Phase editing as a signal pre-processing step for automated bearing fault detection

L. Barbini[a,b,*], A. P. Ompusunggu[b], A. J. Hillis[a], J. L. du Bois[a], A. Bartic[b]

[a] *The University of Bath Department of Mechanical Engineering, Claverton Down, Bath, BA2 7AY, UK*
[b] *Flanders Make VZW, Celestijnenlaan 300, B-3001 Heverlee (Leuven), Belgium*

Abstract

Scheduled maintenance and inspection of bearing elements in industrial machinery contributes significantly to the operating costs. Savings can be made through automatic vibration-based damage detection and prognostics, to permit condition-based maintenance. However automation of the detection process is difficult due to the complexity of vibration signals in realistic operating environments. The sensitivity of existing methods to the choice of parameters imposes a requirement for oversight from a skilled operator. This paper presents a novel approach to the removal of unwanted vibrational components from the signal: phase editing. The approach uses a computationally-efficient full-band demodulation and requires very little oversight. Its effectiveness is tested on experimental data sets from three different test-rigs, and comparisons are made with two state-of-the-art processing techniques: spectral kurtosis and cepstral pre- whitening. The results from the phase editing technique show a 10% improvement in damage detection rates compared to the state-of-the-art while simultaneously improving on the degree of automation. This outcome represents a significant contribution in the pursuit of fully automatic fault detection.

Keywords: Phase spectrum, automated diagnostics of defective bearings, comparison with cepstrum pre-withening, comparison with spectral kurtosis

1. Introduction

Research into fully automated algorithms for the application of condition based maintenance on an industrial scale has gained much attention in recent years. There is a practical need for reliable diagnostic methods to help avoid human errors and to allow affordable implementations of multi-sensor architectures without employing expert users. In this field, vibration based condition monitoring is a well established approach, present in the literature since the 1980s, and a variety of vibration transducers are currently used: for example, accelerometers, acoustic emission and ultrasound [1, 2]. It is well known that bearings play a vital role in the health of many industrial machines and that they are particularly prone to failure [3]. This paper is concerned with automated algorithms for the assessment of bearing health using vibrational data.
The mathematical formulation of the second order cyclostationary nature of defective bearings signal and the clarification of its relationship with the high-frequency resonance technique [4] has

*Corresponding author
Email address: leo.barbini@bath.ac.uk (L. Barbini)

allowed the introduction of a widely accepted model of bearing faults [5, 6]. According to this model, features from defective bearings can be extracted by the squared envelope spectrum (SES) of the vibrational signal, as an estimation of the integrated cyclic coherence. The main difficulty in such feature estimation is that studying the SES of the full band raw signal usually does not provide a reliable assessment of the bearing condition and can result in defective bearings being classified as healthy. In complex machines, such as those used in industrial applications, this is due to the presence of vibration signals from other sources which are superimposed on those from the bearing itself. For instance, gears can contribute non-trivial vibration signals masking the presence of bearing defects. In such cases the SES of the raw vibrational signal shows various peaks, and the results are difficult to classify without the analysis of an expert user. From the point of view of an automated feature estimation there is then the need to find reliable methods to process the raw vibration signals, before the computation of the features from the SES, in order to simplify the diagnostic of a defective bearing.

Attempts have been made to automate common methods for the enhancement of bearing signals, however the results are not sufficiently reliable for industrial applications. For example the well established spectral kurtosis (SK) [7] is often used to select the frequency band containing bearing impact signals, but when automated there is a high risk of selecting the wrong demodulation band [8]. One solution to this problem is to edit the vibration signal such that the full frequency band is considered. This approach can be found in a method introduced by Borghesani [8], known as the zero cepstrum (ZC), where all the cepstral coefficients are set to zero and the SES is calculated for the full band edited signal. However this methods results highly sensitive to noise levels, as will be investigated in the following section. In the paper the focus is on an automated algorithms and cepstral approaches were preferred to other methods for removal of gear signals, as time synchronous averaging or discrete random separation, because somewhere else [9, 10] they performed comparably well but require a faster implementation and result easier to automate.

This paper presents a novel automated signal processing method for feature detection in bearing vibration signals, based on phase editing (PE). As with ZC, the PE method has the advantage of using the entire band for demodulation, so avoids the risk of selecting the wrong frequency band. The novelty lies in the procedure for the removal of vibrational signals not originating from bearings. In many enhancement techniques only the amplitude spectrum of a noisy signal is modified while the phase spectrum is unchanged. The PE approach on the other hand keeps the amplitude unchanged and recombines the signal with an edited phase spectrum. The result is a modified spectrum where small amplitude components are attenuated more than large amplitude components. It is well established that in a vibrational signal the energy from bearings is low compared to that from the other components [11], hence in order to discard the misleading information and enhance the weak signal from the bearings, the final step is to calculate the residual between the original signal and the PE signal. To the best of authors' knowledge the present paper is the first to propose the use of phase editing in the field of condition monitoring. The effectiveness of this novel technique has been tested by comparing its performance on experimental data sets with that of two state-of-the-art methods. The first consists of two steps: removal of discrete components, implemented by means of an automated cepstral editing procedure (ACEP) [12] and SK for selection of the band for demodulation; the second method tested in the comparison is the ZC.

The paper is structured as follows. Section 2 presents details and discussion of the three methods. Section 3 describes the experimental rigs. Results are presented in Section 4 and conclusions are drawn in Section 5.

2. Methods

In the signal acquired by an accelerometer positioned on a gearbox casing at least three components superimpose: vibration induced by gears, vibration from bearings and noise. Usually a suggested operation is that of order tracking for removal of speed fluctuations [11], however in this paper machines are operating at a constant speed and this step will be omitted. All the three methods described below are be intended as enhancement procedures implemented before bearings fault feature estimation by means of the SES. They are automated such that the user needs to provide only the bearing defect frequencies and the rotation speed of the shaft where the bearing is mounted. Diagrams of the three methods ACEP+SK, ZC, PE are shown respectively in Fig. 1-2-4.

2.1. Existing Method (I): ACEP+SK

Method I consists of two blocks, ACEP and SK. The first one minimises the components from the gears, which are likely to mask the weak signal from bearings. This is achieved using an edited version of the cepstrum. Advantages on editing the cepstrum are that it can be automated [12], without the need of adjustment from the user, and that it has performed well when compared with other methods for gears signal removal as time synchronous averaging, auto regressive modelling, self-adaptive noise cancellation, discrete/random separation [9, 10]. In more details three steps are involved in the ACEP block: spectral subtraction is applied to the high-pass liftered real cepstrum of the raw signal and afterwards a comb lifter is automatically generated for removal of impulses in

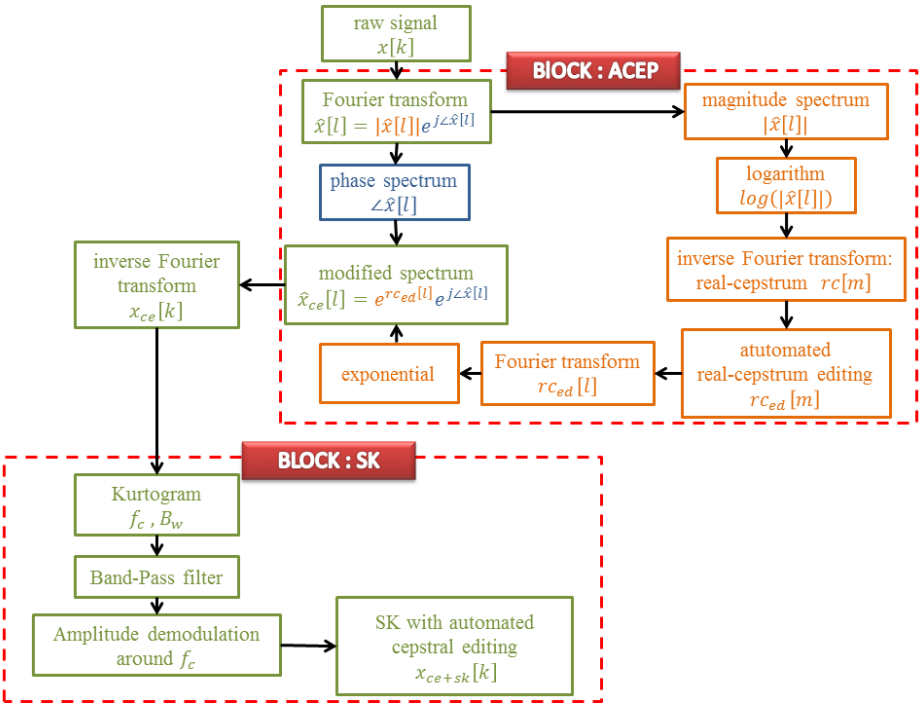

Figure 1: Scheme of method I: automated cepstral editing procedure (ACEP)+ spectral kurtosis (SK). Indexes k, l, m are for the time, frequency and quefrency domains. The green colour is for a signal where both magnitude and phase information are present while orange and blue are magnitude only and phase only.

the quefrency domain related to gear vibrations. This procedure provides an edited real cepstrum which is recombined with the original phase spectrum of the raw signal to obtain an edited vibration signal with enhanced bearing components x_{ce}.
The second step of Method I is the selection of the band of the spectrum where are present mainly vibrations from bearings, this is achieved using the SK. Its fast implementation was introduced by Antoni in [13], the algorithm gives as results: the bandwidth B_w and central frequency f_c of the demodulation band where the signal has the highest kurtosis and also the demodulated signal $x_{ce+sk}[k]$. From the perspective of automation two problems can be identified. Firstly the possible presence of vibration sources characterised by a kurtosis level higher then that of the defective bearings; secondly, the presence of peaks from gear harmonics in the selected narrow band for the demodulation. It should be noticed that the pre-processing with the ACEP algorithm increases the efficacy of the SK removing harmonics from gear components so to address the second of this problems. However the first problem is left unsolved and usually only human analysis of the kurtogram allows the right selection of demodulation band.

