A dynamic independent component analysis approach to fault detection with new statistics

M. Teimoortashloo[1] **and A. Khaki Sedigh**[2,*]

[1]*Department of Electrical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran Tel.: +989153196220, Fax: +985842249420, E-mail address: mazdak1978@bojnourdiau.ac.ir Address: Ghiyam, Number 409, Bojnourd, Iran*
[2]*Industrial Control Center of Excellence, Department of Electrical and Computer Engineering, K.N.Toosi University of Technology, Tehran, Iran*

This paper presents a fault detection method based on Dynamic Independent Component Analysis (DICA) with new statistics. These new statistics are statistical moments and first characteristic function that surrogate the norm operator to calculate the fault detection statistics to determine the control limit of the independent components (ICs). The estimation of first characteristic function by its series is modified such that the effect of series remainder on estimation is reduced. The advantage of using first characteristic function and moments, over second characteristic function and cumulants, as fault detection statistics is also presented. It is shown that the proposed method can detect a class of faults that the former methods cannot; in particular faults with small amplitude ICs that have either different probability density function or identical probability density function of the ICs, but different low order moments of the ICs compared with the normal performance. Simulation results are used to show the effectiveness of the proposed method.

Keywords: DICA, fault detection, statistical moment, statistical cumulants, characteristic function

1. INTRODUCTION

Many industrial plants encompass a large number of control loops and sensors are employed to measure the variable in each individual loop. The real measured variables are sampled periodically and the collected data is used for investigation the remote monitoring system [1] and fault detection strategies. However, after a short period of time the large data sets are assembled that need to be analyzed, and processing these large data sets in a short time requires powerful computational facilities which may not be available in all industrial site. Consequently, it is crucial to use strategies with low computational cost and fast to extract the important data from such large data sets and the statistical approaches for data analysis and classification can be used to achieve this objective. Require to these approaches to be sensitive, because the number of large scale systems increase by proposed the method of filtering of sensor reading in sensor networks such as [2].

In plants such as chemical industries and nuclear power plants, where faults can lead to biological and human tragedies, fault detection at the early stages, before they develop into major defects, is crucial. Being detected at the early stages, problems can be solved by replacing exhausted components and consequently major failures can be prevented. Moreover, fault detection at the early stages prevents failure extension and hence reduces the need for more costly repairs, improves the system reliability and reduces the system downtime [3, 4, 5, 6, and 7]. However, fault detection and root cause diagnosis are challenging tasks. In most industrial plants, the waste exchange lines are installed to preclude raw material loss. This leads to systems with highly interactive loops; in particular, when the buffer tanks are bypassed to save energy. In such systems, the influence of a fault, occurred in one of the loops, propagates through all the loops and makes it is impossible for the operator to identify the root cause of the fault. Therefore, fault detection and root cause diagnosis at the early stages, when the fault amplitude is small, is of critical importance.

*Corresponding Author: E-mail address: sedigh@kntu.ac.ir, Tel.: +982184062317 Address: Sayyed Khandan, P.O.Box 16315-1355 Tehran, Iran

Thornhill and Horch reviewed and categorized the suggested methods in plant-wide disturbance detection and diagnosis [8]. Decomposition is one of the proposed strategies which includes three methods: Principal Component Analysis (PCA), Independent Component Analysis (ICA), and Non-Negative Matrix Factorization (NNMF). PCA has been widely used to solve the stated problems. This approach projects the data space onto a lower dimensional subspace such that the correlation between the projected variables is minimized. It also creates some new unrelated variables by which faults root cause can be detected. The PCA method is well adapted to the Gaussian latent variables. It is worth mentioning that in Gaussian distribution, uncorrelation leads to independence. However, Martin and Morris showed that latent variables of most real systems do not follow a Gaussian distribution [9]. The other method to dimension reduction are proposed by Seungdo in[10].

Lee et al. introduced a new statistical process monitoring approach using ICA [11]. In the ICA method, statistical dependence between latent variables is minimized and it is shown that it reveals more useful information from the observation data in comparison with PCA. However, in fault detection methods based on the ICA, the system is assumed to be static. Lee et al. extended the approach and proposed Dynamic ICA (DICA) to overcome this disadvantage [12]. The idea behind this method is adopted from [13, 14]. The drawback of this approach is the drastic increase in dimensions as a result of expanding data vector to augmented matrix. Teimoortashloo et al. constructed the new observation matrices to improve the drawbacks of dealing with high dimensional systems, such as reduction in the running time, in the number of arithmetic operations, and in the error of orthogonalization estimation, as well as the overall improvement of capability of fault detection [15]. Another assumption in fault detection methods using ICA/DICA is that the variables do not follow a Gaussian distribution. Ge and Song combined PCA and ICA and proposed a new approach to deal with Gaussian and non-Gaussian latent variable [16]. Ge and Song proposed Maximum-likelihood mixture factor analysis model for process monitoring [17]. Ge et al. Improved kernel PCA-based monitoring approach for nonlinear processes [18]. Brys et al. showed some of the weak points of ICA-based fault detection; one of which is the assumption that the offline part observations have no outliers, although this does not always true[19]. In particular circumstances, if a fault occurs in the offline part, it is assumed as normal performance and hence remains undetected in the online part. Therefore, outlier rejection method is proposed by [20]. They also showed that using square of 2-norm for calculating the control limit, which produces elliptical control limit, is not suitable if ICA/DICA method leads to a joint distribution that the mutual independent components have rectangular-shape. Thus they suggested adjusted outlier (AO) approach that can produce a rectangular-shape control limit.

Ge et al. proposed Batch process monitoring based on the support vector data description method [21]. Miao et al. developed a new nonlinear fault detection method based on locally linear embedding [22]. Ge and Song proposed a distribution-free method for process monitoring [23]. Ge et al. reviewed on data-based process monitoring [24]. Mu-Chen et al. used Durbin Watson index to choose the dominant ICs and support vector data description as fault detection statistics [25]. Huang and Yan proposed DPCA-DICA and Bayesian Inference to put

variable into a block on the basis of the variable normality and then used Jarque-Bera test [26]. Zvokelj et al. combined ensemble empirical mode decomposition with ICA to decompose signal to difference time scales [29].

Statistical moments and cumulants have been widely used in signal processing, blind system identification and blind source separation [30]. Statistical moments and cumulants as well as first and second characteristic function, moment generating function, and cumulants generating function are well documented in the literature [30, 31, 32, 33 and 13]. In signal processing, if probability density function (pdf) to be existed and their distribution to be continued, finding the moments can be lead to finding their pdf, and the moments can be calculated from the moment generating function [13]. The characteristic function is also a moment generating function [31]. Utilizing characteristic function is more appropriate than distribution functions as it can be estimated by smaller sample size [34]. Moreover, characteristic function always exists but moment generating function does not necessarily exist [31]. In fact, it is argued that when the sample size is small, analyzing the model by two first moments and a selection of higher order moments yields better results, in terms of stability and robustness, rather than approaches in which the use of full set of higher order moments is suggested [35, 36]. While increasing the number of moments leads to better approximation of distribution function, in practice moments higher than the fourth order are rarely used or needed; that is due to the fact that higher-order moments are sensitive to outliers [13].

The main contributions of this paper can be summarized as follows. The new statistics are introduced in order to detect a class of faults with the small amplitude ICs that the former methods fail to detect. These new statistics are the lower statistical moments and first characteristic function that surrogate the 2-norm or AO operators to calculate the control limit of the independent components (ICs). The fault detection occurs when the moments or value of the first characteristic function of ICs, at certain frequency, are out of the normal performance interval. This means that the pdf of ICs or some moments differ from those of the normal performance.

1.1 Motivating example

To clarify the benefits of the proposed method over the former ones, see Figure 1. This figure shows the joint distribution of the two IC vectors. The red dots are normal operation ICs and the green squares are faulty ICs. Using 2-norm or AO approach leads to an elliptical control limit or a rectangular-shape control limit, respectively, and neither of them can detect the fault (see black dashed circle or blue dashed rectangular in Figure 1). It is worth mentioning that the fault detection occurs when the 2-norm or AO operator of ICs are out of the normal performance interval, control limit. On the other hand, low order moments of ICs detect this fault because the density of ICs increases in the left hand side of joint distribution plot.

The rest of the paper is organized as follows. Section two presents a brief review of DICA and fault detection using DICA. New statistics to determine the control limit are introduced in section three, and section four describes the proposed algorithm. Simulation results and conclusion are included in sections five and six, respectively.
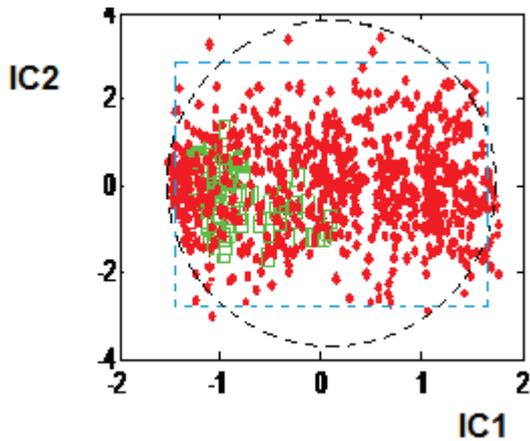
Figure 1 Benefit of proposed method over the former ones. The red dots are normal operation ICs and the green squares are faulty ICs. The black dashed circle and blue dashed rectangular are square of 2-norm and AO control limit.

2. FAULT DETECTION

This section provides a brief review of the DICA algorithm and DICA-based fault detection method.

2.1 The ICA algorithm

In the ICA algorithm, the idea is to find m ICs (s_r $r = 1, \ldots, m$ as elements of random vector $\mathbf{s}(\mathbf{k}) \in R^{m \times 1}$) as source from the m observed variables (x_r $r = 1, \ldots, m$ as element of random vector $\mathbf{x}(\mathbf{k}) \in R^{m \times 1}$) that are sampled from the sensors. The number of variables can be bigger than the ICs, and vice versa. In general, in fault detection methods, the number of variable and ICs are assumed to be equal. In this method, N vectors of $\mathbf{s}(\mathbf{k}) \in R^{m \times 1}$ and $\mathbf{x}(\mathbf{k}) \in R^{m \times 1}$, respectively, are constructed matrices $\mathbf{x} \in R^{m \times N}$ and $\mathbf{S} \in R^{m \times N}$ (N is number of samples of each variable). The ICA algorithm searches to find the mixing matrix $\mathbf{A} \in R^{m \times m}$ and the ICs matrix $\mathbf{S} \in R^{m \times N}$, which are unknown matrices, from the known matrix $\mathbf{x} \in R^{m \times N}$, where

$$\mathbf{X} = \mathbf{AS}. \tag{1}$$

The observed variables are typically mean centered. By centering the observed variables, i.e., subtracting their sample mean simplifies the theory and algorithms. It is made in the rest of this manuscript. Equation

(1) can be written as $\mathbf{S} = \mathbf{W}.\mathbf{X}$, where $\mathbf{W} = \mathbf{A}^{-1}$. Matrix $\mathbf{x}$ is whitened using $\mathbf{Q} \in R^{m \times m}$ where $\mathbf{Q} = \Lambda^{-\frac{1}{2}} \mathbf{U}^T$, $\Lambda \in R^{m \times m}$, is a diagonal matrix with the eigenvalues of matrix $\mathbf{R_x} = E(\mathbf{x}(\mathbf{k})\mathbf{x}^T(\mathbf{k}))$ as its diagonal elements($\mathbf{x}(\mathbf{k})$ is the random vector at sample k with covariance $\mathbf{R_x}$ where E represents expectations see [11]). $\mathbf{U} \in R^{m \times m}$ is a matrix constructed by eigenvectors of $\mathbf{R_x} \in R^{m \times m}$ as its columns. The whitening matrix $\mathbf{Z}$ is obtained by:

$$\mathbf{Z} = \mathbf{QX} = \mathbf{QAS} = \mathbf{BS} \tag{2}$$

Where $\mathbf{B} = \mathbf{QA}$. Whitening the matrix $\mathbf{X}$ means that its components to be uncorrelated and their variances to be equaled to unity. In other words, the covariance matrix of $\mathbf{X}$ to be equaled

the identity matrix. The whitening is a good idea to reduce the complexity of the ICA algorithm because it reduces the number of elements that should be estimated (see [13]). It is easy to show that $\mathbf{B}^{-1} = \mathbf{B}^T$. Therefore, the problem of finding a full rank matrix $\mathbf{A}$ is reduced to finding an orthogonal matrix $\mathbf{B} \in R^{m \times m}$. Equation (2) is a matrix equation with two unknown matrices, $\mathbf{S}$ and $\mathbf{B}$, which can be rewritten as $\mathbf{S} = \mathbf{B}^T \mathbf{Z}$, with two constrains: first, $\mathbf{B} \in R^{m \times m}$ is orthonormal matrix and second, the statistical dependence between elements of vector $\mathbf{s}(\mathbf{k}) \in R^{m \times 1}$ are minimized.

