Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

Principal Components Structured Models for Fault Isolation

C. Verde and J. Mina *

* Instituto de Ingeniería-UNAM, Coyoacán DF, 04510, México
(e-mail: verde@servidor.unam.mx, jminaa@iingen.unam.mx)

Abstract: This work proposes a method for fault isolation by means of the structured model characterization with isolation capability, and the residual generation through dynamic principal component analysis. Specifically, the characterization is obtained using graph theory tools, and is expressed in terms of known variables and subsets of constraints. Thus, in the absence of analytical explicit models, the fault isolation task can be solved if the structured models satisfy isolability conditions and a set of nominal historical data from the process is available to carry out the dynamic principal component analysis based monitoring with adaptive standardization. Simulation results for the three tanks system show the effectiveness of the solution for fault isolation tasks.

Keywords: Faults Isolation; Dynamic Principal Components Analysis; Residual Structure; Redundancy Relation Evaluation.

1. INTRODUCTION

It is known that the possibilities for solving a fault detection and isolation (FDI) problem, depend on the structural properties associated to the relation of internal and external process variables, this means that the existence of a residual generator is determined by the system structure. The search of conditions to solve a FDI problem has been tackled with a variety of tools by De Persis and Isidori [2001], Düstegör et al. [2006], Massoumnia et al. [1989], Krysander and Nyberg [2002]. In particular, Blanke et al. [2003], Gentil et al. [2004] proposed the structural analysis based on graph theory, to obtain redundancy relations between known variables for structurally equivalent systems.

On the other side, for large scale systems the FDI problem based in analytical models is not a simple task and requires a considerable design effort, Isermann [2006]. In these conditions the data driven based approaches are good alternatives, Chiang et al. [2001]. In particular, the dynamic principal component analysis (DPCA), which implicitly describes the linear correlation structure of a multivariate process from historical data, has been successfully used to solve fault detection tasks, Kresta et al. [1991], Raich and Çinar [1996], however, it has limitations for the fault isolation task. To improve the fault isolation capacity, Gertler and Cao [2005] propose from partial DPCA models to generate structured residuals, however, this proposal does not allow to know, *a priori*, if there exist a correlation structure for each partial model. The idea of using *a priori* knowledge of the process operation has been considered by some authors. Wang et al. [2002] propose, by means of graph tools, to optimize the sensor locations in order to improve the performance of PCA; Groenewald et al. [2006] suggest, for the monitoring of a mineral processing plant, to use statistical analysis together with fundamental knowledge of the process operation in order to reduce the computational load in the data processing. Even when these methods take into account some functional knowledge of the process, they don't tackle the isolation issue.

The above described facts motivated this paper in which it is shown that the weakness of DPCA for fault isolation, can be overcome when this method is complemented with structural information of the redundancy relations by graph tools. Because the variables associated to a redundancy relation are correlated, the idea is to determine from the system structure the known variables subsets which characterize the redundancy relations for an specific isolation task and from each obtained subset of correlated variables a DPCA model is performed with adaptive standardization.

The outline of this paper is as follows. Section 2 describes, in the framework of structural analysis (SA), how to determine sets of correlated known variables $\mathcal{Z}_i$ involved in the primary redundancy relations considering the fault isolability. Section 3 is dedicated to describe the method used to generate residuals for a given set $\mathcal{Z}_i$ using a DPCA algorithm with adaptive standardization. Section 4 shows the potentiality of the integration of DPCA with SA considering an academic example in which the maximum fault isolability is achieved. Finally, the conclusions are presented in section 5.

2. SYSTEM STRUCTURE

Two basic concepts in a FDI issue are the redundancy relation and the residual generator. Frisk and Åslund [2005] use consistency relation instead of redundancy relation, but the idea in both definitions remains the same.

Definition 1. Let $\mathbf{z}$ be a vector of known signals. The scalar expression $\boldsymbol{RR}(z, \dot{z}, \ddot{z}, \ldots)$ is a *redundancy relation*

[1] Supported by IN112208-DGAPA-UNAM

if for all $\mathbf{z}$ consistent with the fault-free model it holds that

$$\boldsymbol{RR}(z, \dot{z}, \ddot{z}, \ldots) = \mathbf{0} \tag{1}$$

Definition 2. Let $\mathbf{z}$ be a vector of known signals. A dynamic system, with input $\mathbf{z}$ and a scalar signal $\rho(t)$ as output, is a *residual generator* if $\mathbf{z}$ is consistent with the fault-free model implies $\lim_{t\to\infty} \rho(t) = 0$, where vector $\mathbf{z}$ includes both sensor data and known control signals.

The redundancy relations depend on the system structural properties and are determined by the equations which relate internal and external variables. Thus, these relations can be characterized for a family of systems with the same structure by graph theory, Blanke et al. [2003].

