

DETECTION AND DIAGNOSIS OF SYSTEM NONLINEARITIES USING HIGHER ORDER STATISTICS

M. A. A. S. Choudhury* S. L. Shah ^{*,1}
N. F. Thornhill**

* *Department of Chemical and Materials Engineering
University of Alberta, Edmonton AB, Canada, T6G 2G6*

** *Department of Electronic and Electrical Engineering
University College, London, UK WC1E 7JE*

Abstract: This paper is concerned with the statistical analysis of closed loop data for diagnosing the causes of poor control loop performance. Higher Order Statistical (HOS) techniques have been developed over the last two decades, but until now have not been applied to the area of process monitoring. The main contribution of this work is to utilize the higher order statistical tools such as cumulants and their frequency domain counterparts (bispectrum, bicoherence, trispectrum) to detect and quantify the non-Gaussianity and nonlinearity of regulated processes or control error variables which are sometimes the main contributors to the poor performance of many of the control loops. The bicoherence index together with the process and manipulated variable plots are used to diagnose the sources of system nonlinearities. Successful application of the proposed method is demonstrated on simulated as well as industrial data.

Keywords: stiction, bispectrum, cumulants, bicoherence, control loop performance

1. INTRODUCTION

The field of controller performance monitoring has received much attention in the engineering research literature. However, the diagnosis of poor performance remains an open area. Performance diagnosis requires identification of the causes of poor performance, as for example due to process problems, poor controller tunings, presence of disturbances, process and/or actuator nonlinearities. If there are some non-linearities in the control loop, the controller may not perform at the desired level. Non-linearities degrade the performance of the controller in several ways, for example, they may produce oscillations in process variables, decrease the life of the control valve, may upset process stability, and in most cases

lead to inferior quality end-products thus causing larger rejection rates and reduced profitability. The non-linearities may be present in the process itself or in the actuators or control valves. This study is concerned with both types of nonlinearities, which are defined here as, "system nonlinearities". Process nonlinearity means that the process can not be modelled by a linear transfer function. Actuator or valve non-linearities are typically due to stiction, backlash, saturation, deadzone, ruptured diaphragm, corroded or eroded valve seats.

Classical signal processing tools utilize only the first and second order moments, i.e., the mean and variance. Such tools are mainly useful for analyzing signals from linear processes. The implicit assumption in using these tools is that the distribution of the data is normal. In case of non-linear signals, one needs to look at other

¹ author to whom all correspondence should be addressed

methods of characterizing their statistical properties. This necessitates the use of higher order statistical tools. The third and fourth order moments or cumulants and their frequency domain counterparts are found to be useful in analyzing non-linearities in communication signals and mechanical machine condition monitoring (Nikias and Petropulu, 1993), (W. B. Collis and Hammond, 1998), (Rivola and White, 1999). In this work, these higher order statistical techniques are used to detect and diagnose actuator or valve non-linearities. The method uses only routine operating data such as process variables (*pv*), controller outputs (*op*), and set points (*sp*). Since the method is based on regulatory control, it does not require any additional process excitation.

2. HIGHER ORDER STATISTICS

The first and second order statistics (e.g., mean, variance, autocorrelation, power spectrum) are popular signal processing tools and have been used extensively for analysis of process data. But such second order statistics are only sufficient for describing linear and Gaussian processes. In practice, there are many situations when the process deviates from Gaussianity and linearity, e.g., when it exhibits nonlinear behavior. These type of processes can conveniently be studied using Higher Order Statistics (HOS). There are three main reasons for using HOS: to extract information due to deviations from Gaussianity, to recover the true phase character of the signals, and to detect and quantify nonlinearities in time series (Nikias and Petropulu, 1993). The following sections briefly describe some of the HOS terms.

2.1 Cumulants

For any random variable x , the moment generating function can be defined as the expectation of the transformation, e^{tx} , where $t \in R$, i.e.,

$$M_x(t) = E[e^{tx}] \quad (1)$$

Moments can be obtained from the coefficients of the Taylor's series expansion of the moment generating function about the origin. In a similar fashion, cumulants can be derived from the coefficients of the Taylor's series expansion of the cumulant generating function, $C_x(t)$, which is defined as the logarithmic of the moment generating function:

$$C_x(t) \triangleq \ln(M_x(t)) \quad (2)$$

Cumulants are another set of statistical measures which can be used instead of moments because of their excellent noise suppressing properties. Moments and cumulants are very closely related to each other. For example, the following relationships hold:

$$\begin{aligned} c_1 &= m_1 \\ c_2 &= m_2 - m_1^2 \\ c_3 &= m_3 - 3m_2m_1 + 2m_1^3 \\ c_4 &= m_4 - 4m_3m_1 - 3m_2^2 + 12m_2m_1^2 - 6m_1^4 \end{aligned} \quad (3)$$

where, m_i is the i^{th} order moment defined as $m_i = E[x^i]$ and c_i is the i^{th} order cumulant. A detailed description of this derivation is given in (Choudhury, 2001).

