Diagnosis of nonlinear systems using kernel principal component analysis
Maya Kallas, Gilles Mourot, Didier Maquin, José Ragot

To cite this version:
Maya Kallas, Gilles Mourot, Didier Maquin, José Ragot. Diagnosis of nonlinear systems using kernel principal component analysis. 11th European Workshop on Advanced Control and Diagnosis, ACD 2014, Nov 2014, Berlin, Germany. 10.1088/1742-6596/570/7/072004. hal-01078401

HAL Id: hal-01078401
https://hal.science/hal-01078401v1
Submitted on 26 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.

Distributed under a Creative Commons Attribution 4.0 International License
Diagnosis of nonlinear systems using kernel principal component analysis

M. Kallas, G. Mourot, D. Maquin and J. Ragot
Centre de Recherche en Automatique de Nancy (CRAN), CNRS UMR 7039, Université de Lorraine, France
E-mail: {maya.kallas, gilles.mourot, didier.maquin, jose.ragot}@univ-lorraine.fr

Abstract. Technological advances in the process industries during the past decade have resulted in increasingly complicated processes, systems and products. Therefore, recent researches consider the challenges in their design and management for successful operation. While principal component analysis (PCA) technique is widely used for diagnosis, its structure cannot describe nonlinear related variables. Thus, an extension to the case of nonlinear systems is presented in a feature space for process monitoring. Working in a high-dimensional feature space, it is necessary to get back to the original space. Hence, an iterative pre-image technique is derived to provide a solution for fault diagnosis. The relevance of the proposed technique is illustrated on artificial and real dataset.

1. Introduction
The last century has seen an increasingly emergence of manufacturing and process industries, with the rise of energy costs and environmental regulations. In these environments, processes are highly automated. Therefore, the monitoring algorithms are important to detect any fault that might occur.

Principal component analysis (PCA) has been widely used for sensors’ fault detection and isolation [8, 18] and more generally for process monitoring [1, 15, 10]. Taking for example the detection methods using a model known a priori, one of the most successful 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. The PCA considers the most relevant eigenvectors of the data covariance matrix. Therefore, this reconstruction uses the eigenvectors that span the subspace called secondary or residual corresponding to the smallest eigenvalues. As shown in [8], it is possible to extend this reconstruction to n variables, subject to observability, where we can reconstruct simultaneously p variables out of the n – p remaining ones.

Despite the proven performances of this technique, it is necessary to mention that it is applied on a system whose variables are mainly linearly related. In fact, the PCA is a linear projection and identifies only linear structures in a given dataset. Different techniques have been introduced to learn the nonlinearities leading to the nonlinear PCA or more generally using the kernel machines, the so-called kernel PCA (KPCA). In process monitoring or diagnosis, KPCA has been used for fault identification [3, 9, 6] and more generally for process monitoring [7, 14].

KPCA has been widely used for specific applications such as missing data for partially observed systems, like the case of image processing [2, 16]. Therefore, the association of KPCA and discriminant analysis, for fault identification, was proposed in data processing in order to increase the performance of each technique [4, 11].

It is important to mention that KPCA approach considers an overall system, since it attempts to explain the maximum variance of the data taken all together, but without taking into account
any existing local relationships between neighboring data. Therefore, the authors of [5] propose the
Local Kernel Principal Component Analysis (LKPCA) technique for preserving to some extent the local structure of data, then apply it on the benchmark TE [1]. In the same spirit, in [12] Orthogonal Neighborhood Preserving Embedding (ONPE) technique is also supposed to preserve the proximity of points of the original space, during data processing in the feature space.

In order to isolate the fault affecting some given data, diagnosis techniques developed in a linear framework rely heavily on the idea of structuring indicators of failure. The extension of this structure to the nonlinear case remains difficult because of the pre-image problem, but has been the subject of several studies. Fu et al. in [6] applied the KPCA on several subsets of data, each using only a part of the original variables, the difficulty is obviously to build these subsets. In [5], the contribution evaluation, widely used in PCA, is extended to the nonlinear case to provide fault isolation.

Multiscale approaches have also been proposed to try to take into account the existence of nonlinear correlations between variables at different scales. This resulted in MultiScale KPCA (MSKPCA) techniques with applications for fault detection [20, 1].

While few results making use of KPCA on industrial applications have been published, some studies about sets of simulated data or relative to international benchmarks or dealing with measures from pilot laboratory have been presented in the field of air [8], water [13, 14], combustion [17], batch type process [10], and propulsion system [19].