Fast ICA, which employs fixed point optimization algorithm, was used by [11, 12] for fault detection is as follows:

1. Choose m and Set counter $i = 1$

2. Choose a random initial vector $\mathbf{b}_i \in R^{m \times 1}$ with unit norm ($\mathbf{b}_i$ is a ith column of matrix $\mathbf{B} \in R^{m \times m}$).

3. Calculate $\mathbf{b}_i \leftarrow E\{\mathbf{z}\dot{G}(\mathbf{b}_i^T \mathbf{z})\} - E\{\ddot{G}(\mathbf{b}_i^T \mathbf{z})\}\mathbf{b}_i$ where $\dot{G}$ and $\ddot{G}$ are the first and second derivative of G. Function $G(\mathbf{b}_i^T \mathbf{z})$ is defined as $\tanh(\mathbf{b}_i^T \mathbf{z})$ or $-\exp(-(\mathbf{b}_i^T \mathbf{z}^2)/2)$ or etc. Determination of G is discussed in [37].

4. Orthogonalize $\mathbf{b}_i$ using the Gram-Schmidt method:

$$\mathbf{b}_i \leftarrow \mathbf{b}_i - \sum_{j=1}^{i-1}(\mathbf{b}_j^T \mathbf{b}_i)\mathbf{b}_j$$

5. Normalize $\mathbf{b}_i$ as:

$$\mathbf{b}_i \leftarrow \frac{\mathbf{b}_i}{||\mathbf{b}_i||}$$

6. If $\mathbf{b}_i$ does not converge, go back to step 3.

7. If $\mathbf{b}_i$ converges, it is i^{th} column of $\mathbf{B}$. Then, if $i < m$ set $i = i + 1$ and go back to step 2. $\mathbf{b}_i$. Is converged if dot product of old and new values of $\mathbf{b}_i$ is equal around 1. To summarize, the matrix $\mathbf{B}$ is determined by the above-mentioned method. The independent component matrix, $\mathbf{s}$, then is calculated as $\mathbf{s} = \mathbf{B}^T \mathbf{Z}$, and finally the de-mixing matrix, $\mathbf{w} \in R^{m \times m}$, is calculated as $\mathbf{W} = \mathbf{B}^T \mathbf{Q}$.

2.2 Fault detection based on the DICA

The fault detection algorithm proposed by Lee et al. has two parts, in rest of this paper this method is called former method [11, 12]. An Offline part in which the normal operating data is obtained by an offline training; and an online part, which an online monitoring is performed to detect deviation from the normal operating data. It is worth mentioning that in fault detection, the main idea of ICA is blind source separation to diagnosis the root cause of fault.

2.2.1 The offline part

The offline part includes the following steps:

Step1: The observation matrix $\mathbf{X} \in R^{m(N+1)}$ is obtained from the assembled data and in order to carry out DICA, the time lag l (proposed as $l = 2$ by [12]) is chosen and the augmented matrix $\mathbf{x}_L$ ($\mathbf{X}_L \in R^{(l+1)m \times N}$) is constructed by augmenting columns

$\mathbf{X}(\mathbf{k}+1), \ldots, \mathbf{X}(\mathbf{k}+l)$ to column $\mathbf{X}(\mathbf{k})$ for $k = 1, \ldots, N$. Then the augmented matrix is normalized using the mean and standard deviation of each row.

Step 2: The normalized matrix must be whitened to obtain matrix $\mathbf{Z} \in R^{(l+1)m \times N}$ (see Equation (2)).

Step 3: The ICA algorithm is carried out to calculate the matrix $\mathbf{B} \in R^{m(l+1) \times m(l+1)}$ and de-mixing matrix $\mathbf{W} \in R^{m(l+1) \times m(l+1)}$.

Step 4: Rearrange and divide matrix $\mathbf{W} \in R^{m(l+1) \times m(l+1)}$ into two parts: the dominant part, $\mathbf{W}_d \in R^{d \times m(l+1)}$, and the excluded part, $\mathbf{W}_e \in R^{(m(l+1)-d \times m(l+1))}$, i.e. $\mathbf{W} = \begin{bmatrix} \mathbf{W_d} \\ \mathbf{W_e} \end{bmatrix}$. d is number of rows of dominant part of $\mathbf{W}$ (see [11]). Note that the norm of rows of matrix $\mathbf{W}$ is used for this division; higher norm construct $\mathbf{W}_d$ and lower norm construct $\mathbf{W}_e$. Rearrange and divide matrix $\mathbf{B} \in R^{m(l+1) \times m(l+1)}$ into two matrices $\mathbf{B}_d \in R^{m(l+1) \times d}$ and $\mathbf{B}_e \in R^{m(l+1) \times (m(l+1)-d)}$ corresponding to $\mathbf{W}_d$ and $\mathbf{W}_e$, respectively, i.e. $\mathbf{B} = \mathbf{B_d}\ \mathbf{B_e}$ or calculate by $\mathbf{B_d} = (\mathbf{W_d Q^{-1}})^T$ and $\mathbf{B}_e = (\mathbf{W}_e \mathbf{Q^{-1}})^T$. Then the IC matrix space can be decomposed into two subspaces; the dominant part, $\mathbf{S}_d \in R^{d \times N}$, and the excluded part $\mathbf{S_e} \in R^{(m(l+1)-d) \times N}$ and can be written as $\mathbf{S} = \begin{bmatrix} \mathbf{S_d} \\ \mathbf{S_e} \end{bmatrix}$ where $\mathbf{s_d} = \mathbf{w_d X_L}$ and $\mathbf{S_e} = \mathbf{W_e X_L}$.

Step 5: Three fault detection statistics are calculated to monitor the process:
The systematic dominant part $I_d^2(k) = \mathbf{s_e^T}(\mathbf{k})\mathbf{s_d}(k)\ k = 1, \ldots, N$
The systematic excluded part $I_e^2(k) = \mathbf{s_e^T}(\mathbf{k})\mathbf{s_e}(k)\ k = 1, \ldots, N$
The non-systematic part $SPE(k) = \mathbf{e^T}(\mathbf{k})\mathbf{e}(\mathbf{k})\ k = 1, \ldots, N$
where $\mathbf{e}(\mathbf{k}) = \mathbf{x_L}(\mathbf{k}) - \mathbf{\hat{x}_L}(\mathbf{k})$ and $\mathbf{\hat{x}_L}(\mathbf{k}) = \mathbf{Q^{-1}B_d W_d x_L}(\mathbf{k})$.

Step 6: Confidence limits of the three statistics are calculated by KDE (99% or 95%) to determine the control limit. KDE is a technique that appropriately determines the control limits via these statistics, while it simultaneously satisfies the confidence bounds. A univariate kernel estimator with kernel k is defined as $\hat{f}(x) = \frac{1}{Nh} \sum_{i=1}^{N} k\left(\frac{x-x_i}{h}\right)$, where h is a contant value namely window width, x is the considered data point, x_i is the observation and k is kernel function that usually choses as Gaussian density function [11].

2.2.2 The online part

The online part consists of the following steps:

Step1: For a new observation data vector (new samples of each variable) the augmented vector $\mathbf{x_{newL}}(\mathbf{k}) \in R^{(l+1)m \times 1}$ is constructed by augmenting columns $\mathbf{X_{new}}(\mathbf{k}+1), \ldots, \mathbf{XX_{new}}(\mathbf{k}+1)$ to column $\mathbf{X_{new}}(\mathbf{k})$. Then it is normalized by the offline part mean and standard deviation of each row as in step 1 of the offline part.

Step 2: The dominant and excluded parts of ICs are calculated by offline $\mathbf{W}_d$ and $\mathbf{W}_e$ as:

$$\mathbf{s_{dn}}(\mathbf{k}) = \mathbf{W_d x_{newL}}(\mathbf{k})$$
$$\mathbf{s_{en}}(\mathbf{k}) = \mathbf{W_e x_{newL}}(\mathbf{k})$$

Step 3: As in step 5 of the offline part three fault detection statistics are calculated as:

The systematic dominant part $I_{dn}^2(k) = \mathbf{s_{dn}^T}(\mathbf{k})\mathbf{s_{dn}}(\mathbf{k})$

The systematic excluded part $I_{en}^2(k) = \mathbf{s_{en}^T}(\mathbf{k})\mathbf{s_{en}}(\mathbf{k})$

The non-systematic part $SPE_n(k) = \mathbf{e_n^T}(\mathbf{k})\mathbf{e_n}(\mathbf{k})$ where $\mathbf{e}_n(\mathbf{k}) = \mathbf{x_{newL}}(\mathbf{k}) - \mathbf{\hat{x}_{newL}}(\mathbf{k})$ and $\hat{x}_{\mathbf{newL}}(\mathbf{k}) = \mathbf{Q^{-1}B_d W_d x_{newL}}(\mathbf{k})$.

If these new statistics are out of the control limit, a fault has occurred.

Step 4: Root causes of the faults are detected by using the contribution value. Contribution value reveals the ICs that most influence the fault detection statistics.

3. NEW FAULT DETECTION STATISTICS TO DETERMINE THE CONTROL LIMIT

In the former fault detection methods, control limit is a vector norm (2-norm) of ICs which creates a geometrical shape such that to embed the normal values of ICs acquired in the offline part [11]. Operation condition in the online part is then defined in accordance with this shape. The square of 2-norm value of ICs vectors being outside of this shape in the online part indicates that a fault/faults has/have occurred. This means that these methods cannot detect faults if the square of 2-norm of the ICs vectors in the online part is inside the control limit shape, see Figure 1. In other words, faults that lead to short amplitude ICs remain undetected until they increase to a crucial level and hence the norm of ICs vector exceeds the control limit shape, which can be too late for any further action.

To overcome this drawback, this paper presents a new approach, which the new fault detection statistics are used to define the control limit. In the proposed method, faults will be detected if the pdf, providing that its change can be estimated by low order moments of the first characteristic function, or the low order moments differ from those in the normal operating data.

3.1 Characteristic function

Consider a scalar random variable s with the pdf $f(s)$ (usually the pdf is unknown) and there is available, a set of $m(l+1)$ samples $s_1, \ldots, s_r \ldots, s_{m(l+1)}$ from s that construct the vector $s(k)$. According to Hyvarinen et al., the first characteristic function can be defined as the continues Fourier transform of pdf [13]:

$$\phi(\omega) = E\{e^{js\omega}\} = \int_{-\infty}^{+\infty} e^{js\omega} f(s)ds = \sum_{i=0}^{\infty} \mu_i' \frac{(j\omega)^i}{i} \quad (3)$$

And the second characteristic function can be expressed as follows (see [31]):

$$\psi(\omega) = \log(E\{e^{js\omega}\}) = -\sum_{n=1}^{\infty} \frac{1}{n}\left(-\sum_{i=1}^{\infty} \mu_i' \frac{(j\omega)^i}{i}\right)^n$$
$$= \sum_{i=0}^{\infty} \kappa_i \frac{(j\omega)^i}{i} \quad (4)$$

Where, μ_i' is the ith *moment* and κ_i is the ith cumulant. In principle, the first characteristic function, $\phi(\omega)$, is presented as a Taylor series and the second characteritic function, $\psi(\omega)$, is

presented as a power series in which each term is a Taylor series. μ'_i is defined as:

$$\mu'_i = E\{s^i\} = \int_{-\infty}^{+\infty} s^i f(s) ds \tag{5}$$

However, deu to the fact that the pdf is difficult to obtain in practice, the moment are often estimated by the sample mean as [38]:

$$\mu'_i = E\{s^n\} = \frac{1}{m(l+1)} \sum_{r=1}^{m(l+1)} s_r^n \tag{6}$$

Which s_r is the rth sample of scalar random variable s. According to Kendall and Stuart, nevertheless, the cumulants as well as the moments cannot be calculated directly by a summation operator [31]. Thus, it is necessary to, first, calculate the moments and then substituted the calculated moments in Equation (7) to acquire the cumulants. Providing that the mean of s is nonzero, first cumulant (mean), second cumulant (variance), third cumulant (skewness), and fourth cumulant (kurtosis) can be calculated as follows [13]:

$$\begin{aligned}
\kappa_1 &= \mu'_1 \\
\kappa_2 &= \mu'_2 - \mu'^2_1 \\
\kappa_3 &= \mu'_3 - 3\mu'_2\mu'_1 + 2\mu'^3_1 \\
\kappa_4 &= \mu'_4 - 4\mu'_3\mu'_1 - 3\mu'^2_2 + 12\mu'_2\mu'^2_1 - 6\mu'^4_1
\end{aligned} \tag{7}$$

It is easy to show that if the ith moment of scalar random variable s is shown by μ'_1 then the ith moment of scalar random variable ls is calculated by $l^i\mu'_1$. Where is a constant. It means that if the scale of samples decreases ($l < 1$) the higher order moments will decrease more than the lower order moments. While if the ith cumulant of scalar random variable s is shown by κ_i then the cumulant of scalar random variable ls is calculated by κ_i, too. It means that if the scale of samples decreases the higher order moments will not decrease more than the lower order moments.

