filtered signal associate to the node selected by Upper Autogram (c) AC of the filtered signal's envelope and the threshold level (yellow line) (d) SES of the filtered signal (e) spectrum of the upper part (f) spectrum of the lower part

Figure 9a shows the filtered signal for this node which mainly contains impulses belonging to non-stationarity portions of the raw signal. The non-stationarity around 3.8 seconds causes the high value of kurtosis and it is independent from the specified bearing fault.

The Autogram is shown in Figure 8c and the maximum value of the kurtosis (Equation 5) is assigned to node (4, 4), with center frequency 1312.5 Hz and bandwidth 375 Hz. The frequency band selected by FK is not noticeable, which indicates elimination of the impulsive noise in the AC. The SES of the signal for node (4, 4) is displayed in Figure 8e where the harmonics of BPFI can be detected. The filtered signal and its zoomed part are shown in Figure 9b and, in contrast to the filtered signal related to the node selected by FK, a series of transients is present which, based on the spectrum, is indeed related to the bearing fault.

Furthermore, another interesting frequency band can be identified in the right side of the Autogram. In this branch, node (4,12) with center frequency 4312.5 Hz has the highest kurtosis and its spectrum, not reported here for brevity, provides an additional successful diagnosis.

Other data sets provided by CWRU bearing data center, such as 275 BA, 276 DE with inner race defects and 284 BA with rolling element defect (see ref. [24]), have a similar behaviour. As a consequence, they are not diagnosable with method 3 (N1) [17] but are correctly classified by Autogram. Moreover, the defect frequency (ball spin frequency - BSF) could be detected after applying lower threshold for data 292 FE, which is not diagnosable with any of the applied methods in the benchmark study.

These results show the capability of Autogram in dealing with signals containing impulsive (non-Gaussian) noise.

4.3. Case 3: corrupted signal

In this case, record 177 FE with an inner race defect is studied and the vibration data are plotted in Figure 10a. From the time waveform, it can be seen that the record is corrupted with patches of electrical noise around the 0.3-

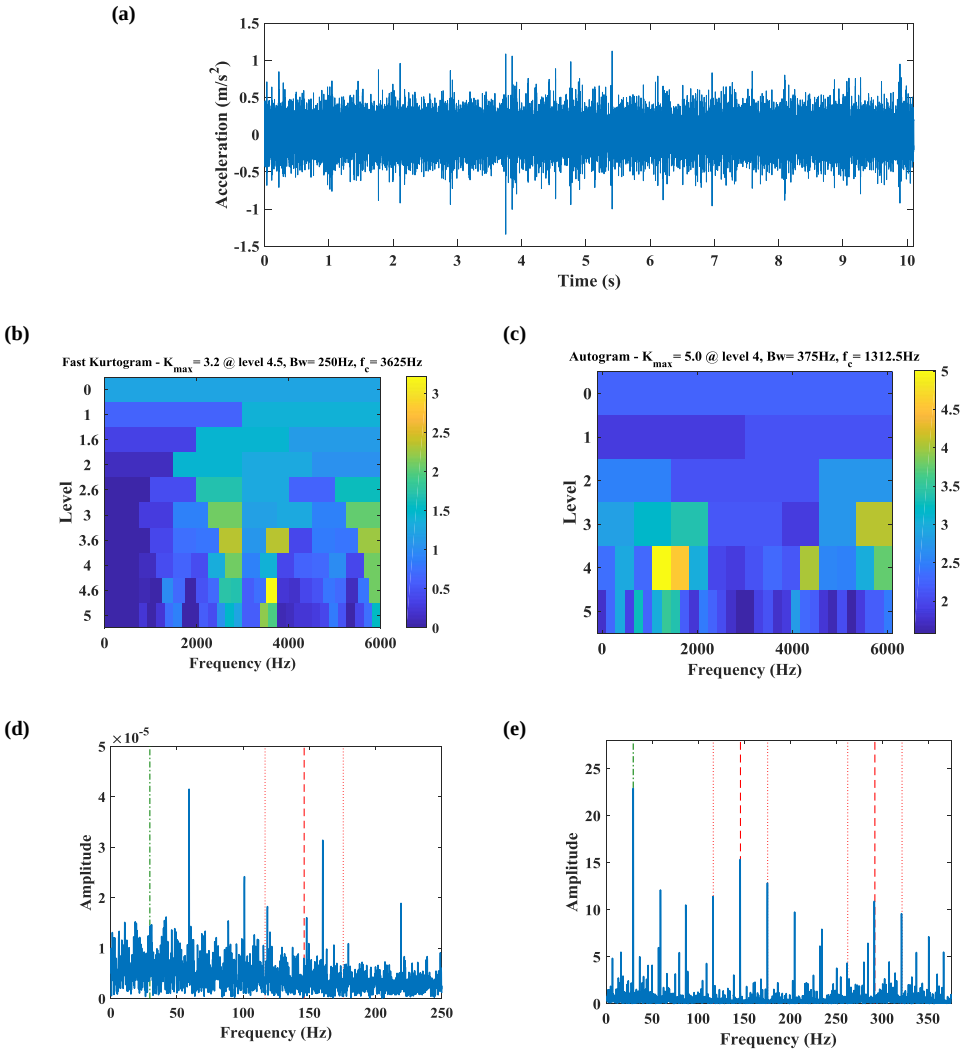

Figure 8: Case 2 (a) Time domain signal: 275 DE (b) FK (c) Autogram, SES of the node selected by (d) FK (e) Autogram