Definition 3. A dynamic system can be described by a bipartite graph $\mathcal{G} = \{\mathcal{C}, \mathcal{V}, \mathcal{E}\}$ where $\mathcal{C}$ is the set associated with the system equations or constraints with $|\mathcal{C}| = n_c$. The set of variables in the graph is defined by $\mathcal{V} = \mathcal{X} \cup \mathcal{Z}$ with $|\mathcal{V}| = n_v$; where $\mathcal{X}$ is the unknown variables set with cardinality $|\mathcal{X}| = n$; $\mathcal{Z} = \mathcal{U} \cup \mathcal{Y}$ is the known variables set, $\mathcal{U}$ the exogenous variables set with $|\mathcal{U}| = l$, and the set $\mathcal{Y}$ of endogenous with $|\mathcal{Y}| = s$. Additionally, $\mathcal{E}$ is the set of edges such that

$$e_{ij} = \begin{cases} (c_i, v_j) & \text{if } v_j \text{ appears in } c_i \\ 0 & \text{on the contrary} \end{cases}$$

A bipartite graph can be obtained using cause-effect relationships and not necessarily using accuracy explicit analytical models.

There are different options to take into account the time derivative or time shift operator in a bipartite graph, Nyberg [2006]. Here, for each state x_i an extra constraint of the form:

$$x_k = dx_i = \dot{x}_i$$

or

$$x_k(t) = qx_i(t) = x_i(t+1)$$

is included for a continuous or discrete variable, respectively.

The basic process to get the structure of $\mathcal{G}$ is the matching, which is based in the calculability property and associates variables with constraints from which the unknown variables can be eliminated. Once a matching is obtained, the involved constraints can be interpreted as operators from one variables set to other generated by constraints concatenation or as paths which links variables following the oriented graph.

For example, in the graph of Fig. 1 if $\{x_i, x_j, x_k\}$ are known, two relations can be found by eliminating the internal variable x_z in the following way

- Relate x_z in terms of measurements $\{x_i, x_k\}$ using $\{c_1, c_2\}$ and then determine the measurement x_j through c_3.
- Relate x_z in terms of x_j using c_3 and then determine the two measurements $\{x_i, x_k\}$ through c_1 and c_2, respectively.

A variety of matching algorithms exist to obtain from $\mathcal{G}$ the possible paths between variables which characterize the primary redundancy relations as concatenated functions of known variables, Lorentzen et al. [2003], Gentil et al. [2004].

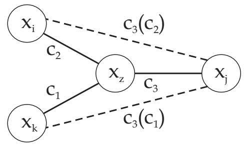

Fig. 1. Elimination of the unknown variable x_z

Due to the graph bidirectional property, one can redefine an endogenous variable as exogenous, which is named *pseudo-exogenous*; this is the case in the above example where x_j has been considered as exogenous. So, similar to *Definition* 1 for an analytical redundancy relation, a redundant graph can be defined as follows:

Definition 4. Let $\mathcal{Z}_i = \mathcal{U}_{si} \cup y_i$ be a subset of known variables perfectly matched through the subset of restrictions $\mathcal{C}_i$, then

$$\mathcal{GR}_i(\mathcal{C}_i, \mathcal{U}_{si}, y_i) \tag{2}$$

is a *redundant graph* which establish, by means of $\mathcal{C}_i$, a consistency between the pseudo-exogenous subset $\mathcal{U}_{si} \subset \mathcal{Z} \setminus y_i$ and the target variable y_i.

This definition together with the concept of *pseudo-exogenous* variables, which are assumed independent in an oriented subgraph, simplifies the analysis of subgraphs and the search of redundancy relations with short paths or short constraints concatenations, maximizing the system fault isolability.

Starting from the bipartite graph $\mathcal{G}$ for a given set of faults $\mathcal{F}$ of interest, the following example clarifies how to obtain and characterize the redundant graphs $\mathcal{GR}_i$ and their corresponding subsets $\mathcal{Z}_i = \mathcal{U}_{si} \cup y_i$ of correlated variables and subsets of constraints $\mathcal{C}_i$. It is also shown that $\mathcal{GR}'s$ with short paths have more isolation capability.

