COMPARISON BETWEEN TWO VERY EFFICIENT SIGNAL PROCESSING APPROACHES FOR VIBRATION-BASED CONDITION MONITORING OF ROLLING ELEMENT BEARINGS: THE MED-SK AND CS-BASED APPROACHES

Dany Abboud[(1)] and Mohammed Elbadaoui[(1)(2)]
[(1)]SafranTech
Rue des Jeunes Bois
Châteaufort 78772 Magny –les- Hameaux, France
Telephone: (+33) 6 95222573
dany.abboud@safrangroup.com

[(2)]Univ Lyon
UJM-St-Etienne, LASPI
EA3059, F-42023, Saint-Etienne, France

Abstract: This paper compares the most efficient signal processing approaches used for vibration-based condition monitoring of rolling element bearings. The first is based on pre-processing the vibration signal through the maximum entropy deconvolution method (MED) followed by the spectral kurtosis (SK), before analyzing the spectrum of the signal envelope. The MED aims at maximizing the signal impulsivity by deconvolving the system transfer function through an optimization approach that maximizes the kurtosis of the output. Then, the spectral kurtosis (SK) is applied to conceive the optimal filter to be applied before computing the envelope spectrum. The second approach is based on a cyclostationary modeling of the bearing signal. It applies the spectral coherence to the signal with a special attention on setting the estimation parameters. The spectral coherence is a bi-variable map of the cyclic frequency, α, and the spectral frequency, f. The former variable describes the cyclic content of the modulations, while the former describes the properties of the carrier. The improved envelope spectrum is then computed by simply projecting its squared-magnitude with respect to the f-variable. These methods are evaluated according to their potentiality to detect the fault in its earliest stage. The comparison is be made on real bearing vibration signals in run-to-failure tests.

Key words: Condition monitoring; spectral coherence; cyclostationarity; improved envelope spectrum; incipient fault detection; minimum entropy deconvolution; spectral kurtosis

1. Introduction: Rolling element bearings (REBs) are perhaps the most widely used elements in rotors. Because of their fragility, bearings are likely to be exposed to sudden failures causing system outage. Therefore, special care growths in last decades to develop appropriate diagnosis technique for incipient fault detection. Due to their non-invasive nature and their high reactivity to incipient faults, the development of vibration based techniques spiked the interest of the scientific committee. In this context, envelope analysis has been recognized for long as a powerful technique for REB diagnosis. Typically, it consists of a bandpass filtering step in a frequency band wherein the impulsive excitation is amplified followed by a demodulation that extracts the signal envelope. The spectrum of the envelope— known as the envelope spectrum— is expected to contain the desired diagnostic information, including the repetition rate of the fault and potential modulations [1]. It has been shown in Ref. [2] that it is preferable to use the squared envelope instead of the envelope inasmuch as the latter introduces extraneous components that appears as misleading peaks in the envelope spectrum. Since that time, the squared envelope spectrum (SES) has probably become the benchmark technique for bearing diagnostics and efforts were directed on how to choose the most suitable band for demodulation. Historically, this band was estimated through a hammer tap testing by finding the bearing housing resonances. Today, this problem is almost solved by the use of the spectral kurtosis (SK) [3]. Being based on a statistical approach, the SK provides a way to determine which frequency bands are informative according to their impulsivity. It can be actually seen as the kurtosis of the signal bandpass filtered with a filterbank covering its frequency content. However, the SK has proven ineffective in high-speed applications. The reason is that the fault period is too small that the impulse responses triggered by the chocks overlap. This significantly reduces the signal impulsivity and jeopardize the effectiveness of the SK. Reference [1] has proposed to preprocess the signal before applying the SK by means of the "minimum entropy deconvolution" (MED). This technique is based on finding the optimal inverse filter which compensates the transfer function of the transmission path, with the aim of separating the excitation impulses. The combination of the MED and the SK has proven very effective in incipient fault detection in REBs and has become a leading procedure in this framework. Another interesting approach for REB diagnosis is based on the cyclostationary (CS) theory. In details, a new model of REB vibrations has been introduced in Refs. [4], providing an insightful understanding of the REB fault signature within the CS framework. Accordingly, it has been shown that the mechanical signature generated by a faulty REB is random in nature and has symptomatic properties that can be detected by means of second-order CS tools such as the spectral correlation, the spectral coherence, the cyclic modulation spectrum and others [5]. These techniques have also proven efficient and their use is in continuous growth. This paper comes in this context aiming at comparing these two efficient approaches on real REB vibration signals in run-to-failure tests. Section 2 recalls the structure of a vibration signal emitted by an incipient fault in a REB. Section 3 and 4 briefly review the MED-SK-based and CS-based approaches, respectively. Section 5 compares these methods on real vibration signals and discuss the results and the paper is sealed with a conclusion in section 6.

