Bearing diagnostics under widely varying speed conditions

Robert Randall[1], Wade Smith[1], and Michael Coats[1]

[1] School of Mechanical and Manufacturing Engineering,
University of New South Wales, Australia
{b.randall, wade.smith}@unsw.edu.au
m.coats@yahoo.com.au

Abstract. The most powerful bearing diagnostic technique, envelope analysis, uses the fact that bearing faults typically excite high frequency resonances. Order tracking is often used to remove the effects of the speed variations so as to be able to use diagnostic techniques developed for constant speed machines. With bearings, these often rely on finding resonances where the impulsive bearing signal is enhanced, or in some cases reducing the impulse response time to ensure separation of the response pulses. This paper proposes a method whereby the signal measured over a wide speed range is first filtered in the time domain, for example to enhance the higher frequency resonances excited by the bearing faults, but also possibly curtail impulse responses whose length approaches their spacing. A speed could often then be found where the subsequent diagnostics, based on envelope analysis of the order tracked signal, is optimised.

Keywords: Bearing diagnostics; order tracking; wide speed range; spectral kurtosis; minimum entropy deconvolution; envelope analysis.

1 Introduction

The most powerful bearing diagnostic technique, envelope analysis, is based on the fact that bearing faults typically excite high frequency resonances. Because the repetition rates are not completely periodic, even at constant speed, but have a small random variation in period, the spectral content in the vicinity of the high frequency resonances is smeared, and often does not contain information about the repetition rates. Envelope analysis solves this problem by obtaining the envelope of the high frequency impulse responses, and the frequency analysis of this envelope signal gives the repetition rates even if they vary by a small percentage. The best way of finding the optimum frequency band (centre frequency and bandwidth) to enhance the impulsive bearing signals, for constant speed machines, is by using the “kurtogram” to maximize spectral kurtosis (SK) [1]. Another useful diagnostic method, in particular for high speed machines, is Minimum Entropy Deconvolution (MED), which can be used to reduce the length of the impulse responses (IRs) excited by bearing faults so that they do not overlap and reduce the effectiveness of envelope analysis. It is based on finding an inverse filter [2] to reverse the effect of the transmission path of the signal from the source to the measurement point.

When the speed of the machine varies considerably, the characteristic bearing frequencies vary with it, and a common approach is to use order tracking to remove the effects of the speed variations, with a view to being able to use the diagnostic techniques developed for constant speed machines. However, changing to a rotation angle basis means that constant frequency resonances are now smeared. Another potential problem is that even though order tracking can remove the frequency modulation effects associated with variable speed, over wider speed ranges there are also amplitude modulations, which are not removed.

One of the most powerful order tracking methods is based on phase demodulation of a speed reference signal to give a map of phase (rotation angle) vs time, and thus allow resampling from constant time intervals to constant phase intervals. One reason for performing analyses over a wide speed range is that some bearing faults might show up dominantly at certain speeds, for example found during a run-up or run-down. The phase demodulation based method of order tracking is limited to speed ranges with a maximum ratio of 2:1, but it has recently been shown that analysis of complete run-ups or run-downs can be done by dividing the signal into overlapping segments, in each of which the speed range is < 2:1, and then recombining the processed segments [3].

This paper proposes a method whereby the signal measured during a run-up is first filtered in the time and/or frequency domain for a range of purposes, such as to retain the higher frequency resonances excited by the bearing faults, or use MED to reduce IR length. Spectral kurtosis and MED can still be used even if the bearing fault pulses vary in spacing. Only after this filtration is the signal order tracked using the wide speed range method. The variation of the envelope spectrum with speed can then be studied to find where the bearing faults best show up. If necessary, the signal can be analysed in more detail in a narrow speed range around this sensitive area.

