Detection, isolation and fault estimation of nonlinear systems using a directional study
Maya Kallas, Gilles Mourot, Didier Maquin, José Ragot

To cite this version:
Maya Kallas, Gilles Mourot, Didier Maquin, José Ragot. Detection, isolation and fault estimation of nonlinear systems using a directional study. 12th European Workshop on Advanced Control and Diagnosis, ACD 2015, Nov 2015, Pilsen, Czech Republic. 10.1088/1742-6596/659/1/012032. hal-01221044

HAL Id: hal-01221044
https://hal.science/hal-01221044v1
Submitted on 25 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Detection, isolation and fault estimation of nonlinear systems using a directional study

Maya Kallas, Gilles Mourot, Didier Maquin, José Ragot

Abstract. In terms of system diagnosis, several studies are generally performed. The diagnosis is composed of three different parts: detecting, isolating and estimating the value of the faults. If many results have been obtained for linear systems with a known model, the situation is quite different in the case of nonlinear systems behavior, especially when the model is not known a priori. This paper proposes to discuss the latter case using a study of the dissimilarities between data. The dissimilarities are evaluated by a nonlinear function of the Euclidean distances. To this end, a radial basis function is used, and a directional study is introduced for fault diagnosis. The relevance of the proposed technique is illustrated on simulated data.

1. Introduction

The supervision of systems covers various additional aspects, among which the diagnosis occupies a privileged place. Based on this statement, there exists a set of techniques used to know the status of a given system. In the state estimation step, the very important aspect is the detection of functioning changes due for example to sensors failure or actuators or even system components. Early detection of these defects can be used to react on the commands of the system so as to minimize the impact of the faults. Numerous techniques have been developed in recent decades, which differ depending on whether a model of the system is known a priori or not. This paper is concerned with the second situation, where only input and output measures of the system are available. Diagnosis strategies are thus based on the analysis of these data. The principal component analysis (PCA) is an example of a well-known technique for diagnosis. PCA and its variants have been the subject of many developments.

The PCA has been widely used for the detection and isolation of sensor failure [11], [13] and more generally for system supervision [2], [15]. As in the case of detection method using a model known a priori, one of the best performing PCA technique is based on the comparison between the value of a variable at a given time and the reconstructed value of this variable from the other variables. This reconstruction uses the eigenvectors of the covariance matrix of a dataset of observations, these eigenvectors belonging to the subspace called secondary or residual obtained by PCA. As indicated in [7], it is possible to extend this reconstruction to a system with \( n \) variables, subject to observability, where we can reconstruct simultaneously \( p \) variables out of the \( n - p \) remaining ones, based on the fact that we have redundancy between some data in the whole dataset defined by the \( n \) variables.

Despite the proven performance of this technique, it is necessary to recall that it is applied on a system whose variables are mainly related linearly. Indeed, the PCA is a method of linear projection that cannot consider or reveal nonlinearities. The nonlinearity issue has motivated a
number of work leading to the nonlinear PCA, or the use of kernel machines such as the kernel PCA (KPCA). In the field of diagnosis, the KPCA has been widely used to detect outliers from faulty sensors for example [6]. KPCA has been used to analyze partially observed systems, in applications of missing data, such as in image processing [3], [12]. The combination, for fault detection, of KPCA and discriminant analysis techniques, was proposed in data processing to increase the performance of each method [5], [10].

To address the difficult problem of isolating faults affecting the data, diagnosis techniques developed in a linear framework rely heavily on the concept of structuring fault indicators. The extension to the nonlinear case remains delicate due to the difficulty of the problem called pre-image (where the existence of a solution is not guaranteed a priori) but has been the subject of several works [1]. The principle used in [6] consists on applying KPCA technique to several subsets of data, each using a part of the original variables. The difficulty of such proposition is to build these subsets.

Based on the literature, it is clear that the detection of outliers is widely studied. However, only few papers propose new nonlinear technique for isolation and fault amplitude estimation affecting one or more variables. In [1], [11] and [8], the pre-image technique offers methods for isolation and estimation of faults.

