Reliable Fault Detection and Diagnosis of Large-Scale Nonlinear Uncertain Systems Using Interval Reduced Kernel PLS

RADHIA FEZAI1, KAMALELDIN ABOYADH2, MAJDI MANSOURI1, (Member, IEEE), ABDELMALEK KOUADRI3, MOHAMED-FAOUZI HARKAI4, HAZEM NOUNOU1, (Senior Member, IEEE), MOHAMED NOUNOU5, (Senior Member, IEEE), AND HASSANI MESSAOUD6

1Electrical and Computer Engineering Program, Texas A&M University at Qatar, Doha 23874, Qatar
2Department of Mathematical Sciences, Prince Sultan University, Riyadh 66833, Saudi Arabia
3Signals and Systems Laboratory, Institute of Electrical and Electronics Engineering, University M Hamed Bougara of Boumerdes, Boumerdes 35000, Algeria
4LASMA, Badji Mokhtar - Annaba University, Annaba 23000, Algeria
5Chemical Engineering Program, Texas A&M University at Qatar, Doha 23874, Qatar
6Research Laboratory of Automation, Signal Processing and Image (LARATSI), National School of Engineering Monastir, Rue Ibn ELJazzar, Monastir 5019, Tunisia

Corresponding author: Majdi Mansouri (majdi.mansouri@qatar.tamu.edu)

This work was made possible by National Priorities Research Program grant NPRP9-330-2-140 from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors.

ABSTRACT Kernel partial least squares (KPLS) models are widely used as nonlinear data-driven methods for faults detection (FD) in industrial processes. However, KPLS models lead to irrelevant performance over long operation periods due to process parameters changes, errors and uncertainties associated with measurements. Therefore, in this paper, two different interval reduced KPLS (IRKPLS) models are developed for monitoring large scale nonlinear uncertain systems. The proposed IRKPLS models present an interval versions of the classical KPLS model. The two proposed IRKPLS models are based on the Euclidean distance between interval-valued observations as a dissimilarity metric to keep only the more relevant and informative samples. The first proposed IRKPLS technique uses the centers and ranges of intervals to estimate the interval model, while the second one is based on the upper and lower bounds of intervals for model identification. These obtained models are used to evaluate the monitored interval residuals. The aforementioned interval residuals are fed to the generalized likelihood ratio test (GLRT) chart to detect the faults. In addition to considering the uncertainties in the input-output systems, the new IRKPLS-based GLRT techniques aim to decrease the execution time when ensuring the fault detection performance. The developed IRKPLS-based GLRT approaches are evaluated across various faults of the well-known Tennessee Eastman (TE) process. The performance of the proposed IRKPLS-based GLRT methods is evaluated in terms of missed detection rate, false alarms rate, and execution time. The obtained results demonstrate the efficiency of the proposed approaches, compared with the classical interval KPLS.

INDEX TERMS Kernel PLS (KPLS), interval KPLS (IKPLS), interval reduced KPLS (IRKPLS), fault detection (FD), uncertain systems.

I. INTRODUCTION Fault detection (FD) becomes increasingly important to improve the quality product, ensure process safety and decrease energy and material consumption [1]. As a large amount of process data are gathered in historical databases in modern industrial processes, data-driven FD techniques have been discussed widely by researchers [2]–[9]. Partial least squares (PLS) regression is the most used data-driven technique in modeling and FD of industrial processes, and it has proved excellent performance [10], [11]. The main idea of PLS is to build the correlation between the process and quality variables [12]. By establishing a model through normal data, PLS can be applied for predicting and FD purposes.
However, the PLS technique assumes that the process data are linear. This assumption makes it inappropriate for nonlinear industrial processes. In order to address this issue, a nonlinear version of PLS model, so-called kernel PLS (KPLS), has been developed in literature [13]–[15]. The KPLS technique aims to map the nonlinear input data into a high-dimensional space in which a linear PLS model is applied. According to the Covers theorem [14], the nonlinear relationship among variables in the original space is most probably linear after a high-dimensional nonlinear mapping. The KPLS can efficiently determine the regression coefficients in feature space using nonlinear kernel functions and improve the prediction performance. It has been successfully implemented for modeling and FD of many industrial processes [2], [16]–[18].

Generally, the quality of the available measurements is essential in the reliability of the identified model as well as in the FD performance. Using the KPLS technique, the measured process data are described by a single value representation. The single-valued data representation is a result of simplification during the data mining procedure. The necessity for interval-valued data can arise in connection with the imprecision of measurement devices, process uncertainties, and with the data fluctuations in the case of collected samples during a specific interval of time.

In actual practice, taking into account the minimum and maximum of the collected sample values gives a complete overview of the measured phenomenon than taking into account the average values. Thus, for more accuracy in the representation of real data, the uncertainty of the data may be represented by considering data at intervals representation [19]–[21]. In literature, several linear techniques for interval-valued data have been developed [21]–[27]. However, up to now there is no studies on the problem of nonlinear uncertain systems using interval kernel PLS.

The interval KPLS (IKPLS) or KPLS-based interval-valued data method is an extension of the single-valued KPLS model in presence of imprecision, variation, and data confidence intervals. For example, the IKPLSUL model aims to build two KPLS models on the lower and upper (UL) bounds of interval values [21]. Another interval technique, so-called Bivariate Center and Range (CR) KPLS (IKPLSCR) consists to create a KPLS model on the new numerical input and output matrices constructed by the concatenation of center and range matrices [21]. However, the major limitation of KPLS-based fault detection techniques is their computational complexity. Using KPLS model for FD leads to a high computational cost when the training data set is large since the recorded and stored data are applied for both modeling and monitoring. This is mainly explained by the fact that the kernel methods depend on the collected observations of the whole process variables. To reduce the computational complexity of the IKPLS models, two interval reduced KPLS (IRKPLS) models are proposed in the current work. The objective is to enhance data-driven monitoring methods, especially in experimental industrial applications with imperfect measurements which significantly deteriorate the detection performance. The proposed IRKPLS schemes do not depend on data outliers and model incompleteness and have the ability to improve both fault detection robustness and sensitivity while maintaining a satisfactory and stable performance over long periods of process operation. Both IRKPLS methods (IRKPLSUL and IRKPLSCR) are based on the Euclidean distance between interval-valued samples as a dissimilarity metric to extract only the non-redundant ones and keep only one informative sample for every small and similar distance.

In addition to their abilities to efficiently monitor uncertain processes, the IRKPLS models are able to reduce the execution time and memory space.