2 Experimental Measurements

2.1 Test rig

The test rig used in this study was a Bearing Prognostics Simulator (BPS) provided by SpectraQuest, Inc. [4], and is shown in Figure 1. It can be seen in the floor plan that the shaft weight is taken by two support bearings, creating a cantilever arrangement for the test bearing, which has a floating housing through which a purely radial force can be applied. It was decided to do a complete run-up in speed from 0-40 Hz over 120s, controlled by the BPS speed control unit, with a faulty bearing mounted in the test housing. Each recording also included a few seconds at constant speed at the end of the run-up. To obtain a reasonably strong signal over the full speed range while remaining within the static load limit of the bearing (7.9 kN), the radial load was kept to 7.0 kN (50% rated dynamic load). Because from experience it was known that the signal measured directly on the bearing housing (position A in Fig.1) was very clear, testing was also

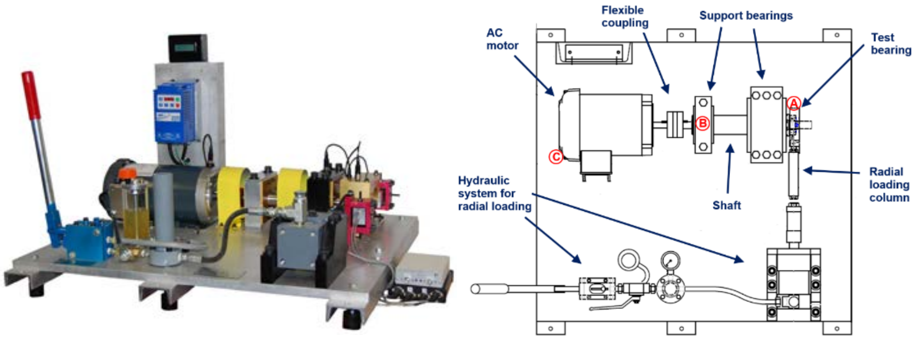

Fig. 1. SpectraQuest Bearing Prognostics Simulator. (a) general view; (b) floor plan [4].

conducted with accelerometers mounted at two other positions marked B and C on the floor plan, to give signals with more interference from back-ground noise, and a longer transmission path. The accelerometer at position A was mounted horizontally (radially), while those at positions B and C were mounted vertically. A Brüel & Kjær PULSE frequency analyser was used to record the data at a sample rate of 16,384 Hz. A once-per-rev tacho pulse was recorded from the shaft using an optical detector on a patch of reflecting tape.

2.2 Bearings and seeded faults

The test bearings used in this study were Nachi 6205-2NSE9 single-row deep-groove ball bearings, containing nine balls of d = 7.94 mm diameter, with a bearing pitch diameter of Dp = 39 mm. Two fault types were examined, both inserted into the bearing using wire-cut electro-discharge machining (EDM) without disassembling the bearing, in the outer race on one bearing (Figure 2) and inner race on the other. To simulate localised bearing faults, notches of 0.4 mm width were inserted across the entire race such that they were approximately 0.5 mm deep at the centre, sufficient to prevent the balls contacting the bottom of the notch. The outer race fault was carefully positioned in the centre of the load zone. The outer race signals are the only ones discussed in this paper.

