Proceedings of the 9th International Symposium on
Dynamics and Control of Process Systems (DYCOPS 2010),
Leuven, Belgium, July 5-7, 2010
Mayuresh Kothare, Moses Tade, Alain Vande Wouwer, Ilse Smets (Eds.)

A Branch and Bound Method for Fault Isolation through Missing Variable Analysis

Vinay Kariwala * **Pabara-Ebiere Odiowei** ** **Yi Cao** **
Tao Chen *

* *School of Chemical & Biomedical Engineering,*
Nanyang Technological University, Singapore 637459
(e-mails: vinay@ntu.edu.sg, chentao@ntu.edu.sg)
** *School of Engineering, Cranfield University,*
Cranfield, Bedford MK43 0AL, UK
(e-mails: p.odiowei@cranfied.ac.uk, y.cao@cranfield.ac.uk)

Abstract: Fault detection and diagnosis (FDD) is a critical approach to ensure safe and efficient operation of manufacturing and chemical processing plants. Multivariate statistical process monitoring (MSPM) has received considerable attention for FDD since it does not require a mechanistic process model. The diagnosis of the source or cause of the detected process fault in MSPM largely relies on contribution analysis, which is ineffective in identifying the joint contribution of multiple variables to the occurrence of fault. In this work, a missing variable analysis approach based on probabilistic principal component analysis is proposed for fault isolation. Furthermore, a branch and bound method is developed to handle the combinatorial nature of the problem involving finding the variables, which are most likely responsible for the occurrence of fault. The efficiency of the method proposed is shown through a case study on the Tennessee Eastman process.

Keywords: Branch and bound, combinatorial optimization, multivariate contribution analysis, multivariate statistical process monitoring, probabilistic principal component analysis.

1. INTRODUCTION

Fault detection and diagnosis (FDD) is a critical approach to ensure safe and efficient operation of manufacturing and chemical processing plants. Within this broad area of FDD, multivariate statistical process monitoring (MSPM) has received considerable attention in terms of both methodological research and industrial applications [Qin, 2003, Venkatasubramanian et al., 2003]. The success of MSPM may be attributed to the fact that a large amount of historical process data is usually available, and thus the developed model attains high accuracy in detecting any deviation from established normal operating conditions (NOC).

In traditional MSPM, principal component analysis (PCA) or partial least squares (PLS) are applied to model the data collected under NOC. Subsequently, monitoring statistics like Hotelling's T^2 and squared prediction error (SPE) are used for fault detection. The subsequent step of MSPM, *i.e.* the diagnosis of the source or cause of the detected fault, is relatively more difficult. If additional information, such as historical data of separate faulty modes, is collected, pattern recognition technique can be used to classify the detected fault. In absence of such information, the primary tool used for fault diagnosis in MSPM is contribution analysis, which quantifies the contribution of individual variables to T^2 and SPE [Miller et al., 1998]. However, the contribution plots based on T^2 and SPE can identify different sets of faulty variables making the decision subjective to operators' experience. Furthermore, contribution analysis investigates individual variables one by one and is ineffective in isolating multiple variables which jointly contribute to the occurrence of fault.

More recently, the need to use a unified monitoring statistic, as opposed to using T^2 and SPE individually, has been well recognized. One way is to combine T^2 and SPE algorithmically [Chen et al., 2004, Yue and Qin, 2001]; the other approach is to rely on fully probabilistic model to provide a single likelihood-based statistic [Chen et al., 2006, Chen and Sun, 2009]. A single monitoring statistic offers a clearer interpretation of the contribution plots. In addition, these studies suggested an alternate approach to quantifying the contribution by using the idea of missing variables [Yue and Qin, 2001, Chen and Sun, 2009]. However, the major difficulty is the large number of possible variable combinations that are required to be evaluated. Therefore, Yue and Qin [2001] made a restrictive assumption that the combinations of faulty variables are specified *a priori* according to known faulty modes, and thus contribution analysis only needs to be conducted on these combinations. Due to lack of an appropriate tool to tackle the combinatorial issue, Chen and Sun [2009] focussed on the analyzing the contribution of individual variables towards the faults. Therefore, the multivariate contribution problem was still left unsolved.