As aforementioned, the PCA and its nonlinear extension KPCA have been studied in all their aspects in terms of fault detection, isolation and estimation. However, there exists a more general technique than PCA, the principal coordinate analysis (PCO) or the well-known multidimensional scaling (MDS) [16]. The main difference between the PCA and PCO can be given by three points. First, PCA seeks patterns in the variables, while PCO searches for similarities between cases. Second, PCA uses an eigen decomposition of a covariance matrix in order to reduce variable dimensionality, while PCO analyzes a distance matrix. Many different distance matrices can be used as long as they are metric. Third, the result of a PCO is a set of coordinates on a number of derived axes such that similar cases are close together. It is not possible to associate these axes with any variables. This technique was not used for diagnosis purposes. Therefore, we propose to study similarities between data. The aim of the scaling is to find the distances that well approximate the dissimilarities of data [14]. In [4], Critchley suggested a power transformation of the distances, if the dissimilarities are derived from Euclidean distances.

Based on the suggestion of Critchley, we propose in this paper a modified MDS study in order to evaluate the dissimilarities of data. We examine the possibility to detect and isolate the sensors providing faulty measures. Based on real non faulty dataset, the proposed technique is to compare the test dataset to these reference data. Because of the nonlinear behavior of systems that are of interest, this comparison will be performed in a space of suitable structure. More precisely, after selecting the appropriate structure of the kernel, the initial reference dataset are used to construct a Gram matrix, in order to represent the original data in a space of larger dimension. The new observations to be analyzed are then projected into this space and their comparison with the reference observations is done through their projections. From this comparison, we can discriminate new healthy observations from faulty ones (section 2.2). In the case where an observation is found abnormal, an isolation of the faulty variable is proposed in (section 2.3), and finally an estimate of the fault amplitude is examined in (section 2.4).

In this paper, the Gaussian kernel is applied. It is important to mention a difficulty when using this kind of kernels. It is a kernel with a decreasing bandwidth depending on the distance to its centers. Therefore, a faulty measure may no longer be in the area of influence of the kernel and thus may have near zero projection in the kernel space, making it difficult to locate the fault. To overcome this difficulty, a sequential search of one or more directions of fault is proposed.

As many fault detection methods, the proposed approach has a number of hyperparameters. It is necessary to adjust them depending on the problem to be treated. The hyperparameters,
in this paper, are the number of principal components to retain from the decomposition of the Gram matrix, and finally the bandwidth of the kernel. The influence of these parameters is studied through the example in section 3, but in the present form of this paper, their choice is done experimentally.

2. Detection, isolation and estimation of the faults

2.1. Notations

Let us consider a set of healthy measures, denoted \( x_i \). They are taken from the system and they form the reference database:

\[
X = \begin{bmatrix} x_1^T \\ \vdots \\ x_N^T \end{bmatrix}, \quad x_i \in \mathbb{R}^n, \quad X \in \mathbb{R}^{N \times n}
\]  

(1)

Based on using the idea of dissimilarities, and from the suggestion of Critchley [4], let us consider a nonlinear function \( \kappa \), also called a kernel. By kernel, we mean a function from \( X \times X \) into \( \mathbb{R} \) with a Hermitian symmetry, that is \( \kappa(x_i, x_j) = \kappa(x_j, x_i) \). Using this definition, we can mention the Radial Basis Function (RBF), and more precisely, the gaussian kernel, which is widely used. The form of this gaussian kernel is characterized by its bandwidth, or hyper-parameter \( c \),

\[
\kappa(x_i, x_j) = \exp \left( -\frac{\| x_i - x_j \|^2}{c} \right)
\]

(2)

During an online diagnosis, we receive a new observation \( x \in \mathbb{R}^n \) to be tested. The analysis of this observation is done in order to detect the possible presence of errors in one or more variables of this measure. To perform this analysis, we need to evaluate the following expressions:

\[
\kappa(x) = \begin{bmatrix} \kappa(x_1, x) \\ \vdots \\ \kappa(x_N, x) \end{bmatrix}, \quad \kappa(x) \in \mathbb{R}^N
\]

\[
\kappa(x, x_i) = \exp \left( -\frac{\| x_i - x \|^2}{c} \right), \quad \kappa(x_i, x) \in \mathcal{R}
\]

