Brief paper

Detecting and isolating multiple plant-wide oscillations via spectral independent component analysis

Chunming Xia, John Howell, Nina F. Thornhill

Center for Mechatronics Engineering, East China University of Science & Technology, Shanghai 200237, China
Department of Mechanical Engineering, University of Glasgow, Glasgow G12 8QQ, Scotland, UK
Department of Electronic and Electrical Engineering, University College London, Torrington Place, London WC1E 7JE, UK

Received 10 January 2003; received in revised form 12 January 2005; accepted 16 February 2005
Available online 3 October 2005

Abstract

Disturbances that propagate throughout a plant can have an impact on product quality and running costs. There is thus a motivation for the automated detection of plant-wide disturbances and for the isolation of the sources. A new application of independent component analysis (ICA), multi-resolution spectral ICA, is proposed to detect and isolate the sources of multiple oscillations in a chemical process. Its key feature is that it extracts dominant spectrum-like independent components each of which has a narrow-band peak that captures the behaviour of one of the oscillation sources. Additionally, a significance index is presented that links the sources to specific plant measurements in order to facilitate the isolation of the sources of the oscillations. A case study is presented that demonstrates the ability of spectral ICA to detect and isolate multiple dominant oscillations in different frequency ranges in a large data set from an industrial chemical process.

Keywords: Chemical industry; Fault diagnosis; Independent component analysis; Multivariate analysis; Oscillation; Plant-wide disturbance; Power spectrum; Principal component analysis; Process control; Spectral analysis

1. Introduction

It is important for process control engineers to detect and diagnose plant-wide oscillations in a chemical process operation (Paulonis & Cox, 2003; Qin, 1998). Oscillations increase variability and can prevent a plant from operating close to optimal constraints so a key requirement of an industrial control engineer is for an automated means for (a) the detection of the presence of a plant-wide oscillation and (b) to give an indication of where to seek the root cause so that maintenance effort can be directed efficiently.

Thornhill, Shah, Huang, and Vishnubhotla (2002) have demonstrated that principal component analysis of power spectra (spectral PCA) provides a means of detecting the presence of a plant-wide oscillation. However, little has been done to isolate the sources of such multiple oscillations. Xia and Howell (2005) proposed the first independent component analysis (ICA) approach to the isolation of multiple oscillations. This paper extends the spectral ICA method in that it: (1) reconstructs the spectral ICA and interpretation procedures; (2) introduces the novel multi-resolution spectral ICA; (3) provides scientific insights that underpin spectral ICA, and (4) broadens the scope of process data analysis because it describes a complete and robust procedure to isolate multiple oscillatory sources.

ICA is a statistical and computational technique for extracting independent sources given only observed data that are mixtures of the unknown sources. What distinguishes ICA from other methods is that it looks for components that are both statistically independent and non-Gaussian. The ICA technique was first introduced in the early 1980s, and several research groups presented algorithms in the mid-1990s. Reported applications of ICA include biochemical signal processing, audio signal separation, telecommunications and feature extraction (Comon, 1994; Delfosse & Loubaton, 1995; Hyvärinen, Karhunen, & Oja, 2001; Hyvärinen & Oja, 2000; Jutten & Herault, 1991; Lee, 1998). Li and Wang (2002) gave the first
example of an application of ICA to the analysis of dynamical chemical process trends and dimension reduction. Kano, Hasebe, Hashimoto, and Ohno (2004), and Lee, Yoo, and Lee (2004) have shown the superior performance of multivariate statistical process control based upon ICA.

Multi-resolution analysis is often needed when several different oscillations are present simultaneously. As observed by Alsberg (1999), multivariate analysis usually uses the full-resolution data set which contains features at all scales and thus all available details are included in the analysis. A multi-resolution analysis focuses on different time scales in the data set in a structured manner in order to detect similarities or differences at that scale which might be swamped by signals at other scales in the full-resolution analysis. Bakshi (1999) gave an overview of methods that enhance the extraction of information from measured data by representing the variables at multiple scales. Authors using a multi-resolution approach to multivariate data from chemical processes include Kosanovich and Piovoso (1997) who performed a PCA analysis on wavelet coefficients and Misra, Yue, Qin, and Ling (2002) who used a multi-resolution wavelet analysis for detecting and identifying process faults.

Section 2 of the paper presents the methods of spectral ICA and a new technique for multi-resolution spectral ICA that aids the detection of high-frequency oscillations of small magnitude. New insights are also given into why spectral ICA is able to find individual narrow-band spectral peaks in a data set that is characterised by multiple oscillations. Scaling of the independent components (ICs) is discussed and two new indices are defined. One of the indexes matches the source of an oscillation to a specific location in the plant and the other gives an indication of the seriousness of the oscillation in terms of its percentage power.

Multi-resolution spectral ICA is applied in Section 3 to an industrial data set in which several independent oscillations are detected and their likely sources are isolated.