2. Vibration-based diagnosis of rolling element bearings: As a defect strikes a mating surface, an abrupt change in the contact stress occurs at the interface (called *shock*) which corresponds to a transfer of kinetic energy into elastic potential energy. These shocks occur on a pseudo-periodic basis due to the rotating kinematics of the bearing and exciting the system resonance. In facts, the generated vibration can be seen as a train of short-time impulses, periodic on average (with a fault period $T_d = 1/f_d$), locked to the period of the fault characteristic frequency. The inter-arrival times of the impacts are subjected to slight random fluctuations, ¢T, due to existing slips. Also, this pseudo-periodic train is likely to be modulated by slower modulation (with a period $T_m = 1/f_m$) due to the change of the load in the impact load. These modulations can either be caused by a possible shaft unbalance or misalignment, or a periodic change in the impulse response as the distance and orientation of the impacts moves towards and backwards the sensor. Figure 1 provides an academic illustration of a typical vibration signal generated by a faulty REB.

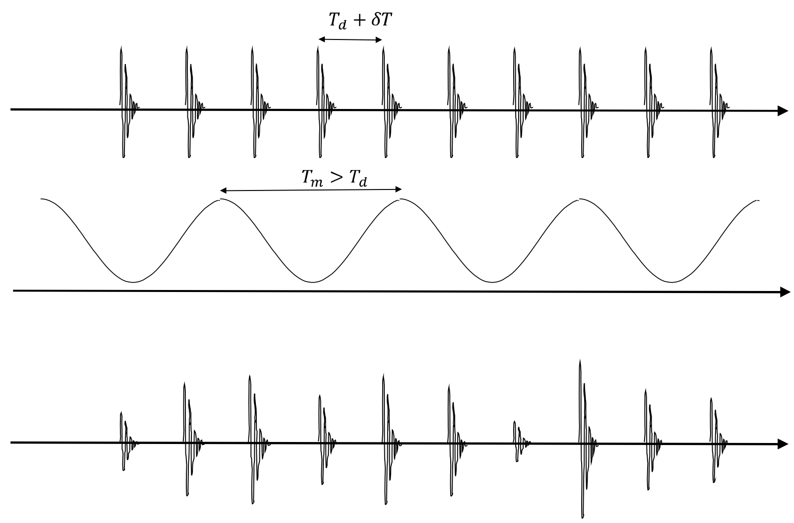

Figure 1: An illustration of the vibration signal emitted by an incipient fault in a REB.

Despite the presence of this periodicity, the power spectrum is continuous and, thus, unable to reveal the mechanical signature of the defect. This is a direct result of the randomness induced by the jitter between the impacts which, despite being in the order of few percent of the fault period, its presence change the whole statistical structure of the signal. The squared magnitude of the envelope, however, is indeed quasi-periodic. Therefore, its spectrum, called the squared envelope spectrum (SES) is discrete and do reveal the fault signature. In details, the fault signature appears as discrete harmonics located at the fault frequency f_d and its multiples, and modulated by the load change modulation frequency f_m and its multiples. This is clearly illustrated in Figure 2. These facts explain why the envelope analysis has been historically used for vibration-based fault detection in REBs.