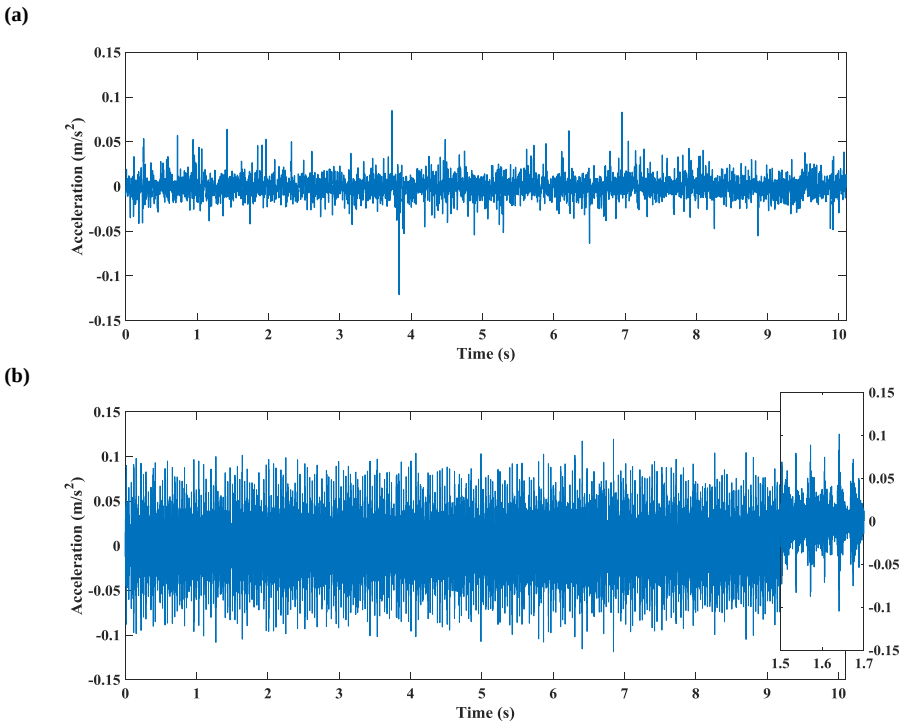

Figure 9: Case 2: Filtered signal associated with the selected node by (a) FK (b) Autogram

0.7 seconds. Many defect repetitive transients are visible in the time waveform and the fault diagnosis result is categorized as P1 and Y2 by methods 1 and 2 of the benchmark study but method 3 failed to find the defect (N1) [17]. Figure 10b indicates that the FK is vulnerable to this high frequency electrical noise and the highest kurtosis is detected at center frequency 23250 Hz with bandwidth 500 Hz.

Although Autogram has a high potential to automatically deal with corrupted signals, a more guided approach is suggested to even increase this capability. When the delay (τ in Equation 4) is very low, the correlation between the currupted part and its delayed version generates high values in the AC. On the contrary, when the delay exceeds the correlation length of the corrupted part with itself, only the relationship between the defect signals will be recognised in the AC. Since AC of the envelope signal is employed to calculate the kurtosis, this provides a flexible tool to select the appropriate part of the AC without losing the defect information. This feature is used to manage the corrupted signal of record 177 FE, disregarding the first 10% of the AC.

The Autogram is shown in Figure 10c and the maximum value is assigned to node (6, 15), with center frequency 9187.5 Hz and bandwidth 375 Hz. The SESs of the filtered signals, selected by the FK and Autogram, are displayed in Figure 10d and Figure 10e respectively. In Figure 10e inner race ballpass frequency (BPFI) and its second harmonic, together with their sidebands at the shaft frequency, can clearly be spotted.

As just discussed, Autogram has the potential to find the proper frequency band of demodulation even when part of a signal is corrupted. Similarly, successful diagnosis can be achieved for data sets 177 DE with inner race defect and 158 BA with outer race defect.

4.4. Case 4: multiple defects

In this case, data set 222 DE is studied. It presumably has a defect on a rolling element but there is often indication of inner/outer race fault for several records [17]. Data are plotted in Figure 11a and reveals a highly non-stationary

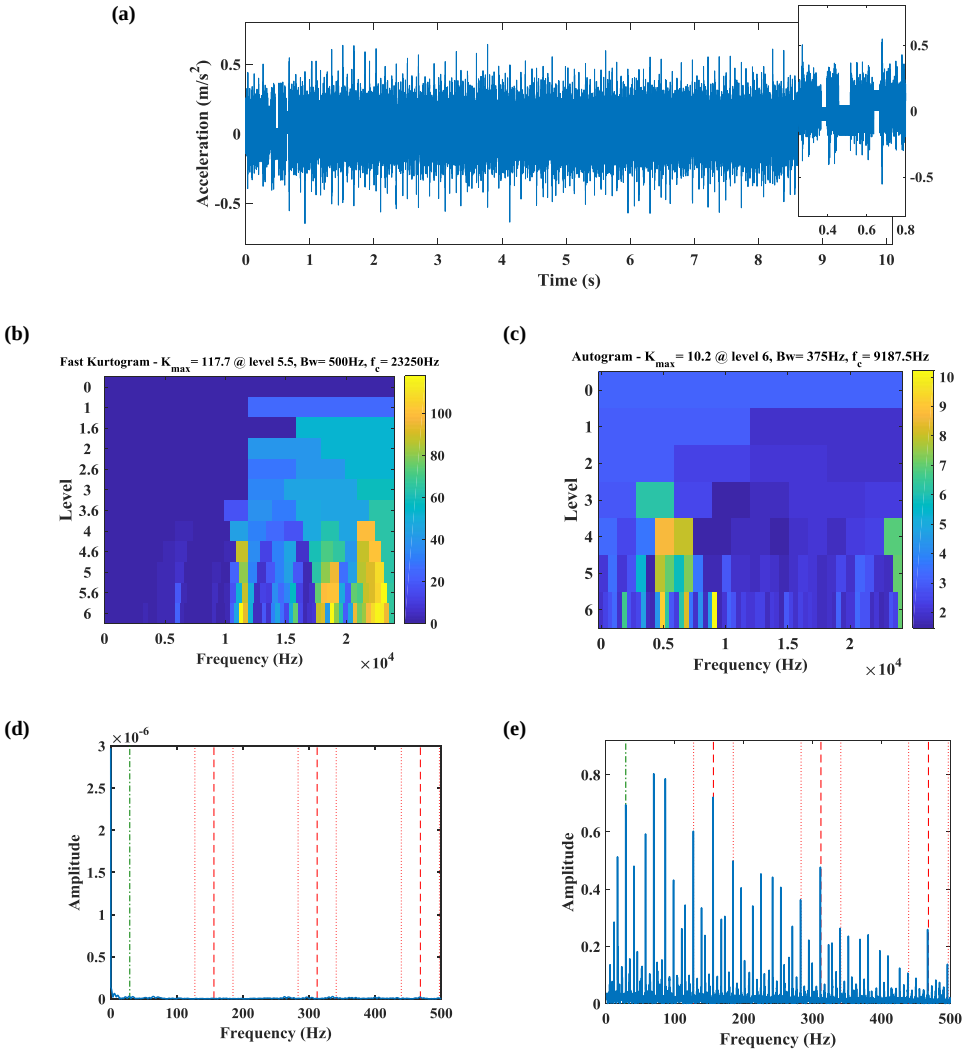

Figure 10: Case 3 (a) Time domain signal: 177 FE (b) FK (c) Autogram, SES of the node selected by (d) FK (e) Autogram