3.2 Comparison between the first and the second Characteristic function as fault detection statistics

In this paper, the first characteristic function is used as fault detection statistics to determine the control limit. More precisely, the first characteristic function is estimated by the first four terms of the series given in Equation (3), and then it is used to determine a new control limit. The first characteristic function is chosen because its series (providing all terms of series be able to estimated) has an infinite radius of convergence, whereas the radius of convergence of the second characteristic function series is finite (even if all terms of series be able to estimated). Furthermore, the propagation of error in estimation the series terms of second characteristic function is bigger than the error of series terms of first characteristic function. Moreover, the error of series remainder can be reduced by changing the scale of all ICs in first characteristic function, while it is not true in second characteristic function.

Radius of convergence of the first and second characteristic function

According to Equation (3), the first characteristic function is expanded as a Taylor series of exponential function. The radius of convergence in exponential function is infinite. While according to Equation (4), the second characteristic function is expanded as a power series of the logarithm function in which each term is a Taylor series with infinite convergence radius. As a result, the second characteristic function can be written as:

$$\begin{aligned}
\psi(\omega) &= \log(E\{e^{js\omega}\}) = \log\left(\sum_{i=0}^{\infty} \mu'_1 \frac{(j\omega)^i}{i}\right) \\
&= \log\left(1 + \sum_{i=1}^{\infty} \mu'_1 \frac{(j\omega)^i}{i}\right)
\end{aligned} \tag{8}$$

Since $\mu'_0 = 1$, see [31]. It converges if

$$\left|\sum_{i=1}^{\infty} \mu'_1 \frac{(j\omega)^i}{i}\right| < 1 \tag{9}$$

Inequality (9) shows that Equation (8) may diverge for certain values of s and ω.

The propagation of error

Propagation of error is inevitable in indirect measurement. According to Equation (7), cumulants are calculated by using the moments and hence the cumulant computation suffers from propagation of error. In Appendix A shows that propagation of error in estimation the cumulants is bigger than the error of moments.

Thus, estimation of the first characteristic function has smaller error in comparison with the second characteristic function due to propagation of error. Furthermore, Appendix A shows the moments can be estimated more precisely in comparison with cumulants and hence, the moments are preferable.

Reduce the error of series remainder in first characteristic function

The error of series remainder can be reduced by changing the scale of all ICs in first characteristic function, while it is not true in second characteristic function. In the rest of this section the proposed method to reduce the error of series remainder is explained.

Estimation of the first characteristic function series only by its first four moments introduces a new type of error which is a direct result of the remainder, sum of the terms of order higher than four, of the series. Equation (3) can be rewritten as:

$$\psi(\omega) = \sum_{i=0}^{4} \mu'_1 \frac{(j\omega)^i}{i} + R(\omega) \tag{10}$$

Where $R(\omega)$ is the remainder of series. It is shown in the Appendix B that $0.125\omega^4 s_{max}^4$ is a supremum for the summation of high order moments. Therefore changing the scale of all ICs matrix elements to values smaller than one can be used to reduce the remainder of first characteristic function (It is worth mentioning that the moments are affected by changing the scale

of ICs values but the cumulants are not affected). This change is also useful in establishment of the inequalities (B-3) and (B-4) in the Appendix B, since it can satisfy the conditions of the inequalities. Thus,

$$s_{\text{new}}(k) = as_{\text{old}}(k) \quad \text{for } k = 1, 2, \ldots, N \qquad (11)$$

Where

$$a = \frac{1}{\lambda s_{\text{max offline}}} \qquad (12)$$

$\lambda > 1$ is a real value. Changing the scale result in $s_{\text{max new}} < 1$ and hence decreases the value of $0.125\omega^4 s_{\text{max new}}^4$ as quartic. Consequently, remainder (error) decreases as quartic. However, it can be seen in the last paragraph of section 3.1, that decline in lower order moments is smaller than that of higher order moments (decline in the first, second, and third moments are proportional, quadratic, and cubic, respectively). Thus the first characteristic function is more influenced by the lower order moments as the scale declines.

It should be mentioned that choosing a large λ, helps the constraints of inequalities (B-3) and (B-4) to be satisfied. But choosing a large λ yields to an increase in calculation error, round-off error, and a decline in the weights of the third and fourth moments in characteristic function estimation. Choosing a large ω, on the other hand, results in a dramatic increases in the supremum magnitude. It may also cause the constraints of inequalities (B-3) and (B-4) not to be satisfied. But choosing a small ω increases the calculation error, round-off error and decreases the weights of the third and fourth moments in characteristic function estimation. In practice, fault frequency, if known, would be a reasonable choice for ω and then λ can be chosen such that:

$$\omega < \lambda \qquad (13)$$

The inequality (13) is a result of substituting Equation (12) in Equation (11) and satisfying the constraints of inequalities (B-3) and (B-4). Notice that choosing a λ much bigger than ω increase the round of error. Therefore, λ should be chosen close to ω.

4. THE PROPOSED ALGORITHM

In this section the proposed fault detection algorithm is discussed.

4.1 The offline part

Step1: The observation matrix $\mathbf{X} \in R^{m \times (N+l)}$ is obtained by the observed data. Then, the time lag l is chosen and the augmented matrix $\mathbf{X}_L(\mathbf{X}_L \in R^{(l+1)m \times N})$ is constructed by augmenting columns $X(\mathbf{k}+1), \ldots, X(\mathbf{x}+l)$ to column $X(\mathbf{k})$ for $k = 1, \ldots, N$. Then the augmented matrix normalized by the mean and standard deviation of each row.

Step2: Compute the whitening matrix $\mathbf{Z} \in R^{m(l+1) \times N}$ (see Equation (2)) from normalized augmented matrix.

Step3: Compute matrices $\mathbf{B} \in R^{m(l+1) \times m(l+1)}$, $\mathbf{W} \in R^{m(l+1) \times m(l+1)}$ and $\mathbf{S} \in R^{m(l+1) \times N}$ using the ICA algorithm and find $S_{\text{max offline}}$, the largest absolute elements of matrix $\mathbf{S}$.

Step4: Choose ω_0 and λ which satisfies $\omega_0 < \lambda$ and close together and calculate a by Equation (12).

Step5: Rearrange and divide matrix $\mathbf{W} \in R^{m(l+1) \times m(l+1)}$ into two parts: the dominant part, $\mathbf{W}_d \in R^{d \times m(l+1)}$, and the excluded part, $\mathbf{W}_e \in R^{(m(l+1)-d) \times m(l+1)}$, i.e. $\mathbf{W} = \begin{bmatrix} \mathbf{W_d} \\ \mathbf{W_e} \end{bmatrix}$. d is number of rows of dominant part of $\mathbf{W}$ (see [11]). Note that the norm of rows of matrix $\mathbf{W}$ is used for this division; higher norm construct $\mathbf{W}_d$ and lower norm construct $\mathbf{W}_e$. Rearrange and divide matrix $\mathbf{B} \in R^{m(l+1) \times m(l+1)}$ into two matrices $\mathbf{B_d} \in R^{m(l+1) \times d}$ and $\mathbf{B_e} \in R^{m(l+1) \times (m(l+1)-d)}$ as:

$$\begin{cases} \mathbf{B_d} = (\mathbf{W_d} Q^{-1})^T \\ \mathbf{B_e} = (\mathbf{W_e} Q^{-1})^T \end{cases} \qquad (14)$$

Then the IC matrix space can be decomposed into two subspaces; the dominant part, $\mathbf{S}_{New\ \mathbf{d}} \in R^{d \times N}$, and the excluded part $\mathbf{S}_{New\ \mathbf{e}} \in R^{(m(l+1)-d) \times N}$ and can be written as $\mathbf{S} = \begin{bmatrix} \mathbf{S}_{New\ \mathbf{d}} \\ \mathbf{S}_{New\ \mathbf{e}} \end{bmatrix}$ Where $\mathbf{s}_{New\ \mathbf{d}} = a\mathbf{W_d}\mathbf{X_L}$ and $\mathbf{S}_{New\ \mathbf{e}} = a\mathbf{W_e}\mathbf{X_L}$.

Step6: chose $k = 1$ and evaluate the following fifteen fault detection statistics to monitor the process:

Calculate the first, second, third, and fourth moment statistics and the first characteristic function statistic at frequency ω_0 for the elements of ICs vector $\mathbf{s}_{New\ d}(k)$, $I\phi_{d\omega_0}(k)$ and $I\mu'_{id}(k)$ $i = 1, 2, 3, 4$, as follows:

$$I\mu'_{i_d}(k) = |\mu'_{i_d}(k)| \ i = 1, \ldots, 4 \qquad (15)$$

$$I\phi_{d\omega_0}(k) = |1 - |\phi_{d\omega_0}(k)|| \qquad (16)$$

Where

$$\mu'_{i_d}(k) = \frac{1}{d} \sum_{r=1}^{d} (s_{newd}(k)_r)^i \ i = 1, \ldots, 4 \qquad (17)$$

$$|\phi_{d\omega_0}(k)| = \left| 1 + j\mu'_{1_d}(k)\omega_0 + \mu'_{2_d}(k)\frac{(j\omega_0)^2}{2} \right. \\ \left. + \mu'_{3_d}(k)\frac{(j\omega_0)^3}{3} + \mu'_{4_d}(k)\frac{(j\omega_0)^4}{4} \right| \qquad (18)$$

Which $s_{Newd}(k)$, is the rth element of vector $\mathbf{s_{Newd}}(\mathbf{k})$

Calculate the first, second, third, and fourth moment statistics and first characteristic function statistic at frequency ω_0 for the elements of ICs vector $\mathbf{s}_{Newe}(k)$, $I\phi_{e\omega_0}(k)$ and $I\mu'_{i_e}(k)$ i=1,2,3,4, as follows:

$$I\mu'_{i_e}(k) = |\mu'_{i_e}(k)| \ i = 1, \ldots, 4 \qquad (19)$$

$$I\phi_{e\omega_0}(k) = |1 - |\phi_{e\omega_0}(k)|| \qquad (20)$$

Where

$$\mu'_{i_e}(k) = \frac{1}{m(l+1) - d} \sum_{r=1}^{m(l+1)-d} (s_{newe}(k)_r)^i \ i = 1, \ldots, 4 \qquad (21)$$

$$|\phi_{e\omega_0}(k)| = \left| 1 + j\mu'_{1_e}(k)\omega_0 + \mu'_{2_e}(k)\frac{(j\omega_0)^2}{2} + \mu'_{3_e}(k)\frac{(j\omega_0)^3}{3} \right. \\ \left. + \mu'_{4_e}(k)\frac{(j\omega_0)^4}{4} \right| \qquad (22)$$

Which $s_{Newe}(k)_r$ is the rth element of vector $\mathbf{s_{Newe}}(k)$.