(3)

We already have the reference database. Thus, we can easily create the Gram matrix \( K(X) \) using

\[
K(X) = \begin{bmatrix} \kappa(x_1, x_1) & \ldots & \kappa(x_1, x_N) \\ \vdots & \ddots & \vdots \\ \kappa(x_N, x_1) & \ldots & \kappa(x_N, x_N) \end{bmatrix}, \quad K(X) \in \mathbb{R}^{N \times N}
\]

(4)

As we can see from the expressions of \( \kappa(x_i, x_j) \), the bandwidth \( c \) of the kernel influences the results. It is relative to the proximity between two observations \( x_i \) and \( x_j \).

If a fault affects a specific variable for a given observation \( x \), it can be defined by:

\[
x = x^* + \xi f
\]

(5)

where \( x \in \mathbb{R}^n \) is the faulty measure, \( x^* \in \mathbb{R}^n \) is the true value of the variable, \( f \in \mathbb{R}^{n_f} \) is a \( n_f \)-dimensional vector of faults amplitude, and \( \xi \in \mathbb{R}^{n_f \times n_f} \) their directions. For this paper, we take into consideration one fault, so \( n_f = 1 \), making it easier to explain. For example, \( \xi = (0 \ 0 \ 1 \ 0)^\top \) is a situation where we have a system with four variables, where a fault is occurring on the third one.

Diagnosis represents different tasks to be done. It is to decide whether the measure \( x \) is faulty, and if so, to determine the direction \( \xi \) of the detected fault, namely the faulty variable or component, and finally to estimate the fault amplitude \( f \). In our case, where \( n_f = 1 \), \( f \) is considered to be a scalar.
2.2. Analysis of $\kappa(x)$ for fault detection

The analysis of the validity of a new observation $x$ is performed in the kernel space. We have evaluated the matrix $K(X)$ given by (4), which is the image of the matrix $X$ in the kernel space. Therefore, it is quite natural to compare the gap between the new observation $x$ and the database $X$. Working in the kernel space, we evaluate $\kappa(x)$ and $K(X)$ representing the healthy dataset. To facilitate this comparison, we reduce the complexity of $K(X)$ by considering its spectral decomposition

$$K(X) = P \Lambda P^\top$$  \hspace{1cm} (6)

where $P$ and $\Lambda$ contain respectively the eigenvectors and eigenvalues of the matrix $K(X)$, sorted by descending order of the magnitude of the eigenvalues. By neglecting the very small values of the eigenvalues, we take the $\ell$ eigenvectors, denoted $P_\ell$, associated to the $\ell$ largest eigenvalues represented by $\Lambda_\ell$. We propose an approximation of the matrix $K(X)$ using the following expression

$$K_\ell(X) = P_\ell \Lambda_\ell P_\ell^\top$$  \hspace{1cm} (7)

We seek to decompose $\kappa(x)$ given by (3) on the basis of the $P_\ell$ eigenvectors. To this end, we consider the criterion $\psi$ defined by

$$\psi = \| \kappa(x) - P_\ell \theta \|^2$$  \hspace{1cm} (8)

We seek its minimum, for a certain value of $\theta$

$$\hat{\theta} = \left( P_\ell^\top P_\ell \right)^{-1} P_\ell^\top \kappa(x)$$  \hspace{1cm} (9)

Thus, the minimum is obtained by replacing (9) in (8), as follows

$$\hat{\psi} = \kappa^\top(x) \left( I - P_\ell \left( P_\ell^\top P_\ell \right)^{-1} P_\ell^\top \right) \kappa(x)$$  \hspace{1cm} (10)

The value of the criterion $\hat{\psi}$ is closely related to those of $\kappa(x)$ and thus the measure $x$ itself. One can study for an observation $x$, on the one hand the value taken by $\kappa(x_i, x)$ and on the other hand, the value taken by $\hat{\psi}$. In particular, an observation $x$ for which a fault affects one variable, will be revealed by the values close to zero of the $\hat{\psi}$ criterion and the component of the $\kappa(x)$ vector. For the latter, a more comprehensive test is considered based on the sum of its parts

$$S = \sum_{i=1}^{N} \kappa(x_i, x)$$  \hspace{1cm} (11)

