A Review on Basic Data-Driven Approaches for Industrial Process Monitoring

Shen Yin, *Member, IEEE*, Steven X. Ding, Xiaochen Xie, *Student Member, IEEE*, and Hao Luo

***Abstract*—Recently, to ensure the reliability and safety of modern large-scale industrial processes, data-driven methods have been receiving considerably increasing attention, particularly for the purpose of process monitoring. However, great challenges are also met under different real operating conditions by using the basic data-driven methods. In this paper, widely applied data-driven methodologies suggested in the literature for process monitoring and fault diagnosis are surveyed from the application point of view. The major task of this paper is to sketch a basic data-driven design framework with necessary modifications under various industrial operating conditions, aiming to offer a reference for industrial process monitoring on large-scale industrial processes.**

***Index Terms*—Data-driven, fault diagnosis, industrial operating conditions, process monitoring.**

I. Introduction

With the increasing demands on production quality, system performance, and economic requirement, modern industrial processes are more complicated in both structure and automation degrees. The reliability and safety issues on these complicated industrial processes become the most critical aspects for system design and are receiving considerably increasing attention nowadays. Due to the various operational requirements and constraints of the underlying industrial processes, the design of model-based process monitoring and fault diagnosis systems has been a remarkable research topic during the past several decades. Based on the physical and mathematical knowledge of the industrial processes, the well-established model-based techniques have been successfully applied on plenty of processes for industrial electronics [1]–[8], automatic control systems [9]–[14], etc. Papers dedicated to review and analyze the state of the art of the model-based process monitoring and fault diagnosis techniques can be found in [15]–[20].

Manuscript received October 7, 2013; revised November 29, 2013; accepted December 13, 2013. Date of publication January 21, 2014; date of current version June 6, 2014. This work was supported in part by the National Natural Science Foundation of China under Grant 61304102 and in part by the National Natural Science Foundation of Liaoning Province, China, under Grant 2013020002.

S. Yin is with Research Center of Intelligent Control and Systems, Harbin Institute of Technology, Harbin 150001, China (e-mail: shen.yin2011@googlemail.com).

S. X. Ding and H. Luo are with the Institute for Automatic Control and Complex Systems (AKS), University of Duisburg–Essen, 47057 Duisburg, Germany (e-mail: steven.ding@uni-due.de; hao.luo@uni-due.de).

X. Xie is with the Research Institute of Intelligent Control and Systems, Harbin Institute of Technology, Harbin 150001, China (e-mail: xiaochenxie.helen@gmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIE.2014.2301773

The model-based schemes require *a priori* physical and mathematical knowledge of the process. After the construction of the process model based on the first principles, the well-established model-based techniques could be successfully applied. On the other hand, the data-driven schemes serve as an efficient alternative way, in which the necessary process information can be extracted directly from huge amounts of the recorded process data [21]–[23]. The multivariate statistical process monitoring methods, which utilize input and output information of the process, are very popular nowadays for the purpose of process monitoring and fault diagnosis, particularly principal component (PC) analysis (PCA) [24]–[29] and partial least squares (PLS) [30]–[34]. Most importantly, the ability to tackle large number of highly correlated variables shows the multivariate statistical methods' significant advantage. Because of their simple forms and low design efforts compared with the model-based techniques, the multivariate statistical methods are widely used for process monitoring in numerous industrial applications [35]–[40].

On the other hand, the industrial applications are generally typical dynamic processes and might operate under different operating conditions in the industrial environment. For these cases, the basic multivariate statistical approaches could not be successfully applied, due to their basic assumption under stationary ideal operating conditions. Based on these observations, it is essential to review basic data-driven process monitoring approaches particularly under industrial operating conditions to understand their original ideas, conditions for successful application, and limitations as well as necessary modifications to achieve better performance for the underlying processes. An efficient integration of data-driven and model-based approaches shall be necessary by considering carefully the pros and cons of model-based and data-driven approaches. Without sophisticated system design efforts, the advanced data-driven approaches should improve the applicability, capacity, and efficiency of the basic data-driven process monitoring methods under various industrial operating conditions. The main objective of the paper is to review the latest results in the literature in order to provide a reference for researchers and engineers focused on process monitoring issue within industrially oriented data-driven framework.

The rest of this paper is organized as follows. The basic process monitoring approaches are reviewed in their standard form in Section II, while the modifications of the basic approaches under ideal operation conditions will be further studied in Section III. Section IV is dedicated to the issues of system dynamics and uncertainties met under industrial conditions. The combined use of the data-driven and model-based methods

and the application of the adaptive techniques will be addressed in detail. The conclusions and discussions are given in the last section.

II. Basic Data-Driven Process Monitoring Approaches

The most popular multivariate statistical approaches for process monitoring like standard PCA and PLS will be generally studied in this section. Remarks on the recent development of statistical approaches will be also addressed.

A. Basic Application Form of PCA

PCA is a basic multivariate statistical method which could preserve the significant variability information extracted from the process measurements and proposed originally for the purpose of dimensionality reduction of huge amount of correlated data. Because of its simple form and its ability to handle large number of process data, PCA has been widely and successfully applied in many computational areas, such as image analysis, feature extraction, pattern recognition, data compression, and time series prediction. PCA has been accepted as an effective and powerful tool in the multivariate statistical framework. In addition, its numerous variants have been popularly used in the process industry for the purpose of process monitoring and fault diagnosis in the last decade.

The training data, which are used in the off-line training phase of PCA-based process monitoring schemes, can be obtained either from the process or from a simulation platform which could imitate the desired process behavior. Suppose that a recorded process measurement data set for training purposes, which contains N samples of m-dimensional measured variables of an industrial process, is given and then normalized to zero mean and unit variance, i.e., denoted as

$$Z = \begin{bmatrix} z_1^{\mathrm{T}} \\ z_2^{\mathrm{T}} \\ \vdots \\ z_N^{\mathrm{T}} \end{bmatrix} = \begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1m} \\ z_{21} & z_{22} & \cdots & z_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ z_{N1} & z_{N2} & \cdots & z_{Nm} \end{bmatrix} \in \mathcal{R}^{N\times m}. \quad (1)$$

The covariance matrix can thus be approximated as

$$\Phi = \frac{1}{N-1} Z^{\mathrm{T}} Z. \quad (2)$$

In order to obtain the significant variability information from the preprocessed data, perform singular value decomposition (SVD) on the approximated covariance matrix Φ, i.e.,

$$\frac{1}{N-1} Z^{\mathrm{T}} Z = P \Lambda P^{\mathrm{T}} \quad (3)$$

where $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$ with $\lambda_1 \geq \cdots \geq \lambda_m \geq 0$. According to the magnitude of the singular values, P and Λ could be divided into

$$P = [P_{\mathrm{pc}} \quad P_{\mathrm{res}}] \qquad \Lambda = \begin{bmatrix} \Lambda_{\mathrm{pc}} & 0 \\ 0 & \Lambda_{\mathrm{res}} \end{bmatrix}$$

respectively. The decomposition of P, Λ into $P_{\mathrm{pc}}, P_{\mathrm{res}}$ as well as $\Lambda_{\mathrm{pc}}, \Lambda_{\mathrm{res}}$ results in two subspaces in the m-dimensional measurement subspace. The subspace spanned by $P_{\mathrm{pc}}^{\mathrm{T}}$ and the subspace spanned by $P_{\mathrm{res}}^{\mathrm{T}}$ are called the principal subspace and the residual subspace, respectively. Let β denote the number of the PCs. Then, $P_{\mathrm{pc}} \in \mathcal{R}^{m\times\beta}$ contains the singular vectors corresponding to the first β large singular values in Λ, and $P_{\mathrm{res}} \in \mathcal{R}^{m\times(m-\beta)}$ contains the singular vectors corresponding to the last $m-\beta$ small singular values. Projecting the m-dimensional measured variables z onto the aforementioned two orthogonal subspaces, it follows that

$$\hat{z} = P_{\mathrm{pc}}^{\mathrm{T}} z \in S_p \equiv \mathrm{Span}\{P_{\mathrm{pc}}\} \quad (4)$$

$$\tilde{z} = P_{\mathrm{res}}^{\mathrm{T}} z \in S_r \equiv \mathrm{Span}\{P_{\mathrm{pc}}\}^{\perp} \quad (5)$$

in which $\mathrm{Span}\{P_{\mathrm{pc}}\}$ denotes all linear combinations of column vectors in $P_{\mathrm{pc}}^{\mathrm{T}}$ and $\mathrm{Span}\{P_{\mathrm{pc}}\}^{\perp}$ is the orthogonal complement of $\mathrm{Span}\{P_{\mathrm{pc}}\}$. For a successful process monitoring purpose, the abnormal changes of the process in the aforementioned two subspaces should both be detected. Generally, in the online detection phase, two test statistics, i.e., squared prediction error (SPE) and T^2, which are respectively formed by means of the projections $\tilde{z}$ and $\hat{z}$, are defined and evaluated online for process monitoring

$$\mathrm{SPE} = z^{\mathrm{T}} P_{\mathrm{res}} P_{\mathrm{res}}^{\mathrm{T}} z \quad (6)$$

$$T^2 = z^{\mathrm{T}} P_{\mathrm{pc}} \Lambda_{\mathrm{pc}}^{-1} P_{\mathrm{pc}}^{\mathrm{T}} z. \quad (7)$$