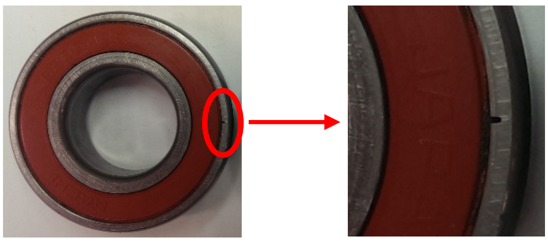

Fig. 2. Nachi 6205-2NSE9 test bearing showing seeded outer race fault

2.3 Initial Processing

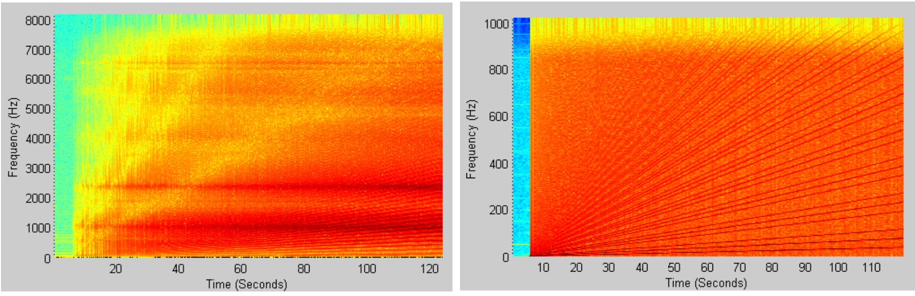

Fig. 3. Spectrograms (a) Acceleration point A (b) tacho signal, decimated by 8

Spectrograms of the various signals were first produced to check the main characteristics of the signals. Figure 3 shows spectrograms of the acceleration signal at point A over the full frequency range, and the tacho signal, decimated to a sample rate of 2048 Hz, to show the low harmonics, of which the first was used for order tracking. It is evident from the spectrograms that the signals are weak at the start, most marked with the tacho signal, so the analysis was only performed over the range from 10-100% of full speed. Horizontal bands in the acceleration spectrogram indicate resonances, of which there are two dominant ones near 1000 Hz and 2400 Hz in this signal. Moreover, the dominant harmonics in this signal are in fact the bearing outer race frequency (BPFO), which is unusual at low orders, but explained in this case because the bearing housing support has low resonance frequencies. On the other hand every fourth harmonic of the tacho signal is suppressed, indicating that the reflective tape spanned 25% of the circumference, giving the pulses an aspect ratio of 1:4.

PSD spectra were also made of the truncated run-up signals and the results are shown in Fig. 4 for the signals measured at points A and C.

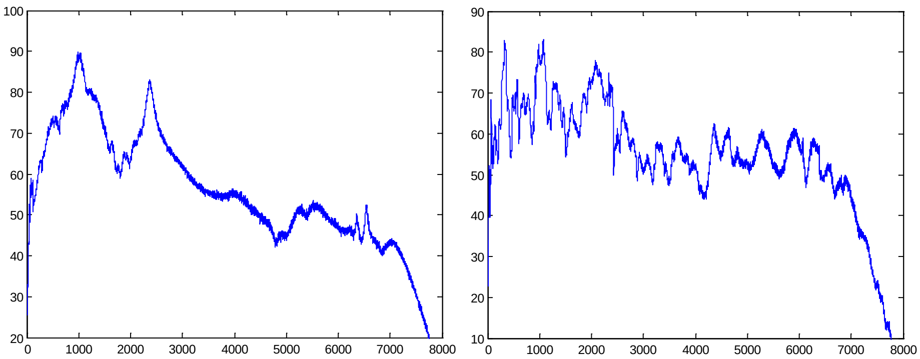

Fig. 4. PSD spectra over run-up range (a) Acceleration point A (b) Acceleration point C

It is seen that most frequency content is below 4 kHz, so the signals were all decimated by 2 to a sampling rate of 8192 Hz.

3 Signal Processing and Results

3.1 Pre-processing in Time Domain

Three methods of pre-processing were applied to enhance the extraction of the bearing signals from the background, as follows:

1. Fast kurtogram [1], to find the optimum passband to maximize the impulsiveness of the bandpass filtered signal.
2. Minimum Entropy Deconvolution (MED) [2], where an inverse filter is used to counteract the smearing effect of the transmission path from original source impulses to response signals.
3. Shortpass cepstral liftering, with an exponential window, to vastly reduce the effects of order related components while retaining modal information so that resonances carrying bearing fault signals would still be preserved. This is somewhat similar to cepstral pre-whitening [5], where the spectrum is made completely white, so that even resonances are suppressed. Fig. 5 shows the original FFT spectrum for acceleration point A, and the amplitude generated by an exponential lifter with a time constant of 5 ms (3 dB bandwidth 63.7 Hz). Time signals are generated with this amplitude and the original phase.