Regarding the three main phases of diagnosis, namely fault presence detection, fault isolation and estimation of its amplitude, the first point has been the subject of various studies. The other two points are less addressed justifying our proposed approach.

The rest of the paper is organized as follows: while section 2 introduces the KPCA technique, section 3 presents a detailed method for fault detection, isolation and estimation of its magnitude. Finally, section 4 illustrates the efficiency of the proposed technique on real and artificial examples, where the sensitivity of the square prediction error (SPE) due to the presence of the fault, and the number of retained principal components (PCs) is detailed and a discussion for choosing the parameters is discussed.

2. Diagnosis using KPCA

Principal component analysis has been widely used for fault detection. However, such technique is limited by its linear nature. Many studies in the past two decades take into consideration the nonlinear models, and introduce the kernel machines due, on the one hand to the development of the statistical learning theory, and on the other hand to the computational efficiency of the corresponding algorithms. This is illustrated with the kernel PCA, the nonlinear version of PCA.

2.1. Kernel PCA and fault detection

Let the kernel \( \kappa: \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R} \) be positive semi-definite, where \( \mathcal{X} \) is an input space, namely \( \sum_{i,j} \alpha_i \alpha_j \kappa(x_i, x_j) \geq 0 \) for all \( \alpha_i, \alpha_j \in \mathbb{R} \) and all \( x_i, x_j \in \mathcal{X} \). Let \( \Phi(\cdot) \) denotes the mapping function from the input space \( \mathcal{X} \) into a high dimensional feature space \( \mathcal{H} \), where the inner product is done. The kernel is given by \( \kappa(x_i, x_j) = \Phi^\top(x_i)\Phi(x_j) \).

Using this concept, a given data matrix, defined by \( n \) variables describing \( m \) measures,

\[
X = \begin{bmatrix}
  x_1^\top \\
  \vdots \\
  x_m^\top
\end{bmatrix}, \quad x_i \in \mathbb{R}^n, \quad X \in \mathbb{R}^{m \times n}
\]

is mapped to \( \mathcal{H} \), more precisely \( x_i \xrightarrow{\Phi} \phi_i = \Phi(x_i), \phi_i \in \mathbb{R}^h \). In a matrix form, we can write

\[
\bar{X} = \begin{bmatrix}
  \phi_1^\top \\
  \vdots \\
  \phi_m^\top
\end{bmatrix}, \quad \bar{X} \in \mathbb{R}^{m \times h},
\]
where $h >> n$ is the dimension of the feature space.

We evaluate the covariance matrix, in the feature space $H$ using:

$$ S = \frac{1}{m - 1} \sum_{i=1}^{m} \phi_i \phi_i^\top $$

$$ S = \frac{1}{m - 1} X^\top X. \quad (1) $$

It is important to mention that $\Phi(\cdot)$ is not explicitly defined, one can evaluate the Gram matrix $X^\top X$ using the kernel function $\kappa(x_i, x_j) = \Phi^\top(x_i)\Phi(x_j)$. For the rest of the paper, the radial type of kernels is studied, in particular the Gaussian kernel defined by:

$$ \kappa(x_i, x_j) = \exp \left( -\frac{(x_i - x_j)^\top(x_i - x_j)}{c} \right), \quad (2) $$

where the influence of $c$ is studied lately. This type of kernels induces infinite-dimensional feature space, namely $h \to \infty$, and it is represented by $m$ dimensions. Let $K$ be the matrix with $\kappa(x_i, x_j)$ elements, we obtain:

$$ K = X X^\top = \begin{bmatrix} \phi_1^\top \phi_1 & \ldots & \phi_1^\top \phi_m \\ \vdots & \ddots & \vdots \\ \phi_m^\top \phi_1 & \ldots & \phi_m^\top \phi_m \end{bmatrix} \quad (4) $$

As in the case of the conventional PCA, KPCA seeks to resolve the eigenvector equation in the feature space, by evaluating the eigenvectors $\nu_i$ and eigenvalues $\lambda_i$ of the covariance matrix $S$. Let $\alpha$ and $\lambda$ be an eigenvector and its corresponding eigenvalue of the matrix $K$:

$$ \nu = \lambda^{-1} X^\top \alpha \quad (5) $$

In the feature space, the eigenvectors $\nu_i$ form a matrix $P = [\nu_1 \cdots \nu_\ell, \nu_{\ell+1} \cdots \nu_m]$, where $\ell$ is the number of principal components (PC). The first part $[\nu_1 \cdots \nu_\ell]$ will be denoted as $P_f \in \mathbb{R}^{m \times \ell}$ and the second $[\nu_{\ell+1} \cdots \nu_m]$ as $P_f \in \mathbb{R}^{m \times (m - \ell)}$ representing the eigenvectors defining respectively the principal and residual spaces, where the two are complementary subspaces. We can write

