# Monitoring gears by vibration measurements: Lempel-Ziv complexity and Approximate Entropy as diagnostic tools

Mourad Kedadouche, Marc Thomas[a] , Antoine Tahan and Raynald Guilbault

Department of Mechanical Engineering, École de Technologie Supérieure.1100, Notre-Dame street West, Montreal, H3C 1K3, Quebec, CANADA

**Abstract.** Unexpected failures of industrial gearboxes may cause significant economic losses. It is therefore important to detect early fault symptoms. This paper introduces signal processing methods based on approximate entropy ( *ApEn* ) and Lempel-Ziv Complexity ( *LZC* ) for defect detection of gears. Both methods are statistical measurements exploring the regularity of a vibratory signal. Applied to gear signals, the parameter selection of *ApEn* and *LZC* calculation are first numerically investigated, and appropriate parameters are suggested. Finally, an experimental study is presented to investigate the effectiveness of these indicators. The results demonstrate that *ApEn* and *LZC* provide alternative features for signal processing. A new methodology is presented combining both Kurtosis and *LZC* for early detection of faults. The results show that this proposed method may be used as an effective tool for early detection of gear faults.

## 1 Introduction

Gearboxes play an important role in industrial applications and unexpected failures often result in significant economic losses. Numerous papers considering gear condition monitoring through vibration measurements were published over the years. Compared to classical techniques such as statistical time indicators or Fast Fourier Transform, advanced signal processing techniques like time-frequency analysis (STFT, Wigner-Ville) [1-4] or wavelet transform [5, 6] have shown to be more efficient for gear defect detection. Baydar *et al.* [7-10] proposed various methods such as the instantaneous power spectrum, Wigner–Ville distribution and the wavelet transform method for local tooth fault detection from vibration and acoustic signals. Yesilyurt [11] applied the spectrogram and scalogram approach for gearbox fault detection. The Lempel-Ziv complexity ( *LZC* ) and approximate entropy ( *ApEn* ) present alternative tools for signal analysis involving nonlinear dynamics. These methods are becoming popular and have found wide applications in various disciplines, especially in the field of biomedical engineering. *ApEn* has recently received more attention. Yan [12] investigated the application of *ApEn* for the health monitoring of rolling element bearings. Y. He and X. Zhang [13] applied the *ApEn* for monitoring acoustic emission signals from defects in rolling element Bearings. Fu *et al.* [14] used the *ApEn* method to fault signal analysis in electric power system. Using *ApEn* , Xu *et al.* [15] detected the looseness of the bearing bushing in turbo generator. In all these works, *ApEn* is used as a nonlinear feature parameter for analysing the vibration signal and effectively identifying the health conditions of the mechanical system. On the other hand, Yan and Gao [16] investigated the application of Lemp-Ziv complexity ( *LZC* ) for the health monitoring of rolling element bearings. Wang *et al.* [17] compare and analyse quantitative diagnosis methods based on Lempel-Ziv complexity for bearing faults by using continuous wavelet transform (CWT), Empirical Mode Decomposition (EMD) method, and wavelet packet method for decomposition of vibration signal. Kedadouche *et al.* [18] combined *LZC* and EMD for early detection of gears cracks. As described above, *LZC* and *ApEn* are becoming more and more attractive in the field of detection and fault diagnosis. However, no work has been found to apply *ApEn* or *LZC* measurements for diagnosis of gear faults. Therefore, in this paper, *ApEn* and *LZC* are introduced for analysing vibration signals from gears and investigating their efficiency for the defect detection and severity evaluation of gears faults.

# 2 THEORETICAL BACKGROUND

## 2.1 Approximate entropy

Consider a time series $S(i)$, $i = 0...N$. Its ''regularity'' may be measured by $ApEn$ in a multiple dimensional space so that a series of vectors are constructed and expressed as follows:

$$X(N-m+1) = \{x(N-m+1), x(N-m+2), \dots, x(N)\} \quad (1)$$

Each vector is composed of m consecutive and discrete point data of the time series $S$. The distance between two vectors $X(i)$ and $X(j)$ can be defined as the maximum difference in their respective corresponding elements:

$$d(X(i), X(j)) = max_{k=1,2,..m}(|x(i+k-1) - x(j+k-1)|) \quad (2)$$

where $i = 1,2...,N-m+1$, $j = 1,2...,N-m+1$, $N$ is the number of data points contained in the time series.