2.2. Existing Method (II): ZC

This method was described and exhaustively compared with other techniques by Borghesani et al. in [8]. ZC enhances the raw vibrational signal by means of editing the real cepstrum, however two differences with ACEP+SK can be found: firstly not only some impulses but the whole real cepstrum is set to zero; secondly no band pass filtering is implemented enabling full band demodulation. As remarked in [8] the computation of the $x_{zc}[k]$ can be accomplished without the calculation of the real cepstrum, as shown in Fig. 2, but simply modifying the Fourier transform

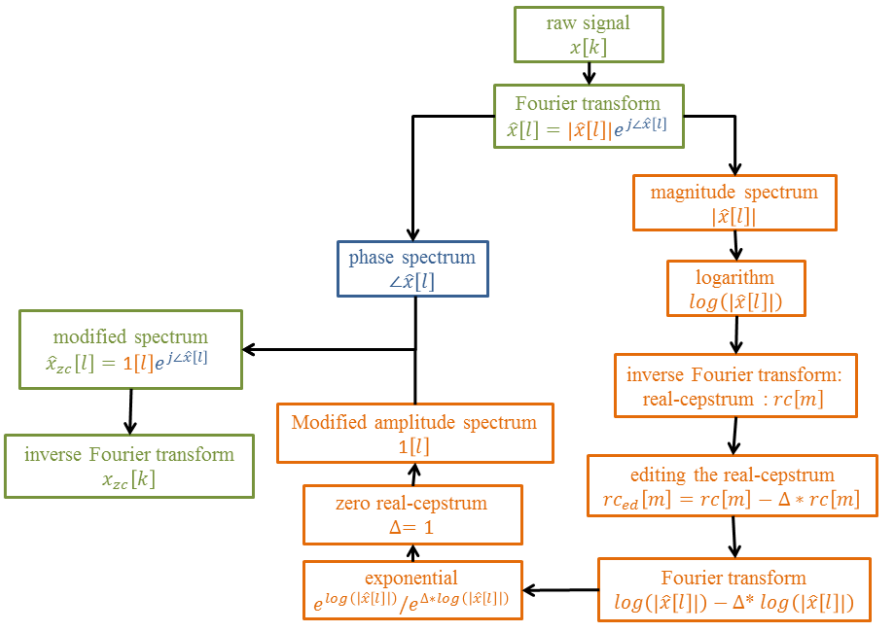

Figure 2: Scheme of method II: zero cepstrum (ZC). Indexes k, l, m are for the time, frequency and quefrency domains. The green colour is for a signal where both magnitude and phase information are present while orange and blue are magnitude only and phase only.

of the raw signal. The result of this procedure gives the so called 'phase only signal' [14].
In this framework the ZC edited signal is a filtered version of the raw signal, specifically the filter $g[k]$ is a function of the input and gives as output a signal with flat amplitude spectrum [15] and unchanged phase:

$$g[k] = IFT\left\{\frac{1}{|\hat{x}[l]|}\right\} \tag{1}$$

This procedure will perform equally well for a signal with no impulses in the real cepstrum. It can be interpreted as an amplitude distortion, implemented by means of the zero-phase filter $g[k]$, yielding to the enhancement of narrow time events. A numerical simulation is presented below in order to demonstrate the effect of this operation.

The simulated signal is the superposition of three unevenly spaced sinusoids at $11, 23$ and 31 $[Hz]$ with unity amplitudes, an outer race defective bearing with impact frequency $f_{od} = 80$ $[Hz]$ and a weak white noise of variance $\sigma = 0.05$. In the signal from the bearings a random jitter is considered for both the occurrence times and amplitudes of impacts, according to [5], and the excited resonance frequencies are in the band $15 - 17[kHz]$. The signal has a length of $7.5[s]$ and is sampled at $51kHz$. Figures 3(a) and (b) show respectively: a segment of the raw signal with seven recognisable impacts, the amplitude spectrum of the whole raw signal. The real cepstrum is shown in Fig.3(d) where the range of quefrencies is selected as 1 / $2f_{od}$ and 3 / f_{shaft} with the shaft frequency taken as the lowest of the sinusoids $f_{shaft} = 11$ $[Hz]$. As expected no clear peak is present in the real cepstrum at 0.1 $[1/Hz]$. Figure 3(c) shows the envelope of the zero

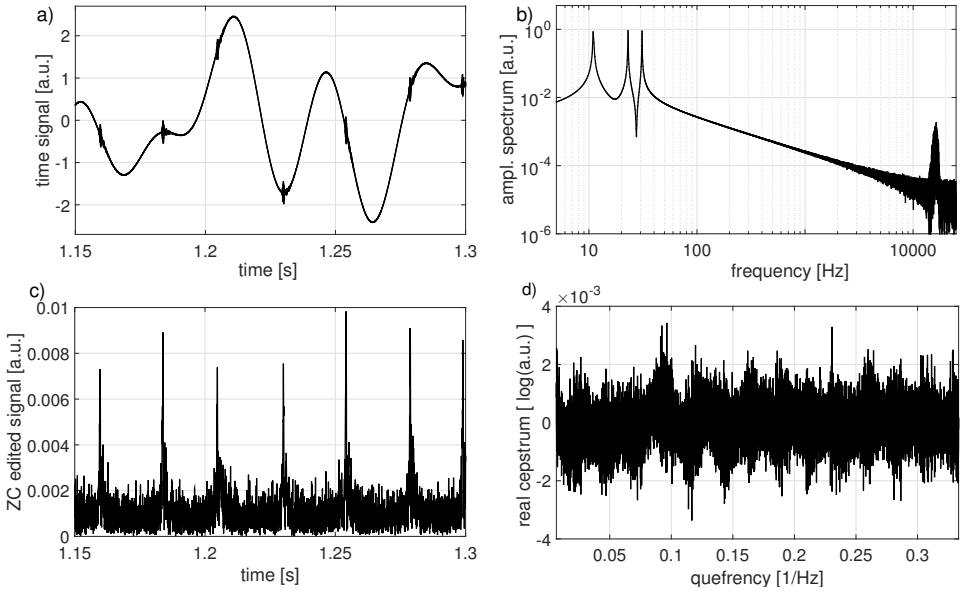

Figure 3: Application of the ZC method to a numerical simulation. The tested signal is realised as the superposition of three sinusoids at $11, 23$ and 31 $[Hz]$, a defective bearing signal and white noise. (a) Time domain: 7 impacts can be identified. (b) Amplitude spectrum of the raw signal: the tonal components have higher power than the defective bearing signal, the structural resonance excited by the impacts is in the band $15 - 17[kHz]$. (c) Absolute value of the analytic representation of the zero cepstrum edited signal $x_{zc}[k]$: sinusoidal components are removed while the impacts are enhanced. (d) Real cepstrum of the input signal: no clear peak can be identified in the quefrency range containing $1/10$ $[1/Hz]$.

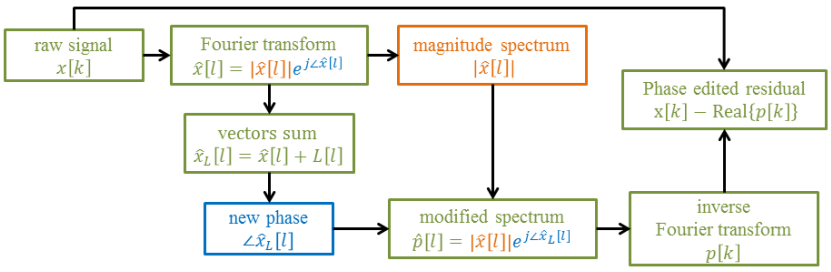

Figure 4: Scheme of method III: phase editing (PE). Indexes k, l are for the time and frequency domains. The green colour is for a signal where both magnitude and phase information are present while orange and blue are magnitude only and phase only.

cepstrum edited signal, calculated as the absolute value of the analytic representation of $x_{zc}[k]$. It is possible to identify the seven impacts and no presence of the sinusoids, confirming as anticipated the usefulness of method II for a signal with real cepstrum without impulses.