$$ P_f = \begin{bmatrix} \frac{1}{\lambda_1} X^\top \alpha_1 & \cdots & \frac{1}{\lambda_\ell} X^\top \alpha_\ell \end{bmatrix}, \quad (6) $$

where $\alpha_1 > \cdots > \alpha_\ell$. The choice of the number $\ell$ of PCs has been the subject of many studies where [21] describes some of them.

2.2. Analysis of a new observation

During the learning phase of the KPCA model, we seek the $\ell$ value, and the different hyperparameters of the model. However, it is important to validate the model through new observations. Let us denoted by $x$ such new observation, its image in the feature space is

$$ \phi = \Phi(x), \quad \phi \in \mathbb{R}^m, $$

for which, the projection on the principal and residual spaces are given by:

$$ \begin{cases} 
  t = P_f^\top \phi, & \in \mathbb{R}^\ell \\
  \tilde{t} = P_f^\top \phi, & \in \mathbb{R}^{m - \ell} 
\end{cases} \quad (7) $$
We evaluate the statistics $SPE$, where the dot product in $\mathcal{H}$ is replaced by the kernel function as given

\[
\begin{align*}
SPE &= \mathbf{t}^\top \mathbf{t} \\
&= \phi^\top P_f^\top \phi \\
&= \phi^\top (I - P_f P_f^\top) \phi \\
&= \phi^\top \theta - \phi^\top P_f P_f^\top \phi \\
&= \kappa(x, x) - \kappa^\top(x) C \kappa(x)
\end{align*}
\]

(8)

where

\[
\begin{align*}
\kappa(x) &= [\kappa(x_1, x), \ldots, \kappa(x_m, x)]^\top \\
C &= P \Lambda^{-1} P^\top
\end{align*}
\]

(9)

It is preferred to have centered data in the feature space. Thus, the statistics $SPE$ will be transformed as

\[
\begin{align*}
\overline{SPE} &= \kappa_c(x, x) - \kappa_c^\top(x) C \kappa_c(x) \\
\kappa_c(x) &= (I - E)(\kappa(x) - K u_m) \\
\kappa_c(x, x) &= 1 - 2u_m^\top \kappa(x) + u_m^\top K u_m \\
u_m &= \frac{1}{m} [1 \ldots 1]^\top \\
E &= m u_m u_m^\top
\end{align*}
\]

(10)

The analysis of $\overline{SPE}$ indicates the presence of a fault in the measure $x$, based on a detection threshold defined during the learning stage.

### 3. Fault isolation and estimation with KPCA

Any new observation $x$ to be analyzed, subject to a possible additive fault, can be expressed in terms of the true value $x^*$ as:

\[
x = x^* + \xi f, \quad x \in \mathbb{R}^n, \xi \in \mathbb{R}^{n \times d},
\]

(11)

where $\xi$ is the fault’s direction and $f$ its amplitude. The true value to be estimated is given by:

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

(12)

#### 3.1. Analysis of the validity of an observation

By substituting any observation $x$ with its real value defined by (12), the $\overline{SPE}$ in (10) is

\[
\overline{SPE} = \kappa_c(x - \xi f, x - \xi f) - \kappa_c^\top(x - \xi f) C \kappa_c(x - \xi f)
\]

(13)

If $\overline{SPE}$ is greater than a predefined threshold, a fault is detected. However in order to isolate and estimate its value, we seek a direction $\hat{\xi}$ and a value of $f$ that make the $\overline{SPE}$ less than the threshold. In (11) and (13), only the value $x$ of the measure is known. The fault direction $\xi$ and the amplitude of the fault $f$ must be determined. Therefore, we seek $\xi$ and $f$ that minimize the index $\overline{SPE}$. To this end, for each $\xi$ of the possible directions, the fixed point iterative technique is used to resolve the optimization problem with

\[
\hat{f} = \frac{\xi^\top B(x - \xi \hat{f}) [u_m + (I - E) C \kappa_c(x - \xi \hat{f})]}{\kappa^\top(x - \xi \hat{f}) [u_m + (I - E) C \kappa_c(x - \xi \hat{f})]}
\]

