

Stochastic-resonance based fault diagnosis for rolling element bearings subjected to low rotational speed

Agusmian Partogi Ompusunggu¹, Steven Devos¹, and Frederik Petre¹

¹ Flanders' Mechatronics Technology Centre (FMTC), Celestijnenlaan 300D, Heverlee, 3001 Belgium

agusmian.ompusunggu@fmtc.be

steven.devos@fmtc.be

frederik.petre@fmtc.be

ABSTRACT

Despite been introduced about more than 30 years ago, the Stochastic-Resonance (SR) theory has only been gaining considerable attention in the field of condition based maintenance (CBM) in recent years. SR is a nonlinear physical phenomenon where weak signals (*i.e.* signals with low signal-to-noise ratio) can be enhanced by a cooperative interaction of noise and periodic excitation (stimulus) in particular nonlinear systems, *e.g.* “bistable” systems. In other words, SR serves as a non-linear filter that can amplify a (periodic) signal of interest heavily masked by a large background noise by adjusting both the parameters of the nonlinear bi-stable system and the intensity of *artificially generated* noise to be added to the measurement signal. This paper discusses an improvement of SR filtering with multi scale noise tuning recently published for multi-fault diagnosis of a rolling element bearing subjected to low rotational speed and low load. Prior to application of the SR filtering, gear-related signals are removed from a measured acceleration vibration signal (*i.e.* signal pre-whitening) by means of the cepstra-based discrete components removal. The pre-whitened signal is further processed by employing an optimized SR filter in which the filter parameters are pre-determined based on the faulty data. Finally, features defined as the *ratio* between the peaks around the corresponding bearing fault frequencies and the background noise level are extracted from the spectrum of the output signal obtained from the SR filtering. A number of experiments have been carried out on a gearbox dynamics simulator with healthy and faulty bearings in order to demonstrate and verify the effectiveness of the proposed approach for bearing fault detection under low shaft rotational speed (348 rpm) and low load (1.2 Nm). For comparison purpose, the experimental data have been analyzed both with the well known “envelope” method and the proposed framework. The

experimental results show that the proposed framework outperforms the envelope method.

1. INTRODUCTION

1.1. Background

Rolling element bearings are one of critical components often reported as the major cause of total breakdowns in rotating machineries. In order to avoid any total breakdown in such machineries, an accurate and robust bearing condition monitoring and diagnostic system is therefore necessary. Nowadays, a number of bearing diagnostic techniques ranging from physically-inspired to data-driven approach have been developed and reported in literature. Among those existing techniques, the optimized High Frequency Demodulation (HFD)/envelope technique comprising several advanced signal processing steps, *e.g.* cepstra based discrete component removal (prewhitening), minimum entropy deconvolution (MED) and spectral kurtosis analysis, is one of the most successful technique. Interested readers are referred to (Randall & Antoni, 2011) for a comprehensive tutorial of this optimized envelope technique.

When a damage occurs on one of the elements (*e.g.* outer race, inner race or rolling element) of a bearing subjected to low rotational speed and low load, it can be argued that the damaged-induced vibration signal (*i.e.* impulsive signal) is severely masked by a large background noise originating from other machine components and the operating environment. As the signal-to-noise ratio (SNR) of the damaged-induced vibration signal is low, bearing fault detection under such circumstances is a very challenging task. Although the envelope technique has been successfully applied for a wide range of applications, it appears that the envelope technique has some limitations when applied on bearings under *low speed* and *low load*, as will be shown in the next section. Therefore, an alternative approach needs to be developed for fault detection of rolling element bearings under low speed and low load.

Agusmian Partogi Ompusunggu et.al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 United States License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

One possible solution for fault detection of bearings subjected to low speed and low load is stochastic-resonance based signal enhancement. The concept of stochastic resonance (SR) was originally introduced about more than 30 years ago by Benzi *et al.* (Benzi, Sutera, & Vulpiani, 1981) and has been matured by different researchers (McNamara & Wiesenfeld, 1989; Jung & Hänggi, 1991; Gammaitoni, Hänggi, Jung, & Marchesoni, 1998) to explain an amplification of signals with low SNR resulting from a cooperative effect of noise and periodic excitation (stimulus) in nonlinear systems (*i.e.* bistable systems). From the signal processing point of view, SR can thus be seen as a *non-linear signal filtering* where the filter characteristics depend on the parameters of a bistable system and the input signal consisting of periodic driving signal and artificially generated noise. In contrast to *linear filtering* where any increase in the input noise will result in a decrease in the output SNR, an important property of SR is that the output SNR first increases with increased input noise, then reaches the maximum and finally decreases again, so it is similar to a “resonance-like phenomenon”.

Since the last decade, this signal processing tool has been attracting the attention of many researchers in the CBM/PHM field because of its nature to enhance fault-induced signals embedded in large background noise and increasing computational power (H.-L. He, Wang, Leng, Zhang, & Li, 2007; Tan et al., 2009; Q. He, Wang, Liu, Dai, & Kong, 2012; Lei, Han, Lin, & He, 2013; Li, Chen, Du, Fang, & He, 2013). However, the performance of SR filtering for fault detection strongly depends on the parameters of the bistable system and the noise characteristics/intensity. Different strategies have been proposed in recent years to determine the optimal parameters for realizing SR, namely (i) simultaneous optimization of system parameters and noise intensity (Xu, Li, & Zheng, 2003; Wu, Jiang, & Repperger, 2006), (ii) frequency-shifted and rescaling (Tan et al., 2009), (iii) multi scale noise tuning (Q. He et al., 2012) and (iv) adaptive method based on sliding window (Li, Chen, & He, 2013). Among the aforementioned strategies, the multi scale noise tuning approach (Q. He et al., 2012) is particularly interesting since there is only *one parameter* to be tuned and the authors believe that this approach seems to be appropriate for practical implementation.

1.2. Problem statement

The bearing diagnostic technique based on SR with multi-scale noise tuning was originally developed by He and co-workers (Q. He et al., 2012). They experimentally verified their method on faulty bearings subjected to moderate rotational shaft-speed (700 and 1700 rpm) and high load. They have shown that this method outperforms other state-of-the-art methods. The SR based diagnostic method was developed for a case where only a single fault is present. However, in practice, multiple faults can simultaneously occur, so their

method is not readily applied for multiple-fault diagnosis. Moreover, they also derived the diagnostic conclusion based on visual interpretation on the spectrum of the enhanced signal. Hence, an automatic diagnostics cannot be generated in this manner.

1.3. Objectives

To remedy the three aforementioned gaps, this study aims at mainly improving the bearing diagnostic technique based on the SR with multi scale noise tuning, described in (Q. He et al., 2012), such that multiple-fault diagnosis becomes possible. Secondly, the objective of this study is to verify whether SR with multi scale noise tuning is relevant and robust for fault diagnosis of bearings subjected to low rotational speed and low load. Finally, features have been developed in this study in order to objectively diagnose different possible faults that may occur in a bearing. In this way, an automatic diagnostics can be realized.

1.4. Paper organization

The remainder of the paper is organized as follows. Section 2 briefly reviews the theory of classical SR and then zooms in the theory of SR with multi scale noise tuning. Section 3 discusses a new procedure for determining the optimal tuning parameter in case of multiple faults and proposes features allowing automatic bearing fault detection. Section 4 discusses the experimental method for verification of the performance of the proposed framework. Section 5 discusses the results obtained from the proposed framework and then compares the results with the ones obtained from the envelope technique. Finally, some important remarks of this study that can be taken up for future work are summarized in Section 6.

2. SR WITH MULTISCALE NOISE TUNING

To provide some basic understanding of the concept, this section first reviews briefly the classical SR theory based on a bistable system. Subsequently, the SR based on multi scale noise tuning is discussed.