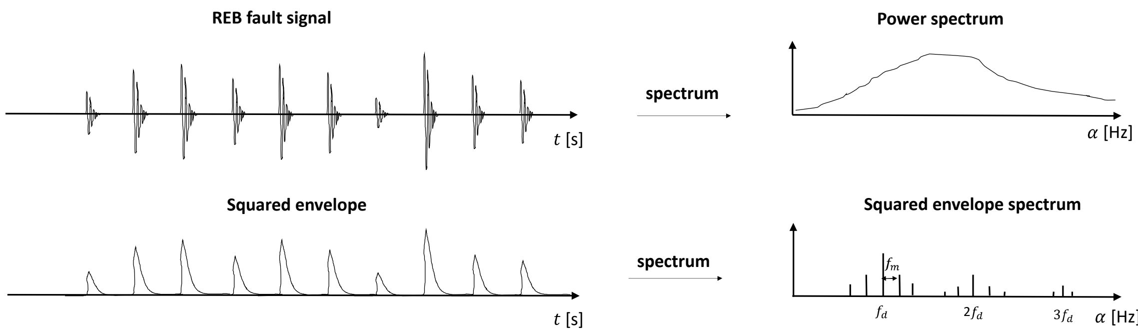

Figure 2: Spectral properties of REB signals.

In practice, the fault component is often masked by multiple contributions generated by multiple sources. These sources can be classified in two major groups. First, the deterministic contribution, being characterized by a discrete spectrum, can contribute in the envelope spectrum and mask the REB fault component. These contributions are typically generated by gears and are generally eliminated by numerous signal processing techniques such as the synchronous average, blind filters, Fourier calculation and others. The elimination of the deterministic component is out of the scope of the current paper. Even after removal of discrete frequency ‘‘noise’’, the bearing signal will often be masked in many frequency bands by other random noise. This effect leads to a reduction in the signal impulsiveness as the fault pulses are modified by the passage through a transmission path with a long impulse response. This is typically the case of high-speed application in which the fault period are too small that the impulse responses triggered by the chocks overlap. This jeopardizes the effectiveness of the classical envelope analysis and calls for different approaches to promote the fault signature when embedded in strong noise. The enhancement of the REB fault component and signature can be made through two very efficient approaches discussed in the following sections.

3. MED-SK-based approach: The first efficient approach is based on pre-processing the vibration signal by combining two efficient methods, namely the “minimum entropy deconvolution” (MED) followed by “the spectral kurtosis” (SK). The MED was first proposed in Ref. [6] to sharpen the reflections from different subterranean layers in seismic analysis. This technique is based on finding the optimal inverse filter which compensates the transfer function of the transmission path, assuming an impulsive excitation. This process is illustrated in Figure 3. The impulsive REB fault signal, $e[n]$, is likely to be corrupted by background noise, $b[n]$, before being convolved with the transfer function of the transmission path from the excitation source to the transducer. The obtained signal is denoted by $x[n]$. The MED aims at estimating the filter, $b[n]$, that inverts the system impulse response function, $g[n]$, such that

$$(h \otimes g)[n] = ¢[n - l_m] \tag{1}$$

where $\otimes$ denotes the numerical convolution, ¢$\left[n-l_m\right]$ is the Kronecker delta function. Note that l_m is allowed to make the inverse filter causal. The filter $g[n]$ is modeled as a FIR filter with L coefficients such that:

$$y[n]=\sum_{l=1}^{L} g[l]x[n-l]. \tag{2}$$

In vibration analysis, the only method used for filter coefficient estimation is the objective function method provided in Ref. [7], where the objective function to be maximized is the kurtosis of the output signal $y[n]$:

$$O_k(g)=\frac{\sum_{n=0}^{N-1} y[n]^4}{\left(\sum_{n=0}^{N-1} y[n]^2\right)^2}. \tag{3}$$

The optimal filter is that whose coefficients maximize $O_k(g)$ or, said differently, set to zero its first derivative, i.e.