This paper extends the missing-variable based contribution analysis to consider the joint effect of multiple variables, namely multivariate contribution analysis. In this study, we choose probabilistic PCA (PPCA) [Tipping and Bishop, 1999, Kim and Lee, 2003] for modelling normal operating data and on-line process monitoring. Based on the PPCA under NOC, a statistical criterion is derived to quantify contribution of multiple missing variables. Using the criterion, when a fault condition is identified through on-line process monitoring, the fault diagnosis can then be conducted by solving a series of subset selection problems. The combinatorial difficulty of the subset selection problem is tackled by a numerically efficient branch and bound (BAB) algorithm developed to isolate the set of faulty variables based on the criterion derived. The efficiency of the proposed method is demonstrated by its application to the Tennessee Eastman Process Plant [Downs and Vogel, 1993].

2. PROBABILISTIC PRINCIPAL COMPONENT ANALYSIS

Principal component analysis (PCA) [Jolliffe, 2002] is a general multivariate statistical projection technique for dimension reduction. The central idea of PCA is to project the original r-dimensional data, $\mathbf{x}$, onto a space where the variance is maximized: $\mathbf{x} = \mathbf{W}\mathbf{t} + \bar{\mathbf{x}} + \mathbf{e}$. Here $\mathbf{W}$ refers to the eigenvectors of the sample covariance matrix corresponding to the q ($q \leq r$) largest eigenvalues, $\mathbf{t}$ is the q-dimensional scores, $\bar{\mathbf{x}}$ is the mean of the data, and $\mathbf{e}$ is the noise term.

Recently, Tipping and Bishop [1999] proposed a probabilistic formulation of PCA (PPCA) from the perspective of a Gaussian latent variable model. Specifically the noise is assumed to be Gaussian: $\mathbf{e} \sim G(\mathbf{0}, \sigma^2\mathbf{I})$, which implies $\mathbf{x}|\mathbf{t} \sim G(\mathbf{W}\mathbf{t} + \bar{\mathbf{x}}, \sigma^2\mathbf{I})$. Furthermore, by adopting a Gaussian distribution for the scores, $\mathbf{t} \sim G(\mathbf{0}, \mathbf{I})$, the marginal distribution of the data is also Gaussian: $\mathbf{x} \sim G(\bar{\mathbf{x}}, \mathbf{C})$, where the covariance matrix is $\mathbf{C} = \mathbf{W}\mathbf{W}^{\mathrm{T}} + \sigma^2\mathbf{I}$. The model parameters, $\{\bar{\mathbf{x}}, \mathbf{W}, \sigma^2\}$, can be estimated using the maximum likelihood algorithm; see [Tipping and Bishop, 1999] for details. Later, PPCA was applied for process monitoring with improved fault detection capability [Kim and Lee, 2003].

The probabilistic framework of PPCA provides a single statistic for fault detection, as opposed to T^2 and SPE in traditional PCA. It was shown by Chen and Sun [2009] that the data point $\mathbf{x}$ should be considered as out-of-control when the monitoring statistic

$$M^2 = (\mathbf{x} - \bar{\mathbf{x}})^{\mathrm{T}}\mathbf{C}^{-1}(\mathbf{x} - \bar{\mathbf{x}}) > \chi_r^2(\beta) \tag{1}$$

where $\chi_r^2(\beta)$ is the β-fractile of the χ-square distribution with r degrees of freedom.

For the purpose of contribution analysis, Chen and Sun [2009] suggested a missing variable approach in the PPCA framework. In particular, consider that the measurement vector $\mathbf{x}$ is partitioned into n-dimensional $\mathbf{x}_o$ and d-dimensional $\mathbf{x}_m$ as

$$\mathbf{x}^T = \begin{bmatrix}\mathbf{x}_o^T & \mathbf{x}_m^T\end{bmatrix} \tag{2}$$

where the subscripts o and m refer to observed and missing variables, respectively. Similarly, let the mean ($\bar{\mathbf{x}}$) and covariance matrix ($\mathbf{C}$) be partitioned as

$$\bar{\mathbf{x}} = \begin{bmatrix}\bar{\mathbf{x}}_o \\ \bar{\mathbf{x}}_m\end{bmatrix}; \quad \mathbf{C} = \begin{bmatrix}\mathbf{C}_{oo} & \mathbf{C}_{om} \\ \mathbf{C}_{mo} & \mathbf{C}_{mm}\end{bmatrix} \tag{3}$$

The conditional mean ($\bar{\mathbf{x}}_{m|o}$) and covariance matrix ($\mathbf{C}_{m|o}$) of $\mathbf{x}_m$ given $\mathbf{x}_o$ are

$$\bar{\mathbf{x}}_{m|o} = \bar{\mathbf{x}}_m + \mathbf{C}_{mo}\mathbf{C}_{oo}^{-1}(\mathbf{x}_o - \bar{\mathbf{x}}_o) \tag{4}$$

$$\mathbf{C}_{m|o} = \mathbf{C}_{mm} - \mathbf{C}_{mo}\mathbf{C}_{oo}^{-1}\mathbf{C}_{om} \tag{5}$$