For each vectors $X(i)$, a measurement that describes the similarity between the vectors $X(i)$ and all other vectors $X(j)$ can be defined as:

$$C_i^m(r) = \frac{1}{N-(m-1)} \sum_{j \neq i} \Theta\{r - d(X(i), X(j))\} \quad (3)$$

where

$$\Theta\{x\} = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (4)$$

The symbol '$r$' represents a predetermined tolerance value, with $k$ a positive constant ($k > 0$) and $std_S$ the standard deviation of the time series $S$.

$$r = k \times std_S \quad (5)$$

A finite time series consisting of $N$ data points is used to estimate the $ApEn$ value of the time series, which is defined as:

$$ApEn(m,r,N) = \varnothing^m(r) - \varnothing^{m+1}(r) \quad (6)$$

$$\varnothing^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} ln\left[C_i^m(r)\right] \quad (7)$$

## 2.2 Complexity analysis

The complexity analysis is based on the Lempel-Ziv definition [19]. This approach transforms the analysed signal into a data sequence. To illustrate the procedure, consider a gear vibration signal with a known mean value. A new sequence ($S$) is reconstructed by comparing the value of each sample of the previous sequence within the mean value. If the value of the sample is larger, it is set to one (1), otherwise to zero (0). Therefore, only two binary symbols are present in the new data sequence. This S is subsequently scanned from its first sample to its end. When a subsequence that is not encountered in the previous scanning process is discovered, the complexity value is increased by one (1). Thus, the Lempel-Ziv complexity reflects the number of all different sub-sequences contained in the original sequence. Figure 1 (reproduced from [16]) described the algorithm. For generality sake, normalized complexity $C(n)$ is often used to obtain a measure independent of the sequence length.

$$C(n) = c(n) / b(n) \quad (8)$$

$$b(n) = N / log2(N) \quad (9)$$

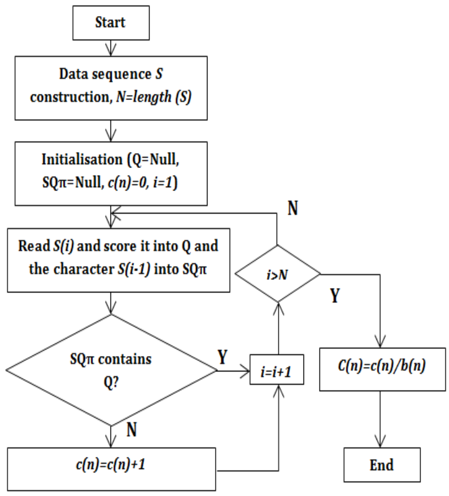


**Figure 1.** The Flow chart of $LZC$ algorithm (from [16]).

## 2.3 Parameters Selection of $ApEn$ and $LZC$ for gear Signals

From above described algorithm of $ApEn$, it can be seen that the calculated $ApEn$ value depends on two parameters, which are the embedded dimension $m$ and the tolerance $r$. However, no guideline exists for optimizing their values. In order to simulate the vibratory signals of gearbox, a gear multiplicative model whose the meshing is modulated in amplitude has been used. The gear model as defined in [20] is used:

$$\begin{aligned} x(t) = & \sum_{m=-\infty}^{+\infty} S_{r1}(t - m\tau_{r1}) + \sum_{m=-\infty}^{+\infty} S_{r2}(t - m\tau_{r2}) \\ & + \sum_{n=-\infty}^{+\infty} S_e(t - n\tau_e).\left(1 + \sum_{m=-\infty}^{+\infty} S_{r1}(t - m\tau_{r1})\right) \end{aligned} \quad (10)$$

where $\tau_e$, $\tau_{r1}$ and $\tau_{r2}$ represent the meshing period and the rotational periods. $S_e(t)$, $S_{r1}(t)$ and $S_{r2}(t)$ represents the meshing signal and its modulation.