2.3. New Method (III): PE

The use of an edited phase spectrum, combined with the original magnitude spectrum, was first introduced by Wojcicki *et al.* [16] as a novel approach to speech enhancement. The method suppresses low-energy signal components through manipulation of the phase of the conjugate signal components in the frequency domain. In the original application it was used to clarify the high-energy spectral components and reduce noise. The methods are adapted here to use the residual from the original signal, thus rejecting the high-energy components and retaining the low-energy bearing signals of interest.

The schematic of Method III is shown in Fig.(4). The raw signal is first subject to a discrete Fourier transform. A real value $L[l]$ is then added to each complex component in the frequency domain, and normalised to the original component amplitude. The value $L[l]$ is of constant amplitude λ, and is positive for frequency components in the upper half of the imaginary plane and negative for those in the lower half-plane:

$$L[l] = \begin{cases} +\lambda & 0 \leq l < N/2 \\ -\lambda & N/2 \leq l \leq N-1 \end{cases} \tag{2}$$

with λ a positive real constant. The only parameter to tune in this procedure is λ, and this determines the amplitude threshold of the filter. The effect is shown in Figure 5(a). When the two complimentary signal components are recombined it serves to reduce the real part of the signal for small amplitude signal components. Upon transformation back to the time domain, the imaginary portion of the signal is discarded and the low-energy components are suppressed accordingly. The following analysis demonstrates the principle of operation.

The contribution of a frequency component $l = n \leq \frac{N}{2}$ to the real signal $x[k]$ having DFT $\hat{x}[l]$ with $k, l = 0, \ldots, N-1$ and $\hat{x}[N-n] = \hat{x}^*[n]$, is:

$$x_n[k] = \frac{1}{N}\left\{\hat{x}[n]e^{j\frac{2\pi}{N}nk} + \hat{x}[N-n]e^{j\frac{2\pi}{N}(N-n)k}\right\} = \frac{2\,|\hat{x}[n]|}{N} cos\left(\frac{2\pi}{N}nk + \angle(\hat{x}[n])\right) \tag{3}$$

For the phase edited signal the contribution of a frequency component n is:

$$\begin{aligned} Re\left\{p_n[k]\right\} &= \frac{|\hat{x}[n]|}{N} Re\left\{ \frac{\hat{x}_L^+[n]}{|\hat{x}_L^+[n]|} e^{+j\frac{2\pi}{N}nk} + \frac{\hat{x}_L^-[n]}{|\hat{x}_L^-[n]|} e^{-j\frac{2\pi}{N}nk} \right\} \\ &= \frac{|\hat{x}[n]|}{N} \Bigg[cos\left(\angle\hat{x}_L^+[n]\right) cos\left(\frac{2\pi}{N}nk\right) - sin\left(\angle\hat{x}_L^+[n]\right) sin\left(\frac{2\pi}{N}nk\right) + \\ &\quad + cos\left(\angle\hat{x}_L^-[n]\right) cos\left(\frac{2\pi}{N}nk\right) - sin\left(\angle\hat{x}_L^-[n]\right) sin\left(\frac{2\pi}{N}nk\right) \Bigg] \\ &= \frac{2\,|\hat{x}[n]|}{N} cos\left(\frac{2\pi}{N}nk + \frac{\angle\hat{x}_L^+[n] - \angle\hat{x}_L^-[n]}{2}\right) cos\left(\frac{\angle\hat{x}_L^+[n] + \angle\hat{x}_L^-[n]}{2}\right) \end{aligned} \tag{4}$$

where $\hat{x}_L^+[l] = \hat{x}[l] + L[l]$ and $\hat{x}_L^-[l] = \hat{x}^*[l] - L[l]$. For the enhancement of low-energy vibrational signals, in contrast to the case of speech signal denoising [17], the last step is the calculation of the residual. Using Eqs. (3) and (4), the new phase edited signal is given by

$$x_{pe,n}[k] = x_n[k] - Re\left\{p_n[k]\right\} = K[n] \frac{2\,|\hat{x}[n]|}{N} cos\left(\frac{2\pi}{N}nk + \angle\hat{x}[n] + \theta[n]\right). \tag{5}$$

where $K[n]$ and $\theta[n]$ are the gain and phase, respectively, of the PE residual filter, given by

$$\begin{aligned} K[n] &= \left[1 + cos^2\left(\frac{\alpha[n] + \beta[n]}{2}\right) - 2cos\left(\frac{\alpha[n] + \beta[n]}{2}\right) cos\left(\frac{\beta[n] - \alpha[n]}{2}\right)\right]^{1/2} \\ \theta[n] &= cos\left(\frac{\alpha[n] + \beta[n]}{2}\right) sin\left(\frac{\beta[n] - \alpha[n]}{2}\right) \Bigg/ \left[cos\left(\frac{\beta[n] - \alpha[n]}{2}\right) cos\left(\frac{\alpha[n] + \beta[n]}{2}\right) - 1\right] \end{aligned} \tag{6}$$

The angles $\alpha[n]$ and $\beta[n]$ are shown in Fig. 5(a) and defined as $\alpha[n] = \angle\hat{x}[n] - \angle\hat{x}_L^+[n]$, $\beta[n] =$

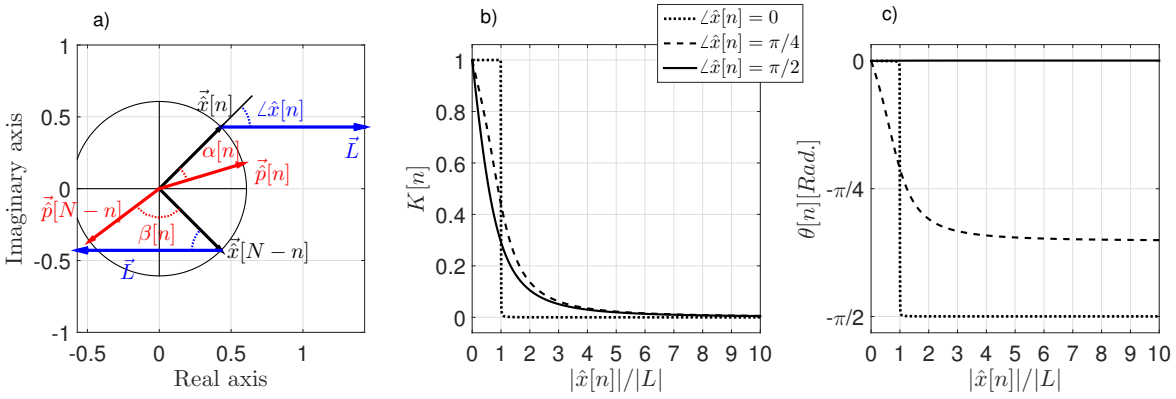

Figure 5: Effect of the proposed PE residual method on a single frequency component of a signal. (a) Example of vector rotation in the complex plane. In black the original vector and its complex conjugate, in blue the added vectors and in red the rescaled resultants. The angles $\alpha[n]$ and β in red, positive direction counterclockwise. (b) Variation of gain $K[n]$ with signal component amplitude. (c) Variation of phase modification $\theta[n]$ with signal component amplitude.

$\angle\hat{x}^*[n] - \angle\hat{x}_L^-[n]$ and for $0 \leq \angle\hat{x}[n] \leq \pi$

$$\alpha[n] = cos^{-1}\left[\frac{|\hat{x}[n]|/|L| - cos\left(\pi - \angle\hat{x}[n]\right)}{\left(1 + |\hat{x}[n]|^2/|L|^2 - 2cos\left(\pi - \angle\hat{x}[n]\right)|\hat{x}[n]|/|L|\right)^{1/2}}\right]$$
$$\beta[n] = cos^{-1}\left[\frac{|\hat{x}[n]|/|L| - cos\left(\angle\hat{x}[n]\right)}{\left(1 + |\hat{x}[n]|^2/|L|^2 - 2cos\left(\angle\hat{x}[n]\right)|\hat{x}[n]|/|L|\right)^{1/2}}\right] \tag{7}$$

while for $\pi < \angle\hat{x} < 2\pi$, $\alpha = -\alpha$ and $\beta = -\beta$. Figures 5(b-c) show respectively the gain $K[n]$ and the phase modification $\theta[n]$ as a function of the signal component amplitude $|\hat{x}[n]|/|L|$, for three different values of $\angle\hat{x}[n]$.
According to Eq.(5)

$$\begin{aligned} x_{pe,n}[k] &\simeq x_n[k] && for\ |\hat{x}[n]| \ll |L| \\ x_{pe,n}[k] &\simeq 0 && for\ |\hat{x}[n]| \gg |L| \end{aligned} \tag{8}$$

therefore PE can be interpreted as a thresholding in the amplitude of the spectral components where the level of the threshold is controlled setting $\lambda = |L|$. Spectral components with amplitude smaller than the threshold are decreased, independently of their frequency, since Eq. 5 is valid $\forall n$.