The two proposed IRKPLSUL and IRKPLSCR models are used to generate the monitored interval residuals. The aforementioned interval residuals are fed to generalized likelihood ratio test (GLRT) chart as input for detecting faults. GLRT is considered the most used method for detection of industrial faults [28], [29].

Thus, in the proposed IRKPLS-based GLRT approaches, the IRKPLS methods are used for identifying the interval-valued models and the GLRT chart is applied for faults detection.

The detection efficiency of the proposed IRKPLS-based GLRT approaches is evaluated using a Tennessee Eastman (TE) process. The false alarm rate (FAR), missed detection rate (MDR), and execution time (ET) are used as metrics to evaluate the detection and monitoring abilities.

The remainder of this paper is structured as follows. In the Preliminaries section, PLS and kernel PLS are briefly introduced. The developed IRKPLS models are presented in Section 3. The fault detection using GLRT is detailed in Section 4. In section 5, the application of the developed IRKPLS-based GLRT approaches to a Tennessee Eastman (TE) process is presented. The conclusions are presented in Section 6.

II. PRELIMINARIES

A. PARTIAL LEAST SQUARES

Partial least squares (PLS) is a multivariate statistical analysis method. PLS models the relation between input variables (regressors) and output variables (responses). The PLS decomposition of input $X \in \mathbb{R}^{N \times m}$ and response $Y \in \mathbb{R}^{N \times p}$ have the following form [30], [31]

$$
\begin{align*}
X &= \sum_{i=1}^{\ell} t_ip_i^T + E = TP^T + E \\
Y &= \sum_{j=1}^{\ell} u_jq_j^T + F = UQ^T + F,
\end{align*}
$$

where $T = [t_1, t_2, \cdots, t_\ell] \in \mathbb{R}^{N \times \ell}$ and $U = [u_1, u_2, \cdots, u_\ell] \in \mathbb{R}^{N \times \ell}$ represent the input and output score matrices, respectively. $P = [p_1, p_2, \cdots, p_\ell] \in \mathbb{R}^{m \times \ell}$ and $Q = [q_1, q_2, \cdots, q_\ell] \in \mathbb{R}^{N \times \ell}$ indicate loading matrices, respectively. $E \in \mathbb{R}^{N \times m}$ and $F \in \mathbb{R}^{N \times p}$ present the residual matrices.
The score vector \( t_i \), which denotes the column of the matrix \( T \), can be determined as,

\[
t_i = x_i c_i
\]

(2)

where \( c_i \) is the weight vector which is obtained by the optimization of the following problem \([32]\),

\[
\max c_i^T x_i^T y_i q_i \quad \text{subject to} \quad ||c_i||=1 \quad ||q_i||=1
\]

Likewise, with \( t \), the score vector \( u_t \) can be estimated as,

\[
u_t = y_i q_i
\]

(4)

The PLS regression model can be represented by using a regression coefficient matrix \( B \) and a residual matrix \( R \) as,

\[
\begin{align*}
Y &= XB + R \\
B &= X^T U (T^T X X^T U)^{-1} T^T
\end{align*}
\]

(5)

**B. KERNEL PARTIAL LEAST SQUARES**

PLS is a linear regression approach which assumes that the process variables are linear. In many applications, this assumption is not realistic. To solve this issue, a nonlinear regression technique so-called kernel PLS (KPLS) is developed \([33]-[35]\). KPLS is the most widely used nonlinear input-output statistical methods. KPLS consists to map the input data into high-dimensional space. Then, it aims to apply linear PLS in the feature space. Given an input data matrix \( X = [x_1 \ldots x_N]^T \in \mathbb{R}^{N \times m} \) with \( N \) observations and \( m \) process variables \( x_i \in \mathbb{R}^m \), and output matrix \( Y = [y_1 \ldots y_N]^T \in \mathbb{R}^{N \times q} \) regrouping \( N \) observations with \( q \) quality variables \( y_i \in \mathbb{R}^q \). We assume that the original input data \( x_i \in \mathbb{R}^m \) are mapped into the high dimension features space by using a nonlinear projection function \( \phi \), so that \( x_i \in \mathbb{R}^m \rightarrow \phi(x_i) \in \mathbb{R}^p \). Note that \( p \) is the dimensionality of the feature space \( \mathcal{F} \) which can be arbitrarily large or even infinite. The mapped input data matrix \( \Phi \) is given by

\[
\Phi = [\phi(x_1), \ldots, \phi(x_N)]^T
\]

(6)

where \( \sum_{i=1}^{N} \phi(x_i) = 0 \).

The KPLS model of \((\Phi, Y)\) is defined as \([36]\):

\[
\begin{align*}
\Phi &= TP^T + \Phi_r \\
Y &= UQ^T + Y_r
\end{align*}
\]

(7)

where \( T \in \mathbb{R}^{N \times A} \) and \( U \in \mathbb{R}^{N \times A} \) represent the input and output score matrices, respectively. The loading matrices of \( \Phi \) and \( Y \) are denoted by \( P \in \mathbb{R}^{m \times A} \) and \( Q \in \mathbb{R}^{N \times A} \), respectively. \( \Phi_r \) and \( Y_r \) present the residuals. Denote \( A \) as the number of latent variables in high-dimensional space selected by the cumulative percent variance (CPV) criterion \([36]\). In KPLS, using the kernel function \( k(x_i, x_j) = \phi(x_i)^T \phi(x_j) \), we can avoid to carry out explicit nonlinear mapping and determine dot products in space \( \mathcal{F} \) using a kernel functions. Radial basis kernel is the most used kernel functions and it is defined as: \( k(x_i, x_j) = \exp(-\frac{(x_i-x_j)^T(x_i-x_j)}{2\sigma^2}) \). Denote by \( K = \Phi^T \), an \( N \times N \) kernel matrix. The KPLS algorithm is outlined in Algorithm 1

**Algorithm 1 KPLS Algorithm**

1. Set \( i = 1 \), \( K_1 = K \), and \( Y_1 = Y \).
2. Initialize the vector \( u_i(N \times 1) \) of the latent variable \( u_i \) as the maximum-variance column of \( Y_i \).
3. Calculate the score vector \((t_i(N \times 1)\) of the latent variable \( t_i \) of \( \Phi_i \) as \( t_i = K_i u_i / \|K_i u_i\|, \|t_i\| = 1 \).
4. Regress the columns of \( Y_i \) on \( t_i \); \( c_i = Y_i t_i \), where \( c_i \) is a weighting vector.
5. Compute the new score-vector: \( u_i = Y_i c_i / \|Y_i c_i\|, \|u_i\| = 1 \).
6. Iterate the steps (3) to (5) until the convergence of \( t_i \).
7. Deflate the residuals: \( K_{i+1} = (I - t_i t_i^T)K_i (I - t_i t_i^T) \)
\( Y_{i+1} = Y_i - t_i t_i^T Y_i \).
8. Retain the data in the matrices: \( T \leftarrow t_i, U \leftarrow u_i \).
9. Set \( i = i + 1 \), and return to step (2). Stop when \( i > \ell \), with \( \ell \) being the extracted number of latent variables.