Then, the conditional mean ($\bar{\mathbf{x}}_{|o}$) and covariance matrix ($\mathbf{C}_{|o}$) of whole $\mathbf{x}$ given $\mathbf{x}_o$ are

$$\bar{\mathbf{x}}_{|o} = \begin{bmatrix}\mathbf{x}_o \\ \bar{\mathbf{x}}_{m|o}\end{bmatrix}; \quad \mathbf{C}_{|o} = \begin{bmatrix}0 & 0 \\ 0 & \mathbf{C}_{m|o}\end{bmatrix} \tag{6}$$

Therefore, the re-calculated monitoring statistic with missing values is given by

$$E[M^2] = \mathrm{tr}(\mathbf{C}^{-1}((\bar{\mathbf{x}}_{|o} - \bar{\mathbf{x}})(\bar{\mathbf{x}}_{|o} - \bar{\mathbf{x}})^T + \mathbf{C}_{|o})) \tag{7}$$

where $E[\cdot]$ denotes the expectation operator.

The criterion (7) is applicable to any number of missing variables. In the original method of Chen and Sun [2009], each individual variable of $\mathbf{x}$ is regarded as missing, and the monitoring statistic in (7) is re-calculated. If a variable contributes significantly to the data being detected as faulty, then the re-calculated statistic will be dramatically reduced. Furthermore, if $E[M^2]$ is smaller than the confidence bound, then we can say that by removing the corresponding variable, the process would return to the normal operating region. However, this approach is not capable of studying the joint contribution of multiple variables.

From the fault diagnosis point of view, the original source of a fault should correspond to a small set of measurements. Therefore, the objective becomes to select a minimum number of missing variables, whose re-calculated monitoring statistic is below the confidence bound. The flow of the method is given by the following algorithm:

Algorithm 1. Initially set $d = 1$.

(1) Select d missing variables such that the re-calculated monitoring statistic is minimized.
(2) If the minimum statistic is below the confidence bound, then the corresponding variables are isolated as the source of fault, and the algorithm can be terminated.
(3) Otherwise, set $d = d + 1$, and return to Step 1.

The selection of d-dimensional $\mathbf{x}_d$ from r-dimensional $\mathbf{x}$ in Step 1 of Algorithm 1 is a combinatorial optimization problem, which is NP-hard. One of the main contributions of this work is the development of a branch and bound (BAB) algorithm to solve this optimization problem efficiently, as discussed in the next section.

3. BRANCH AND BOUND METHOD

3.1 General principle

Let $X_r = \{x_i | i = 1, 2, \cdots, r\}$, be an r-element set. A subset selection problem with the selection criterion ϕ involves finding the optimal solution, X_n^*, such that

$$\phi(X_n^*) = \min_{X_n \subset X_r} \phi(X_n) \tag{8}$$

For this problem, the number of alternatives is $\mathcal{C}_r^n = \frac{r!}{(r-n)!n!}$, which grows very quickly with r and n rendering exhaustive search unviable. BAB approach can provide globally optimal solution for the subset selection problem in (8) without exhaustive search. In this approach, the original problem (node) is divided (branched) into several non-overlapping subproblems (sub-nodes). If any of the n-element solutions of a sub-problem cannot lead to the optimal solution, the sub-problem is not evaluated further (pruned), else it is branched again. The pruning of sub-problems allows the BAB approach to gain efficiency in comparison with exhaustive search.

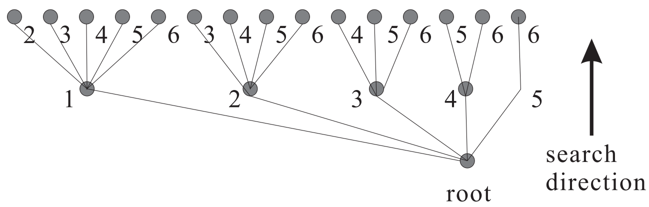

Fig. 1. Solution tree for selecting 2 out of 6 elements

The solution tree required for implementation of BAB approach for subset selection based on the monitoring statistic in (7) is shown in Figure 1. In this solution tree, each node has a fixed set F_f and a candidate set C_c, which have f and c elements, respectively. The relationship between the fixed and candidate sets of a node and its i^{th} sub-node (branching rule) is given as follows:

$$F_{f+1}^i = F_f \cup x_i; \quad C_{c-i}^i = C_c \setminus \{x_1, \cdots, x_i\} \tag{9}$$

