IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received November 20, 2018, accepted December 17, 2018, date of publication January 4, 2019, date of current version January 23, 2019.

Digital Object Identifier 10.1109/ACCESS.2018.2890301

Quality Weakly Related Fault Detection Based on Weighted Dual-Step Feature Extraction

BO ZHAO, HONGBO SHI, BING SONG, AND SHUAI TAN
Key Laboratory of Advanced Control and Optimization for Chemical Processes, Ministry of Education, East China University of Science and Technology, Shanghai 200237, China

Corresponding author: Hongbo Shi (hbshi@ecust.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61374140 and Grant 61673173, in part by the Fundamental Research Funds for the Central Universities under Grant 222201714031, and in part by the China Postdoctoral Research Fund under Grant 2017M611472.

ABSTRACT Process variables can be separated into quality strongly related variables (QSRVs) and quality weakly related variables (QWRVs) based on the difference of correlation between process variables and quality. When the fault occurs on different types of process variables, the degree of impact on quality is different. The fault which occurs on QWRV is defined as the quality weakly related fault. This paper presents a weighted dual-step feature extraction (WDSFE) method for quality weakly related fault detection. In the first step, a QSRV block and a QWRV block are separated from original process variables. In the second step, the two variable blocks are further decomposed into a quality-direct-related subspace (QDRS), a quality-potential-related subspace (QPRS), and two quality unrelated subspaces (QUSs). Subsequently, the weighted T^2 statistic is established based on the statistics in both QDRS and QPRS to detect the quality-related fault. The *BIC* statistic is established based on the statistics in two QUSs to detect the quality unrelated fault. Finally, a numerical example and a simulated industrial process are used to illustrate the effectiveness of WDSFE.

INDEX TERMS Process monitoring, fault detection, feature extraction, quality, statistical analysis.

I. INTRODUCTION

Some process variables are strongly related to quality variables and are defined as quality strongly related variables (QSRVs). The remaining process variables which are weakly related to quality variables are defined as quality weakly related variables (QWRVs). When the fault occurs on QSRV, the effect on quality is significant. However, when the fault occurs on QWRV, the impact on quality may be inconspicuous. Such fault is defined as quality weakly related fault (QWRF). If the amplitude of the QWRF is small and the impact on quality is acceptable, then the QWRF is regarded as quality unrelated fault. On the contrary, if the amplitude of the QWRF is large enough to cause significant anomalies of quality, then the QWRF is considered as quality related fault. Therefore, QWRF detection is a challenge due to the insensitivity of QWRF and the diversity of results.

In order to monitor the abnormality of the process state in time, the process monitoring method based on multivariate statistical analysis has been extensively studied [1]–[6]. Rapidly, research on process monitoring methods has made remarkable achievements [7]–[10].With further research, some scholars have found that not all faults can affect product quality [11]–[15]. If only the quality abnormality need to be monitored, some process monitoring methods like principal component analysis (PCA) would raise massive false alarms, which leads to unnecessary downtime and maintenance [10], [16], [17]. Therefore, quality monitoring method has become one of the research hotspots [18]–[23].

Given that product quality can not be measured online, it is necessary to establish models between process variables and quality [24]–[26]. Some recent methods merely focus on several postprocessing techniques under supervision [27]–[29]. Typically, partial least squares (PLS) is considered as an oblique projection technique. Li *et al.* [30] first proposed a geometric interpretation of PLS. The contribution of Li provides a theoretical basis for some PLS-based methods. Zhou *et al.* [31] argued that the changes in the score matrix and the residual matrix of the standard PLS could make influence on the output. Hence, Zhou proposed the total PLS (TPLS) to furtherly decompose the score matrix and the residual matrix of the standard PLS. Inspired by TPLS, Qin and Zheng [32] gave a concurrent method called concurrent

PLS (CPLS) for the decomposition of process variable matrix and quality matrix. Based on PLS and singular value decomposition (SVD), Yin *et al.* [33] established a quality monitoring method. In addition, Wang *et al.* [34] and Peng *et al.* [35] proposed principal component regression (PCR)-based quality monitoring methods in the past few years. These methods mainly focused on the subsequent decomposition process, but ignored the effect of different process variables on the quality. It may cause the information of the quality weakly related variables to be submerged by the quality strongly related variables. Some other methods only consider the classification of process variables, and then the classified process variables are monitored based on an unsupervised process monitoring methods. A typical example is the mutual information-kernel principal component analysis (MI-KPCA) proposed by Huang and Yan [36]. However, some quality weakly related process variables may be classified into the quality unrelated variable block. In this case, when the fault occurs in these variables, MI-KPCA may obtain the wrong monitoring results. Thus, an effective quality monitoring method should consider both the classification of process variables and orthogonal decomposition under supervision.

In addition, compared to the linear process, the study of nonlinear quality related process monitoring methods is still limited. By extending TPLS to nonlinear processes, Peng *et al.* [37] first provided total KPLS (TKPLS) for nonlinear quality monitoring. Shortly afterwards, Zhang *et al.* [38] extended CPLS to nonlinear situations and proposed KCPLS for nonlinear quality monitoring. Recently, Jia and Zhang [39] proposed a new nonlinear monitoring method based on KPLS and SVD. Subsequently, kernel direct decomposition (KDD) and kernel least square (KLS) were put forward by Wang *et al.* [40], [41].

However, in multi-subspace process monitoring, excessive statistics can lead to difficulties in the interpretation of monitoring results. An easy way is to integrate the statistics. The commonly adopted methods for integration are: direct summation (DS) and Bayesian integration (BI). Different quality related subspaces have different degrees of influence on quality, and this difference should be expressed in order to obtain better monitoring performance. However, direct summation cannot reflect the importance of different subspaces. Although the Bayesian integration is based on the sum of probability weights, it cannot reflect the degree of correlation between different subspaces and quality. In order to achieve the integration of quality related subspace statistics, the weight T^2 statistics based on the degree of correlation with quality are established. Quality unrelated subspace do not need to reflect the impact on quality, thus Bayesian integration can achieve satisfactory results in the quality unrelated subspaces.