Nevertheless, for a healthy observation $x$, the values of $S$ and $\hat{\psi}$ are significantly greater than zero. To decide the inconsistency of a measure, that is to say, the presence of a fault on one of its components, it is proposed to use the following two tests

$$\hat{\psi} < \delta_1$$  \hspace{1cm} (12a)
$$S < \delta_2$$  \hspace{1cm} (12b)

where $\delta_1$ and $\delta_2$ are two thresholds, having small values. They are chosen based on the original dataset during the learning stage. These thresholds are evaluated as the mean of different tests on many healthy observations.
2.3. Directional analysis of $\kappa(x)$ for fault isolation

As indicated in the definition (3) of $\kappa(x)$, the same weight is assigned to all components of the vectors involved in calculating the distances. In the case where one of the components of a test observation $x$ is corrupted by a fault, the distance between this observation and those of reference $x_i$ can become important. This can make $\kappa(x_i, x)$ close to zero, which provides little information, except for the fact to say that this observation is faulty.

To overcome this problem, the solution would be to increase the bandwidth of the kernel function. However, this solution is hardly possible because it leads to a very particular Gram matrix, from a numerical point of view, because the obtained matrix would be nearly an identity matrix. Therefore, the main idea of the kernel methods is not represented, and it will be as we are in a linear case. However, using a rather restrictive assumption on the limited number of faulty variables on the observation $x$, the evaluation of the distance can be performed with a particular metric by not taking into account some variables. To this end, $\kappa(x_i, x)$ is changed to this form

$$
\kappa_d(x_i, x) = \exp\left(-\frac{(x_i - x)\top W_d (x_i - x)}{c}\right)
$$

where the diagonal matrix $W_d$ contains 1 everywhere on its diagonal, except the component of rank $d$ that is chosen to be zero. Thus, the distance between the observations $x$ and $x_i$ takes into account all the variables except the one of rank $d$. Therefore, if this variable is faulty, it will not be taken into account in the evaluation of the distance and the difficulty aforementioned about the null value of $\kappa(x_i, x)$ is overlapped. Obviously, as the faulty variable is not known a priori, and in the case of just one faulty variable, it is necessary to examine all $n$ matrices $W_d$, where $d = 1, \ldots, n$.

We also note that the weighting matrix $W_d$ is directly related to the direction of a fault. Each direction, introduced in (5), corresponds in fact to a vector $\xi_d$ of the canonical basis of $\mathbb{R}^n$

$$
\xi_d = (\delta_{1,d} \ \delta_{2,d} \ \ldots \ \delta_{n,d}), \quad d = 1, \ldots, n
$$

where $\delta_{i,d}$ is the Kronecker delta, defined by

$$
\delta_{i,j} = \begin{cases} 
0 & \text{if } i \neq j \\
1 & \text{if } i = j 
\end{cases}
$$

Therefore, the weighting matrix $W_d$, that can overcame the variable of rank $d$ can be defined as

$$
W_d = I_n - \xi_d \xi_d\top
$$

The rest of the procedure, proposed in section 2.2, is applied with these new definitions of kernel space given in (13). We can check for each test case $x$, the values taken by $\kappa_d(x_i, x)$ and $\hat{\psi}_d$ for different weighting matrices $W_d$ where $d = 1, \ldots, n$. In particular, an observation $x$ whose $p$ rank variable is faulty, is identified by near-zero values of the components of $\kappa_d(x)$ and a value close to zero of $\hat{\psi}_d$ for $d = 1, \ldots, n, d \neq p$. As before, a more general evaluation of $\kappa_d(x)$ is suggested by defining the scalar indicator

$$
S_d = \sum_{i=1}^{N} \kappa_d(x_i, x), \quad d = 1, \ldots, n
$$

Given the different expressions of $S_d$ according to the variables used, the truth table 1 summarizes the possible situations. In this table, the quantities $\delta x_i$, for $i = 1, \ldots, n$ represent
Table 1. Theoretical fault signatures

| $\delta x_1$ | $\delta x_2$ | $\delta x_3$ | \ldots | $\delta x_n$ |
|-------------|-------------|-------------|---------|-------------|
| $S_1$       | $\times$    | 0           | 0       | \ldots      | 0           |
| $S_2$       | 0           | $\times$    | 0       | \ldots      | 0           |
| $\vdots$    | $\vdots$    | $\vdots$    | $\vdots$| \ldots      | $\times$    |
| $S_n$       | 0           | 0           | 0       | \ldots      | $\times$    |

the faults affecting the components of the observation $x$, where the 0 symbol indicates that a fault makes the indicator $S$ a very small value.