2.2 Bispectrum and Bicoherence

Not all the information content of a signal may be easily obtained from statistical analysis of the data in the time domain. Transforming the signal from time to frequency domain can expose the periodicities of the signal and can also aid in understanding the signal generating process.

Just as the power spectrum is the frequency domain representation of the second order moment, the bispectrum is the frequency domain counterpart of the third order cumulants. The bispectrum is defined as

$$\begin{aligned} B(f_1, f_2) &\triangleq DDFT[c_3(\tau_1, \tau_2)] \\ &\equiv E[X(f_1)X(f_2)X^*(f_1 + f_2)] \end{aligned} \quad (4)$$

where, $X(f)$ is the Fourier transformation of the data series $x(t)$ and DDFT stands for double discrete Fourier transformation. Equation 4 shows that it is a complex quantity having both magnitude and phase. The bispectrum can be plotted against two independent frequency variables, f_1 and f_2 in a three dimensional plot. Each point in the plot represents the bispectral content of the signal at the bifrequency, (f_1, f_2) . In fact, the bispectrum at point $(B(f_1, f_2), f_1, f_2)$ measures the interaction between frequencies f_1 and f_2 . This interaction between frequencies can be related to the non-linearities present in the signal generating systems and therein lies the core of its usefulness in the detection and diagnosis of non-linearities.

It can be shown that the bispectral estimates are asymptotically unbiased and the variance of the estimator depends on the second order spectral properties (Hinich, 1982). That is,

$$var(\hat{B}(f_1, f_2)) \propto P(f_1)P(f_2)P(f_1 + f_2) \quad (5)$$

where, $P(f)$ is the power of the signal at frequency, f . Since the estimate depends directly on the bifrequency, the variance of the estimate will be higher at a bifrequency where the signal energy is high and will be lower where the energy is low. This causes serious problems in its estimation. In this work, the bispectrum is normalized in a such way that it gives a measure whose variance is independent of the signal energy. This is termed

as bicoherence and is defined by the following equation :

$$bic^2(f_1, f_2) \triangleq \frac{|E[B(f_1, f_2)]|^2}{E[|X(f_1)X(f_2)|^2]E[|X(f_1 + f_2)|^2]} \quad (6)$$

where ‘bic’ is known as the bicoherence function. It has been shown by (Kim and Powers, 1979) that the variance of the bicoherence estimator satisfies the following expression:

$$var[\hat{bic}^2(f_1, f_2)] \approx \frac{1}{M}[1 - bic^2(f_1, f_2)] \quad (7)$$

where M is defined as the number of segments used in the estimation. A useful feature of bicoherence function is that it is bounded between 0 and 1. There are also other normalization methods but they are not popular because their properties have not been extensively studied (Fackrell, 1996).

2.3 Test of Linearity Based on the Bispectrum

Let $x(k)$ be a discrete stationary time series. We call it linear, if it can be represented by

$$x(k) = \sum_n h(n)e(k - n) \quad (8)$$

where, $e(k)$ is a sequence of independent identically distributed random variables with $E[e(k)] = 0$, $\sigma_e^2 = E[e^2(k)]$, and $\mu_3 = E[e^3(k)]$. For this case, the following frequency domain relationships can be obtained: the power spectrum,

$$P(f) = \frac{\sigma_e^2}{2\pi} |H(f)|^2 \quad (9)$$

and the bispectrum,

$$B(f_1, f_2) = \frac{\mu_3}{(2\pi)^2} H(f_1)H(f_2)H^*(f_1 + f_2) \quad (10)$$

where, $H(f) = \sum_n h(n)e^{-inf}$. Equation 6 can be rewritten as

$$\begin{aligned} bic^2(f_1, f_2) &\triangleq \frac{|E[B(f_1, f_2)]|^2}{E[|X(f_1)X^*(f_1)||X(f_2)X^*(f_2)|]} \\ &\quad \frac{E[|X(f_1 + f_2)X^*(f_1 + f_2)|]}{|E[B(f_1, f_2)]|^2} \\ &\equiv \frac{|E[B(f_1, f_2)]|^2}{E[|P(f_1)||P(f_2)||E[|P(f_1 + f_2)|]} \end{aligned} \quad (11)$$