After selecting the desired kernel latent variables, the regression coefficient matrix between \( \Phi \) and \( Y \) has the following form \([37], [38]\):

\[
B = \Phi^T U (T^T K U)^{-1} T^T Y
\]

(8)

The prediction outputs on training samples is given by

\[
\hat{Y} = \Phi B = K U (T^T K U)^{-1} T^T Y
\]

(9)

The output prediction of the testing samples \( x^* \) with the trained model is written as:

\[
\begin{align*}
\hat{y} &= B^T \phi(x^*) = Y^T T \left[U (T^T K U)^{-1}\right]^T \Phi \phi(x^*) \\
&= Y^T T \left[U (T^T K U)^{-1}\right]^T k(x^*)
\end{align*}
\]

(10)

where \( k(x^*) \) is the kernel vector,

\[
k(x^*) = \Phi \phi(x^*) = [k(x_1, x^*) \ldots k(x_N, x^*])^T
\]

(11)

**III. INTERVAL REDUCED KERNEL PLS METHOD**

This section presents the interval kernel PLS (IKPLS) methods. In effect, interval-valued data presents a method that takes into account the available sensors information with uncertainty and variability.

The dimension of the IKPLS model depends on the number of samples of the training data set. Reducing the structure of the IKPLS model leads to reduce the number of samples.

In this paper, therefore, two novel interval reduced kernel PLS (IRKPLS) models are developed to handle large data sets and reduce the computational complexity of the IKPLS models. In the proposed IRKPLS models, first, the Euclidean distance between samples is used as a dissimilarity metric to keep only the more relevant and informative observations in
the input space. Then, they apply the new reduced data set to build the IRKPLS models. Next, two proposed IRKPLS models based on Euclidean distance (IRKPLS UL and IRKPLS CR) are presented, respectively.

A. EUCLIDEAN DISTANCE METRIC FOR INTERVAL-VALUED DATA

Let $[X] = \{[x_1(1)], [x_2(1)], \ldots, [x_j(N)]\}$ and $[Y] = \{[y_1(1)], [y_2(1)], \ldots, [y_j(N)]\}$ be the interval valued data vector of inputs and outputs, where $[x_j(k)] = [\underline{x}_j(k), \overline{x}_j(k)] \forall k \in \{1, \ldots, N\}$, with $\underline{x}_j(k) \leq \overline{x}_j(k)$ and $[y_j(k)] = [\underline{y}_j(k), \overline{y}_j(k)]$ with $\underline{y}_j(k) \leq \overline{y}_j(k)$.

The new interval input matrix $[X]$ is given by

$$[X] = \left[ \begin{array}{cccc} [\underline{x}_1(1), \overline{x}_1(1)] & \ldots & [\underline{x}_m(1), \overline{x}_m(1)] \\ \vdots & \ddots & \vdots \\ [\underline{x}_1(N), \overline{x}_1(N)] & \ldots & [\underline{x}_m(N), \overline{x}_m(N)] \end{array} \right]$$

(12)

The interval valued data matrix $[Y]$ of outputs is expressed as

$$[Y] = \left[ \begin{array}{cccc} [\underline{y}_1(1), \overline{y}_1(1)] & \ldots & [\underline{y}_q(1), \overline{y}_q(1)] \\ \vdots & \ddots & \vdots \\ [\underline{y}_1(N), \overline{y}_1(N)] & \ldots & [\underline{y}_q(N), \overline{y}_q(N)] \end{array} \right]$$

(13)

The Euclidean distance between observations is used as dissimilarity metric. The idea behind dissimilarity metric is adopted to extract a most relevant subset from all available interval-valued samples. To do so, data consisting of measures of dissimilarity between all pairs of interval-valued observations can be designated using a dissimilarity matrix $D$ as

$$D = \begin{bmatrix} d_{11} & d_{12} & \ldots & d_{1N} \\ d_{21} & d_{22} & \ldots & d_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ d_{N1} & d_{N2} & \ldots & d_{NN} \end{bmatrix}$$

(14)

where $d_{ij}$ is the Euclidean distance between the rows $[X_i]$ and $[X_j]$ of the data matrix $[X]$. So, $d_{ij}$ is given by:

$$d_{ij} = \sum_{k=1}^{m} ([x_k(i)] - [x_k(j)])^2$$

(15)

with

$$([x_k(i)] - [x_k(j)])^2 = ([\underline{x}_k(i)] - [\underline{x}_k(j)])^2 + ([\overline{x}_k(i)] - [\overline{x}_k(j)])^2$$

(16)

Hence, the dissimilarity matrix $D$ is symmetric and its diagonal elements are null. Then, all redundancy in observations are eliminated from the input data matrix based on a pick out dissimilarity distance. This plays an important role in reducing the number of observations.

The resulting reduced interval matrix of inputs can be defined as:

$$[X'] = \left[ [x'(1)], \ldots, [x'(N_r)] \right]^T$$

(17)

where $N_r$ represents the number of reduced samples.

The corresponding reduced interval matrix of outputs is given by

$$[Y'] = \left[ [y'(1)], \ldots, [y'(N_r)] \right]^T$$

(18)

Once, the new reduced input and output matrices are evaluated, the interval-valued data matrices are transformed on a numerical data matrices, then, IKPLS model is applied on the computed matrices.

B. IRKPLS BASED ON INTERVAL UPPER AND LOWER BOUNDS (IRKPLS UL)

One of the first proposed approaches is IRKPS based on interval upper and lower bounds IRKPLS UL. It consists to build two KPLS models using the reduced inputs and outputs data matrices formed, respectively, by the lower and upper bounds of interval measurements. Let $x'(i) = [\underline{x}'_1(i), \ldots, \underline{x}'_m(i)]$, with $i \in \{1, \ldots, N_r\}$, be the lower interval input vector with lower bound $\underline{x}'(i)$. Let consider $y'(i) = [\underline{y}'_1(i), \ldots, \underline{y}'_m(i)]$ the lower bounds of interval output vector.