Consider the dynamic system given by

$$\dot{x}_1 = -ax_1 + x_2 + u_1 \tag{c1}$$

$$x_3 = dx_1 \tag{c2}$$

$$\dot{x}_2 = x_1 + bx_2 \tag{c3}$$

$$x_4 = dx_2 \tag{c4}$$

$$y_1 = x_1 + x_2 \tag{c5}$$

$$y_2 = 5 + x_2 \tag{c6}$$

with no concurrent faults in sensors, actuator and process; the latest ones specified as deviations in parameters a and b. Starting in the control framework, u_1 is an exogenous signal and $\{y_1, y_2\}$ are endogenous signals. Under these conditions the oriented graph described in Fig. 2 is obtained, where two paths between known variables are identified which connect u_1 with y_1 and y_2, respectively:

- $\mathcal{GR}_1(\mathcal{C}_1, \mathcal{U}_{s1}, y_1)$ connects the exogenous variable $\mathcal{U}_{s1} = u_1$, with the target variable y_1, by way of constraints $\mathcal{C}_1 = \{c1, c2, c3, c4, c5\}$.
- $\mathcal{GR}_2(\mathcal{C}_2, \mathcal{U}_{s2}, y_2)$ connects the exogenous variable $\mathcal{U}_{s2} = u_1$, with the target variable y_2, by way of constraints $\mathcal{C}_2 = \{c1, c2, c3, c4, c6\}$.

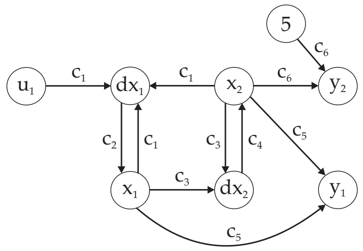

Fig. 2. Matched Graph with u_1 as exogenous variable

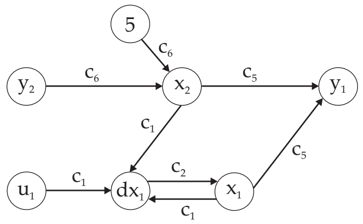

Fig. 3. Matched Graph with the set $\{u_1, y_2\}$ as pseudo-exogenous variables

Taking into account the variables and constrains involved in each graph, the corresponding fault sensitivity can be identified and is shown in the following faults signature matrix where f denotes a fault and the subindex defines the specific fault.

$\backslash$	$\mathcal{Z}_i$	$\lvert\mathcal{C}_i\rvert$	f_{y1}	f_{y2}	f_{u1}	f_a	f_b
$\mathcal{GR}_1$	$\mathcal{U}_{s1}, y_1$	5	1	0	1	1	1
$\mathcal{GR}_2$	$\mathcal{U}_{s2}, y_2$	5	0	1	1	1	1

Since, $\mathcal{GR}_1$ and $\mathcal{GR}_2$ respond to all faults except one, their isolation capability is low; only sensor faults can be isolated considering these graphs. Moreover, since the three last fault columns have the same signature, the actuator fault f_{u1}, and process faults $\{f_a, f_b\}$ are detectable, but not isolable with these $\mathcal{GR}'s$.

Considering y_2 as pseudo-exogenous in the bipartite graph, the matching process generates the oriented graph $\mathcal{GR}_3(\mathcal{C}_3, \mathcal{U}_{s3}, y_1)$ given in Fig. 3 which connects the set $\mathcal{U}_{s3} = \{u_1, y_2\}$ with the target y_1 by way of constraints set $\mathcal{C}_3 = \{c1, c2, c5, c6\}$. Given that $c3$ is not in the path, this graph is insensitive to faults in $c3$; $c4$ and the dynamic of x_2 are not considered. As result the faults $\{f_{u1}, f_a\}$ can be isolated of f_b. In other words, a correlation between $\mathcal{U}_{s3}$ and y_1 has been detected by $\mathcal{GR}_3$ which improves the isolability with respect to the above cases.

$\backslash$	$\mathcal{Z}_3$	$\lvert\mathcal{C}_3\rvert$	f_{y1}	f_{y2}	f_{u1}	f_a	f_b
$\mathcal{GR}_3$	$\mathcal{U}_{s3}, y_1$	4	1	1	1	1	0

Another option is to assume $\mathcal{U}_{s4} = \{y_1\}$ as pseudo-exogenous resulting the matched graph of Fig. 4. One path from $\mathcal{U}_{s4}$ to target y_2 is identified by way of $\mathcal{C}_4 = \{c3, c4, c5, c6\}$. The redundant graph $\mathcal{GR}_4(\mathcal{C}_4, \mathcal{U}_{s4}, y_2)$ is insensitive to f_{u1} and f_a and can also be considered in the faults signature matrix

$\backslash$	$\mathcal{Z}_4$	$\lvert\mathcal{C}_4\rvert$	f_{y1}	f_{y2}	f_{u1}	f_a	f_b
$\mathcal{GR}_4$	$\mathcal{U}_{s4}, y_2$	4	1	1	0	0	1

Since, $\mathcal{GR}_3$ and $\mathcal{GR}_4$ both isolate f_b from $\{f_a, f_{u_1}\}$, they have the same isolation capability and one of these can be eliminated in the fault signature matrix. Considering that $\mathcal{GR}_3$ involves more known variables than $\mathcal{GR}_4$, the latest is selected.