2. Methods

2.1. Time-domain ICA

The instantaneous mixing model (IMM) is a basic ICA model. It has the following structure:

\[ X = AS, \]

where matrix \( X = [x_1 \ x_2 \ \ldots \ x_m] \) represents all sensor observations and \( S = [s_1 \ s_2 \ \ldots \ s_n] \) is the matrix containing all independent, non-Gaussian source outputs. Each row, is termed independent component (IC). \( A \) is the mixing matrix \( a_{ij} \) \( (i = 1, \ldots , m, j = 1, \ldots , n) \). Each sensor observation is decomposed into a linear combination of a set of ICs, i.e.,

\[ x'_i = a_{i1}s'_1 + a_{i2}s'_2 + \cdots + a_{in}s'_n, \quad i = 1, \ldots , m. \]

ICA involves the estimation of both the mixing matrix, \( A \) and the independent source matrix, \( S \). This goal is equivalent to finding a separating matrix \( W \) that satisfies

\[ \hat{S} = WX = \begin{bmatrix} w'_1 \\ w'_2 \\ \vdots \\ w'_n \end{bmatrix} X, \]

where \( \hat{S} = [\hat{s}_1 \ \hat{s}_2 \ \ldots \ \hat{s}_n] \) is the estimate of \( S \) and \( w_j \) is the \( j \)-th normalised, separating vector. One way of achieving this is to estimate the first separating vector \( w_1 \), which maximises the non-Gaussianity of \( \hat{s}_1' = w'_1X \), and then to estimate the second separating vector on the same basis, and so on.

2.2. Basic spectral ICA model

Introduction: In spectral ICA, different ICs usually contain different spectral peaks because, and as will be shown later, a component having a single spectral peak is more dominant in terms of non-Gaussianity than a component having multiple peaks. The maximisation of non-Gaussianity in spectral ICA results in the preferential extraction of single-peak, narrow-band ICs rather than multiple-peak, or wide-band components. Thus, a spectrum can be decomposed into a combination of spectrum-like and single-peak ICs by means of spectral ICA.

Benefits of spectral ICA: The extraction of dominant narrow-band peaks from a multivariate set of power spectra is superior to the extraction of oscillatory sources by time-domain ICA, because spectral ICA is invariant to time delays and phase lags, which is not the case with time-domain ICA.

Spectral analysis also allows the determination of the locations and strengths of the various frequencies and gives insight into their propagation through a plant.

Spectral ICA can be thought of as an extension or improvement of spectral PCA as proposed by Thornhill et al. (2002). The disadvantage of spectral PCA is that more than one peak may appear in the spectrum-like principal components (PCs) and different PCs can contain the same peaks, or peaks at the same position but of different sign.

Formulation: In the spectral ICA model, the rows of the data matrix, \( X \), are single-sided power spectra \( P(f) \) of the observations over a range of frequencies up to the Nyquist frequency (one-half of the sampling frequency):

\[ X = \begin{pmatrix} P_1(f_1) & \cdots & P_1(f_N) \\ \vdots & \ddots & \vdots \\ P_m(f_1) & \cdots & P_m(f_N) \end{pmatrix} \]

The power spectrum of a signal can be estimated by applying discrete Fourier transform (DFT) methods, or the averaged periodogram method (Welch, 1967) to the mean centred time trends. Each power spectrum is normalised, i.e.,

\[ \sum_{k=1}^{N} P_i(f_k) = 1, \quad i = 1, \ldots , m. \]
The task is to extract narrow band ICs by analysing the spectral signatures using spectral ICA. To achieve this, \( \mathbf{X} \) is first decomposed into \( n \) dominant ICs:

\[
\mathbf{X} = \mathbf{A}_{m \times n} \begin{bmatrix} \mathbf{y}_1' \\ \vdots \\ \mathbf{y}_n' \end{bmatrix} = \mathbf{A}_{m \times n} \mathbf{Y},
\]

where \( \mathbf{y}_j' \) is the \( j \)th IC, \( \mathbf{A} \) is an \( m \times n \) real mixing matrix and \( m \geq n \). Alternatively, this can be described in terms of a separating matrix \( \mathbf{W} \)

\[
\begin{bmatrix} \mathbf{y}_1' \\ \vdots \\ \mathbf{y}_n' \end{bmatrix} = \mathbf{WX} = \begin{bmatrix} \mathbf{w}_1' \\ \vdots \\ \mathbf{w}_n' \end{bmatrix} \mathbf{X}.
\]

ICs can be estimated by finding those vectors \( \mathbf{w}_j' \mathbf{X} \) (\( j = 1 \ldots n \)), which maximise the kurtosis of each normalised independent vector in turn

\[
kurt(y) = (1/N) \sum_{j=1}^{N} y^4 - 3((1/N) \sum_{j=1}^{N} y^2)^2 .
\]

Computer programmes that perform ICA are available in the public domain (Hurri, Gävert, Särelä, & Hyvärinen, 1998). ICA methods do not provide a means of determining the number of ICs which should be included in the model. An initial condition of the mixing matrix is also required for the maximisation process. A truncated PC model \( \mathbf{X} = \mathbf{TW} \) can be used to provide both of these: the number of dominant ICs is set to the number of dominant PCs and the initial value of the mixing matrix of the ICA model is the matrix \( \mathbf{T} \).