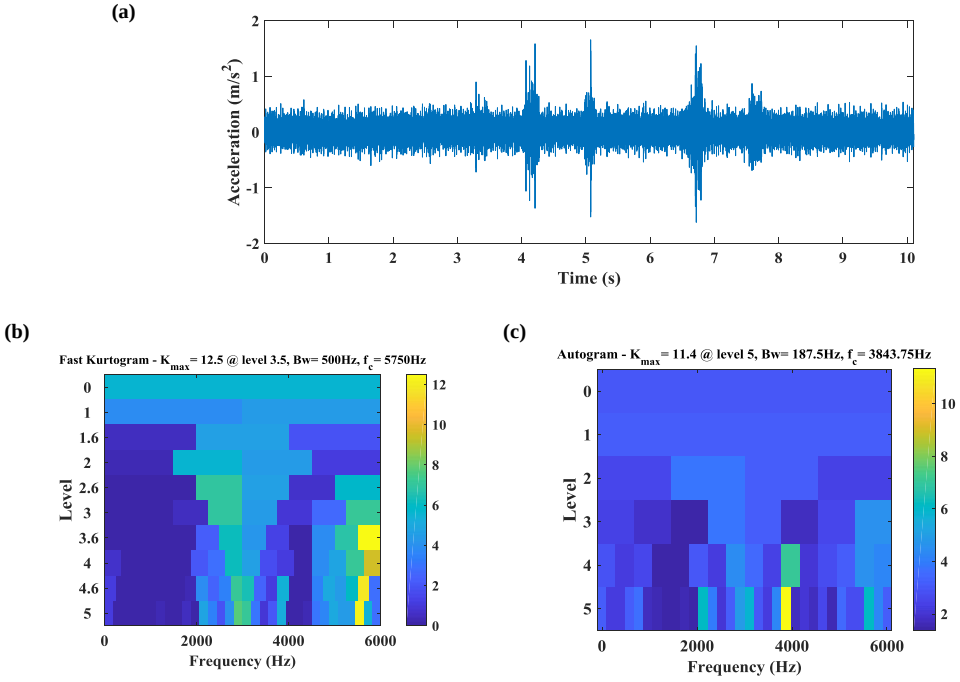

Figure 11: Case 4 (a) Time domain signal: 222 DE (b) FK (c) Autogram

vibration signal with many large impulses. In regard to the ball defect, the fault diagnosis is categorized as P1, Y2 and Y2 by methods 1, 2 and 3 but neither method 2 nor method 3 are able to diagnose the unintentional inner race fault [17].

The FK is plotted in Figure 11b and the node with center frequency 5750 Hz and bandwidth 500 Hz has the highest kurtosis. The SES of the signal related to this node is depicted in Figure 12a where smeared peaks can be detected at the first and second harmonics of the BSF, particularly at 2×BSF. High slippage is usually responsible for smeared components in frequency domain but sources of these smeared peaks are ascribed to random and impulsive amplitude modulation of the signal rather than ball slip [17].

The Autogram is plotted in Figure 11c, where an optimal band with center frequency 3843.75 Hz and bandwidth 187.5 Hz is found. In Figure 12b, the

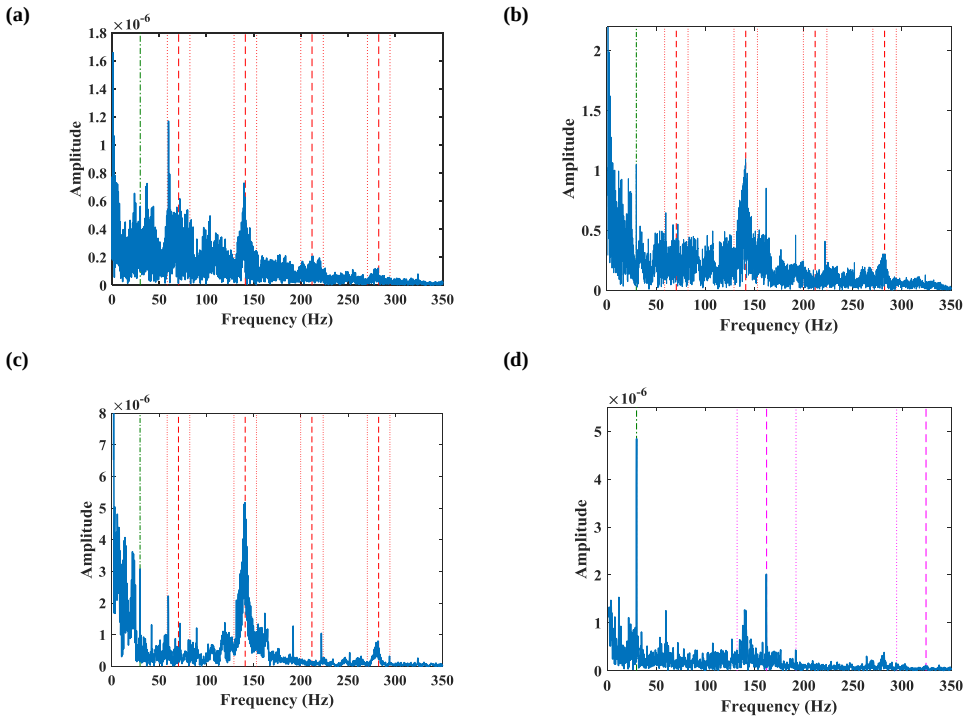

Figure 12: Case 4, SES of the node selected by (a) FK (b) Autogram (c) upper part spectrum (d) lower part spectrum

SES of the filtered signal provides ball fault signatures for bearing fault diagnosis. It shows smeared components that correspond to BSF and its harmonics, specifically even harmonics (2×BSF and 4×BSF). Being BSF the frequency at which the fault is engaged with the same race (outer or inner race), the even harmonics of BSF are often dominant in the envelope spectrum.

The large impulses in time are attributed to the ball defect and FK enhances only the largest pulses related to BSF, which causes high kurtosis. Conversely, Autogram tries to find the frequency band which contains a series of impulses and, as a result, detects a different frequency band, with more dominant even harmonics of the BSF. An interesting observation is that a discrete frequency, the BPFI at 162.2 Hz, is present in Figure 12b but does not exist in Figure 12a. This indicates better performance of Autogram over FK in presence of multiple defects, especially when they are not subjected to the same random and impulsive amplitude modulation.

The SESs after applying upper and lower threshold are depicted in Figure 12c and d respectively and the level of noise is reduced drastically. Moreover, as the Upper Autogram mainly contains the highly amplitude modulated impulses associated to the ball defect, the BSF is more evident in (Figure 12c) than in (Figure 12b). On the contrary, since the lower part is less affected by modulation, in Figure 12d the dominancy of the BSF is decreased and now the shaft frequency and BPFI have the largest amplitude in the spectrum. In Figure 12d the pink dashed line cursor is tuned at the first two harmonics of the expected BPFI, and dotted lines demonstrate the first order shaft modulation sidebands around the fault frequency and its harmonics.

