THE APPLICATION OF SPECTRAL KURTOSIS TO BEARING DIAGNOSTICS

Nader Sawalhi, Robert B. Randall

School of Mechanical and Manufacturing Engineering, The University of New South Wales, Sydney 2052, Australia

Abstract

The choice of demodulation band for envelope analysis of faulty bearings is often made by spectrum comparison with a healthy bearing, to choose resonance frequencies where the largest change occurred as a result of the fault. It has recently been established that the so-called "spectral kurtosis" gives a very similar indication of the band to be demodulated without requiring historical data. The kurtosis is a statistical parameter based on the fourth moment of a signal, which is close to zero for gaussian noise and other stationary signals, but large for impulsive signals containing series of short transients, such as a bearing fault signal. The spectral kurtosis (SK) is obtained by calculating the kurtosis for each frequency line in a time-frequency diagram. It has also been found that the spectral kurtosis can be used to form a filter to select out that part of the signal that is most impulsive, considerably reducing the background noise and improving the diagnostic capability. The initial definition of the SK used the short time Fourier transform (STFT) for the time-frequency analysis, but this does give some artifacts and anomalies in the results, and the paper discusses the potential use of other time/frequency analyses such as Wigner-Ville and wavelets. The paper illustrates the use of the SK for bearing diagnostics, including a number of extensions and improvements in the basic technique and the choice of optimum analysis parameters. The results are illustrated using simulated and actual bearing fault signals.

Introduction

Spectral kurtosis (SK) is a statistical parameter indicating how the impulsiveness of a signal varies with frequency. Since faults in rolling element bearings give rise to a series of short impulse responses as the rolling elements strike faults on the races, the SK is potentially useful for determining the frequency bands dominated by the bearing fault signals, usually containing resonance frequencies excited by the faults. This information is required in order to choose the optimum frequency band to demodulate to perform "envelope analysis" on the bearing signals. It can be shown that the raw spectrum of a bearing fault signal contains little diagnostic information, being dominated by the resonance frequencies excited, whereas the envelope signal, obtained by amplitude demodulation, contains the required information about the repetition frequency of impacts, and any modulation caused by passage of the fault through the load zone.

There has long been a debate about the best method of choosing the frequency band to demodulate for envelope analysis. Some have recommended analyzing the response to hammer taps on the bearings, but the most logical method is based on comparing the spectra before and after the occurrence of a change thought to be due to a bearing fault. This usually occurs at very high frequencies compared with the rate of excitation, and small random fluctuations in the latter cause the response spectrum to smear any harmonics that could otherwise indicate the repetition frequencies. However, if the change is due to a bearing fault, the dB difference (ie $20\log_{10}(A_s/A_n)$, where A_s is the amplitude of the spectrum value with fault, and A_n is the original spectrum value, or masking noise) would indicate where the signal/noise ratio of the bearing signal to background is highest, and this gives a good guide as to the optimum centre frequency of the demodulation band. In Ref. [1] it is shown that the width of the band demodulated directly determines the highest frequency in the resulting envelope spectrum, and this will be sufficient to diagnose any fault type if it is at least four times the inner race fault frequency.

Spectral Kurtosis - Background

The SK was introduced in the eighties for detecting transients in sonar signals [2]. After this isolated application, it was recently reintroduced in the signal processing community for distinguishing between different types of signals [3].

In both of these applications, the SK was defined in terms of the Short Time Fourier Transform (STFT), a way of obtaining a time/frequency diagram of a signal. As illustrated in Figure 1, the STFT is obtained by moving a short time window, such as Hanning, along the record in overlapping steps. The spectra for each window position are arranged in a 3-D diagram with time on one axis and frequency on the other. The SK is then calculated for each frequency line in the time direction. In [3], the STFT is made up of complex spectra, and the SK is calculated in a special way which makes it possible to distinguish between a value of –1 for a sinusoid, and zero for a gaussian random signal, but for the current application it is only necessary to distinguish between high values (of the order of 10-30), for an impulsive signal, and near zero for stationary signals. It is thus convenient to use the normal definition of kurtosis as the fourth moment of the signal $X(t)$, normalized by the square of the second moment minus 3 to make it zero for