(14)

where

\[
\begin{align*}
B(x - \xi \hat{f}) &= \begin{bmatrix} \kappa(x - \xi \hat{f}, x_1)(x_1 - x)^\top \\ \vdots \\ \kappa(x - \xi \hat{f}, x_m)(x_m - x)^\top \end{bmatrix} \\
\kappa(x - \xi \hat{f}) &= [\kappa(x_1, x - \xi \hat{f}), \ldots, \kappa(x_m, x - \xi \hat{f})]^\top \\
\kappa_c(x - \xi \hat{f}) &= (I - E)(\kappa(x - \xi \hat{f}) - K u_m) \\
\kappa(x_i, x_j) &= \exp\left(\frac{-||x_i - x_j||^2}{c} \right)
\end{align*}
\]

(15)
When the magnitude \( f \) is estimated, the measure \( x \) can be corrected with (12) and \( \text{SPE} \) is recalculated with the corrected measure using (13). A meaningful assessment of the fault direction and amplitude results in a significant decrease of \( \text{SPE} \).

This is done for any given direction \( \xi \). However, it is worth nothing that the direction in not defined, since the diagnosis aims to localize the fault among the component of \( x \). KPCA technique, proposed in [6], operates on different subsets of variables, here, we propose an alternative in order to identify the fault’s direction. In case of a single fault, affecting the \( j^{th} \) component of the observation \( x \), the direction is represented by \( \xi_j = [0 \ldots 1 \ldots 0] \) where 1 is at the \( j^{th} \) position. Therefore, for all \( j = 1 \ldots n \), the optimal amplitude \( \hat{f}_j \) is the one that minimizes \( \text{SPE} \), and makes it below a given threshold.

**Remark 1. Interpretation of the equation (14)**

Let \( w \) of elements \( w_i \) be a weighting vector:

\[
    w = u_m + (I - E)C\kappa_c(x - \hat{x})
\]

Thus, we have

\[
    \kappa^\top(x - \hat{x})w = \sum_{i=1}^{m} \kappa(x_i, x - \hat{x})w_i
\]

By considering a fault direction \( \xi_j \), the term \( \xi_j^\top B^\top(x - \hat{x}) \) of (14) is used to select the \( j^{th} \) element of \( x_i - x \). Therefore, we can write:

\[
    \xi_j^\top B^\top(x - \hat{x})w = \sum_{i=1}^{m} \kappa(x_i, x - \hat{x})(x_i - x)^\top \xi_j w_i
\]

Therefore, the equation (14) is written as

\[
    \hat{f} = \frac{\sum_{i=1}^{m} \kappa(x_i, x - \hat{x})w_i (x_i - x)^\top \xi_j}{\sum_{i=1}^{m} \kappa(x_i, x - \hat{x})w_i}
\]

The term \( \kappa(x_i, x - \hat{x})w_i \) provides a weight for the scalar \( (x_i - x)^\top \xi_j \), leading to the \( j^{th} \) element of the distance \( x_i - x \) between the new observation and the \( i^{th} \) one of the dataset.

### 3.2. Algorithm

As aforementioned, two parameters indicate the right fault detection. The first one is the amplitude estimate of the fault. As the direction of the fault is not known, it is necessary to test all possible directions \( \xi_j = (0, \ldots, 1, 0, \ldots, 0) \), where solving (14) provides an estimate of the fault \( \hat{f}_j \). A first indicator \( I_{f,j} \) is defined by the value of the estimated of the fault over the \( j^{th} \) variable with respect to a predefined threshold \( \delta_f \).

The second parameter takes into consideration the \( \text{SPE} \). It must be either zero for fault-free measures or corrected ones, or more precisely for robustness issues with respect to uncertainties, must be less than a predefined threshold \( \delta_s \). In practice, this value is defined by the user and corresponds to the minimal magnitude needed to detect a fault. Effectively, the combination of the two tests is given in the following procedure steps:

(i) Define the detection thresholds \( \delta_f \) and \( \delta_s \).
(ii) Define the fault directions \( \xi_j, j = 1, \ldots, n \) when all the components are null except the \( j^{th} \).
(iii) Estimate the fault magnitude \( \hat{f}_j \) at each instant for a specific direction \( \xi_j, j = 1, \ldots, n \), then evaluate \( \text{SPE}_j \) after correcting the faulty measure.
(iv) Define the amplitude validity parameter by

$$I_{f,j} = \frac{1}{2} \left(1 + \text{sgn}(\hat{f}_j > \delta_f)\right)$$

(v) Define the $SPE_j$ validity parameter by:

$$I_{s,j} = \frac{1}{2} (1 - \text{sgn}(SPE_j < \delta_s))$$

(vi) Evaluate the fault presence indicator

$$I_j = I_{f,j} \& I_{s,j}$$

The adjustment of thresholds $\delta_s$ and $\delta_f$ affects the performance of fault detection and isolation. It can be chosen carefully at the learning stage. Based on different tests, we decided to use an automatic adjustment based on the comparison of different estimated amplitudes $\hat{f}_j$ and a comparison of different $SPE_j$ criteria. Thus, these thresholds are defined as follows:

$$\begin{cases} 
\delta_s = \frac{1}{n} \sum_{j=1}^{n} SPE_j \\
\delta_f = \frac{1}{n} \sum_{j=1}^{n} \hat{f}_j
\end{cases}$$

(20)

4. Experiments

In this section, we study the relevance of the proposed technique by applying it on simulated and real data.

4.1. Example on simulated dataset

The data in this example were generated from a model with $n = 4$ variables $x_i$, $i = 1, \ldots, 4$ linked by two nonlinear redundancy equations, namely:

$$\begin{cases} 
x_3 - x_1 x_2 = 0 \\
x_4 - x_1^2 = 0
\end{cases}$$

(21)

As we can see from table 1, variables $x_1$ and $x_4$ can be isolated due the difference of the signature, where $r_j$ denotes the redundancy equations with respect to faults $f_i$ affecting variable $x_i$ represented by $\delta_i$.

To construct the KPCA model, 50 measures were used to form the training data, while the testing set was limited to 25 samples, a fault of constant amplitude affects $x_2$. Figure 1 illustrates the data, where the shaded area represents the fault’s presence. In order to assess the relevance of estimating the fault’s magnitude, $SPE$ defined with (10) is evaluated.

Let us consider the case when a fault of magnitude $-1.20$ occurs on a variable, as for example the situation presented within figure 1.

As aforementioned, four fault estimates were performed according to directions $\xi_i$, where $i = 1, \ldots, 4$. Columns 3 and 4 of table 2 indicate the estimated fault and $SPE$ when the

| Table 1. Fault signatures’ table |
|----------------------------------|
| $\delta_1$ | $\delta_2$ | $\delta_3$ | $\delta_4$ |
| $r_1$   | $\times$ | $\times$ | $\times$ |
| $r_2$   | $\times$ | $0$    | $0$    | $\times$ |
measures are corrected. Data collected at $t = 5$ are therefore not affected by any fault, since $SPE$ and $\hat{f}$ are null. At $t = 15$, we note that in directions $\xi_i$ such that $i = 1, 3, 4$ a fault was found, whose value cannot be accepted as significant, because it does not make $SPE$ null. The $\xi_2$ direction is the one whose fault is on variable $x_2$, which after correction makes null the $SPE$. Hence, the amplitude of the fault has been correctly estimated.

The relevance of the proposed method is demonstrated by the preceding example. However, it requires the setting of some parameters. Therefore, it is important to study the influence, on the one hand, of the number of eigenvalues $\ell$ retained in the feature space (1), and on the other hand, the hyper-parameter $c$ in (3).

Figures 3 and 4 visualize the influence of these two parameters for six values of $c \in \{0.025, 0.1, 0.25, 0.5, 1, 1.5\}$ and two values for PCs in $\{3, 12\}$. Each figure shows the evolution of $SPE$ as a function of the fault amplitude affecting only the variable $x_2$. On each sub-figure, the vertical dashed blue line recalls the fault amplitude on $x_2$ which is $-1.2$. The vertical red line positions the estimated fault value for the considered direction $\xi_2$. These two lines are almost overlapping, reflecting the correct estimate of fault magnitude.

The first sub-figures in 3 and in 4 report an extreme case, as established with a very small hyper-parameter ($c = 0.025$), where the first one uses a reduced number of PCs in the feature space. Despite the particular shape of $SPE$ and its values, the fault is therefore isolated on $x_2$.
and its amplitude is correctly estimated. However, the second one with a larger number of PCs presents a criterion near 0 when the fault amplitude tends to 0.

4.2. **SPE sensitivity due to fault and choice of PCs**

*SPE* is sensitive when a fault is present and it is affected by the number of PCs retained in feature space.

4.2.1. **Sensitivity due to fault presence** Let us start by an example in order to study the sensitivity of the *SPE* in presence of a fault. Figure 2 illustrates the evolution of the *SPE* with respect to *f* for different values of *c*, where *c* ∈ [0.05, 1] with a step of 0.15. The first 15 PCs are considered, using the same configuration of the simulated data, from which we can conclude that

- If *f* >> *c*, the *SPE* reaches a constant.
- If *c* → 0, the *SPE* is not 0 when there is no fault.