The reduced kernel matrix for lower bound data is expressed as,

$$K' = \begin{bmatrix} k(x'_1, x'_1) & \ldots & k(x'_1, x'_m) \\ \vdots & \ddots & \vdots \\ k(x'_N_r, x'_1) & \ldots & k(x'_N_r, x'_m) \end{bmatrix}$$

(19)

In IRKPS technique, the regression coefficients matrix $B'$ for lower bound of interval-valued data can be written as,

$$B' = \Phi'^T U'(T'^T K' U')^{-1} T'^T Y'$$

(20)

Therefore, the output prediction on reduced training samples for lower bound of interval-valued data is

$$\hat{Y}' = \Phi' B' = K' U'(T'^T K' U')^{-1} T'^T Y'$$

(21)

The interval output prediction on test data $x^*$ can be made as,

$$\hat{Y}^* = B'^T \Phi(x^*) = Y'^T T' \left[ U'(T'^T K' U')^{-1} \right] k'(x^*)$$

(22)

where

$$k'(x^*) = \left[ k(x'_1, x^*) \ldots k(x'_N_r, x^*) \right]^T$$

(23)

Define $\overline{x}'(i) = [\overline{x}'_1(i), \ldots, \overline{x}'_m(i)]$ as the upper bound of interval input vector, with upper bound $\overline{x}'(i)$. The upper bound of interval output vector can be represented of the form $\overline{y}'(i) = [\overline{y}'_1(i), \ldots, \overline{y}'_m(i)]$.

In this case, the second reduced kernel matrix $K''$ for upper bound of interval valued data is given as

$$K'' = \begin{bmatrix} k(\overline{x}'_1, \overline{x}'_1) & \ldots & k(\overline{x}'_1, \overline{x}'_m) \\ \vdots & \ddots & \vdots \\ k(\overline{x}'_{N_r}, \overline{x}'_1) & \ldots & k(\overline{x}'_{N_r}, \overline{x}'_m) \end{bmatrix}$$

(24)
In this case, the matrix of regression coefficient \( \hat{B} \) for the upper bound in the IRKPLS technique has the following form,

\[
\hat{B} = \Phi^T U (T^T K' U)^{-1} T^T \bar{Y}
\] (25)

The prediction of interval output variables is determined by

\[
\hat{Y} = \Phi \hat{B} = K' U (T^T K' U)^{-1} T^T \bar{Y}
\] (26)

As a result, the prediction on test subset for upper bound of interval-valued data is written as,

\[
\bar{y}_* = \hat{b}^T \phi(x_*) = \hat{y}_T T' \left[ U (T^T K' U)^{-1} \right]^T k' (\hat{x}_*)
\] (27)

where

\[
k' (\hat{x}_*) = \left[ k(\hat{x}_1, \hat{x}_*) \ldots k(\hat{x}_N, \hat{x}_*) \right]^T
\] (28)

The main steps of IRKPLSUL algorithm are outlined in Algorithm 2.

**Algorithm 2 IRKPLSUL Algorithm**

**Training step:**
- Inputs: \( N \times m \) input data matrices \( X, X \) and \( N \times q \) output data matrices \( Y, \bar{Y} \).
  1. Calculate the elements of the dissimilarity matrix \( D \).
  2. Pick the smallest distance from \( D \) and extract the new reduced data input matrices \( \bar{x} \in R_{N}^{N \times m}, \bar{x} \in R_{N}^{N \times m} \) and output data matrices \( \bar{y} \in R_{N}^{N \times q}, \bar{y} \in R_{N}^{N \times q} \).
  3. Scale the data matrices to zero mean and unit variance.
  4. Compute the kernel matrices \( K' \) and \( K'' \) and scale them.
  5. Determine the number kernel of latent variables, to be kept in the IRKPLS model.
  6. Compute the regression coefficients matrices \( \bar{B}' \) and \( \bar{B}'' \) as given by equations 20 and 25, respectively.

**Testing step:**
- 1. For new testing data vectors \( x*, \bar{x}, y* \) and \( \bar{y} \), map input data vectors into feature space to get \( K'(x*) \) (equation 23) and \( k'(\bar{x}) \) (equation 28).
  2. Compute the output estimations \( \hat{y*} \) and \( \bar{y} \) given by equations 22 and 27.

**C. IKPLS BASED ON INTERVAL CENTERS AND RANGES (IRKPLS)\)**

The proposed IRKPLS model aims first to transform the input [\( X \)] and output [\( Y \)] matrices into numerical matrices based on the interval center and range. Then, it consists to decrease the dimensions of the original data matrices using the Euclidean distance between observations as dissimilarity metric. After that, it uses both the center and the range variables to construct a reduced regression model. In the proposed IRKPLS method, new data matrices are built by the concatenation of center and range data matrices, and a KPLS model is applied to this new reduced data matrices. The new reduced input \( X'_{\text{CR}} \) and output \( Y'_{\text{CR}} \) data matrices are defined as,

\[
\begin{align*}
X'_{\text{CR}} &= [X'_{\text{C}}, X'_{\text{R}}] \in R_{N}^{N \times 2m} \\
Y'_{\text{CR}} &= [Y'_{\text{C}}, Y'_{\text{R}}] \in R_{N}^{N \times 2q}
\end{align*}
\] (29)

where the input center \( X'_{\text{C}} \) and range \( X'_{\text{R}} \) matrices, and output center \( Y'_{\text{C}} \) and range \( Y'_{\text{R}} \) matrices are given by:

\[
\begin{align*}
X'_{\text{C}} &= [x'_{1}, x'_{2}, \ldots, x'_{N}] \in R_{N}^{N \times m} \\
X'_{\text{R}} &= [x'_{1}, x'_{2}, \ldots, x'_{N}] \in R_{N}^{N \times m} \\
Y'_{\text{C}} &= [y'_{1}, y'_{2}, \ldots, y'_{N}] \in R_{N}^{N \times q} \\
Y'_{\text{R}} &= [y'_{1}, y'_{2}, \ldots, y'_{N}] \in R_{N}^{N \times q}
\end{align*}
\] (30)

Then, the input reduced data matrix is mapped to a feature space and the reduced kernel matrix \( K'_{\text{CR}} \) based on center and range approach is defined as,

\[
K'_{\text{CR}} = \left[ \begin{array}{ccc}
    k(x'_{\text{CR}_1, x'_{\text{CR}_1}}) & \cdots & k(x'_{\text{CR}_1, x'_{\text{CR}_N}}) \\
    \vdots & \ddots & \vdots \\
    k(x'_{\text{CR}_N, x'_{\text{CR}_1}}) & \cdots & k(x'_{\text{CR}_N, x'_{\text{CR}_N}})
  \end{array} \right]
\] (31)