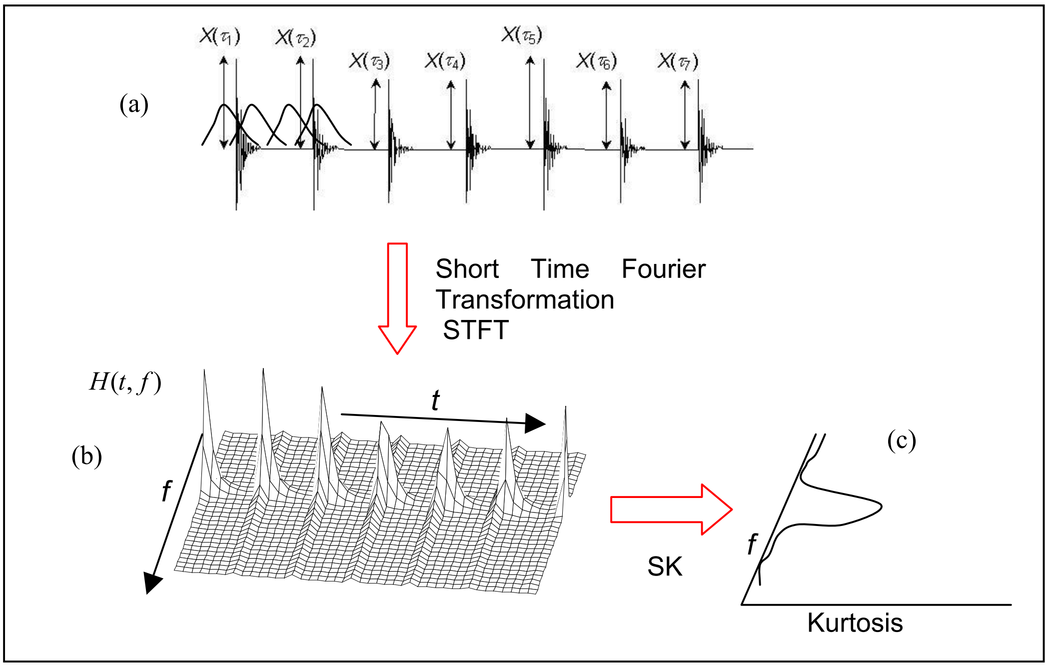

Figure 1: Calculation of SK for a simulated bearing fault signal
(a) Time signal, showing moving time windows. (b) Amplitude of STFT. (c) Spectral kurtosis vs frequency

a gaussian random signal. Thus:

$$K_X = \left\langle E\left\{X^4(t)\right\}\right\rangle / \left\langle E\left\{X^2(t)\right\}\right\rangle^2 - 3 \qquad (1)$$

In Ref. [4], it is shown that a non-stationary signal can be modelled as a stationary gaussian white noise modulated by a time and frequency varying complex envelope function such as $H(t,f)$ in Fig. 1 (modulus shown). The square of this modulus $|H(t,f)|^2$ is proportional to the value of the power spectrum of the STFT for each time slice, which can be shown to be proportional to the second moment $\left\langle E\left\{X^2(t)\right\}\right\rangle^2$ at each point in time in Equ. (1). The square of this again is proportional to the fourth moment in Equ. (1).

Selection of Demodulation Band

Figure 2 compares spectra from a gear test rig with and without an inner race bearing fault. The spectrum range is 0-24 kHz, but it is seen that the fault only causes a spectrum increase above 8 kHz. The faulty spectrum has the appearance of white noise, and thus does not contain diagnostic information as such. Fig. 2(b) shows the spectrum difference in dB, while Fig. 2(c) shows the SK. It is seen that the SK has almost the same shape as the dB spectrum difference, and can therefore equally well be used to select the optimum demodulation bandwidth.

Somewhat surprisingly, the actual values of kurtosis and dB difference are roughly the same. This was also found to be the case for a less extensive outer race fault, where the full-scale values for both dB difference and SK were approximately 6. However, the actual scaling of SK is affected by the size of window used (see later) so it is only remarkable that the two parameters are proportional.