Thus, this paper proposes a weighted dual-step feature extraction (WDSFE) method for quality weakly related fault detection. The dual-step feature extraction can explain the effect of the weakly related variables on the quality well. In the first step, the process variables are divided into two variable blocks based on maximum information coefficient (MIC), which can avoid the information on the quality weakly related variables being submerged. This is because the quality weakly related variables have a lower impact on quality than quality strongly related variables. In the second step, a supervised orthogonal decomposition method based on KLS is adopted to furtherly extract the quality feature in the two variable blocks. If and only if the supervised orthogonal decomposition method is employed in the second step, two mutually orthogonal subspaces can be obtained. As a result, one subspace is related to quality and the other subspace is orthogonal to quality. The dual-step feature extraction method improves the monitoring sensitivity of the quality weakly related variables. However, sometimes the amplitude of the fault occurring on the quality weakly related variables is small, and it is difficult to cause a significant anomaly of the quality. In this case, the fault is quality unrelated, but the false alarm rate will increase as the monitoring sensitivity of the quality weakly related variable is raised. Therefore, the correlation between the subspaces and quality should be considered for the statistic integration of quality related subspaces. Moreover, because of the increase in the numbers of subspaces, it may have trouble in obtaining direct monitoring results. Thus, a weight T^2 statistic is proposed for the statistic integration of quality related subspaces. Given that there is no need to consider the correlation between the quality unrelated subspaces and quality, the *BIC* statistic based on Bayesian integration is applied to integrate the statistics of two quality unrelated subspaces. Compared with the methods that only consider the classification of process variables or the supervised postprocessing techniques, the advantages of the proposed method are shown by diagrammatic sketch in section III. Meanwhile, the merits of the weight T^2 statistic based on the quality correlation are verified in section IV. Finally, the proposed nonlinear quality monitoring method can obtain better monitoring performance in a numerical example and the TE process.

The contributions of this paper are listed as follows:

1. A weighted dual-step feature extraction method is proposed for quality weakly related fault detection.

2. The dual-step feature extraction method improves the monitoring sensitivity of the quality weakly related feature.

3. The weighted T^2 statistic is proposed to improve the accuracy of quality weakly related fault detection.

II. PRELIMINARIES

Quality monitoring based on KLS is briefly introduced in this section. The process variables matrix and quality variables matrix is recorded as follows:

$$\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_m] \in \mathbb{R}^{n\times m} \tag{1}$$

$$\boldsymbol{Y} = [\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_l] \in \mathbb{R}^{n\times l} \tag{2}$$

where n is the number of samples, m is the number of process variables and l is the number of quality variables.

Given a nonlinear projection function, the samples of $\boldsymbol{X}$ are projected into a high-dimensional feature space.

$$\boldsymbol{x}_k \in \boldsymbol{R}^A \rightarrow \boldsymbol{\Phi}(\boldsymbol{x}_k) \in \mathbb{R}^{\Omega} \tag{3}$$

Then $\boldsymbol{X}$ is turned into feature matrix $\boldsymbol{\Phi}$.

$$\boldsymbol{\Phi} = [\boldsymbol{\Phi}(\boldsymbol{x}_1), \boldsymbol{\Phi}(\boldsymbol{x}_2), \cdots, \boldsymbol{\Phi}(\boldsymbol{x}_n)]^T \in \mathbb{R}^{n\times\Omega} \tag{4}$$

Establishing a least square model, it induces:

$$\boldsymbol{Y} = \boldsymbol{\Phi M} + \boldsymbol{E} \tag{5}$$

where $\boldsymbol{M}$ is regression coefficient matrix. Then it holds that

$$\frac{1}{n}\boldsymbol{Y}^T\boldsymbol{\Phi} = \frac{1}{n}\boldsymbol{M}^T\boldsymbol{\Phi}^T\boldsymbol{\Phi} + \frac{1}{n}\boldsymbol{E}^T\boldsymbol{\Phi} \approx \frac{1}{n}\boldsymbol{M}^T\boldsymbol{\Phi}^T\boldsymbol{\Phi} \tag{6}$$

Thereafter $\boldsymbol{M}$ can be calculated as follows.

$$\boldsymbol{M} = \left(\boldsymbol{\Phi}^T\boldsymbol{\Phi}\right)^{\dagger}\boldsymbol{\Phi}^T\boldsymbol{Y} \tag{7}$$

where $\left(\boldsymbol{\Phi}^T\boldsymbol{\Phi}\right)^{\dagger}$is the pseudoinverse of$\left(\boldsymbol{\Phi}^T\boldsymbol{\Phi}\right)$.

The value cannot be calculated directly due to the dimension Ω of $\boldsymbol{\Phi}$ is arbitrarily large or even infinite. The detailed calculation process is introduced in [40]. Thereafter $\boldsymbol{\Phi}$ can be divided into a quality related subspace $\hat{\boldsymbol{\Phi}}$ and a quality unrelated subspace $\tilde{\boldsymbol{\Phi}}$. It can be expressed as:

$$\boldsymbol{Y} = \boldsymbol{\Phi M} = \left(\hat{\boldsymbol{\Phi}} + \tilde{\boldsymbol{\Phi}}\right)\boldsymbol{M} = \hat{\boldsymbol{\Phi}}M \tag{8}$$

The above aim can be implemented by the following steps. Eigenvalue decomposition of $\boldsymbol{MM}^T$is given as:

$$\boldsymbol{MM}^T = \begin{bmatrix}\boldsymbol{V}_1 & \boldsymbol{V}_2\end{bmatrix}\begin{bmatrix}\boldsymbol{\Lambda} & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}\boldsymbol{V}_1^T\\ \boldsymbol{V}_2\end{bmatrix} = \boldsymbol{V}_1\boldsymbol{\Lambda}\boldsymbol{V}_1^T \tag{9}$$

where$\boldsymbol{V}_1 = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{pc}]$, $\boldsymbol{V}_2 = [\boldsymbol{v}_{pc+1}, \boldsymbol{v}_{pc+2}, \ldots, \boldsymbol{v}_n]$, $\boldsymbol{\Lambda} = \mathrm{diag}\left(\lambda_1, \lambda_2, \ldots, \lambda_{pc}\right)$, and pc is the number of nonzero eigenvalues. Projecting $\boldsymbol{\Phi}$ onto $\boldsymbol{V}_1\boldsymbol{V}_1^T$ and $\boldsymbol{V}_2\boldsymbol{V}_2^T$,

$$\hat{\boldsymbol{\Phi}} = \boldsymbol{\Phi}\boldsymbol{V}_1\boldsymbol{V}_1^T \tag{10}$$

$$\tilde{\boldsymbol{\Phi}} = \boldsymbol{\Phi}\boldsymbol{V}_2\boldsymbol{V}_2^T \tag{11}$$

According to the nature of the eigenvalue decomposition, it holds that

$$\boldsymbol{V}_1\boldsymbol{V}_1^T + \boldsymbol{V}_2\boldsymbol{V}_2^T = \boldsymbol{I} \tag{12}$$

$$\boldsymbol{V}_1^T\boldsymbol{V}_2 = 0 \tag{13}$$

$$\boldsymbol{V}_2^T\boldsymbol{M} = 0 \tag{14}$$

It is obvious that

$$\hat{\boldsymbol{\Phi}} + \tilde{\boldsymbol{\Phi}} = \boldsymbol{\Phi}\left(\boldsymbol{V}_1\boldsymbol{V}_1^T + \boldsymbol{V}_2\boldsymbol{V}_2^T\right) = \boldsymbol{\Phi} \tag{15}$$

$$\hat{\boldsymbol{\Phi}}\tilde{\boldsymbol{\Phi}}^T = \boldsymbol{\Phi}\boldsymbol{V}_1\boldsymbol{V}_1^T\boldsymbol{V}_2\boldsymbol{V}_2^T\boldsymbol{\Phi}^T = 0 \tag{16}$$

$$\boldsymbol{Y} = \left(\hat{\boldsymbol{\Phi}} + s\tilde{\boldsymbol{\Phi}}\right)\boldsymbol{M} = \boldsymbol{\Phi}\left(\boldsymbol{V}_1\boldsymbol{V}_1^T + \boldsymbol{V}_2\boldsymbol{V}_2^T\right)\boldsymbol{M} = \hat{\boldsymbol{\Phi}}M \tag{17}$$

Obviously, $\hat{\boldsymbol{\Phi}}$ is orthogonal to $\tilde{\boldsymbol{\Phi}}$ and $\boldsymbol{Y}$ is only related to $\hat{\boldsymbol{\Phi}}$.