The thresholds for the fault detection can be determined respectively with a given confidence level α as follows:

$$J_{\mathrm{th,SPE}} = \theta_1 \left(\frac{c_\alpha \sqrt{2\theta_2 h_0^2}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} \right)^{1/h_0} \quad (8)$$

$$J_{\mathrm{th},T^2} = \frac{\beta(N^2-1)}{N(N-\beta)} F_\alpha(\beta, N-\beta) \quad (9)$$

in which c_α denotes the normal deviate corresponding to the upper $1-\alpha$ percentile, $F(\beta, N-\beta)$ denotes the F-distribution with β, $N-\beta$ degrees of freedom, and

$$h_0 = 1 - \frac{2\theta_1\theta_3}{3\theta_2^2} \qquad \theta_i = \sum_{j=\beta+1}^{m} (\lambda_j)^i; \quad i = 1, 2, 3.$$

To this end, the fault detection logic is given as follows:

$$\begin{cases} \mathrm{SPE} \leq J_{\mathrm{th,SPE}} \quad \text{and} \quad T^2 \leq J_{\mathrm{th},T^2}, & \text{fault-free} \\ \text{otherwise}, & \text{faulty}. \end{cases}$$

Note that the PCA-based process monitoring schemes claim that the assumptions on the multivariate Gaussian distribution of the measurement variables have to be fulfilled. In addition, the normalization in the off-line training phase gives the same weighting for each measurement variable, in which the relationship of the input and output variables has not been taken into consideration. However, the input–output correlation could provide more beneficial information for the purpose of

process monitoring. In the next section, another basic multivariate statistical method with consideration of the input–output correlation will be further introduced.

B. Basic Application Form of PLS

PLS is another basic multivariate statistical method and extensively used for model building, fault detection, and diagnosis. By the use of the off-line trained correlation model and online process measurements, the prediction of key performance indicator of an industrial process can be achieved online. For the purpose of process monitoring, PLS is able to detect the faults which occurred in the process input by the use of the information contained in the input–output correlation.

Suppose that a normalized input data set U, which contains N samples of l-dimensional process input variables, and a normalized output (key performance indicator) data set Y, which contains N samples of m-dimensional process outputs, are given as

$$U = \begin{bmatrix} u_1^{\mathrm{T}} \\ u_2^{\mathrm{T}} \\ \vdots \\ u_N^{\mathrm{T}} \end{bmatrix} \in \mathcal{R}^{N\times l} \qquad Y = \begin{bmatrix} y_1^{\mathrm{T}} \\ y_2^{\mathrm{T}} \\ \vdots \\ y_N^{\mathrm{T}} \end{bmatrix} \in \mathcal{R}^{N\times m}$$

where $u_i \in \mathcal{R}^l$ and $y_i \in \mathcal{R}^m$, for $i = 1, \ldots, N$. Define the so-called latent variables (LVs) as follows:

$$T = [t_1 \quad t_2 \quad \cdots \quad t_\gamma] \in \mathcal{R}^{N\times\gamma}$$

where γ denotes the number of LVs. Based on the projection of the inputs U and outputs Y onto LVs, the input–output correlation could be determined by

$$U = TP^{\mathrm{T}} + \tilde{U} = \hat{U} + \tilde{U} \tag{10}$$

$$Y = TQ^{\mathrm{T}} + E_y = UM + E_y \tag{11}$$

in which $P \in \mathcal{R}^{l\times\gamma}$ and $Q \in \mathcal{R}^{m\times\gamma}$ denote the loading matrices of U and Y, respectively. $\hat{U} = TP^{\mathrm{T}}$ is highly correlated with Y. $\tilde{U}$ and E_y are residual subspaces and assumed to be uncorrelated with U and Y, respectively. According to the input–output correlation given in (10) and (11), T and the coefficient matrix $M \in \mathcal{R}^{l\times m}$ can be determined by

$$T = UR \qquad M = RQ^{\mathrm{T}} \tag{12}$$

where $P^{\mathrm{T}}R = R^{\mathrm{T}}P = I_{\gamma\times\gamma}$ and $R \in \mathcal{R}^{l\times\gamma}$. Since the input variables are highly correlated to the output variables, we project the input variables onto the following two subspaces:

$$\hat{u} = PR^{\mathrm{T}}u \in S_{\hat{u}} \equiv \mathrm{Span}\{P\} \tag{13}$$

$$\tilde{u} = (I_{l\times l} - PR^{\mathrm{T}})u \in S_{\tilde{u}} \equiv \mathrm{Span}\{R\}^{\perp}. \tag{14}$$

With appropriate test statistics to monitor the aforementioned subspaces, successful process monitoring could be achieved consequently. Like PCA, the following T^2 and SPE test statistics are commonly applied for the online monitoring, based on which the faults that occurred in $\hat{u}$ and $\tilde{u}$ could be detected, i.e.,

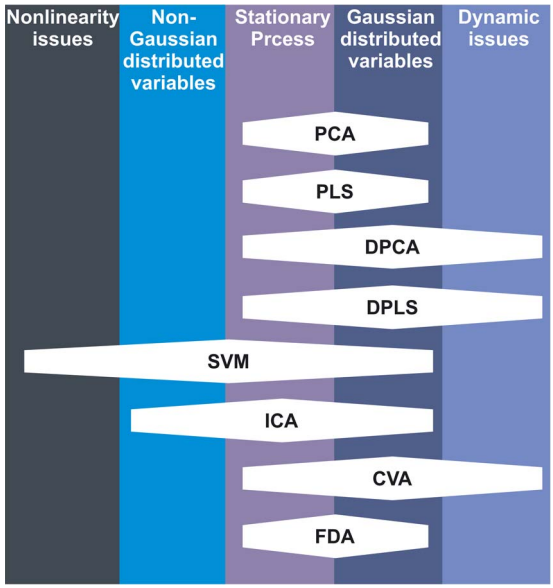


Fig. 1. Popular data-driven approaches for process monitoring.

$$T^2 = u^{\mathrm{T}}R\left(\frac{T^{\mathrm{T}}T}{N-1}\right)^{-1}R^{\mathrm{T}}u \tag{15}$$

$$\mathrm{SPE} = \|\tilde{u}\|^2 = \left\|(I_{l\times l} - PR^{\mathrm{T}})u\right\|^2. \tag{16}$$

Given a confidence level α, the thresholds of the aforementioned two test statistics can be determined by

$$J_{\mathrm{th},T^2} = \frac{\gamma(N^2-1)}{N(N-\gamma)}F_\alpha(\gamma, N-\gamma) \tag{17}$$

$$J_{\mathrm{th,SPE}} = g\chi^2_\alpha(h) \tag{18}$$

in which $\chi^2(h)$ denotes the χ^2-distribution with h degrees of freedom. h and g are determined respectively by $h = 2\mu^2/S$ and $g = S/2\mu$, in which μ denotes the mean value and S stands for the variance of SPE test statistic.

It is evident that, similar to the PCA-based schemes, for successful application of the PLS-based approaches, the input and output variables should also follow the multivariate Gaussian distribution.

C. Recent Developments on Data-Driven Approaches for Process Monitoring

The data-driven approaches, which are popularly applied in industry for process monitoring, are shown in Fig. 1. Within the multivariate statistical framework, data-driven approaches like PCA, PLS, and their variants are commonly applied for process monitoring. To successfully apply the standard PCA and PLS schemes, a linear process that runs under stationary operating conditions is a primary assumption. However, the process dynamics and nonlinearity are very important aspects in industrial processes. In order to design a robust process monitoring system which is able to cope with the real industrial environment, more and more attention has been paid to the PCA- and PLS-based process monitoring methods. A

PLS-based scheme named total projection to latent structure (TPLS) is proposed in [41] to handle the oblique decomposition problem contained in standard PLS. Recently, a modified PLS approach is proposed in [42], which provides a better process monitoring performance while less computation cost is required compared with all the PLS-based process monitoring methods. In [43], the so-called concurrent projection to latent structures (CPLS) method is proposed. Based on the concurrent projection of input and output data spaces, CPLS delivers a complete monitoring of faults that happen in the predictable output subspace and the unpredictable output-residual subspace. To tackle the dynamic issues of the industrial process, variants like dynamic PCA and dynamic PLS are proposed [44]–[48]. Recently, the kernel-based PCA and PLS approaches [49]–[52] and support vector machine [53], [54] have been intensively studied in order to deal with the nonlinearity in the industrial process.