Now substituting the expressions from equation 9 and 10, it can be shown that

$$bic^2(f_1, f_2) = \frac{\mu_3^2}{2\pi\sigma_e^6} \quad (12)$$

From this equation, it can be seen that the squared bicoherence is constant for a linear process. It is also independent of the frequency. So, if the bicoherence is not a non-zero constant in the principal domain of the bispectrum, it is concluded that the process is nonlinear. If μ_3 is zero, then the squared bicoherence is also zero. In that case, the signal is Gaussian.

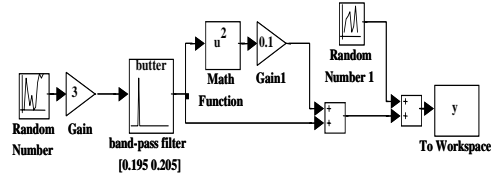


Fig. 1. Block diagram of mixed signal system

3. ILLUSTRATIVE EXAMPLES

3.1 Bicoherence of a mixed signal

A mixed signal was generated using the simulink block diagram shown in figure 1. The input signal was white noise with zero mean and unit variance. The amplified signal was passed through a bandpass filter with the frequency band [0.195 to 0.205]. The filtered signal was squared in order to introduce a small nonlinearity. Only 1/10 of it was added to the filtered signal to produce a mixed signal. White noise with variance 0.001 was added to the mixed signal to produce the final signal of interest.

This analysis is based on 4096 samples with FFT segment length of 128. The objective in this example is to demonstrate the power of the bicoherence in the detection of nonlinearity. First, a linear signal, y_{linear} , was generated without introducing any nonlinearity in figure 1, i.e., the squaring block was absent. Thereafter a non-linear signal, $y_{nonlinear}$, was generated using figure 1. By merely looking at the time trend of the signals (figures 2a and 2b), it is not possible to distinguish the two signals. Also, the power spectrums (figures 2c and 2d) or the second order moments look alike and are unable to detect the non-linearity present in the second signal. Figures 2e and 2f show the three dimensional bicoherence plot. For y_{linear} , the maximum bicoherence value is 0.08 while for $y_{nonlinear}$ it is 0.54. A large bicoherence value is an indication of significant nonlinearity. From the bicoherence plot, the peak position in the nonlinear case is at (0.1,0.1) Hz bifrequency. This means that the nonlinearity in the signal is due to interaction of these two frequencies. Going back to the signal generating system shows that the band pass filtered signal has the frequency range [0.195 to 0.205]. This signal was squared to introduce a nonlinearity. This is due to the multiplication of two signals, each of them having a frequency of approximately 0.2 Hz. In the bicoherence contour plot the frequencies identified are 0.1 and 0.1 Hz. This is because in this case the frequency is normalized with the Nyquist frequency scale which is one-half of the sampled frequency. Therefore, it correctly identifies the nonlinear frequency interactions. The frequency interactions are the result of nonlinearity present in the signal due to the system or the process.

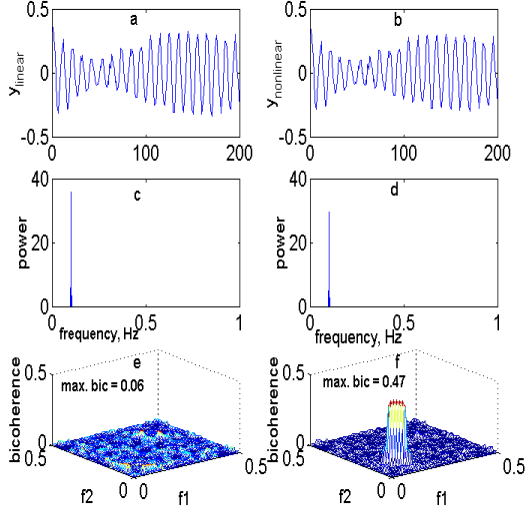


Fig. 2. Results for y_{linear} (left) and $y_{nonlinear}$ (right)

3.2 Bicoherence of a nonlinear sinusoid signal with noise

For this example, an input signal was constructed adding two sinusoids, each of them having a different frequency and phase. That is,

$$x'(k) = \sin(2\pi f_1 k + \pi/3) + \sin(2\pi f_2 k + \pi/8) \quad (13)$$

$$x(k) = x'(k) + d(k) \quad (14)$$

where, $f_1 = 0.12$, $f_2 = 0.30$, $k = 1$ to 4096 and $d(k)$ is a white noise sequence with a standard deviation of 0.2 .