2.1. Classical SR based on a bistable system

Let us consider the bistable double-well potential of a particle illustrated in Fig. 1(a) in the absence of any external stimulus (excitation), which is expressed as

$$U(x) = -\frac{a}{2}x^2 + \frac{b}{4}x^4, \quad (1)$$

where a and b denote the barrier parameters with positive real values, the potential minima are at $\pm c = \pm\sqrt{a/b}$ and the barrier height is $\Delta U = a^2/4b$. Suppose that the particle is heavily damped so that it will come to rest at one of the two minima of the potential which are located at $x = \pm c$. In the presence of a moderate amount of random excitation, the

particle will still spend most of its time near $x = \pm c$, but will occasionally cross the potential barrier ΔU located at the center $x = 0$. When the random excitation, D , is increased, the rate of transition (jumping) from one well to another one, ρ , increases. Typically, ρ grows very rapidly with D at first, but once D is large enough the barrier becomes relatively easy to overcome, ρ grows more slowly as D is further increased. This transition rate can be approximated with the Kramers rate ρ_K defined as

$$\rho_K = \frac{a}{\sqrt{2\pi}} \exp\left(-\frac{2\Delta U}{D}\right). \quad (2)$$

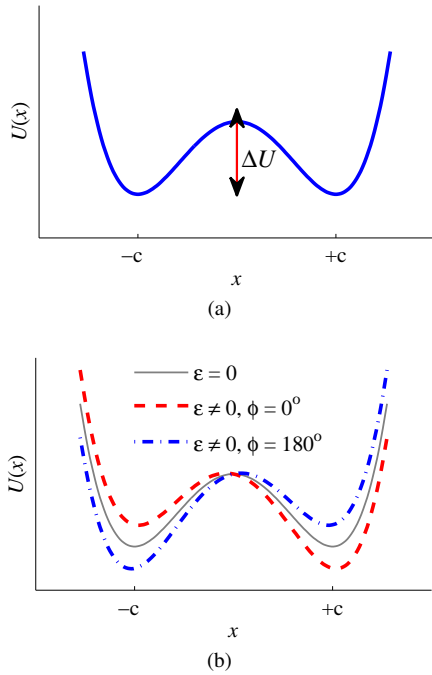


Figure 1. (a) Double-well potential with two minima at $\pm c$ and barrier height ΔU . (b) Sinusoidally modulated potential ($\epsilon \neq 0$) at two times 180° out of phase.

Assume now that a periodic (sinusoidal) excitation with the frequency of ω_0 is applied to the particle. As shown in Fig. 1(b), this additional excitation has the effect of tilting the potential first *clockwise* ($\phi = 0^\circ$), so the particle is more likely to be located at $+c$, and then *counter-clockwise* in half a cycle later ($\phi = 180^\circ$). Since it is presumed throughout of this paper that the amplitude of periodic excitation is small enough that, in the absence of any noise, it is insufficient to force the particle to move from one well to another one. It is also assumed that the period of excitation is longer than some characteristic intra-well relaxation time for the system (McNamara & Wiesenfeld, 1989). The periodic excitation has the effect on modulating the jumping (transition) rate, making the rate out of the $+$ well (located at $x = +c$), ρ^+ ,

oscillate out of phase with the rate out of the $-$ well (located at $x = -c$), ρ^- . For very small random excitation, these rates are still too slow for there to be appreciable hopping and the small periodic excitation does not play a significant role. The characteristic rate associated with this bistable system is given by $\rho = \rho^+ + \rho^-$. As D is increased from a very low value, both ρ^+ and ρ^- increase, thus consequently increasing ρ . When D is increased further such that $\rho \approx \omega_0$, there is a cooperative phenomenon taking place where incoherent noise power is feeding into the coherent motion of the particle. It becomes likely that, for example, the particle in the $+$ well to jump to the $-$ well in the first half a cycle and it is likely to jump back to the $+$ well in another half a cycle. This way, the motion of the particle from one well to another one that is coherent with the periodic excitation is considerably pronounced. With too much noise excitation, the cooperative phenomenon subsides, and the noise in feeds the noise out.

In the presence of a sinusoidal excitation, the double-well potential $U(x, t)$ of Fig. 1(b) has the form of (McNamara & Wiesenfeld, 1989)

$$U(x, t) = -\frac{a}{2}x^2 + \frac{b}{4}x^4 - x\epsilon \cos(\omega_0 t + \phi_0), \quad (3)$$

where ϵ , ω_0 , and ϕ_0 are respectively amplitude, driving frequency and initial phase of the periodic excitation. In the case of large damping (*i.e.* the inertia effect is negligible, $d^2x/dt^2 \rightarrow 0$), the equation of motion of the particle under random (white noise) excitation is

$$\frac{dx}{dt} = -\frac{\partial U(x, t)}{\partial x} + \xi(t). \quad (4)$$

Here $\xi(t)$ represents Gaussian white noise such that $\langle \xi(t) \rangle = 0$ and $\langle \xi(t)\xi(t+T) \rangle = 2D\delta(T)$ with the operator $\langle \bullet \rangle$ standing for statistical expectation, and D denotes the noise intensity. From Eqs. (1) and (4), the equation of overdamped motion of a particle under random excitation (also called as a Brownian particle) in sinusoidally modulated double-well potential can then be rewritten as

$$\frac{dx}{dt} = ax - bx^3 + \epsilon \cos(\omega_0 t + \phi_0) + \xi(t). \quad (5)$$

From signal processing point-of-view, the system described above can be seen as a non-linear filtering process applied on a noisy periodic signal. Here, three essential equivalences are revealed, where (i) the periodic excitation $\epsilon \cos(\omega_0 t + \phi_0)$ is equivalent to a periodic signal of interest, (ii) white noise excitation $\xi(t)$ is equivalent to measurement noise and (iii) the barrier parameters of a, b are equivalent to a non-linear filter parameters. When a noisy periodic signal is inputted to such a non-linear filter, the asymptotic output response $\langle x(t) \rangle_{as}$ can be approximated as (Gammaitoni et al., 1998)

$$\langle x(t) \rangle_{as} = X \cos(\omega_0 t - \phi), \quad (6)$$

where the amplitude X and the phase lag ϕ in function of noise intensity D can be expressed as

$$X(D) = \frac{\epsilon \langle X^2 \rangle_0}{D} \frac{2\rho_K}{\sqrt{4\rho_K^2 + \omega_0^2}}, \quad (7)$$

$$\phi(D) = \phi_0 + \arctan\left(\frac{\omega_0}{2\rho_K}\right), \quad (8)$$

with $\langle X^2 \rangle_0$ denoting the D -dependent variance of the stationary unperturbed system (*i.e.* $\epsilon = 0$).

Figure 2 schematically illustrates the amplitude and phase response of SR filtering applied on a periodic signal with the frequency ω_0 at different noise intensity D according to Eqs. (7) and (8). As seen in Figure 2(a), a resonance-like phenomenon is revealed, where the amplitude of the periodic signal first increases, then reaches its maximum value at $D = D_r$, finally decreases for $D > D_r$. Regarding the phase response, it is seen that the phase decreases asymptotically when the noise intensity is further increased.

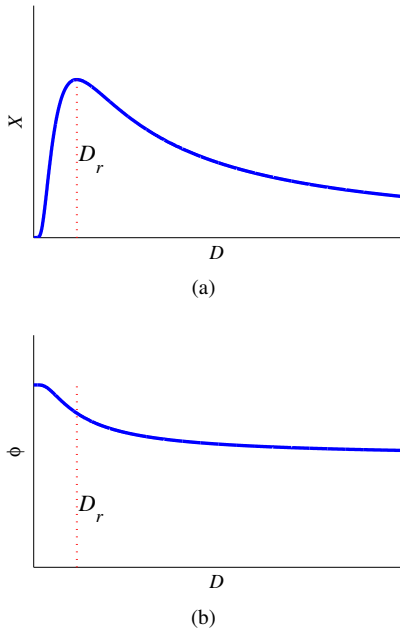


Figure 2. (a) Amplitude response and (b) phase response of SR filtering of a periodic signal with the frequency ω_0 as a function of input noise intensity D . The resonance-like phenomenon occurs at $D = D_r$.

In the above theoretical analysis of the bistable SR model, it is assumed that the frequency and the amplitude of a periodic signal as well as the noise intensity are all smaller than one (McNamara & Wiesenfeld, 1989; Gammaitoni et al., 1998). Nevertheless, through an appropriate transformation on the above-discussed bistable system, a parameter tuning approach can be derived for making the classical SR fil-

tering capable of enhancing a high frequency signal heavily masked by background noise (Leng, Leng, Wang, & Guo, 2006; Tan et al., 2009). Let us introduce two scaled variables $y = x\sqrt{b/a}$, and $\tau = at$, and define a scaling factor $K = \sqrt{a^3/b}$, hence Eq. (5) can be rewritten in a normalized form as

$$\frac{dy}{d\tau} = y - y^3 + \frac{1}{K} \left[\epsilon \cos\left(\frac{\omega_0\tau}{a} + \phi_0\right) + \xi\left(\frac{\tau}{a}\right) \right]. \quad (9)$$

From the normalized expression, it is obvious that the frequency of the signal is scaled down to $1/a$ times of the original frequency ω_0 in the new model and the frequencies of the noise are also normalized in the same form. By choosing an appropriate value for the parameter a , a high frequency (> 1 Hz) can be scaled down to be much smaller than one, thus allowing detection of a weak signal containing high frequency components.