Figure 2 and Figure 3 represent the simulated signal and its spectrum, respectively.

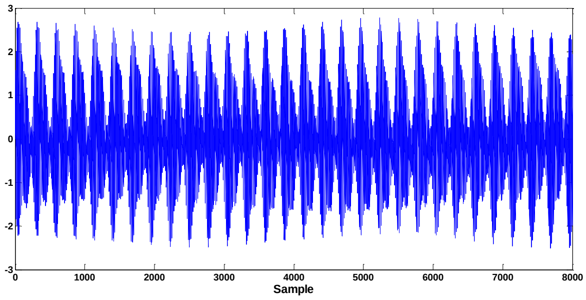

**Figure 2.** The simulated signal

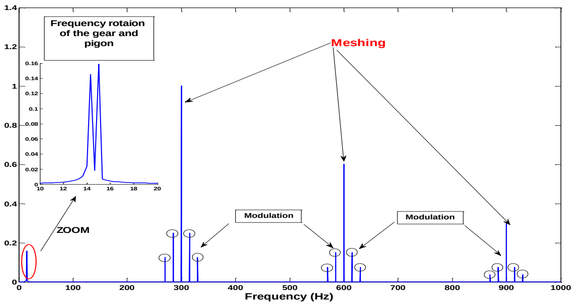


**Figure 3.** Spectrum of the simulated signal.

Figure 4 shows the variation $ApEn$ depending on $m$ and $K\ values$. It can be seen that when $m=1$, $ApEn$ has poorest convergence property. With $m$ increasing, the convergence property of $ApEn$ becomes better. However, a larger 'm' will lead to much higher computational cost. The value $m=2$ has been found as a good compromise since the convergence property of $ApEn$ is already good enough. For a small $K\ value$, one usually achieves poor conditional probability estimates, while for large values, too much detailed system information will be lost. Meanwhile, when $K\ value = 0.5$, $ApEn$ calculated by various $m$ values converge to a satisfactory state. By this investigation, $m=2$ and $K\ value = 0.5$ times the standard deviation are selected for the $ApEn$ calculation of the vibration signal in this paper.

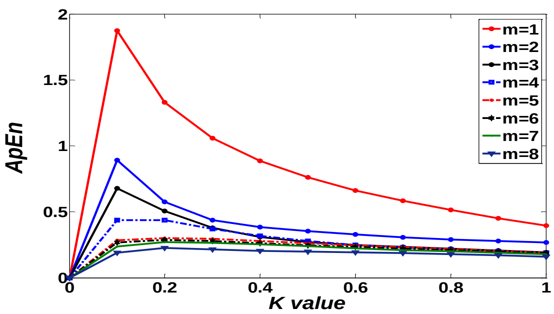


**Figure 4.** The calculated $ApEn$ values by different parameters.( $Kvalue$ and $m$ ).

The relationship between the $ApEn$, $LZC$ value and the data length is illustrated in Figure 5 and Figure 6, where seven simulated signals are comparatively displayed, under sampling rates of 4, 8, 12, 16, 20 and 24 kHz, respectively. It is seen, in both cases ( $ApEn$ and $LZC$ ), that when the data length is greater than 5000 points, the variation of $ApEn$ and $LZC$ with respect to each frequency sampling rate become insignificant. Figure 7 shows a comparison of computing time between $LZC$ and $ApEn$. It is clear that the $LZC$ is better than $ApEn$.