2.3.1. Application of PE to vibrational signals

This section describes the mechanism behind the efficacy of the proposed approach for enhancing bearing signals in gearbox vibrations. The vibrations from gears can be considered as amplitude and phase modulations of the fundamental tooth meshing frequency and its harmonics. In the frequency domain they will contribute a series of peaks [18] among which is possible to select

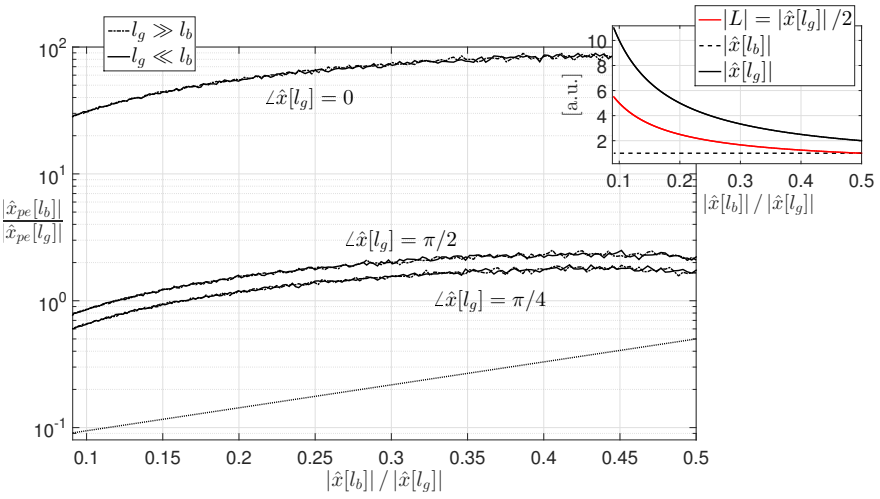

Figure 6: Performance of PE for different values of $\angle\hat{x}[l_g]$, $|\hat{x}[l_g]|/|L|$ and $|\hat{x}[l_b]|/|L|$. The dotted line is at the bisector and shows the level of no increase in the ratio $|\hat{x}[l_b]|/|\hat{x}[l_g]|$

the frequency l_g giving the peak with the smallest amplitude:

$$l_g = \min_{0<l\leq N/2} |\hat{x}[l]|$$

In contrast vibrations from a defective bearing will contribute a continuous spectrum around the resonant frequency of the impacts, among which the frequency l_b corresponds to the component with the biggest amplitude:

$$l_b = \max_{0<l\leq N/2} |\hat{x}[l]|$$

Most commonly, the gear meshing vibrations dwarf the signals from the bearings, and in the case of $|\hat{x}[l_g]| \gg |\hat{x}[l_b]|$ it is possible to select λ such that

$$|\hat{x}[l_g]| \gg \lambda \gg |\hat{x}[l_b]| \tag{9}$$

It is evident from Eqs. (8) that the result of applying the proposed PE residual filter in this case to almost completely suppress the high energy components from the gears while the low energy signal from the defective bearing is almost unchanged.

If, however, instead of Eq.(9) the vibrational composition is such that $|\hat{x}[l_g]| > \lambda > |\hat{x}[l_b]|$ the performance of PE depends on the phases of the spectral components of interest and on their ratios in respect to λ. A numerical investigation of the performance of PE for such a case is presented in Fig.(6). Three different values of $\angle\hat{x}[l_g]$ are considered: $0, \pi/4$ and $\pi/2$. For each of these phases the component from the bearing is considered to have a phase such that $\angle\hat{x}[l_b] = \angle\hat{x}[l_g] + \Delta\phi$. As shown in the inset plot, the amplitude of $|\hat{x}[l_b]|$ is kept constant while $|\hat{x}[l_g]|$ varies and $|L|$ is taken as $|\hat{x}[l_g]|/2$. The plot shows the mean of a Monte Carlo simulation comprised of 10^3 cases of $\angle\hat{x}[l_b]$ analysed for each value of $\angle\hat{x}[l_g]$. In the worst case, the ratio of the bearing signal over the gear signal, after the application of PE, is increased of a factor 4. Furthermore as expected from Eq.(5) the method does not depend on the frequency of the peaks and it shows the same performance for $l_b \gg l_g$ and $l_b \ll l_g$. Thus it is expected that the proposed method will provide useful signal conditioning wherever a separation in amplitudes exists between gear and bearing signals, and not only where that separation is large.

Finally a numerical investigation of PE on the same data as in Fig. 3 is presented, however in this simulation the level of noise is higher, with $\sigma = 0.2$. Figures 7(a) and (b), in black, show respectively a portion in the time domain and the amplitude spectrum of the raw signal. Firstly method II is implemented in order to check the performance in this noisy scenario. In Fig. 7(c) is shown the envelope of x_{zc}. It is not possible to identify the seven impulses from the bearing signal, unlike Fig. 3(c).
Fig. 7(d) shows the result of the application of the PE procedure as introduced in this paper. The absolute value of the analytic representation of $x_{pe}[k]$ allows the identification of the seven impulses in this noisy condition. The results of the application of method I for this numerical simulation are not shown here: SK selects the correct band for demodulation and the $x_{acep+sk}$ signal is comparable to x_{pe} in 7(d). The amplitude spectrum of x_{pe} is shown in Fig. 7(b) in yellow, as anticipated the phase editing method yields a signal with strongly decreased high energy components. In order to quantify the phase distortion due to the rotation of the Fourier vectors, Fig. 7(e) shows the difference between the unwrapped phases $\angle\hat{x}_{pe}[l]$ and $\angle\hat{x}[l]$, the frequency is zoomed around the resonance band $15 - 17[kHz]$. Phase differences of 2π are mainly identified,

yielding a good precision in the impulses occurrences as can be noticed by comparing Fig. 3(a-c) and Fig. 7(d).

3. Experimental setups

The experimental data sets used for the comparison between the three methods are collected from a gearbox dynamic simulator located in the laboratories of the Flanders Make VZW in Leuven, Belgium. The test bench is employed in two different configurations, namely Flanders Make 1 (FM1) and FM2. Also data sets publicly available on the websites of the Prognostics and Health Management Society (PHM) [19] and Machinery Failure Prevention Technology (MFPT) [20] are analysed.
Figures 8(a) and (b) show the photograph and schematic top view of the test setup FM1. The apparatus consists of two gearboxes, namely a parallel-shaft gearbox (2) and a perpendicular-shaft gearbox (3). The input shaft (8) of the parallel-shaft gearbox is driven through a flexible coupling, by an electric induction motor (1). The shaft rotational speed of the motor can be varied from 0 to 3000 rpm. The output shaft of the perpendicular shaft gearbox (9) is coupled with a magnetic-particle brake (4), where the torque can be adjusted from 0 to 50 Nm. The accelerometer $A\#1$ with sensitivity of $100[mV/g]$ and $\pm 5\%$ response deviation in the frequency range of $0.5[Hz]$ to $5[kHz]$ is used as vibrations transducer. The parallel-shaft gearbox (3) was designed and built with three parallel-shafts. Four helical gears are arranged in the gearbox such that it has two-stage reductions. Two straight bevel gears are assembled on their corresponding shaft, in the perpendicular-shaft

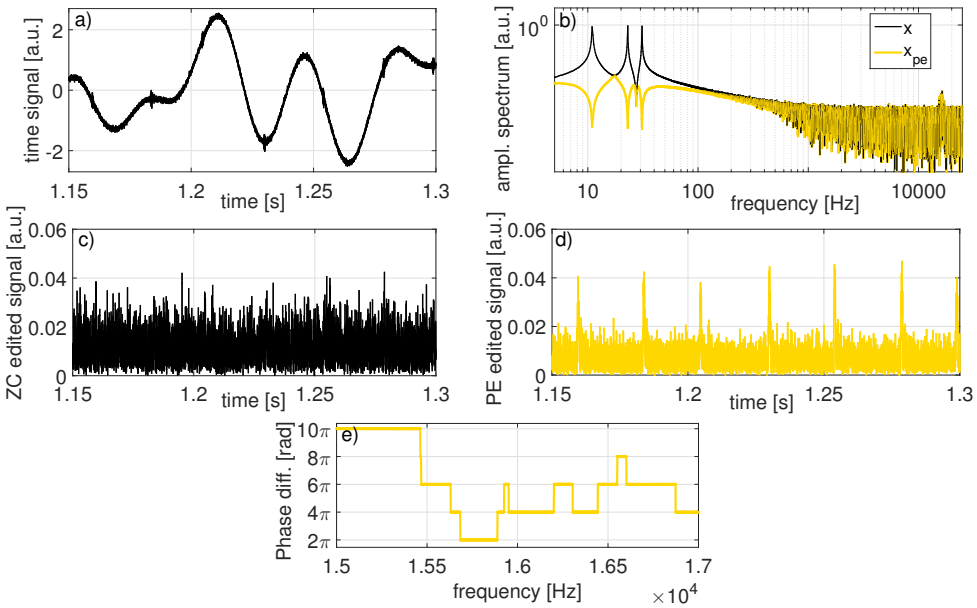

Figure 7: Application of the PE method to a numerical simulation. The tested signal is the same as in Fig. 3 with an increased level of noise. (a) The raw signal in time domain. (b) Amplitude spectrum of the raw and phase edited residual signals in black and yellow respectively. (c) Absolute value of the analytic representation of the ZC edited signal $x_{zc}[k]$: no clear impacts can be identified. (d) Absolute value of the analytic representation of the PE residual signal $x_{pe}[k]$: the seven impacts are clearly identified. (e) Phase difference between $x_{pe}[l]$ and $x[l]$ in the frequency range of the structural resonance excited by bearing impacts.