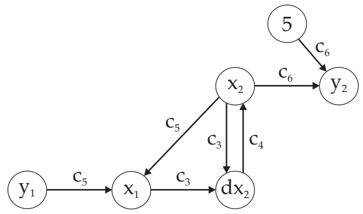

Fig. 4. Matched graph with the set $\{u_1, y_1\}$ as pseudo-exogenous variables

Table 1. Faults Signature for System (c1-c6)

$\backslash$	$\mathcal{Z}_i$	$\lvert\mathcal{C}_i\rvert$	f_{y1}	f_{y2}	f_u1	f_a	f_b
$\mathcal{GR}_1$	$\mathcal{U}_{s1}, y_1$	5	1	0	1	1	1
$\mathcal{GR}_2$	$\mathcal{U}_{s2}, y_2$	5	0	1	1	1	1
$\mathcal{GR}_4$	$\mathcal{U}_{s4}, y_2$	4	1	1	0	0	1

Integrating the selected redundant graphs, the faults signatures matrix of Table 1 is obtained, which characterizes the isolability structure, where only the actuator fault f_{u1} and the process fault f_a are not isolable.

As conclusion, for a given redundant graph $\mathcal{GR}_i$, the higher is the cardinality of its corresponding restrictions set $\mathcal{C}_i$, the lower is its isolation capability. The long paths in the graphs result at considering in $\mathcal{U}_s$ only the conventional control inputs $\{u_1, u_2, \ldots, u_l\}$. On the other hand, increasing the cardinality of $\mathcal{U}_s$ with the elements of $\mathcal{Y}$, the number of constraints involved in the resulting redundant graphs is reduced, which improves the isolability. Thus, the maximal isolability is achieved when, for the selection of any target variable y_i, the pseudo-exogenous set is conformed by $\mathcal{U}_s = \mathcal{U} \cup \{y_k \in \mathcal{Y}, y_k \neq y_i\}$. Compare the fault signature of $\mathcal{GR}_1$ and $\mathcal{GR}_2$ with $\mathcal{GR}_4$.

It is important to note that this methodology allows the search of insensitive $\mathcal{GR}'s$ to a specific fault f_j, by eliminating the set $\mathcal{C}_{fault}$ associated to f_j in the set of constraints of the bipartite graph.

On the other hand, from the $\mathcal{GR}'s$ the analytical redundancy relations (1) can be obtained using symbolic tools and constraints concatenation, as long as the parameters of the analytical model be available, this is, for the residual generation the parameterized models are used, Frisk and Åslund [2005], Staroswiecki and Comtet-Varga [2001]. However, in large scale process this idea for the residual generation could be inadequate because the analytical models are not available or uncertain.

The above described fact suggest the use of data driven based methods as an alternative for the residual generation. The argument is that given the structure of the $\mathcal{GR}'s$, the existence of a correlation between the graph signals, $\mathcal{Z}_i$, is assured. Here, specifically, a DPCA modeling for each redundant graph is performed and the square predictive error as residual is evaluated for each observation. For this proposal it is assumed as known the redundant graphs and the availability of process historical data. The details of this integration are given in the following section.

3. RESIDUAL BY DPCA STRUCTURED MODELS

Given a redundant graph

$$\mathcal{GR}_i\left(\mathcal{C}_i, \mathcal{U}_{si}, y_i\right) \tag{3}$$

which is sensitive to faults set $\mathcal{F}_i$, according to the definition (4), exist a redundancy relation of the form (1) which is also sensitive to $\mathcal{F}_i$.

If the process under supervision is stable, works around an operation point and its model is linearizable, exist a linear redundancy relation

$$\mathbf{z}_{i(w)}(t)\mathbf{a}_i = 0 \tag{4}$$

mapped from (3), with $\mathbf{a}_i \in \Re^{m\times 1}$ and the vector $\mathbf{z}_{i(w)}(t) \in \Re^{1\times m}$ written by

$$\mathbf{z}_{i(w)}(t) = \\ [\mathbf{u}_{si}(t)\ \mathbf{u}_{si}(t-1)\ldots\mathbf{u}_{si}(t-w)\ y_i(t)\ y_i(t-1)\ldots y(t-w)]$$

where: $\mathbf{u}_{si}(t) \in \Re^{1\times l}$ is associated to the pseudo-exogenous subset $\mathcal{U}_{si}$ and y_i is the target variable; $m = (l+1)(w+1)$.

Under these conditions, the residual generation for (4) can be tackled by DPCA, taking into account the auto and cross-correlations of the signals $\mathbf{z}_{i(w)}(t)$, which by simplicity we will rename as $\mathbf{z}(t)$.