$$\frac{\partial O_k(g)}{\partial g}=0. \tag{4}$$

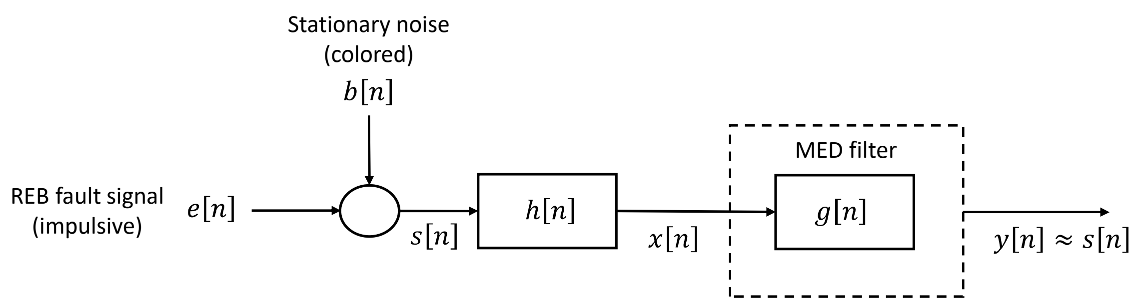

Figure 3: Principle of the MED technique.

The MED has for object to enhance the signal impulsivity, yet there is still a need to enhance the REB signal with respect to residual background noise before analyzing the SES. In other words, there is a need to define a filter which maximizes the signal-to-noise ratio and, consequently, enhances the fault signature in the SES. As previously indicated, the SK is a very good candidate to accomplish this task. Let's denote by $Y(n;f)$ the short-time Fourier transform of $y[n]$, the SK is defined as:

$$K_y(f) = \frac{\left\langle \left| Y(n;\, f) \right|^4 \right\rangle}{\left\langle \left| Y(n;\, f) \right|^2 \right\rangle^2} - 2\,. \tag{5}$$

with $\langle \bullet \rangle$ stands for the average operator over the time index n. Note that the subtraction of 2 is used to enforce a zero value for a complex Gaussian process. Since the SK returns high values in the frequency bands related to an impulsive REB fault, it can be used to conceive a filter which maximizes the signal impulsivity. Since the aim is to demodulate the signal to apply the SES, the matched filter is perhaps the most appropriate as maximizes the SNR of the filtered signal without regard to its shape. Let's denote by $g'[n]$ the optimized matched filter based on the SK, the filtered signal can be simply expressed as:

$$z[n] = \sum_{l=1}^{L} g'[l]\, y[n-l] \tag{6}$$

Eventually, the SES is applied on $z[n]$ to reveal the hidden cyclic content:

$$SES_z(\sim) = DTFT_{n\rightarrow\sim}\left\{ z[n]^2 \right\}. \tag{7}$$

For a faulty REB, $SES_Z(\sim)$ is expected to exhibit harmonics at the fault frequency f_d with sidebands spaced by the load change modulation frequency f_m:

$$SES_z(\sim) = \sum_{i,j} SES_z^{i,j}\text{¢}\left(\sim - if_d + jf_m\right). \tag{8}$$

The SK-based filtering operation followed by the SES is illustrated in Figure 4.

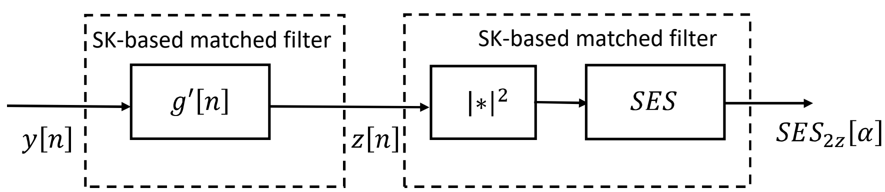

Figure 4: SK-based filtering and envelope analysis (continuation of Figure 3).

4. CS-based approach: As asserted in precursory works [3,4] vibrations produced by faulty bearings are random CS (i.e. second-order CS). This is principally due to the presence of slight random fluctuations in inter-arrival times of the impacts caused by the rolling element slippage. The properties of such a signal are fully preserved in autocorrelation function, defined for a stochastic signal $x[n]$ as:

$$R_{2x}[m,n] = \mathrm{E}\left\{x[n+m/2]x[n-m/2]^*\right\}. \tag{9}$$

where m and n respectively denotes the time and time-lag indexes. The same information can be more prominently represented through the spectral correlation, defined as being the double Fourier transforms of the autocorrelation function;