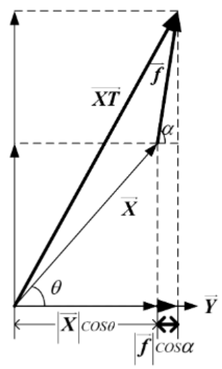


FIGURE 1. The schematic diagram of monitoring result with supervised orthogonal decomposition methods.

III. WEIGHTED DUAL-STEP FEATURE EXTRACTION

A. THE CHALLENGE OF QUALITY WEAKLY RELATED FAULT

Process variables are separated into quality strongly related variables, quality weakly related variables and quality unrelated variables based on correlation with quality, where quality weakly related variables have the low impact on the quality. When the amplitude of the fault occurring on quality weakly related variables is very large, the quality can fluctuate significantly. Some latest methods merely focus on postprocessing techniques on all process variables. It may result in the information in quality weakly related variables being submerged by quality strongly related variables, as the quality weakly related variables have a lower impact on quality than the quality strongly related variables. The geometric interpretation of monitoring result with supervised orthogonal decomposition methods for quality weakly fault detection can be shown in Fig. 1.

In Fig.1, $\overrightarrow{X}$ represents the process variables, and $\overrightarrow{Y}$ represents the quality variables. $\overrightarrow{X}$ are decomposed into two parts by a supervised orthogonal decomposition method like KLS. One part is in the same direction as $\overrightarrow{Y}$, and the other is orthogonal to $\overrightarrow{Y}$. The fault in Fig. 1 occurs on quality weakly related variables (the direction of $\overrightarrow{f}$ is the same as $\overrightarrow{Xw}$ in Fig. 2). The effect of the fault on the quality can be reflected as the projection on the quality direction, that is, $|\overrightarrow{f}|\cos\alpha$. The quality related part of the process variable is the projection of $\overrightarrow{X}$ on $\overrightarrow{Y}$, that is, $|\overrightarrow{X}|\cos\theta$. Due to the influence of the fault, the change rate of the quality related part is $|\overrightarrow{f}|\cos\alpha/|\overrightarrow{X}|\cos\theta$. As shown in Fig. 1, $|\overrightarrow{f}|\cos\alpha/|\overrightarrow{X}|\cos\theta$ is a small value. Obviously, it gives rise to the information in quality weakly related variables being submerged by quality strongly related variables, which presents a challenge for quality monitoring based on multivariate statistical analysis. Therefore, it is necessary to separate quality strongly and weakly related variables.

Furthermore, some latest methods only consider the classification of process variables. Thereafter an unsupervised monitoring method is performed on the classified process variables, such as KPCA. Since the quality weakly related variables may be misclassified, which can cause that the

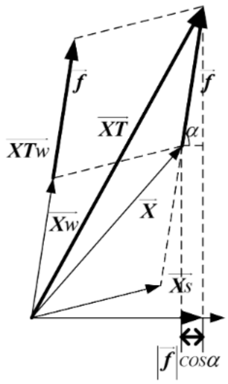


FIGURE 2. The geometric interpretation of monitoring result with variables classification methods.

quality weakly related variables are simply considered to have no effect on quality. The geometric interpretation of monitoring result with variables classification methods for quality weakly fault detection is shown in Fig. 2.

In Fig. 2, $\overrightarrow{X}$ represents the process variables, and $\overrightarrow{Y}$ represents the quality variables. $\overrightarrow{X}$ is decomposed into $\overrightarrow{Xs}$ and $\overrightarrow{Xw}$. Based on variables classification methods, $\overrightarrow{Xs}$ is simply considered as a quality related variable block, and $\overrightarrow{Xw}$ is considered as a quality unrelated variable block. The direction and the amplitude of the fault in Fig. 2 are the same as the fault in Fig.1 (the direction of $\overrightarrow{f}$ is the same as $\overrightarrow{Xw}$ in Fig. 2). From Fig. 2, it can be obtained that the fault $\overrightarrow{f}$ can only be reflected on the variable block $\overrightarrow{Xw}$. Thus, it can conclude that the fault is a quality unrelated fault. However, as shown in Fig. 2, the effect of the fault on quality is $\left|\overrightarrow{f}\right| \cos\alpha$. Obviously, it is a wrong conclusion. Therefore, it is necessary to use a supervised process monitoring method to achieve further quality related feature extraction. The performance of some supervised monitoring methods is significantly affected by the number of primary components, such as KPLS and TKPLS. In addition, subspaces based on KPLS are not orthogonal to each other which inevitably lead to incomplete information decomposition. Inspired by this, a KLS-based quality related feature extraction method is provided for modeling.

B. DUAL-STEP FEATURE EXTRACTION AND MODELING

In order to separate quality strongly and weakly related variables, a maximum information coefficient (MIC) -based quality related feature extraction method is provided in the first step. The MIC is briefly introduced as follows.

The MIC is an indicator of the correlation between two-dimensional variables. On the basis of mutual information, MIC is optimized and corrected by unequal interval. The MIC can effectively reflect any functional relationship (including linear or nonlinear relationship) between variables [42].

The process variable x and quality y constitute a two-dimensional variable data set D. Thereafter, The x value in D is divided into x units, and the y values in D is divided into y units, each of which is allowed to be empty. As a result, the data set D is separated into the grid of $x \times y$. Given a grille G, let $D|_G$ be the distribution of the unit of Gin the set D. For a fixed set, different grille G forms a distribution $D|_G$ of different points. The MIC can be defined as [43]:

$$\begin{aligned} I^*(D, x, y) &= \max I(D|_G) \\ &= \max\left(\int p(x, y) \log_2 \tfrac{p(x,y)}{p(x)p(y)} dxdy\right) \end{aligned} \tag{18}$$

where $D \in \mathbb{R}^2$ is a finite set, (x, y) is a set of elements in the set D. $I(D|_G)$ represents the mutual information of $D|_G$, $p(x, y)$ is the joint probability density distribution function of x and y. Thereafter, $M(D)$ is given as:

$$M(D)_{x,y} = \frac{I^*(D, x, y)}{\log \min\{x, y\}} \tag{19}$$

$M(D)$ can be implemented in all $x \times y$ grilles with the highest normalized mutual information, and MIC is the maximum in the matrix. The MIC of D is defined as:

$$MIC(D) = \max_{xy<B(n)} \{M(D)_{x,y}\} \tag{20}$$

where n is the number of samples.