2.2. SR with multi scale noise tuning

The classical SR filtering discussed in the previous section assumes that the measurement noise or noise to be added is white Gaussian noise. As pointed out by He and co-workers (Q. He et al., 2012), white noise adjustment for SR filtering in some cases cannot effectively induce the SR effect. To remedy this, they proposed the use of a band-limited noise at different scales, also called multi scale noise, as both types of noise result in spectra with the Lorentzian distribution (*i.e.* $1/f$ distribution) when applied to a SR filtering. Multiscale noise with $1/f$ distribution can be constructed from white noise by means of orthonormal wavelet decomposition such that the variance of the detailed wavelet coefficients d_j at the scale j respects the following rule (Flandrin, 1992)

$$\text{Var}(d_j) = \sigma^2 2^j, \quad (10)$$

with $\text{Var}(\bullet)$ and σ^2 respectively being the variance operator and a positive constant.

For enhancing a noisy periodic signal $x(t)$ with multi scale noise tuning, the signal is first decomposed into J scales by discrete wavelet transform (DWT) as proposed by (Q. He et al., 2012). The decomposition of the signal is formulated as follows

$$a_j(k) = \int x(t)v_{j,k}(t) dt, \quad (11)$$

$$d_j(k) = \int x(t)w_{j,k}^*(t) dt, \quad (12)$$

where $v_{j,k}$ and $w_{j,k}$ respectively represent the scaling function and the primary wavelet function, while a_j and d_j respectively denote the *approximation coefficients* and the *detailed coefficients*, with j being the decomposition scale. The level of decomposition J is predetermined such that d_J covers the frequency of interest in the signal $f_0 = \omega_0/2\pi$. Math-

ematically, the decomposition level J is calculated according to the following formula

$$J = \arg \left\{ j \in \mathbb{Z}^+ : \frac{f_s}{2^{j+1}} \leq f_0 \leq \frac{f_s}{2^j} \right\}, \quad (13)$$

with f_s denoting the sampling frequency and \mathbb{Z}^+ denoting a set of positive integers.

Once the signal is decomposed onto level J , a set of detail coefficients $\Delta_s = \{d_{s,1}, d_{s,2}, \dots, d_{s,J}\}$ is obtained. In a similar way, *artificially generated* white noise $\xi(t)$, with the same length as the signal and with unit standard deviation, is also decomposed onto level J resulting in another set of detail coefficients $\Delta_n = \{d_{n,1}, d_{n,2}, \dots, d_{n,J}\}$. For multi scale noise realization, the two sets of detail coefficients are merged to become a new one $\tilde{\Delta} = \{\tilde{d}_1, \tilde{d}_2, \dots, \tilde{d}_J\}$ such that the requirement for $1/f$ noise expressed in Eq. (10) can be satisfied. Let us transform Eq. (10) into the 2-base logarithmic form as below

$$\log_2 \left[\text{Var} \left(\tilde{d}_j \right) \right] = j + \alpha, \quad (14)$$

with $j = \{1, 2, \dots, J\}$ and $\alpha = \log_2(\sigma^2)$. Here, α is the tuning parameter of the multi scale noise. Thus, a new detail coefficients \tilde{d}_j at scale j can be computed as

$$\tilde{d}_j = \rho_{1,j} d_{s,j} + \rho_{2,j} d_{n,j}, \quad (15)$$

where $\rho_{1,j}$ and $\rho_{2,j}$ are weighting factors being defined as

$$\rho_{1,j} = \begin{cases} \left[\frac{2^{j+\alpha}}{\text{Var}(d_{s,j})} \right]^{\frac{1}{2}}, & \text{if } j = J; \\ 1, & \text{if } j < J, \end{cases} \quad (16)$$

and

$$\rho_{2,j} = 2^{\frac{j+\alpha}{2}} - [\text{Var}(d_{s,j})]^{\frac{1}{2}}. \quad (17)$$

Note that the approximate coefficients $a'_j s$ (*i.e.* lower frequency band) are not adjusted since they do not contain any information corresponding to the frequency of interest.

A multi scale-noise tuned signal can be reconstructed based on the adjusted detail coefficients \tilde{d}_j and the approximate coefficients of the signal a_j . This reconstructed signal is then inputted to the normalized SR filter, described in Eq. (9) with the barrier parameters of $a = 1$ and $b = 1$, for signal enhancement. To determine an optimal tuning parameter α_{opt} that can optimally enhance the signal corresponding to the frequency of interest f_0 , a grid search method is employed. Here, the following objective value η has been defined in (Q. He et al., 2012) as the criterion for determining the optimal parameter

$$\eta = \frac{X(\alpha, f_0)}{X(\alpha, f_p)}, \quad (18)$$

where $X(\alpha, f_0)$ represents the spectrum magnitude of the

output signal at the frequency of interest f_0 and at given parameter α and $X(\alpha, f_p)$ represents the highest magnitude (peak) in the spectrum of the output signal at the frequency f_p and at given parameter α . Hence, the optimal parameter α_{opt} can now be determined by identifying the parameter α_k giving the largest value on $\eta(\alpha_k)$ in a predefined finite set of reasonable values $\alpha_l \leq \alpha_k \leq \alpha_u, k = 1, 2, \dots, N$. Mathematically, identification of the optimal parameter α_{opt} can be formulated as

$$\alpha_{opt} = \text{argmax} \{ \eta(\alpha_k) : \alpha_l \leq \alpha_k \leq \alpha_u \}. \quad (19)$$

3. METHODOLOGY

This section first discusses the overall methodology proposed in this work. Subsequently, a procedure for determining the optimal tuning parameters required in the proposed method is zoomed in. Finally, features corresponding to different localized faults in a bearing (*i.e.* inner race fault, outer race fault and rolling element fault) are formulated.

Let us assume that the optimal tuning parameters, $\alpha_{opt,n}; n = 1, 2$, the decomposition level J , and artificially generated white noise ξ with the record length of the vibration signal, are given for a certain operating condition. Based on the theory of SR with multi scale noise tuning discussed in the previous section, a procedure for multiple fault diagnosis of a bearing subjected to low speed can be developed, which is schematically described in Figure 3. It is seen in the figure, the overall methodology consists of *three main steps*.

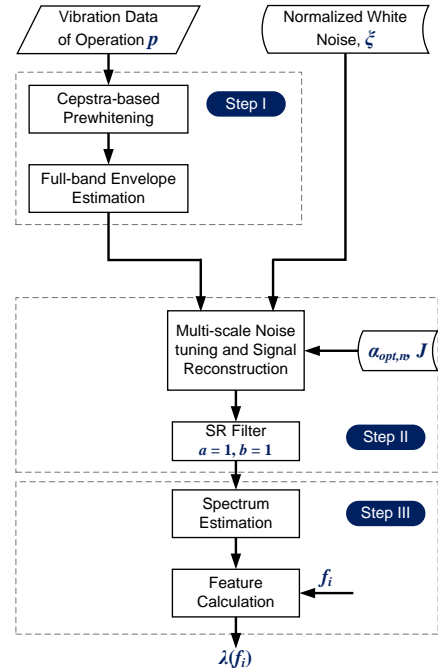


Figure 3. Signal processing steps and feature calculation given the optimal parameter α_{opt}

In the first step, a raw vibration signal is prewhitened in order to remove strong periodic components (typically coming from the dynamics of gears and shafts) in the signal. Due to its capability of removing periodic components and the sidebands with a less computational burden, the cepstra-based technique (Randall & Sawalhi, 2011) is employed in this paper. Note that the prewhitening is a necessary step since the frequencies of the periodic components (harmonics and sidebands) in some cases are close to the theoretical bearing fault frequencies. Then, the full-band enveloped signal is estimated by means of Hilbert transformation.

In the second step, given $\alpha_{opt,n}; n = 1, 2, J$ and ξ , SR filtering with multi-scale noise tuning as discussed in the previous section is then applied to the enveloped signal for the signal enhancement. Note that the level of wavelet decomposition J is calculated using Eq. (13) based on the target frequency (*i.e.* frequency of interest) f_0 , Daubechies 4 wavelet (db4) is employed for the multi scale noise tuning and the barrier parameters are fixed to one, *i.e.* $a = b = 1$.

In the third step, bearing fault features that will be discussed in Section 3.2 are calculated from the spectrum of the enhanced signal. Here, the feature values can be seen as the representation of different localized faults in a bearing (*i.e.* inner race fault, outer race fault and rolling element fault). In this way, multi fault diagnosis of a bearing can be realized in an automatic way.

3.1. Parameter tuning

As in practice the inner race fault frequency (BPFI), outer race fault frequency (BPFO) and ball damage frequency (BDF) are close to each other (typically $BPFO < BDF < BPFI$), while the cage damage frequency (FTF) is typically much lower than those three frequencies (*i.e.* $FTF \approx BPFO/n$, with n denoting the number of rolling elements), two target frequencies can therefore be proposed for multi fault diagnosis purpose. The first target frequency $f_{0,1}$ corresponds to FTF and the second one $f_{0,2}$ corresponds to the average of the other fault frequencies (BPFO, BDF and BPFI). Hence, two tuning parameters $\alpha_{opt,1}$ and $\alpha_{opt,2}$ are required for multiple fault diagnosis.