**Insight:** To explain why the maximisation of kurtosis can extract a narrow band IC from spectra data, the kurtosis of the power spectrum of Gaussian noise may be compared with either the power spectrum of a pure sine wave or the power spectrum of a compound sinusoidal signal composed of a fundamental and several harmonics. A proof is given in the appendix that the power spectrum of a finite sample of Gaussian noise has a kurtosis of 6, whereas the power spectrum of a pure sinusoid has a kurtosis of \( (\bar{N}/2) - 3 \) (\( \bar{N} \) is the number of DFT data points) and the power spectrum of the compound signal has a kurtosis that ranges from \( (\bar{N}/2k) - 3 \) to \( (\bar{N}/2) - 3 \) where \( k \) is the total number of sinusoidal components. If \( \bar{N} \) is chosen large enough (> 1000), the kurtosis of a spectrum with narrow-band peaks is found to be significantly greater than 6, especially when a hidden source is generating a relatively pure sinusoidal signal.

**Post-ICA algorithm:** The ICs obtained from the ICA formulation above are not unique in terms of sign and magnitude, so additional constraints are needed to arrive at a solution. A post-ICA stage is applied in which all ICs are adjusted to have positive peak values for enhanced visualisation, and the mixing matrix is also scaled so that the dominant relationships between the ICs and the spectral signatures can be easily identified.

The signs of the ICs are first manipulated so that the dominant spectrum-like peak in each component is positive, to make them resemble spectrum-like features. The algorithm to achieve the sign constraints is as follows:

1. find the sign of the maximum absolute value, i.e., the sign of the dominant peak, for every IC, \( \mathbf{y}_j' \), denoted by \( SN_j \) where \( j = 1 \ldots n \).
2. adjust the mixing matrix, \( \mathbf{A} \), and the IC matrix, \( \mathbf{Y} \):

\[
\mathbf{B} = \mathbf{A} \text{diag}(SN_1, SN_2, \ldots, SN_n),
\]

\[
\mathbf{S} = \text{diag}(SN_1, SN_2, \ldots, SN_n) \mathbf{Y},
\]

then

\[
\mathbf{X} = \mathbf{BS},
\]

where \( \text{diag}(\cdot) \) is an \( n \times n \) diagonal matrix.

A new term **significance index** has been created to underline the importance of the elements of the resulting mixing matrix. A significance index of 1 represents the strongest influence from an IC onto a power spectrum signature, whereas a smaller significance index implies that the influence is weaker. Note that there is no equivalent relationship in PCA.

The algorithm to achieve the unit upper bound of significance indices for each IC is equivalent to the adjustment of the maximum value of each column of the mixing matrix

\[
\mathbf{B} = [b_1, b_2, \ldots b_n] \]

to be unity:

1. find element of each column of \( \mathbf{B} \) with the maximum absolute value (i.e. the vector \( \infty \)-norm of each column of \( \mathbf{B} \))

\[
\lambda_j = \|b_j\|_\infty, \quad j = 1 \ldots n.
\]

2. scale the mixing matrix \( \mathbf{B} \) and the IC matrix \( \mathbf{S} \)

\[
\mathbf{A} = \mathbf{B} \text{diag}(\lambda_1^{-1}, \lambda_2^{-1}, \ldots, \lambda_n^{-1}),
\]

\[
\mathbf{C} = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \mathbf{S},
\]

\[
\mathbf{X} = \mathbf{AC}
\]

\[
= [a_1, a_2, \ldots, a_n] \begin{bmatrix} \mathbf{c}_1' \\ \vdots \\ \mathbf{c}_n' \end{bmatrix} = \begin{bmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{m,1} \end{bmatrix} \mathbf{c}_1' + \begin{bmatrix} a_{1,2} \\ a_{2,2} \\ \vdots \\ a_{m,2} \end{bmatrix} \mathbf{c}_2' + \cdots + \begin{bmatrix} a_{1,n} \\ a_{2,n} \\ \vdots \\ a_{m,n} \end{bmatrix} \mathbf{c}_n',
\]