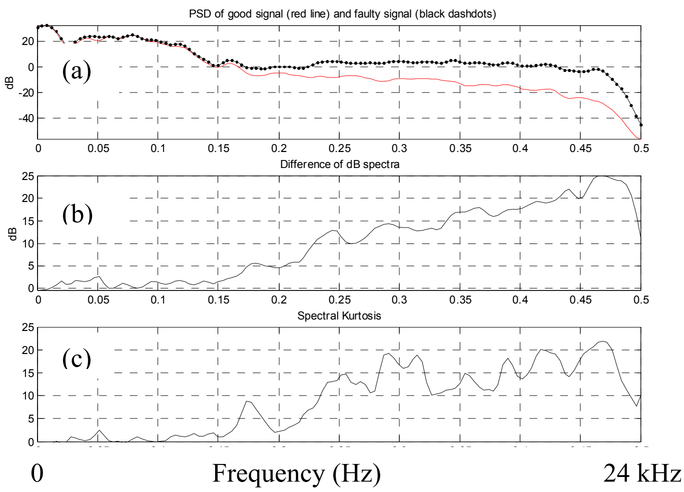

Figure 2 (a) dB spectrum comparison with and without the inner race fault (b) dB spectrum difference (c) Spectral kurtosis

Figure 3 shows a typical envelope spectrum obtained by demodulating a band of width 150 Hz in the vicinity of 14 kHz, where there are local peaks in both the dB difference and SK. It is typical of an inner race fault, with harmonics of BPFI (ballpass frequency, inner race) and harmonics and sidebands with a spacing equal to the shaft speed (the rate at which the inner race fault passes

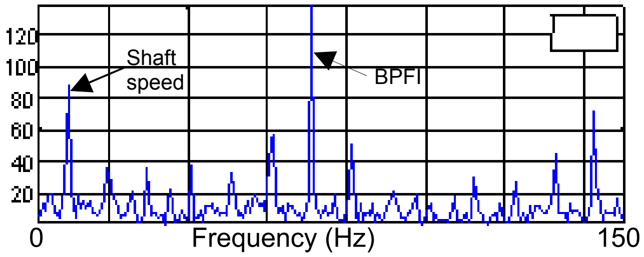

Figure 3. Envelope spectrum from demodulating a band centred on 14 kHz, where there are local peaks in both the dB difference and SK of Fig. 2.

through the load zone). Similarly good results were obtained for outer race and rolling element faults.

SK as a Filter

Since the SK is large in frequency bands where the impulsive bearing fault signal is dominant, and effectively zero where the spectrum is dominated by stationary components (see Fig. 2) it makes sense to use it as a filter function to filter out that part of the signal with the highest level of impulsiveness. In Ref. [5] it is shown that it is possible to define both Wiener and matched filters in terms of the SK for an impulsive signal, such as generated by bearing faults, masked by stationary random noise. It is shown that the Wiener filter, which minimises the difference from the signal targeted for extraction by the filter, is proportional to the square root of the SK, while the matched filter, which maximises the similarity with the targeted signal can also be estimated. If $W(f)$ is the Wiener filter, it can be shown that the equivalent matched filter $M(f)$ is given by:

$$M(f) = \frac{W(f)}{1 - W(f)} \qquad (2)$$

In Ref. [5] it is shown that this corresponds to a bandpass filter centred on the maximum value of SK.

Figure 4 shows the result of applying the optimum Wiener and matched filters to a signal for a weak outer race bearing fault, masked by gear noise. In this case the matched filter produces the best result, but the result of envelope analysis in either case clearly reveals the outer race fault frequency visible in the filtered time signal.

A recent development, developed by Antoni and described in Ref. [5], is the "kurtogram", which is a diagram indicating the optimum centre frequency and bandwidth combination of a bandpass filter to maximize the kurtosis of the filter output.

Figure 5 shows the kurtogram for the outer race fault of Fig. 4. Figure 6 shows the optimum bandpass filter derived from it, compared against the SK for all frequencies and finally Figure 7 shows the filtered time signal resulting from this signal.