The prediction of the output variables-based center and range method can be expressed as

\[
\hat{Y}'_{\text{CR}} = \Phi'_{\text{CR}} B'_\text{CR}
\]

\[
= K'_{\text{CR}} U'_{\text{CR}} (T'_{\text{CR}} K'_{\text{CR}} U'_{\text{CR}})^{-1} T'_{\text{CR}} Y'_{\text{CR}}
\] (32)

where the new regression coefficient \( B'_\text{CR} \) is given by

\[
B'_\text{CR} = \Phi'_{\text{CR}} U'_{\text{CR}} (T'_{\text{CR}} K'_{\text{CR}} U'_{\text{CR}})^{-1} T'_{\text{CR}} Y'_{\text{CR}}
\] (33)

For a new observation \( x'^* \) using the reduced data set, the interval output estimation \( \hat{y}'_{\text{CR}} \) based on the IRKPLS model is represented by,

\[
\hat{y}'_{\text{CR}} = B'_\text{CR} T'_{\text{CR}} \phi(x'^*)
\]

\[
= Y'_{\text{CR}} T'_{\text{CR}} \left[ U'_{\text{CR}} (T'_{\text{CR}} K'_{\text{CR}} U'_{\text{CR}})^{-1} \right] \phi(x'^*)
\] (34)

The IRKPLS model based center and range of interval algorithm is presented in Algorithm 3.

**IV. FAULT DETECTION BASED ON IRKPLS**

Once the IRKPLS models are built, the monitored interval residuals are computed and used later for process monitoring purposes. To monitor the behavior of a process, different detection indices are developed in literature [16], [28], [34], [36], [39]-[43]. Among the fault detection charts, the GLRT statistic is a hypothesis testing scheme [36], [44]-[47]. Let \( E \in R_{N}^{N} \) be a measured vector follows one of the two Gaussian distributions: \( \mathcal{N}(0, \sigma^2 I_N) \) or \( \mathcal{N}(\theta \neq 0, \sigma^2 I_N) \), where \( \theta \) is the mean vector and \( \sigma^2 \) is the known variance. The hypothesis testing problem is given by,

\[
\begin{align*}
H_0 &= \{ E \sim \mathcal{N}(0, \sigma^2 I_N) \} \quad \text{(null hypothesis)}; \\
H_1 &= \{ E \sim \mathcal{N}(\theta, \sigma^2 I_N) \} \quad \text{(alternative hypothesis)}.
\end{align*}
\] (35)

Using the GLRT approach, the unknown parameter \( \theta \) is changed by its maximum likelihood estimate. The GLRT G(E) decision function is computed as

\[
G(E) = 2 \log \left( \frac{\sup_{\theta \in R_{2r}} f_\theta(E)}{f_{\theta=0}(E)} \right)
\] (36)
Algorithm 3 IRKPLS\(_{CR}\) Algorithm

Training step:

- **Inputs**: \(N \times m\) input interval data matrix \([X]\) and \(N \times p\) output interval data matrix \([Y]\)
- 1. Compute the center and ranges matrices \(X^c, X^c\) and \(Y^c\), respectively,
- 2. Form the new data matrices \(X_{CR} = [X^c X^c]\) and \(Y_{CR} = [Y^c Y^c]\), respectively,
- 3. Calculate the elements of the dissimilarity matrix \(D\)
- 4. Pick the smallest distance from \(D\) and extract the new reduced input data matrices \(X'_{CR} \in \mathbb{R}^{N_r \times 2m}\) and \(Y'_{CR} \in \mathbb{R}^{N_r \times 2p}\), \(\bar{Y}' \in \mathbb{R}^{N_r \times q}\).
- 5. Standardize the data matrices to zero mean and unit variance, \(X'_{CR}\) and \(Y'_{CR}\).
- 6. Determine the kernel matrix \(K'_{CR}\) using data matrix \(X'_{CR}\).
- 7. Scale the data matrix \(K'_{CR}\).
- 8. Determine the number of kernel latent variable, to be kept in the IRKPLS model,
- 9. Determine the regression coefficients matrix \(B'_{CR}\) as given by equation 33,
- 10. Compute the estimations of output matrix \(\hat{Y}'_{CR}\) using equation 32.

Testing step:

- 1. For new testing data vectors \(x_{CR}^*\) and \(y_{CR}^*\), map the input data vector into feature space to get \(k(x_{CR}^*)\)
- 2. Compute the output estimations \(\hat{\gamma}_{CR}\) using equation 34.

where \(f_0(\varepsilon)\) is the probability density function which is given by,

\[
f_0(\varepsilon) = \frac{1}{(2\pi)^{\frac{N_r}{2}}\sigma^{N_r}} \exp\left(-\frac{1}{2\sigma^2} \|\varepsilon - \theta\|_2^2\right) \tag{37}
\]

The multivariate GLRT chart is determined for upper and lower bounds of interval-valued data in the feature space as

\[
\begin{align*}
\mathbf{G} &= \frac{1}{\hat{\sigma}^2} \|\mathbf{E}\|_2^2 \\
\bar{\mathbf{G}} &= \frac{1}{\hat{\sigma}^2} \|\bar{\mathbf{E}}\|_2^2
\end{align*}
\tag{38}
\]

where \(\sigma\) and \(\bar{\sigma}\) are the variance of the residuals \(\mathbf{E}\) and \(\bar{\mathbf{E}}\), respectively.

According to equations 22 and 27, the residuals based on the lower and upper bounds of interval-valued data are defined as

\[
\begin{align*}
\mathbf{E} &= y - \hat{y}_x \\
\bar{\mathbf{E}} &= \bar{y} - \hat{\bar{y}}_x
\end{align*}
\tag{39}
\]

A fault is detected using GLRT chart, if both upper and lower indices, \(\mathbf{G}\) and \(\bar{\mathbf{G}}\), are out of their respective thresholds, \(\mathbf{G}_\alpha\) and \(\bar{\mathbf{G}}_\alpha\),

\[
\begin{align*}
\mathbf{G} &> \mathbf{G}_\alpha \\
\bar{\mathbf{G}} &> \bar{\mathbf{G}}_\alpha
\end{align*}
\tag{40}
\]

The control limits \(\mathbf{G}_\alpha\) and \(\bar{\mathbf{G}}_\alpha\) are computed as

\[
\begin{align*}
\mathbf{G}_\alpha = \frac{g}{\hat{\sigma}^2} \chi^2_{k,\alpha} \\
\bar{\mathbf{G}}_\alpha = \frac{\bar{g}}{\hat{\sigma}^2} \chi^2_{k,\alpha}
\end{align*}
\tag{41}
\]

where,

\[
\begin{align*}
g &= \frac{b}{b - a} \quad \bar{g} = \frac{\bar{b}}{\bar{b} - \bar{a}} \\
\hat{h} &= \frac{2\alpha^2}{\chi^2_{k,\alpha}}
\end{align*}
\tag{42}
\]

with \(a\) and \(b\) are the mean and variance of the \(\mathbf{G}\) index, and \(\bar{a}\) and \(\bar{b}\) are the mean and variance of the \(\bar{\mathbf{G}}\) index.