Figures 3a and 3b show the time series while figures 3c and 3d show the power spectrum of the signal x and y , respectively. Neither of these plots helps in distinguishing the two signals. However, the use of higher order statistics can successfully detect the nonlinearities present in y . Figures 3e and 3f show the three dimensional bicoherence plot of x and y respectively. For the linear x signal, the maximum bicoherence value is 0.05 and there are no sharp peaks in the bicoherence plots. On the other hand, the bicoherence plot of y shows two large peaks and three small peaks with a maximum bicoherence value of 0.47. These values are not constant at all bifrequencies. This clearly identifies the nonlinearities present in this signal. This example illustrates that the bicoherence index is very sensitive to even a very small amount of nonlinearity (only 5% of the quadratic nonlinearity in y). The interpretation of the peaks in the bifrequency plane can be understood by rewriting the expression for y using some trigonometric identities:

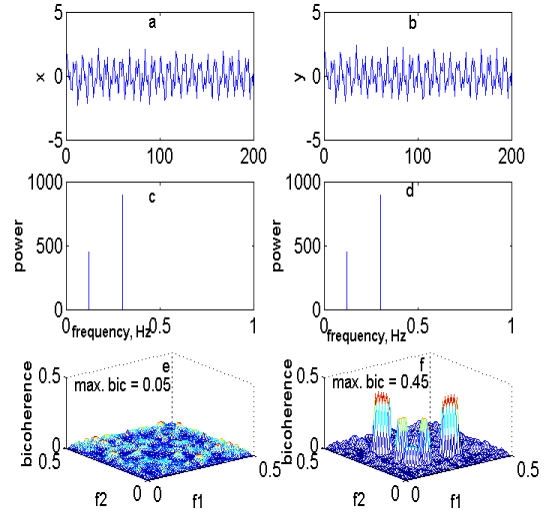


Fig. 3. Diagnostic plots of the nonlinear sinusoid signal with noise

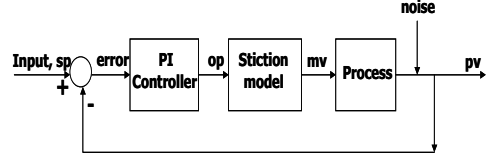


Fig. 4. Block diagram of a simple SISO process with stiction nonlinearity in the valve

$$y(k) = \sin(2\pi f_1 k + \pi/3) + \sin(2\pi f_2 k + \pi/8) + 0.05(1 - \cos(2(2\pi f_1 k + \pi/3)) - \cos(2(2\pi f_2 k + \pi/8))) + 0.05 \cos(2\pi(f_1 - f_2)k + \pi/3 - \pi/8) - 0.05 \cos(2\pi(f_1 + f_2)k + \pi/3 + \pi/8) + d(k) \quad (15)$$

The nonlinearities may be caused by the interactions of any two of the signals with frequencies f_1 , f_2 , $2f_1$, $2f_2$, $f_1 - f_2$, and $f_1 + f_2$. For the output signal y , the bicoherence plots show peaks at $(0.12, 0.12)$, $(0.18, 0.12)$, $(0.30, 0.30)$, $(0.30, 0.12)$, and $(0.4, 0.3)$ Hz bifrequencies. These frequencies correspond to (f_1, f_1) , $(f_1 - f_2, f_1)$, (f_2, f_2) , (f_1, f_2) , and $(2f_2, f_2)$, respectively. The bicoherence plot correctly identifies frequency interactions which result in nonlinearities in this signal.

4. EXAMPLES TO ILLUSTRATE THE DETECTION OF VALVE NONLINEARITIES

4.1 Simulation results

The simple single input single output (SISO) system given in figure 4 has been used for generating simulated data. The second order process with time delay is given by the following transfer function:

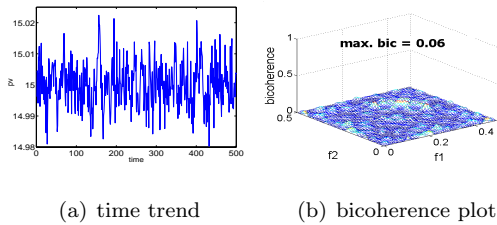


Fig. 5. Analysis of simulated data without stiction

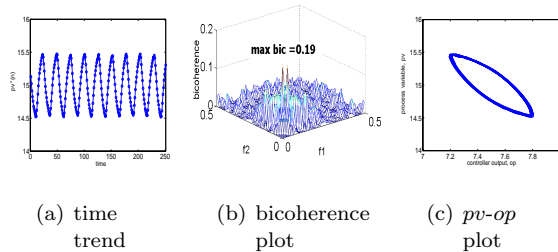


Fig. 6. Analysis of simulated data with 3% stiction

$$\frac{2}{(3s+1)(10s+1)}e^{-3s} \quad (16)$$

The controller is a PI with $K_c = 0.5$ and $\tau_i = 12.3$. Random noise with variance 0.01% was filtered with a low pass filter and added to the process variable. The simulation was performed for 10000 seconds using varying amount of stiction in the model. For a detailed description of the stiction model please refer to (Choudhury, 2001). The raw data was filtered with a low pass filter to give the filtered $pv'(n)$:

$$pv'(n) = \frac{pv(n) + pv(n+1)}{2} \quad (17)$$

Then, every other sample was discarded to yield the subsampled $pv''(n)$. Finally, in order to avoid the effect of transients the last 4096 steady state data points were used for analysis. The resulting process variable (pv'') was analyzed to detect the nonlinearity present in system.

4.1.1. Without stiction For this case, the non-linear ‘stiction model’ block was removed from the simulation block diagram. Figure 5 shows the time trend of the process variable (pv) and the corresponding bicoherence plot. The maximum bicoherence value for this case was 0.06, which is very small, confirming the absence of nonlinearities in the system.

4.1.2. Presence of stiction Figure 6 shows the time series of $pv''(n)$, the bicoherence plot and the pv - op plot. The presence of 3% stiction produces very distinct limit cycles in the process variables. The maximum value of the bicoherence, 0.19, for this case clearly detects the presence of nonlinearity in the process signal. For real cases, this maximum bicoherence value is expected to be higher. For symmetric distributions, the bicoherence value is usually low because it takes into

account the skewness of the distribution. In addition to the bicoherence index, the process variable versus controller output plot, i.e., pv - op plot is useful in identifying the type of nonlinearity. Figure 6(c) shows this characteristic plot. The presence of distinct loops with sharp turn around points is an indication of stiction in the valve.

4.2 Case Study of Industrial Data

The higher order statistical techniques have been successfully applied in detection and isolation of process/actuator faults for some of the control loops of an industrial unit of a chemical complex in Edmonton, Canada. Data was collected with a sampling interval of 1 min over two time periods: from April 10 to 17, 2001 and following the annual maintenance shutdown of the plant, from July 1 to 15, 2001. Results of the analysis of two loops are discussed below:

Flow loop 1: The first one is a recycle flow control loop. The detailed diagnostic plots are shown in figure 7. Time series of the data collected in April and July are shown in figures 7a and 7b. The op trend in figure 7c shows that the valve movement was very slow and insignificant compared to the change in the error signal, (pv - sp). The bicoherence plot in figure 7e has a maximum bicoherence value of 0.17 and it has three peaks in the non-redundant region. This indicates the presence of nonlinearity in the error signal. The op and pv time trend in figure 7a and 7c show that a little change in op caused a big change in pv value (note the range of y-axis for op , 49.4 to 50). Therefore it was suggested that this valve had nonlinearities and it was most likely due to an oversized valve. This 6 inch valve was replaced by a 3 inch valve during the annual maintenance shutdown of the plant (in May, 2001). In order to confirm the result of the analysis, additional data was collected in July and the results of the ‘post-maintenance’ data analysis are shown in the right half part of the figure 7. The maximum bicoherence value is reduced to 0.06 indicating linear system characteristics.