$$S_{2x}(f,\sim) = \lim_{N\to\infty} \frac{\Delta_t}{N} \sum_{n=0}^{N-1} \sum_{m=0}^{N-n} R_{2x}[m,n] e^{j2\circledR\Delta_t(mf+n\sim)} \text{ .} \tag{10}$$

where Δ_t stands for the sampling period, N is the signal length, f is the spectral frequency variable in Hz, $\sim$ is the cyclic frequency variable in Hz. Alternatively, the spectral correlation can be expressed as:

$$S_{2x}(f,\sim) = \lim_{N\to\infty} \frac{1}{N\Delta_t} \mathrm{E}\left\{X_N[f-\sim/2]X_N[f+\sim/2]^*\right\}. \tag{11}$$

where $X_N[f] = \Delta_t \sum_{n=0}^{N-1} x[n] e^{j2\circledR fn\Delta_t}$ is the signal discrete time Fourier transform over N samples. This expression provides another interpretation of the SC, which actually justifies its name: it is a measure of the correlation between the frequency components of the signal at $f-\sim/2$ and $f+\sim/2$. A related operator is the spectral coherence (SCoh) which has the property of measuring the degree of correlation between two spectral components independently of the signal power spectrum. This quantity is defined as being the coherence (instead of the correlation) between the frequency components located at f and spaced by $\sim$:

$$¥_{2x}(f,\sim) = \lim_{N\to\infty} \frac{\mathrm{E}\left\{X_N[f-\sim/2]X_N[f+\sim/2]^*\right\}}{\left(\mathrm{E}\left\{X_N[f-\sim/2]^2\right\}\mathrm{E}\left\{X_N[f+\sim/2]^2\right\}\right)^{1/2}} \text{ .} \tag{12}$$

Just like the regular correlation coefficient, the squared magnitude of the SCoh, $|¥_{2x}(f,\sim)|^2$, is normalized between 0 and 1: the closer it is to the unity, the stronger the CS component at cyclic frequency $\sim$. Interestingly, when a CS signal of cyclic frequency $\sim$ is embedded in stationary noise, its $|¥_{2x}(f,\sim)|^2$ can be seen as an indication of the signal-to-noise ratio. Often, vector indicators are preferred in industry as their readability is generally simple and does not require advanced skills in signal processing. Accordingly, one may sacrifice the optimality of the SCoh by integrating over the spectral frequency variable preferably in a high SNR frequency band (or simply over the whole frequencies). Precisely, one may devise a suboptimal quantity, henceforth called the “improved envelope spectrum” (IES)—as:

$$IES_{2X}\left(\sim\right)=\frac{1}{F_2-F_1}\int_{F_1}^{F_2}\left|\yen_{2x}\left(f,\sim\right)\right|^2 df\,, \tag{13}$$

where F_1 and F_2 are respectively the lower and upper bounds of the integration frequency domain. For a faulty REB, $IES_{2X}\left(\sim\right)$ is expected to exhibit harmonics at the fault frequency f_d with sidebands spaced by the load change modulation frequency f_m:

$$IES_{2x}\left(\sim\right)=\sum_{i,j} IES_{2x}^{i,j}\text{¢}\left(\sim - if_d + jf_m\right). \tag{14}$$

This approach is illustrated in Figure 5.