where F_{f+1}^i and C_{c-i}^i denote the fixed and candidate sets of the i^{th} sub-node and $i = 1, 2, \cdots, f + c - n + 1$. Based on (9), it can be noted that F_f is gradually expanded using the elements of C_c, until the dimension of F_f reaches the target subset size n. An example of the solution tree obtained by recursively applying the branching rule in (9) is shown in Figure 1. For the root note in this solution tree, we have $F_f = \emptyset$ and $C_c = X_r$. The label of the nodes denote the element being moved from C_c to F_f. The solution tree has $\mathcal{C}_n^r$ terminal nodes, which represent different n-element subsets of X_r. As the subset size is gradually increased in this solution tree, the corresponding BAB method is called upwards BAB method [Cao and Kariwala, 2008, Kariwala and Cao, 2009, 2010]. The reader is referred to [Narendra and Fukunaga, 1977, Yu and Yuan, 1993, Cao and Saha, 2005] for details on downwards BAB method.

To describe the pruning principle, let $\mathcal{X}$ denote the ensemble of all n-element subsets, which can be obtained by expanding F_f using (9), *i.e.*,

$$\mathcal{X} = \{\{F_f, X_{n-f}\} | X_{n-f} \subset C_c\} \tag{10}$$

and $\underline{\phi}(F_f)$ be the lower bound on ϕ computed over all elements of $\mathcal{X}$, *i.e.*

$$\underline{\phi}(F_f) = \min_{X_n \in \mathcal{X}} \phi(X_n) \tag{11}$$

Assume that B is an upper bound of the globally optimal criterion, *i.e.* $B \geq \phi(X_n^*)$. Then,

$$\phi(X_n) > \phi(X_n^*) \, \forall X_n \in \mathcal{X}, \text{if } \underline{\phi}(F_f) > B \tag{12}$$

Hence, any $X_n \in \mathcal{X}$ cannot be optimal and can be pruned without further evaluation, if $\underline{\phi}(F_f) > B$.

3.2 Application to fault isolation problem

Through lengthy but straightforward manipulations, it can be shown that the monitoring statistic in (7) can be alternately expressed as

$$E[M^2] = (\mathbf{x}_o - \bar{\mathbf{x}}_o)^T \mathbf{C}_{oo}^{-1} (\mathbf{x}_o - \bar{\mathbf{x}}_o) + d \tag{13}$$

which is more amenable to the application of BAB method.

In Algorithm 1, $d = r - n$ is constant during every iteration. Thus, variable selection can be carried out by minimizing the first term in (13). Furthermore, $\bar{\mathbf{x}}_o$ can be subtracted in a preprocessing stage to simplify calculations. By defining $\mathbf{y} = \mathbf{x} - \bar{\mathbf{x}}$, the n observed variables can be selected (equivalent to selecting d missing variables) by solving the following problem:

$$\min_{X_n \subset X_r} \phi(X_n) = \mathbf{y}_{X_n}^T (\mathbf{C}_{X_n, X_n})^{-1} \mathbf{y}_{X_n} \tag{14}$$

where $X_r = \{1, 2, \cdots, r\}$, $\mathbf{y}_{X_n}$ denotes the elements of $\mathbf{y}$ with indices in X_n and $\mathbf{C}_{X_n,X_n}$ represents the principal submatrix of $\mathbf{C}$ with rows and columns indexed by X_n.

The use of BAB for solving the optimization problem in (14) requires a lower bound on the selection criteria, calculated over all $\mathcal{X}$ in (10). This lower bound is derived in the next proposition.

Proposition 1. Consider a node with fixed set F_f and candidate set C_c. For $\mathcal{X}$ in (10),

$$\phi(F_f) \leq \min_{X_n \in \mathcal{X}} \phi(X_n) \tag{15}$$

Proposition 1 implies that the non-optimal nodes can be pruned based on $\phi(F_f)$. Although this lower bound allows the development of an efficient BAB algorithm, further efficiency can be gained by pruning the sub-nodes directly [Cao and Kariwala, 2008, Kariwala and Cao, 2009, 2010]. The next proposition relates the selection criteria of a node with its sub-nodes, which in turn enables pruning of sub-nodes.