where \( \mathbf{a}_j = [a_{1,j}, a_{2,j}, \ldots, a_{m,j}]' \), \( j = 1 \ldots n \), is the \( j \)th column of the scaled mixing matrix \( \mathbf{A} \) and \( \mathbf{c}_j' \) is the \( j \)th spectrum-like IC. It can be seen that each power spectrum or \( P_i (i = 1 \ldots m) \) in \( \mathbf{X} \) is represented by a linear combination of the ICs. The \( j \)th column of the scaled \( \mathbf{A} \) matrix relates to the \( j \)th spectrum-like independent component \( \mathbf{c}_j' \), and value \( a_{i,j} \), i.e. the significance index, represents the influence from the \( j \)th
The period of each oscillation can be estimated from the peak position of the IC. The total energy of all the power spectra in $\mathbf{X}$ can be approximated by $\|\mathbf{A}c\|_\text{sum}$, which can also be interpreted as the total energy carried by all dominant oscillatory ICs, where $\|\cdot\|_\text{sum}$ is the sum of the absolute values of all its elements. The total energy related to the $j$th IC is $\|a_j c_j\|_\text{sum}$. Each dominant IC may be associated with a ratio, termed the dominance ratio (DR), which reflects the influence or dominance of that IC. Dominance is specified as the ratio of the total energy carried by that IC over the total energy carried by all dominant oscillatory ICs:

$$\text{DR}(j) = \frac{\|a_j c_j\|_\text{sum}}{\|\mathbf{A}c\|_\text{sum}} = 100 \left\{ \frac{\|a_j c_j\|_\text{sum}}{\sum_{j=1}^n (\|a_j c_j\|_\text{sum})} \right\} \%.$$  

(9)

The larger the $\text{DR}(j)$ value, the greater the influence of the $j$th IC in the perturbed plant. An IC is said to be dominant if its $\text{DR}$ value is not too small. Thus the new method also gives the period and energy percentage, i.e., $\text{DR}$ value, for each IC.

Isolation of sources: Each dominant IC is assumed to represent a separate source. Its propagation around the plant can be found by looking at the significance indices, either in tabular form or as a plot. For the $j$th dominant IC, a large significance index value of $a_{i,j}$, in the $j$th column of the $\mathbf{A}$ matrix, represents a significant linkage between the $j$th IC and the $i$th power spectrum, and hence the $i$th process variable. In particular, a value of $a_{i,j} = 1$ suggests that the source of the $j$th IC oscillation, i.e., the root cause of this oscillation, is most probably closest to the $i$th process variable. If this oscillation were to propagate to the $k$th process variable, resulting in a secondary oscillation in that process variable, its significance index $a_{k,j}$ will also be much bigger than 0. A significance index in the $k$th process variable close to 1 would indicate a strong direct link between the $i$th and $k$th process variables. Examples of such deductions are given in Section 3. Significance indices associated with one or more ICs are compared by plotting them as stacked bar charts like that shown in Fig. 3. Each bar represents a separate process variable and is formed on the basis of the contributing ICs, where the total length is the sum of the contributing IC significance indices.

2.4. Multi-resolution spectral ICA

Experience has shown that some oscillations in either the middle-frequency or high-frequency ranges might not be detectable, because oscillations at low frequencies can dominate the total energy. In order to improve the detection resolution in the middle-frequency to high-frequency ranges, a multi-resolution spectra ICA is preferred. Alsberg (1999) has also commented on this problem and compared methods for multi-resolution analysis. Following his guideline, a multi-resolution spectral analysis rather than wavelet analysis is called for because persistent oscillations are present with no time-localisation. The multi-resolution spectral ICA is achieved by three steps:

Step 1: The power spectra are partitioned into a number of different frequency ranges, for example low-frequency, middle-frequency and high-frequency range,

i.e., $\mathbf{X} = (\mathbf{X}_1 : \mathbf{X}_2 : \mathbf{X}_3)$

where

$$\mathbf{X}_1 = \begin{pmatrix} P_1(f_{N_1}) & \cdots & P_1(f_{N_2}) \\ \vdots & \vdots & \vdots \\ P_m(f_{N_1}) & \cdots & P_m(f_{N_2}) \end{pmatrix},$$

$$\mathbf{X}_2 = \begin{pmatrix} P_1(f_{N_1}+1) & \cdots & P_1(f_{N_3}) \\ \vdots & \vdots & \vdots \\ P_m(f_{N_1}+1) & \cdots & P_m(f_{N_3}) \end{pmatrix},$$

$$\mathbf{X}_3 = \begin{pmatrix} P_1(f_{N_2}+1) & \cdots & P_1(f_{N}) \\ \vdots & \vdots & \vdots \\ P_m(f_{N_2}+1) & \cdots & P_m(f_{N}) \end{pmatrix},$$

where $f_{N_1}$ and $f_{N_2}$ separate the frequency ranges: $0 < f_{N_1} < f_{N_2} < f_N$. Before applying spectral ICA on each of the subset $\mathbf{X}_1$, $\mathbf{X}_2$ and $\mathbf{X}_3$, each row of each sub-set matrix should be normalised as described in Eq. (4). The normalised version is then

$$\mathbf{X}_k = \mathbf{\Omega}_k^{-1} \mathbf{X}_k \quad (k = 1, 2, 3),$$

(10)

where $\mathbf{\Omega}_k$ is an $m \times m$ diagonal matrix with the element at position $(m,m)$ being equal to the sum of the $m$th row of $\mathbf{X}_k$.