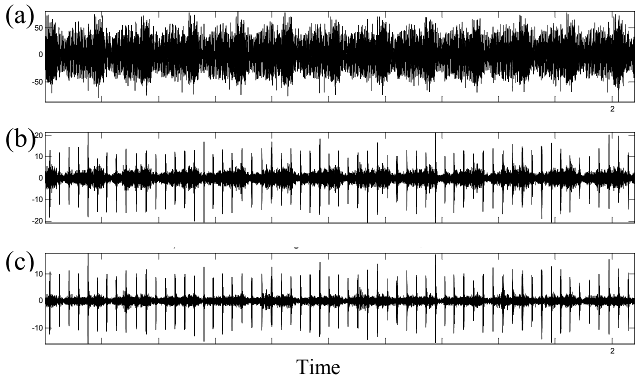

Figure 4. Use of SK as a filter on a signal from a gear test rig with a weak outer race bearing fault. (a) Total signal (b) Output from Wiener filter (c) Output from matched filter (from [5])

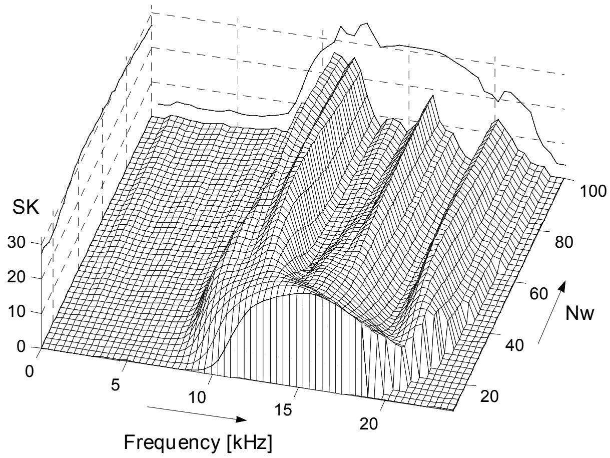

Figure 5. Kurtogram for signal of Fig. 4 (from [5]) N_W is window length defining spectral resolution. SK is spectral kurtosis. Maxima are projected onto each plane.

The kurtosis of the signal in Fig. 7 is 46.7, compared with 9.9 and 25.8, respectively, for the Wiener and matched filters of Fig. 4.

Choice of Optimum Window

As mentioned above, the calculated SK does depend on the choice of analysis parameters, in particular the length of time window, which has to give a balance between separating the individual bursts and encompassing them within the window. Figure 8 gives an example for a rolling element fault of SK calculated for window lengths of 16, 32, 64, 128, 256 samples, with consequent inverse frequency resolution (seen in the accompanying PSD spectra averaged over the whole record). It is evident that 16 gives insufficient frequency resolution, while > 64 reduces the values of SK, so in this case either 32 or 64 should be chosen.

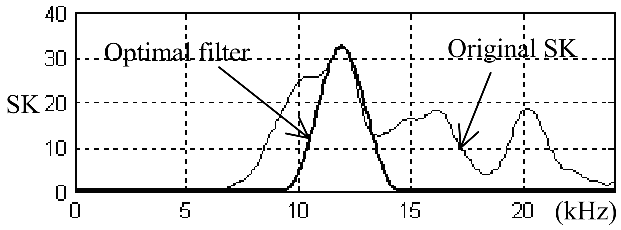

Figure 6. Optimal bandpass filter compared with SK (from [5]).

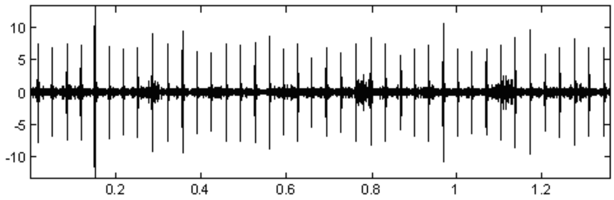

Figure 7. Outer race fault signal obtained using the filter of Fig. 6 (from [5]).

In order to investigate the question of window length in more detail, a study was made using simulated impulse responses, as illustrated in Figure 9, with a range of effective lengths and spacings. This uncovered an anomaly, as also shown in Fig. 9, that at the resonance frequency, with the peak spectral response there was actually a drop in SK, in particular for shorter window lengths. On investigation this could be seen to be due to the phenomenon shown in Fig. 10, which compares slices of the STFT diagram at the resonance and on the flank at one side. Even though the peak value is greatest at the resonance frequency, this is more than compensated by the shorter length of the smaller pulse, which when averaged across the total time between pulses gives a greater value of kurtosis. This does not seem to give problems in practice, at least when the bearing fault manifests itself in a frequency range with overlapping resonances (as in Fig. 2), but in any case it led to an investigation of how to mitigate it. It was found that by pre-whitening the spectrum the anomaly can be avoided (because the pulses become shorter at all frequencies). Pre-whitening can be achieved by using an autoregressive (AR) model to fit the signal The residual error of this model is in fact a representation of the noise and the nonstationary part of the signal, and the spectrum of this residual represents the white spectrum.