Proposition 2. Consider a node with fixed set F_f and candidate set C_c. For $x_i \in C_c$, $i = 1, 2, \cdots, c$,

$$\phi(F_f \cup i) = \phi(F_f) + \alpha_i \tag{16}$$

where

$$\alpha_i = \frac{\left(\mathbf{y}_i - \mathbf{y}_{F_f} \mathbf{C}_{F_f,F_f}^{-1} \mathbf{C}_{F_f,i}\right)^2}{(\mathbf{C}_{i,i} - \mathbf{C}_{F_f,i}^T \mathbf{C}_{F_f,F_f}^{-1} \mathbf{C}_{F_f,i})^2} \tag{17}$$

The evaluation of (16) requires inversion of only one matrix $\mathbf{C}_{F_f,F_f}$, which is the same for all $i \in C_c$. Thus, the use of (16) to obtain the selection criteria for all sub-nodes together is computationally more efficient than directly evaluating the selection criteria for every node. In summary, the following BAB algorithm can be used as Step 1 of Algorithm 1 for fault isolation purposes.

Algorithm 2. Initialize $f = 0$, $F_f = \emptyset$, $C_c = X_r$, $\phi(F_f) = 0$ and $B = \infty$. Call the following recursive algorithm:

(1) If $\phi(F_f) > B$, prune the current node and return, else perform the following steps.
(2) Calculate α_i in (17) $\forall i \in C_c$. Prune the subsets with $\phi(F_f) + \alpha_i > B$.
(3) If $f = n$, go to next step. Otherwise, generate the c sub-nodes according to the branching rule in (9) and call the recursive algorithm in Step 1 for each sub-node. Return to the caller after the execution of the loop finishes.
(4) Find $J_{\min} = \phi(F_f) + \min_{i \in C_c} \alpha_i$. If $J_{\min} < B$, update $B = J_{\min}$. Return to the caller.

4. TENNESSEE EASTMAN PROCESS

To demonstrate its effectiveness and efficiency, the proposed approach is applied for fault isolation of the Tennessee Eastman (TE) process [Downs and Vogel, 1993]. This process has 5 main units, which are the reactor, condenser, separator, stripper and compressor. Streams of the plant consists of 8 components; A, B, C, D, E, F, G and H. Components A, B and C are gaseous reactants which are fed to the reactor to form products G and H. The flowsheet of the TE process is shown in Figure 2.

For fault isolation, the TE process is considered under closed-loop control as described by Downs and Vogel [1993]. The 41 available measurements and 11 out of 12 manipulated variables (MVs) are collected for 21 operational modes, which correspond to the normal and 20 faulty operation modes. These 52 measured variables are listed in Table 1. Eleven out of these 20 operational faults, listed in Table 2, can be clearly identified using the PPCA approach [Chen and Sun, 2009, Tipping and Bishop, 1999, Kim and Lee, 2003] and are adopted for the case study. The identified variables for these eleven faults studied are listed in Table 3.

Table 3. Possible Fault responsive variables detected by the branch and bound algorithm

Fault IDs	Variables eliminated
1	$\{x_{16}\}$
2	$\{x_{24}\}$
4	$\{x_9, x_{51}\}$
5	$\{x_{22}\}$
6	$\{x_1, x_{44}\}$
7	$\{x_4, x_6, x_9, x_{16}, x_{22}, x_{45}, x_{51}\}$
8	$\{x_{37}\}$
11	$\{x_{51}\}$
12	$\{x_{22}\}$
13	$\{x_{37}\}$
14	$\{x_9, x_{51}\}$

As indicated in Table 3, most of the faults result in only one and two responsible variables. An exception is Fault 7, for which BAB method indicates that there are up to seven measured variables responsive to this fault. To select 7 out of 52 measured variables, there are $C_{52}^7 = 133,784,560$ alternatives. If one had to evaluate all alternatives to find the minimum criterion value, it would take several days to get the conclusion even if each evaluation takes only one millisecond. However, it only takes about 3.37 seconds for the BAB algorithm to find the subset on a PC with Intel® Core™2 CPU T5200 (1.60GHz, 2038MB, 32-bit Operating System) using MATLAB® R2006. This indicates that the proposed algorithm is suitable for on-line FDD to identify the root cause of a fault under investigation in a short time so that the consequent loss due to the fault can be minimised.

To better appreciate the proposed fault diagnosis algorithm, the variation of the monitoring statistic with the number of variables to be observed n is shown in Figure 3. The horizontal line in Figure 3 is the upper control limit calculated as $\chi_r^2(\beta)$ with $r = 52$ and $\beta = 0.99$. It is shown that for $n \leq 45$ (left to the dashed line), the minimum criterion value is less than the upper control limit, whilst for $n \geq 46$, (right to the dotted line) the criterion is above the upper control limit. Therefore, the maximum number of non-responsive variables is 45 and the number of possible responsive variables to Fault 7 is $52 - 45 = 7$. The actual deviations from corresponding means of these 7 variables are $x_4 = -14.5015$, $x_6 = -3.8761$, $x_9 = -3.1665$, $x_{16} = -4.2085$, $x_{22} = -3.4847$, $x_{45} = 4.4334$, and $x_{51} = -4.4255$. The corresponding locations of these 7 variables and the fault are marked in Figure 2.