Step 2: Basic spectral ICA is performed on each of the three normalised data sub-sets. Each analysis will result in $n_k$ ICs

$$\mathbf{X}_k = \mathbf{A}_k \mathbf{C}_k \quad (k = 1, 2, 3),$$

(11)

where $\mathbf{A}_k$ is the $m \times n_k$ mixing matrix and $\mathbf{C}_k$ is the $n_k \times r_k$ IC matrix, where $r_1 = N_1$, $r_2 = N_2 - N_1$ and $r_3 = N - N_2$. Substitution of Eq. (11) into Eq. (10) gives

$$\mathbf{X}_k = \mathbf{\Omega}_k \mathbf{A}_k \mathbf{C}_k \quad (k = 1, 2, 3).$$

(12)

Step 3: The results are combined, ensuring that the final result has the same form as Eq. (8), by padding out matrices with zeros. The alignment is made on both $\mathbf{X}_k$ and $\mathbf{C}_k$

$$[\mathbf{X}_1, 0_{m \times (N-N_1)}]_{m \times N} = \mathbf{\Omega}_1 \mathbf{A}_1 [\mathbf{C}_1, 0_{n_1 \times (N-N_1)}]_{n_1 \times N},$$

(13)

$$[0_{m \times N_1}, \mathbf{X}_2, 0_{m \times (N-N_2)}]_{m \times N} = \mathbf{\Omega}_2 \mathbf{A}_2 [0_{r_2 \times N_1}, \mathbf{C}_2, 0_{n_2 \times (N-N_2^2)}]_{n_2 \times N},$$

(14)

$$[0_{m \times N_2}, \mathbf{X}_3]_{m \times N} = \mathbf{\Omega}_3 \mathbf{A}_3 [0_{r_3 \times N_2}, \mathbf{C}_3]_{n_3 \times N}.$$  

(15)
The synthetic result will then be

\[ X = Z_1 + Z_2 + Z_3 \]

\[ = \left[ \Omega_1 A_1, \Omega_2 A_2, \Omega_3 A_3 \right] \begin{bmatrix} E_1 \\ E_2 \\ E_3 \end{bmatrix} = AC. \quad (16) \]

which is exactly the same model as that of the basic spectral ICA. The post-ICA algorithm for determination of significance indexes should be performed on this synthetic result.

3. Case study

*Industrial data*: The data set was provided courtesy of the Eastman Chemical Company, Kingsport, Tennessee. It consists of 30 process variables (PVs) including 15 controlled variables and 15 uncontrolled variables from indicators. Controller output (OP) records were also available for each controller. The time trends and their power spectra are shown in Fig. 1. The sampling interval was 20 s and data records of 8192 samples were used to generate the power spectra by DFT with \( N = 8192 \). Thus the rows of the data matrix \( X \) had 4096 elements up to the Nyquist frequency.

Detection of oscillatory disturbances: Fig. 1 shows that many spectra contain sharp peaks that correspond to oscillatory trends in the time domain. It is clear that there is a dominant, plant-wide oscillation corresponding to peaks at about 0.0028–0.0032 on the normalised frequency axis (356–315 samples/minute). This oscillation had already been examined and a sticking valve was found in the loop with Tag 22, identified by means of a classical plot of the valve demand versus flow which showed a deadband due to excessive static friction, and valve tests by Eastman engineers. The fault lay with the level control valve on the outflow of a tank.

It is easy to see the oscillatory behaviour when the data are presented in high-density plots such as Fig. 1. However, such plots are not a normal part of a plant operator’s display panel. The aim of the method discussed in this article is to provide analysis in the background that will alert a plant control engineer to a problem and provide a rapid solution.

Spectral ICA should detect and isolate not only this dominant oscillatory source but also other oscillations in the entire frequency range. Multi-resolution spectral ICA was applied because some of the less dominant peaks such as those in Tags 16 to 18 appear in the medium-to-high frequencies. Fig. 2 shows the results of 10 ICs (IC1–IC5 in the low-frequency range, IC6–IC8 in the middle-frequency range and IC9–IC10 in the high-frequency range). The dominance ratio and estimated periods are listed in Table 1. Significance index plots for different frequency ranges are shown in Figs. 3 and 4.