Some other multivariate statistical techniques are also recently applied on industrial processes; an observation could be taken from the applications of Fisher discriminant analysis (FDA) and independent component (IC) analysis (ICA). FDA is originally proposed for dimensionality reduction and has been recently widely applied in the multivariate statistic fields and pattern classification [55], [56]. For process monitoring, FDA is an efficient tool for the purpose of fault classification [36]. Moreover, with consideration of an additional class of data denoting the normal operating conditions, FDA can also be applied on industrial processes for fault detection [57]–[59]. Another multivariate statistical technique, ICA is a well-known technique for solving blind source separation problems [60]–[62] by separation of ICs from process measurements. ICA has been recently extended and applied for the purpose of fault detection, particularly for the processes whose measurements are non-Gaussian distributed [63], [64]. Although the calculation involved in ICA is more complicated than the standard PCA and PLS approaches, it is still worthy to investigate whether extra bonus for process monitoring could be delivered from the independence among ICs. An alternative modified ICA algorithm is proposed in [65], which provides a unique solution of ICs and needs less computation cost than the standard approach. Recently, an attractive research field could be observed in the literature, in which the process monitoring is achieved by using subspace identification method (SIM)-based schemes. One of the popular applied SIMs is the so-called canonical variate analysis (CVA). CVA is originally proposed as a dimensionality reduction technique [66], [67], based on which the fault detection could be achieved by using the pairs of variables selected from the input–output data set that maximize the correlation [46], [68], [69]. A brief comparison of the computation complexity among the previously discussed data-driven process monitoring methods is given in Table I.

Although lots of complicated variants of the aforementioned approaches have been proposed to tackle different issues in industrial processes, a simple and convenient scheme, which could reduce the design efforts, particularly without special knowledge and high-cost computations, should always be an interesting topic from the industry point of view. Moreover, the necessary modifications on basic approaches are still essential and have been paid more attention nowadays.

TABLE I
BRIEF COMPARISON AMONG BASIC DATA-DRIVEN METHODS

Method	Computation complexity
PCA	Low: 1 SVD on $m \times m$ matrix.
DPCA	Medium: 1 SVD on $hm \times hm$ matrix.
PLS	Medium: γ times SVDs on $m \times m$ matrix.
DPLS	Medium: γ times SVDs on $hm \times hm$ matrix.
TPLS	Medium: cost of PLS + 2 SVDs on $m \times m$ matrix + 1 SVD on $\alpha \times \alpha$ matrix.
CPLS	Medium: cost of PLS + 1 SVDs on $m \times m$ matrix + 1 SVD on $\alpha \times \alpha$ matrix.
MPLS	Low: 2 SVDs on $m \times m$ matrix.
FDA	Medium: generalized EVD (eigenvalue decomposition) on $m \times m$ matrix.
ICA/MICA	High: cost of PCA + iterative constraint optimization problems.
CVA	Medium: 1 SVD on $(m+l)s \times (s_f+1)m$ matrix.

III. MODIFICATIONS UNDER IDEAL OPERATING CONDITION

A. On PCA and Its Test Statistic

As aforementioned, PCA is able to retain the significant variability information during the dimension reduction. However, for process monitoring, it does not lighten the computational load since the process data are projected on both principal and residual subspaces according to (4) and (5). From the mathematical viewpoint, PCA-based approaches are based on the SVD and seem numerically reliable. In order to ensure a better fault detection property, the thresholds of the standard test statistics should be reconsidered.

Denote the mean-centered measurement as $z \sim \mathcal{N}(0, \Phi)$; the projections of z onto $P_{\rm pc}$ and $P_{\rm res}$ are evaluated according to (4) and (5) with $z = \hat{z} + \tilde{z}$, $P_{\rm pc}^{\rm T} z \sim \mathcal{N}(0, \Lambda_{\rm pc})$, and $P_{\rm res}^{\rm T} z \sim \mathcal{N}(0, \Lambda_{\rm res})$. Assuming that the number of the collected data N is large enough, instead of F-distribution given in (9), it is more reasonable to apply the χ^2-distribution for the threshold setting, which could be determined by

$$J_{{\rm th},T^2} = \chi_\alpha^2(\beta) \tag{19}$$

where α stands for the confidence level defined by the user and β denotes the number of PCs. Similar to the scheme indicated by (7), the following Hawkin's test statistic has also been proposed and applied for monitoring the residual subspace:

$$T_H^2 = z^{\rm T} P_{\rm res} \Lambda_{\rm res}^{-1} P_{\rm res}^{\rm T} z. \tag{20}$$

However, the Hawkin's test statistic may cause a computational problem because the singular values of $\lambda_{\beta+1}, \ldots, \lambda_m$ might be very small and close to zero that leads $\Lambda_{\rm res}^{-1}$ to be ill-conditioned. To deal with this difficulty, SPE test statistic is usually applied, while the corresponding threshold setting could be determined by (8). However, instead of the SPE test statistic given in (6), an alternative scheme has been proposed in [70], which could solve the possible ill-conditioning problems of $\Lambda_{\rm res}^{-1}$ with reliable numerical property. Let

$$\Xi = {\rm diag}\left(\frac{\lambda_m}{\lambda_{\beta+1}}, \ldots, \frac{\lambda_m}{\lambda_{m-1}}, 1\right) \in \mathcal{R}^{(m-\beta)\times(m-\beta)}.$$

It turns out that $\Xi^{1/2}P_{\rm res}^{\rm T}z \sim \mathcal{N}(0, \lambda_m I_{(m-\beta)\times(m-\beta)})$ and $z^{\rm T}P_{\rm res}\Xi P_{\rm res}^{\rm T}z = \lambda_m z^{\rm T}P_{\rm res}\Lambda_{\rm res}^{-1}P_{\rm res}^{\rm T}z$. Therefore, an alternative SPE test statistic and its corresponding threshold for monitoring the residual subspace are given as follows:

$$T_{\rm new}^2 = z^{\rm T}P_{\rm res}\Xi P_{\rm res}^{\rm T}z \tag{21}$$

$$J_{{\rm th},T_{\rm new}^2} = \lambda_m \chi_\alpha^2(m-\beta). \tag{22}$$

Since a fault might occur either in the PC subspace spanned by $P_{\rm pc}^{\rm T}$ or in the residual subspace spanned by $P_{\rm res}^{\rm T}$, the projections of the mean-centered measurement vector z onto both subspaces, namely, (4) and (5), should be both monitored. It is well known that, based on the generalized likelihood ratio (GLR) method, optimal detection results could be achieved for a given N [13]. It is easy to verify that the T^2 test statistic is indeed the GLR test statistic when the parameter N for online evaluation is equal to one. In other words, T^2 and $T_{\rm new}^2$ provide an optimal solution for monitoring the principal and residual subspaces, respectively. Different from conventional PCA-based process monitoring approach whose thresholds are determined based on an approximation of the covariance matrix, $T_{\rm new}^2$ follows χ^2-distribution, and its corresponding threshold could be directly determined without using any off-line training results. Furthermore, the computation cost of (21) and (22) is considerably reduced compared with the standard techniques.

Due to the nature of SVD, the two subspaces, spanned respectively by $P_{\rm pc}$ and $P_{\rm res}$, are mutually orthogonal. As a result, the fault that occurred in one subspace cannot be detected by using the detection method for the other. A combined index [26], [38] which simultaneously uses both test statistics is proposed, i.e.,

$$T_c^2 = \beta_1 T^2 + \beta_2 T_{\rm res}^2 \tag{23}$$

with known constants $\beta_1, \beta_2 > 0$. Rewrite (23) into $T_c^2 = z^{\rm T}P\Psi P^{\rm T}z$, where

$$\Psi = \begin{bmatrix} \beta_1\Lambda_{\rm pc}^{-1} & 0 \\ 0 & \beta_2 Q \end{bmatrix} \qquad Q = \begin{cases} \Lambda_{\rm res}^{-1}, & T_{\rm res}^2 = T_H^2 \\ I, & T_{\rm res}^2 = {\rm SPE} \\ \Xi, & T_{\rm res}^2 = T_{\rm new}^2. \end{cases}$$

Recall that $P^{\rm T}z \sim \mathcal{N}(0,\Lambda)$, $z^{\rm T}P\Lambda^{-1}P^{\rm T}z \sim \chi^2(m)$, the corresponding combined test statistic is

$$T_{\rm comb}^2 = \lambda_m\left(T^2 + T_H^2\right) = z^{\rm T}P\bar{\Xi}P^{\rm T}z \tag{24}$$

with $\bar{\Xi} = {\rm diag}(\lambda_m/\lambda_1, \ldots, \lambda_m/\lambda_{m-1}, 1)$. The corresponding threshold could be determined with a user-defined confidence level α by

$$J_{\rm th,comb} = \lambda_m \chi_\alpha^2(m). \tag{25}$$

The aforementioned combined test statistic (24) is in a quadratic form of the observation projection $P^{\rm T}z$ with a weighting factor $\bar{\Xi}$. It is worth to point out that the direction of projection in $P^{\rm T}$ corresponding to a larger singular value is less weighted while the direction corresponding to a smaller singular value has a stronger weighting.