4.5. Case 5: several non-periodic impulses

In this case, record 291 FE which has a defect on its rolling element is examined. The vibration signal of the record is plotted in Figure 13a, showing high level of non-stationarity due to several large but non-periodic impulses. This data is part of the fan end measurements. The important characteristic of this data sets is that the bearing characteristic frequencies are almost integer

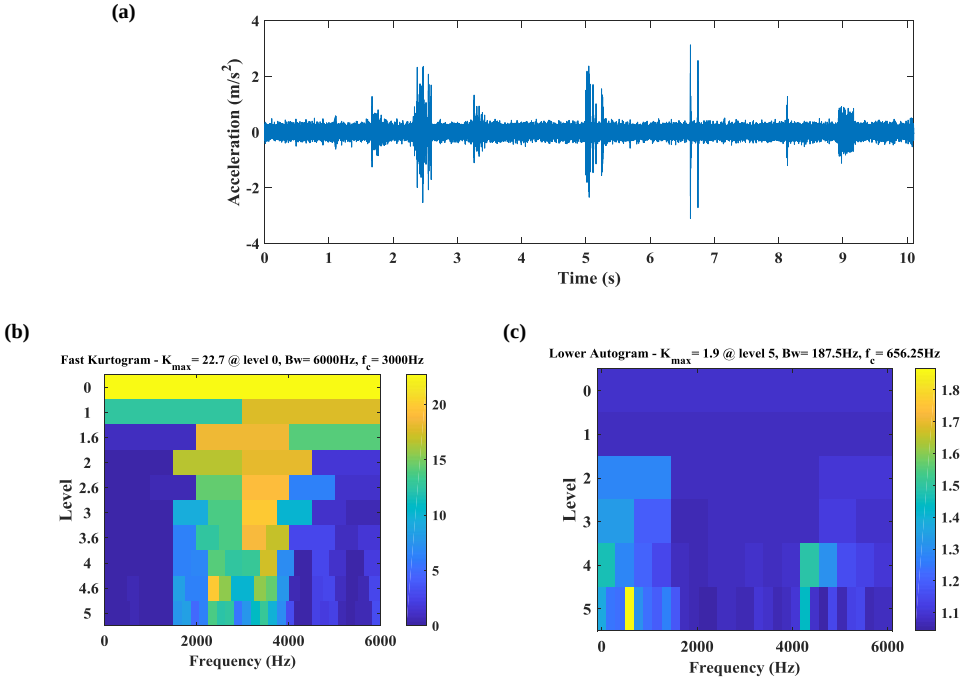

Figure 13: Case 5 (a) Time domain signal: 291 FE (b) FK (c) Autogram

multiples of the shaft frequency which, in many cases, makes it difficult to differentiate between the defect and shaft frequencies. The bearing fault is diagnoseable (P2) by methods 1 and 3 of the benchmark study but method 2 does not determine the defect (N1) [17].

Figure 13b illustrates that the FK selects the frequency band with center frequency 3000 Hz and bandwidth 6000 Hz which is associated with the raw signal. The SES of the raw signal is shown in Figure 14a. Though it is categorized as P2 there is no evidence of the BSF, possibly because in this paper, contrary to [17], no DRS filtering has been applied.

Very large impulses in the signal cause high kurtosis and the effectivity of the FK and to some extent Autogram is reduced. This drawback can nonetheless be overcome by taking advantage of Lower Autogram.

The Lower Autogram is shown in Figure 13c and the maximum value is

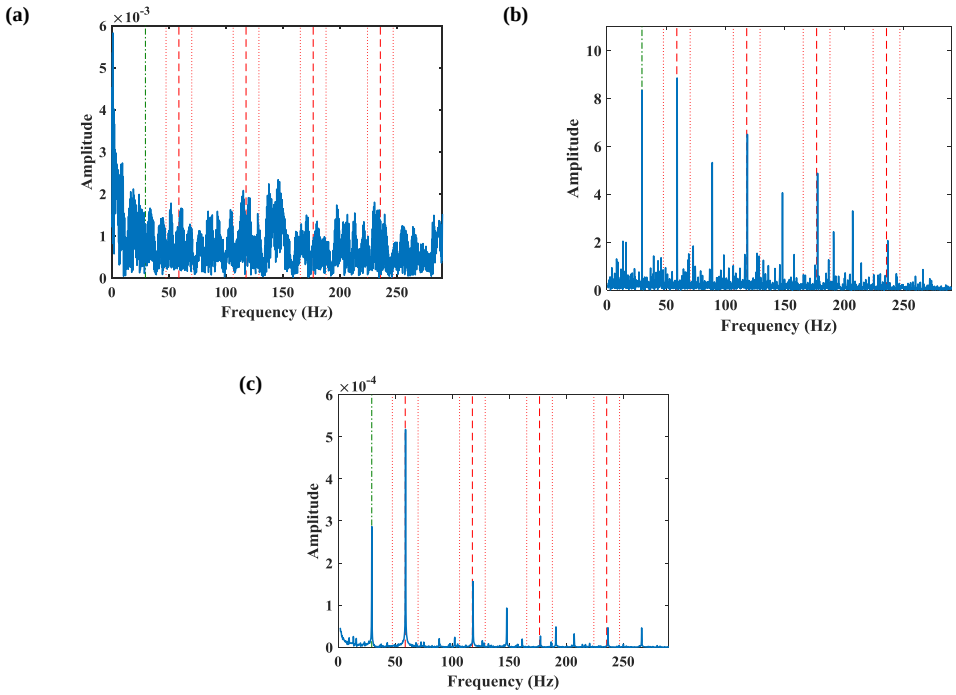

Figure 14: Case 5, SES of the node selected by (a) FK (b) Autogram (c) lower part spectrum

assigned to the node (5, 4). In the present case, the defect frequency is almost twice the shaft frequency. Figure 14b displays the envelope spectrum of the filtered signal which clearly provides the ball fault signature, as the BSF has larger amplitude than the shaft frequency. As it was discussed, the effect of modulation could be alleviated by computing the spectrum of the lower part of the envelope's AC. The outcome is depicted in Figure 14c in which the defect frequency is even more relevant than in Figure 14b. Also, the third harmonic of the shaft frequency is completely canceled which means it was a sideband of the defect frequency rather than a harmonic of the shaft frequency.

By using Lower Autogram other data sets such as 291 BA, 283 DE, 284 DE/FE, 285 DE can be diagnosed similarly.

4.6. Case 6: low signal to noise ratio

As the last case of the first part, record 204 FE with a defect on its outer race is studied and Figure 15a presents the vibration data in the time domain. The SNR is very low and the data set is not diagnosable with any of the applied methods of the benchmark study (N1/N1/N2) [17].