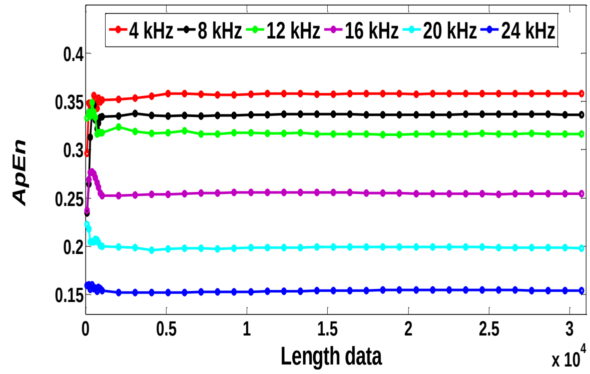


**Figure 5.** The calculated $ApEn$ values by different parameters. (Length of the data and the sample frequency).

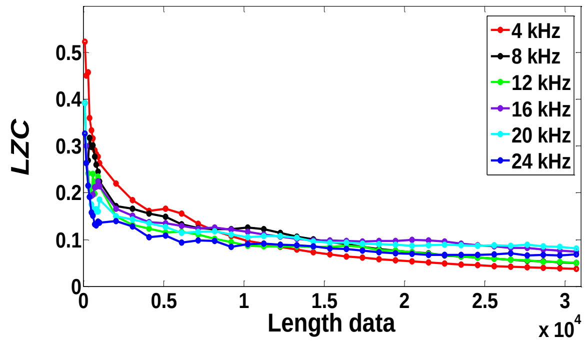


**Figure 6.** The calculated $LZC$ values by different parameters. (Length of the data and the sample frequency).

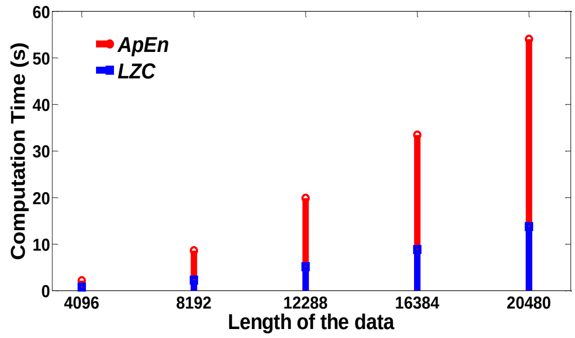


**Figure 7.** Time Computation of both $LZC$ and $ApEn$ for different length of data.

## 2.4 Influence of noise

It is well known that a white noise contains most abundant frequency components compared with other kinds of signal. If a signal is contaminated by a white noise, calculated $ApEn$ and $LZC$ values will also be. Using the simulated signal, the $ApEn$ and $LZC$ values corresponding to different SNRs were calculated, as listed in Table 1

**Table 1.** *ApEn* and *LZC* for different SNR of gear signals

| Cases | *ApEn* | *LZC* |
|---|---|---|
| The simulated signal | 0.3529 | 0.1294 |
| SNR= 80 dB | 0.3547 | 0.1364 |
| SNR=60dB | 0.3648 | 0.1692 |
| SNR=40 dB | 0.4495 | 0.3056 |
| SNR=0 dB | 1.3163 | 0.8909 |

It may be noticed that the *ApEn* and *LZC* values increase as the SNR decreases, which corresponds to a degradation of the data quality. This is analogous to the deterioration of a machine system where defects initiate and propagate through the structure. This simulation result confirms that the *ApEn* and *LZC* values provide a quantitative measurement for characterising the severity of degradation, which represents the deterioration of a machine's health condition. From Table 1, it can be seen that *ApEn* is very sensitive to large noises; however, *ApEn* is nearly unaffected by the noise which stays low. Otherwise, we can see that the *LZC* is more sensitive to noise than *ApEn* .

## 3 Experimental study

The recordings of vibration signals were carried out at CETIM, France on a gear system with a train of gearing, with a ratio of 20/21 functioning continuously until its destruction [20-22]. Table 2 gives the details of the gear test rig parameters. The test duration was 13 days with a daily mechanical appraisal; measurements were collected every 24 h except at the first day. Table 3 gives a description of the state of the gear at each 24 h. The acceleration signals for days 2, 5, 7, 9, 10 and 12 are shown in Figure 10.

**Table 2.** Geared system data

| Parameter | Pinion | Gear |
|---|---|---|
| Number of tooth | 20 | 21 |
| Speed(rpm) | 1000 | 952.38 |
| Drive torque (Nm) | 200 | |

Pareya *et al.* [21] use the same signal for their own research. Only the Kurtosis and Crest Factor were considered. The kurtosis values for the experimental signal were calculated from day 2 to 13 and are shown in Figure 8 and Figure 9 (on day 1, no signal was taken). It can be seen that the kurtosis increases greatly after the day 11. This indicates that the signal become impulsive.