For the special case that the covariance matrix Φ is not full rank, we have

$$\Phi = \frac{1}{N-1}Z^{\rm T}Z = [P \quad P_\perp]\begin{bmatrix}\Lambda & 0 \\ 0 & 0\end{bmatrix}\begin{bmatrix}P^{\rm T} \\ P_\perp^{\rm T}\end{bmatrix}.$$

The combined test statistic (24) and its corresponding threshold (25), together with the parity verification

$$P_\perp^{\rm T}z = 0 \tag{26}$$

can be applied to detect the faults that occurred both in the principal and residual subspaces. The corresponding detection logic is

$$\begin{cases} T_{\rm comb}^2 \le J_{\rm th,comb} \text{ and (26) is true}, & \text{fault-free} \\ \text{otherwise}, & \text{faulty}. \end{cases}$$

B. On PLS and Its Modification

PLS extracts the correlation model from the process inputs and outputs for further prediction and fault diagnosis purposes. Unlike PCA, PLS inclines to discover the faults that occurred in process inputs, which might influence the key performance indicator. Recently, PLS and its variants have been comprehensively studied. It reported that the standard PLS may cause problems for fault diagnosis since the oblique projection between two subspaces $\hat{U}$ and $\tilde{U}$ [30]. In addition, the standard PLS approach results a partial correlation of $\hat{U}$ and Y due to the oblique projection; thus, to use the orthogonal variations in $\hat{U}$ to predict the process outputs seems problematic. Furthermore, $\tilde{U}$ may contain large variability of U; then, it cannot be treated as a residual subspace for process monitoring purposes. In [41], TPLS is proposed to tackle the aforementioned problems, which performs decomposition on certain subspaces based on the results of standard PLS.

To simplify the task, a desired relation can be described as

$$Y = UM + E_y = \hat{Y} + E_y \tag{27}$$

in which the coefficient matrix M contains the input–output correlation. E_y denotes the residual subspace of Y and is uncorrelated with the process inputs, i.e., ${\rm cov}(e_y, u) = 0$, where u and e_y are vectors in U and E_y, respectively. Generally, M is assumed to be of full-column rank. According to (27), when $N \gg \max\{l, m\}$, it yields

$$\frac{1}{N-1}Y^{\rm T}U = \frac{1}{N-1}M^{\rm T}U^{\rm T}U + \frac{1}{N-1}E_y^{\rm T}U \approx M^{\rm T}\frac{U^{\rm T}U}{N-1}.$$

If $U^{\rm T}U$ is full rank, we have $M = (U^{\rm T}U)^{-1}U^{\rm T}Y$; else, if $U^{\rm T}U$ is singular, it follows that $M = (U^{\rm T}U)^{\dagger}U^{\rm T}Y$, where $\dagger$ denotes the pseudoinverse: $(U^{\rm T}U)^{\dagger} = P_{u,{\rm pc}}\Lambda_{u,{\rm pc}}^{-1}P_{u,{\rm pc}}^{\rm T}$. Thus, Y is divided into two desired uncorrelated components, as expressed in (27).

In order to decompose U into two uncorrelated parts. An orthogonal projection of U onto the subspaces $S_{\hat{u}} \equiv span\{M\}$ and $S_{\tilde{u}} \equiv span\{M\}^\perp$ is performed. As a result, we have

$$\tilde{U}M = 0 \qquad \hat{Y} = UM = \hat{U}M \tag{28}$$

$$\hat{U} \in S_{\hat{u}} \qquad \tilde{U} \in S_{\tilde{u}}. \tag{29}$$

The calculation of the aforementioned equations requires following steps.

1) Do SVD on MM^{T}

$$MM^{\mathrm{T}} = [P_M \quad \tilde{P}_M] \begin{bmatrix} \Lambda_M & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} P_M^{\mathrm{T}} \\ \tilde{P}_M^{\mathrm{T}} \end{bmatrix} \tag{30}$$

with $P_M \in \mathcal{R}^{l\times m}$, $\tilde{P}_M \in \mathcal{R}^{l\times(l-m)}$, and $\Lambda_M \in \mathcal{R}^{m\times m}$.

2) Construct Π_M and $\Pi_M^{\perp}$, based on which U could be projected perpendicularly onto two orthogonal subspaces $S_{\hat{u}}$ and $S_{\tilde{u}}$, respectively, i.e.,

$$\Pi_M = P_M P_M^{\mathrm{T}} \qquad \Pi_M^{\perp} = \tilde{P}_M \tilde{P}_M^{\mathrm{T}}. \tag{31}$$

3) Project U onto two orthogonal subspaces $S_{\hat{u}} \equiv span\{M\}$ and $S_{\tilde{u}} \equiv span\{M\}^{\perp}$, respectively. U could thus be divided into two perpendicular parts, i.e.,

$$U = \hat{U} + \tilde{U} \tag{32}$$

such that

$$\hat{U} = U\Pi_M = UP_M P_M^{\mathrm{T}} \qquad \tilde{U} = U\Pi_M^{\perp} = U\tilde{P}_M \tilde{P}_M^{\mathrm{T}}.$$

To this end, the following T^2 test statistic is utilized for monitoring subspace $\hat{U}$:

$$T_{\hat{u}}^2 = u^{\mathrm{T}} P_M \left(\frac{P_M^{\mathrm{T}} U^{\mathrm{T}} U P_M}{N-1} \right)^{-1} P_M^{\mathrm{T}} u \tag{33}$$

while the corresponding threshold is given by

$$J_{\mathrm{th},T_{\hat{u}}^2} = \frac{m(N^2-1)}{N(N-m)} F_\alpha(m, N-m) \tag{34}$$

in which $F(m, N-m)$ denotes the F-distribution with m and $N-m$ degrees of freedom and α is the confidence level defined by the user. If N is large enough, then the threshold (34) can be determined by using χ^2-distribution table, i.e.,

$$J_{\mathrm{th},T_{\hat{u}}^2} = \chi_\alpha^2(m). \tag{35}$$

The fault detection logic is

$T_{\hat{u}}^2 > J_{\mathrm{th},T_{\hat{u}}^2} \Rightarrow$ a fault contained in inputs with influences on outputs is detected.

Similarly, for monitoring the subspace $\tilde{U}$, we have

$$T_{\tilde{u}}^2 = u^{\mathrm{T}} \tilde{P}_M \left(\frac{\tilde{P}_M^{\mathrm{T}} U^{\mathrm{T}} U \tilde{P}_M}{N-1} \right)^{-1} \tilde{P}_M^{\mathrm{T}} u. \tag{36}$$

The corresponding threshold for $T_{\tilde{u}}^2$ is

$$J_{\mathrm{th},T_{\tilde{u}}^2} = \frac{(l-m)(N^2-1)}{N(N-l+m)} F_\alpha(l-m, N-l+m). \tag{37}$$

Similarly, the threshold can be determined by $J_{\mathrm{th},T_{\tilde{u}}^2} = \chi_\alpha^2(l-m)$ when N is large enough. The associated fault detection logic is

$T_{\tilde{u}}^2 > J_{\mathrm{th},T_{\tilde{u}}^2} \Rightarrow$ a fault contained in inputs without influences on outputs is detected.

IV. Dynamic and Uncertainty Issues Under Industrial Operating Conditions

As introduced in previous sections, the basic multivariate statistical methods, like PCA and PLS, are effective to monitor the industrial processes when the processes operate under ideal stationary conditions. However, the operating conditions of industrial processes could vary according to the production requirements. In the off-line training procedure of the basic multivariate statistical methods, only the stationary behaviors are considered. Therefore, the basic multivariate statistical methods are not suitable to handle the process dynamics and operation condition changes. Although more complicated variants have been proposed to handle the process dynamics [22], [71], simple schemes with low computation cost that could reduce the design efforts are still of great interest from the industrial application viewpoint.

On the other hand, the well-established model-based fault diagnosis methods have the advantage of handling process dynamics, based on which the observer-based fault detection systems are well studied [72], [73]. However, their applications on industrial processes are often unrealistic due to the sophisticated modeling procedure based on the first principles. Recently, based on the intimate relationship between the well-established SIM and parity vectors [27], [74], [75], a new model-free data-driven design procedure is proposed in [76], whose core is a residual generator in the form of the diagnostic observer (DO) that is directly constructed using the parameters which are identified from the process measurements. Based on the new model-free data-driven design procedure, the fault detection and isolation (FDI) and the fault-tolerant control (FTC) systems could be directly constructed. The whole design procedure of the fault diagnosis systems becomes much simpler, easier, and efficient. A schematic description of the SIM-aided residual generation and the new model-free method is shown in Fig. 2.