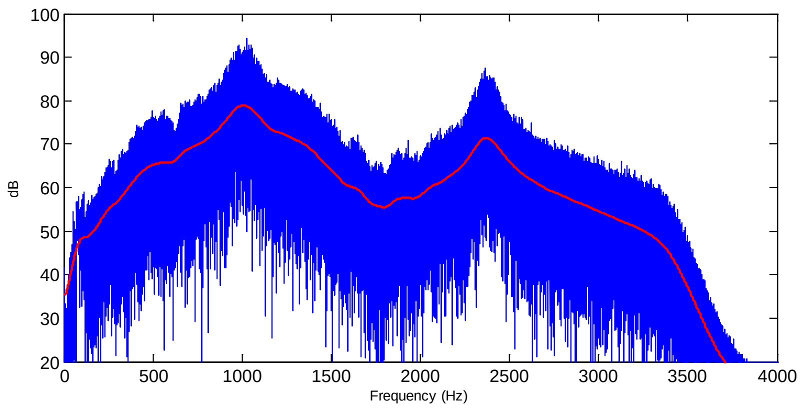

Fig. 5. FFT amplitude spectrum (blue) for Acceleration point A, and amplitude spectrum (red) used for Method 3, using an exponential “lifter” to suppress orders and enhance resonances

Figure 6 compares short sections of the envelopes of the time signals for measurement points A and C, at 25% full speed and full speed. This is done for the untreated signal and those processed by the three methods above.

It is seen that the original bearing fault pulses are reasonably well separated over the full speed range for acceleration point A, though they are starting to close over at the highest speed. At point C, the IRs are longer; still separated at low speed, but overlapping at full speed. On detailed examination, the long IRs could be seen to be due to the 300 Hz resonance in Fig. 4(b).

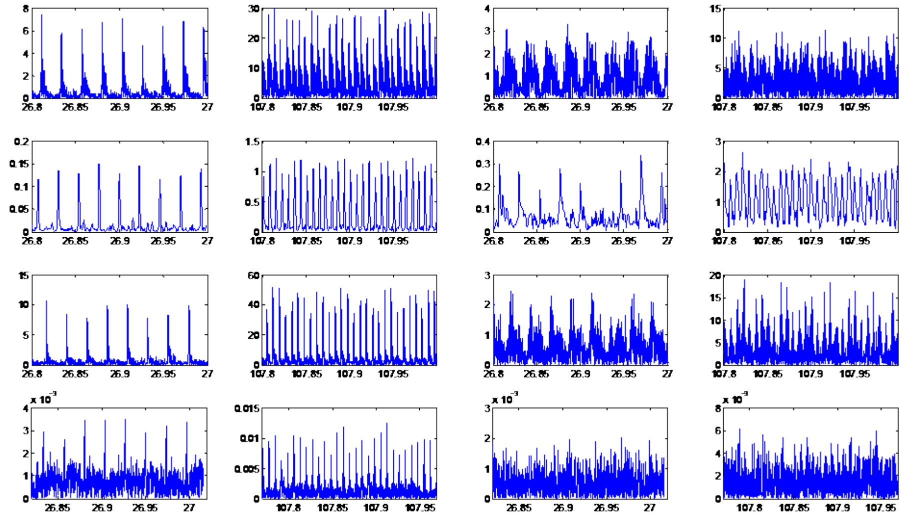

Fig. 6. Signal envelopes for different methods (Row 1) Original (Row 2) SK filtered (Row 3) MED (Row 4) Exponential liftered (Column 1) A - 25% (Column 2) A - end (Column 3) C - 25% (Column 4) C - end

The SK based filtering using the kurtogram gives excellent results for both signals, with both low and high speed signals slightly better separated for point A and a marked improvement for point C. The kurtograms for the two signals are shown in Fig. 7, from which it can be seen that the optimum bands are different. For point A it is 1024 Hz centred on 3584 Hz, and for point C it is 512 Hz centred on 1792 Hz. Note that the maximum kurtosis for the former is 14, against only 3.5 for the latter.