The starting point to get a DPCA modeling is a set of N nominal historical data of $\mathbf{z}(t)$ which satisfy (4) and are written as a matrix

$$\mathbf{Z}(t) = \begin{bmatrix} \mathbf{z}(t) \\ \mathbf{z}(t+1) \\ \vdots \\ \mathbf{z}(t+N-1) \end{bmatrix} \in \Re^{N\times m}$$

Usually, data $\mathbf{Z}(t)$ are standardized with respect to their means $\mu_{\mathbf{z}}$, and standard deviations $\sigma_{\mathbf{z}}$ and the resulting matrix is denoted as $\widetilde{\mathbf{Z}}(t)$.

The DPCA based implicit model is derived from the eigenstructure of the correlation matrix $\mathbf{R} = \frac{1}{N-1}\widetilde{\mathbf{Z}}^T\widetilde{\mathbf{Z}}$, which can be written by

$$\mathbf{RQ} = \mathbf{Q\Lambda} \tag{5}$$

where $\mathbf{Q} \in \Re^{m\times m}$ is the orthonormal eigenvectors matrix and $\mathbf{\Lambda} \in \Re^{m\times m}$ is the diagonal matrix of the corresponding eigenvalues ordered in decreasing form: $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m$ with $m = p + r$.

Following the calculation given by Gertler and Cao [2005], and considering the linearity of the redundancy relation (4), p eigenvalues of $\mathbf{\Lambda}$ are significative and $r = w+1$ are close to zero. Thus matrix $\mathbf{Q}$ can be decomposed by

$$\mathbf{Q} = [\,\mathbf{Q}_p\ \mathbf{Q}_r\,] \tag{6}$$

where $\mathbf{Q}_p \in \Re^{m\times p}$ is the eigenvectors subset associated to the most significative eigenvalues which are a base for the named principal component subspace; and $\mathbf{Q}_r \in \Re^{m\times r}$ is the complementary eigenvectors subset associated to the zero eigenvalues which are a base for the named residual subspace.

Thus, for a given nominal time series vector $\mathbf{z}(t)$ satisfying (4), which is standardized with respect to the historical means $\mu_{\mathbf{z}}$, and standard deviations $\sigma_{\mathbf{z}}$, its projection to the residual subspace yields the residual vector

$$\mathbf{r}_i(t) = \widetilde{\mathbf{z}}(t)\mathbf{Q}_r = \mathbf{0} \in \Re^{1\times r} \tag{7}$$

This means that consistent observations are orthogonal to the residual subspace.

On the other side, considering a standardized inconsistent observation

$$\widetilde{\mathbf{z}}_f(t) = \widetilde{\mathbf{z}}(t) + \mathbf{f}_{\mathbf{z}}(t) \tag{8}$$

where $\|\mathbf{f}_{\mathbf{z}}\| \neq 0$ represents the faults effect. The projection of (8) to the residual subspace is given by

$$\mathbf{r}_i(t) = \widetilde{\mathbf{z}}_f(t)\mathbf{Q}_r = \widetilde{\mathbf{z}}(t)\mathbf{Q}_r + \mathbf{f}_{\mathbf{z}}(t)\mathbf{Q}_r = \mathbf{f}_{\mathbf{z}}(t)\mathbf{Q}_r \neq \mathbf{0} \tag{9}$$

As conclusion, the projection of any new observation $\widetilde{\mathbf{z}}(t)$ to the residual subspace can be used as a residual of (4), even if this explicit relation is unknown.

To generate a scalar residual from the projection $\mathbf{r}_i(t)$, the square predictive error is used

$$\rho_i = \mathbf{r}_i(t)\mathbf{r}_i^T(t) \tag{10}$$

Thus, the residual evaluation for each redundancy relation given in (4) can be carried out by DPCA based implicit modeling.

Because the redundancy relation (4) is an independent subsystem and $\mathbf{u}_{si}(t)$ is an external signals vector to (4), which can contain system output variables, then, in the framework of DPCA it is necessary to consider the variations of $\mathbf{u}_{si}(t)$ as 'normal', even if they are induced by changes in the operating point or by faults in other subsystems. This consideration prevents false alarms and allows to carry out the residual generation according to the fault signature matrix. This means, the DPCA algorithm, requires a standardization process with on-line estimated statistical parameters ($\mu_{\mathbf{z}}$, $\sigma_{\mathbf{z}}$), obtained from $\mathbf{u}_{si}(t)$, instead of fixed statistical values calculated off-line.