A. On Dynamic Issue

Assume that the process under consideration could be realized into the following state-space representation:

$$x(k+1) = Ax(k) + Bu(k) + w(k) \tag{38}$$

$$y(k) = Cx(k) + Du(k) + v(k) \tag{39}$$

where $w(k) \in \mathcal{R}^n$ and $v(k) \in \mathcal{R}^m$ are assumed to be zero-mean Gaussian-distributed white noise. The system matrices A, B, C, and D and system order n are unknown *a priori*.

Construct the following Hankel structures using state variable x and output y:

$$X_i = [x(k+1) \quad x(k+2) \quad \cdots \quad x(k+N)]$$

$$Y_p = [y_s(k) \quad y_s(k+1) \quad \cdots \quad y_s(k+N-1)]$$

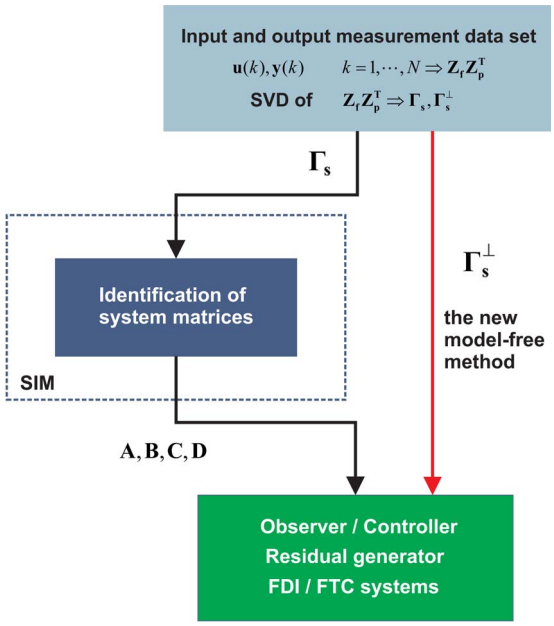


Fig. 2. Schematic description of SIM-aided residual generation and the new model-free method.

$$Y_f = [y_s(k+s) \quad \cdots \quad y_s(k+N+s-1)]$$
$$y_s(k) = \left[y^{\mathrm{T}}(k-s+1) \quad y^{\mathrm{T}}(k-s+2) \quad \cdots \quad y^{\mathrm{T}}(k)\right]^{\mathrm{T}}$$

in which N denotes the number of sampled data, the subscript p stands for the data in the "past," and f denotes the data in the "future." The integer s represents the delay which could be defined by the user, and usually, $s > n$ is chosen. The input and noise terms, i.e., U_p, U_f, V_p, V_f, W_p, and W_f, are gathered in the similar Hankel structures. Construct past and future Hankel structures as

$$Z_p = \begin{bmatrix} Y_p \\ U_p \end{bmatrix} \qquad Z_f = \begin{bmatrix} Y_f \\ U_f \end{bmatrix}.$$

From the system dynamic equations (38) and (39), we have

$$Y_f = \Gamma_s X_i + H_{u,s} U_f + H_{w,s} W_f + V_f \tag{40}$$

where Γ_s, $H_{u,s}$, and $H_{w,s}$ are quasi-standard forms. Based on the aforementioned notations, the data matrix Z_f can be formulated as

$$Z_f = \begin{bmatrix} \Gamma_s & H_{u,s} \\ 0 & I \end{bmatrix} \begin{bmatrix} X_i \\ U_f \end{bmatrix} + \begin{bmatrix} H_{w,s} W_f + V_f \\ 0 \end{bmatrix}. \tag{41}$$

In order to extract the reliable system information from the data sets Z_f and U_f, the influences of the last term in the aforementioned equation should be eliminated or compensated. For this purpose, the so-called instrumental variables can be applied, which utilize the statistical independence of the noises to compensate the noise effects. Following the assumptions, the noise in the "past" and "future" should be statistically independent. Therefore, Z_p could be chosen as an instrumental variable to eliminate the effects of noise contained in (41). To this end, we have, for large N, $\lim_{N\to\infty} (1/N)(H_{w,s} W_f + V_f) Z_p^{\mathrm{T}} = 0$. Therefore,

$$Z_f Z_p^{\mathrm{T}} \approx \begin{bmatrix} \Gamma_s & H_{u,s} \\ 0 & I \end{bmatrix} \begin{bmatrix} X_i \\ U_f \end{bmatrix} Z_p^{\mathrm{T}}. \tag{42}$$

If the following persistent excitation condition holds, i.e.,

$$\operatorname{rank}\left(\begin{bmatrix} X_i \\ U_f \end{bmatrix} Z_p^{\mathrm{T}} \right) = n + sl \tag{43}$$

then it yields

$$\operatorname{rank}\left(Z_f Z_p^{\mathrm{T}}\right) = \operatorname{rank}\left(\begin{bmatrix} \Gamma_s & H_{u,s} \\ 0 & I \end{bmatrix} \right) = sl + n \tag{44}$$

which guarantees $\Gamma_s^{\perp}$ and $\Gamma_s^{\perp} H_{u,s}$ can be successfully identified from the data sets. Recalling $\operatorname{rank}(\Gamma_s^{\perp}) = sm - n$ and (43), the following results are valid:

$$Z_f^{\perp} \begin{bmatrix} \Gamma_s & H_{u,s} \\ 0 & I \end{bmatrix} = 0 \qquad \operatorname{rank}\left(Z_f^{\perp}\right) = sm - n.$$

The model order can thus be determined based on an SVD on $(1/N) Z_f Z_p^{\mathrm{T}}$, i.e.,

$$\frac{1}{N} Z_f Z_p^{\mathrm{T}} = U_z \begin{bmatrix} \Lambda_{z,1} & 0 \\ 0 & \Lambda_{z,2} \end{bmatrix} V_z^{\mathrm{T}} \tag{45}$$

where $U_z \in \mathcal{R}^{s(l+m)\times s(l+m)}$ and $V_z \in \mathcal{R}^{s(l+m)\times s(l+m)}$. According to the condition of (43), $\Lambda_{z,1}$ is a diagonal matrix which contains $sl + n$ nonzero singular values, i.e.,

$$\operatorname{rank}(\Lambda_{z,1}) = \operatorname{rank}\left(\begin{bmatrix} \Gamma_s & H_{u,s} \\ 0 & I \end{bmatrix} \right) = sl + n \tag{46}$$

and the diagonal matrix $\Lambda_{z,2}$ has exactly $sm - n$ zero singular values. Thus, divide U_z as follows:

$$U_z = \begin{bmatrix} U_{z,11} & U_{z,12} \\ U_{z,21} & U_{z,22} \end{bmatrix}$$

where $U_{z,11} \in \mathcal{R}^{sm\times(sl+n)}$, $U_{z,12} \in \mathcal{R}^{sm\times(sm-n)}$, $U_{z,21} \in \mathcal{R}^{sl\times(sl+n)}$, and $U_{z,22} \in \mathcal{R}^{sl\times(sm-n)}$. Finally, $\Gamma_s^{\perp}$ and $\Gamma_s^{\perp} H_{u,s}$ can be calculated by

$$\Gamma_s^{\perp} = U_{z,12}^{\mathrm{T}} \qquad \Gamma_s^{\perp} H_{u,s} = -U_{z,22}^{\mathrm{T}}. \tag{47}$$

Based on the aforementioned algorithm, $\Gamma_s^{\perp}$ and $\Gamma_s^{\perp} H_{u,s}$ are directly identified from the training data sets Z_f and Z_p. According to the one-to-one relationship between the DO and the parity-vector-based residual generator [13], given

$$\upsilon_s = [\upsilon_{s,0} \quad \upsilon_{s,1} \quad \cdots \quad \upsilon_{s,s-1}] \tag{48}$$
$$\beta_s = [\beta_{s,0} \quad \beta_{s,1} \quad \cdots \quad \beta_{s,s-1}] \tag{49}$$

where $\upsilon_s \in \Gamma_s^{\perp}$, $\beta_s \in \Gamma_s^{\perp} H_{u,s}$ and $\upsilon_{s,i} \in \mathcal{R}^{1\times m}$, $\beta_{s,i} \in \mathcal{R}^{1\times l}$ for $i = 1, \ldots, s-1$, then the DO can be constructed in the following form:

$$z(k+1) = A_z z(k) + B_z u(k) + L_z y(k) \tag{50}$$
$$r(k) = g_z y(k) - c_z z(k) - d_z u(k) \tag{51}$$

where

$$A_z = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \ddots & \vdots & 0 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \qquad L_z = -\begin{bmatrix} \upsilon_{s,0} \\ \upsilon_{s,1} \\ \vdots \\ \upsilon_{s,s-2} \end{bmatrix}$$

$$B_z = \begin{bmatrix} \beta_{s,0} \\ \beta_{s,1} \\ \vdots \\ \beta_{s,s-2} \end{bmatrix} \qquad c_z = [0 \quad \cdots \quad 0 \quad 1]$$

$$g_z = \upsilon_{s,s-1} \qquad d_z = \beta_{s,s-1}.$$

Different from the well-developed SIM whose aim is to identify the whole process model, in the aforementioned procedure, only the parameters relevant to the construction of the residual generator are identified. Recent study can be found in [77] and [78] to extend the single residual generator to multiple case. The multiple residual generator delivers a better fault detection performance, which could be further used for process control and FTC purpose.