The Crest Factor observes a little growth after the day 9 (2.98 to 3.60). This is due to the evolution of the spalling of the teeth 15/16 as observed in the day 9.

**Table 3.** Daily mechanical appraisal [22]

| Day | Observation |
|---|---|
| 1 | No acquisition |
| 2 | No anomaly |
| 3 | No anomaly |
| 4 | // // |
| 5 | // // |
| 6 | // // |
| 7 | Tooth spalling ½ |
| 8 | No evolution |
| 9 | Tooth 1/2 : no evolution;<br>Tooth 15/16: spalling beginning |
| 10 | Spalling evolution the teeth 15/16 |
| 11 | // // |
| 12 | // // |
| 13 | Spalling on all the width of the tooth 15/16 |

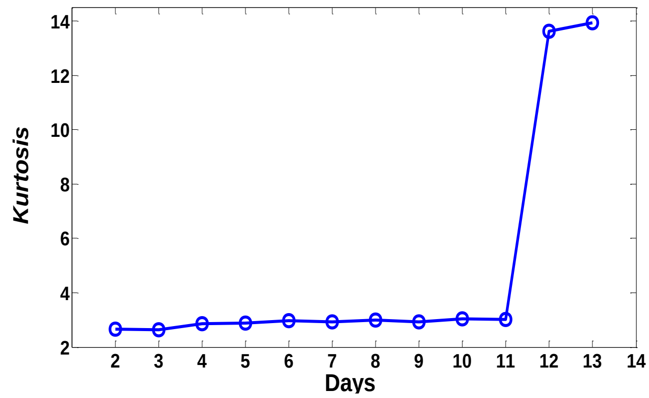


**Figure 8.** Kurtosis evolution during the 13 days [from 21]

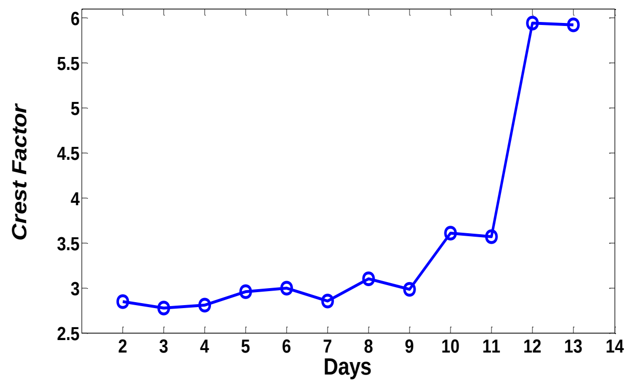


**Figure 9.** Crest Factor evolution during the 13 days [from 21]

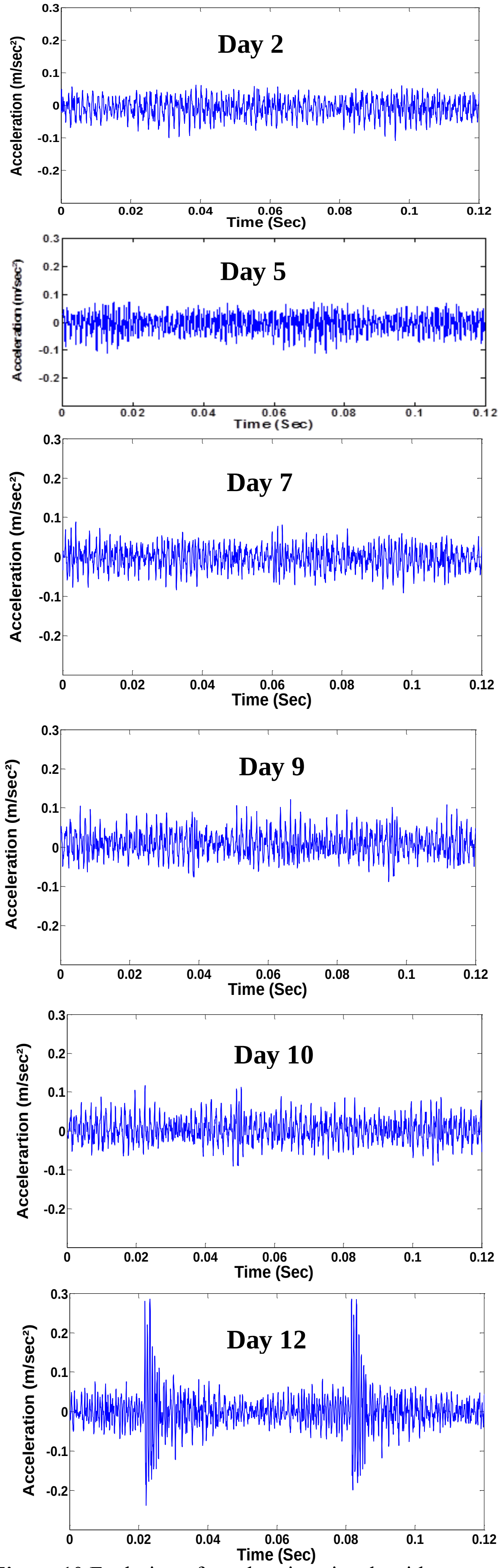

**Figure 10** Evolution of acceleration signals with wear.

The values of *ApEn* and *LZC* for all these signals are plotted in Figure 11 and Figure 12. The *ApEn* method presents a significant increase from days 5 to 11. After