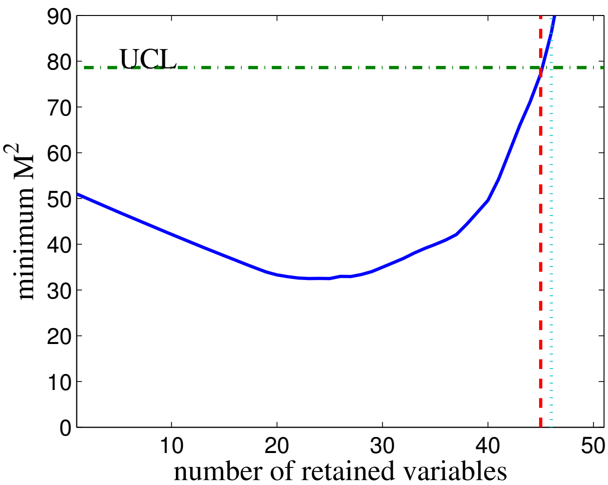

Fig. 3. The minimum M^2 against number of variables (n) observed for Fault 7.

A physical explanation of the propagation of fault to the identified variables is as follows: Fault 7 reflected in Figure 2 involves C Header Pressure Loss - reduced availability (step change in stream 4), resulting in a decrease in the Total Feed in stream 4 (x_4) and a corresponding increase in the MV to Total Feed Flow in stream 4 (x_{45}) to counter the effect of the fault occurrence through the corresponding flow control loop of the plant. The decrease in the Total Feed in stream 4 (x_4) results in a corresponding decrease in the Recycle back to the reactor through stream 6, thereby reducing the Reactor Feed Rate in stream 6 (x_6) as well as the Reactor Temperature (x_9). This decrease in the Reac-