B. On Uncertainty Issue

During the industrial production, the industrial processes may operate under different operating conditions due to various production requirements or control strategies. These changes may significantly influence all process variables. In order to ensure that the predesigned process monitoring systems still function properly and effectively in this case, the online identification techniques could be applied. For this reason, the well-developed adaptive control methods, which guarantee the stability and convergence properties, have drawn lots of attention. For our purpose, to tackle the uncertainty issues, it might be advantageous to design an adaptive residual generator by combining the subspace-based data-driven design scheme and the adaptive control technique.

As mentioned in the previous section, the DO (50) and (51) could be directly constructed based on the identified vectors $\upsilon_s \in \Gamma_s^{\perp}$ and $\beta_s \in \Gamma_s^{\perp} H_{u,s}$, which implies that the parameter changes could be observed in the changes of υ_s and β_s, namely, in B_z, L_z, g_z, and d_z. Therefore, extend DO into

$$z(k+1) = A_z z(k) + B_z u(k) + L_z r(k) + L_0 r(k) \tag{52}$$

$$r(k) = g_z y(k) - c_z z(k) - d_z u(k) \tag{53}$$

where L_0 is a design parameter which should guarantee the stability of $\bar{A}_z = A_z - L_z c_z - L_0 c_z$. Let

$$\theta = \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} \in \mathcal{R}^{s(m+l)} \qquad \theta_1 = \begin{bmatrix} \mathrm{col}(B_z) \\ \mathrm{col}(L_z) \end{bmatrix}$$

$$\theta_2 = \begin{bmatrix} \mathrm{col}(g_z) \\ \mathrm{col}(d_z) \end{bmatrix}$$

$$Q\left(u(k), y(k)\right) = \begin{bmatrix} \mathcal{U}(k) & \mathcal{Y}(k) & L_0 y^{\mathrm{T}}(k) & -L_0 u^{\mathrm{T}}(k) \end{bmatrix}$$

$$\mathcal{U}(k) = \begin{bmatrix} u_1(k) \times I_{(s-1)\times(s-1)} & \cdots \\ & u_l(k) \times I_{(s-1)\times(s-1)} \end{bmatrix}$$

$$\mathcal{Y}(k) = \begin{bmatrix} y_1(k) \times I_{(s-1)\times(s-1)} & \cdots \\ & y_m(k) \times I_{(s-1)\times(s-1)} \end{bmatrix}$$

where $\mathrm{col}(\cdot)$ denotes the columnwise reordering of a matrix

$$P = [\, p_1 \quad \cdots \quad p_\alpha \,] \in \mathcal{R}^{\beta \times \alpha} \qquad \mathrm{col}(P) = \begin{bmatrix} p_1 \\ \vdots \\ p_\alpha \end{bmatrix} \in \mathcal{R}^{\beta\alpha}.$$

Therefore, (52) and (53) can be rewritten into

$$z(k+1) = \bar{A}_z z(k) + Q\left(u(k), y(k)\right)\theta \tag{54}$$

$$r(k) = g_z y(k) - c_z z(k) - d_z u(k). \tag{55}$$

The residual generator (54) and (55) are the standard form for the adaptive observer design. Moreover, the residual generator should ensure that the residual signal $r(k)$ exponentially converges to zero if the operating condition is unchanged. Based on the well-established adaptive control theory, the adaptive residual generator scheme contains the following three subsystems can be further established.

Residual Generator:

$$\hat{z}(k+1) = \bar{A}_z \hat{z}(k) + Q\left(u(k), y(k)\right)\hat{\theta}(k) + V(k+1)\left(\hat{\theta}(k+1) - \hat{\theta}(k)\right) \tag{56}$$

$$r(k) = \hat{g}_z(k) y(k) - c_z \hat{z}(k) - \hat{d}_z(k) u(k). \tag{57}$$

Auxiliary Filter:

$$V(k+1) = \bar{A}_z V(k) + Q\left(u(k), y(k)\right) \tag{58}$$

$$\varphi(k) = c_z V(k) - \begin{bmatrix} 0 & \cdots & 0 & y^{\mathrm{T}}(k) & -u^{\mathrm{T}}(k) \end{bmatrix}. \tag{59}$$

Parameter Estimator:

$$\hat{\theta}(k+1) = \gamma(k)\varphi^{\mathrm{T}}(k) r(k) + \hat{\theta}(k) \tag{60}$$

$$\gamma(k) = \frac{\mu}{\delta + \varphi(k)\varphi^{\mathrm{T}}(k)}, \qquad \delta \geq 0; \quad 0 < \mu < 2. \tag{61}$$

The proof of the stability and the convergence property of (56)–(61) can be found in [79].

During the aforementioned design procedure, instead of the identification of the complete process model, only the key parameters relevant to process monitoring are identified from process measurement data. Based on it, to solve the various operating conditions and the possible uncertain change issues during industrial production, an adaptive residual generation scheme with low online computation cost is proposed. Therefore, the addressed methods are highly suitable to monitor industrial large-scale processes while the dynamic and uncertainty issues could be handled.

V. Conclusion and Discussion

Due to the uncertainties and complexities of modern industrial systems, it is very difficult, or impossible, to construct

process models based on the first principles in practice. Therefore, the most popular data-driven process monitoring schemes, i.e., PCA and PLS, within the multivariate statistical framework are briefly introduced due to their simplicities and excellent ability to handle large amounts of data. Necessary modifications are further reviewed under stationary operating conditions to ensure better performance and a successful implementation on industrial applications. For the dynamic and uncertainty issues in the industrial processes, the subspace aided data-driven design scheme is reviewed. Moreover, an adaptive residual generation is integrated in the subspace-based data-driven design framework to further handle the uncertainty issues. Compared with the standard system identification techniques, the online computation cost is significantly reduced, and the design procedure is simple yet efficient and suitable for large-scale industrial applications.

In contrast to the well-developed model-based process monitoring design framework, the data-driven design framework is still in its embryonic form.

1) To improve its feasibility and effectiveness in the industrial environment, more comprehensive and detailed studies are still needed to solve the nonlinear and non-Gaussian problems for large-scale industrial processes.
2) Since the incompleteness of data that is usually in terms of missing or erroneous values is quite common in large-scale industrial processes, the resulting computations, i.e., projections, estimations, and predictions, might be inaccurate. It is still a challenging topic in the data-driven design framework to deal with incomplete data.
3) It is natural and engineering way to realize an integrative design of FDI system and FTC system, in which appropriate FTC strategies should be performed when a process fault has been successfully detected. Moreover, the lifetime management system should be built in order to indicate the status of the process, in which the propagation of the process faults will be monitored and analyzed.

REFERENCES

[1] D. U. Campos-Delgado and D. R. Espinoza-Trejo, "An observer-based diagnosis scheme for single and simultaneous open-switch faults in induction motor drives," *IEEE Trans. Ind. Electron.*, vol. 58, no. 2, pp. 671–679, Feb. 2011.

[2] A. Akrad, M. Hilairet, and D. Diallo, "Design of a fault-tolerant controller based on observers for a PMSM drive," *IEEE Trans. Ind. Electron.*, vol. 58, no. 4, pp. 1416–1427, Apr. 2011.

[3] T. A. Najafabadi, F. R. Salmasi, and P. Jabehdar-Maralani, "Detection and isolation of speed-, DC-link voltage-, and current-sensor faults based on an adaptive observer in induction-motor drives," *IEEE Trans. Ind. Electron.*, vol. 58, no. 5, pp. 1662–1672, May 2011.

[4] L. M. Capisani, A. Ferrara, A. Ferreira de Loza, and L. M. Fridman, "Manipulator fault diagnosis via higher order sliding-mode observers," *IEEE Trans. Ind. Electron.*, vol. 59, no. 10, pp. 3979–3986, Oct. 2012.

[5] S. Huang, K. K. Tan, and T. H. Lee, "Fault diagnosis and fault-tolerant control in linear drives using the Kalman filter," *IEEE Trans. Ind. Electron.*, vol. 59, no. 11, pp. 4285–4292, Nov. 2012.

[6] D. He, R. Li, and J. Zhu, "Plastic bearing fault diagnosis based on a two-step data mining approach," *IEEE Trans. Ind. Electron.*, vol. 60, no. 8, pp. 3429–3440, Aug. 2013.

[7] Y. Gritli, L. Zarri, C. Rossi, F. Filippetti, G. Capolino, and D. Casadei, "Advanced diagnosis of electrical faults in wound-rotor induction machines," *IEEE Trans. Ind. Electron.*, vol. 60, no. 9, pp. 4012–4024, Sep. 2013.