The FK is depicted in Figure 15b, and Figure 15d shows the SES for the indicated optimal band, which does not contain useful diagnostic information.

As the SNR is very low, the Upper Autogram is preferable (see subsubsection 4.1.3) and it is plotted in Figure 15c. The SES of the node selected by Upper Autogram reveals fault related frequencies (ballpass frequency outer race - BPFO) in Figure 15e, although they are not the most dominant components.

Records 199 DE, 203 DE and 204 DE are other cases which can be diagnosed successfully.

This part demonstrates that the proposed method not only can provide comprehensive information regarding probably diagnosable (P) data sets but also successfully diagnoses data sets in the N categories.

4.7. Computational Time

Before ending the first part, it is worth mentioning the computational time of the Autogram for these 6 cases. Cases 1, 3 and 6 consist of roughly 480000

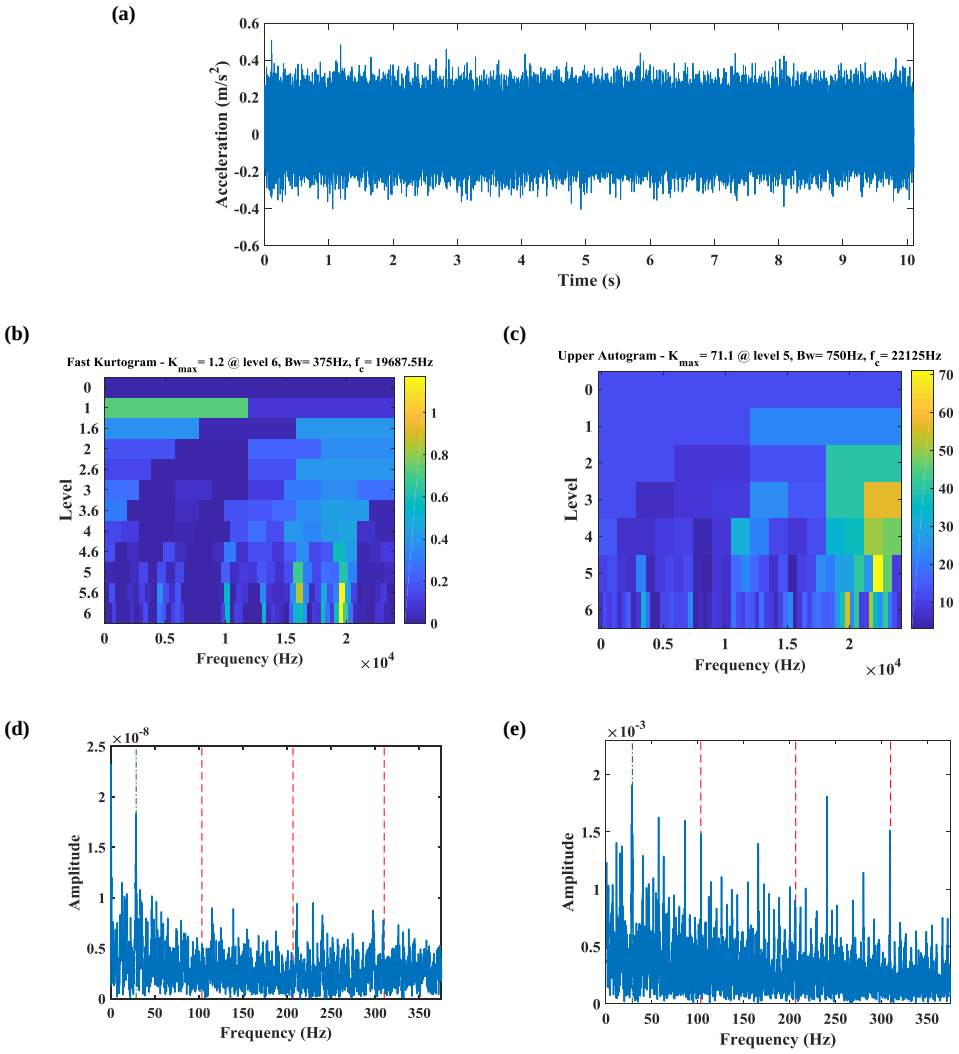

Figure 15: Case 6 (a) Time domain signal: 204 FE (b) FK (c) Autogram, SES of the node selected by (d) FK (e) Autogram

Table 2: CPU time in seconds required to compute the Autogram for 6 cases studied in this section

		Case 1	Case 2	Case 3	Case 4	Case 5	Case 6
Length of Signal (sample)		481e3	121e3	482e3	121e3	121e3	480e3
Level of Decomposition		6	5	6	5	5	6
CPU Time (sec)	Autogram	32	4.5	39	4.5	4.3	32

samples and a 6 levels decomposition is performed. Cases 2, 4 and 5 include roughly 120000 samples and the signals are decomposed in 5 levels. Table 2 reports the required CPU time on a laptop computer (Intel Core i7-4810HQ Processor 2.50 GHz). The computational time mainly consists the time for filtering the signal, calculation the ACs and their envelopes. Since all the filtered signals share the same length of the original data, each level of decomposition requires twice the computational time of the preceding one.

On the other hand, less than 1 second is needed to compute the FK as it exploits multirate filtering process and only kurtosis values are calculated for each level and frequency band (but not the corresponding time series).

Although Autogram is not as fast as FK, based on the reported computational time in Table 2, it could still be considered for online condition monitoring of industrial systems.

In the following two cases, the performance of Autogram and CSES will be shown and for the sake of comparison, the Fast-SC and the Enhanced Envelope Spectrum (EES) (see Ref. [15]) are also presented.

It should be noted that the EES and CSES produce different quantities and cannot then be compared directly. The aim of these last examples is to show the capability of the proposed method in providing diagnosis information similar to Fast-SC and EES.