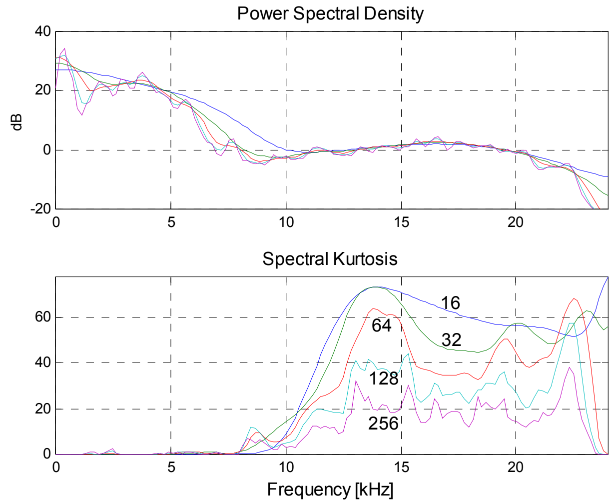

Figure 8. Figure 5: (a) PSD spectra with the different window lengths. (b) SK calculated for the indicated window length in samples

In an AR model, a value at time t is based upon a linear combination of prior values as in:

$$Y_k = \sum_{i=1}^{p} a_i x_{i+k} + \varepsilon_v \tag{3}$$

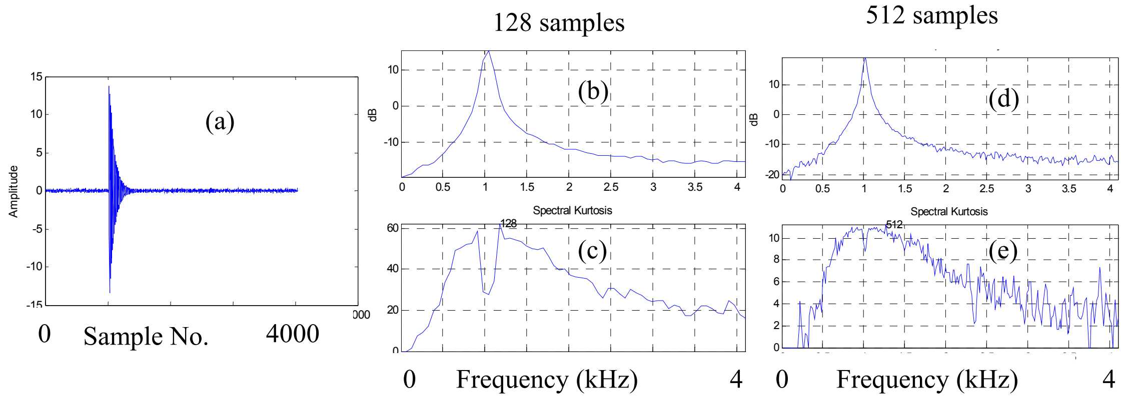

Figure 9: Effect of window length on PSD and SK for a simulated fault pulse
(a) Single pulse with additive noise (SNR 20 dB) (b) PSD (128 samples)
(c) SK (128 samples) (d) PSD (512 samples) (e) SK (512 samples

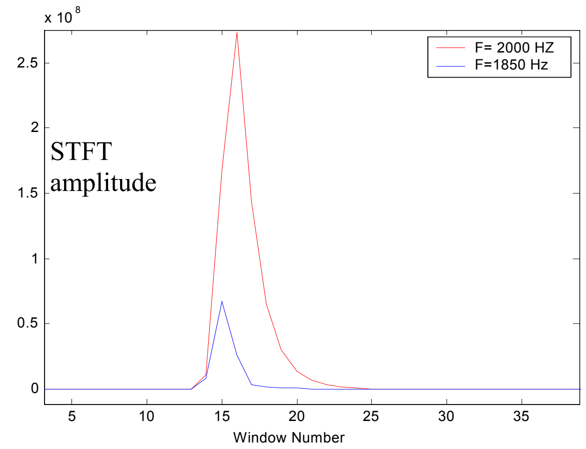

Figure 10. Comparison of slices through the STFT diagram at the resonance frequency (2000 Hz) and a frequency on the flank (1850 Hz).