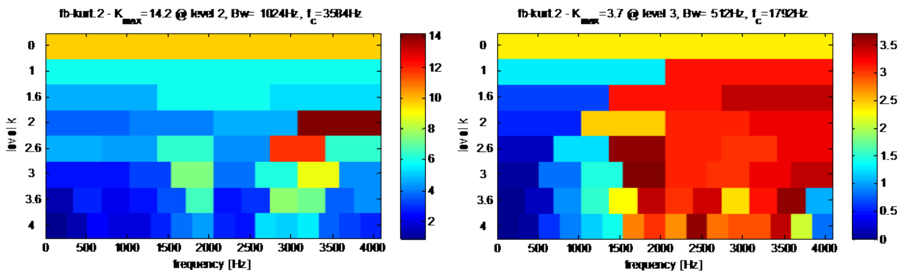

Fig. 7. Kurtograms for acceleration (a) Point A (b) Point C

The MED method does at least as well for the signal from point A, but not as well as the SK filtering for point C. It is possible that the extremely wide range of resonances in the latter signal (see Fig. 4(b)) made it difficult to generate an inverse filter for the whole range.

Somewhat similarly, the cepstral liftering method also performs well for the signal from point A, but is possibly worse than the original for point C. It seems that it is only useful when the resonances carry information about the bearing faults, and that is evidently not the case for point C, remote from the bearing. However, most bearing monitoring is carried out using accelerometers mounted on the bearings themselves, and this method has been found useful in other cases.

3.2 Order Tracking

The method used is that described in [3], where the signal is divided up into sections, in each of which the speed range is less than 2:1 (as a maximum, also limited by other considerations). This is so that phase demodulation can be used on a speed reference signal (preferably first harmonic, to give the greatest speed range) to give a map of phase (rotation angle) vs time to allow resampling from equal steps in time to equal steps in rotation angle, which is at the heart of order tracking. The advantage of the phase demodulation method is that if the order being demodulated has no overlap in frequency with other orders or disturbing components, the phase vs time curve can be reconstructed to any degree of resolution. The sections are windowed with half-Hanning windows in the overlap regions to allow seamless joining after processing. Even though the windows are distorted in transforming from the time to angle domain, they still add to unity everywhere so the scaling is unchanged.

It is often possible to use a separated harmonic of the vibration signal as a reference signal, but here the first harmonic of the tacho signal was used. Figure 8 contains spectrograms of the order tracked signals from points A and C, for comparison with Fig. 3(a). The orders are now horizontal lines and the resonances hyperbolic curves. Note that as the sampling frequency reduces with speed, it is necessary to apply a lowpass filter (fixed frequency in each segment, but progressively lower) to avoid aliasing. This was done by an ideal filter in the frequency domain, corresponding to half the sampling frequency at the lowest speed for each segment. In the case shown, 100 orders would be valid for the whole speed range, even though the Nyquist “frequency” is 150 orders.

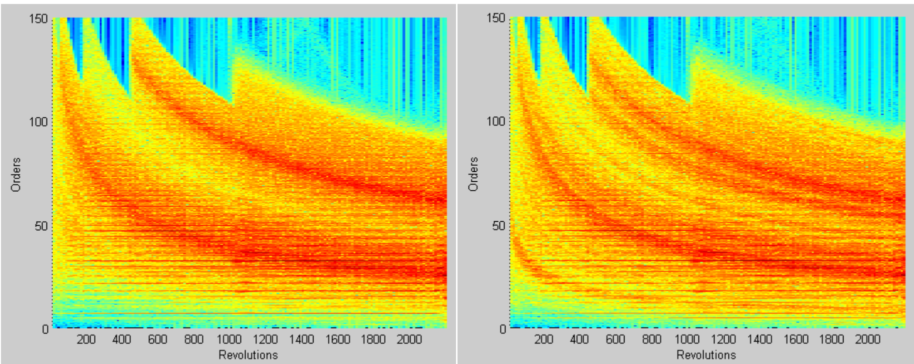

Fig. 8. Spectrograms (a) Point A (b) Point C

In order to check how the bearing envelope analyses might change with speed, spectrograms were also produced for the envelope signal over the whole speed range. However, since only the lower harmonics were of interest, the envelope signals were decimated by 3 (to give a maximum of 33 valid orders) and only 20 are displayed (Fig. 9).