It is important to mention that for the second case, where *c* is very small, we must choose more PCs, and the value of the *SPE* decreases.

The *SPE* is thus affected by the fault *f*, the choice of *c* and the considered *x*<sup>*</sup>. We will study analytically this sensitivity. Let *x* be as defined in (11), namely *x* = *x*<sup>*</sup> + *ξ*<sub>j</sub>*f*. In order to evaluate *κ(x, x<sub>i</sub>)*, the expression of the exponential is developed using

\[
(x - x_l)^\top (x - x_l) = (x^* - x_l + \xi_j f)^\top (x^* - x_l + \xi_j f)
\]

\[
= (x^* - x_l)^\top (x^* - x_l) + 2f(x^* - x_l)^\top \xi_j + f^2
\]

\[
= \sum_{t=1}^{m} (x^* - x_{lt})^2 + (x^* - x_{lj})^2 + 2f(x^* - x_{lj}) + f^2
\]  \hspace{1cm} (22)

In a general manner, the kernel *κ(x, x<sub>i</sub>)* can be defined as

\[
κ(x, x_l) = \prod_{t=1}^{m} \exp \left\{ -\frac{(x^*_t - x_{lt})^2}{c} \right\} \exp \left\{ -\frac{f^2}{c} \left( 1 + \frac{(x^*_j - x_{lj})}{f} \right)^2 \right\}
\]  \hspace{1cm} (23)

Let us study the influence of *f* and *x*<sup>*</sup><sub>j</sub> - *x*<sub>lj</sub>.
• If $f$ and $x_j - x_{ij}$ are of the same sign:

$$1 < 1 + \frac{x_j^* - x_{ij}}{f} \Rightarrow \exp \left\{ - \frac{f^2}{c} \left( 1 + \frac{(x_j^* - x_{ij})}{f} \right)^2 \right\} < \exp \left\{ - \frac{f^2}{c} \right\}$$

(24)

Therefore

$$\exp \left\{ - \frac{f^2}{c} \right\} \rightarrow 0$$

(25)

By taking the gaussian distribution, if $f > 3\sqrt{\frac{2}{\pi}}$, the $\exp \left\{ - \frac{f^2}{c} \right\} < 0.01$. In this case, if $\kappa(x) \rightarrow 0$ then $\text{SPE}$ in (10) is written as

$$\text{SPE} = 1 + u_m^\top K u_m - (K u_m)^\top (I - E)(I - E)(K u_m).$$

This expression does not depend on $\kappa(x)$, thus it is a constant which is its maximum.

• If $f$ and $x_j - x_{ij}$ are of opposite sign:

- $\left| \frac{x_j^* - x_{ij}}{f} \right| \ll 1$, then if $|f| > \max_x(|x_j^* - x_{ij}|)$, we are in the same situation described in (25) and $\exp \left\{ - \frac{f^2}{c} \right\} \rightarrow 0$. In this case, we might say that $c$ should have a small value. However from figure 2, $\text{SPE}$ is not null in absence of fault. Therefore, in order to detect the presence of such fault, $f$ should have an impact greater than the threshold obtained for $f = 0$.

- $\left| \frac{x_j^* - x_{ij}}{f} \right| \gg 1$, then if $|f| < \min_x(|x_j^* - x_{ij}|)$, then

$$\exp \left\{ - \frac{f^2}{c} \left( 1 + \frac{(x_j^* - x_{ij})}{f} \right)^2 \right\} < \exp \left\{ - \frac{1}{c} (x_j^* - x_{ij})^2 \right\} \quad \forall f$$

In this case, $\kappa(x) \rightarrow \kappa(x^*)$, where $x^*$ is the fault-free value. $\text{SPE}$ is given by:

$$\text{SPE} = 1 - 2u_m^\top \kappa(x^*) + u_m^\top K u_m - (\kappa(x^*) - K u_m)^\top (I - E)(I - E)(\kappa(x^*) - K u_m).$$

The obtained value is $\text{SPE}(x^*)$, therefore, the fault $f$ will be considered as noise and cannot be detected for any value of $c$.