Similar to the aforementioned procedure, the first four moment statistics and the first characteristic functions statistic at frequency ω_0 for the elements of vector $\mathbf{e}(k)$, $I\phi_{ns\omega_0}(k)$ and $I\mu'_{ins}(k)$ $i = 1, 2, 3, 4$, need to be calculated (For the nonsystematic part) as:

$$I\mu'_{ins}(k) = |\mu'_{ins}(k)| \; i = 1, \ldots, 4 \qquad (23)$$

$$I\phi_{ns\omega_0}(k) = |1 - |\phi_{ns\omega_0}(k)|| \qquad (24)$$

Where

$$\mu'_{ins}(k) = \frac{1}{m(l+1)} \sum_{r=1}^{m(l+1)} (e(k)_r)^i \; i = 1, \ldots, 4 \qquad (25)$$

$$|\phi_{ns\omega_0}(k)| = \left| 1 + j\mu'_{1ns}(k)\omega_0 + \mu'_{2ns}(k)\frac{(j\omega_0)^2}{2} \right.$$
$$\left. \mu'_{3ns}(k)\frac{(j\omega_0)^3}{3} + \mu'_{4ns}(k)\frac{(j\omega_0)^4}{4} \right| \qquad (26)$$

where $\mathbf{e}(k) = \mathbf{x_L}(k) - \hat{\mathbf{x}}_\mathbf{L}(k)$ and $\hat{\mathbf{x}}_\mathbf{L}(k) = \mathbf{Q}^{-1}\mathbf{B_d}\mathbf{W_d}\mathbf{x_L}(k)$ and $e(k)_r$ is the rth element of vector $\mathbf{e}(k)$.

if $k < N$ set $k = k + 1$ and go back to step 6.

Step7: Calculate confidence limits of the fifteen fault detection statistics by KDE (99% or 95%) to acquire the control limit.

4.2   The online part

Step1: Construct the new augmented observation vector and it is normalized by the offline part mean and standard deviation of each row as in step 1 of the offline part to obtain $\mathbf{x}_{online}(k)$.

Step2: Compute vectors $\mathbf{s_{online}}(k)$ and $\mathbf{s_{eonline}}(k)$ by prior $\mathbf{W}_d$ and $\mathbf{W}_e$ as:

$$\mathbf{s_{don}}(k) = a\mathbf{W_d}\mathbf{x_{online}}(k) \qquad (27)$$

$$\mathbf{s_{eon}}(k) = a\mathbf{W_e}\mathbf{x_{online}}(k) \qquad (28)$$

Step3: Compute the fifteen fault detection statistics as in step 6 of the offline part for the new vector of ICs as:

- Calculate the first, second, third, and fourth moment statistics and the first characteristic function statistic at frequency ω_0 for the elements of ICs vector $\mathbf{s_{don}}(k)$, $I\phi_{ond\omega_0}$ and $I\mu'_{i\,ond}(k)$ $i = 1, 2, 3, 4$, as follows:

$$I\mu'_{ion_d}(k) = |\mu'_{i\,ond}(k)| \; i = 1, \ldots, 4 \qquad (29)$$

$$I\phi_{ond\omega_0}(k) = |1 - |\phi_{ond\,\omega_0}(k)|| \qquad (30)$$

Where

$$\mu'_{i\,ond}(k) = \frac{1}{d} \sum_{r=1}^{d} (s_{don}(k)_r)^i \; i = 1, \ldots, 4 \qquad (31)$$

$$|\phi_{ond\,\omega_0}(k)| = \left| 1 + j\mu'_{1\,ond}(k)\omega_0 + \mu'_{2\,ond}(k)\frac{(j\omega_0)^2}{2} \right.$$
$$\left. + \mu'_{3\,ond}(k)\frac{(j\omega_0)^3}{3} + \mu'_{4ond}(k)\frac{(j\omega_0)^4}{4} \right| \qquad (32)$$

Which $s_{don}(k)_r$ is the rth element of vector $\mathbf{s_{don}}(k)$.

- Calculate the first, second, third, and fourth moment statistics and first characteristic function statistic at frequency for the elements of ICs vector $\mathbf{s_{eon}}(k)$, $I\phi_{one\omega_0}(k)$ and $I\mu'_{ione}(k)$ $i = 1, \ldots, 4$

$$I\mu'_{i\,one}(k) = |\mu'_{i\,one}(k)| \; i = 1, \ldots, 4 \qquad (33)$$

$$I\phi_{one\omega_0}(k) = |1 - |\phi_{one\omega_0}(k)|| \qquad (34)$$

Where

$$\mu'_{i\,one}(k) = \frac{1}{m(l+1) - d} \sum_{r=1}^{m(l+1)-d} (s_{eon}(k)_r)^i$$
$$i = 1, \ldots, 4 \qquad (35)$$

$$|\phi_{one\omega_0}(k)| = \left| 1 + j\mu'_{1\,one}(k)\omega_0 + \mu'_{2\,one}(k)\frac{(j\omega_0)^2}{2} \right.$$
$$\left. + \mu'_{3\,one}(k)\frac{(j\omega_0)^3}{3} + \mu'_{4\,one}(k)\frac{(j\omega_0)^4}{4} \right|$$
$$(36)$$

Which $s_{eon}(k)_r$ is the rth element of vector $\mathbf{s_{eon}}(k)$.

- Similar to the aforementioned procedure, the first four moment statistics and the first characteristic functions statistic at frequency ω_0 for the elements of vector $\mathbf{e}_{on}(k)$, $I\phi_{onms\,\omega_0}(k)$ and $I\mu'_{i\,omns}(k)$ $i = 1, 2, 3, 4$, need to be calculated (For the non-systematic part) as:

$$I\mu'_{i\,onns}(k) = |\mu'_{ionns}(k)| \quad i = 1, \ldots, 4 \qquad (37)$$

$$I\phi_{onns\omega_0}(k) = |1 - |\phi_{onns\omega_0}(k)|| \qquad (38)$$

Where

$$\mu'_{i\,onns}(k) = \frac{1}{m(l+1)} \sum_{r=1}^{m(l+1)} (e_{on}(k)_r)^i \quad i = 1, \ldots, 4$$
$$(39)$$

$$|\phi_{onns\omega_0}(k)| = \left| 1 + j\mu'_{1\,onns}(k)\omega_0 + \mu'_{2\,onns}(k)\frac{(j\omega_0)^2}{2} \right.$$
$$\left. + \mu'_{3\,onns}(k)\frac{(j\omega_0)^3}{3} + \mu'_{4\,onns}(k)\frac{(j\omega_0)^4}{4} \right|$$
$$(40)$$

where $\mathbf{e}_{on}(k) = \mathbf{x_{online}}(k) - \hat{\mathbf{x}}_{online}(k)$ and $\hat{\mathbf{x}}_{online}(k) = \mathbf{Q}^{-1}\mathbf{B_d}\mathbf{W_d}\mathbf{x_{online}}(k)$ and $e_{on}(k)_r$ is the rth element of vector $\mathbf{e}_{on}(k)$.

If these new fault detection statistics are out of the control limit, a fault has occurred.

Step4: Detect the root cause of the fault by calculating contribution's values of each sample in fault detection statistics.

4.3   Root cause diagnosis

Root cause diagnosis is as important as fault detection and needs to be done before the fault spreads through all loops and develops to a crucial level. In the ICA-based fault detection methods, the dependence between the ICs, in each vector, is minimized. As

a result, the ICs which lead to increase in the fault detection statistics can be identified. In other words, contribution of the ICs in the value of statistics, by which the fault is being detected, can be determined which ICs is root cause of the fault. In other words:

1. If the first moments detect the fault in sample k, the ICs with bigger absolute value in sample is the root cause of the fault.

2. If the second moments detect the fault in sample k, the ICs that have bigger squared value in sample k is the root cause of the fault.

3. If the third moments detect the fault in sample k, the ICs with bigger cubed value in sample k is the root cause of the fault.

4. If the forth moments detects the fault in sample k, the ICs with bigger ICs to the power of four value in sample k is the root cause of the fault.

5. If the first characteristic function detects the fault in sample k, the ICs sample that have bigger value in Equations (18) or (22) in sample k is root cause of the fault.

Notice that all absolutely integrable functions of two random variables became independent if these two random variables are independent, also expectation have linearity property.

Next step, after finding the ICs that is the root cause of the fault, is to find the loop which made the ICs. It is of crucial importance to find the loop without disturbing the system normal operation. Knowledge of the system properties is the key to answer this question. For example, the loop can be found if the operator can detect a characteristic, such as frequency, that shows which loop made which ICs in the offline part.

4.4 Analysis of the simultaneous application of the moments and the first characteristic function

The first characteristic function statistic is not sufficient for fault detection because:

- The faults will be detected when the first characteristic function statistic in the online part differs from that of the offline part. In certain circumstances, however, increase in some terms and decrease in other terms of the first characteristic function cancel out each other's effect on the first characteristic function and hence the first characteristic function statistic does not change significantly. As a result, the faults cannot be detected.

- The first characteristic function statistic is a function of frequency ω. Hence, the first characteristic function statistic changes if the value of ω changes. In other words, two completely different Fourier transform functions can have identical values in some frequencies. Thus, faults will not be detected when, at one chosen frequency, the first characteristic function in the online part is equal to that of the offline part.

The first four moment statistics are not sufficient for fault detection because:

1. The moments-based fault detection methods fail to detect the occurrence of the faults if all four moments are inside the control limit. To improve the effectiveness of the fault detection methods in such situations, the first characteristic function along with the first four moments can be used as statistics to determine the control limit. The reason is that, according to Equations (18) and (22), the first characteristic function will not reach its maximum if all four first moments are at their maximum values. Therefore, the maximum value of the first characteristic function in the offline part is defined as new constraint to determine the control limit.

4.5 Limitation of the proposed method

While the proposed method is powerful in detecting a class of faults which the former methods cannot detect, it has two limitations:

2. If the faulty ICs vector in the online part have elements with values greater than $\lambda s_{\max\ offline}$, it is possible that constraints of inequalities (B-3) and (B-4) not being satisfied. As a result, the effects of higher order moments overweight the effects of lower order moments in first characteristic function estimation and hence fault detection may fail. In such circumstances, the effectiveness of the method can be improved by choosing a small frequency ω_0. However, that leads to small elements in the ICs vector and therefore reduces the calculation accuracy.

3. Applying the proposed method in a short time requires more powerful computational facilities.

5. SIMULATION RESULTS

The former and proposed fault detection methods are applied to two systems and the simulation results are used to validate the algorithm.

5.1 The simple system

The proposed fault detection method based on DICA with new statistics, described in section 4, is applied to the system suggested by [39] and the simulation results are used to validate the algorithm.

The system is comprised of two interfering loops and hence faults propagation is inevitable. Each loop consists of a plant and a PI controller. To model the sensor fault, a block with a limited time output is added to the lower loop. In DICA-based fault detection methods, oscillations that occur in both offline and online parts cannot be detected by fault detection statistics. It, in basically, means that faults that occur in the offline part, outliers, cannot be detected in the online part because it is assumed as normal condition. To show such a situation, two sinusoidal signals are added to the feedback signals. Both signals are of amplitude

1 and their frequencies are set to 0.7Hz for the lower loop and to 5Hz for the upper loop. Also a white noise with power of approximately 0.001 is added to the upper loop. Step signals of amplitude 10 are chosen as input signals. The simulation is run for 200 seconds with sample time of 0.05 which means 4000 samples (200/0.05) are accumulated. The first 2000 samples are discarded as they represent the transient response and the last 5 samples are excluded to make the augmented observation matrix. Thus, about 2000 samples are shown in the figures. About half of these samples are chosen as the offline part signals which are shown with red dots in Figures 2, 3, 4, 5 and the rest is chosen as the online part signals represented by blue dots in these figures. Occurrence of two faults is assumed; sensor fault and actuator fault. The sensor fault amplitude is chosen to be 0.025 of the feedback signal tolerance amplitude. This fault occurs between 155 and 160 seconds which corresponds to samples from 1100 to 1200; notice that the first 2000 samples are omitted. In order to model the actuator fault, a random exponential pdf fault is used as an additive distortion to the control signal in the upper loop. The amplitude of this fault is chosen about 0.02 of the control signal. This fault occurs between 170 and 175 seconds which corresponds to samples 1400 to 1500. Lag time of one ($l = 1$) is chosen to construct the augmented matrix. The observation matrix is constructed by both controllers and system outputs. In Figures 2, 3, 4 and 5 only dominant ICs are shown.