gearbox, so there is only one-stage reduction. Two different bearing types are assembled in the test setup, namely MB ER-14K and MB ER-16K. Two bearings ($B\#1$ and $B\#2$) of the former type are mounted in the parallel-shaft gearbox to support the input shaft (8) and the other bearings (from $B\#3$ to $B\#8$) of the latter type are mounted to support the other shafts. Details of the gears assembled in the two gearboxes and the bearings theoretical fault frequencies are listed in Table 1. For the setup FM1 two different types of defective bearings are tested separately, firstly a bearing with an inner race fault, secondly one with a rolling element fault, both are mounted on the input side of the second shaft of the parallel-shaft gearbox, namely ($B\#3$). Two different load conditions are tested for each configuration, $6[Nm]$ and $20[Nm]$ on the output shaft (9). Furthermore four constant speeds are investigated, on the shaft with the defective bearing they correspond to: $f_{shaft} = \{2.9;\ 8.7;\ 11.6;\ 14.5\}\ [Hz]$. The accelerometer of FM1 provides a total of 24 sets, 8 in healthy condition, 8 with inner race fault and 8 with rolling element faults. The length of each recorded vibrational signal is 5 seconds and the sampling rate is $51.2[kHz]$.
Figure 8(c) shows the schematic top view for the test setup FM2. The differences between FM1 and FM2 are that the perpendicular-shaft gearbox is removed and that the load is applied directly on the output shaft of the parallel-shaft gearbox. Gears and bearings are the same as in the FM1 setup and details are listed in Table 1. The accelerometer $A\#2$ has same specifications of $A\#1$ used in FM1, but is placed on the left side of the gearbox. For the setup FM2 a defective bearing

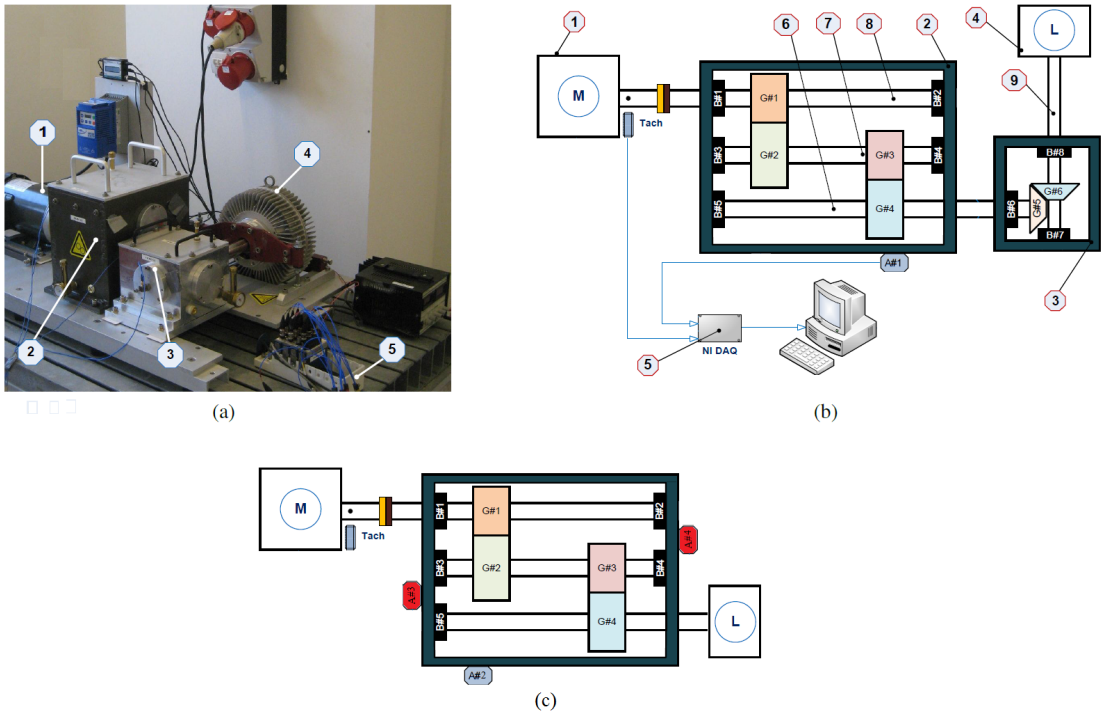

Figure 8: Experimental setups: (a) and (b) are respectively a photograph and the schematic top view of the setup Flanders Make 1 (FM1), the defective bearing is $B\#3$. (c) Schematic top view of FM2 and PHM. In blue the accelerometer used in the FM1 setup, with $B\#3$ as defective bearing. In red the accelerometers used in the PHM setup, with $B\#1$ as defective bearing.

Component	# of teeth	Type
G#1	29	Helical
G#2	100	Helical
G#3	36	Helical
G#4	90	Helical
G#5	20	Straight bevel
G#6	40	Straight bevel

Bearing defect	Cyclic Order $\times f_{shaft}$
Fundamental train freq. (FTF)	0.402
Ballpass freq. outer race (BPFO)	3.572
Ballpass freq. inner race (BPFI)	5.430
Ball spin freq. (BSF)	2.322

Table 1: Gears and defective bearing characteristics of the FM setups

MB ER-16K with an inner race fault is mounted on the input side of the second shaft of the parallel-shaft gearbox, namely ($B\#3$) and again two different load conditions are tested $6[Nm]$ and $20[Nm]$. Seven constant speeds are investigated, specifically on the shaft with the defective bearing they correspond to: $f_{shaft} = \{0.29;\ 0.58;\ 0.87;\ 1.16;\ 1.45;\ 1.74;\ 2.9\}\ [Hz]$. The length of the recorder vibrational signal is selected to obtain 15 revolutions for each f_{shaft} and the sampling rate is $51.2[kHz]$. The accelerometer of FM2 provides a total of 28 data sets, and 14 in healthy condition and 14 with inner race faults.
Figure 8(c) shows as well the schematic top view for the test setup used by the PHM. In this apparatus two accelerometers ($A\#3$-$A\#4$) are used as vibrations transducers with sensitivity of $10mV/g$, $\pm 1\%$ response deviation and resonant frequency $> 45[kHz]$. Furthermore all the bearings (from $B\#1$ to $B\#5$) assembled in the PHM test setup are MB ER-10K. Details of the gears and the bearings theoretical fault frequencies are listed in Table 2. Several combinations of faulty gears and bearings are present in the PHM dataset, here the following case is analysed: defective bearing with an inner race fault mounted on the input side of the first shaft, namely ($B\#1$) tested in two different load conditions. In this condition the gear on the output shaft, namely ($G\#4$), has a sheared key-way. Five constant speeds are investigated: $f_{shaft} = \{30;\ 35;\ 40;\ 45;\ 50\}[Hz]$, the length of each recorded vibrational signal is 4 seconds and the sampling rate is $200/3[kHz]$. Each accelerometer of the PHM setup provides a total of 20 data sets, 10 in healthy condition and 10 with inner race faults, only the data from the accelerometer on the input side, namely ($A\#3$), will be analysed.

Data sets from different types of defective bearings are stored in the MFPT database. Here three vibration signals from a bearing with an outer race defect and three from an healthy bearing are analised. The shaft rotational speed of the apparatus is $25[Hz]$, the sample rate is $97.656[kHz]$ and the length of each record is 6 seconds, only one load conditions is tested. The defective bearing is a NICE and the theoretical fault frequencies are listed in Table 3. The speed of the shaft with the defective bearings is $f_{shaft} = 25\ [Hz]$.