Figure 4 shows a flowchart for determining the optimal parameters $\alpha_{opt,1}$ and $\alpha_{opt,2}$ for a certain operating condition p , where vibration signal obtained from measurements in a faulty state is used as training data. In general, the flowchart can be divided into four steps where the first two steps of the parameter tuning are similar to the first two steps of the overall methodology (compare Figures 3 and 3). In the third step, a learning curve is generated by computing the objective value η for the predefined finite set of $\alpha_k \in [\alpha_l, \alpha_u]$. In the fourth step, the optimal parameters can be identified according to Eq. (19).

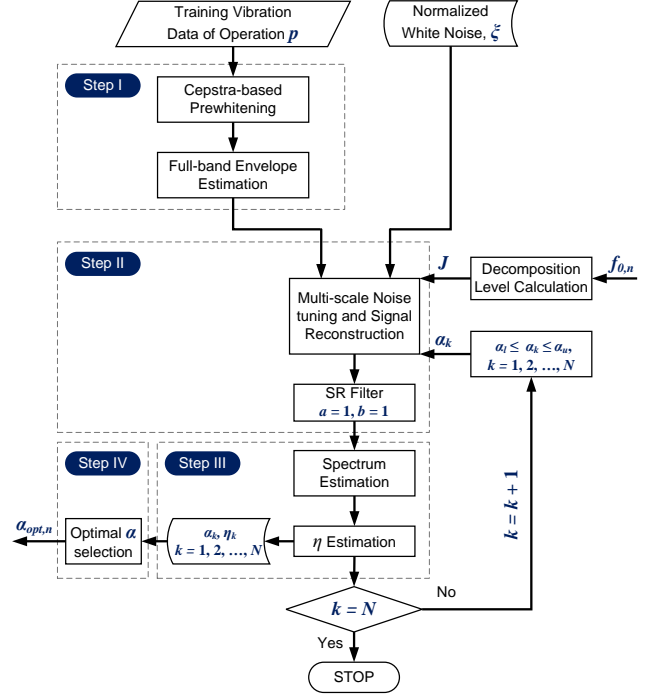


Figure 4. Flowchart for determining the optimal tuning parameter α_{opt} at a certain operating condition.

3.2. Feature definition and calculation

In a healthy state, the magnitudes of the spectrum around the theoretical fault frequencies (BPFO, BDF, BPFI, and FTF) are *ideally* zero. However, due to unavoidable noise, the spectrum magnitudes around those frequencies are more or less the same as the floor noise level. Depending on the type of fault present in a bearing, a high peak around the corresponding fault frequency in the frequency spectrum is expected in a faulty state. Moreover, peaks around some harmonics of the fault frequencies may also appear in the spectrum. To quantitatively distinguish between the healthy and faulty state, a feature dedicated to a specific fault is necessary. Based on the preceding rationale, the feature can be defined as the ratio between the spectrum magnitude around the corresponding frequency (BPFO, BDF, BPFI, or FTF) and the averaged floor noise level

$$\lambda(f_i) = \sum_{k=1}^3 \left[\frac{X_p(kf_i)}{X_n} \right], \quad (20)$$

where f_i denotes the actual fault frequency (BPFO, BDF, BPFI, or FTF), $\lambda(f_i)$ denotes the feature value corresponding to a certain fault frequency f_i , $X_p(kf_i)$ denotes the spectrum magnitude at kf_i with $k = 1, 2, 3$ and X_n denotes the averaged floor noise level. It can be argued that the actual fault frequencies in the spectrum typically deviate from the theoretical ones (1-2%) due to slip of the rolling elements. The floor noise level X_n is computed as follows.

First, the frequency range of interest is determined, which is considered in this study as two times higher than BPF1 (*i.e.* $0 < f < 2 \times \text{BPF1}$). Consecutively, all the local maxima in this frequency range are identified and then removed from the spectrum. Finally, the floor noise level is estimated by taking the average of the residual spectrum in the frequency range of interest.

4. EXPERIMENTAL METHOD

To experimentally verify the proposed bearing diagnostic method, vibration data have been collected from a gearbox dynamic simulator (GDS) setup with healthy and faulty bearings. The GDS is a dedicated test setup that enables one to simulate a healthy and faulty gearbox with different type of faults, namely localized bearing faults (*i.e.* inner race, outer race and rolling element faults) and gear faults (eccentricity, chipped tooth, missing tooth, cracked tooth, backlash, and profile error). The test setup and the experimental procedure are described in the forthcoming subsections.

4.1. Apparatus

Figure 5 shows the photograph and the schematic-top-view of the test setup used in this study. The GDS consists of two gearboxes connected to each other, namely a parallel-shaft gearbox (3) and a perpendicular-shaft gearbox (4). The input shaft (10) of the parallel-shaft gearbox is driven through a flexible coupling by an induction electric motor (2) which is controlled by a variable-frequency-drive (1). The shaft rotational speed of the motor can be varied from 0 to 3000 rpm, with either a stationary mode or a transient mode (run-up/run-down). The output shaft (13) of the perpendicular-shaft gearbox is coupled with a magnetic-particle brake (5), where the torque applied to the output-shaft can be adjusted by the brake controller (8) from 0 to 50 Nm. The control signals to the motor drive (1) and to the brake controller (8) are sent out by a PC (7) using dedicated programs. Vibration signals in the axial (perpendicular to the gearboxes wall) and horizontal-radial directions are sensed by three triaxial piezoelectric (ICP type) accelerometers, which are mounted at different positions; two of them (A#1 and A#2) respectively on the drive side and non-drive side of the parallel-shaft gearbox and the third one (A#3) on the drive side of the perpendicular-shaft gearbox. The three accelerometers have the same sensitivity of 100 mV/g with $\pm 5\%$ response deviation in the frequency range of 0.5 Hz to 5 kHz. An optical tachometer system is used to measure the shaft rotational speed of the motor where the light from the sensor head is directed to a reflective tape attached on the motor shaft, thus generating one pulse signal per revolution. All the signals are synchronously acquired by a National Instruments (NI) data acquisition system (6) and then stored to the PC with a Labview program.

The parallel-shaft gearbox (3) was designed and built with

three parallel-shafts as seen in Figure 5(b). 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 gearbox, so there is only one-stage reduction. Some details of the gears assembled in the two gearboxes are summarized in Table 1. Note that the total reduction factor from the motor shaft to the output shaft of the perpendicular-shaft gearbox can be easily calculated based on the number of teeth of the gears, that is equal to $(100/29) \times (90/36) \times (40/20) = 18.2$.

Component	Number of teeth	Comment
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

Table 1. Specifications of the assembled gears.

Two different bearing types are assembled in the test setup, namely MB ER-14K and MB ER-16K. Two identical bearings (B#1 and B#2) of the former type are mounted in the parallel-shaft gearbox to support the input shaft (10) and the other bearings (from B#3 to B#8) of the latter type are mounted to support the other shafts, see Figure 5(b). They all are deep-groove ball bearings with the same geometrical specifications, thus the same theoretical fault frequencies as listed in Table 2.

Number of rolling element	9
Diameter of rolling element	0.3125 inch
Pitch diameter	1.516 inch
FTF	$0.402 \times f_r$ Hz
BPFO	$3.572 \times f_r$ Hz
BDF	$4.644 \times f_r$ Hz
BPFI	$5.430 \times f_r$ Hz

Table 2. Specifications of the two bearing types and the shaft-speed-dependent fault frequencies, with f_r denoting the shaft rotational speed in Hz. Note that the BDF is equal to two times of the ball spin frequency (BSF).

4.2. Test Procedure

Faulty state simulation

In a faulty state, two bearings with localized damages were concurrently mounted on the intermediate shaft (11) of the parallel-shaft gearbox. One bearing with combined faults, *i.e.* inner race fault, outer race fault and rolling element fault, was mounted on the intermediate shaft (12) of the drive side (B#3) while another one with single fault, *i.e.* rolling element fault, was mounted on the same shaft but on the non-drive side (B#4).