Based on the normalized process variables $\boldsymbol{X}$ and quality variables $\boldsymbol{Y}$, the MIC matrix can be built as

$$\boldsymbol{MIC} = \begin{bmatrix} MIC_{11} & MIC_{12} & \cdots & MIC_{1m} \\ MIC_{21} & MIC_{22} & \cdots & MIC_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ MIC_{l1} & MIC_{l2} & \cdots & MIC_{lm} \end{bmatrix} \in \mathbb{R}^{l\times m} \tag{21}$$

where MIC_{ij} is the MIC between $\boldsymbol{x}_i$ and $\boldsymbol{y}_j$. The correlation between each process variables $\boldsymbol{x}_i$ and quality is calculated as

$$c_i = MIC_{1i} + MIC_{2i} + \cdots + MIC_{li} \tag{22}$$

A classification criterion is given in Eq. (23) for each process variable

$$\text{No.variable} = \begin{cases} \text{strongly related variable} & c_i \geq Th \\ \text{weakly related variable} & c_i < Th \end{cases} \tag{23}$$

If c_i is greater than the threshold Th, the ith variable x_i is assigned to the strongly related variable block. The threshold Th is a preset constant, which can be set by the MIC gradient descent rate. After all $c_i\ (i \in 1, 2, \cdots, m)$is sorted in descending order. Referring to the PCs selection method of PCA, the threshold calculation method is given. The MIC gradient descent rate is defined as follows:

$$rate = \frac{c_i - c_{i+1}}{c_i} \tag{24}$$

where c_i represents the MIC of the ith sorted process variable and quality. The largest MIC gradient descent rate means that c_i and c_{i+1} have a cliff-like fall. Therefore, MIC_i can be used as a segmentation threshold (Th) for quality strongly related variables and quality weakly related variables.

Process variables are divided into two blocks: the quality strongly related variable block $\boldsymbol{X}_s \in \mathbb{R}^{n\times A}$ and the quality weakly related variable block $\boldsymbol{X}_w \in \mathbb{R}^{n\times(m-A)}$. Thereafter,

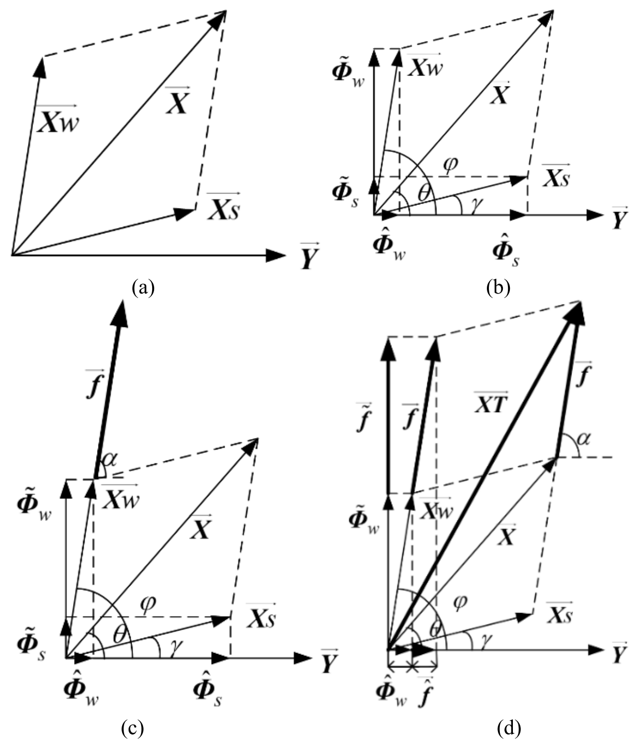


FIGURE 3. The geometric interpretation of the advantages of DSFE.

two KLS models are separately set up on $\boldsymbol{X_s}$ and $\boldsymbol{X_w}$ in the second step. The strongly related variable block $\boldsymbol{X_s}$ is decomposed into the quality-direct-related subspace (QDRS) $\hat{\boldsymbol{\Phi}}_s$ and the quality unrelated subspace (QUS) $\tilde{\boldsymbol{\Phi}}_s$. Similarly, the quality weakly related variables variable block $\boldsymbol{X}_w$ is decomposed into the quality-potential-related subspace (QPRS) $\hat{\boldsymbol{\Phi}}_w$ and the quality unrelated subspace (QUS) $\tilde{\boldsymbol{\Phi}}_w$.

C. THE GEOMETRIC INTERPRETATION OF THE ADVANTAGES OF DUAL-STEP FEATURE EXTRACTION

In this paper, the dual-step feature extraction (DSFE) method is proposed for quality weakly related fault detection. One obvious advantage of DSFE is that it can improve the monitoring sensitivity of quality weakly related variables. The geometric interpretation of the advantages of DSFE is analyzed in this section.

In Fig. 3(a), $\overrightarrow{X}$ represents the process variables, and $\overrightarrow{Y}$ represents the quality variables. $\overrightarrow{X}$ is decomposed into $\overrightarrow{Xs}$ and $\overrightarrow{Xw}$ based on MIC, where $\overrightarrow{Xs}$ is considered as a quality strongly related variable block, and $\overrightarrow{Xw}$ is considered as a quality weakly related variable block. Then, the orthogonal decomposition under supervision is performed on two variable blocks as shown in Fig. 3(b). In Fig. 3(c), the same fault as Fig.1 and Fig.2 occurs on quality weakly related variables. As shown in Fig.3(c), the variable block $\overrightarrow{Xs}$ can not be affected by the fault. In other words, $\hat{\boldsymbol{\Phi}}_s$ and $\tilde{\boldsymbol{\Phi}}_s$ are in the normal state. Thereafter, the impact of the fault on the variable block $\overrightarrow{Xw}$ is analyzed in detail in Fig.3(d). Because $\overrightarrow{f}$ occurs on $\overrightarrow{Xw}$, the projection change of $\overrightarrow{Xw}$ on $\overrightarrow{Y}$ is the effect of the fault $\overrightarrow{f}$ on $\overrightarrow{Y}$. Therefore, with the influence of $\overrightarrow{f}$, the change rate of quality related part in $\overrightarrow{Xw}$ can be calculated as

$$rate = \left|\overrightarrow{f}\right| \Big/ \left|\hat{\boldsymbol{\Phi}}_w\right| = \left|\overrightarrow{f}\right| \cos\alpha \Big/ \left|\overrightarrow{Xw}\right| \cos\varphi \tag{25}$$

Obviously, $\left|\overrightarrow{f}\right| \cos\alpha / \left|\overrightarrow{Xw}\right| \cos\varphi \;>\; \left|\overrightarrow{f}\right| \cos\alpha / \left|\overrightarrow{X}\right| \cos\theta$ (Fig.1). Compared with the orthogonal decomposition method under supervision, DSFE method improves the sensitivity of quality weakly related fault detection. Meanwhile, owning to the twice decomposition, the method proposed does not result in the wrong monitoring because of the misclassification of the quality weakly related variables.