Results: Each detected IC in Fig. 2 contains a single dominant peak, which indicates an oscillatory source in the data. It can be seen from Table 1 that

- IC2 and IC4 should be examined first because they represent 47.4% of the total energy (with dominance ratios of 23.5%
Table 1
Dominance ratio (DR) and period of oscillation

<table>
<thead>
<tr>
<th></th>
<th>DR (%)</th>
<th>Period (samples/cycle)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Low-f range</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IC1</td>
<td>14.0</td>
<td>8192</td>
</tr>
<tr>
<td>IC2</td>
<td>23.5</td>
<td>315</td>
</tr>
<tr>
<td>IC3</td>
<td>14.7</td>
<td>4096</td>
</tr>
<tr>
<td>IC4</td>
<td>23.9</td>
<td>356</td>
</tr>
<tr>
<td>IC5</td>
<td>2.1</td>
<td>512</td>
</tr>
<tr>
<td><strong>Mid-f range</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IC6</td>
<td>0.4</td>
<td>62</td>
</tr>
<tr>
<td>IC7</td>
<td>0.8</td>
<td>66</td>
</tr>
<tr>
<td>IC8</td>
<td>0.2</td>
<td>76</td>
</tr>
<tr>
<td><strong>High-f range</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IC9</td>
<td>16.2</td>
<td>18</td>
</tr>
<tr>
<td>IC10</td>
<td>4.2</td>
<td>32</td>
</tr>
</tbody>
</table>

- IC2 and IC4 contain oscillatory sources with periods of 315 and 356 samples/cycle.
- IC10 has a period of 32 samples/cycle oscillation, DR = 4.2%.
- IC9 has a period of 18 samples/cycle oscillation, DR = 16.2%.
- IC6, IC7, and IC8 have periods of 62, 66, and 76 samples/cycle, respectively.

Fig. 3. Significance index plot for low frequency IC2, IC4 and IC5.

Fig. 4. Other significance index plots for ICs in different frequency ranges.
• IC5 with an energy of 2.1% is small but not negligible. It also should be investigated. The period of oscillation is 512 samples per cycle;
• IC9 and IC10 are high-frequency components that represent 20.4% of the total energy (16.2% + 4.2%); they contain oscillatory sources with periods of 18 and 32 samples/cycle; IC9 is most dominant in this frequency range;
• IC6, IC7 and IC8 are middle-frequency components with periods of 62, 66 and 76 samples/cycle, and contain only 1.4% of the total energy (0.4% + 0.8% + 0.2%). In spite of their lack of dominance, they are still detectable by means of multiresolution analysis.

Isolation of the oscillatory sources: Isolation of these oscillatory sources is based on an examination of the significance index plots. It is assumed that a large significance index for a given tag means that the IC in question is closely associated with that tag, which is therefore a candidate for the source.

Fig. 3 shows the low-frequency components IC2, IC4 and IC5. The figure suggests that Tag 22 is the source of IC2 and IC4 (315 and 356 samples/cycle) and that those oscillations are linked together and have propagated widely. The faulty control valve in tag 22 discussed earlier was causing a limit cycle oscillation. The reasoning behind these conclusions is as follows:

• the most dominant oscillatory components, i.e. IC2 and IC4, always appear together in almost all of the perturbed process measurements (both PVs and OPs);
• IC2 and IC4 appear in 21 tags;
• tags 22.PV and 22.OP have the biggest significance indices for the combination of IC2 and IC4.

The conclusion that the IC2 and IC4 oscillations are linked may also be inferred from inspection of the original spectra. Many of the spectra have a double peak at about 0.003 on the frequency axis. The 315 and 356 samples/cycle oscillations are the two components within that double peak. These two oscillations can also be seen in the time domain plots. For instance, in Tag 22 the peaks are closer together in the time period from 14–18 h and wider apart between 18 and 24 h. It is possible that the plant operators working on different shifts chose to run the plant with slightly different settings.

Further observations focus on the low energy oscillation in IC5 at 512 samples per cycle:

• Tag 10.OP is different from others, it has a big significance index value for IC5;
• no other process measurements are significant for IC5.

These findings show that the oscillation of IC5 (512 samples/cycle) is located only in the control loop of Tag 10, and this oscillation does not propagate. Moreover, it is present only in the OP of Tag 10 and not in the PV thus suggesting the control loop for Tag 10 is successfully rejecting an external disturbance. This demonstrates the importance of analysing OP data.

Fig. 4(a) shows the very low-frequency components IC1 and IC3. Both can be assigned to a source and both have propagated:

• the source of IC1 most directly links to the process variable 3.PV, because of its largest significance index for IC1. Process measurements with PV tags 2, 4 and 9 and OP tags 1 and 2 are also influenced by this slow oscillatory source;
• the source of IC3 most directly links to the process variable 21.PV, and its effect can also be seen in the data pertaining to PV tags 2, 14, 24 and 27 and OP tags 1, 2, 10, 14 and 17.

Fig. 4(b) shows the high-frequency components IC9, with 18 samples per cycle, and IC10, with 32 samples per cycle. IC9 is a dominant oscillation having a dominance ratio of 16.2%. Its significance index suggests it is associated with Tag 16. The oscillation with 32 samples per cycle (IC10) is most strongly linked to Tag 17. Both oscillations have propagated to a few other locations (e.g. tags 15 and 18).

Fig. 4(c) shows the middle-frequency components IC6, IC7 and IC8. These middle-frequency oscillations are less dominant because they have small dominance ratio (DR) values. Isolation of these sources is still feasible by means of the significance indices, as follows:

• the control loop for Tag 17 is responsible for the IC6 oscillation (see 17.PV);
• the control loop for Tag 7 is responsible for the IC7 and IC8 oscillation (see 7.OP).