Flow loop 2: The second loop is also a flow control loop at the outlet of a pump located at the bottom of a distillation column. Analysis of the April, 2001 data of this loop revealed that this loop had severe nonlinearity problems with a maximum bicoherence value of 0.772. The diagnostic plots are shown in the left side of the figure 8. The 3D bicoherence plot correctly detected the presence of a significant nonlinearity. The pv - op characteristic plot indicated a type of nonlinear characteristic in the process or the valve that had not been observed before. During

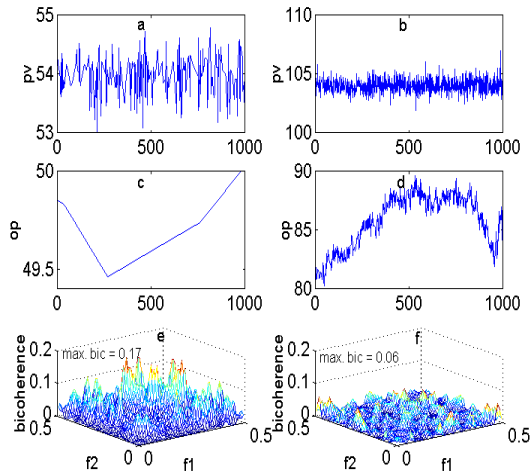


Fig. 7. Results for flow loop 1, April (left) and July (right)

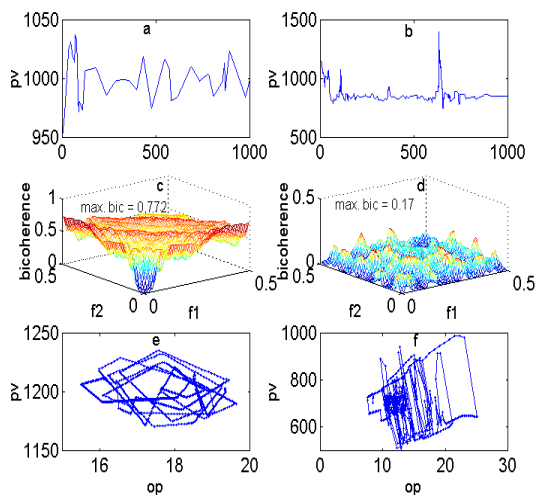


Fig. 8. Results for flow loop 2, April (left) and July (right)

the annual maintenance, the plant instrument personnel noticed that the valve seat was severely corroded. This valve was also replaced. The results of the ‘post-maintenance’ analysis are shown in the right side of figure 8. The bicoherence index now is 0.17, still indicating the presence of some nonlinearities, but reduced substantially from the previous value of 0.772. The $pv - op$ plot still shows some unfamiliar patterns for unknown sources of process nonlinearities.

5. CONCLUDING REMARKS

Higher order statistical measures such as cumulants and bicoherence are used in this paper to quantify the nonlinearities present in the regulated control error signals, and thus indicate the presence or absence of system nonlinearity. The

bicoherence magnitude threshold limit for nonlinearity detection is chosen as 0.1, which is an ad hoc value and is based on the experience of using this tool in process performance diagnosis. However, the choice of the appropriate threshold limits (statistical based) to detect and raise the alarm for process or actuator nonlinearities is an area that is currently under investigation. Additional diagnosis tools such as $pv - op$ plots are combined with the bicoherence index to diagnose the potential source of the nonlinearity. The method is evaluated by successful application to two industrial data sets in which the presence of valve problems were confirmed and resolved during the annual plant maintenance shutdown.

6. ACKNOWLEDGEMENTS

Financial support to the first author, in the form of a CIDA scholarship from the Canadian government, is gratefully acknowledged. The project has also been supported by the Natural Sciences and Engineering Research Council of Canada (NSERC), Matrikon Inc., and the Alberta Science and Research Authority (ASRA) in the form of an NSERC-Matrikon-ASRA Industrial Research Chair Program at the University of Alberta.

7. REFERENCES

- Choudhury, M. A. A. S. (2001). Detection and diagnosis of system nonlinearities using higher order statistics. Candidacy report, Department of Chemical and Materials Engineering, U of A, Canada.
- Fackrell, Justin W. A. (1996). Bispectral Analysis of Speech Signals. PhD thesis. The University of Edinburgh. UK.
- Hinich, Melvin J. (1982). Testing for gaussianity and linearity of a stationary time series. *Journal of time series analysis* **3**, 169–176.
- Kim, Y. C. and E. J. Powers (1979). Digital bispectral analysis and its applications to nonlinear wave interactions. *IEEE Transactions on Plasma Science* **PS-7**, 120–131.
- Nikias and A. Petropulu (1993). *Higher-Order Spectra: A nonlinear signal processing framework*. Prentice Hall. New Jersey.
- Rivola, A. and P. R. White (1999). Use of higher order spectra in condition monitoring: Simulation and experiments. In: *Proceedings of the DETC99*.
- W. B. Collis, P.R. White and J. K. Hammond (1998). Higher-order spectra: The bispectrum and trispectrum. *Mechanical Systems and Signal Processing* **12**, 375–394.