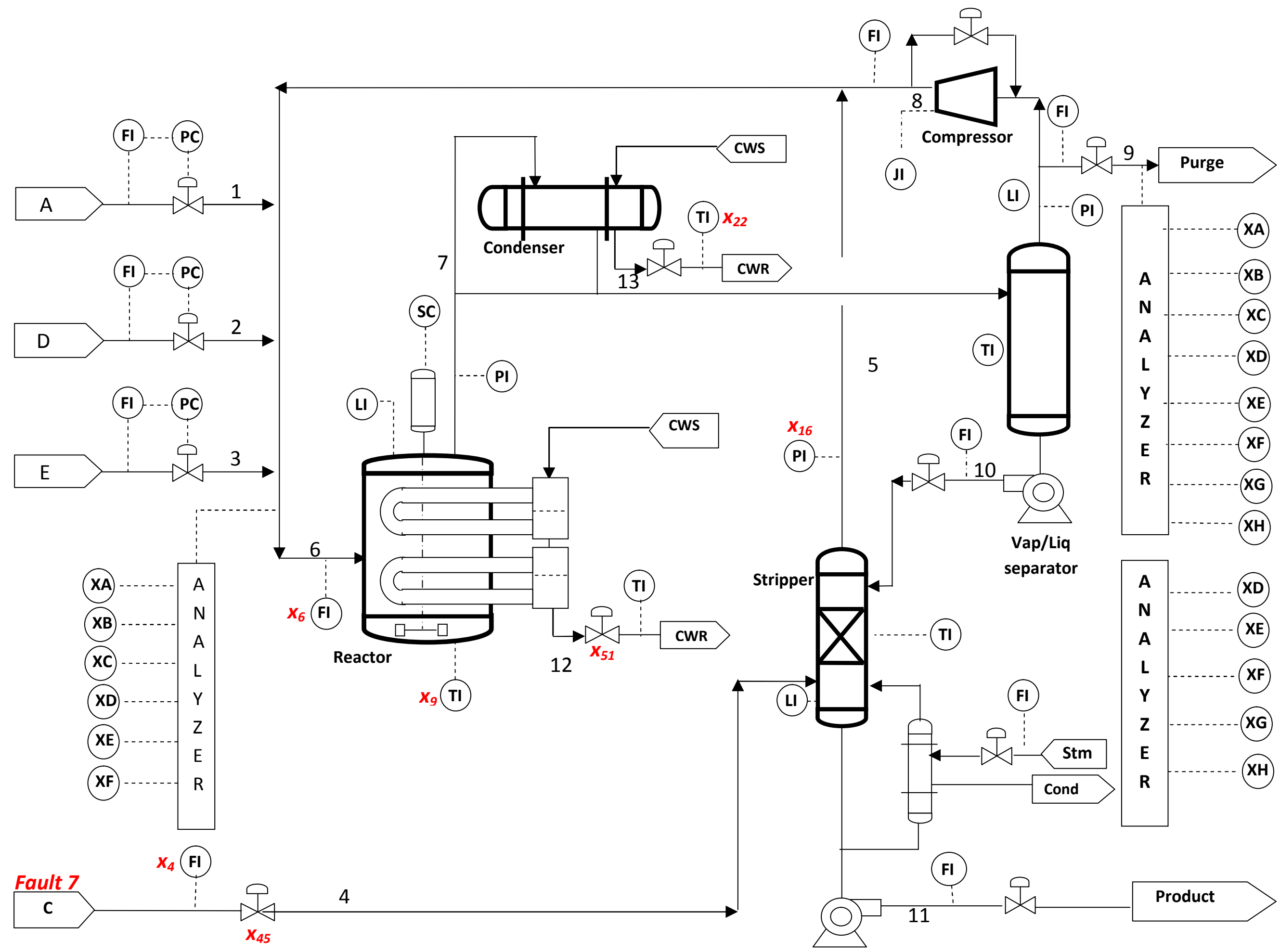

Fig. 2. Flowsheet of the TEP Plant showing detected variables for Fault 7

Table 1. Measured variables

ID	Description	ID	Description
x_1	A Feed (Stream 1)	x_{27}	Component E (Stream 6)
x_2	D Feed (Stream 2)	x_{28}	Component F (Stream 6)
x_3	E Feed (Stream 3)	x_{29}	Component A (Stream 9)
x_4	Total Feed (Stream 4)	x_{30}	Component B (Stream 9)
x_5	Recycle Flow (Stream 8)	x_{31}	Component C (Stream 9)
x_6	Reactor Feed Rate (Stream 6)	x_{32}	Component D (Stream 9)
x_7	Reactor Pressure	x_{33}	Component E (Stream 9)
x_8	Reactor Level	x_{34}	Component F (Stream 9)
x_9	Reactor Temperature	x_{35}	Component G (Stream 9)
x_{10}	Purge Rate (Stream 9)	x_{36}	Component H (Stream 9)
x_{11}	Product Separator Temperature	x_{37}	Component D (Stream 11)
x_{12}	Product Separator Level	x_{38}	Component E (Stream 11)
x_{13}	Product Separator Pressure	x_{39}	Component F (Stream 11)
x_{14}	Product Separator Underflow (Stream 10)	x_{40}	Component G (Stream 11)
x_{15}	Stripper Level	x_{41}	Component H (Stream 11)
x_{16}	Stripper Pressure	x_{42}	MV to D Feed Flow (Stream 2)
x_{17}	Stripper Underflow (Stream 11)	x_{43}	MV to E Feed Flow (Stream 3)
x_{18}	Stripper Temperature	x_{44}	MV to A Feed Flow (Stream 1)
x_{19}	Stripper Steam Flow	x_{45}	MV to Total Feed Flow (Stream 4)
x_{20}	Compressor Work	x_{46}	Compressor Recycle Valve
x_{21}	Reactor Cooling Water Outlet Temperature	x_{47}	Purge Valve (Stream 9)
x_{22}	Separator Cooling Water Outlet Temperature	x_{48}	MV to Separator Pot Liquid Flow (Stream 10)
x_{23}	Component A (Stream 6)	x_{49}	MV to Stripper Liquid Product Flow (Stream 11)
x_{24}	Component B (Stream 6)	x_{50}	Stripper Steam Valve
x_{25}	Component C (Stream 6)	x_{51}	MV to Reactor Cooling Water Flow
x_{26}	Component D (Stream 6)	x_{52}	MV to Condenser Cooling Water Flow