D. STATISTICS INTEGRATION

Four subspaces are obtained in section III(B). If the statistics are established in the four subspaces directly, it may be hard to directly get explicit monitoring results. Hence, a strategy is given to reduce the monitoring statistics with the integration of similar subspace statistics. The amplitude of some faults occurring on quality weakly related variables is small, which may not cause the anomaly of quality. In this case, the DSFE method will raise the alarm by error. If direct summation (DS) or Bayesian integration (BI) is adopted, this problem can not be solved. In this paper, a weighted T^2 statistic based on quality correlation is proposed to integrate the statistics of quality related subspace.

The traditional T^2 statistics in the subspaces corresponding to $\hat{\boldsymbol{\Phi}}_s$ and $\hat{\boldsymbol{\Phi}}_w$ are calculated as follows:

$$T_{rs}^2 = \boldsymbol{t}_{sk}^T \boldsymbol{\Sigma}_s^{-1} \boldsymbol{t}_{sk} \tag{26}$$

$$T_{rw}^2 = \boldsymbol{t}_{wk}^T \boldsymbol{\Sigma}_w^{-1} \boldsymbol{t}_{wk} \tag{27}$$

Let

$$\boldsymbol{T_s} = \begin{bmatrix} \boldsymbol{t}_{s1}^T \\ \boldsymbol{t}_{s2}^T \\ \vdots \\ \boldsymbol{t}_{sn}^T \end{bmatrix} = [\boldsymbol{T}_{s1}, \boldsymbol{T}_{s2}, \cdots, \boldsymbol{T}_{sl}] \tag{28}$$

$$\boldsymbol{T}_w = \begin{bmatrix} \boldsymbol{t}_{w1}^T \\ \boldsymbol{t}_{w2}^T \\ \vdots \\ \boldsymbol{t}_{wn}^T \end{bmatrix} = [\boldsymbol{T}_{w1}, \boldsymbol{T}_{w2}, \cdots, \boldsymbol{T}_{wl}] \tag{29}$$

The correlation matrix can be built as

$$\boldsymbol{MIC}_s = \begin{bmatrix} MIC_{s11} & MIC_{s12} & \cdots & MIC_{s1l} \\ MIC_{s21} & MIC_{s22} & \cdots & MIC_{s2l} \\ \vdots & \vdots & \ddots & \vdots \\ MIC_{sl1} & MIC_{sl2} & \cdots & MIC_{sll} \end{bmatrix} \in \mathbb{R}^{l\times l} \tag{30}$$

$$\boldsymbol{MIC}_w = \begin{bmatrix} MIC_{w11} & MIC_{w12} & \cdots & MIC_{w1l} \\ MIC_{w21} & MIC_{w22} & \cdots & MIC_{w2l} \\ \vdots & \vdots & \ddots & \vdots \\ MIC_{wl1} & MIC_{wl2} & \cdots & MIC_{wll} \end{bmatrix} \in \mathbb{R}^{l\times l} \tag{31}$$

where MIC_{sij} is the correlation between $\boldsymbol{T}_{sj}$ and $\boldsymbol{y}_i$, and MIC_{wij} is the correlation between $\boldsymbol{T}_{wj}$ and $\boldsymbol{y}_i$. The weight matrix can be calculated as

$$\boldsymbol{W}_s = [w_{s1}, w_{s2}, \cdots, w_{sl}] = \begin{bmatrix} MIC_{s11} + MIC_{s21} + \cdots + MIC_{sl1} \\ MIC_{s12} + MIC_{s22} + \cdots + MIC_{sl2} \\ \vdots \\ MIC_{s1l} + MIC_{s2l} + \cdots + MIC_{sll} \end{bmatrix} \tag{32}$$

$$\boldsymbol{W}_w = [w_{w1}, w_{w2}, \cdots, w_{wl}] = \begin{bmatrix} MIC_{w11} + MIC_{w21} + \cdots + MIC_{wl1} \\ MIC_{w12} + MIC_{w22} + \cdots + MIC_{wl2} \\ \vdots \\ MIC_{w1l} + MIC_{w2l} + \cdots + MIC_{wll} \end{bmatrix} \tag{33}$$

Based on the new matrix $\boldsymbol{T}_r = [\boldsymbol{T}_s, \boldsymbol{T}_w] = \left[\boldsymbol{t}_{r1}^T, \boldsymbol{t}_{r2}^T, \cdots \boldsymbol{t}_{rn}^T\right]^T$ and the weight matrix $\boldsymbol{W}_r = [\boldsymbol{W}_s, \boldsymbol{W}_w]$, the weighted T^2 (WT^2) statistic can be established as:

$$WT^2 = \boldsymbol{t}_{rk}^T diag\left(\boldsymbol{W}_r\right) \left(\frac{\boldsymbol{T}_r^T \boldsymbol{T}_r}{n-1}\right)^{-1} diag\left(\boldsymbol{W}_r\right) \boldsymbol{t}_{rk} \tag{34}$$

The threshold Th_{WT^2} of WT^2 can be estimated by the kernel density estimation (KDE) [44]. To avoid the numerical problem in the inverse process, Q_{unrs} and Q_{unrw} statistics are established in the two quality unrelated subspace $\tilde{\boldsymbol{\Phi}}_s$ and $\tilde{\boldsymbol{\Phi}}_w$.

A comprehensive monitoring index BIC is established based on the Bayesian inference as

$$BIC = \frac{p\left(Q_{unrs}|F\right) p\left(F|Q_{unrs}\right)}{p\left(Q_{unrs}|F\right) + p(Q_{unrw}|F)} + \frac{p(Q_{unrw}|F) p(F|Q_{unrw})}{p[Q_{unrs}|F] + p(Q_{unrw}|F)} \tag{35}$$

where

$$p\left(Q_{unrs}\right) = p\left(Q_{unrs}|N\right) p(N) + p\left(Q_{unrs}|F\right) p(F) \tag{36}$$

$$p\left(Q_{unrw}\right) = p\left(Q_{unrw}|N\right) p(N) + p\left(Q_{unrw}|F\right) p(F) \tag{37}$$

$$p\left(Q_{unrs}|N\right) = \exp\left[-\frac{Q_{unrs}}{(Q_{unrs})_{\lim}}\right] \tag{38}$$

$$p\left(Q_{unrs}|F\right) = \exp\left[-\frac{(Q_{unrs})_{\lim}}{Q_{unrs}}\right] \tag{39}$$

$$p\left(Q_{unrw}|N\right) = \exp\left[-\frac{Q_{unrw}}{(Q_{unrw})_{\lim}}\right] \tag{40}$$

$$p\left(Q_{unrw}|F\right) = \exp\left[-\frac{(Q_{unrw})_{\lim}}{Q_{unrw}}\right] \tag{41}$$

where N and F represent normal and fault, respectively. In addition, $p(N)$ is equal to the confidence level α and $p(F)$ is equal to $1 - \alpha$.