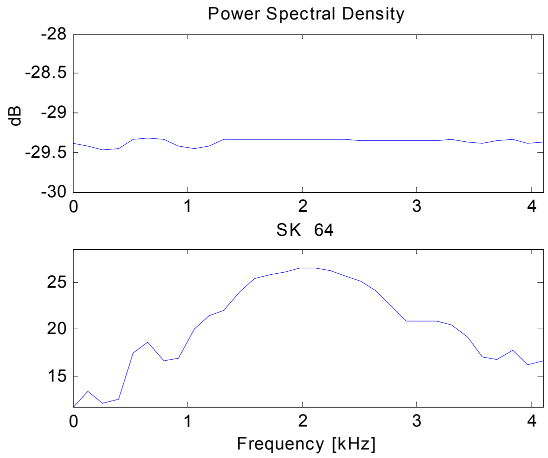

Figure 12. PSD and SK for a window length of 64 after pre-whitening the PSD

The error term ε_V represents the difference between the actual and linearly predicted values, and is white. It contains additive white noise and nonstationarities in the form of impulses, also with a white spectrum. In this study we used the linear prediction filter in Matlab® to make the AR model and obtain the residual.

Figure 11 shows the synthesised signal of Fig. 9 after prewhitening, where the impulse response is seen to have become a delta function. Figure 12 shows the corresponding power spectrum (almost white), and the SK which is almost unchanged, except that the dip at the former resonance frequency has disappeared. The shorter window length of 64 samples gives better separation of the impulses now that they are shorter.

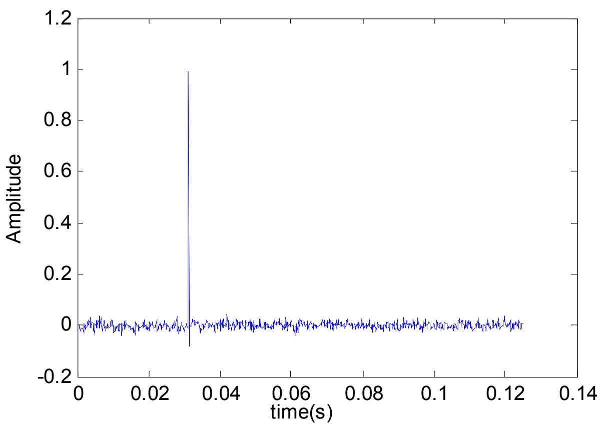

Figure 11. Result of prewhitening the transient of Fig. 9

In Ref. [6] it is shown that prewhitening like this also gives a better diagnostic capability in a difficult case where a weak outer race fault could not be detected before prewhitening, but the latter increased the impulsiveness (and kurtosis) and facilitated the diagnosis.

Alternatives to the STFT

Alternative time/frequency analysis techniques to the STFT were considered, such as the Wigner-Ville distribution (WVD) [7] and wavelet analysis. The former theoretically gives better simultaneous time and frequency resolution than the STFT, at the cost of (non-physical) interference components which are difficult to distinguish from genuine components. A brief investigation showed that the WVD suffered at least as much as the STFT from the anomalous effect illustrated in Fig. 9, and so this was not pursued further.

In Ref. [8], bearing signals were analysed using wavelet analysis as an alternative to the STFT. Complex continuous Morlet wavelets were used, which can be considered as a kind of filter bank with constant proportional bandwidth (uniform resolution on a logarithmic frequency scale). The scale used was approximately the equivalent of 1/5-octave filters.