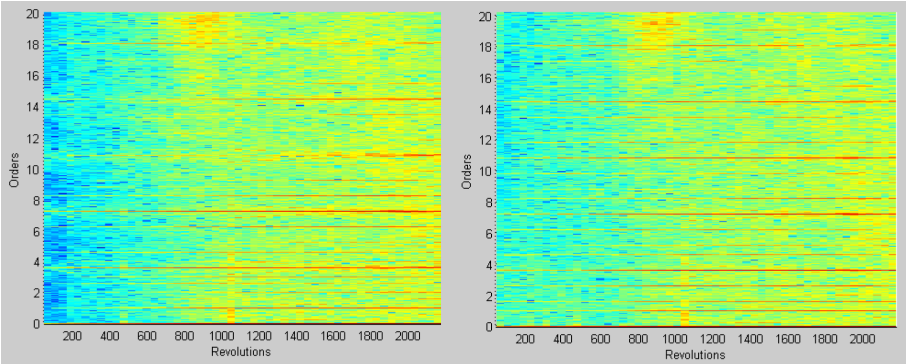

Fig. 9. Envelope spectrograms (a) Point A (b) Point C

In interpreting these order tracked diagrams, it should be kept in mind that with speed increasing linearly with time, the rotation angle goes up with the square of speed, so for example, the revolution range from 1100-2200 would represent the speed range from 70-100%, and from 550-2200 revs, the speed range (and time) from 50-100%.

Both spectrograms indicate that a satisfactory diagnosis could be made over most of the range, with several harmonics of BPFO (order 3.58) but with some low harmonics and side bands spaced at shaft speed. This is somewhat unusual for an outer race fault, which would only be modulated at shaft speed if there were a rotating force such as an unbalance.

Figure 10 shows spectrograms of the envelopes of the point C signal pre-processed by MED and exponential liftering.

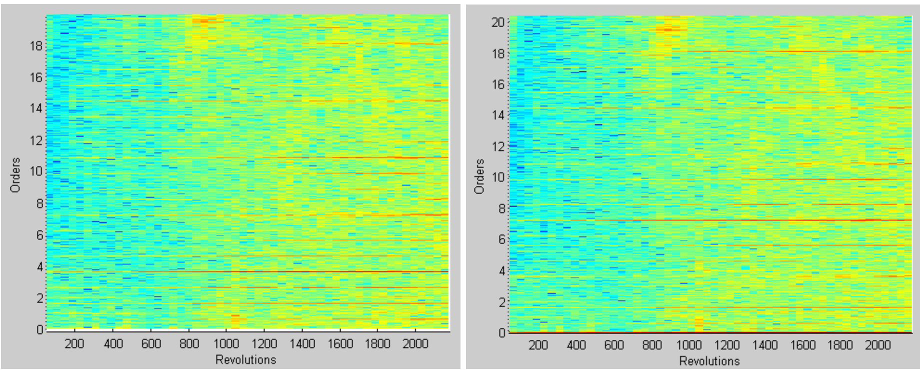

Fig. 10. Envelope spectrograms Point C (a) After MED (b) After exponential liftering

It appears that this data is so clean that there is not much difficulty to diagnose it over quite a range of speeds, even using the measurement from Point C, which appears a bit noisy in the time domain (Fig. 6).

Figure 11 shows (squared) envelope spectra for some of the cases depicted in Figs. 9 and 10, from two different sections of the record. Although the lower speed sections are taken from ¼ of the distance along the order tracked record (500-600 revs), they correspond roughly to half way along the time records, rather than ¼ as in Fig. 6.