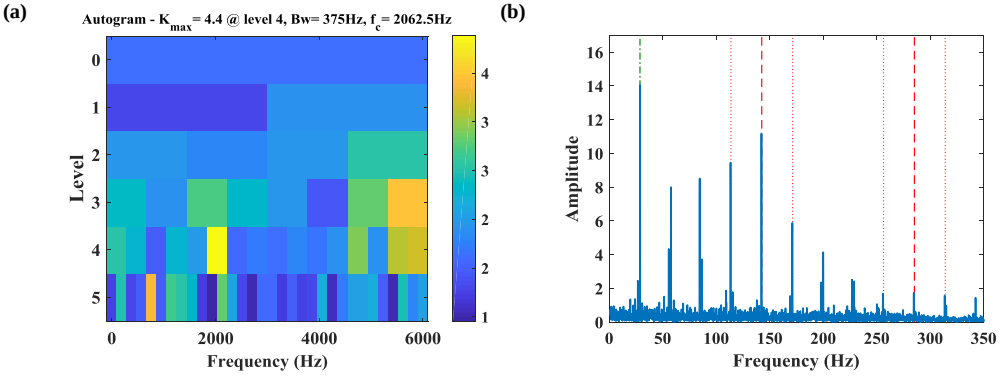

Figure 16: Case 7 (a) Autogram (d) SES of the node selected by Autogram

4.8. Case 7: two faulty bearings

In this case, record 277 DE is examined. The bearing located on the fan end is reported to be defective on inner race but the signal is collected on the drive end bearing. As it is illustrated in Figure 16a, the Autogram selects the frequency band with center frequency 2062 Hz and bandwidth 375 Hz. The SES of the signal related to this node is shown in Figure 16b where the modulation or cyclic frequency associated with the BPFI of the fan end bearing can be clearly detected. Other nodes have high kurtosis in the Autogram so that, to investigate their frequency components, CSES (subsection 3.3) is calculated. The CSES result is depicted in Figure 17a and Figure 17b represents the average of the CSES for these six levels. The BPFI of the drive end bearing (156 Hz), that was not highlighted in the SES of the selected node by Autogram, can now be detected (pink dashed line). This example shows the improved performance of the Autogram by utilizing the CSES instead of the SES.

For the sake of comparison, the Fast-SC and EES in full band ([15]) are shown in Figure 18a and Figure 18b respectively. In this case, the EES and CSES generate almost comparable results and both defects frequencies can be detected in their spectra.

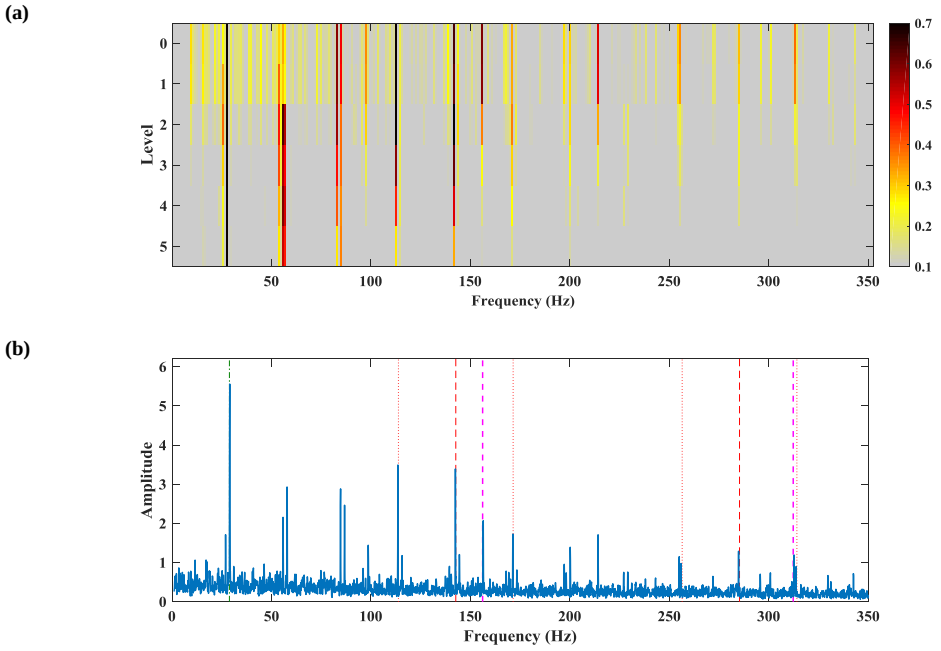

Figure 17: Case 7 (a) Combined Squared Envelope Spectrum (CSES) (b) average Combined Squared Envelope Spectrum (CSES) for all levels

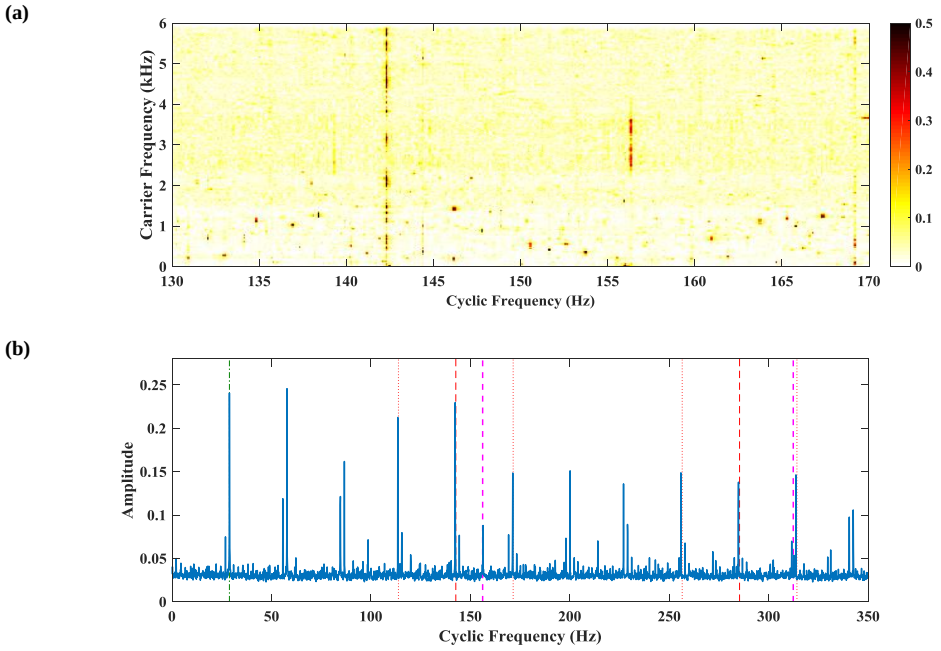

Figure 18: Case 7 (a) Fast Spectral Correlation (b) full band Enhanced Envelope Spectrum (EES)