Figure 13 shows the envelope of the resulting wavelet coefficients for a signal with a bearing inner race fault (similar to that analysed in Figs. 2 and 3). It is seen that

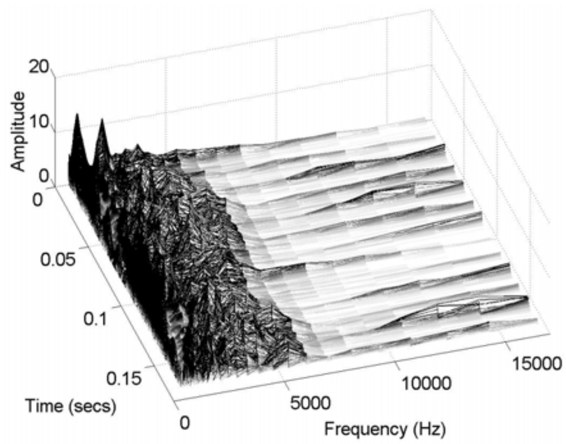

Figure 13. Modulus of wavelet coefficients for a signal with an inner race bearing fault. [9]

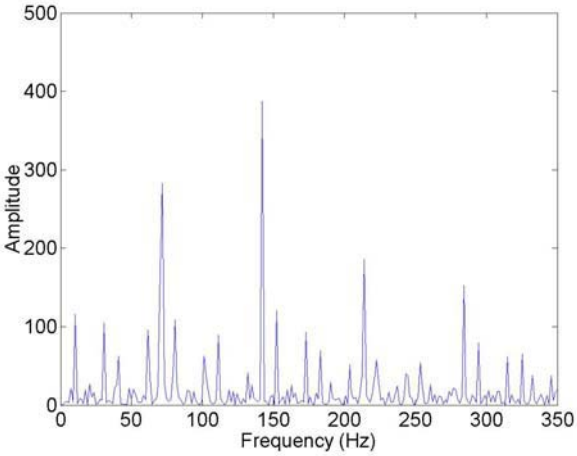

Figure 14. Envelope spectrum of wavelet coefficients from Fig. 13 centred on 14 kHz. [9].

the wavelet analysis gives increasingly better time resolution at higher frequencies, at the expense of poorer frequency resolution, and this is the main difference with respect to the STFT.

Figure 14 shows the envelope spectrum obtained by FFT analysis of the wavelet coefficients (moduli) centred on frequency 14,158 Hz. It is seen to be very similar to the results of Figure 3. Thus at this stage it does not appear that wavelet analysis offers any particular advantage over the STFT (except the possibility of automatically adjusting time resolution at higher frequencies) and does not lend itself as readily to use for generating filters.

Conclusions

Spectral kurtosis is an exciting new analysis technique for use in the diagnostics of rolling element bearing faults. It can be used to indicate the best band to demodulate for envelope analysis without requiring historical data as for the previously best method involving spectrum comparison. It can also be used to determine the optimum bandwidth as well as centre frequency to maximize the kurtosis of the filter output. The spectral kurtosis can also be used to generate filters to separate the most impulsive part of a signal from stationary masking components and aid diagnosis. It can be an advantage to use prewhitening of the signal before analysis in order to best extract the bearing fault signal from the background noise.

Acknowledgements

This research is supported by DSTO through their Centre of Expertise scheme.

References

[1] D. Ho and R.B. Randall, "Optimisation of Bearing Diagnostic Techniques Using Simulated and Actual Bearing Fault Signals", *Mechanical Systems and Signal Processing*, **14** (5), September, 763-788, 2000.

[2] R.F. Dwyer, "Use of the kurtosis statistic in the frequency domain as an aid in detecting random signals", *IEEE Journal of Oceanic Engineering*, **OE-9**, No.2, 85-92, Apr. 1984.

[3] V.D. Vrabie, P. Granjon, C. Servière, "Spectral kurtosis: from definition to application", IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing, Grado, Italy, June 8-11, 2003.

[4] J.Antoni, "The Spectral Kurtosis: A useful tool for characterising nonstationary signals", *Mechanical Systems and Signal Processing*, submitted for publication, 2004.

[5] J.Antoni and R.B.Randall, "The Spectral Kurtosis: Application to the Surveillance and Diagnostics of Rotating Machines", *Mechanical Systems and Signal Processing* submitted for publication, 2004.

[6] N. Sawalhi and R.B. Randall, "Spectral Kurtosis Enhancement using Autoregressive Models". Paper submitted to ACAM2005 conference, Melbourne, February 2005.

[7] L. Cohen, *Time-frequency Analysis*, Prentice-Hall, NJ, 1995.

[8] L. Shi, J. Antoni, and R.B. Randall, "Rolling Element Bearing Fault Detection using Envelope Spectrum based on Wavelet Transform and Spectral Kurtosis". Submitted to IMechE Conference on Vibrations in Rotating Machines, Swansea, September 2004.