Here, it is adopted the adaptive standardization procedure proposed by Mina and Verde [2007] based in the fact that the correlation structure $\mathbf{R}$ of subsystem (4) is invariant in nominal conditions. So, the means and variances of $\mathbf{u}_{si}(t)$ are estimated on-line, respectively, using exponentially weighted moving average and exponentially weighted moving covariance; and the mean and variance of the target variable y_i is estimated from means and variances of $\mathbf{u}_{si}(t)$. Fig. 5 describes the integration of a $\mathcal{GR}_i$ with its respective residual generator based in DPCA with adaptation.

The results of the integration of DPCA and SA are demonstrated for the residual generation of a three tanks hydraulic system, in the following section.

4. THREE TANKS SYSTEM

The three tanks hydraulic system, described in Fig. 6, is composed of cylindrical tanks, interconnected at the bottom by pipes with valves V_1 in the link between tanks 1 and 3, V_3 in the link between tanks 3 and 2, and V_2 in the link between tank 2 and the outside, which can be manipulated to emulate faults (e.g. pipe blockage). The

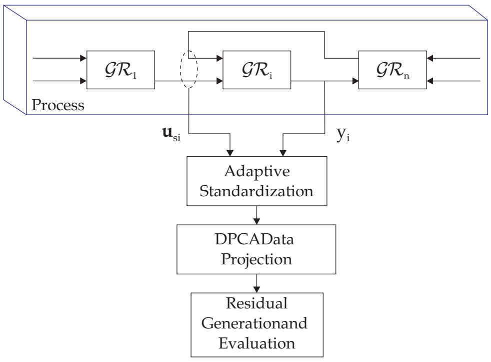

Fig. 5. FDI based in DPCA Structured Models with Adaptive Standardization

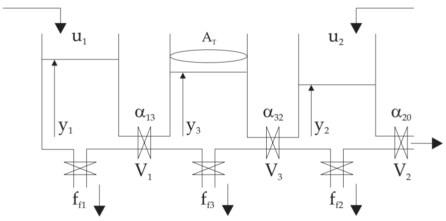

Fig. 6. Three Tanks System

system is fed by two flows to the tanks 1 and 2; these flows are controlled by pumps whose respective inputs u_1 and u_2 are known. The tanks levels y_1, y_2 and y_3 are considered as measured output variables.

In this system eleven faults are considered: three associated with the sensors $\{f_{y_1}, f_{y_2}, f_{y_3}\}$; two associated with the actuators $\{f_{u_1}, f_{u_2}\}$; three associated with blockages in the pipelines between the tanks $\{f_{\alpha_{13}}, f_{\alpha_{32}}, f_{\alpha_{20}}\}$ and three associated to leaks in the tanks $\{f_{f1}, f_{f2}, f_{f3}\}$. This FDI problem was tackled by Alcorta and Frank [1999] and solved using three non-linear observers.

The general mathematical model of this system is given by

$$\dot{x}_1 = f_1\left(x_1, x_3, u_1, f_{f1}, \alpha_{13}\right) \tag{t1}$$

$$\dot{x}_2 = f_2\left(x_3, x_2, u_2, f_{f2}, \alpha_{32}, \alpha_{20}\right) \tag{t2}$$

$$\dot{x}_3 = f_3\left(x_1, x_3, x_2, f_{f3}, \alpha_{13}, \alpha_{32}\right) \tag{t3}$$

$$x_4 = dx_1 \tag{t4}$$

$$x_5 = dx_2 \tag{t5}$$

$$x_6 = dx_3 \tag{t6}$$

$$y_1 = x_1 \tag{t7}$$

$$y_2 = x_2 \tag{t8}$$

$$y_3 = x_3 \tag{t9}$$

where α_{ij} is the flow constant for the corresponding pipe between tanks i and j.

The bipartite graph of model (t1-t9) is represented in Fig. 7. Eliminating the unknown variables $\{x_1, x_2, x_3, \dot{x}_1, \dot{x}_2, \dot{x}_3\}$ by means of a matching process for subsets of system measurable signals, ten graphs $\mathcal{GR}_i$ are obtained. Due to space limitations, only the four graphs, which achieve the maximum isolability, are presented

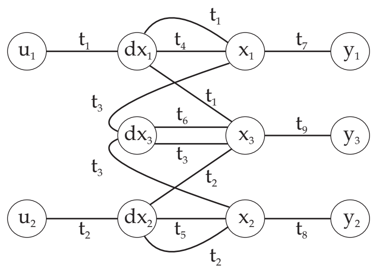

Fig. 7. Bipartite graph of the system (t1-t9)