The IRKPLS\(_{UL}\)-based GLRT fault detection strategy is shown in Figure 1.

From the IRKPLS\(_{CR}\) model and using equation 34, the residual is given by

\[
E_{CR}(x_{CR}^*) = \gamma_{CR} - \hat{\gamma}_{CR}
\tag{43}
\]

The multivariate fault detection GLRT chart, \(\mathbf{G}_{CR}(x_{CR}^*)\) of a new sample \(x_{CR}^*\), is computed in the feature space as:

\[
\mathbf{G}_{CR}(x_{CR}^*) = \frac{1}{\sigma^2} \|E_{CR}(x_{CR}^*)\|_2^2
\tag{44}
\]

Let \(\mathbf{G}_{CR,\alpha}\) be the corresponding control limit. A fault is detected if \(\mathbf{G}_{CR}(x_{CR}^*) > \mathbf{G}_{CR,\alpha}\).

The IRKPLS-based GLRT statistic algorithm is illustrated schematically in Figure 2.

V. CASE STUDY USING THE TENNESSEE EASTMAN (TE) PROCESS

To evaluate the detection efficiencies of the proposed techniques, an industrial benchmark of Tennessee Eastman (TE) process is used. [48]–[51]. The monitoring performance of both schemes is appraised based on false alarms rate (FAR) which is the percentage of false alarms in a healthy set, missed detection rate (MDR) which represents the percentage...
FIGURE 2. Diagram of IRKPLS-based GLRT statistic for fault detection.

A. TE PROCESS DESCRIPTION

The TE process has been widely used by the process monitoring community as a source of data for comparing various methods [48], [52]–[55]. TE process contains, G and H are the main products, A, C, D, and E are reactants (gaseous raw material), F is a by-product, and B is an inert gas. The irreversible exothermic chemical reactions that away place in the reactor are:

$$A(g) + C(g) + D(g) \rightarrow G(\text{liq})$$
$$A(g) + C(g) + E(g) \rightarrow H(\text{liq})$$
$$A(g) + E(g) \rightarrow F(\text{liq})$$
$$3D(g) \rightarrow 2F(\text{liq})$$

(45)

The TE process includes five principal units: a reactor, a compressor, a condenser, a stripper, and a vapor/liquid separator. More information about a process including the reactions are broadly detailed in the literature [56], [57] (Figure 3).

B. FAULT DETECTION RESULTS

In this study, a training dataset of 1024 samples is collected from the TE process under normal operating conditions, while 21 different testing datasets, each contains 1024 samples, are generated at a sampling rate of 3 minutes over 21 benchmark faults. Each sample of each variable is associated with an uncertainty bound of $\delta = 10\%$ to generate an interval-valued sample for the training and testing datasets. Figure 4 illustrates the time evolution of the interval-valued variable XMEAS(1) during the healthy and faulty statuses.

Two conventional and two developed IKPLS models are built using CPV criterion with 95% as an explained variance threshold. The two conventional IKPLS models, referred to IKPLS$_{CR}$ and IKPLS$_{UL}$, are structured by 13 and 12 retained kernel latent variables, respectively. The two IRKPLS$_{CR}$ and IRKPLS$_{UL}$ models are identified by only 585 extracted samples thanks to the Euclidean distance similarity measure from which 12 kernel latent variables are maintained for each of them. The nonlinear uncertain process monitoring is achieved using the different obtained IKPLS models, from which the fault detection index GLRT is assessed for comparison. Its threshold is established at $\alpha = 5\%$ as a significance level.

The obtained FAR, MDR and ET values contributed by IKPLS$_{CR}$, IKPLS$_{UL}$, IRKPLS$_{CR}$, and IRKPLS$_{UL}$ techniques based on GLRT statistic are summarized in Table 1.
TABLE 1. Summary of FAR, MDR and ET.

| Faults | FAR | MDR | ET |
|--------|-----|-----|----|
| Fault 1 | 0.005 | 0.015 | 0.8 |
| Fault 2 | 0.004 | 0.018 | 0.9 |
| Fault 3 | 0.006 | 0.012 | 0.7 |
| Fault 4 | 0.007 | 0.013 | 0.8 |
| Fault 5 | 0.008 | 0.014 | 0.9 |
| Fault 6 | 0.009 | 0.015 | 0.7 |

where the best of them are shown in bold. Whereas the underlined value represents the best value in all techniques for each performance metric. It can clearly seen that the FARs contributed by IRKPLS_CR-based GLRT method for almost faults are slightly improved compared to those resulting from the IKPLS_CR-based GLRT technique. Unlike, the IKPLS_UL-based GLRT contributes to a fairly acceptable practical FAR and ensures higher robustness compared to all algorithms. While the two developed IRKPLS models provide better results compared to the conventional ones in terms of MDR which is successfully improved for the most faults, ensuring reasonably correct detection. Furthermore, they also give an important ET enhancement in all testing cases. Thus, more than 57% of ET is gained by the IRKPLS_CR and IRKPLS_UL based GLRT techniques, as a result, memory requirements are satisfactorily reduced to 55%.

VI. CONCLUSIONS

In this work, we considered the problem of fault detection in large-scale nonlinear uncertain systems using interval input-output reduced kernel PLS (IRKPLS). The IRKPLS model presented an interval version of the conventional
KPLS model. It was developed to consider errors and imprecision of the measurements devices and uncertainties in the system. In addition to its ability to efficiently monitor these processes, the IRKPLS model had significantly reduced the execution time and memory space. The developed IRKPLS used the Euclidean distance between interval-valued samples as a dissimilarity measure to extract the more relevant and efficient samples. Then, the monitored interval residuals, evaluated using the IRKPLS models, were introduced to the generalized likelihood ratio test (GLRT) as input for fault detection purposes. The detection performance and effectiveness of the developed IRKPLS-based GLRT techniques were assessed through different kinds of faults of the well-known Tennessee Eastman (TE) process in terms of missed detection rate, false alarms rate and execution time. In addition, the execution time, and consequently the memory space had been remarkably reduced. Moreover, the proposed IRKPLS-based GLRT methods were able to easily manage the execution time and memory requirements while ensuring a required fault detection performance.