5.1.1 Performance of first characteristic function and its terms in comparison with square of 2-norm as statistics to determine the control limit

In this section performance of the proposed method is compared with that of the former method in which square of 2-norm is used to determine the control limit. In Figure 2, parts a, b, c, and d show, respectively, the first, second, third and fourth moment statistics of the ICs and Part e shows the first characteristic function statistic of the ICs. The values of ω_0 and λ are chosen as 1 and 10, respectively. Part f presents the square of 2-norm of ICs as the former statistic. Parts g and h are two control signals sampled from system and Part i shows the joint distribution of the dominant ICs in the online part. In this part the normal operation ICs are shown by red dots and the faulty ICs are shown by green squares. Fault(s) will be detected if amplitude of each statistic exceeds the value of corresponding statistic in the normal condition that made control limit. It means fault/faults will be detected if some samples of either of statistics, the first four moment statistics and the first characteristic function statistic in the proposed method, or square of 2-norm in the former method, are bigger than the same statistics in the offline part obtained from normal operation condition. A dashed line represents the maximum value of statistics in the online part.

The first, second and the fourth moment statistics, Part a, b and d, although not easily, detect the sensor fault but fail to detect the actuator fault. Whereas, the third moment statistic, Part c, detects both sensor and actuator faults. Also, the first characteristic function statistic only detects the sensor fault, Part e. On the other hand, it is clear that the former method detects none of those faults, see Part f. Part i clearly shows that the faulty ICs, green squares, are located among the normal ICs shown by red dots. Also it is noticeable that the faults do not have a significant influence on the first and second control signals; see

Parts g and h. In other words, the faults cannot be detected by monitoring the signals that are not processed by the proposed fault detection method. The Simulation results emphasis that the proposed method is a more effective fault detection method than the former method.

5.1.2 The root cause diagnosis

Simulation results are used to verify the ability of the proposed method in root cause diagnosis. For better visibility, the amplitudes of both faults are increased by 200%, the actuator fault is shifted by 10 seconds. In other words, the actuator fault is assumed to occur between 180 and 185 seconds which correspond to samples from 1600 to 1700.

The results are shown in Figure 3. In this Figure, Parts a, d, g, j and m are, respectively, the first four moments statistics and the first characteristic function statistic of the ICs. Parts b, e, h, k and n are, respectively, contribution values of the first dominant part of ICs, corresponding to the control signal of upper loop that has higher frequency, in the first four moments statistics and the first characteristic function statistic. Also Parts c, f, i, l and o are, respectively, contribution values of the second dominant part of ICs, corresponding to the control signal of lower loop that has lower frequency, in the first four moments statistics and the first characteristic function statistic.

It is clear that the sensor fault that occurs in the lower loop is detected by all five statistics about sample 1200, see Parts a, d, g, j and m. And the root cause of this fault is detected in the second dominant part of ICs, see Parts c, f, i, l and o. Comparing the offline parts of c and b shows that the second dominant part of ICs has a lower frequency than the first dominant part of ICs. Since, it is known, from the system characteristics, that the lower loop has a lower frequency, it can be concluded that the second dominant part of ICs corresponds to the lower loop; hence the sensor fault is occurred in the lower loop.

Also the actuator fault that occurs in the upper loop is detected by all five statistics about sample 1650, see Parts a, d, g, j and m; although, the occurrence of this fault is not clear in the Part j. Parts b, e, h, k and n clearly show that the root cause of the actuator fault is in the first dominant part of ICs. Comparing the offline parts of c and b shows that the first dominant part of ICs has a higher frequency than the second dominant part of ICs. The system characteristics show that the upper loop has a higher frequency; as a result, the first dominant part of ICs corresponds to the upper loop. Therefore the actuator fault is occurred in the upper loop.

5.1.3 Comparison between the first and second characteristic functions

In this section the effects of using the first characteristic function statistic and the first four moment statistics on the performance of the fault detection algorithm is compared with that of the second characteristic function statistic and the first four cumulant statistics. Absolute value of $1 - |\psi_{\omega_0}(k)|$ and absolute value of first four cumulants are, respectively, called second characteristic function statistic and cumulant statistics, see Equations (15) and (16). The amplitude of the sensor fault and the actuator fault are, respectively, decreased to 0.001 of the tolerance amplitude of the feedback signal and to 0.0003 of the control signal. Also,
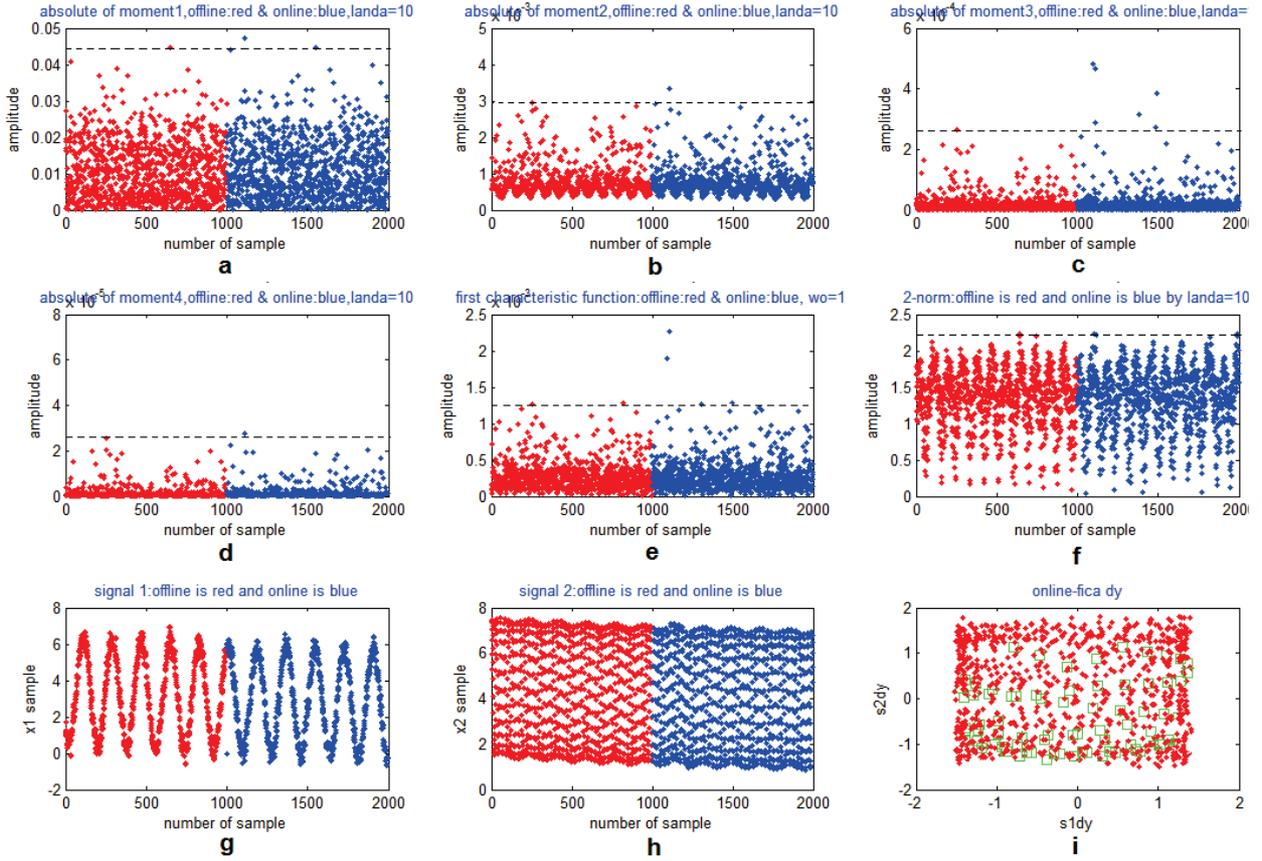
Figure 2 Comparison between the proposed control limit indices and the former control limit. In Parts a, b, c, d, e, f, g and h the offline part samples are shown with red dots and the online part samples represented by blue dots. In Part i, the normal operation ICs are shown by red dots and the faulty ICs are shown by green squares.

λ is set to 2.

In Figure 4, Parts a, c, e, and g are, respectively, the first, second, third, and fourth moment statistics of ICs and Part i presents the first characteristic function statistic of the ICs. Parts b, d, f, and h are absolute values of the first, second, third and fourth cumulant of ICs, respectively and Part j shows the second characteristic function of ICs.

Parts a and b are identical. However, in Part c the amplitude of the faulty statistics is slightly bigger than the statistics in Part d and exceeds the corresponding statistics maximum in the offline part (dashed line). Comparing Parts e and g with f and h, respectively, shows that the difference in the amplitude of faulty statistics becomes more obvious as the order of moments and cumulants increases. It is obvious that by using the moments as the statistics to determine the control limits, fault detection method is able to identify both faults, whereas by using the cumulant as the statistic to determine the control limit, the fault detection method only detects the sensor fault. Thus it is obvious that detecting faults by moments yields much better result than by cumulants. Moreover, comparison between Parts i and j reveals the benefit of using the first characteristic function over the second characteristic function.

5.1.4 Effect of frequency ω_0 and λ on the first characteristic function statistic to determine the control limit

In this section, the effect of frequency ω_0 and λ on the performance of first characteristic function statistics is investigated. In

Figure 5, the values of frequency ω_0 in first column, parts a, e, i and m, is chosen 0.33 and in second column, parts b, f, j and n, is chosen 1 and in third column, parts c, g, k and o, is chosen 10 and in forth column, parts d, h, l and p, is chosen 100. the values of in first row, parts a, b, c and d, is chosen 2 and in second row, parts e, f, g and h, is chosen 20 and in third row, parts i, j, k and l, is chosen 200 and in forth row, parts m, n, o and p, is chosen 500. Figure 5 clearly illustrates that the first characteristic function statistic values change when the faults occur. It can be seen, moving in each row from left to right in Figure 5, that with a fixed λ, choosing a bigger ω_0 leads to a bigger change in amplitude of first characteristic function statistics and hence a smaller round of error. Also Figure 5, parts c, d and h, shows that choosing λ and ω_0 in a way that inequality $\omega_0 < \lambda$ is not satisfied, yield to first characteristic function estimation with amplitude greater than 1 which is an invalid estimation because

$$0 < |\phi(w)| = \left| \int_{-\infty}^{+\infty} e^{js\,\omega} f(s)ds \right| \leq \int_{-\infty}^{+\infty} |e^{js\omega}||f(s)|ds = 1$$

$$\Leftrightarrow -1 \leq \phi(\omega) \leq 1.$$

(41)