Component	# of teeth	Type
G#1	32	Spur
G#2	96	Spur
G#3	48	Spur
G#4	80	Spur

Bearing defect	Cyclic Order $\times f_{shaft}$
FTF	0.382
BPFO	3.052
BPFI	3.984
BSF	2.474

Table 2: Gears and defective bearing characteristics of the PHM setup

Bearing defect	Cyclic Order $\times f_{shaft}$
FTF	0.406
BPFO	3.245
BPFI	5.109
BSF	2.378

Table 3: Defective bearing characteristics of the MFPT setup

4. Results & discussion

This section presents the comparison of the three methods described in Section(1), the analysis is carried out on data sets taken from the setups detailed in Section(2). A fault feature value is calculated from the normalised $S\tilde{E}S$ of the three edited signals x_{ce+sk}, x_{zc}, x_{pe}. From each experimental configuration two feature values are calculated, one for the healthy case and one for the corresponding defective case. Features are calculated as the sum of the values of the $S\tilde{E}S[l]$ in a band around the theoretical fault frequency and up to a selected harmonic: $[k \cdot f_{defect} - 0.5 \cdot f_b,\ k \cdot f_{defect} + 0.5 \cdot f_b]$, with $k = 1, 2, \ldots, n$ with n =number of harmonics.
Specifically the SES is normalized by means of the ratio:

$$S\tilde{E}S[l] = SES[l]/SES[0] \tag{10}$$

and for the ACEP+SK signal the SES is the Fourier transform of the square of the absolute value of x_{ce+sk}, while for the zero-cepstrum and phase-edited methods the SES is implemented as the Fourier transform of the square of the absolute value of the analytic representation of x_{zc} and x_{pe}. The maximisation of the objective value $r(\lambda)$ has been defined as the criterion for determining the optimal parameter λ of Eq.2 in method III:

$$r(\lambda) = \frac{F_d(\lambda)}{F_h(\lambda)} \tag{11}$$

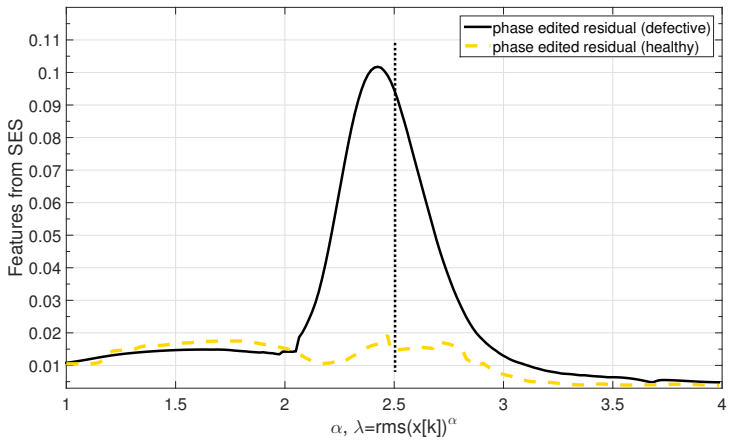

Figure 9: Tuning of the parameter λ in the PE method. Black and yellow are the features estimated from the squared envelope spectrum of the phase edited residual signal in healthy and faulthy bearing cases as a function of λ. Black vertical line is plotted at λ giving the max of the ratio $r(\lambda)$ between the features.

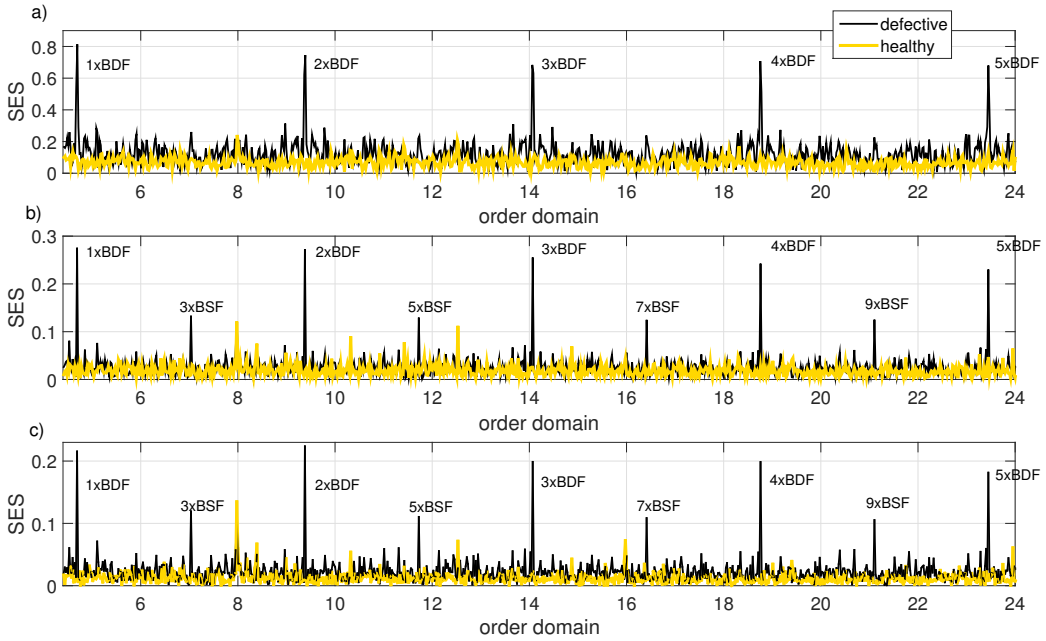

Figure 10: Performance comparison of the three methods on FM1 data set with a ball defect and $f_{shaft} = 8.7[Hz]$. SES of signal edited with (a) ACEP+SK, (b) ZC, (c) PE. Yellow and black are respectively healthy and defective conditions.

where $F_{h,d}(\lambda)$ are respectively the feature values calculated for an healthy and defective case. This is implemented in an automated way varying λ as a power α of the root mean square of the raw signal $x[k]$:

$$\lambda = rms\left(x[k]\right)^{\alpha} \tag{12}$$

Figure 9 shows an example of this procedure on a data set taken from FM1 setup with a shaft speed of $8.7[Hz]$ and a ball defect. Yellow and black lines are the features calculated from the SES of x_{pe} as a function of λ, respectively for the faulty and healthy conditions, while the vertical black line at the value of $\alpha = 2.5$ gives $max\left\{r\left(\lambda(\alpha)\right)\right\}$. For each couple of defective-healthy data sets the optimal parameter λ is computed in an automated way. In a practical application $F_h(\lambda)$ is a baseline, measured when the equipment is first commissioned and from this point onwards the signal processing can operate in a fully automated way with the estimation of features from each case to diagnose and the respective $r(\lambda)$. From all the 43 data sets tested in this paper the mean value of the tuning parameter is $\bar{\lambda} = 2.29$ with standard deviation $\sigma_\lambda = 0.13$.
Fig. 10-11-12-13 show the normalised SES in order domain for data sets from each of the four different setups, yellow is healthy case and black is defective case. The plots from top to bottom are the SES of methods I II and III, namely of x_{ce+sk}, x_{zc} and x_{pe} with the tuned λ.
In Fig. 10 the data is from the FM1 setup with a shaft speed of $8.7[Hz]$ and a ball defect, with all the methods peaks can be distinguished up to the fifth harmonic of the theoretical fault frequency $BDF = 2BSF$, as in Table 1. For the zero cepstrum and the phase edited residual signals peaks are also present at the odd harmonics of BSF, these are less clear in the SES of x_{ce+sk}. The ball spin frequency is the frequency with which the fault strikes the same race (inner or outer)

hence in general there are two impacts at BDF per basic period so that its expected that the even harmonics of BSF are dominant [11].
In Fig. 11 the data is from the FM2 setup with a shaft speed of $0.87[Hz]$ and inner race defect, with all the methods peaks can be distinguished up to the fifth harmonic of the theoretical fault frequency $BPFI$ as in Table 1. Typical sidebands due to the passing of the inner race fault into and out of the load zone are not clearly detected. This is attributable to the low speed and load conditions: the load on the shaft with the defective bearing is $20[Nm]$
Figure 12 shows the SES of the edited data from the MFPT setup. The shaft speed is $25[Hz]$ and the defect is on the outer race of the bearing. The ACEP+SK and phase editing methods, top and bottom respectively, allow the detection up to the fifth harmonic of the theoretical BPFO, Table 2, while x_{zc} up to the third harmonic. Furthermore can be noticed some spurious peaks in the ACEP+SK and that for the zero cepstrum edited signal the amplitude of peaks is one order of magnitude smaller. The SES is normalized as in Eq.10 hence the latter is the consequence of an higher value of the DC component and less pronounced impacts in the time domain envelope of x_{zc}.
In Fig. 13 the data is from the PHM setup with a shaft speed of $35[Hz]$ and an inner race defect, peaks up to the third harmonic of the theoretical fault frequency $BPFI$ as in Table 3 can be distinguished for the three methods, for this data set also the harmonics due to the load are present. The smearing of spectral peaks could be attributable to the presence of a sheared gear, or to speed fluctuations; the data is processed without applying order tracking.
Fig. 14 presents the comparison of the performance on all the available 43 data sets. It shows