REFERENCES

[1] Z. He, Y. A. W. Shardt, D. Wang, B. Hou, H. Zhou, and J. Wang, “An incipient fault detection approach via detrending and denoising,” Control Eng. Pract., vol. 74, pp. 1–12, May 2018.

[2] Y. Zhang, W. Du, Y. Fan, and L. Zhang, “Process fault detection using directional kernel partial least squares,” Ind. Eng. Chem. Res., vol. 54, no. 9, pp. 2509–2518, Mar. 2015.

[3] H. Hao, K. Zhang, S. X. Ding, Z. Chen, and Y. Lei, “A data-driven multiplicative fault diagnosis approach for automation processes,” ISA Trans., vol. 53, no. 5, pp. 1436–1445, Sep. 2014.

[4] C. F. Alcala and S. Joe Qin, “Analysis and generalization of fault diagnosis methods for process monitoring,” J. Process Control, vol. 21, no. 3, pp. 322–330, Mar. 2011.

[5] S. J. Qin and Y. Zheng, “Quality-relevant and process-relevant fault monitoring with concurrent projection to latent structures,” AIChE J., vol. 59, no. 2, pp. 496–504, Feb. 2013.

[6] M. Lucke, A. Stief, M. Chioia, J. R. Oettle, and N. F. Thomhill, “Fault detection and identification combining process measurements and statistical alarms,” Control Eng. Pract., vol. 94, Jan. 2020, Art. no. 104195.

[7] T. Cheng, A. Dairi, F. Harrou, Y. Sun, and T. Leiknes, “Monitoring influential conditions of wastewater treatment plants by nonlinear data-based techniques,” IEEE Access, vol. 7, pp. 108827–108837, 2019.

[8] B. Khaldi, F. Harrou, F. Cherif, and Y. Sun, “Flexible and efficient topological approaches for a reliable robots swarm aggregation,” IEEE Access, vol. 7, pp. 96372–96383, 2019.

[9] F. Palumbo and C. N. Lauro, “A PCA for interval-valued data based on midpoints and radii,” in New Developments in Psychometrics. Tokyo, Japan: Springer, 2010, pp. 641–648.

[10] P. D’Urso and P. Giordani, “A least squares approach to principal component analysis for interval valued data,” Chemomercial Intell. Lab. Syst., vol. 70, no. 2, pp. 179–192, Feb. 2004.

[11] C. Bote, M. Mansouri, M. Nounou, H. Nounou, M. M. Karim, “Kernel PLS-based GLRT method for fault detection of chemical processes,” J. Loss Prevention Process Industries, vol. 43, pp. 212–224, Sep. 2016.

[12] M. Mansouri, M. N. Nounou, and H. N. Nounou, “Improved statistical fault detection technique and application to biological phenomena modeled by S-systems,” IEEE Trans. Nanobiosci., vol. 16, no. 6, pp. 504–512, Sep. 2017.

[13] R. Rospial and N. Krämer, “Overview and recent advances in partial least squares,” in International Statistical and Optimization Perspectives Workshop Subspace, Latent Structure and Feature Selection. Berlin, Germany: Springer, 2005.

[14] P. W. Wu, Z. Liu, and X. He, “Learning-based super resolution using kernel partial least squares,” Image Vis. Comput., vol. 29, no. 6, pp. 394–406, May 2011.

[15] B. S. Doyal and J. F. MacGregor, “Improved PLS algorithms,” J. Chemometrics, vol. 11, no. 1, pp. 73–85, Jan. 1997.

[16] S. Yin, C. Wang, and S. Zhang, “Performance modeling of centrifugal compressor using kernel partial least squares,” Appl. Thermal Eng., vol. 44, pp. 90–99, Nov. 2012.

[17] Y. Zhang, T. Huang, H. He, W. Zhou, Q. Han, and T. Li, “A novel framework for fault diagnosis using kernel partial least squares based on an optimal preference matrix,” IEEE Trans. Ind. Electron., vol. 64, no. 5, pp. 4315–4324, May 2017.

[18] Y. Zhang, Y. Teng, and Y. Zhang, “Complex process quality prediction using modified kernel partial least squares,” Chem. Eng. Sci., vol. 65, no. 6, pp. 2153–2158, Mar. 2010.

[19] B. M. Nikolai, K. I. Theron, and J. Lammertyn, “Kernel PLS regression on wavelet transformed NIR spectra for prediction of sugar content of apple,” Chemomeric Intell. Lab. Syst., vol. 85, no. 2, pp. 243–252, Feb. 2007.

[20] S. Yin, X. Zhu, and O. Kaynak, “Improved PLS focused on key-performance-indicator-related fault diagnosis,” IEEE Trans. Ind. Electron., vol. 62, no. 3, pp. 1651–1658, May 2015.
[40] J. Shang, M. Chen, and H. Zhang, “Fault detection based on augmented kernel mahalanobis distance for nonlinear dynamic processes,” Comput. Chem. Eng., vol. 109, pp. 311–321, Jan. 2018.

[41] S. Wang and F. Xiao, “AHU sensor fault diagnosis using principal component analysis method,” Energy Buildings, vol. 36, no. 2, pp. 147–160, Feb. 2004.

[42] Y. Zhang and C. Ma, “Fault diagnosis of nonlinear processes using multiscale KPCA and multiscale KPLS,” Chem. Eng. Sci., vol. 66, no. 1, pp. 64–72, Jan. 2011.

[43] M. Mansouri, M. N. Nounou, and H. N. Nounou, “Multiscale kernel PLS-based exponentially weighted-GLRT and its application to fault detection,” IEEE Trans. Emerg. Topics Comput. Intell., vol. 3, no. 1, pp. 49–58, Feb. 2019.

[44] T. Ferguson, Mathematical statistics: A decision theoretic approach. New York, NY, USA: Academic, 1967.

[45] E. Lehmann, Testing Statistical Hypotheses. New York: Chapman & Hall, 1996.

[46] T. Severini, Likelihood Methods in Statistics. London, U.K.: Oxford Univ. Press, 2000.

[47] Y. Pawitan, All Likelihood: Statistical Modeling and Inference Using Likelihood. London, U.K.: Oxford Univ. Press, 2001.