E. MONITORING LOGIC

The monitoring logics are listed as follows:

$$\begin{cases} WT^2 < Th_{WT^2} \\ BIC < 1 - \alpha \end{cases} \Rightarrow faultfree \tag{42}$$

$$\begin{cases} WT^2 < Th_{WT^2} \\ BIC > 1 - \alpha \end{cases} \Rightarrow quality - unrelated\ fault \tag{43}$$

$$WT^2 > Th_{WT^2} \Rightarrow quality - related\ fault \tag{44}$$

The flowchart of fault detection can be shown in Fig. 4.

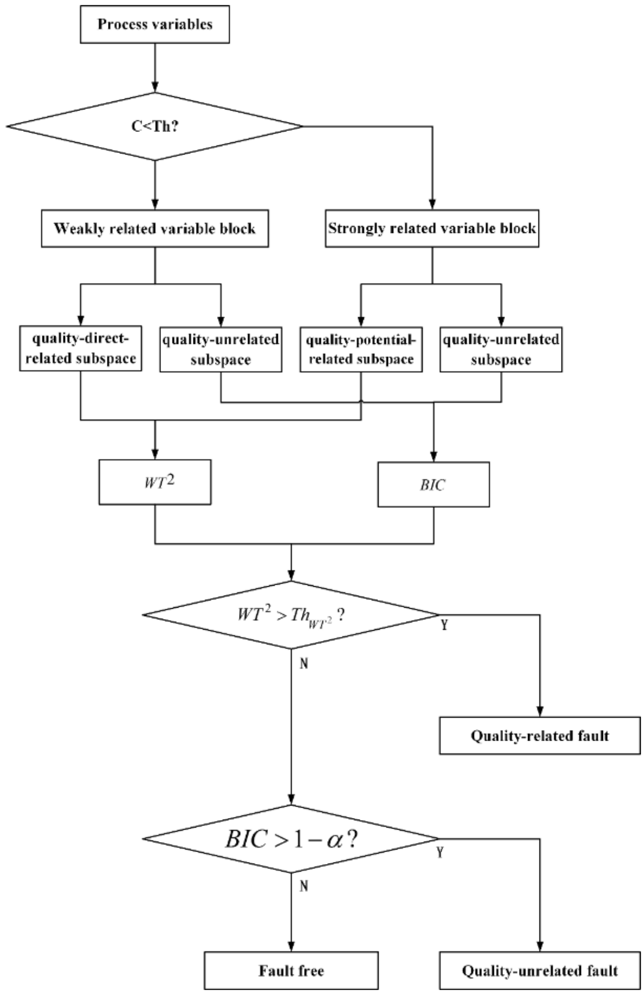


FIGURE 4. **The flowchart of fault detection.**

IV. EXAMPLES AND APPLICATIONS

A. NUMERICAL EXAMPLE

A numerical example is introduced as follows:

$$\begin{cases} x_1, x_4, x_7, x_8 \sim U\left(1, 2\right) \\ x_2 = x_1^2 + 3x_1 + 4 \\ x_3 = x_2^2 + 3x_2 + 4 \\ x_5 = x_4^2 + 3x_4 + 4 \\ x_6 = x_5^2 + 3x_5 + 4 \\ y = 10x_2^2 + 10x_2x_3 + 10x_1 + x_5^2 + x_5x_6 + x_4 \end{cases} \tag{45}$$

The MIC between x_i and y is shown in Fig. 5. Obviously, x_1, x_2 and x_3 should be divided into the quality strongly related variable block, and x_4, x_5, x_6, x_7 and x_8 should be divided into the quality weakly related variable block. 1000 normal

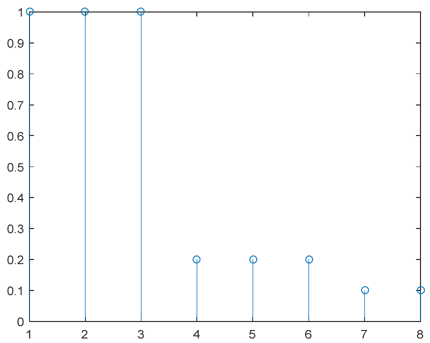


FIGURE 5. The MIC between x_i and y.

TABLE 1. The fault scenarios.

NO.	Fault style	Fault location	Fault amplitude	Quality related /unrelated fault
1	Step	x_1	2	Quality related fault
2	Step	x_7	2	Quality unrelated fault
3	Step	x_4	2	Quality related fault
4	Step	x_4	0.5	Quality unrelated fault

samples are generated as the training dataset, and 500 normal samples and 500 fault samples are generated as the testing dataset. The fault scenarios are shown in Table 1. The monitoring results of mutual information-kernel principal component analysis (MI-KPCA), kernel partial least squares (KPLS), kernel Least Squares (KLS), Dual-Step feature extraction (DSFE), direct summation DSFE (DSDSFE), Bayesian integration DSFE (BIDSFE) and weighted DSFE (WDSFE) can be shown in Table 2.

As can be seen in Table 1, fault 1 occurs on the variable x_1, and it is a quality related fault. It is worth noting that x_1 is strongly related to y, which causes the fault have a significant effect on y. From the monitoring result in Table 2, MI-KPCA, KPLS, KLS, DSFE, DSDSFE, BIDSFE and WDSFE can obtain the satisfactory monitoring performance for fault 1. However, both MI-KPCA and DSFE have two statistics for quality related fault detection, which makes it difficult to obtain direct monitoring results.

For fault 2, it occurs on the variable x_7. Since x_7 is unrelated to y, fault 2 is a quality unrelated fault. In other words, fault 2 does not affect the quality. The seven methods can keep low fault detection rates in the quality related subspaces. In the quality unrelated subspaces, all methods can keep alarm all the time. Therefore, all of them can be concluded that fault 2 is a quality unrelated fault. However, MI-KPCA and DSFE cannot be directly concluded as the other four methods.

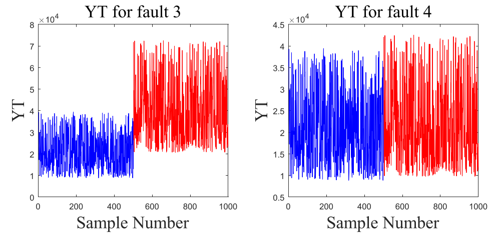


FIGURE 6. The effects of fault 3 and fault 4 on y.

From Table 1, both fault 3 and fault 4 occur on the variable x_4. Different from fault 1, x_4 is weakly related to y. It means that when the magnitude of the fault is large, it is a quality related fault. However, when the magnitude of the fault is small, the effect of the fault on the quality can be negligible. In other words, fault 3 is a quality related fault and fault 4 is a quality unrelated fault. In order to further explain, the effects of fault 3 and fault 4 on y are shown in Fig. 6.