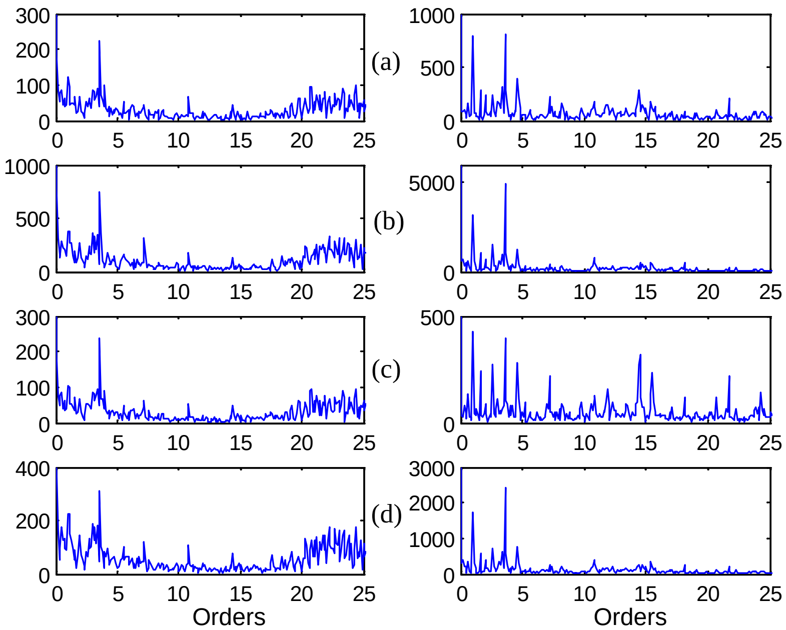

Fig. 11. Envelope spectra at different positions along record; Left: 25%, Right: near end (a) Point A (b) Point C (c) C after MED (d) C After exponential liftering

In the end there is little to choose between the various methods, and measurement points, despite the time signals appearing different in Fig. 6. It seems the envelope analysis method is very robust. There is evidently more modulation at shaft speed at the end of the record, where the speed is higher, which points to the possibility of unbalance being a factor. This is most marked at Point C.

It can perhaps be remarked that the MED method has given stronger high harmonics, possibly because of making things more impulsive, in particular at high speed where the original pulses tend to run together.

4 Discussion and Conclusion

This paper has shown that in variable speed situations, illustrated here by a run-up over a wide speed range, it is advisable to treat bearing fault signals in the time domain first, because many characteristics are related to resonance frequencies which are constant in the time/frequency domains (as opposed to rotation angle/order). After enhancement in the time domain, the signals can then be order tracked to allow diagnostics by envelope analysis, the bearing fault “frequencies” being related to harmonic orders.

It should be noted that the bandpass signals found by the kurtogram were not order tracked here, as this would require a little further software development. The bandpass filtered signals are 1-sided in frequency, and thus complex in time, and it still has to be investigated whether the order tracking can be carried out on the frequency shifted complex signals, or possibly on their envelopes at a lower sampling rate, which would be very efficient. The least efficient possibility would be to generate full bandwidth signals, but bandpass filtered in the band shown by the kurtogram, and then apply order tracking to these real signals.

The various techniques used were shown to work well in this situation, but the data was relatively straight forward to diagnose, and gave good results even with simple methods. Even so, it is thought that other data will be found where the more complex signal processing, such as MED and exponential liftering, will give an improvement over the untreated data, and this has been shown to be possible with the methods illustrated.

Acknowledgements. This research was supported by the Australian Research Council and SpectraQuest, through Linkage Project [LP110200738].

References

1. Antoni, J., Fast Computation of the Kurtogram for the Detection of Transient Faults, Mech. Syst. and Sig. Proc., 21(1), pp. 108–124, (2006).
2. Sawalhi, N., Randall R.B., Endo, H., The Enhancement of Fault Detection and Diagnosis in Rolling Element Bearings using Minimum Entropy Deconvolution Combined with Spectral Kurtosis. Mech. Syst. and Sig. Proc., 21(6), pp 2616-2633, (2007).
3. Coats, M., Randall R.B., Order Tracking over a Wide Speed Range such as a Run-up or Run-down, IFToMM ICORD 2014 conference, Milan, September (2014).
4. SpectraQuest, Bearing Prognostics Simulator; Available from: http://spectraquest.com.
5. Sawalhi, N., Randall, R.B., Signal Prewhitening using Cepstrum Editing (Liftering) to Enhance Fault Detection in Rolling Element Bearings, in COMADEM. 2011: Stavanger, Norway (2011).