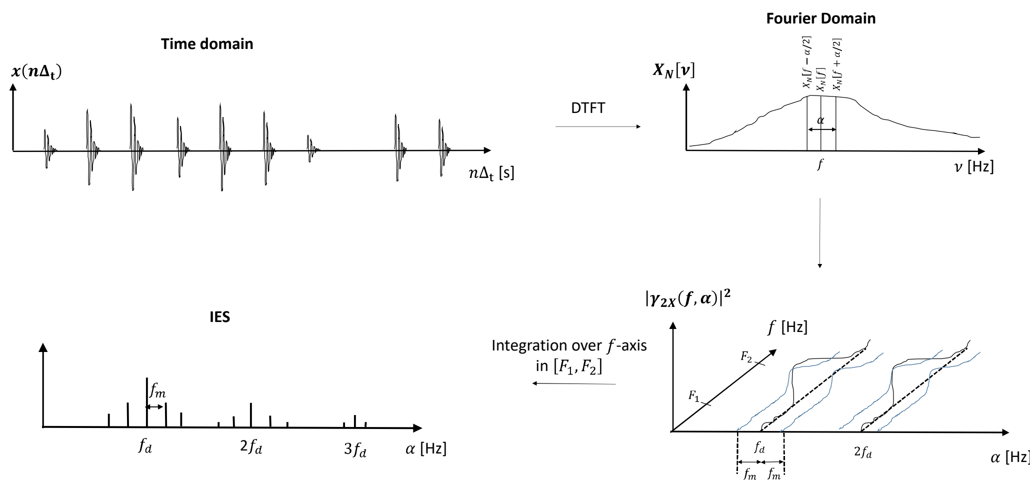

Figure 5: Illustration of the cyclostationary-based approach for REB fault detection.

5. Application: The approaches discussed above are tested on real vibration signal emitted from a benchmark including faulty REB. The data are generated by the NSF I/UCR Center for Intelligent Maintenance Systems (IMS –www.imscenter.net) with support from Rexnord Corp. in Milwaukee, W [8]. The test rig consists of four bearings installed on a shaft. The rotation speed was kept constant at 2000 RPM by an AC motor coupled to the shaft via rub belts. A radial load of 6000 lbs is applied onto the shaft and bearing by a spring mechanism. All bearings are force lubricated. Rexnord ZA-2115 double row bearings were installed on the shaft as shown in Figure 6. PCB 353B33 High Sensitivity Quartz ICP accelerometers were installed on the bearing housing. All failures occurred after exceeding designed life time of the bearing which is more than 100 million revolutions. More details can be found in Ref. [9].

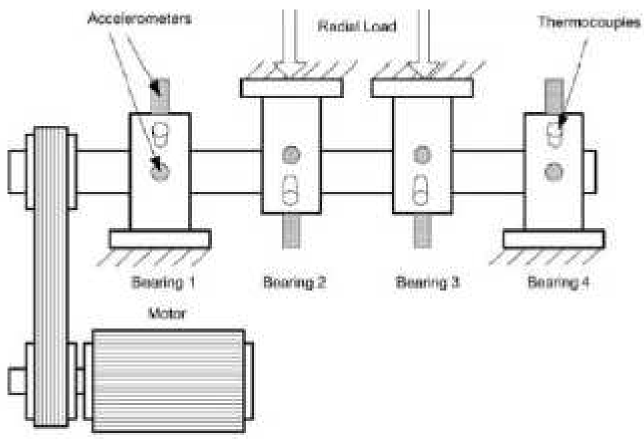

Figure 6: Bearing Test-rig.

Each of the REB is a double row bearing consisting 16 rollers, having a pitch diameter of 2.815 in., roller diameter of 0.331 in., and a tapered contact angle of 15.17 degrees. The data sampling rate is 20 kHz and the data length is 20480 points. For every record of 1s duration, the RMS value and the kurtosis are computed for a general evaluation of the test. The SK-MED based approach is evaluated by evaluated by taking the maximum of the SES around the theoretical fault frequencies (outer race fault frequency BPFO, inner race fault frequency BPFI, ball fault frequency BSF) as indicators. The CS approaches is evaluated through the same indicators applied on the IES instead of the SES. These values are computed in the vicinity of the theoretical frequency as the actual fault frequency values are likely to deviate from the theoretical one.
During the test, an outer race fault in bearing 1 occurs at the end. The indicators are calculated on the vibration signals captured from Acc1 (located at the bearing 1 housing). The indicators are computed with a threshold equal to the mean of the first 200 records (assuming a healthy bearing state) plus six of their standard deviation. The obtained results are displayed in Figure 7. The RMS was able to detect a defect at the record number 542, whereas the kurtosis detects at the record 702 but subjects significant fluctuations after the record that may jeopardize the detectability (notice that the value of the kurtosis has decreased under the threshold at the record 778). It is worth noting that the RMS was efficient in this detection due to the near position of the accelerometer to the defect source, however, the RMS can’t specify the fault type. The MED-SK-SES was able to detect the fault at the same record as the RMS, yet only the BPFO value has exceeded the threshold indicating the fault type (i.e. outer race fault). Interestingly, the IES was able to detect the fault earlier at the record 533.