day 11, the *ApEn* decreases significantly. This due to the fact that the signal becomes impulsive and the SNR becomes higher as explained in section 2. On the other hand, *LZC* shows a significant increase after the day 4, revealing a clear anomaly due to gear wear.

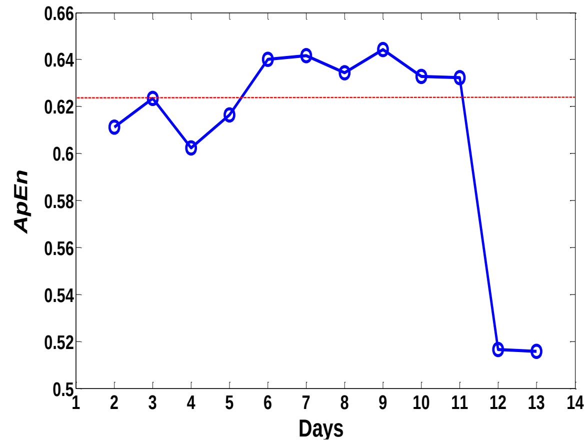

**Figure 11.** *ApEn* Crest Factor Value during the test for all days

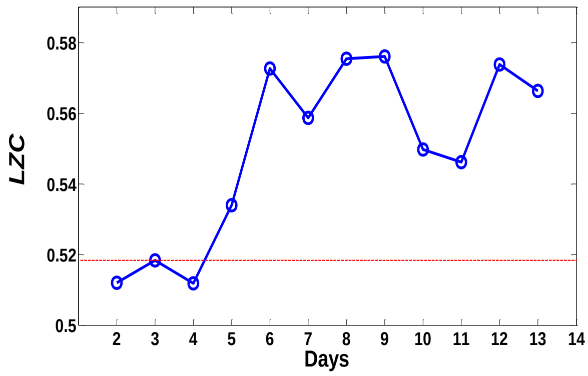

**Figure 12.** *LZC* Crest Factor Value during the test for all days

Figure 13 presents both the kurtosis and *LZC* of the day 12. The Kurtosis is reported in the abscise X and the *LZC* on Y. This representation gives both information on the impulsiveness of the signal and on effect of the number of frequency components and noise into the signal.

It can be seen that this plot divides the twelve days into three regions. The first region contains the three first days. The second enclose all days from 6 to 11. The second region is separated from the first by the day 5. The day 5 present the change in the characteristics of the gear signal.