Fig. 7 shows the monitoring results of the fault 3 with the six methods. Fig. 7 (a) is the monitoring result of MI-KPCA, which is a method that only takes the classification of process variables into account. Since the x_4 is classified into a quality unrelated variable block by error, the fault 3 can only be reflected in the quality unrelated subspace. As a result of this, MI-KPCA got the wrong result. Fig. 7(b) and Fig. 7(c) are the monitoring result of KPLS and KLS, which only focuses on postprocessing technology. Because the quality weakly related information is covered by quality strongly related information, the fault detection rates of KPLS and KLS in the quality related subspace are not very high. Hence, it is difficult to obtain satisfactory results. Fig. 7(d) presents the monitoring result of DSFE, which considers both the classification of process variables and the supervised orthogonal decomposition. It not only overcomes difficulties on classification of the quality weakly related variables, but also enhances the sensitivity of the quality weakly related information. Compared with MI-KPCA and KLS, the monitoring results of DSFE are the most satisfactory. However, DSFE has four subspaces, and it is not easy to get direct results. Fig. 7(e) shows the monitoring results of the three statistics integration method. It can be seen that when the fault occurred in the quality weakly related variables causes significant quality anomalies, the three methods can get satisfactory monitoring results.

TABLE 2. The fault detection rates of MI-KPCA, KLS, DSFE, DSDSFE, BIDSFE and DWDEFE.

NO.	MI-KPCA				KPLS		KLS		DSFE				DSDSFE		BIDSFE		WDSFE	
	T^2_{rel}	SPE_{rel}	T^2_{unr}	SPE_{unr}	T^2_{rel}	SPE_{unr}	T^2_{rel}	Q_{unr}	T^2_{srel}	Q_{sunr}	T^2_{wrel}	Q_{wunr}	T^2_{rel}	Q_{unr}	BIC_{rel}	BIC_{unr}	WT^2	BIC
1	99.2	99.2	0.8	0.8	100	7.2	100	6.8	100	100	0.8	0.8	100	100	100	100	100	100
2	0.8	1.4	100	0	0	94.6	0.2	100	1	0.2	7.6	100	1.8	100	6.6	100	0.4	100
3	0	0	100	100	18.6	100	23.8	100	0.4	0.4	100	100	100	0.6	100	100	100	100
4	0.2	0.6	38.4	46.6	6.6	46.2	8.6	19.8	0.4	0.4	46.6	45.8	35.6	0.6	46.6	45.8	4.8	45.8

TABLE 3. Notes for statistic names.

Method	Statistic names	Meaning
MI-KPCA	T^2_{rel}	The T^2 statistic in the quality related variable set
	SPE_{rel}	The SPE statistic in the quality related variable set
	T^2_{unr}	The T^2 statistic in the quality unrelated variable set
	SPE_{unr}	The SPE statistic in the quality unrelated variable set
KPLS	T^2_{rel}	The T^2 statistic for quality related features
	SPE_{unr}	The SPE statistic for quality unrelated features
KLS	T^2_{rel}	The T^2 statistic for quality related features
	Q_{unr}	The Q statistic for quality unrelated features
DSFE	T^2_{srel}	The T^2 statistic in QDRS for QRSV set
	Q_{sunr}	The Q statistic in QUS for QRSV set
	T^2_{wrel}	The T^2 statistic in QPRS for QWSV set
	Q_{wunr}	The Q statistic in QUS for QWSV set
DSDSFE	T^2_{rel}	The integrated T^2 statistic for QDRS and QPRS based on DS
	Q_{unr}	The integrated Q statistic for two QUSs based on DS
BIDSFE	BIC_{rel}	The integrated BIC statistic for QDRS and QPRS based on BI
	BIC_{unr}	The integrated BIC statistic for two QUSs based on BI
WDSFE	WT^2	The integrated WT^2 statistic for QDRS and QPRS based on weights
	BIC	The integrated BIC statistic for two QUSs based on BI

Fig. 8 shows the monitoring results of fault 4 with the six methods. In Fig. 8(a), it is fluky that x_4 is classified into an unrelated variable block, so that MI-KPCA gets the correct results. The false alarm rates in the quality related subspace of the KPLS in Fig. 8(b) and the KLS in Fig. 8(c) are not very high, thus obtain the available monitoring results. Nevertheless, for DSFE in Fig. 8(d), it improves the monitoring sensitivity of the quality weakly related variables. When the amplitude of the fault on the quality weakly related variable has no effect on the quality, DSFE still has a high fault detection rate in the quality related subspace. This may cause unnecessary downtime and maintenance. The monitoring results of the three statistics integration methods are displayed in Fig. 8(e). DSDSFE and BIDSFE cannot change the high false alarm rates in the quality related subspace. On the contrary, WDSFE takes the correlation between each latent variable in the quality related subspace and y into account. Accordingly, the embodiment of the fault in the statistic can be eliminated by the weight. Obviously, WDSFE can get the most satisfactory monitoring results.

In a word, the advantage of WDSFE is to explain the fault on the quality weakly related variables. It not only overcomes the difficulties on classification of the quality weakly related variable, but also solves the problem that the information of the quality weakly related variable is easily concealed by the quality strongly related variable. Then the WT^2 statistics can