On the other hand, moving upward in each column of figure 5, choosing a smaller λ while keeping the ω_0 fixed, results in a bigger change in amplitude of the first characteristic function statistics and thus a decrease in the round of error. A brief review of Figure 5 emphasizes that choosing λ bigger but close to ω_0, leads to more observable changes in the amplitude of
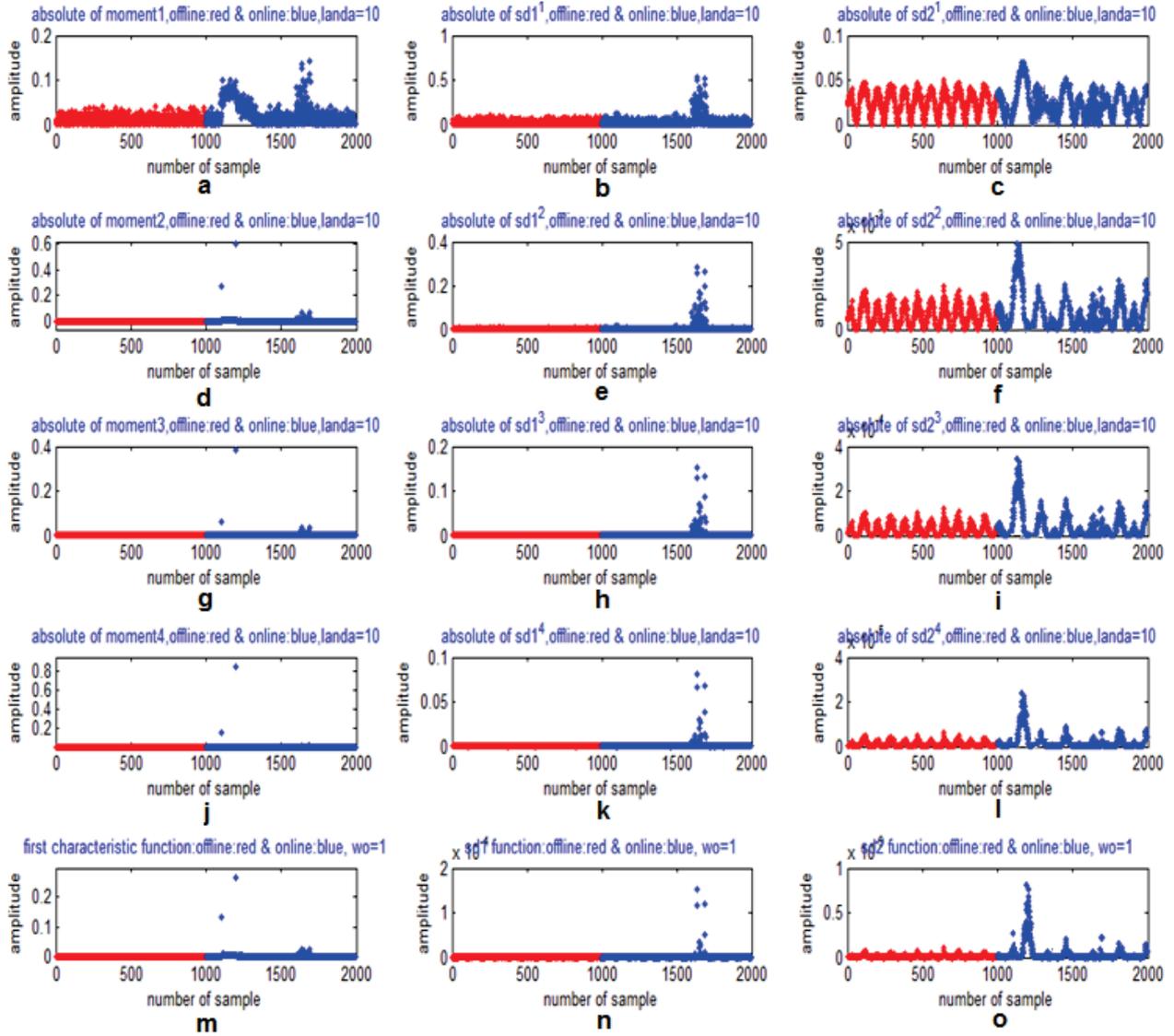
Figure 3 The root cause detection of the fault in each statistics of the control limit. In Parts $a, b, c, d, e, f, g, h, i, j, k, l, m, n$ and o the offline part samples are shown with red dots and the online part samples represented by blue dots.

first characteristic function statistics and hence smaller round of error.

In summary, first ω_0 need to be selected, not too large value to lead to a decrease in the value of moments statistics as λ should be a supremum for ω_0. Then λ should be chosen bigger than but close to ω_0 to decrease the round of error.

5.2 The Couple tank system

A simple nonlinear model of the coupled tank is provided in Ko et al. as [40]:

$$\frac{dH_1}{dt} = \frac{Q_{i1}}{A_1} - \frac{\alpha_1}{A_1}\sqrt{H_1} - \frac{\alpha_3}{A_1}\text{sgn}(H_1 - H_2)\sqrt{|H_1 - H_2|}$$
(42)

$$\frac{dH_2}{dt} = \frac{Q_{i2}}{A_2} - \frac{\alpha_2}{A_2}\sqrt{H_2} - \frac{\alpha_3}{A_2}\text{sgn}(H_2 - H_1)\sqrt{|H_1 - H_2|}$$
(43)

Where H_1, H_2 are the height of fluid in tank 1 and 2, respectively. Q_{i1} and Q_{i2} are the pump flow rate into tank 1 and 2, respectively. α_1, α_2 and α_3 are constants. A_1, A_2 are the cross sectional area of tank 1 and 2, respectively. Equations (42) and (43) make a system with two interfering loops, thus the fault propagates through the loops. In order to decrease the error, a lag controller is designed for each loop; $G_{C1} = \frac{s+0.9}{s+0.002}$ and $G_{C2} = \frac{s+0.8}{s+0.001}$ for tank 1 and 2, respectively.

Two sinusoidal signals are added to output signals in order to model the waves on the surface. The amplitude of the sinusoidal signal is 0.03 and its frequency is 0.2Hz in the control loop of tank 1. The amplitude of sinusoidal signal in the control loop of tank 2 is 0.8 and its frequency is 0.5Hz. Also, a white noise with the power of 0.0002 is added to each loop. Step input signals with amplitude 17 and 15 are chosen as desired height of fluid in tank 1 and 2, respectively. $\alpha_1 = 10.78$, $\alpha_2 = 11.03$ and $\alpha_3 = 11.03$ are set and $A_1 = 3m^2$, $A_2 = 32m^2$ are chosen.

The simulation runs for 200 seconds with a sample time of 0.05 which means 4000 samples (200/0.05) are accumulated.
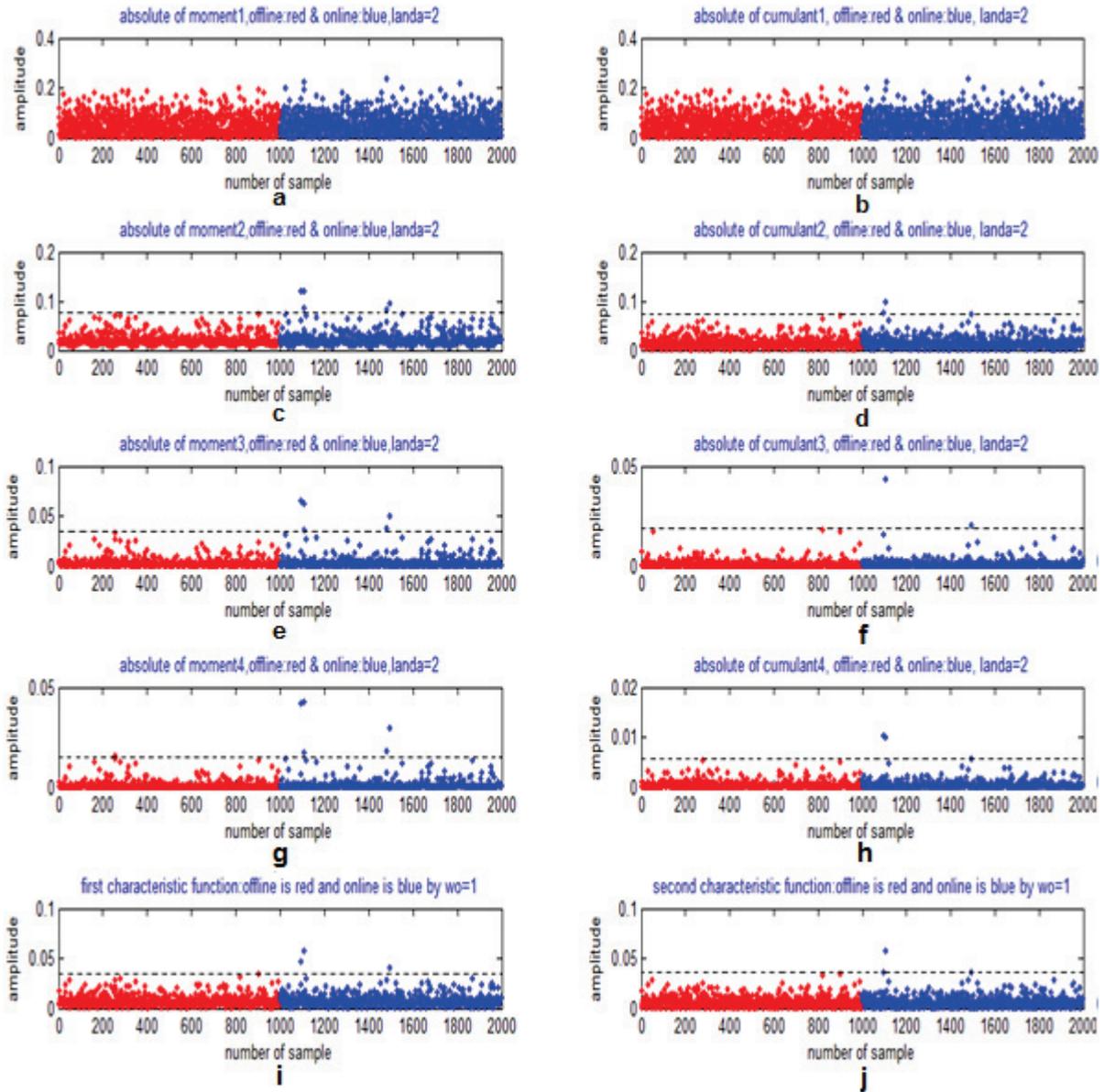
Figure 4 Comparison between the first and second characteristic functions and their terms. In Parts a, b, c, d, e, f, g, h, i and j the offline part samples are shown with red dots and the online part samples represented by blue dots.

The first 2000 samples are neglected as they represent the transient response and the last 5 samples are excluded to make the augmented observation matrix. Thus, less than 2000 samples are shown in the Figures 7 and 8.

5.2.1 Comparison Results

To model the stochastic fault of level sensor in tank 2, three random numbers are added to the feedback signal of tank 2 control loop. The first, second and third sensor fault occurs between samples 1150 to 1180 with exponential distribution, 1500 to 1530 with uniform distribution and 1850 to 1880 with Gama distribution, respectively. First 1000 samples are chosen as the offline part signals which are shown with red dots in Figures 7 and 8, representing the system normal performance. The rest of the samples is chosen as the online part and is represented by blue dots in these figures. KDE (%99) is used to calculate the confidence bound of control limit which is shown by dashed line

in Figure 7, parts a, b, c, d, e and f. The values of $l = 2$ is chosen in both the former and the proposed fault detection methods. The observation matrix is constructed by error signals.

Figure 7 part f shows the 2-norm fault detection statistics in the online and the offline part, obtained from the former fault detection method. It is clear that the square of 2-norm value of faulty IC vectors, blue dots between samples 1850 to 1880, do not exceed from normal performance area, the dashed line, in the former fault detection method. Figure 7, parts a, b, c, and d show, respectively, the first, second, third and fourth moment statistics of the ICs and Part e shows the first characteristic function statistic of the ICs. The values of ω_0 and λ are chosen as 1 and 2, respectively.