### 4.2.2. Sensitivity due to principal functional space dimension tuning

If we choose $\ell$ PCs, the $\text{SPE}$ criterion is given by:

$$\left\{ \begin{array}{l}
\text{SPE}_\ell = \kappa_c(x - \xi f, x - \xi f) - \kappa_c^\top (x - \xi f) C_\ell \kappa_c(x - \xi f) \\
C_\ell = P_\ell \Lambda_\ell^{-1} P_\ell^\top
\end{array} \right.$$  

(26)

If we use $\ell + 1$ PCs, we obtain:

$$\left\{ \begin{array}{l}
C_{\ell+1} = P_{\ell+1} \Lambda_{\ell+1}^{-1} P_{\ell+1}^\top \\
P_{\ell+1} = [P_\ell \mid p_{\ell+1}] \\
\Lambda_{\ell+1}^{-1} = \begin{bmatrix} \Lambda_\ell^{-1} & 0 \\ 0 & \lambda_{\ell+1}^{-1} \end{bmatrix}
\end{array} \right.$$  

(27)

Therefore, by combining (26) and (27), the expression of $\text{SPE}$ is:

$$\left\{ \begin{array}{l}
C_{\ell+1} = C_\ell + p_{\ell+1} \lambda_{\ell+1}^{-1} P_{\ell+1}^\top \\
\text{SPE}_{\ell+1} = \kappa_c - \kappa_c^\top C_{\ell+1} \kappa_c \\
\text{SPE}_{\ell+1} = \kappa_c - \kappa_c^\top (C_{\ell+1} + p_{\ell+1} \lambda_{\ell+1}^{-1} P_{\ell+1}^\top) \kappa_c \\
\text{SPE}_{\ell+1} = \frac{\text{SPE}_\ell - \kappa_c^\top P_{\ell+1} \lambda_{\ell+1}^{-1} P_{\ell+1}^\top \kappa_c}{\text{SPE}_\ell - \kappa_c^\top P_{\ell+1} \lambda_{\ell+1}^{-1} P_{\ell+1}^\top \kappa_c}
\end{array} \right.$$  

(28)
Thus, the variation of the $\overline{SPE}$ due to an increase of PCs is given by:

$$\Delta \overline{SPE}_\ell+1 = -\frac{(\kappa_p^+ p_{\ell+1})^2}{\lambda_{\ell+1}}$$ (29)

This expression is used to adjust the contribution of different eigenvectors for $\overline{SPE}$, making it possible to obtain the number of PCs to retain. Figure 5 was established for a faulty observation, while figure 6 was established for fault-free observation. Both of them show the evolution of $\overline{SPE}$ and its variations as a function of the number of PCs. In both situations, 15 PCs are enough to obtain an acceptable $\overline{SPE}$.

4.3. Chemical reactor
For the CSTR considered reactor, the two state variables have a concentration $C_a$ of a specie $A$ and a temperature $T$ in the reactor. Data were generated by varying the concentration $C_{af}$ of $A$ and the volumetric flowrate $q$. The properties of the reactor are summarized in table 3.

For a steady state, and for a supposed homogeneous reactor system, conservations of the mass of $A$ specie and the energy result in the following two relationships:

$$\begin{align*}
0 &= q (C_{af} - C_a) - k_0 V \exp(-E/T) C_a \\
0 &= q \rho C_p (T_f - T) + V H k_0 \exp(-E/T) C_a + U_a (T_c - T)
\end{align*}$$ (30)

Two fault sensors are considered, the first of amplitude $-3$ affecting the measurement of variable $T_c$ between 10 and 20, the second of amplitude 2 affecting the measurement of variable $T$ between 30 and 40. Figure 7 includes 50 collected measures for different steady states of the reactor. After identifying the KPCA model for these 50 measures, the procedure for fault detection and estimation of their amplitudes was made on 50 other measures. Recall that for each measure of 5 variables at a given time, the procedure examines five possible situations of fault presence, the latter may affect any of the 5 variables.

Figure 8 shows the evolution over 50 observations of the five $\overline{SPE}_i$ criteria assessed after measures correction of their respective estimated faults. For the case of fault absence or corrected
Table 3. CSTR. Variables, parameters and their nominal values

| Variable | Value | Description |
|----------|-------|-------------|
| $T_c$    | 270   | Temperature outside CSTR |
| $q$      | 100   | Volumetric Flowrate ($m^3/sec$) |
| $V$      | 140   | Volume of CSTR ($m^3$) |
| $\rho$   | 1000  | Density of A-B Mixture ($kg/m^3$) |
| $C_p$    | 0.239 | Heat capacity of A-B Mixture ($J/kg - K$) |
| $H$      | 5e4   | Heat of reaction for A->B ($J/mol$) |
| $E$      | 8750  | Activation energy |
| $k_0$    | 7.2e10| Pre-exponential factor (1/sec) |
| $U_a$    | 5e4   | Overall Heat Transfer Coefficient |
| $C_{af}$ | 1     | Feed Concentration ($mol/m^3$) |
| $T_f$    | 370   | Feed Temperature (K) |
| $C_a$    | 0.9   | Concentration of A in CSTR ($mol/m^3$) |
| $T$      | 305   | Temperature in CSTR (K) |