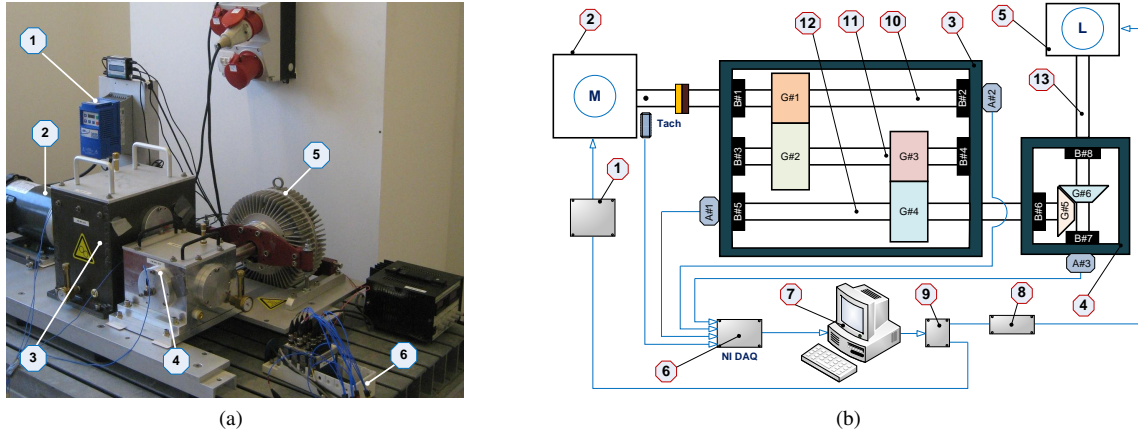


Figure 5. Experimental apparatus: (a) photograph and (b) schematic top view

Operating condition

A clear distinction between a high speed and low speed machinery has been made in literature (Kim, Tan, Mathew, & Yang, 2006; Kosse & Tan, 2006; Lin, Kim, & Tan, 2013), where the speed border is around 600 rpm. To simulate a machinery subjected to low speed, in this study, the shaft rotational speed of the electric motor of the gearbox setup was set to 1200 rpm. As the reduction ratio between the input shaft and the intermediate shaft is of 100/29, the rotational speed of the intermediate shaft, where the faulty bearings were mounted, was of $1200 \times 29/100 = 348$ rpm (*i.e.* $f_r = 5.8$ Hz). This way, the criterion for low speed machinery was satisfied. In addition, all the experiments were performed under a constant load, where the output torque of the brake was set to 6 Nm. Hence, the acting torque on the intermediate shaft was of $6 \times (20/40) \times (36/90) = 1.2$ Nm. According to the aforementioned test specification (shaft rotational speed and seeded faults), one may expect distinctive peaks around three frequencies in the spectrum of the enhanced signal of the faulty state measurement. These frequencies are listed in Table 3.

Frequency	Value [Hz]	Comment
f_1	20.72	BPFO
f_2	26.94	BDF
f_3	31.50	BPFI

Table 3. Expected peak frequencies in the spectrum of the enhanced signal.

Measurement settings

Prior to signal digitizing, each measured signal was low-pass filtered with an anti-aliasing filter embedded in each channel of the used NI data acquisition system. This way, potential aliasing problems resulting from high frequency noise can be avoided. With this data acquisition system, the cut-off

frequency of the anti-aliasing filter is automatically selected depending on the used sampling frequency. Later on, the filtered signals were sampled at 51.2 kHz with a duration of 60 seconds. Finally, the digital data were stored in the PC and then processed with Matlab programs as will be discussed in the next section. Note that two datasets were recorded from two successive runs (dataset#1 and dataset#2), where the time difference between the two runs was 120 seconds.

5. RESULTS AND DISCUSSION

For benchmarking purpose, the diagnostic results based on the optimized envelope method are first given. All the technical details of this envelope method are not discussed here but they can be found in (Randall & Antoni, 2011) for a complete overview. In short, all the signal processing steps of the envelope method are summarized in Figure 6. Consecutively, the

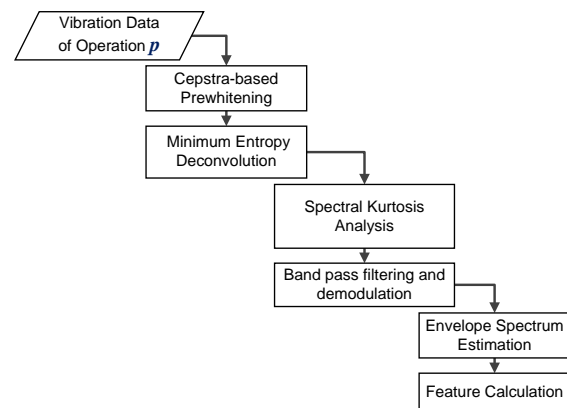


Figure 6. Flowchart of the signal processing steps and feature calculation of the optimized envelope method.

diagnostic results based on the SR with multiscale noise tuning method are discussed and then compared to those of the envelope method. As will be shown later, the SR with multiscale noise tuning based method outperforms the envelope

based method.

5.1. Signal segmentation and analysis procedure

All the recorded vibration signals from the two datasets are segmented with the window length of 6 seconds without overlapping, so there are 20 segments in total for the two datasets. From now on, each segment is referred to the segment index S_i ($i = 1, 2, \dots, 20$). The signal segmentation of dataset#1 is illustrated in Figure 7, where the first window (solid line) is employed both for identification of the optimal parameters and feature calculation, and the other windows (dashed lines) are used only for feature calculation. This way, the repeatability and robustness of extracted features can be evaluated and verified. It should be noted here that the raw vibration signal is re-sampled differently for the two methods. The sampling frequency of 25.6 kHz is used for the envelope method, while the sampling frequency of 3.2 kHz is used for the SR method.

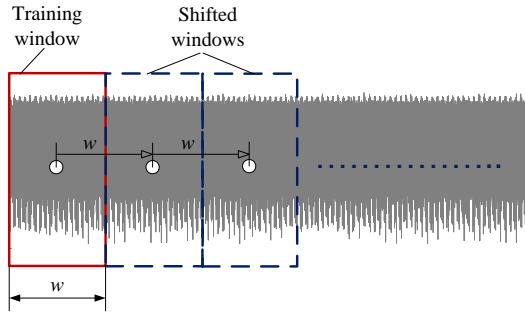


Figure 7. Illustration of the signal segmentation.

5.2. Envelope method

Figure 9 shows the signals and spectra of the faulty state vibration data resulting from different signal processing steps according to the flowchart the envelope method shown in Figure 6. It should be mentioned that all the plots are obtained from the analysis on the first window of dataset#1. As seen in Figure 9(a), expected peaks around the theoretical fault frequencies (f_1, f_2, f_3 , see Table 3) in the spectrum of the raw vibration signal are not obvious. After applying the cepstra-based prewhitening on the raw signal, low-frequency components of the signal are enhanced as can be seen in the spectrum of Figure 9(b). The low-frequency components are more pronounced after applying MED filtering as shown by the spectrum in Figure 9(c). Subsequently, the MED-filtered signal is bandpass-filtered around the frequency band of 4233 - 4833 Hz and demodulated at the central frequency of 4533 Hz. Here, the frequency band is determined based on the spectral kurtosis analysis (Antoni & Randall, 2006) that eventually results in the kurtogram given in Figure 8.

It is seen from the kurtogram that the spectral kurtosis of the

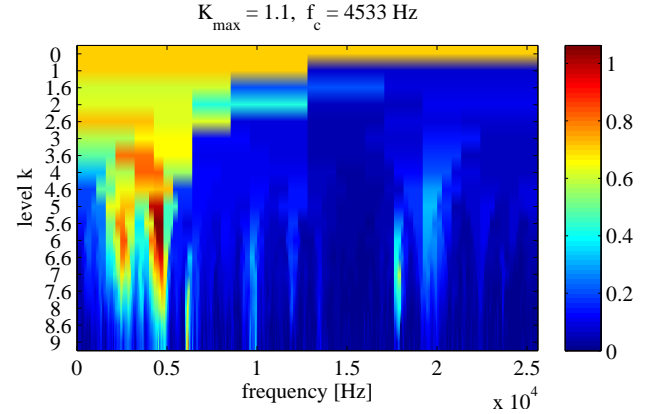


Figure 8. Kurtogram of the signal processed after MED operation

latter signal is maximum when the band of 4233 - 4833 Hz is selected. Finally, the squared envelope signal is obtained by simply taking the square of the resulting signal as shown in Figure 9(d). The lower panel of the latter figure shows the normalized envelope spectrum, where the normalization is performed by dividing the spectrum magnitudes by the DC magnitude. It is seen from the figure that peaks around f_1 (BPFO) and the corresponding harmonics are evident, but it is not the case for the other expected frequencies f_2, f_3 and their harmonics.

Figure 10 shows the results obtained from the envelope analysis on the vibration data of the healthy state. Note that a damage-induced impact is theoretically not generated in any healthy state, so it can be argued that none of resonance frequencies of the gearbox system is excited. In order to compare the results of the healthy state and faulty state in a fair manner, the frequency band obtained from the spectral kurtosis analysis on the faulty data is used for the envelope analysis on the healthy data. It is seen in the figure that peaks around the theoretical fault frequencies (f_1, f_2, f_3 , see Table 3) and the harmonics in the envelope spectrum are not evident. Instead, the magnitudes around those frequencies and the corresponding harmonics are in the same level as the floor noise.