[48] J. J. Downs and E. F. Vogel, “A plant-wide industrial process control problem,” Comput. Chem. Eng., vol. 17, no. 3, pp. 245–255, Mar. 1993.

[49] M. Kano, S. Hasebe, I. Hashimoto, and H. Ohno, “Statistical process monitoring based on dissimilarity of process data,” AIChE J., vol. 48, no. 6, pp. 1231–1240, Jun. 2002.

[50] L. Wang and H. Shi, “Improved kernel PLS-based fault detection approach for nonlinear chemical processes,” Chin. J. Chem. Eng., vol. 22, no. 6, pp. 657–663, Jun. 2014.

[51] P. R. Lyman and C. Georgakis, “Plant-wide control of the tennessee eastman problem,” Comput. Chem. Eng., vol. 19, no. 3, pp. 321–331, Mar. 1995.

[52] S. Yin, S. X. Ding, A. Haghani, H. Hao, and P. Zhang, “A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark tennessee eastman process,” J. Process Control, vol. 22, no. 9, pp. 1567–1581, Oct. 2012.

[53] M. Mansouri, H. Nounou, M. F. Harkat, and M. Nounou, “Fault detection of chemical processes using improved generalized likelihood ratio test,” in Proc. 22nd Int. Conf. Digit. Signal Process. (DSP), Aug. 2017, pp. 1–5.

[54] H. Chen, P. Tião, and X. Yao, “Cognitive fault diagnosis in tennessee eastman process using learning in the model space,” Comput. Chem. Eng., vol. 67, pp. 33–42, Aug. 2014.

[55] T. Guo, D. Zhou, J. Zhang, M. Chen, and X. Tai, “Fault detection based on robust characteristic dimensionality reduction,” Control Eng. Pract., vol. 84, pp. 125–138, Mar. 2019.

[56] E. L. Russell, L. H. Chiang, and R. D. Braatz, Data-Driven Methods for Fault Detection and Diagnosis in Chemical Processes. Springer, 2012.

[57] M. Ammiche, A. Kouadri, and A. Bakdi, “A combined monitoring scheme with fuzzy logic filter for plant-wide tennessee eastman process fault detection,” Chem. Eng. Sci., vol. 187, pp. 269–279, Sep. 2018.

**RADHIA FEZAI** holds a postdoctoral position at the Electrical Engineering Program, Texas A&M University at Qatar. Her work focuses on the use of applied mathematics and statistics concepts to develop statistical data and model-driven techniques and algorithms for modeling, fault detection, and monitoring with the aim of improving industrial systems.

**KAMALELDIN ABDAYEH** received the M.Sc. degree in functional analysis from University College Dublin and the Ph.D. degree from University College Cork, Ireland, in 1997. He had his Post-doctoral at the Department of Process Engineering, University College Cork. Since 2001, he has been with Prince Sultan University, Saudi Arabia. He published more than 40 articles in various areas of pure and applied mathematics. His research areas include functional analysis, theoretical physics, discrete potential theory, fixed point theory, and quality monitoring and statistical hypothesis testing.

**MAJDI MANSOURI** (Member, IEEE) received the B.E. degree in electrical engineering from the Higher School of Communication of Tunisia (SUPCOM), Tunisia, in 2006, the master’s degree in electrical engineering from the School of Electronic, Informatique and Radiocommunications in Bordeaux (ENSEIRB), France, in 2008, the Ph.D. degree in electrical engineering from the University of Technology of Troyes (UTT), France, in 2011, and the H.D.R. degree in electrical engineering from the University of Orleans, France, in December 2019. He joined the Electrical Engineering Program, Texas A&M University at Qatar, in 2011, where he is currently an Associate Research Scientist. His work focuses on statistical signal processing. He is the author of more than 150 refereed journal and conference publications and book chapters, and has worked on several projects as a Lead Principal Investigator (LPI) and a Principal Investigator (PI). He served as the technical committee member of several international journals and conferences.

**ABDELMALEK KOUADRI** is currently a Professor of electrical engineering with the Institute of Electrical and Electronics Engineering, University M’Hamed Bougara of Boumerdés, Algeria. He has more than ten years of combined academic and industrial experience. He has published more than 50 refereed journal and conference publications and book chapters. He served as the technical committee member of several international journals and conferences. His research interests are in the area of systems engineering and control, with emphasis on process modeling, monitoring, and estimation.

**MOHAMED-FAOUZI HARKAT** received the B.E. degree in automatic control from Annaba University, Algeria, in 1996, and the Ph.D. degree from the Institut National Polytechnique de Lorraine (INPL), France, in 2003. From 2002 to 2004, he was an Assistant Professor with the School of Engineering Sciences and Technologies of Nancy (ESSTIN), France. In 2004, he joined the Electronics Department, Badji Mokhtar – Annaba University, Annaba, Algeria, where he is currently a Professor. He has over twenty years of research and practical experience in systems engineering and process monitoring. He is the author of more than 100 refereed journal and conference publications and book chapters. He served as an Associate Editor and in technical committees of several international journals and conferences.

**HAZEM NOOUNOU** (Senior Member, IEEE) is currently a Professor of electrical and computer engineering with Texas A&M University at Qatar. He has more than 19 years of academic and industrial experience. He has significant experience in research on control systems, databased control, system identification and estimation, fault detection, and system biology. He has been awarded several national priorities research program (NPRP) research projects in these areas. He has published more than 200 refereed journal and conference papers and book chapters. He has successfully served as the lead PI and a PI on five QNRF projects, some of which were in collaboration with other PIs in this proposal. He has served as an Associate Editor and on the technical committees of several international journals and conferences.
MOHAMED NOUNOU (Senior Member, IEEE) is currently a Professor of chemical engineering with TAMU-Texas A&M University at Qatar. He has more than 19 years of combined academic and industrial experience. He has published more than 200 refereed journal and conference publications and book chapters. His research interests are in the area of systems engineering and control, with emphasis on process modeling, monitoring, and estimation. He has successfully served as the lead PI and a PI on several QNRF projects (six NPRP projects and three UREP projects). He is a Senior Member of the AIChE (American Institute of Chemical Engineers).

HASSANI MESSAOUD has prepared the Ph.D. thesis at the University of Nice-Siophia Antipolis, France, in 1993, and the habilitation thesis at the School of Engineers, Tunis, in June 2001. Actually, he is currently a Professor and the Head of the Research Lab LARATSI (Code:LR13ES13), School of Engineers, Monastir, Tunisia. His main majoring is process identification and control as well as signal and image processing. He has published more than 200 refereed journal and conference publications and book chapters. He served as the technical committee member of several international journals and conferences.

* * *