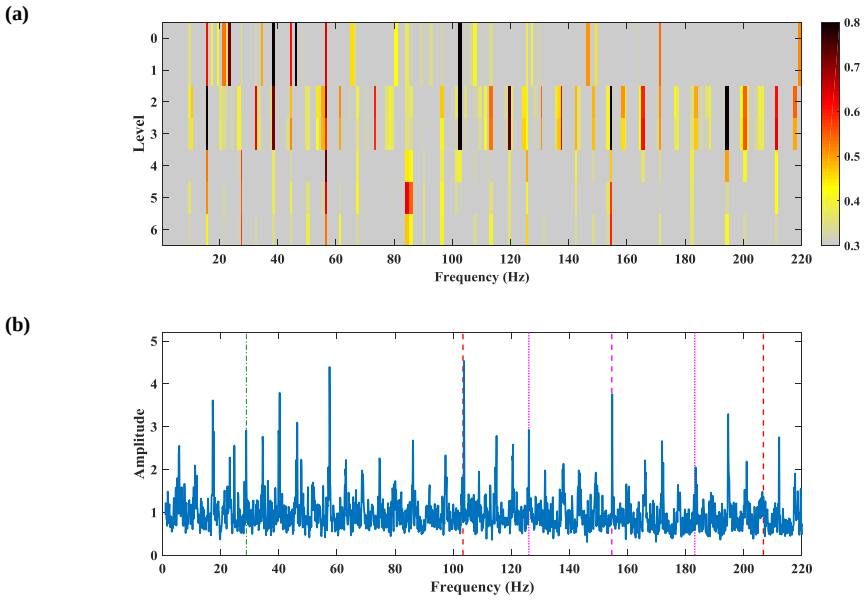

Figure 19: Case 8 (a) Combined Squared Envelope Spectrum (CSES) (b) average Combined Squared Envelope Spectrum (CSES) for all levels

4.9. Case 8: bearing with two defected races

Record 204 DE, with an outer race fault on the drive end bearing, is studied as the last case. The CSES is depicted in Figure 19a for the Upper Autogram and the BPFO of the drive end bearing (103.4 Hz) is the dominant frequency through levels 0 to 3. The average CSES for these 7 levels of decomposition is shown in Figure 19b. Figure 20a illustrates the Fast-SC and its zoomed portion around the BPFO, and the EES in full band is shown in Figure 20b. Although no defect on the inner race is reported, the BPFI (156.1 Hz) can also be observed in the CSES (Figure 19) and EES (Figure 20b). The EES computed in the frequency band [2000 4000] Hz, in which the BPFO of the drive end bearing shows high values, is shown in Figure 21a and comparing to full band EES Figure 20b, the defect frequency is more dominant. Finally, Figure 21b displays the average CSES for levels 0 to 3 where the BPFO has high values.

The SC has the benefit of automatically showing the carrier frequency re-

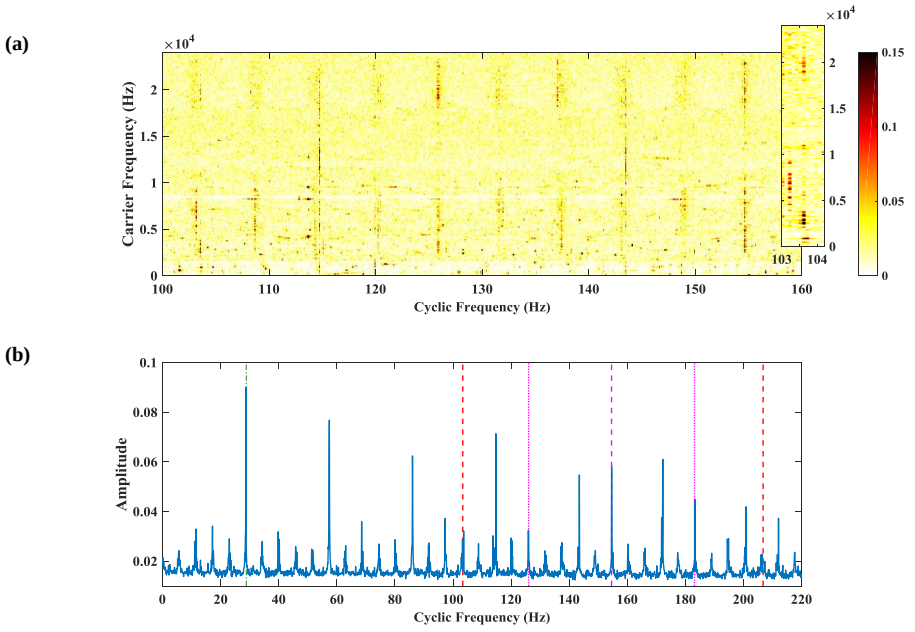

Figure 20: Case 8 (a) Fast Spectral Correlation (b) full band Enhanced Envelope Spectrum (EES)

lated to each modulation frequency. Moreover, computing EES for selected bands can enhance its diagnosis capabilities, because the spectrum can reveal many harmonics of the defect frequencies as well as their sidebands. Also the proposed method (Autogram + CSES) is able to automatically select the carrier frequencies containing useful information, even if the signal processing approach is different. According to the so far presented examples, Autogram can indeed achieve valuable results, at least comparable to other state-of-the art procedures.

5. Conclusion

This paper proposes a new method to find the proper frequency band of demodulation for bearing faults diagnosis. At first, undecimated wavelet packet transform (MODWPT) is adopted to split a signal in different frequency bands and central frequencies. Second, unbiased autocorrelation of the squared envelope of theses signals is calculated to take advantage of the 2nd order cyclo-

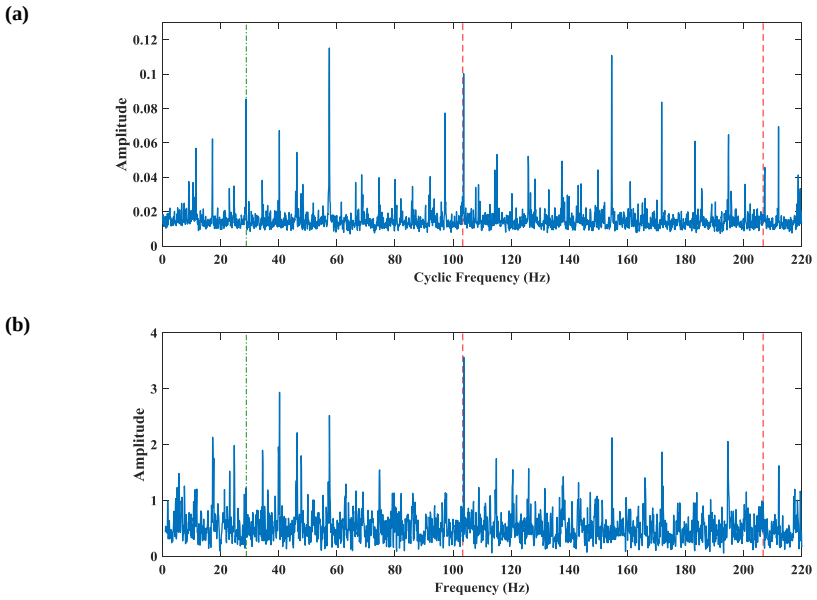

Figure 21: Case 8 (a) Enhanced Envelope Spectrum (EES) in selected frequency band [2 4] kHz (b) average Combined Squared Envelope Spectrum (CSES) for levels 0 to 3

stationarity of bearing faults signals, to reduce the level of uncorrelated random noise and enhance the fault related peaks. Third, kurtosis of the ACs is computed and a two dimensional colormap, named Autogram, is presented in order to locate the optimal frequency band for demodulation. Additionally, two modified versions of kurtosis equation are introduced which lead to two special cases of Autogram, namely Lower and Upper Autogram. The advantages and characteristics of each proposed Autogram and the conditions in which using Upper/Lower Autogram is beneficial are also discussed. Finally, the node with the highest kurtosis is chosen and Fourier transform is used to obtain a frequency domain representation of the envelope signal so to identify the defect frequencies of bearings.