Furthermore, to quantitatively evaluate the performance of the envelope method for diagnosing low-rotational speed bearing, features are extracted from the envelope spectra of the faulty and healthy state according to the feature definition given in Eq. (20). The comparisons between the extracted features for different faults obtained from the faulty and healthy state are shown in Figure 14. It is evident from the figure that the feature values corresponding to all the seeded faults extracted from both the healthy and fault state are not well separated. These results suggest that the features extracted using the envelope method are not robust for diagnosis of a bearing subjected to low speed and load.

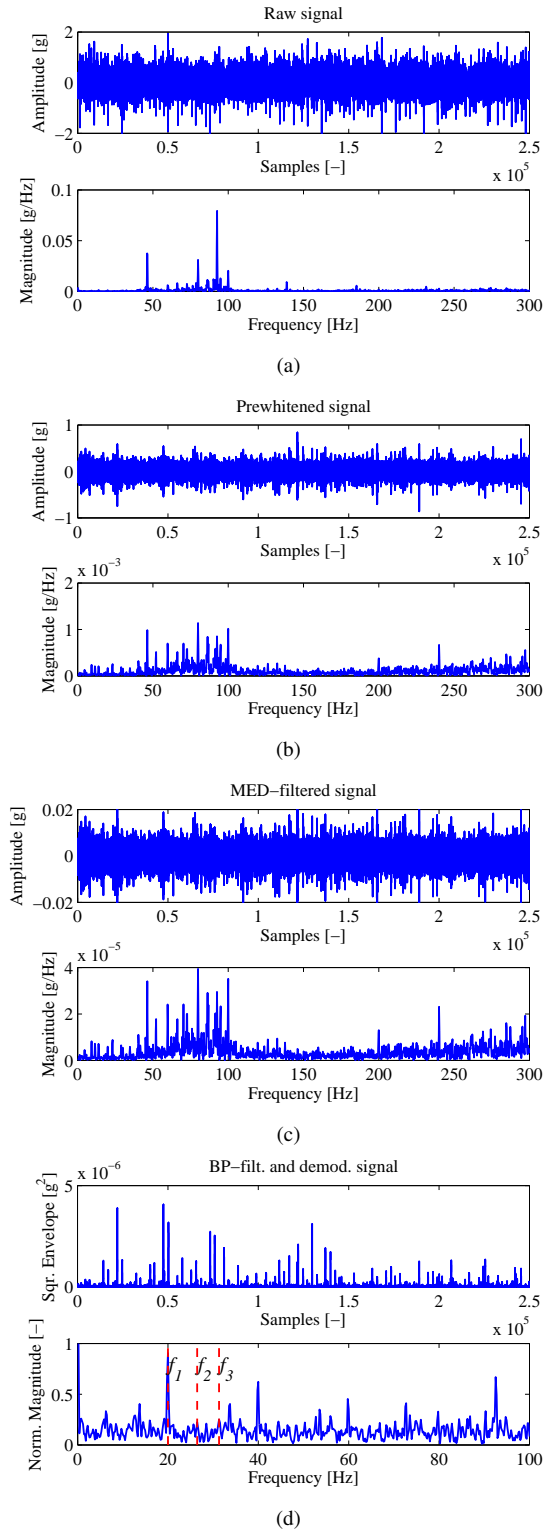


Figure 9. Resulting signals and the corresponding spectra obtained with the envelope method applied on the vibration data of the **faulty state**. (a) raw signal and the spectrum, (b) prewhitened signal and the spectrum, (c) MED filtered signal and the spectrum, and (d) the squared envelope and the *normalized* spectrum.

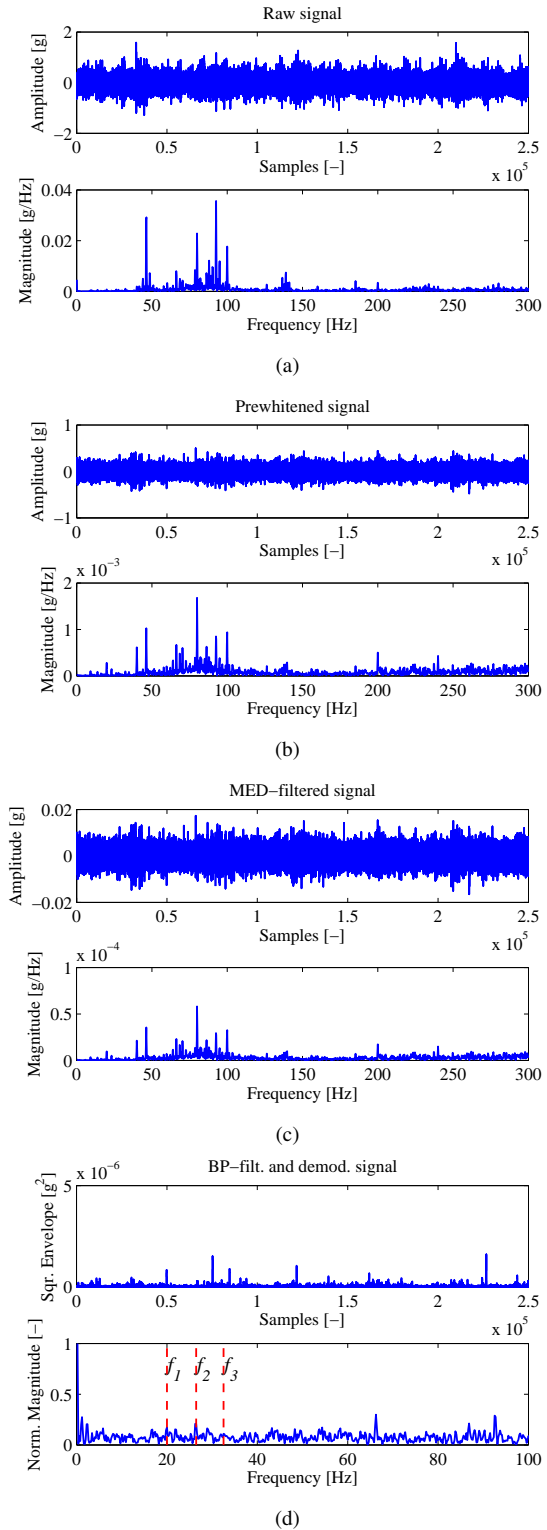


Figure 10. Resulting signals and the corresponding spectra obtained with the envelope method applied on the vibration data of the **healthy state**. (a) raw signal and the spectrum, (b) prewhitened signal and the spectrum, (c) MED filtered signal and the spectrum, and (d) the squared envelope and the *normalized* spectrum.

5.3. SR based method

Prior to feature calculation based on the SR with multi scale noise tuning method, the optimal tuning parameter need to be identified. As the seeded faults in the experiments were outer race fault f_1 , rolling element fault f_2 and inner race fault f_3 , only $\alpha_{opt,2}$ corresponding to the second target frequency $f_{0,2}$ (*i.e.* averaged frequency of f_1 , f_2 and f_3) is required. To this end, a search range of possible values $\alpha_k \in [0, 38]$ is predefined. By following the procedure described in Section 3.1, the learning curve shown in Figure 11 can be generated. It is worthwhile to emphasize here that the learning curve is generated based on the training (segmented) signal described in Figure 7. As can be seen from Figure 11, the objective value η reaches its maximum when α_k is about 16. Therefore, the optimal parameter to enhance the bearing signals corresponding the three fault frequencies (f_1 , f_2 and f_3) was set to $\alpha_{opt,2} = 16$.

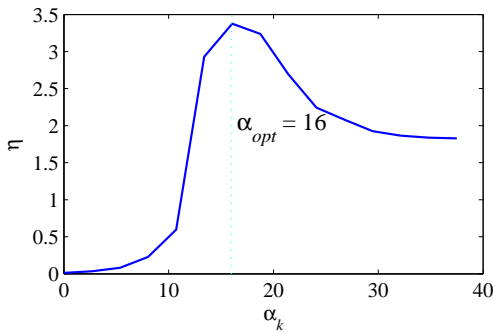


Figure 11. The learning curve for determining the optimal tuning parameter $\alpha_{opt,2}$.

Figure 12 shows the resulting signals and the spectra obtained by the SR with multi scale noise tuning method of the faulty case. For convenience, the pre-processed signals and spectra (Cepstra-based prewhitened and full-band enveloped signals) are not shown here. One can notice from Figure 12(a) that the multi scale noise tuning with wavelet decomposition already enhances the signal components corresponding to f_1 (BPFO) and f_3 (BPFI). The two components are further enhanced after applying the SR filtering on the multi scale noise tuned signal as can be seen in Figure 12(b). However, the signal component corresponding to f_2 (BDF) is not pronounced in the spectra.