In summary, the steps of isolating a faulty variable for a given observation $x$ can be summarized as

- Create the matrices $W_d$ (15) for all possible directions $\xi_d, d = 1, \ldots, n$
- Evaluate $\kappa_d(x_i, x)$ given by (13) for all fault directions and any reference observation $x_i$
- Compute the indicator $S_d$ for $d = 1, \ldots, n$ given by (16) for fault presence of observation $x$
- If one of the indicators $S_d$ for $d = 1, \ldots, n$, only the $p$ rank is greater than the selected threshold, then the $p^{th}$ variable of the observation $x$ is faulty

2.4. Estimation of the fault amplitude

This step finalizes the diagnosis since it completes the detection and fault isolation by estimating their amplitudes. The direction $\xi_d$ of the fault is known, it is determined in the previous step. The objective is to estimate its amplitude $f$. To this end, we use the redundancy related to the kernel space. We know that any vector $K(x)$ can be formed from the column vectors of the matrix $K(X)$, or its approximate form $K_{\ell}(X)$ given by (7), where $x$ is a vector of measures and $X$ the reference healthy database. In reality, once the observation $x$ is considered to be a faulty measure, we evaluate its estimation with:

$$x^* = x - \xi f$$

(17)

Thus, we can evaluate the image of the corrected observation in the kernel space, using:

$$\kappa(x^*) = \left[ \kappa(x_1, x - \xi f) \right]$$

$$\kappa(x_i, x - \xi f) = \exp\left(-\frac{\|x_i - x + \xi f\|^2}{c}\right)$$

(18)

As aforementioned, the best value of $f$ is the one that helps explaining $\kappa(x^*)$ from the column vectors of the matrix $K_{\ell}(X)$. From this statement, the value of $f$ is evaluated by maximizing with respect to $f$ the criterion:

$$\psi = \|K_{\ell}(X)\ k(x - \xi f)\|^2$$

(19)

The optimal value $\hat{f}$ of $f$ is the solution to the following equation:

$$\frac{\partial\kappa^\top(x - \xi f)}{\partial f} K_{\ell}^2(X) \kappa(x - \xi f) = 0$$

(20)
Since the derivative of \( \kappa \) with respect to \( f \) is given by

\[
\frac{\partial \kappa(x_i, x - \xi f)}{\partial f} = -\frac{2}{c} \kappa(x_i, x - \xi f) \xi^\top (x_i - x + \xi f)
\]  

(21)

Therefore, we can write

\[
\frac{\partial \kappa(x - \xi f)}{\partial f} = -\frac{2}{c} (b(x - \xi f) + \kappa(x - \xi f) f)
\]  

(22)

where

\[
b(x - \xi f) = \begin{pmatrix}
\kappa(x_1, x - \xi f) \xi^\top (x - x_1) \\
\vdots \\
\kappa(x_N, x - \xi f) \xi^\top (x - x_N)
\end{pmatrix}
\]  

(23)

Using these expressions, we can write (20) as:

\[
b^\top (x - \xi f) K_f^2 \kappa(x - \xi f) = \kappa^\top (x - \xi f) K_f(X)^2 \kappa(x - \xi f) f
\]  

(24)

Using (24), we can derive \( f \) with a fixed point iterative technique \( f = g(f) \) starting with an initial point as mentioned in [1] and [9]. The final expression \( \hat{f} \) of the estimated value of \( f \) can be derived, and the estimation is given by:

\[
\hat{f}^{(i+1)} = \frac{b^\top (x - \xi f^{(i)}) K_f^2(X) \kappa(x - \xi f^{(i)})}{\kappa^\top (x - \xi f^{(i)}) K_f^2(X) \kappa(x - \xi f^{(i)})}
\]  

(25)

3. Experiments
In this section, we present the results obtained on an example deliberately chosen to be of small dimensions. It is used to illustrate the above principles of detection, isolation and fault estimation of outliers given by faulty sensors.