By combining these observations it can be inferred that Tag 17 is the possible source of two separate oscillations, IC10 at 32 samples per cycle and IC6 at 62 samples per cycle. Given that 32 and 62 are close to having a ratio of 2:1 it may be the case that the IC10 oscillation is a harmonic of a non-sinusoidal oscillation at about 62 samples per cycle.

Some of the significance indices in Figs. 3 and 4 are very slightly negative, so slightly negative that they are unlikely to influence any conclusions that would be drawn. They arise because ICA performs a search resulting in ICs, which can have both negative as well as positive parts (e.g. see Fig. 3) and hence significance indices can be both positive and negative. The sign correction algorithm described in Section 2.2 does not correct for the negative noise effects because it focuses on the dominant peaks.

4. Conclusion

The paper has presented a new way of applying independent component analysis (ICA) for the detection and isolation of multiple oscillations in measurements from a chemical process. The proposed multi-resolution spectral ICA is based on the ICA analysis of spectra derived via a discrete Fourier transform from time-domain process data. Spectral ICA is able to extract dominant spectrum-like ICs each of which has a narrow-bank peak that captures the behaviour of one of the oscillation sources. Proofs were presented to underpin the theory of spectral ICA, in particular it was shown that, for a data set having multiple
oscillatory sources, extraction of ICs with narrow-band spectral peaks can be guaranteed by an ICA algorithm that maximises the kurtosis of the ICs.

The periods of the detected oscillations may be determined from the positions of the spectral peaks in the ICs and two indexes have been defined and shown to be effective in the isolation of the oscillatory sources. One is termed the dominance ratio (DR) that indicates the severity of an oscillation as a percentage of the total power. The other is a significance index that isolates an oscillatory source to a particular location in the plant. A bar chart was used to aid visualisation and comparison of the significance indexes. The benefits of these two indexes are that they help plant control engineers and maintenance personnel to prioritise the faults and to focus effort in the right part of the plant.

This novel method has been applied to plant data obtained from an industrial chemical plant. Results demonstrate its ability to detect and isolate multiple dominant oscillations in different frequency ranges.

The starting point for spectral ICA is a spectral principal component analysis (spectral PCA). The benefit to be gained from the additional computational steps of spectral ICA is that the ICs each contain one narrow band peak compared to spectral PCA where the principal components may contain more than one peak. The isolation of the oscillatory sources is thus aided and enhanced by adding a spectral ICA stage to a multivariate spectral analysis.

Acknowledgements

The authors gratefully acknowledge John Cox and Michael Paulonis from Eastman Chemical Company, Kingsport, TN, USA, for the supply of real plant data, the Committee of Vice-Chancellors and Principals (ORS Awards) and the University of Glasgow (Faculty Scholarship).

Appendix: The kurtosis of some typical power spectra

The definition of the normalised kurtosis of a random variable $x$ is

$$kurt(x) = \frac{\mu_4 - 3\mu_2^2}{\mu_2^3},$$

where $\mu_4$ is the fourth-order central moment of $x$, and $\mu_2$ is the second-order central moment, i.e., the variance of $x$.

**Lemma 1.** The kurtosis of the power spectrum of a finite sample of Gaussian noise is 6.

**Proof 1.** Based on the fact that both the real and imaginary part of the Gaussian-noise $\bar{N}$-point FFT result have the same normal distribution and are independent (Ljung, 1999), it can be found that the 2-sided power spectrum of Gaussian noise has a special Gamma distribution that is exponential

$$f(t) = \frac{1}{\beta^r \Gamma(r)} t^{r-1} e^{-t/\beta} = e^{-t/\beta}$$

with $\alpha = 1$ and $\beta = 2\bar{N}/\sigma^2$, where $\Gamma(\cdot)$ is the Gamma function (Fisz, 1963), and $r$ represents power spectrum $P_{xx}$. The moment generating function for the exponential distribution is

$$M(t) = (1 - \beta t)^{-1}$$

and the $r$th moment can be given by $\mu'_r = \left. \frac{\partial^r M(t)}{\partial t^r} \right|_{t=0}, r = 1, 2, 3, 4, \ldots$, e.g. $\mu'_1 = \beta$, $\mu'_2 = 2\beta^2$, $\mu'_3 = 6\beta^3$, $\mu'_4 = 24\beta^4$.

Any central-moment $\mu_4$ can be evaluated from moments of order up to $r$ (Papoulis, 1991) for example,

$$\mu_2 = \mu'_2 - (\mu'_1)^2, \quad \mu_4 = \mu'_4 - 4\mu'_1 \mu'_2 + 6(\mu'_1)^2 \mu'_2 - 3(\mu'_1)^4.$$  

The kurtosis of $P_{xx}$ is then

$$kurt(P_{xx}) = \frac{\mu_4 - 3\mu_2^2}{\mu_2^3} = \frac{9\beta^4 - 3(2\beta^2)^2}{(2\beta^2)^3} = 6. \quad \Box$$

**Lemma 2.** The kurtosis of the power spectrum of a pure sine wave, $x = A \sin(\omega t)$, is $(\bar{N}/2) - 3$, where $\bar{N}$ is the number of FFT data points.