The main advantage of the algorithm is its ability in limiting the influence of non-periodic impulses and noise from the raw time data, which are not related to any actual defects of bearings, thus enhancing the selection of correct frequency band for the ensuing spectrum calculation. This valuable achievement

is effectively reached by computing the autocorrelation of the envelope signal. Recognition of faults is based on defect frequencies identification so that it is not necessary to rely on undamaged conditions to diagnose the bearing status. On the other hand, Autogram is specifically designed to improve the detection of periodic impulses and therefore it is not suited to discover other sort of damages like pitting or corrosion.

The procedure can be fully automated to exploit the combined capabilities of Autogram and CSES, in order to generate a single spectrum where defect frequencies are clearly indicated. A major advantage of FK over Autogram is related to the computational time since Autogram, in contrast to FK, does not take advantage of multirate filtering.

The method has been tested on experimental data provided by the Case Western Reserve University bearing data center, and compared with both literature results and other effective procedures so to assess its performances in REBs diagnosis. The results are indeed positive and Autogram allows to identify damaged bearings also in case of non-Gaussian noisy data.

Acknowledgement

The first author would like to thank professor Jérôme Antoni for sharing the Fast Kurtogram and Fast-SC codes publicly.

External resources: The algorithm for computing the Autogram can be obtained from the authors, or downloaded at https://it.mathworks.com/matlabcentral/fileexchange/65413-autogram.

References

[1] L. Renforth, P. S. Hamer, D. Clark, S. Goodfellow, R. Tower, Continuous, remote on-line partial discharge (olpd) monitoring of hv ex/atex motors in the oil and gas industry, in: Petroleum and Chemical Industry Technical Conference (PCIC), 2013 Record of Conference Papers Industry Applications Society 60th Annual IEEE, IEEE, 2013, pp. 1–8.

[2] M. S. Darlow, R. H. Badgley, G. Hogg, Application of high-frequency resonance techniques for bearing diagnostics in helicopter gearboxes, Tech. rep., Mechanical Technology Inc Latham NY (1974).

[3] R. B. Randall, J. Antoni, Rolling element bearing diagnostics–a tutorial, Mechanical systems and signal processing 25 (2) (2011) 485–520.

[4] J. Antoni, The spectral kurtosis: a useful tool for characterising non-stationary signals, Mechanical Systems and Signal Processing 20 (2) (2006) 282–307.

[5] J. Antoni, Fast computation of the kurtogram for the detection of transient faults, Mechanical Systems and Signal Processing 21 (1) (2007) 108–124.

[6] Y. Lei, J. Lin, Z. He, Y. Zi, Application of an improved kurtogram method for fault diagnosis of rolling element bearings, Mechanical Systems and Signal Processing 25 (5) (2011) 1738–1749.

[7] T. Barszcz, A. JabŁoński, A novel method for the optimal band selection for vibration signal demodulation and comparison with the kurtogram, Mechanical Systems and Signal Processing 25 (1) (2011) 431–451.

[8] D. Wang, W. T. Peter, K. L. Tsui, An enhanced kurtogram method for fault diagnosis of rolling element bearings, Mechanical Systems and Signal Processing 35 (1) (2013) 176–199.

[9] J. Antoni, The infogram: Entropic evidence of the signature of repetitive transients, Mechanical Systems and Signal Processing 74 (2016) 73–94.

[10] Y. Wang, J. Xiang, R. Markert, M. Liang, Spectral kurtosis for fault detection, diagnosis and prognostics of rotating machines: A review with applications, Mechanical Systems and Signal Processing 66 (2016) 679–698.

[11] J. Antoni, Cyclic spectral analysis in practice, Mechanical Systems and Signal Processing 21 (2) (2007) 597–630.

[12] J. Antoni, Cyclic spectral analysis of rolling-element bearing signals: facts and fictions, Journal of Sound and vibration 304 (3) (2007) 497–529.

[13] J. Antoni, D. Hanson, Detection of surface ships from interception of cyclostationary signature with the cyclic modulation coherence, IEEE Journal of Oceanic Engineering 37 (3) (2012) 478–493.

[14] P. Borghesani, The envelope-based cyclic periodogram, Mechanical Systems and Signal Processing 58 (2015) 245–270.

[15] J. Antoni, G. Xin, N. Hamzaoui, Fast computation of the spectral correlation, Mechanical Systems and Signal Processing 92 (2017) 248–277.

[16] J. Antoni, Fast kurtogram.
URL https://mathworks.com/matlabcentral/fileexchange/48912-fast-kurtogram

[17] W. A. Smith, R. B. Randall, Rolling element bearing diagnostics using the case western reserve university data: A benchmark study, Mechanical Systems and Signal Processing 64 (2015) 100–131.

[18] J. Antoni, Cyclic spectral analysis in practice, Mechanical Systems and Signal Processing 21 (2) (2007) 597–630.

[19] R. B. Randall, J. Antoni, S. Chobsaard, The relationship between spectral correlation and envelope analysis in the diagnostics of bearing faults and other cyclostationary machine signals, Mechanical systems and signal processing 15 (5) (2001) 945–962.

[20] J. Antoni, R. Randall, Differential diagnosis of gear and bearing faults, Transactions-American Society of Mechanical Engineers Journal of Vibration and Acoustics 124 (2) (2002) 165–171.

[21] A. T. Walden, Wavelet analysis of discrete time series, in: European Congress of Mathematics, Springer, 2001, pp. 627–641.

[22] Case western reserve university bearing data center website.
URL http://csegroups.case.edu/bearingdatacenter/home

[23] R. Randall, N. Sawalhi, M. Coats, A comparison of methods for separation of deterministic and random signals, International Journal of Condition Monitoring 1 (1) (2011) 11–19.

[24] A. Moshrefzadeh, A. Fasana, L. Garibaldi, Using unbiased autocorrelation to enhance kurtogram and envelope analysis results for rolling element bearing diagnostics, in: International Conference Surveillance 9, Morocco, 2017.