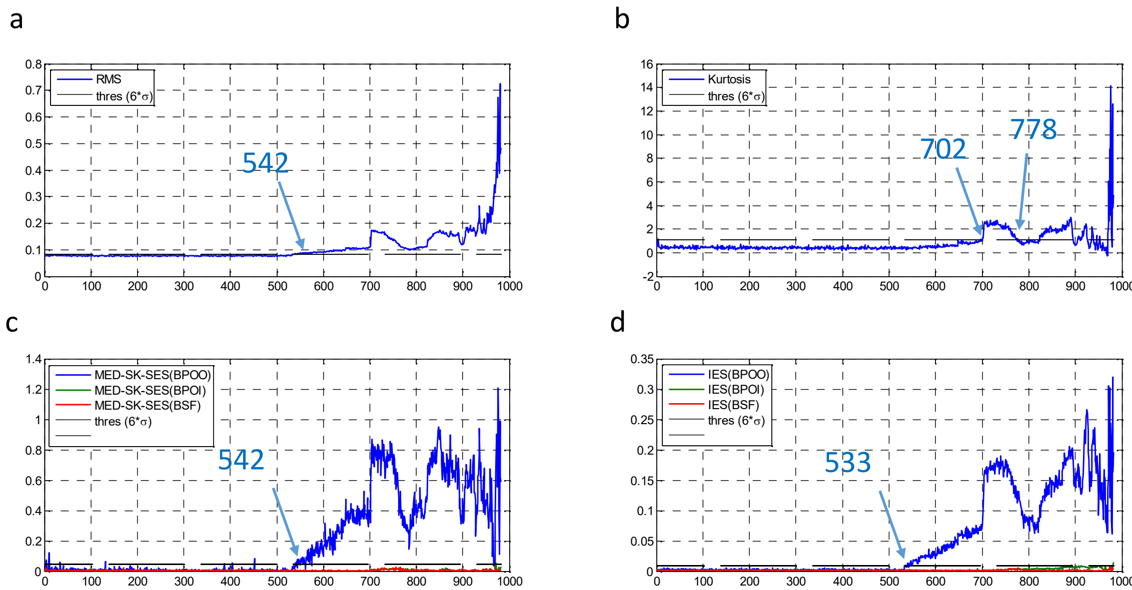

Figure 7: Indicators applied on Acc1 signals. (a) RMS, (b) kurtosis, (c) MED-SK-SES at the fault frequencies, (e) its zoom, (d) IES at the fault frequencies, (f) its zoom.

In the following, these indicators are applied on Acc3 signal (mounted on bearing 3) with the aim of evaluating the fault detectability from a remoted accelerometer. The indicators are exposed in Figure 8. The RMS was able to detect a global energetic change at the record 904 whereas the kurtosis and the MED-SK-SES have totally failed to detect any sign of failure. However, the IES was able to detect the outer race fault at the record 838.