Table 2. Operational Faults

Fault ID	Description
1	Step in A/C Feed Ratio, B Composition Constant (stream 4)
2	Step in B composition while A/C ratio is constant (stream 4)
4	Step in Reactor Cooling Water Inlet Temperature
5	Step in Condenser Cooling Water Inlet Temperature
6	A Feed loss (step change in stream 1)
7	C Header Pressure Loss - reduced availability (step change in stream 4)
8	Random variation in A,B,C Feed Composition (stream 4)
11	Random variation in Reactor Cooling Water Inlet Temperature
12	Random variation in Condenser Cooling Water Inlet Temperature
13	Slow drift in Reaction Kinetics
14	Sticking Reaction Cooling Water Valve

tor Temperature (x_9) results in a decrease in the Reactor Cooling Water Flow (x_{51}). Also, the decrease in the Total Feed in stream 4 (x_4) into the Stripper results in a decrease in the Stripper Pressure (x_{16}) and consequently, a decrease in the Separator Cooling Water Outlet Temperature (x_{22}). From the above explanation, it can be concluded that the results obtained through the BAB algorithm is also supported by the analysis through a physical understanding of the plant. Therefore, this proposed approach will be practically useful for fault isolation associating with condition monitoring.

5. CONCLUSIONS

In contrast to fault detection, the isolation of the variables responsible for the fault using considerable attention paid to multivariate statistical process monitoring techniques can be difficult. It is for the first time that a multivariate missing-variable approach and an efficient branch and bound algorithm are proposed to efficiently identify the joint contribution of multiple variables to the occurrence of the fault. Although the focus of this paper is on using probalistic principal component analysis (PPCA), the framework developed in this work can be easily extended to adopt other criteria for fault diagnosis. The numerical case study on the Tennessee Eastman process shows that the proposed method is able to find the minimum set of variables, which are affected by the fault, in a short time. This computational efficiency can give operators more time to identify and further deal with the fault in order to minimize the consequent loss due to the occurrence of the fault.

ACKNOWLEDGEMENTS

This work was supported by the Royal Society, UK through grant no. IV0871568.

REFERENCES

Y. Cao and V. Kariwala. Bidirectional branch and bound for controlled variable selection: Part I. Principles and minimum singular value criterion. *Comput. Chem. Engng.*, 32(10):2306–2319, 2008.

Y. Cao and P. Saha. Improved branch and bound method for control structure screening. *Chem. Engg. Sci.*, 60(6): 1555–1564, 2005.

Q. Chen, U. Kruger, M. Meronk, and A. Y. T. Leung. Synthesis of T^2 and Q statistics for process monitoring. *Control Engineering Practice*, 12:745–755, 2004.

T. Chen and Y. Sun. Probabilistic contribution analysis for statistical process monitoring: A missing variable approach. *Control Engineering Practice*, 17(4):469–477, 2009.

T. Chen, J. Morris, and E. Martin. Probability density estimation via an infinite Gaussian mixture model: application to statistical process monitoring. *Journal of the Royal Statistical Society C (Applied Statistics)*, 55:699–715, 2006.

J. J. Downs and E. F. Vogel. A plant-wide industrial process control problem. *Comput. Chem. Eng.*, 17(3): 245–255, 1993.

I. T. Jolliffe. *Principal Component Analysis*. Springer, 2nd edition, 2002.

V. Kariwala and Y. Cao. Bidirectional branch and bound for controlled variable selection: Part II. Exact local method for self-optimizing control. *Comput. Chem. Eng.*, 33:1402–1412, 2009.

V. Kariwala and Y. Cao. Bidirectional branch and bound for controlled variable selection: Part III. Local average loss minimization. *IEEE Transactions on Industrial Informatics*, 6:54–61, 2010.

D. Kim and I.-B. Lee. Process monitoring based on probabilistic PCA. *Chemometrics and intelligent laboratory systems*, 67:109–123, 2003.

P. Miller, R. E. Swanson, and C. F. Heckler. Contribution plots: a missing link in multivariate quality control. *International Journal of Applied Mathematics and Computer Science*, 8:775–792, 1998.

P. Narendra and K. Fukunaga. A branch and bound algorithm for feature subset selection. *IEEE Trans. Comput.*, C-26:917–922, 1977.

S. J. Qin. Statistical process monitoring: basics and beyond. *Journal of Chemometrics*, 17(8-9):480–502, 2003.

M. E. Tipping and C. M. Bishop. Probabilistic principal component analysis. *Journal of the Royal Statistical Society B*, 61:611–622, 1999.

V. Venkatasubramanian, R. Rengaswamy, S.N. Kavuri, and K. Yin. A review of process fault detection and diagnosis Part III: Process history based methods. *Computers and Chemical Engineering*, 27(3):327–346, 2003.

B. Yu and B. Yuan. A more efficient branch and bound algorithm for feature selection. *Pattern Recognition*, 26: 883–889, 1993.

H.H. Yue and S.J. Qin. Reconstruction based fault identification using a combined index. *Industrial and Engineering Chemistry Research*, 40:4403–4414, 2001.