[8] J. Seshadrinath, B. Singh, and B. K. Panigrahi, "Vibration analysis based interturn fault diagnosis in induction machines," *IEEE Trans. Ind. Informat.*, vol. 10, no. 1, pp. 340–350, Feb. 2014.

[9] J. Gertler, *Fault Detection and Diagnosis in Engineering Systems*. New York, NY, USA: Marcel Dekker, 1998.

[10] R. Patton, P. Frank, and R. Clark, *Issues of Fault Diagnosis for Dynamic Systems*. Berlin, Germany: Springer-Verlag, 2000.

[11] M. Blanke, M. Kinnaert, J. Lunze, and M. Staroswiecki, *Diagnosis and Fault-Tolerant Control*. New York, NY, USA: Springer, 2003.

[12] R. Isermann, *Fault Diagnosis Systems*. New York, NY, USA: Springer-Verlag, 2006.

[13] S. X. Ding, *Model-Based Fault Diagnosis Techniques*. Berlin, Germany: Springer-Verlag, 2008.

[14] H. Li, J. Yu, C. Hilton, and H. Liu, "Adaptive sliding mode control for nonlinear active suspension vehicle systems using T–S fuzzy approach," *IEEE Trans. Ind. Electron.*, vol. 60, no. 8, pp. 3328–3338, Aug. 2013.

[15] V. Venkatasubramanian, R. Rengaswamy, K. Yin, and S. N. Kavuri, "Review of process fault detection and diagnosis. Part I: Quantitative model-based methods," *Comput. Chem. Eng.*, vol. 27, no. 3, pp. 293–311, Mar. 2003.

[16] Z. Gao, X. Dai, T. Breikin, and H. Wang, "Novel parameter identification by using a high-gain observer with application to a gas turbine engine," *IEEE Trans. Ind. Informat.*, vol. 4, no. 4, pp. 271–279, Nov. 2008.

[17] B. Shen, Z. Wang, H. Shu, and G. Wei, "On nonlinear H_∞ filtering for discrete-time stochastic systems with missing measurements," *IEEE Trans. Autom. Control*, vol. 53, no. 9, pp. 2170–2180, Oct. 2008.

[18] P. Zhang and S. X. Ding, "On fault detection in linear discrete-time, periodic, sampled-data systems (survey)," *J. Control Sci. Eng.*, vol. 2008, pp. 1–19, Jan. 2008.

[19] I. Hwang, S. Kim, Y. Kim, and C. E. Seah, "A survey of fault detection, isolation, reconfiguration methods," *IEEE Trans. Control Syst. Technol.*, vol. 18, no. 3, pp. 636–653, Mar. 2010.

[20] H. Dong, Z. Wang, and H. Gao, "Distributed H_∞ filtering for a class of Markovian jump nonlinear time-delay systems over lossy sensor networks," *IEEE Trans. Ind. Electron.*, vol. 60, no. 10, pp. 4665–4672, Oct. 2013.

[21] M. Ma, D. Wong, S. Jang, and S. Tseng, "Fault detection based on statistical multivariate analysis and microarray visualization," *IEEE Trans. Ind. Informat.*, vol. 6, no. 1, pp. 18–24, Feb. 2010.

[22] Y. Yao, T. Chen, and F. Gao, "Multivariate statistical monitoring of two-dimensional dynamic batch processes utilizing non-Gaussian information," *J. Process Control*, vol. 20, no. 10, pp. 1188–1197, Dec. 2010.

[23] R. Gong, S. Huang, and T. Chen, "Robust and efficient rule extraction through data summarization and its application in welding fault diagnosis," *IEEE Trans. Ind. Informat.*, vol. 4, no. 3, pp. 198–206, Aug. 2008.

[24] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometr. Intell. Lab. Syst.*, vol. 2, no. 1–3, pp. 37–52, Aug. 1987.

[25] I. Jolliffe, *Principal Component Analysis*, 2nd ed. New York, NY, USA: Springer-Verlag, 2002.

[26] S. J. Qin, "Statistical process monitoring: Basics and beyond," *J. Chemometr.*, vol. 17, no. 8/9, pp. 480–502, Aug./Sep. 2003.

[27] J. Gertler and J. Cao, "PCA-based fault diagnosis in the presence of control and dynamics," *AIChE J.*, vol. 50, no. 2, pp. 388–402, Feb. 2004.

[28] V. Choqueuse, M. E. H. Benbouzid, Y. Amirat, and S. Turri, "Diagnosis of three-phase electrical machines using multidimensional demodulation techniques," *IEEE Trans. Ind. Electron.*, vol. 59, no. 4, pp. 2014–2023, Apr. 2012.

[29] C. M. Vong, P. K. Wong, and W. F. Ip, "A new framework of simultaneous-fault diagnosis using pairwise probabilistic multi-label classification for time-dependent patterns," *IEEE Trans. Ind. Electron.*, vol. 60, no. 8, pp. 3372–3385, Aug. 2013.

[30] G. Li, S. J. Qin, and D. Zhou, "Geometric properties of partial least squares for process monitoring," *Automatica*, vol. 46, no. 1, pp. 204–210, Jan. 2010.

[31] Y. Zhang, H. Zhou, S. J. Qin, and T. Chai, "Decentralized fault diagnosis of large-scale processes using multiblock kernel partial least squares," *IEEE Trans. Ind. Informat.*, vol. 6, no. 1, pp. 3–10, Feb. 2010.

[32] D. Wang, J. Liu, and R. Srinivasan, "Data-driven soft sensor approach for quality prediction in a refining process," *IEEE Trans. Ind. Informat.*, vol. 6, no. 1, pp. 11–17, Feb. 2010.

[33] R. Muradore and P. Fiorini, "A PLS-based statistical approach for fault detection and isolation of robotic manipulators," *IEEE Trans. Ind. Electron.*, vol. 59, no. 8, pp. 3167–3175, Aug. 2012.

[34] X. He, Z. Wang, Y. Liu, and D. Zhou, "Least-squares fault detection and diagnosis for networked sensing systems using direct state estimation