Table 2. Faults Signature for the Three Tanks System

\	$\mathcal{GR}_1$	$\mathcal{GR}_2$	$\mathcal{GR}_3$	$\mathcal{GR}_4$
$\mathcal{Z}_i$	u_1, y_3, y_1	u_2, y_3, y_2	y_1, y_2, y_3	u_1, u_2, y_3
$\lvert\mathcal{C}_i\rvert$	4	4	5	7
f_{y1}	1	0	1	0
f_{y3}	1	1	1	1
f_{y2}	0	1	1	0
f_{u1}	1	0	0	1
f_{u2}	0	1	0	1
f_{f1}	1	0	0	1
f_{f3}	0	0	1	1
f_{f2}	0	1	0	1
$f_{\alpha13}$	1	0	1	1
$f_{\alpha32}$	0	1	1	1
$f_{\alpha20}$	0	1	0	1

$$\begin{array}{l} \mathcal{GR}_1(t_1, t_4, t_7, t_9, u_1, y_3, y_1) \\ \mathcal{GR}_2(t_2, t_5, t_8, t_9, u_2, y_3, y_2) \\ \mathcal{GR}_3(t_3, t_6, t_7, t_8, t_9, y_1, y_2, y_3) \\ \mathcal{GR}_4(t_1, t_2, t_3, t_4, t_5, t_6, t_9, u_1, u_2, y_3) \end{array} \tag{11}$$

According to the graphs in (11), the faults signature matrix, given in Table 2, is obtained. Then the isolable faults sets are $\{f_{y_1}\}$, $\{f_{y_3}\}$, $\{f_{y_2}\}$, $\{f_{u_1}, f_{f1}\}$, $\{f_{u_2}, f_{f2}, f_{\alpha_{20}}\}$, $\{f_{f3}\}$, $\{f_{\alpha_{13}}\}$, $\{f_{\alpha_{32}}\}$.

To get the training data to design the DPCA based residual generator, the system is simulated around the following operating point:

Values	Variance
$u_1 = 0.03l/s$	$\sigma^2_{u_1} = 5 \times 10^{-11}$
$u_2 = 0.02l/s$	$\sigma^2_{u_2} = 5 \times 10^{-11}$
$y_1 = 0.310m$	$\sigma^2_{y_1} = 5 \times 10^{-4}$
$y_2 = 0.130m$	$\sigma^2_{y_2} = 5 \times 10^{-4}$
$y_3 = 0.220m$	$\sigma^2_{y_3} = 5 \times 10^{-4}$

with nominal parameters $\alpha_{13} = 1.002 \times 10^{-4}$, $\alpha_{32} = 1.027 \times 10^{-4}$ and $\alpha_{20} = 1.360 \times 10^{-4}$.

Thus, considering the subsets of correlated variables $\mathcal{Z}_i$ indicated in Table 2 four residuals ρ_i are generated using DPCA according with expressions (9) and (10).

To test the diagnosis system, eight conditions are simulated; a running for each independent fault, which are activated at $40 \times 10^3\ s$. The residual response for the faults f_{y_1}, f_{y_2}, f_{y_3}, f_{u_1}, f_{u_2}, f_{f_3}, $f_{\alpha_{13}}$, $f_{\alpha_{32}}$ are shown in Fig. 8. From this figure one verifies that the residuals behavior corresponds with that indicated in Table 2. Similar results are obtained with the remaining non isolable faults.

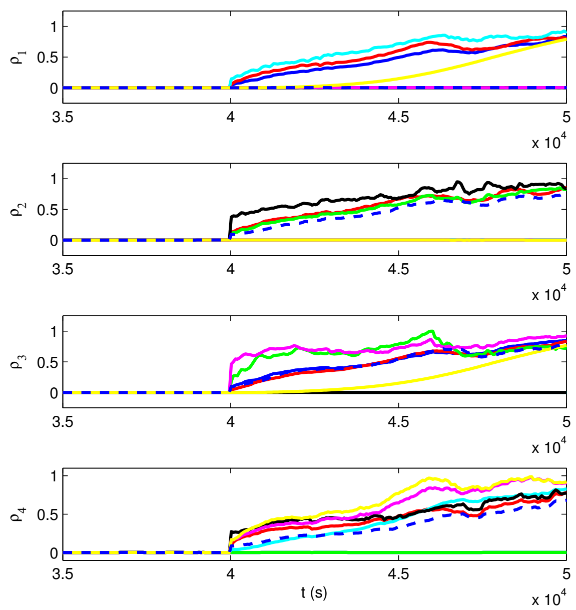

Fig. 8. Residual response for: f_{y_1} (blue –); f_{y_3} (red –); f_{y_2} (green –); f_{u_1} (cyan –); f_{u_2} (black –); f_{f_3} (magenta –); $f_{\alpha_{13}}$ (yellow –); $f_{\alpha_{32}}$ (blue - -)