3.1. Simulated data
A set of 100 observations has been generated from a nonlinear system with \( n = 3 \) variables, where \( x_1 \) are randomly selected in \([-7, 7]\). The system is defined by:

\[
\begin{align*}
x_2 &= 0.15 x_1^2 - 0.5 \\
x_3 &= x_1 + x_2
\end{align*}
\]  

(26)

However, the system is considered to be noisy. Therefore, the variables have been modified with additive noises generated randomly, as follows:

\[
\begin{align*}
\epsilon_i &\sim \mathcal{N}(0, 0.01), & i = 1, 2, 3 \\
x_1 &= x_1 + \epsilon_1 \\
x_2 &= x_2 + \epsilon_2 \\
x_3 &= x_3 + \epsilon_3
\end{align*}
\]  

(27)

Another set of \( N_t = 9 \) observations, denoted \( x'_j, j = 1, \ldots, N_t \), have been generated using the model (26) for testing purpose. The nine observations are divided into two groups: faulty and correct values. Figure 1 illustrates these observations and more precisely the red dot representing the faulty observations 1, 4, 7 and 9. The faulty ones have been corrupted as follows:

- \( x'_1 \): adding 4 on the amplitude of the variable 1
- \( x'_2 \): adding -4 on the amplitude of the variable 2
- \( x'_3 \): adding 4 on the amplitude of the variable 3
- \( x'_4 \): modifying the amplitude of the variables 1, 2 and 3
3.2. Detecting outliers

Figure 1 displays the $N$ measures with the blue dots. It contains three subfigures in order to represent the three variables into three different 2D dimensions, namely, $\{x_1, x_2\}$, $\{x_2, x_3\}$ and $\{x_3, x_1\}$. This database is the reference dataset. It is used to test the proposed technique in order to detect the outliers.

![Figure 1](image_url)

**Figure 1.** Representation of the simulated data in three subfigures. The blue dots represent the original database, green dots are the new non faulty observations, and the red dots are the new faulty ones. Each subfigure represents a different 2D representation of the 3D simulated data.

Of course, visual inspection of this figure allows direct isolation of the outliers from the real measures. For example, the observation 4 is abnormal and this is due to the variable 2 since in the 2D $\{x_3, x_1\}$ representation, the representative point of this observation can be considered as belonging to the field of true observations. However, in the 2D representations based on $\{x_1, x_2\}$ and $\{x_2, x_3\}$, the same point does not belong to the field of non faulty observations.

Figures 2 and 3 visualize changes in criterion $\hat{\psi}$ resulting from the approximation of $\kappa(x)$ in the kernel space. This criterion was evaluated for two different values of the bandwidth of the kernel function (0.5, 0.7), taking into account 30 principal components, for each of the $N_t$ observations representing outliers. They are identified by numbers 1, 4, 7 and 9. For very small values of the bandwidth less than 0.2, the criterion fails to indicate the real presence of fault by making almost all values near zero. For large values greater than 0.8, the threshold is difficult to estimate due to the extreme variations between values.

Given these figures, there is a clear separation between the values of the criteria related to the outliers (1, 4, 7 and 9) and the values of the non faulty observations. This statement is true for different numbers of principal components and bandwidth of the kernel function. This procedure is therefore quite relevant for the detection of outliers. However, no information can be given concerning the faulty variables.
3.3. Isolation of the faulty variables

In order to isolate the faulty variable, the indicator $S_d$ defined by (16), was evaluated for all the directions $d = 1, \ldots, n$ on the test set. To identify the observation, the indicator is denoted $S_{d,j}$ with $j = 1, \ldots, N_t$. To make the interpretation and comparison of $S_{d,j}$ values easier, for the various observations and directions, the values of this indicator have been normalized. First, these values are written as $S_{d,j}$, for $d = 1, \ldots, n$, and $j = 1, \ldots, N_t$. The comparison is based on the normalized value of the indicator given by:

$$
\overline{S}_{d,j} = \frac{S_{d,j}}{\max_j S_{d,j}}
$$

(28)