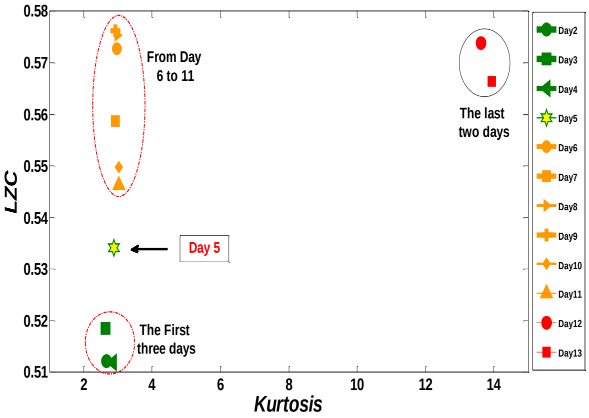


**Figure 13.** Kurtosis versus LZC Value during the test for all days.

Figure 14 compares the spectrum evolution between days 4 and 5. An increase of the amplitude at day 5 may be clearly noticed as compared with the day 4. Effectively, this day presents the beginning of the degradation due to wear. The last region is marked by an increase of the Kurtosis and the *LZC* stays in the same level as the second region. At this stage, the signal becomes impulsive and the gear is damaged

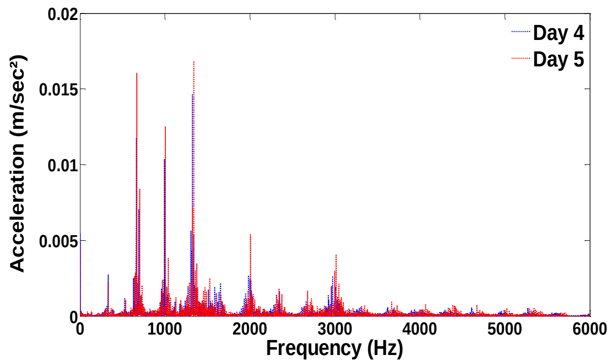


**Figure 14.** Spectrum of the experimental signal.

According to Table 3, the spalling has only visually been observed after the day 6. According to *LZC* measurements, the beginning of the chipping was in fact initiated at the day 5. This initiation is characterised by the growth of the frequency components related to the meshing. Consequently, this representation combining Kurtosis and *LZC* may be used as an efficient tool for early detection of faults

## 4 Conclusion

This paper introduces *ApEn* and *LZC* metrics to analyse the vibration signals recorded from defected gears. With respect to gear signals, the parameter selection of *ApEn* was investigated and the results show that $m = 2$ and $K\ value = 0.5$ times the standard deviation may be considered as suitable and a good compromise for the *ApEn* calculation. In addition, the influence of white noise on the *ApEn* and *LZC* calculation was also investigated. The results show that *ApEn* is nearly unaffected by the noise when staying at a small level. However, *ApEn* is very sensitive to the noise at high levels. The *LZC* is more sensitive to noise as compared with *ApEn*. An experimental study was conducted to evaluate the effectiveness of these parameters. The results show that the *LZC* and *ApEn* can detect the defect of the gears earlier than the classical temporal indicators. However the *LZC* appeared more efficient than *ApEn*. A new representation in a plan (Kurtosis, *LZC*) is proposed as a new tool for effectively monitoring gear defects. Although experimental results look promising, the proposed vibration methodology has to be tested on other test rigs. Research is thus being continued to analyze vibration signals from different defect types and on different types of gears and to validate the efficiency of this technique.

## Acknowledgments

The financial support of NSERC (Natural Sciences and Engineering Research Council of Canada), FQRNT (Fonds Québecois de la Recherche sur la Nature et les Technologies), MITACS Canada, Pratt & Whitney Canada and CETIM which provided the experimental results are gratefully acknowledged.