Figure 13 shows the resulting signals and the spectra obtained by the SR with multi scale noise tuning method of the healthy case. As expected, the spectrum magnitudes around the fault frequencies f_1 , f_2 , f_3 are the same as the floor noise level.

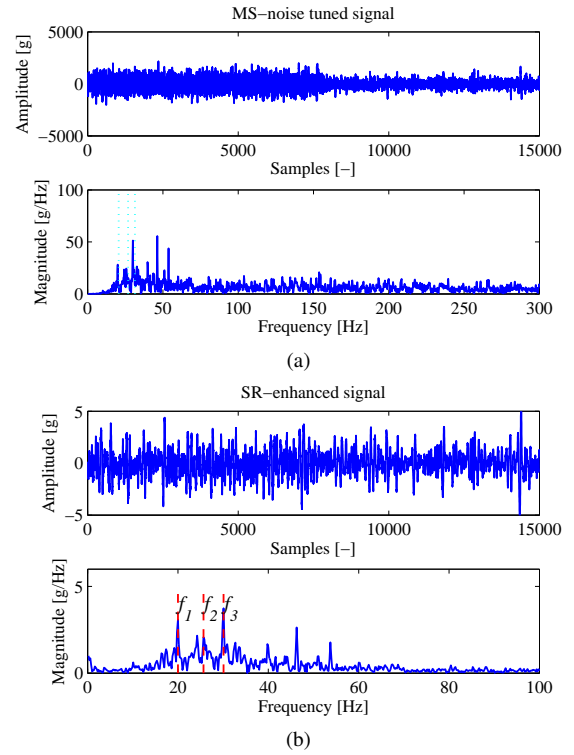


Figure 12. Signal processing results of the **faulty** case. (a) multi scale noise tuned signal of the full band envelope signal and (b) the output signal of the SR filter.

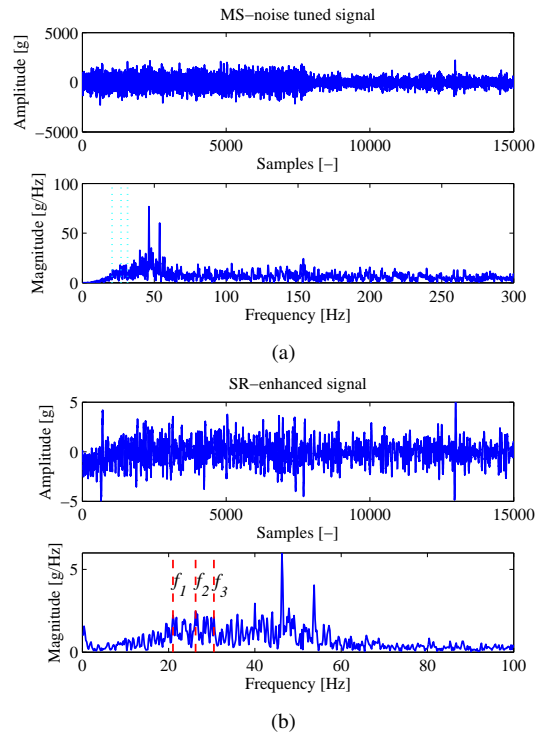


Figure 13. Signal processing results of the **healthy** case. (a) multi scale noise tuned signal of the full band envelope signal and (b) the output signal of the SR filter.

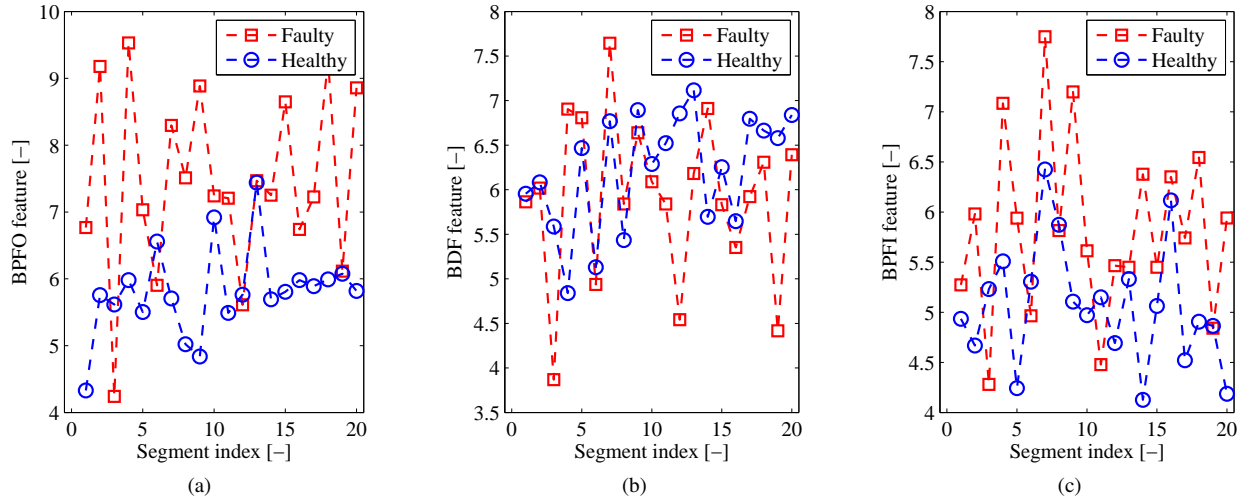


Figure 14. Diagnostic results obtained with the **optimized envelope method** for: (a) the outer race fault, (b) the rolling element fault, and (c) the inner race fault.

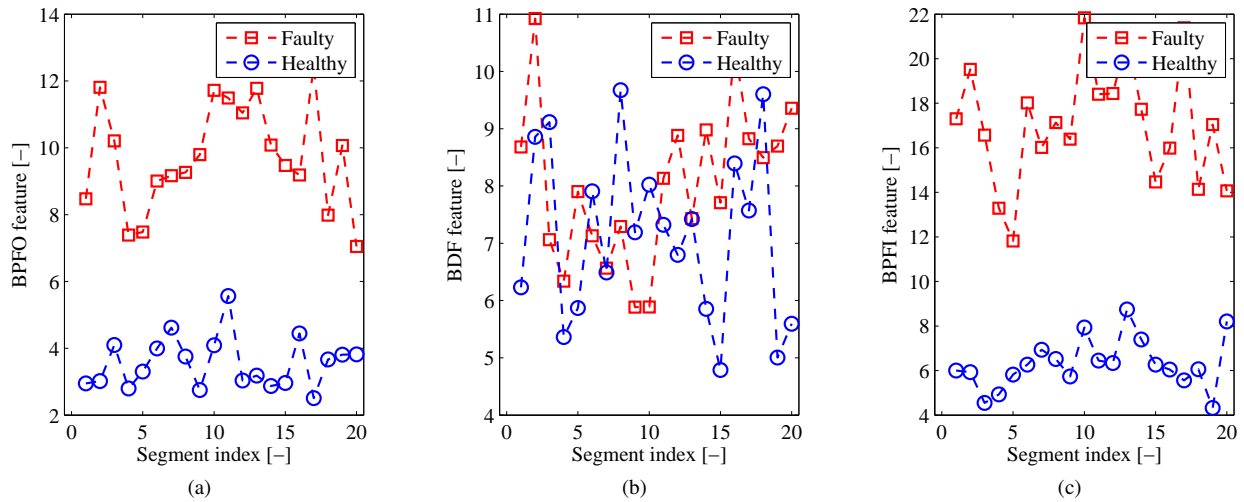


Figure 15. Diagnostic results obtained with the **SR method with multi scale noise tuning** for: (a) the outer race fault, (b) the rolling element fault, and (c) the inner race fault.

For a quantitative comparison, the features corresponding to BPFO, BDF and BPFI defined in Eq. (20) have been extracted from both the faulty and healthy datasets. The feature values obtained from both states, for different faults, are shown in Figure 15. It is seen in the figure that good separations between the healthy and faulty case are revealed for the *outer race* and *inner race* faults. This evidence suggests that the two faults seeded on the outer and inner race are well detected with the proposed method. Nevertheless, the seeded fault on one of the rolling elements cannot be successfully detected since the feature values of both the healthy and faulty state are not well separated.