approach," *IEEE Trans. Ind. Informat.*, vol. 9, no. 3, pp. 1670–1679, Aug. 2013.
[35] E. Russell, L. Chiang, and R. Braatz, *Data-Driven Methods for Fault Detection and Diagnosis in Chemical Processes*. London, U.K.: Springer-Verlag, 2000.
[36] L. Chiang, E. Russell, and R. Braatz, *Fault Detection and Diagnosis in Industrial Systems*. London, U.K.: Springer-Verlag, 2001.
[37] Y. Zhang and M. S. Dudzic, "Online monitoring of steel casting processes using multivariate statistical technologies: From continuous to transitional operations," *J. Process Control*, vol. 16, no. 8, pp. 819–829, Sep. 2006.
[38] S. J. Qin, "Data-driven fault detection and diagnosis for complex industrial processes," in *Proc. IFAC SAFEPROCESS Symp.*, Jul. 2009, pp. 1115–1125.
[39] S. J. Qin, "Survey on data-driven industrial process monitoring and diagnosis," *Annu. Rev. Control*, vol. 36, no. 2, pp. 220–234, Dec. 2012.
[40] Z. Gao and X. Dai, "From model, signal to knowledge: A data-driven perspective of fault detection and diagnosis," *IEEE Trans. Ind. Informat.*, vol. 9, no. 4, pp. 2226–2238, Nov. 2013.
[41] D. Zhou, G. Li, and S. J. Qin, "Total projection to latent structures for process monitoring," *AIChE J.*, vol. 56, no. 1, pp. 168–178, Jan. 2010.
[42] S. Yin, S. X. Ding, P. Zhang, A. Haghani, and A. Naik, "Study on modifications of PLS approach for process monitoring," in *Proc. 18th IFAC World Congr.*, Milan, Italy, 2011, pp. 12 389–12 394.
[43] 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.
[44] W. Ku, R. H. Storer, and C. Georgakis, "Disturbance detection and isolation by dynamic principal component analysis," *Chemometr. Intell. Lab. Syst.*, vol. 30, no. 1, pp. 179–196, Nov. 1995.
[45] R. Luo, M. Misra, and D. M. Himmelblau, "Sensor fault detection via multiscale analysis and dynamic PCA," *Ind. Eng. Chem. Res.*, vol. 38, no. 4, pp. 1489–1495, Apr. 1999.
[46] E. L. Russell, L. H. Chiang, and R. D. Braatz, "Fault detection in industrial processes using canonical variate analysis and dynamic principal component analysis," *Chemometr. Intell. Lab. Syst.*, vol. 51, no. 1, pp. 81–93, May 2000.
[47] J. Mina and C. Verde, "Fault detection using dynamic principal component analysis by average estimation," in *Proc. 2nd Int. Conf. Elect. Electron. Eng.*, Mexico City, Mexico, Sep. 2005, pp. 374–377.
[48] S. W. Choi and I. B. Lee, "Multiblock PLS-based localized process diagnosis," *J. Process Control*, vol. 15, no. 3, pp. 295–306, Apr. 2005.
[49] C. F. Alcala and S. J. Qin, "Reconstruction-based contribution for process monitoring with kernel principal component analysis," *Ind. Eng. Chem. Res.*, vol. 49, no. 17, pp. 7849–7857, Sep. 2010.
[50] J. H. Cho, J. M. Lee, S. W. Choi, D. Lee, and I. B. Lee, "Fault identification for process monitoring using kernel principal component analysis," *Chem. Eng. Sci.*, vol. 60, no. 1, pp. 279–288, Jan. 2005.
[51] R. Rosipal and L. Trejo, "Kernel partial least squares regression in reproducing kernel Hilbert space," *J. Mach. Learn. Res.*, vol. 2, pp. 97–123, Dec. 2001.
[52] A. J. Willis, "Condition monitoring of centrifuge vibrations using kernel PLS," *Comput. Chem. Eng.*, vol. 34, no. 3, pp. 349–353, Mar. 2010.
[53] L. B. Jack and A. K. Nandi, "Support vector machines for detection and characterization of rolling element bearing faults," *Proc. Inst. Mech. Eng. C, J. Mech. Eng. Sci.*, vol. 215, no. 9, pp. 1065–1074, Sep. 2001.
[54] C. Gao, L. Jian, and S. Luo, "Modeling of the thermal state change of blast furnace hearth with support vector machines," *IEEE Trans. Ind. Electron.*, vol. 59, no. 2, pp. 1134–1145, Feb. 2012.
[55] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. Hoboken, NJ, USA: Wiley, 2001.
[56] G. J. McLachlan, *Discriminant Analysis and Statistical Pattern Recognition*. Hoboken, NJ, USA: Wiley, 2004.
[57] L. H. Chiang, E. L. Russell, and R. D. Braatz, "Fault diagnosis in chemical processes using Fisher discriminant analysis, discriminant partial least squares, principal component analysis," *Chemometr. Intell. Lab. Syst.*, vol. 50, no. 2, pp. 243–252, Mar. 2000.
[58] L. H. Chiang, M. E. Kotanchek, and A. K. Kordon, "Fault diagnosis based on Fisher discriminant analysis and support vector machines," *Comput. Chem. Eng.*, vol. 28, no. 8, pp. 1389–1401, Jul. 2004.
[59] Q. P. He, S. J. Qin, and J. Wang, "A new fault diagnosis method using fault directions in Fisher discriminant analysis," *AIChE J.*, vol. 51, no. 2, pp. 555–571, Feb. 2005.
[60] M. Girolami, *Self-Organising Neural Networks: Independent Component Analysis and Blind Source Separation*. London, U.K.: Springer-Verlag, 1999.
[61] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. Hoboken, NJ, USA: Wiley, 2001.
[62] T. Lee, *Independent Component Analysis: Theory and Applications*. Boston, MA, USA: Kluwer, 1998.
[63] M. Kano, S. Tanaka, S. Hasebe, and I. Hashimoto, "Monitoring independent components for fault detection," *AIChE J.*, vol. 49, no. 4, pp. 969–976, Apr. 2003.
[64] J. M. Lee, C. Yoo, and I. B. Lee, "Statistical process monitoring with independent component analysis," *J. Process Control*, vol. 14, no. 5, pp. 467–485, Aug. 2004.
[65] J.-M. Lee, S. J. Qin, and I.-B. Lee, "Fault detection and diagnosis based on modified independent component analysis," *AIChE J.*, vol. 52, no. 10, pp. 3501–3514, Oct. 2006.
[66] W. E. Larimore, "System identification reduced-order filtering and modeling via canonical variate analysis," in *Proc. Amer. Control Conf.*, 1983, pp. 445–451.
[67] W. E. Larimore, "Canonical variate analysis for system identification, filtering, adaptive control," in *Proc. IEEE Conf. Decision Control*, 1990, pp. 596–604.
[68] Y. Wang, D. Seborg, and W. Larimore, "Process monitoring based on canonical variate analysis," in *Proc. ADCHEM*, Banff, AB, Canada, 1997, pp. 523–528.
[69] P.-E. P. Odiowei and Y. Cao, "Nonlinear dynamic process monitoring using canonical variate analysis and kernel density estimations," *IEEE Trans. Ind. Informat.*, vol. 6, no. 1, pp. 36–45, Feb. 2010.
[70] S. X. Ding, P. Zhang, E. L. Ding, S. Yin, A. S. Naik, P. Deng, and W. Gui, "On the application of PCA technique to fault diagnosis," *Tsinghua Sci. Technol.*, vol. 15, no. 2, pp. 138–144, Apr. 2010.
[71] M. Jia, F. Chu, F. Wang, and W. Wang, "On-line batch process monitoring using batch dynamic kernel principal component analysis," *Chemometr. Intell. Lab. Syst.*, vol. 101, no. 2, pp. 110–122, Apr. 2010.
[72] V. Krishnaswami and G. Rizzoni, "A survey of observer based residual generation for FDI," in *Proc. IFAC Symp. Fault Detection, Supervision Safety Tech. Process.*, Oxford, U.K., 1994, pp. 35–40.
[73] P. M. Frank and S. X. Ding, "Survey of robust residual generation and evaluation methods in observer-based fault detection systems," *J. Process Control*, vol. 7, no. 6, pp. 403–424, Dec. 1997.
[74] J. Wang and S. J. Qin, "A new subspace identification approach based on principal component analysis," *J. Process Control*, vol. 12, no. 8, pp. 841–855, Dec. 2002.
[75] J. Wang and S. J. Qin, "Closed-loop subspace identification using the parity space," *Automatica*, vol. 42, no. 2, pp. 315–320, Feb. 2006.
[76] S. X. Ding, P. Zhang, A. Naik, E. Ding, and B. Huang, "Subspace method aided data-driven design of fault detection and isolation systems," *J. Process Control*, vol. 19, no. 9, pp. 1496–1510, Oct. 2009.
[77] S. X. Ding, Y. Shen, Y. Wang, Y. Wang, Y. Yang, and B. Ni, "Data-driven design of observers and its applications," in *Proc. 18th IFAC World Congr.*, Milan, Italy, 2011, pp. 11 441–11 446.
[78] S. X. Ding, Y. Wang, P. Zhang, Y. Yang, and E. L. Ding, "Data-driven design of fault-tolerant control systems," in *Proc. 8th IFAC Symp. Fault Detection, Supervision Safety Tech. Process.*, Mexico City, Mexico, 2012, pp. 1323–1328.
[79] S. X. Ding, S. Yin, P. Zhang, E. L. Ding, and A. Naik, "An approach to data-driven adaptive residual generator design and implementation," in *Proc. 7th IFAC Symp. Fault Detection, Supervision Safety Tech. Process.*, Barcelona, Spain, 2009, pp. 941–946.

Shen Yin (M'12) received the B.E. degree in automation from Harbin Institute of Technology, Harbin, China, in 2004 and the M.Sc. degree in control and information system and the Ph.D. degree in electrical engineering and information technology from the University of Duisburg–Essen, Duisburg, Germany.

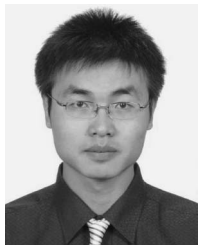
He is currently with the Research Center of Intelligent Control and Systems, Harbin Institute of Technology. His research interests are model-based and data-driven fault diagnosis, fault-tolerant control, and big data focused on industrial electronics applications.

Steven X. Ding received the Ph.D. degree in electrical engineering from the Gerhard-Mercator University of Duisburg, Duisburg, Germany, in 1992.

From 1992 to 1994, he was an R&D Engineer with Rheinmetall GmbH. From 1995 to 2001, he was a Professor of control engineering with the University of Applied Science Lausitz, Senftenberg, Germany, where he served as the Vice-President during 1998–2000. Since 2001, he has been a Professor of control engineering and the Head of the Institute for Automatic Control and Complex Systems (AKS), University of Duisburg–Essen, Duisburg. His research interests are model-based and data-driven fault diagnosis, fault-tolerant systems, and their applications in industry with a focus on automotive systems and chemical processes.

Xiaochen Xie (S'13) received the B.E. degree in automation from Harbin Institute of Technology, Harbin, China, in 2012, where she is currently working toward the M.Sc. degree in control science and engineering in the Research Institute of Intelligent Control and Systems.

Her research interests are robust control, data-driven techniques, delta operator approach, and networked control systems.



Hao Luo received the B.E. degree in electrical engineering from Xi'an Jiaotong University, Xi'an, China, in 2007 and the M.Sc. degree in electrical engineering and information technology from the University of Duisburg–Essen, Duisburg, Germany, in 2012, where he is currently working toward the Ph.D. degree in the Institute for Automatic Control and Complex Systems (AKS).

His research interests include model-based and data-driven fault diagnosis, fault-tolerant systems, and their applications in industrial systems.