**Proof 2.** The 2-sided power spectrum of $x$ is given by

$$P_{xx} = U \delta(\omega - \omega_k) + U \delta(\omega - (\omega_k - \omega)), \quad (21)$$

where $\delta(\cdot)$ is the delta function, $\omega_k$ is the sampling frequency, and $U = (\bar{N}/4)A^2$ is the magnitude of the power spectrum peak. The second- and fourth-order sample central moments can be calculated according to the natural moment estimators:

$$\mu_2 = \frac{2}{N} \times \left( U - \frac{2U}{N} \right)^2, \quad \mu_4 = \frac{2}{N} \times \left( U - \frac{2U}{N} \right)^4.$$  

(22,23)

The kurtosis is then

$$kurt(P_{xx}) = \frac{\mu_4}{\mu_2^3} - 3 = \frac{\bar{N}}{2} - 3. \quad \Box$$

**Lemma 3.** The kurtosis of the power spectrum of a deterministic trend $x$ with several significant harmonics $x = \sum_{i=1}^{k} A_i \sin(\omega_i t)$, where $\omega_i \neq \omega_j$ ($\forall i \neq j$), has a lower bound of $(\bar{N}/2k) - 3$, and an upper bound of $(\bar{N}/2) - 3$, where $\bar{N}$ is the number of FFT data points.

**Proof 3.** The 2-side power spectrum of $x$ is given by

$$P_{xx} = \sum_{i=1}^{k} \left[ x_i \delta(\omega - \omega_i) + x_i \delta(\omega - (\omega_k - \omega_i)) \right],$$

(25)

where $\delta(\cdot)$ is the delta function, $\omega$ is sampling frequency, and $x_i$ ($i = 1 \ldots k$) is the magnitude of the relevant power spectrum.
peak. The second- and fourth-order sample central moments are:

\[
\mu_2 = \frac{2}{N} \times \sum_{i=1}^{k} B_i, \quad \mu_4 = \frac{2}{N} \times \sum_{i=1}^{k} B_i^2,
\]

where \( B_i = (x_i - \sum_{i=1}^{k} 2x_i^2/N)^2, \ B_i > 0, \ (i = 1 \ldots k). \)

The kurtosis is then

\[
kurt(P_{xx}) = \frac{\mu_4}{\mu_2^2} - 3 = \frac{N}{2} \left( \sum_{i=1}^{k} B_i^2 \right)^2 - 3. \tag{26}
\]

It is clear that: \( kurt(P_{xx}) \leq (N/2) - 3, \) and the equality will hold only when \( k = 1. \) The low bound in \( kurt(P_{xx}) \geq (N/2k) - 3 \) can be found by considering the inequality equation (27)

\[
\left( \sum_{i=1}^{k} B_i \right)^2 / k \leq \sum_{i=1}^{k} B_i^2 / k. \tag{27}
\]

For the low bound, the equality will only hold when \( B_1 = B_2 = \ldots = B_k, \) i.e., all dominant sine wave components have the same height of their peaks in the power spectrum. \( \square \)

References


Chuming Xia was educated at the East China University of Science and Technology (B.Sc, Process Control & Instrumentation, 1991; M.Sc. Intelligent Control Equipment, 1994) and received his Ph.D. from the Glasgow University in 2003. He worked for the East China University of Science and Technology (ECUST) as a lecturer from 1994 to 1999 and joined ECUST again in 2004 having been at the Honeywell Technology Solution Lab from 2003 to 2004. His research interests are in performance monitoring and assessment, in fault diagnosis in process industries, and in engineering applications of signal processing.

John Howell was educated at the Sheffield University (B.Eng. Control Systems, 1975) and received his Ph.D. from the Glasgow University in 1991. He worked on the fast reactor project with the United Kingdom Atomic Energy Authority from 1973 to 1984, when he joined the Glasgow University. He has 20 years of experience in nuclear materials safeguards, originally in near real-time materials accountability, but more recently in the monitoring of the progress of solutions through a process plant. His research interests are in anomaly diagnosis in general and in plant-wide performance assessment in particular.

Nina Thornhill was educated at Oxford University (B.A. Physics, 1976) and Imperial College London (M.Sc. Control Systems, 1983). She worked for ICI from 1976 to 1982, joined University College London (UCL) in 1984 and in 1992–1995 held an EPSRC/Royal Society Industry Fellowship with BP. She is currently a Professor in the Department of E&E Engineering and a member of the Imperial College/UCL Centre for Process Systems Engineering. Her research interests are process data analysis and plant-wide performance assessment with applications in oil, chemicals and pharmaceuticals.