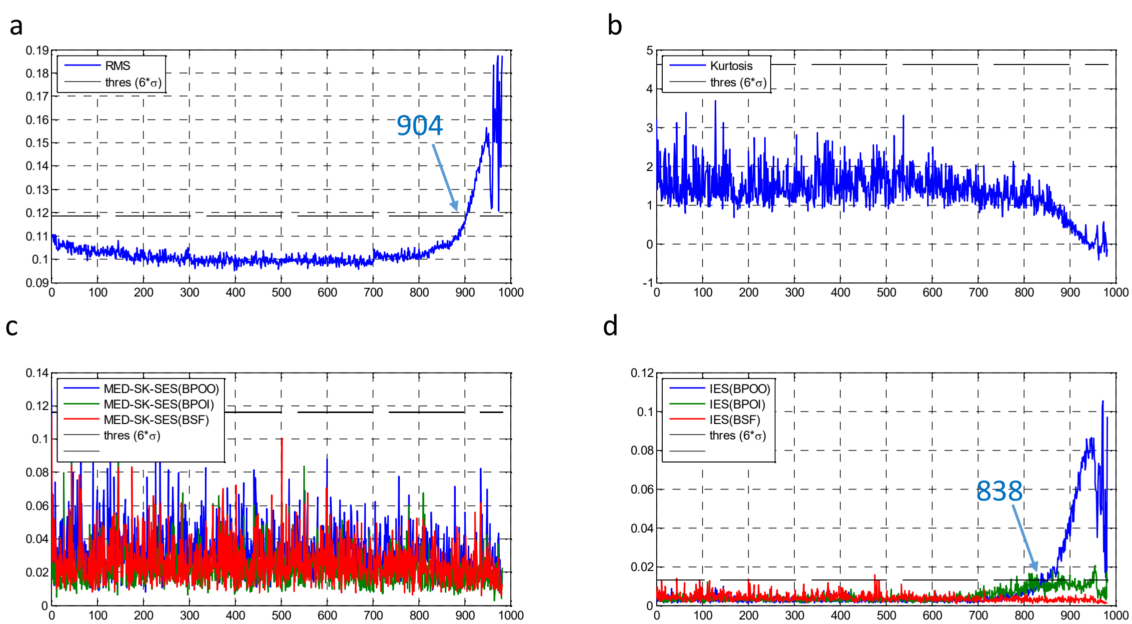

Figure 8: Indicators applied on Acc1 signals. (a) RMS, (b) kurtosis, (c) MED-SK-SES at the fault frequencies, (e) its zoom, (d) IES at the fault frequencies, (f) its zoom.

Conclusion: This paper has addressed the issue of incipient fault detection in REBs. Two leading techniques are tested on a run-to-failure test bench and compared in terms of precocity of detection and fault detectability of a remoted source. The first approach preprocess the signal by an optimized filter based on the minimum entropy deconvolution method, followed by the spectral kurtosis which design another filter that increases the signal-to-noise ratio before computing the squared envelope spectrum. The second approach is based on the cyclostationary theory. It consists of calculating the improved envelope spectrum through a projection over the spectral frequency of the squared-magnitude of the spectral coherence. Condition monitoring indicators are then designed based on these techniques. The results has shown a clear superiority of the cyclostationary approach which has proven very effective in detecting and identifying the bearing fault in its earliest stage, even when vibrations are captured from a remoted accelerometer.

References:

[1] R.B. Randall, J. Antoni, Rolling element bearing diagnostics—A tutorial, Mech. Syst. Sig. Process., Volume 25, Issue 2, p. 485-520.

[2] D. Ho, R.B. Randall, Optimisation of bearing diagnostic techniques using simulated and actual bearing fault signals, Mechanical Systems and Signal Processing 14 (5) (2000) 763–788.

[3] J. Antoni, R.B. Randall, The spectral kurtosis: application to the vibratory surveillance and diagnostics of rotating machines, Mechanical Systems and Signal Processing (2006) 308–331.

[4] R.B. Randall, J. Antoni, S. Chobsaard, The relationship between spectral correlation and envelope analysis for cyclostationary machine signals: application to ball bearing diagnostics. Mechanical Systems and Signal Processing. 15 (5) (2001) 945–962.

[5] J. Antoni, Cyclic spectral analysis of rolling-element bearing signals: Facts and fictions, Journal of Sound and Vibration 304 (2007) 497–529

[6] R.A. Wiggins, Minimum Entropy Deconvolution, Geoexploration, vol. 16, Elsevier Scientific Publishing, Amsterdam, 1978, pp. 21–35.

[7] H. Endo, R.B. Randall, Application of a minimum entropy deconvolution filter to enhance autoregressive model based gear tooth fault detection technique, Mechanical Systems and Signal Processing 21 (2) (2007) 906–919

[8] J. Lee, H. Qiu, G. Yu, J. Lin, and Rexnord Technical Services (2007). IMS, University of Cincinnati. "Bearing Data Set", NASA Ames Prognostics Data Repository (http://ti.arc.nasa.gov/project/prognostic-data-repository), NASA Ames Research Center, Moffett Field, CA

[9] Hai Qiu, Jay Lee, Jing Lin. “Wavelet Filter-based Weak Signature Detection Method and its Application on Roller Bearing Prognostics.” Journal of Sound and Vibration 289 (2006) 1066-1090