Figures 4 and 5 concern the values of the normalized test indicator $\overline{S}_{d,j}$, defined in (28). It is evaluated according to the three directions of the data space. The comparison is based on a single kernel bandwidth with two different values of the number of principal components. For the case, where the number of principal components is correctly chosen, we can evaluate the indicator $\overline{S}_{d,j}$ in order to isolate the faulty direction. If the number of principal components is too small, the approximation of the initial dataset is wrong, and the information is lost. This number should be acceptable in order to maintain the main information needed to estimate the original dataset.

Each figure contains three subfigures showing the values of the fault indicator $\overline{S}_{d,j}$ for the $N_t = 9$ test observations considering the $n = 3$ fault directions. More explicitly, the matrices $W_d$ introduced aforementioned using (13) are defined by:

$$
W_1 = \text{diag}(0, 1, 1)
$$
$$
W_2 = \text{diag}(1, 0, 1)
$$
$$
W_3 = \text{diag}(1, 1, 0)
$$

(29)
Therefore, for the metric $W_1$, the variable 1 is not involved in evaluating distances. Thus, the Gram matrix is insensitive to errors that can affect this variable. This perfectly explains the three values taken by the indicator $\overline{S}_{d,j}$. For example, for the first measure, $\overline{S}_{2,1}$ and $\overline{S}_{3,1}$ are evaluated from the variables 1, 3 and 1, 2 respectively. The fact that $\overline{S}_{2,1}$ and $\overline{S}_{1,1}$ provide practically zero values, whereas $\overline{S}_{3,1}$ is greater than zero, is due to the presence of the variable 1 that is common to the first two indicators, this variable being the faulty one. Moreover, for the observation $x_9$, we can detect the presence of fault, due to the zero values for each direction, however we can’t isolate and indicate which direction is the faulty one. This can be explained by the fact that there exists different simultaneous faults affecting many directions.

Numerous tests were conducted with different values of the number of principal components and with different values of the kernel bandwidth. For a wide range of values for these parameters, the shapes of the preceding graphs are retained.

To analyze all the test observations, we construct the table 2 representing the theoretical signatures. This table shows the influence of an error impacting one of the three variables on the $S_d$ indicators.

**Table 2. Theoretical fault signatures: influence of a unique fault affecting each variable on the $S_d$ indicator**

| Variables | 1 | 2 | 3 |
|-----------|---|---|---|
| $S_1$     | × | 0 | 0 |
| $S_2$     | 0 | × | 0 |
| $S_3$     | 0 | 0 | × |

Then, from figures 4 and 5, we construct the table 3 which describes the experimental fault signatures. Two symbols have been used. The symbol 0 illustrates that a defect causes a near zero value of $\kappa$ and thus the indicator $S$. However, the cross indicates that there is no faulty observation or measure. Once this table is constructed, we compare the experimental signature of each of the nine observations to the three theoretical signatures given in table 2, making it easier to detect in each case the variable presenting a faulty measure. Clearly, for observations 1, 4, and 7, the variables 1, 2, 3 are respectively detected as faulty.

**Table 3. Experimental fault signatures**

| observations | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|--------------|---|---|---|---|---|---|---|---|---|
| $S_1$        | × | × | × | 0 | × | × | 0 | × | 0 |
| $S_2$        | 0 | × | × | × | × | 0 | × | 0 |
| $S_3$        | 0 | × | × | 0 | × | × | × | 0 |

### 3.4. Fault amplitude estimation

Once the faulty observations have been detected (observations 1, 4, 7 and 9), and the variable isolation is performed, we, thus, proceed to estimate the amplitude of the fault. We apply the technique proposed in section 2.4. Table 4 illustrates the results, where we compare the true value of the fault and the estimated one.
Table 4. Real $f$ and estimated $\hat{f}$ fault amplitudes

|       | $f$  | $\hat{f}$ |
|-------|------|-----------|
| obs 1, var 1 | 4.03 | 3.98      |
| obs 4, var 2 | -4 | -4.01     |
| obs 7, var 3 | 4.03 | 3.95      |

4. Conclusion

In this paper, we propose to diagnose a nonlinear system using a directional study. To this end, the technique proposed in order to detect outliers derived from a nonlinear system is based on analysis of the observations in the kernel space. Three aspects are presented: the detection of faulty measures, the variable isolation of these faulty measures and the estimation of their fault amplitude. Therefore, we analyze the projections of the observations onto the kernel space. We take into consideration projections along directions favoring the recognition of healthy variables.