6. CONCLUSION

As reported in literature, the High Frequency Demodulation (HFD)/envelope technique has been successfully applied for fault diagnosis of rolling-element bearings used in different applications. However, it has been shown in this study that the envelope technique seems to fail for fault diagnosis of bearings subjected to low rotational speed and low load. For a bearing subjected to low rotational speed and low load, it can be argued that the damaged-induced vibration signal (*i.e.* impulsive signal) is severely masked by a large background noise originating from other machine components and the operating environment. It appears that a low signal-to-noise ratio inherent in the vibration signals of such an operating condition is one of the possible reasons of the limitation of the envelope technique. In order to remedy the aforementioned substantial gap, we have proposed an alternative method.

This paper has made an improvement on the recently published bearing diagnostic method based on the Stochastic Resonance (SR) with multi noise tuning for multi-fault diagnosis of rolling element bearings subjected to low rotational speed and low load. Features that can be used to quantitatively assess the presence of different bearing faults, *e.g.* inner race, rolling element and outer race fault, have been mathematically formulated. This way, bearing fault diagnosis can be realized in an automatic way given a certain (predefined) threshold.

The proposed method has been experimentally verified on the vibration data collected on a gearbox dynamics simulator with healthy and faulty (*i.e.* combined seeded faults) bearings. Here, the drive shaft of the bearing seeded with combined faults (inner race, rolling element and outer race faults) was run at 348 rpm and under torsional load of 1.2 Nm. Moreover, the results obtained with the proposed method have also been compared with the ones obtained with the well known high frequency demodulation (envelope) technique. It has been shown in this study that the proposed method can successfully detect the two seeded faults (inner race and outer race faults). In contrast, the envelope method is not able to

detect the seeded faults.

ACKNOWLEDGMENT

The authors are thankful for the support from the project partners. A credit goes to Claude Dieleman and Gianni Roose of Dana Spicer Off Highway Belgium for the technical support.

NOMENCLATURE

Abbreviations

DWT	discrete wavelet transform
BPFI	ball pass frequency inner
BPFO	ball pass frequency outer
BDF	ball damage frequency
BSF	ball spin frequency
FTF	fundamental train frequency
GDS	gearbox dynamics simulator
HFD	high frequency demodulation
MED	minimum entropy deconvolution
SNR	signal-to-noise ratio
SR	stochastic resonance

Latin symbols

a, b	barrier parameters of a bistable system
a_j, d_j	approximation and detail coefficients
D	noise intensity
D_r	noise intensity at the resonance
f_0	driving frequency in Hz
f_i	theoretical bearing fault frequency in Hz
f_r	frequency of drive shaft
J	level of wavelet decomposition
j, k	indices
t	time
U	potential
X	spectrum magnitude of the output signal x

Greek symbols

α	tuning parameter
α_{opt}	optimal tuning parameter
ΔU	barrier potential
ϵ	signal amplitude
η	objective value
λ	feature value
ω_0	driving frequency in rad/s
ϕ	phase response
ρ_K	Kramer's rate
ξ	white noise

REFERENCES

Antoni, J., & Randall, R. (2006). The spectral kurtosis: application to the vibratory surveillance and diagnostics

- of rotating machines. *Mechanical Systems and Signal Processing*, 20(2), 308 - 331.
- Benzi, R., Sutera, A., & Vulpiani, A. (1981). The mechanism of stochastic resonance. *Journal of Physics A: Mathematical and General*, 14(11), L453.
- Flandrin, P. (1992). Wavelet analysis and synthesis of fractional brownian motion. *Information Theory, IEEE Transactions on*, 38(2), 910-917.
- Gammaitoni, L., Hänggi, P., Jung, P., & Marchesoni, F. (1998, Jan). Stochastic resonance. *Rev. Mod. Phys.*, 70, 223-287.
- He, H.-L., Wang, T.-Y., Leng, Y.-G., Zhang, Y., & Li, Q. (2007). Study on non-linear filter characteristic and engineering application of cascaded bistable stochastic resonance system. *Mechanical Systems and Signal Processing*, 21(7), 2740 - 2749.
- He, Q., Wang, J., Liu, Y., Dai, D., & Kong, F. (2012). Multi-scale noise tuning of stochastic resonance for enhanced fault diagnosis in rotating machines. *Mechanical Systems and Signal Processing*, 28(0), 443 - 457.
- Jung, P., & Hänggi, P. (1991, Dec). Amplification of small signals via stochastic resonance. *Phys. Rev. A*, 44, 8032-8042.
- Kim, Y.-H., Tan, A., Mathew, J., & Yang, B.-S. (2006). Condition Monitoring of Low Speed Bearings: A Comparative Study of the Ultrasound Technique Versus Vibration Measurements. In J. Mathew, J. Kennedy, L. Ma, A. Tan, & D. Anderson (Eds.), *Engineering asset management* (p. 182-191). Springer London.
- Kosse, V., & Tan, A. (2006). Development of Testing Facilities for Verification of Machine Condition Monitoring Methods for Low Speed Machinery. In J. Mathew, J. Kennedy, L. Ma, A. Tan, & D. Anderson (Eds.), *Engineering asset management* (p. 192-197). Springer London.
- Lei, Y., Han, D., Lin, J., & He, Z. (2013). Planetary gearbox fault diagnosis using an adaptive stochastic resonance method. *Mechanical Systems and Signal Processing*, 38(1), 113 - 124.
- Leng, Y., Leng, Y., Wang, T., & Guo, Y. (2006). Numerical analysis and engineering application of large parameter stochastic resonance. *Journal of Sound and Vibration*, 292(35), 788 - 801.
- Li, J., Chen, X., Du, Z., Fang, Z., & He, Z. (2013). A new noise-controlled second-order enhanced stochastic resonance method with its application in wind turbine drivetrain fault diagnosis. *Renewable Energy*, 60(0), 7 - 19.
- Li, J., Chen, X., & He, Z. (2013). Adaptive stochastic resonance method for impact signal detection based on sliding window. *Mechanical Systems and Signal Processing*, 36(2), 240 - 255.
- Lin, T., Kim, E., & Tan, A. (2013). A practical signal processing approach for condition monitoring of low speed machinery using Peak-Hold-Down-Sample algorithm. *Mechanical Systems and Signal Processing*, 36(2), 256 - 270.
- McNamara, B., & Wiesenfeld, K. (1989, May). Theory of stochastic resonance. *Phys. Rev. A*, 39, 4854-4869.
- Randall, R., & Antoni, J. (2011). Rolling element bearing diagnostics - A tutorial. *Mechanical Systems and Signal Processing*, 25(2), 485 - 520.
- Randall, R., & Sawalhi, N. (2011). Use of the cepstrum to remove selected discrete frequency components from a time signal. In T. Proulx (Ed.), *Rotating machinery, structural health monitoring, shock and vibration, volume 5* (p. 451-461). Springer New York.
- Tan, J., Chen, X., Wang, J., Chen, H., Cao, H., Zi, Y., & He, Z. (2009). Study of frequency-shifted and re-scaling stochastic resonance and its application to fault diagnosis. *Mechanical Systems and Signal Processing*, 23(3), 811 - 822.
- Wu, X., Jiang, Z.-P., & Repperger, D. (2006). Enhancement of stochastic resonance by tuning system parameters and adding noise simultaneously. In *American control conference, 2006* (p. 6 pp.-).
- Xu, B., Li, J., & Zheng, J. (2003). How to tune the system parameters to realize stochastic resonance. *Journal of Physics A: Mathematical and General*, 36(48), 11969.

BIOGRAPHIES

Agusmian Partogi Ompusunggu received a PhD degree in mechanical engineering (i.e. reliability of mechatronic systems) from the KU Leuven, Belgium. He earned his B.Eng. degree in 2004 and his M.Eng. degree in 2006 both from Institut Teknologi Bandung (ITB), Indonesia. Currently, he is a senior research engineer at Flanders' Mechatronics Technology Centre (FMTC), Belgium. His interests include signal processing, state estimation, data mining/pattern recognition, machine condition monitoring/diagnostics/prognostics, dynamics/vibration analysis and measurement, multi-physics modeling and simulation, and tribology.

Steven Devos received a PhD degree in mechanical engineering from the KU Leuven, Belgium. He earned his M.Eng degree in 1999 at Ghent University, Belgium. Currently, he is a project manager at FMTC. His interests include the measurement and simulation of machine dynamics.

Frederik Petre is currently the Program Manager of the R&D program on "Smart Sensors for Condition, Process and Product Quality Monitoring" at the Flanders' Mechatronics Technology Centre (FMTC) in Leuven, Belgium. Over there he leads a team of 11 Project Engineers and 4 Project Managers. Before joining FMTC, Frederik was a Senior Scientist within the Wireless Research Group at imec Belgium, investigating baseband signal processing algorithms and digital architectures for future generation wireless communication sys-

tems. He holds a M.Sc. degree (1997) and a Ph.D. (2003) in Electrical Engineering, both from the KU Leuven, Belgium. In 2009 he obtained an Executive MBA in General International Management from the Vlerick Business School.