5. CONCLUSIONS

This paper proposes, for FDI purposes, to carry out the *a priori* analysis of the system structural possibilities by means of graph theory tools. The resulting redundant graphs describe the isolation capability of the system and summarize the information for the implementation of the corresponding residual generators. In the case that accuracy models are not available or have the sparsity property, as it happens in large scale systems, the correlation between variables associated to the redundant graphs can be exploited to generate the residuals. Here, principal component analysis is used for the implementation of the residual generator for each redundant graph. The advantage of the proposed integration is that it only requires a generic model of the system in order to obtain the redundant graphs, and historical data for the implementation of the residual generators. The simulation results for the three tanks system benchmark show the effectiveness, for fault isolation, of the proposed algorithm with similar results to those obtained using non-linear observers with unknown inputs, however, with less effort in the design.

REFERENCES

E. Alcorta and P. M. Frank. A Novel Design of Structured Observer-Based Residuals for FDI. In *Proceedings of the American Control Conference, ACC'99*, pages 1341–1345, San Diego, CA, USA, June 1999.

M. Blanke, M. Kinnaert, J. Lunze, and M. Staroswiecki. *Diagnosis and Fault Tolerant Control.* Springer Verlag, Berlin, 2003.

L. H. Chiang, E. L. Russell, and R. D. Braatz. *Fault Detection and Diagnosis in Industrial Systems.* Advanced Textbooks in Control and Signal Processing, Springer-Verlag, London, 2001.

C. De Persis and A. Isidori. A Geometric Approach to Nonlinear Fault Detection and Isolation. *IEEE Transactions on Automatic Control*, 46(6):853–865, June 2001.

D. Düstegör, E. Frisk, V. Cocquempot, M. Krysander, and M. Staroswiecki. Structural Analysis of Fault Isolability in the DAMADICS Benchmark. *Control Engineering Practice*, 14(6):597–608, 2006.

E. Frisk and J. Åslund. Lowering Orders of Derivatives in Non-Linear Residual Generation using Realization Theory. *Automatica*, 41(10):1799–1807, October 2005.

S. Gentil, J. Montmain, and C. Combastel. Combining FDI and AI Approches within Causal Model Based Diagnosis. *IEEE Transactions on Systems Man and Cybernetics*, 34(5):2207–2221, 2004.

J. Gertler and J. Cao. Design of Optimal Structured Residuals from Partial Principal Component Models for Fault Diagnosis in Linear Systems. *Journal of Process Control*, 15(5):585–603, August 2005.

J. W. de V. Groenewald, L. P. Coetzer, and C. Aldrich. Statistical Monitoring of a Grinding Circuit: An Industrial Case Study. *Minerals Engineering*, 19(11):1138–1148, 2006.

R. Isermann. *Fault Diagnosis Systems - An Introduction from Fault Detection to Fault Tolerance.* Springer Verlag, Germany, 2006.

J. V. Kresta, J. F. MacGregor, and T. E. Marlin. Multivariate Statistical Monitoring of Process Operating Performance. *The Canadian Journal of Chemical Engineering*, 69:35–47, February 1991.

M. Krysander and M. Nyberg. Structural Analysis Utilizing MSS Sets with Application to a Paper Plant. In *Proceedings of 13th International Workshop on Principles of Diagnosis, DX'02*, Semmering, Austria, 2002.

T. Lorentzen, M. Blanke, and H. Niemann. Structural Analysis. A Case Study of the Romer Satellite. In *Proceedings of the 5th IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes, SAFEPROCESS'03*, pages 187–192, Washington, USA, June 2003.

M. Massoumnia, G. Verghese, and A. Willsky. Fault Detection and Identification. *IEEE Transactions on Automatic Control*, 34(3):316–321, March 1989.

J. Mina and C. Verde. Fault Detection for MIMO Systems Integrating Multivariate Statistical Analysis and Identification Methods. In *Proceedings of the American Control Conference, ACC'07*, pages 3234–3239, New York, USA, 11-13 July 2007.

M. Nyberg. A Generalization of the GDE Minimal Hittings-Set Algorithm to Handle Behavioral Modes. In *Proceedings of the 14th International Workshop on Principles of Diagnosis, DX'2006*, Spain, 2006.

A. Raich and A. Çinar. Statistical Process Monitoring and Disturbance Diagnosis in Multivariable Continuous Processes. *AIChE Journal*, 42(4):995–1009, April 1996.

M. Staroswiecki and G. Comtet-Varga. Analytical Redundancy Relations for Fault Detection and Isolation in Algebraic Dynamic Systems. *Automatica*, 37(5):687–699, May 2001.

H. Wang, Z. Song, and H. Wang. Statistical Process Monitoring Using Improved PCA with Optimized Sensor Locations. *Journal of Process Control*, 12(6):667–751, September 2002.