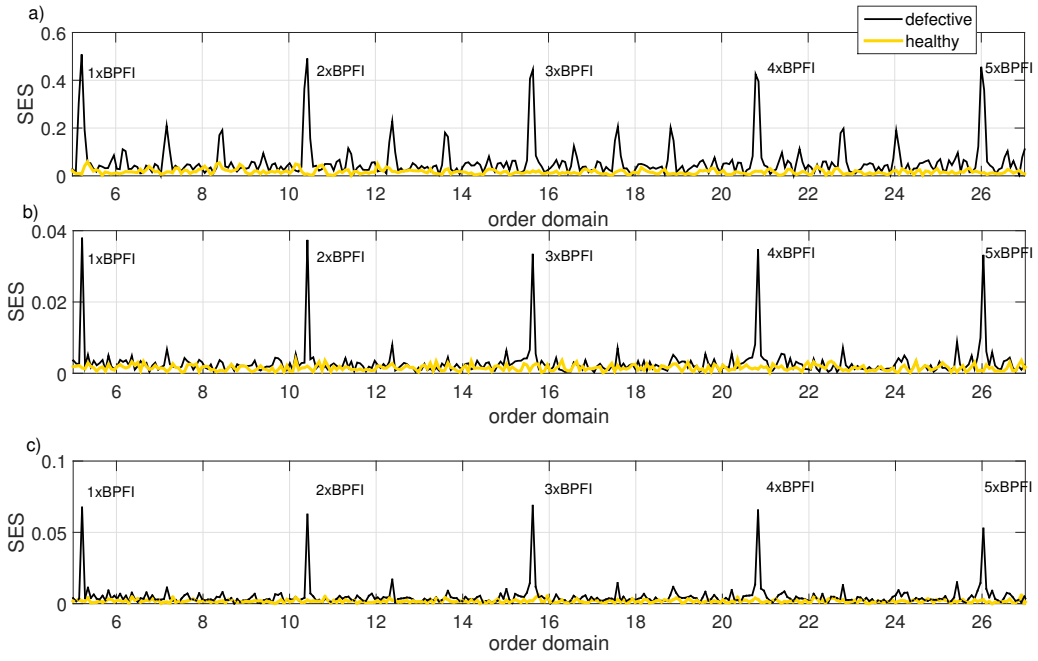

Figure 11: Performance comparison of the three methods on FM2 data set with a inner race defect and $f_{shaft} = 0.87[Hz]$. SES of signal edited with (a) ACEP+SK, (b) ZC, (c) PE. Yellow and black are respectively healthy and defective conditions.

for each case the normalised features of healthy and corresponding defective case in a scatter plot. The ideal detection of a defective bearing corresponds to the point $(1, 0)$. Features are normalised according to each set up type, defect and load. For example for FM1 in the case of a ball defect, with small load we normalise the features corresponding to the four frequencies $f_{shaft} = \{2.9;\ 8.7;\ 11.6;\ 14.5\}\ [Hz]$ and for the high load we normalise separately. Following a recent paper by Smith et al. [15] it is also studied the SES of the full bandwidth raw signals, Fig. 14 (a); as for method II and III the envelope is calculated as the absolute value of the analytic representation of the signal. Fig. 14(b) shows results from method I and Fig. 14(c-d) from method II, III. In green it is highlighted the portion of the square where the faulty features are at least three times bigger than the features of the corresponding healthy cases, providing an high confidence detection of the presence of a defect. In red the portion where the ratio between defective and healthy features is less then one; in orange where the ratio is between one and two and in yellow the portion where the ratio between defective and healthy features is between two and three and the detection is less reliable. The selection of the value three as good discriminant is made by the human inspection of the relative SES. Asterisk and triangle markers correspond respectively to data sets from FM2 and FM1 while circle and cross to PHM and MFPT data sets. The worst performance is of the envelope of the raw signal Fig. 14(a), in several cases the healthy condition has a feature higher than the faulty one, nevertheless 11 points reside in the green zone and the correct detection is accomplished. For the other three methods the results are comparable, with a slightly better performance of the proposed PE. Method I and II accomplish correct detection respectively in 31 and 30 cases, while the proposed method III in 34 cases, as shown in Fig. 14(b-

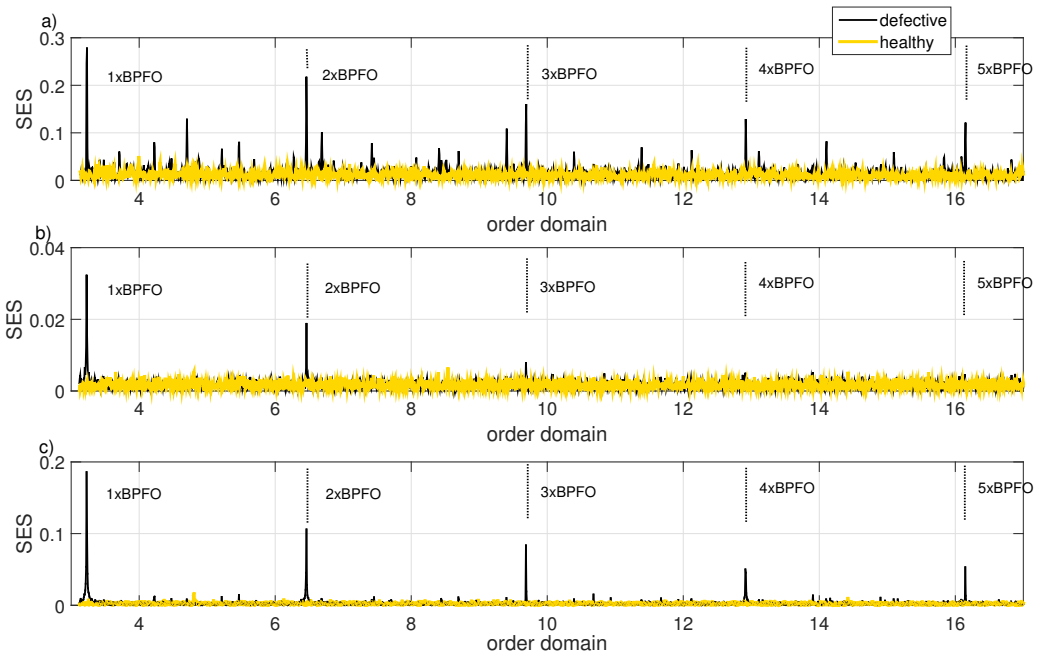

Figure 12: Performance comparison of the three methods on MFPT data set with a outer race defect and $f_{shaft} = 25[Hz]$. SES of signal edited with (a) ACEP+SK, (b) ZC, (c) PE. Yellow and black are respectively healthy and defective conditions.

c-d). Signals with a dominant bearing component can be identified in Fig. 14(a) because these are diagnosed successfully using the raw, unmodified signal only if the damage is severe. The data set from MFPT shows a case of dominant bearing damage signal, where the bearing damage was severe. This can be noticed from Fig. 14(a) with two of the three points in the green area and only one in the yellow one, indicating a good diagnosis. In Fig. 14(d) the disposition of the markers is almost unchanged, therefore when there is no need to suppress disturbing components the proposed method III is shown to be equally effective.

5. Conclusions

This paper proposed and tested a new method for the enhancement of bearing vibration signals to be used in automated condition monitoring: phase editing. The performance of the method is evaluated by means of feature estimation from the SES of the signals, with an overall correct detection rate of 79% when using the enhanced signals, compared with a rate of 25% when using the raw signals. For each case the ratio between the feature calculated from a faulty bearing and the feature from an healthy one was computed. The faulty bearing is considered correctly detected if the value of the ratio is greater than three. The effectiveness of the method compared favourably with that of two other promising contemporary enhancement approaches, which obtained performances of 72% (automated cepstral editing procedure with spectral kurtosis) and 69% (zero cepstrum). The proposed method was applied in an automated way in 43 experimental cases considering inner and outer race faults as well as ball faults, for machines operating at different speeds and loads. For each analysed case there is only one parameter to tune for a correct performance and this was

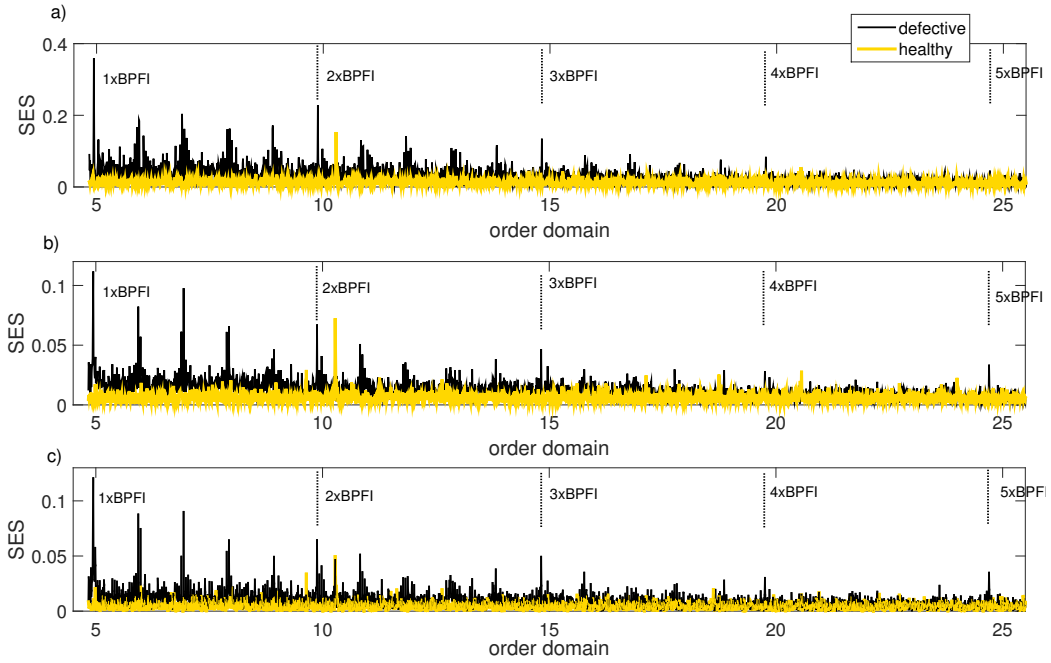

Figure 13: Performance comparison of the three methods on PHM data set with a inner race defect, a spur gear and $f_{shaft} = 35[Hz]$. SES of signal edited with (a) ACEP+SK, (b) ZC, (c) PE. Yellow and black are respectively healthy and defective conditions.