As for future work, a study of simultaneous many faulty variables will be derived, specifying the conditions for isolability. Moreover, an extension to nonlinear principal component regression can be proposed.

References

[1] C.F. Alcala and S.J. Qin. Reconstruction-based contribution for process monitoring with kernel principal component analysis. Industrial & Engineering Chemistry Research, 49(17):7849–7857, 2010.
[2] H.B. Aradhye, B.R. Bakshi, and R. Straus. Process monitoring by PCA, dynamic PCA, and multiscale PCA- theoretical analysis and disturbance detection in the tennessee eastman process. Technical report, AIChe Annual Meeting, 1999.
[3] M. Berar, M. Desvignes, G. Bailly, Y. Payan, and B. Romanui. Missing data estimation using polynomial kernels. In S. Singh, M. Singh, C. Apte, and P. Perner, editors, International Conference on Advances in Pattern Recognition (1), volume 3686 of Lecture Notes in Computer Science, pages 390–399. Springer, 2005.
[4] F. Critchley. Multidimensional scaling: a short critique and a new method. In L.C.A. Corsten and J. Hermans, editors, COMPSTAT, pages 297–303. Physica, 1978.
[5] P. Cui and J. Fang. KPCA plus FDA for fault detection. In D. Liu, S. Fei, Z. Hou, H. Zhang, and C. Sun, editors, 4th International Symposium on Neural Networks : Advances in Neural Networks, volume 4493 of Lecture Notes in Computer Science, pages 597–606. Springer Berlin Heidelberg, 2007.
[6] K. Fu, L. Dai, T. Wu, and M. Zhu. Sensor fault diagnosis of nonlinear processes based on structured kernel principal component analysis. Journal of Control Theory and Applications, 7(3):264–270, 2009.
[7] M.F. Harkat, G. Mourot, and J. Ragot. An improved PCA scheme for sensor FDI: Application to an air quality monitoring network. Journal of Process Control, 16(6):625–634, July 2006.
[8] M. Kallas, G. Mourot, D. Maquin, and J. Ragot. Diagnosis of nonlinear systems using kernel principal component analysis. In 11th European Workshop on Advanced Control and Diagnosis, ACD, Berlin, Germany, 2014.
[9] M. Kallas, G. Mourot, D. Maquin, and J. Ragot. Fault estimation of nonlinear processes using kernel principal component analysis. In 14th European Control Conference, Linz, Austria, 2015.
[10] Z. Li, G. Tan, and Y. Li. Fault diagnosis based on improved kernel fisher discriminant analysis. Journal of software, 7(12):2657–2662, 2012.
[11] S.J. Qin. Statistical process monitoring : basis and beyond. Journal of Chemometrics, 17(8-9):480–502, 2003.
[12] M. Scholz, F. Kaplan, C.L. Guy, J. Kopka, and J. Selbig. Non-linear PCA: a missing data approach. Bioinformatics, 21(20):3887–3895, 2005.
[13] M. Tamura and S. Tsujita. A study on the number of principal components and sensitivity of fault detection using PCA. Computers & Chemical Engineering, 31(9):1035–1046, 2007.
[14] C.K.I. Williams. On a connection between kernel pca and metric multidimensional scaling. In Advances in Neural Information Processing Systems 13, pages 675–681. MIT Press, 2001.
[15] J. Xu, J. Zhao, B. Ma, and S. Hu. Fault diagnosis of complex industrial process using KICA and sparse SVM. In Mathematical Problems in Engineering, 2013.
[16] A.F. Zuur. Principal coordinate analysis and non-metric multidimensional scaling. In *Analysing Ecological Data*, Statistics for Biology and Health, pages 259–264. Springer New York, 2007.