## References

1. Wang WJ and McFadden PD. Early detection of gear failure by vibration analysis –I. Calculation of the time-frequency distribution. MSSP , 7(3): 193–203. (1993).
2. Wang WJ and McFadden PD. Early detection of gear failure by vibration analysis –II. Interpretation of the time–frequency distribution using image processing techniques. MSSP; 7(3): 205–15 ; (1993).
3. Safizadeh M.S., Lakis A.A. and Thomas M., Gear Fault Diagnosis using time-frequency methods, proceedings of 20th seminar on machinery vibration, Canadian Machinery Vibration Association, Quebec, 7.19-7.29, (2002);.
4. Safizadeh M.S., Lakis A.A. and Thomas M., Using Short-Time Fourier Transform in Machinery diagnosis, Proceedings of WSEAS (Brazil) 494-200, (2005).
5. Z.K. Peng, F.L. Chu, Application of the wavelet transform in machine condition monitoring and fault diagnostics: a review with bibliography, Mechanical Systems and Signal Processing 18 199–221, (2004).
6. R. Yan, R. X. Gao and X. Chen, Wavelets for fault diagnosis of rotary machines: A review with applications, Signal Processing 961–15, (2014).
7. Baydar N and Ball A. Detection of gear deterioration under varying load conditions by using the instantaneous power spectrum. Mechanical Systems and Signal Processing ;14(6): 907–21, (2000).
8. Baydar N, Chen Q, Ball A, Kruger U. Detection of incipient tooth defect in helical gears using multivariate statistics. MSSP 15(2): 303–21, (2001).

9. Baydar N., Ball A. A comparative study of acoustic and vibration signals in detection of gear failures using Wigner–Ville distribution. Mechanical Systems and Signal Processing 15(6): 1091–1107, (2001).
10. Baydar N. and Ball A. Detection of gear failures via vibration and acoustic signal using wavelet transform. Mechanical Systems and Signal Processing,17(4): 787–804, (2003).
11. Yesilyurt I., The application of the conditional moments analysis to gearbox fault detection—a comparative study using the spectrogram and scalogram, NDT&E International 37,309–320, (2004).
12. Yan, R., and Gao, R. X., “Approximate Entropy as a Diagnostic Tool for Machine Health Monitoring,” Mechanical Systems and Signal Processing. 21, 824–839, (2007).
13. He Y. and X. Zhang. Approximate Entropy Analysis of the Acoustic Emission From Defects in Rolling Element Bearings, Journal of Vibration and Acoustics, , Vol. 134 / 061012-1, (2012).
14. Fu, L., He, Z. Y., Mai, R. K., and Qian, Q. Q., Application of Approximate Entropy to Fault Signal Analysis in Electric Power System, Proceedings of the Chinese Society of Electric Engineering, 28(28), 68–73, (2008 ).
15. Xu, Y. G., Li, L. J., and He, Z. J. Approximate Entropy and its Applications in Mechanical Fault Diagnosis, Chin. J. Inf. Control, 31(6), 547–551, (2002).
16. R. Yan and R. X. Gao, Complexity as a Measure for Machine Health Evaluation, IEEE transactions on instrumentation and measurement,53 (4),1327-1334, (2004).
17. Wang J., L. Cui, H. Wang and P. Chen , Improved Complexity Based on Time-Frequency Analysis in Bearing Quantitative Diagnosis, Advances in Mechanical Engineering, Article ID 258506, 11 pages http://dx.doi.org/10.1155/2013/258506, (2013)
18. Kedadouche M., Kidar T., Thomas M. and Tahan A. , Combining EMD and Lempel-Ziv Complexity for early detection of gear cracks, Surveillance 7, Chartres, France, pp 100-110, 28-29, (2013).
19. Lempel A. and Ziv J., On the complexity of finite sequences, IEEE Trans. Inform. Theory,Jan. Vol. IT22, 75–81, (1976).
20. El Badaoui, M., Contribution of vibratory diagnostic of gearbox by Cepstral analysis, Ph.D. thesis, Jean Monnet University of St Etienne (FR), p. 141 (1999) (in French).
21. Pareya A., El Badaoui M., Guillet F., Tandon N.. Dynamic modelling of spur gear pair and application of empirical mode decomposition-based statistical analysis for early detection of localized tooth defect, Journal of Sound and Vibration 294, 547–561, (2006).
22. Haloui N., Chikouche D. and Benidir M., Oct 2007, Diagnosis of gear systems by spectral analysis of vibration signals. International Journal of Computer Science and Network security, Vol 7 (1), 285-293