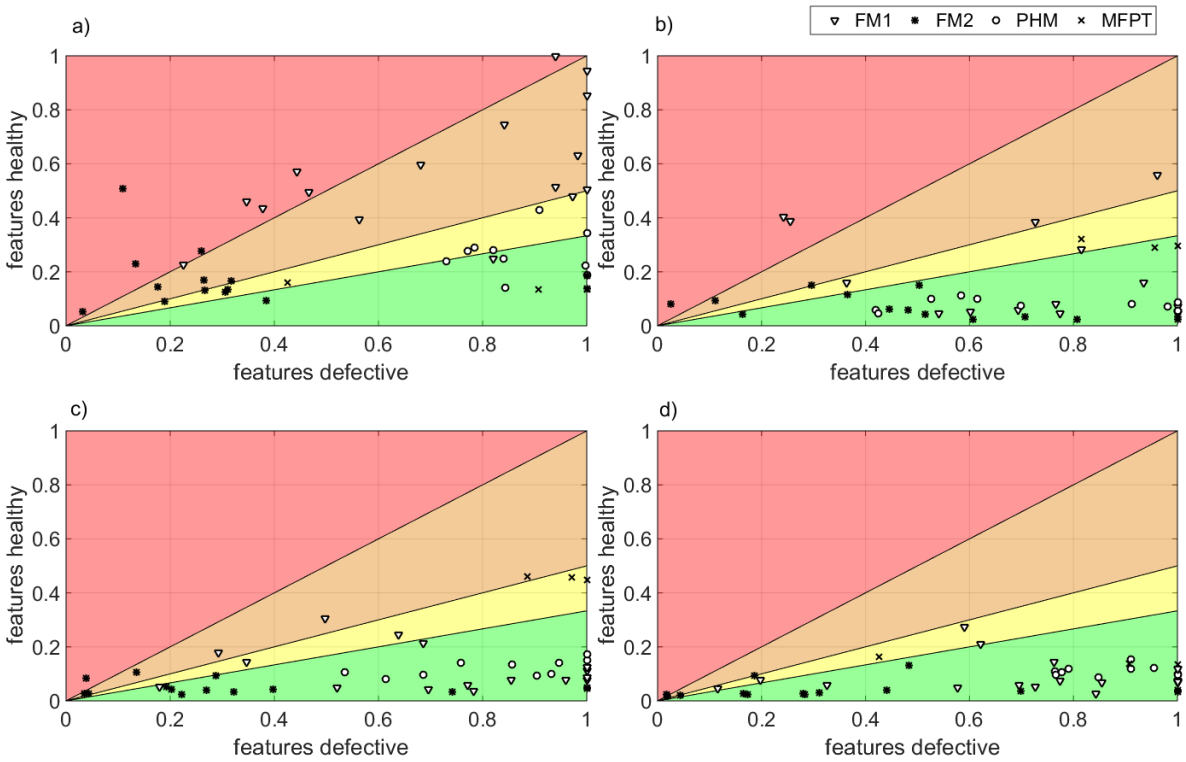

Figure 14: Healthy-defective scatter plots of features estimated from the SES: (a) of the raw signals, (b) of the signals edited with ACEP+SK, (c) ZC and (d) PE. Green represents the area where the ratio between the features is greater than 3 and the detection of the fault is very good. Yellow orange and red are the areas where the ratio is respectively between 2 and 3, 1 and 2 and smaller than 1. Triangle and asterisk pointers are FM1 and FM2 datasets while circle and cross to PHM and MFPT datasets

selected by a fully automated optimisation algorithm. Therefore only the bearing specification characteristics and rotation speed have to be adjusted for each case, so this method is suitable for use without specialist signal processing expertise from an operator. Furthermore phase editing uses the full band of the spectrum for demodulation and is computationally easy to implement, using mainly the FFT algorithm, so that it is a good candidate for an industrial application of vibration based bearing condition monitoring.

References

[1] D. Mba, B. K. N. Raj, Development of acoustic emission technology for condition monitoring and diagnosis of rotating machines: Bearings, pumps, gearboxes, engines, and rotating structures, The Shock and Vibration Digest 38 (1) (2006) 3–16.

[2] Y.-H. Kim, A. C. C. Tan, J. Mathew, B.-S. Yang, Engineering Asset Management: Proceedings of the 1st World Congress on Engineering Asset Management, Springer London, London, 2006, Ch. Condition Monitoring of Low Speed Bearings: A Comparative Study of the Ultrasound Technique Versus Vibration Measurements, pp. 182–191.

[3] I. El-Thalji, E. Jantunen, A summary of fault modelling and predictive health monitoring of rolling element bearings, Mechanical Systems and Signal Processing 6061 (2015) 252 – 272.

[4] P. McFadden, J. Smith, Vibration monitoring of rolling element bearings by the high-frequency resonance technique a review, Tribology International 17 (1) (1984) 3 – 10.
[5] J. Antoni, Cyclic spectral analysis of rolling-element bearing signals: Facts and fictions, Journal of Sound and Vibration 304 (35) (2007) 497 – 529.
[6] R. Randall, J. Antoni, S. Chobsaard, The relationship between spectral correlation and envelope analysis in the diagnostic of bearing faults and other cyclostationary machine signals, Mechanical Systems and Signal Processing 15 (5) (2001) 945 – 962.
[7] J. Antoni, The spectral kurtosis: a useful tool for characterising non-stationary signals, Mechanical Systems and Signal Processing 20 (2) (2006) 282 – 307.
[8] P. Borghesani, P. Pennacchi, R. Randall, N. Sawalhi, R. Ricci, Application of cepstrum pre-whitening for the diagnosis of bearing faults under variable speed conditions, Mechanical Systems and Signal Processing 36 (2) (2013) 370 – 384.
[9] R. B. Randall, N. Sawalhi, M. Coats, A comparison of methods for separation of deterministic and random signals, The International Journal of Condition Monitoring 1 (1) (2011) 11 – 19.
[10] B. Kilundu, A. P. Ompusunggu, F. Elasha, D. Mba, Effect of parameters setting on performance of discrete component removal (dcr) methods for bearing faults detection, in: Second European conference of the prognostics and health management (PHM) society, Nantes, France, 2014.
[11] R. B. Randall, J. Antoni, Rolling element bearing diagnostics a tutorial, Mechanical Systems and Signal Processing 25 (2) (2011) 485 – 520.
[12] A. P. Ompusunggu, T. A. Bartic, Automated cepstral editing procedure (acep) for removing discrete components from vibration signals, in: The twelfth International Conference on Condition Monitoring (CM) and Machinery Failure Prevention Technologies (MFPT), The Oxford Hotel, Oxford,UK, 2015.
[13] J. Antoni, Fast computation of the kurtogram for the detection of transient faults, Mechanical Systems and Signal Processing 21 (1) (2007) 108 – 124.
[14] A. Oppenheim, J. Lim, The importance of phase in signals, Proceedings of the IEEE 69 (5) (1981) 529–541.
[15] 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 6465 (2015) 100 – 131.
[16] K. Wójcicki, M. Milacic, A. Stark, J. Lyons, K. Paliwal, Exploiting conjugate symmetry of the short-time fourier spectrum for speech enhancement, Signal Processing Letters, IEEE 15 (2008) 461–464.
[17] K. Paliwal, K. Wójcicki, B. Shannon, The importance of phase in speech enhancement, Speech communication 53 (4) (2011) 465–494.
[18] P. McFadden, J. Smith, A signal processing technique for detecting local defects in a gear from the signal average of the vibration, Proceedings of the Institution of Mechanical Engineers. Part C. Mechanical engineering science 199 (4) (1985) 287–292.
[19] Prognostics and health management society (phm) datasets, `http://www.phmsociety.org/references/datasets`, accessed: 2016-Feb.
[20] Machinery failure prevention technology (mfpt) datasets, `http://www.mfpt.org/FaultData/FaultData.htm`, accessed: 2016-Feb.