Figure 7. CSTR. Measures of five variables with faults on $T_c$ between instances 10 and 20 and on $T$ between 30 and 40

Figure 8. CSTR. Variation of $SPE$ for the 5 fault directions and maximal fault amplitude

5. Conclusion

In this paper, we propose to diagnose a nonlinear system using the KPCA technique. The latter can be described as the linear PCA but on transformed data using a nonlinear mapping function.
into a feature space.

The term diagnosis is defined as the detection of a fault, its isolation and the estimation of its amplitude. A technique used for diagnosis is the analysis of the SPE. It should have a reduced value when evaluated on raw data or corrected ones. In order to estimate the magnitude of these errors, knowing that several faults can appear on multiple variables, the chosen method considers different fault directions. The performance of this approach has been demonstrated on real and artificial data, in terms of faults detection, isolation and estimation.

Many improvements can be addressed in particular to study the influence of some chosen learning database. The choice of the other parameters, the hyper-parameter and the number of PCs in the feature space is relative to each dataset.

References

[1] 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.

[2] M. Bera, M. Desvignes, G. Bailly, Y. Payan, and B. Romaniuk. Missing data estimation using polynomial kernels. In S. Singh, M. Singh, C. Apté, 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.

[3] M. Bin Shams. Fault identification using kernel principal component analysis. In 18th IFAC World Congress, pages 4320–4325, Milano, Italy, 2011.

[4] 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.

[5] X. Deng, X. Tian, and S. Chen. Modified kernel principal component analysis based on local structure and its application to nonlinear process fault diagnosis. Chemometrics and Intelligent Laboratory Systems, 127:195–209, August 2013.

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

[8] 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.

[9] M. Hehe, H. Yi, and S. Hongbo. Statistics kernel principal component analysis for nonlinear process fault detection. In 9th World Congress on Intelligent Control and Automation, 2011.

[10] J.M. Lee, C.K. Yoo, and I.B. Lee. Fault detection of batch processes using multiway kernel principal component analysis. Computers & Chemical Engineering, 28(9):1837–1847, 2004.

[11] Z. Li, G. Tan, and Y. Li. Fault diagnosis based on improved kernel fisher discriminant analysis. Journal of Software, 7(12):2657–2662, 2012.

[12] A. Miao, Z.H. Song, Q. Wen, and Z. Ge. Process monitoring based on generalized orthogonal neighborhood preserving embedding. In 8th IFAC Symposium on Advanced Control of Chemical Processes, pages 148–153, Singapore, 2012.

[13] J. Ni, C. Zhang, L. Ren, and S.X. Yang. Abrupt event monitoring for water environment system based on KPCA and SVM. IEEE Transactions on Instrumentation and Measurement, 61(4):980–989, 2012.

[14] A. Nowicki, M. Grochowski, and K. Duzinkiewicz. Data-driven models for fault detection using kernel PCA: A water distribution system case study. Applied Mathematics and Computer Science, 22(4):939–949, 2012.

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

[16] M. Scholz, E. Kaplan, C.L. Guy, J. Kopka, and J. Selbig. Non-linear PCA: a missing data approach. Bioinformatics, 21(20):3887–3895, 2005.

[17] D. Sun, G. Lu, H. Zhou, and Y. Yan. Condition monitoring of combustion processes through flame imaging and kernel principal component analysis. Combustion Science and Technology, 185(9):1400–1413, 2013.

[18] M. Tanaka 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.

[19] T. Xu. Sensor Fault Detection and Identification Method with KPCA in the Process of Aero-Engine Ground Testing. Applied Mechanics and Materials, 303:297–301, February 2013.

[20] T. Xuemin and D. Xiaogang. A fault detection method using multi-scale kernel principal component analysis. In 27th Chinese Control Conference, pages 25–29, Kunming, Yunnan, China, 2008.

[21] Y. Zhang, S. Li, and Y. Teng. Dynamic processes monitoring using recursive kernel principal component analysis. Chemical Engineering Science, 72(0):78 – 86, 2012.