The first moment statistic, Part a, only detects the first fault. The second and the fourth moment statistics, Parts b and d, detect the first and second faults. Whereas, the first characteristic function statistic and third moment statistic detect all three sen-
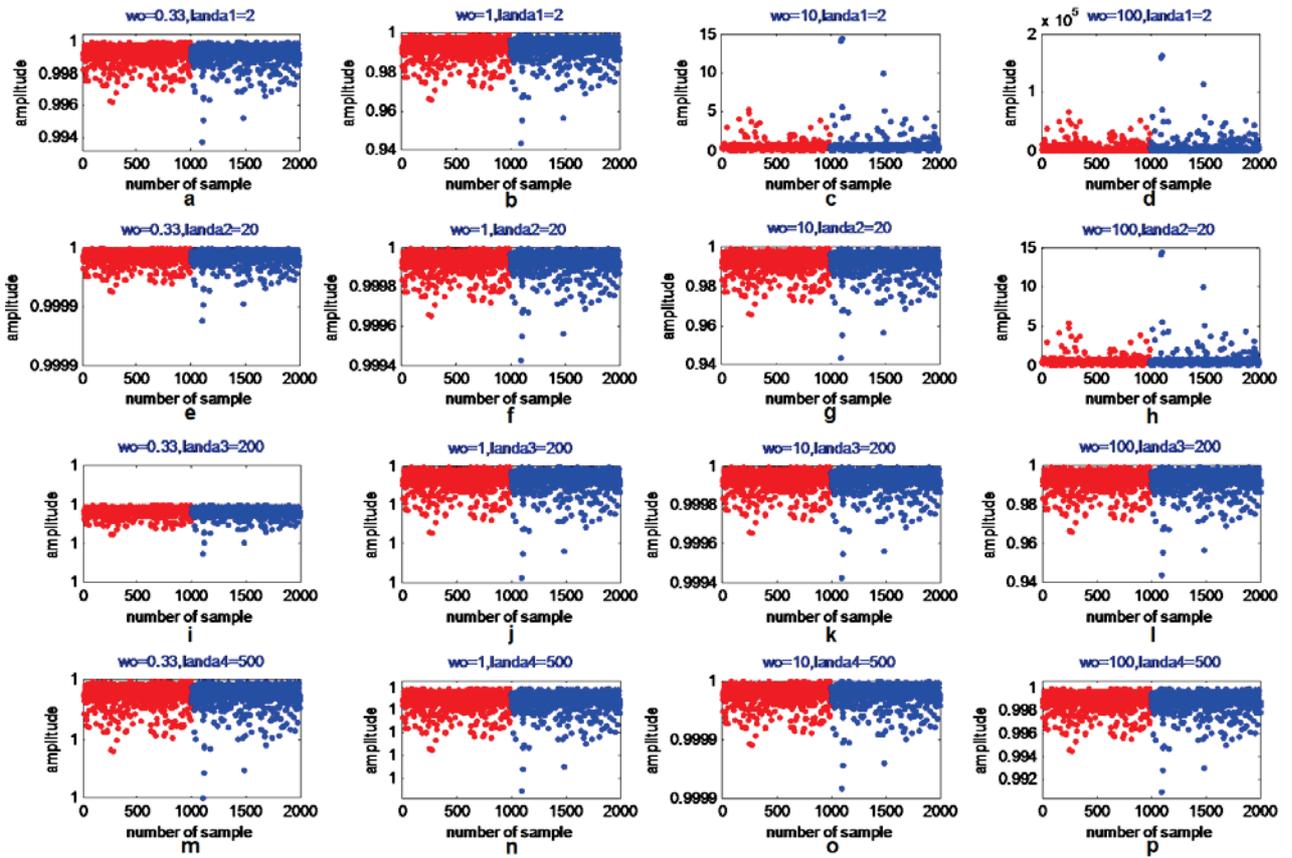
Figure 5 Effect of frequency and λ on the first characteristic function as the statistic of the control limit. In Parts *a* to *o* the offline part samples are shown with red dots and the online part samples represented by blue dots.
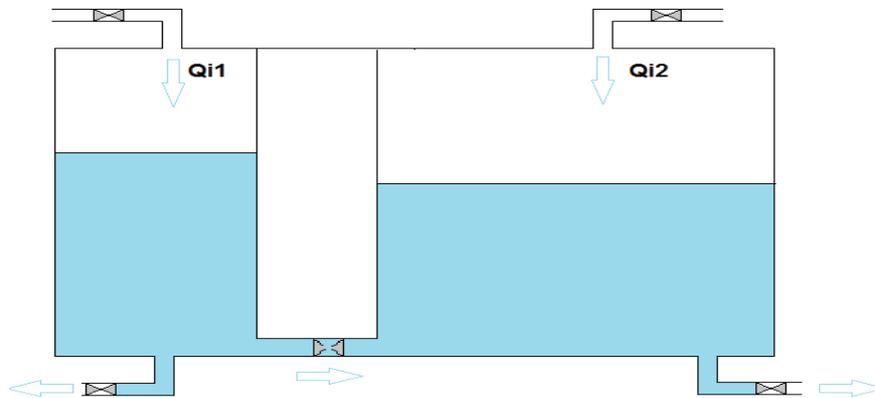
Figure 6 The couple tanks System.

sor faults, Parts c and e. It is worth mentioning that the proposed method detect the faults if any of the five proposed statistics detect the fault.

This simulation is run 30 times without any changes in the parameters and the proposed method detects 84 faults from 90 faults that occurred, whereas the former method detects 76 faults from 90 faults that occurred.

The simulation results emphasis that the proposed method is a more effective fault detection method than the former method.

5.2.2 The root cause diagnosis

Simulation results are used to verify the ability of the proposed method in root cause diagnosis. For better visibility, the amplitudes of all faults are increased by 250%. The results are shown in Figure 8.

In this Figure, Parts a, b, c, d and e are, respectively, the first four moments statistics and the first characteristic function statistic of the ICs, and all of them detected the three faults. Part f
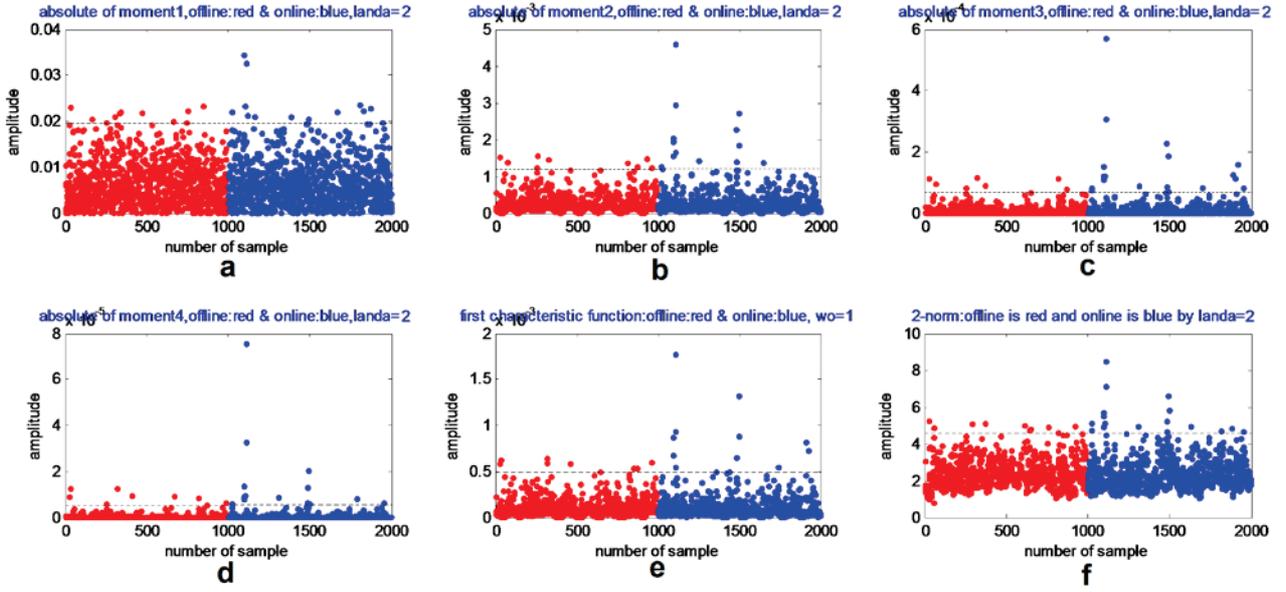
Figure 7 Comparison between the proposed control limit indices and the former control limit (square of 2-norm). The offline part samples are shown with red dots and the online part samples represented by blue dots.

indicates the fault signals. Part g and h are, respectively, absolute values of the first and second dominant parts of ICs. It is clear that absolute values of second ICs have increased in samples where fault is being detected, see part h. As a result, the second ICs are the cause of increase in the fault detection statistics. Basically if it can be identified that tank2 has oscillations with higher frequency in the offline part, one can detect the root cause as the graph in part h has higher frequency.

6.   CONCLUSIONS

This paper has introduced a fault detection method using the DICA with new statistics to determine the control limit. The benefit of the proposed algorithm is the increase in fault detection ability by estimating the first characteristic function and four first moments to determine the control limit indices. It has shown that the proposed method can detect a class of faults that their pdf, providing that its change can be estimated by low order moments of the first characteristic function, or lower moments are different from those of the normal ICs. In contrast to the ICA-based fault detection methods which use square of 2-norm as statistic to determine the control limit, the proposed approach can efficiently detect faulty samples although they are among normal samples. Simulation results are used to verify the effectiveness of the proposed methodology.

7.   APPENDIX

7.1   Appendix A

Propagation of error is calculated by the method described in [41, 42]. A common formula among engineers and experimental scientists to calculate propagation of error, the variance formula (see [43]) is Equation (A-1). For a function of i uncorrelated quan-

tities, $\kappa = f(\mu'_1, \mu'_2, \ldots, \mu'_i)$, or by neglecting the coloration between quantities the propagation of error can be calculated as:

$$\Delta \kappa_i = \sqrt{\left(\frac{\partial f}{\partial \mu'_1} \Delta \mu'_1\right)^2 + \left(\frac{\partial f}{\partial \mu'_2} \Delta \mu'_2\right) + \cdots + \left(\frac{\partial f}{\partial \mu'_1} \Delta \mu'_i\right)}$$

(A-1)

Where $\Delta \kappa_i$ is propagation of error in i^{th} cumulant (κ_i) and $\Delta \mu'_i$ is error in i^{th} moment (μ'_i) ($\Delta \mu'_i$ consist of the error due to calculation of moment and the error due to measurement of signals), providing $\Delta \mu'_i$ and $\Delta \mu'_j$ $i \neq j$ are uncorrelated.

The Equation (7) shows that in calculation of κ_i, always the first term of right hand side of equation is μ'_i. Therefore in Equation (A-1) the term under root is always begin with $\left(\frac{\partial f}{\partial \mu'_i} \Delta \mu'_i\right)^2$ and the other terms are nonnegative. Hence

$$\Delta \kappa_i = \sqrt{(\Delta \mu'_i)^2 + a^2}$$

Where a^2 is positive term. Therefore $\Delta \kappa_i \geq \Delta \mu'_i$.

This shows that propagation of errors in cumulants are bigger than the errors of moment when $i > 1$. It is worth mentioning that there is not any propagation of error in estimation of moments because they are computed directly.

7.2   Appendix B

By rearranging Equations (10) and substituting $\phi(\omega)$ from Equation (3) the remainder can be calculated as follows [31]:

$$R(\omega) = \phi(\omega) - \sum_{i=0}^{4} \mu'_i \frac{(j\omega)^i}{i} = \int_{-\infty}^{+\infty} \left(e^{js\omega} - \sum_{i=0}^{4} \frac{(js\omega)^i}{i}\right) f(s) ds$$

$$= \int_{-\infty}^{+\infty} \left(\cos(s\omega) + j \sin(s\omega) - \sum_{i=0}^{4} \frac{(js\omega)^i}{i}\right) f(s) ds$$
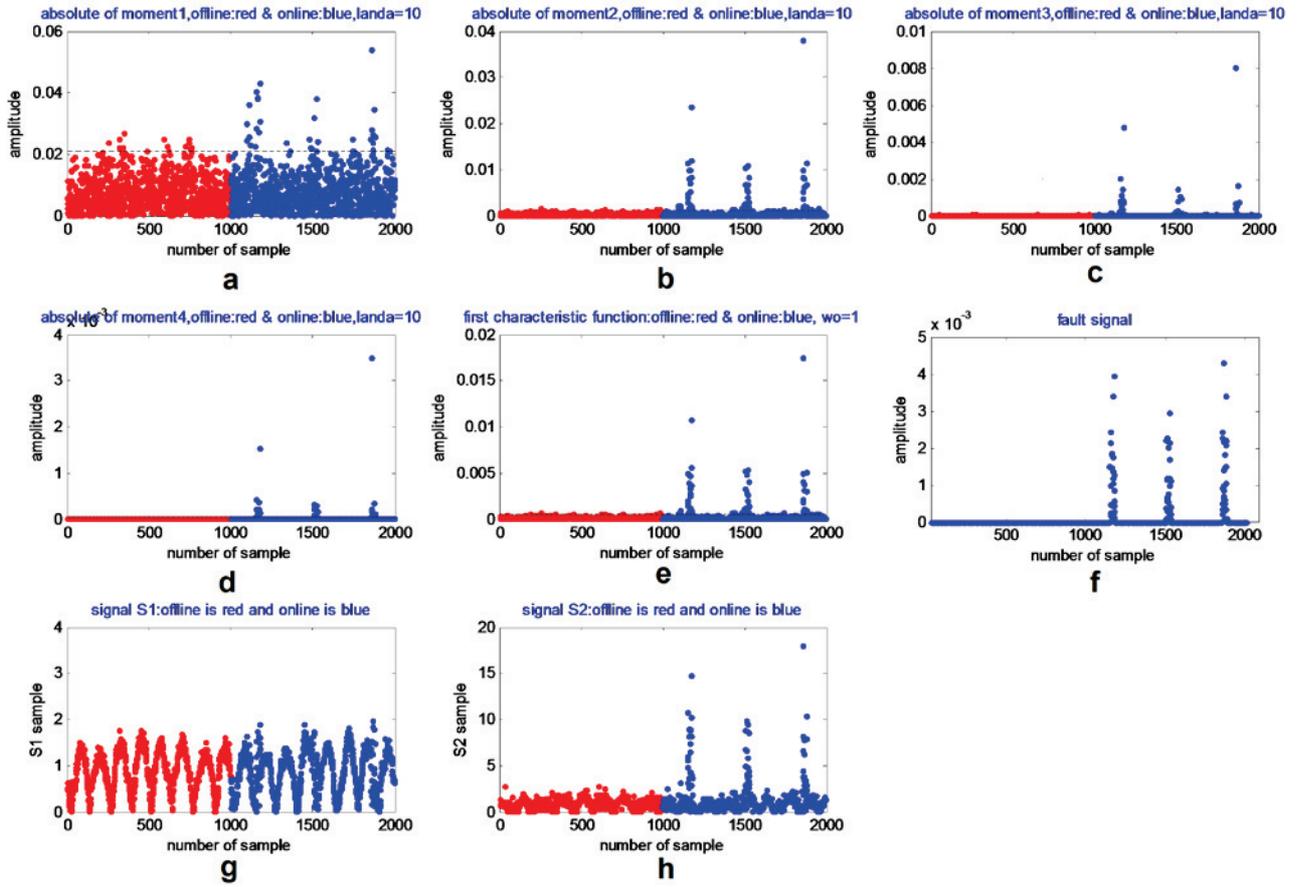
Figure 8 Root cause detection of the fault in each statistics of the control limit.

Equation (B-1) can be rewriting as:

$$
\begin{aligned}
R(\omega) &= \int_{-\infty}^{+\infty} \left(\left(\sum_{i=0}^{\infty} \frac{(js\omega)^{2i}}{2i} + \sum_{i=1}^{\infty} \frac{(js\omega)^{2i-1}}{(2i-1)} \right) \right. \\
&\quad \left. - \left(\sum_{i=0}^{2} \frac{(js\omega)^{2i}}{2i} + \sum_{i=1}^{2} \frac{(js\omega)^{2i-1}}{(2i-1)} \right) \right) f(s)ds \\
&= \int_{-\infty}^{+\infty} \left(\left(\sum_{i=3}^{\infty} \frac{(js\omega)^{2i}}{2i} + \sum_{i=3}^{\infty} \frac{(js\omega)^{2i-1}}{(2i-1)} \right) \right) f(s)ds \\
&= \int_{-\infty}^{+\infty} \left(\sum_{i=3}^{\infty} (-1)^i \frac{(s\omega)^{2i}}{2i} \right) f(s)ds \\
&\quad + j \int_{-\infty}^{+\infty} \left(\sum_{i=3}^{\infty} (-1)^{i-1} \frac{(s\omega)^{2i-1}}{(2i-1)} \right) f(s)ds \\
\Rightarrow |R(\omega)| &\leq \left| \int_{-\infty}^{+\infty} \left(\sum_{i=3}^{\infty} (-1)^i \frac{(s\omega)^{2i}}{2i} \right) f(s)ds \right| \\
&\quad + \left| j \int_{-\infty}^{+\infty} \left(\sum_{i=3}^{\infty} (-1)^{i-1} \frac{(s\omega)^{2i-1}}{(2i-1)} \right) f(s)ds \right| \\
&\leq \int_{-\infty}^{+\infty} \left| \left(\sum_{i=3}^{\infty} (-1)^i \frac{(s\omega)^{2i}}{2i} \right) f(s)ds \right| \\
&\quad + \int_{-\infty}^{+\infty} \left| j \left(\sum_{i=3}^{\infty} (-1)^{i-1} \frac{(s\omega)^{2i-1}}{(2i-1)} \right) \right| f(s)ds
\end{aligned}
$$

Considering inequalities (B-3) and (B-4):

$$
\left| \sum_{i=3}^{\infty} (-1)^i \frac{(s\omega)^{2i}}{2i} \right| \leq \left| - \left(\frac{(s\omega)^4}{4} \frac{\partial^2 (\cos s\omega)}{\partial (s\omega)^2} + \frac{(s\omega)^4}{4} \right) \right|
$$

$$
\leq 2\frac{(s\omega)^4}{4} \text{ if } s\omega < 1
$$

(B-3)

$$
\left| \sum_{i=3}^{\infty} (-1)^i \frac{(s\omega)^{2i-1}}{2i-1} \right| \leq \left| - \left(\frac{(s\omega)^4}{4} \frac{\partial^2 (\sin s\omega)}{\partial (s\omega)^2} \right) \right|
$$

$$
\leq 2\frac{(s\omega)^4}{4} \text{ if } s\omega < 1 \quad \text{(B-4)}
$$

Equation (B-2) can be expressed as:

$$
\begin{aligned}
|R(\omega)| &\leq \int_{-\infty}^{+\infty} \left(\left(2\frac{(s\omega)^4}{4} \right) + \left(\frac{(s\omega)^4}{4} \right) \right) f(s)ds \\
&\leq 3 \int_{-\infty}^{+\infty} \frac{(s\omega)^4}{4} f(s)ds
\end{aligned}
$$

Substituting $s_{\max}$, the biggest s found in the offline part, in Equation (B-5) results in:

$$
\begin{aligned}
|R(\omega)| &\leq 3 \int_{-\infty}^{+\infty} \frac{(s\omega)^4}{4} f(s)ds \\
&\leq \frac{3}{4}\omega^4 s_{\max}^4 \int_{-\infty}^{+\infty} f(s)ds = 0.125\omega^4 s_{\max}^4
\end{aligned}
$$

Consequently, Inequality (B-6) shows that $0.125\omega^4 s_{\max}^4$ is a supremum for the summatory of high order moments.

REFERENCES

1. Wei, Y, Songshan, S, Zhenggang, L and Guoxiang, L, (2015). Investigation on remote monitoring system for hybrid electric bulldozer. International Journal of Computer Systems Science and Engineering. 30(5).

2. Mai, H T, Lee, K Y and Kim, M H, (2015). Dynamic filtering of sensor readings for approximate query processing in sensor networks. International Journal of Computer Systems Science and Engineering. 30(3).

3. Wireman, T (2008). Preventive Maintenance. Industrial Press.

4. Sahin, I and Polatoglu, H (1998). Quality, Warranty and Preventive Maintenance. International Series in Operations Research & Management Science. Springer Science and Business media New York.

5. Patton, J, D (1995). Preventive Maintenance. International Society for Measurement and Control.

6. Levitt, J (2003). Complete Guide to Preventive and Predictive Maintenance. Industrial Press.

7. Hathaway, L and Buckingham, F (1992). Preventive Maintenance. Deere & Co.

8. Thornhill, N, F, Horch, A (2007). Advances and new directions in plant-wide disturbance detection and diagnosis. Control Engineering Practice, 15(10) 1196–1206.

9. Martin, E, B and Morris, A, J (1996). Non-Parametric Confidence Bounds for Process Performance Monitoring Charts. Journal of Process Control 6(6) 349–358.

10. Seungdo, J, Sang-Wook, K and Byung-Uk, C, (2016). Effective indexing and searching with dimensionality reduction on high-dimensional space. International Journal of Computer Systems Science and Engineering. 31(4)

11. Lee, J, M, Yoo, C, K and Lee, I, B (2004). Statistical Process Monitoring With Independent Component Analysis. Journal of Process Control. 14(5) 467–485.

12. Lee, J, M, Yoo, C, K and Lee, I, B (2004). Statistical monitoring of dynamic processes based on dynamic independent component analysis. Chemical Engineering Science. 59(14) 2995–3006.

13. Hyvarinen, A, Karhunen, J and Oja, E (2001). Independent component analysis. John Wiley & Sons Inc.

14. Lee, T, W (1998). Independent component analysis, Theory and application. Springer 1 edition.

15. Teimoortashloo, M and Khaki Sedigh, A (2015). A Modified Independent Component Analysis-Based Fault detection Method in Plant-wide Systems. Control and Cybernetics. 44(2) 287-307.

16. Ge, Z and Song, Z (2007). Process Monitoring Based on Independent Component Analysisâ͈"Principal Component Analysis (ICA-PCA) and Similarity Factors. Industrial and Engineering Chemistry Research. 46(7) 2054–2063.

17. Ge, Z and Song, Z (2010). Maximum-likelihood mixture factor analysis model and its application for process monitoring. Chemometrics & Intelligent Laboratory Systems. 102(1) 53-61.

18. Ge, Z, Yang, C and Song, Z (2009). Improved kernel PCA-based monitoring approach for nonlinear processes. Chemical Engineering Science. 64(9) 2245-2255.

19. Brys, G, Hubert, M and Rousseeuw, P, J (2005). A Robustification of Independent Component Analysis. Journal of Chemo metrics. 19(5-7) 364–375.

20. Chun-Chin, H, Mu-Chen, C and Long-Sheng, C (2010). A Novel Process Monitoring Approach with Dynamic Independent Component Analysis. Control Engineering Practice. 18(3) 242-253.

21. Ge, Z, Gao, F and Song, Z (2011). Batch process monitoring based on support vector data description method. Journal of Process Control. 21(6) 949-959.

22. Miao, A, Song, Z, Ge, Z, Zhou, L and Wen, Q (2013). Nonlinear fault detection based on locally linear embedding. Journal of Control Theory and Applications. 11(4) 615-622.

23. Ge, Z and Song, Z (2011). A distribution-free method for process monitoring. Expert Systems with Applications. 38(8) 9821-9829

24. Ge, Z, Song, Z and Gao, F (2013). Review of Recent Research on Data-Based Process Monitoring. Industrial & Engineering Chemistry Research. 52(10) 3543-3562

25. Mu-Chen, C, Chun-Chin, H, Malhotra, B and Tiwari, M K,(2016). An Efficient ICA-DW-SVDD Fault Detection And Diagnosis Method for Non-Gaussian Processes. International Journal of Production Research. 54(17)5208-5218.

26. Huang, j and Yan, X,(2015). Dynamic Principal Component Analysis, Dynamic Independent Component Analysis and Bayesian inference. Chemometrics and Intelligent Laboratory Systems. 148, 115-127.

27. Zvokelj, M, Zupan, S and Prebil, I,(2016). EEMD-based Multiscale ICA Method for Slewing Bearing Fault Detection and Diagnosis. 370, 394-423.

28. Nandi, Asoke Kumar (Ed.)(1999). Blind Estimation Using Higher-Order Statistics. Kluwer Academic Publishers, Boston.

29. Kendall, M, G and Stuart, A (1969). The Advanced Theory of Statistics. Griffin, London.

30. Peccati, G and Murad, S, T (2011). Wiener Chaos: Moments, Cumulants and Diagrams, A survey with Computer Implementation. Bocconi University Press & Springer.

31. Lukacs, E (1970). Characteristic Functions. Griffin, London.

32. Van Montfort, K and De Leeuw, Y (2004). Factor Analyses for Non-normal Variables by Fitting Characteristic Functions. Department of Statistics Papers, Department of Statistics, UCLA, UC Los Angeles.

33. Mooijaart, A and Satorra, A (2011). Testing for Non-Linear Relationship in Structural Equation Modeling. Department of Statistics Papers, Department of Statistics, UCLA, UC Los Angeles.

34. Satorra, A (2002). Asymptotic Robustness in Multiple Group Linear-latent Variable Models. Econometric Theory. Cambridge University Press. 18(2) 297-312.

35. Hyvarinen, A (1999). Fast and Robust Fixed-Point Algorithms for Independent Component Analysis. IEEE Trans. on Neural Networks. 10(3) 626–634.

36. Makino, S, Lee, T, W and Sawada, H (2007). Blind Speech Separation. Netherlands. Springer.

37. Xia, C, Howell, J (2005). Isolating multiple sources of plant-wide oscillations via independent component analysis. Control Engineering Practice. 13(8) 1027–1035.

38. Ko, C, C, Chen, B, M, Chen, J, Zhuang, Y and Chen Tan, K, (2001). Development of a Web-Based Labortatory for Control Experiments on a Coupled tank Apparatus. IEEE Transactions on Education. 44(1)76-86.

39. Gertsbakh, L (2003). Measurement Theory for Engineers. Springer.

40. Ghilani, C, D (2010). Adjustment Computations Spatial Data Analysis. Fifth Edition. John Wiley & Sons, Inc., Hoboken, New Jersey.

41. Ku, H, H, (1966). Notes on the use of propagation of error formulas. Journal of Research of the National Bureau of Standards 70c(4) 262-273.