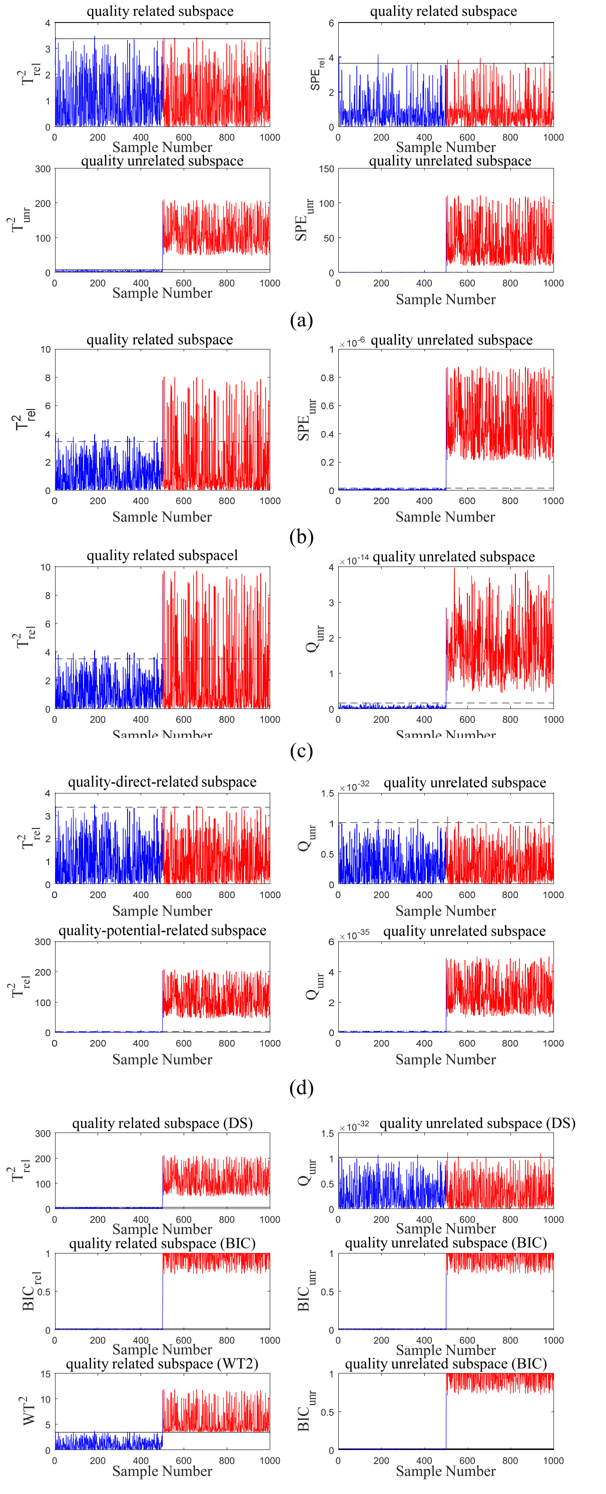


FIGURE 7. The monitoring results of MI-KPCA, KLS, DSFE, DSDSFE, BIDSFE and WDSFE for fault 3. (a) MI -KPCA. (b) KPLS. (c) KLS. (d) DSFE. (e) DSDSFE, BIDSFE and WDSFE.

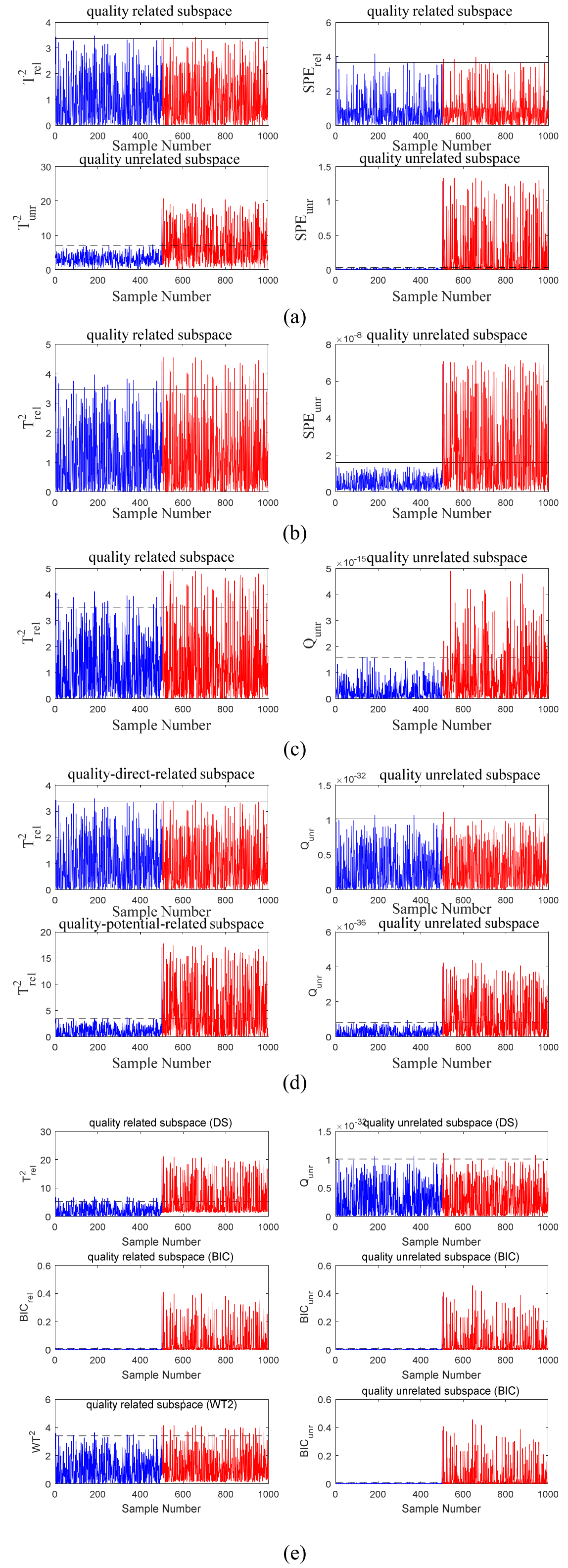


FIGURE 8. The monitoring results of MI-KPCA, KLS, DSFE, DSDSFE, BIDSFE and WDSFE for fault 4. (a) MI –KPCA. (b) KPLS. (c) KLS. (d) DSFE. (e) DSDSFE, BIDSFE and WDSFE.

further improve the accuracy of monitoring results for quality weakly related fault. Finally, the integration of two statistics has settled the deficiency of several statistics.

B. INDUSTRIAL EXAMPLE

Penicillin's fermentation process has serious nonlinearities and uncertainties. In this simulation, 10 readily available measurement variables are used as process variables, and biomass concentration is used as the quality variable. Thereafter, three normal batches are considered in the offline modeling. From the above numerical examples and a large number of scholars' researches, there is no doubt that most of the methods can achieve satisfactory monitoring performance for the faults occurring on quality strongly related variables and quality unrelated variables. Therefore, the quality weakly related faults are analyzed in detail in this section.

1) FAULT SCENARIO 1

The fault occurs on the aeration rate and the fault amplitude is -15%. The influence degree for biomass concentration of the fault can be described as

$$rate = \frac{y_f - y_n}{y_n} = 0.255\% \tag{46}$$

In this simulation, the influence degree below 0.5% is considered normal. Obviously, the fluctuation of biomass concentration is acceptable. Thus the fault is a quality unrelated fault. The monitoring results of MI-KPCA, KPLS, KLS, DSFE, DSDSFE, BIDSFE and WDSFE for the fault scenario are shown in Fig. 9.

It can be seen from the monitoring results that MI-KPCA, KPLS, KLS and WDSFE can get correct results. However, the performances of the other four methods are unsatisfactory. It is noteworthy that MI-KPCA does not always obtain a satisfying monitoring result for the fault on the quality weakly related variables. It is because the aeration rate is properly classified into the quality unrelated variable block. Analogously, since the information on the quality weakly related variables is covered by quality strongly related variables, KPLS and KLS lose the ability to detect the quality weakly related fault. Therefore, they get the correct monitoring results. Given that DSFE has improved the monitoring sensitivity of the quality weakly related fault, the quality-potential-related subspace of the DSFE continues to be alerting when the fault occurs. It is an undesirable situation. For the three statistics integration methods, DSDSFE and BIDSFE cannot distinguish the effect of the quality-directly-related subspace and quality-potential-related subspace on quality. As a result, they will still give the alarm in the quality related subspace. WDSFE takes the correlation between the latent variable and the quality as the weight. When the fault amplitude is small, the influence on the statistics can be eliminated by a smaller weight.

2) FAULT SCENARIO 2

The fault occurs on the aeration rate and the fault amplitude is -30%. As the amplitude of the fault becomes larger, the influence of the fault on the biomass concentration is stronger. The influence degree for biomass concentration of the fault can be

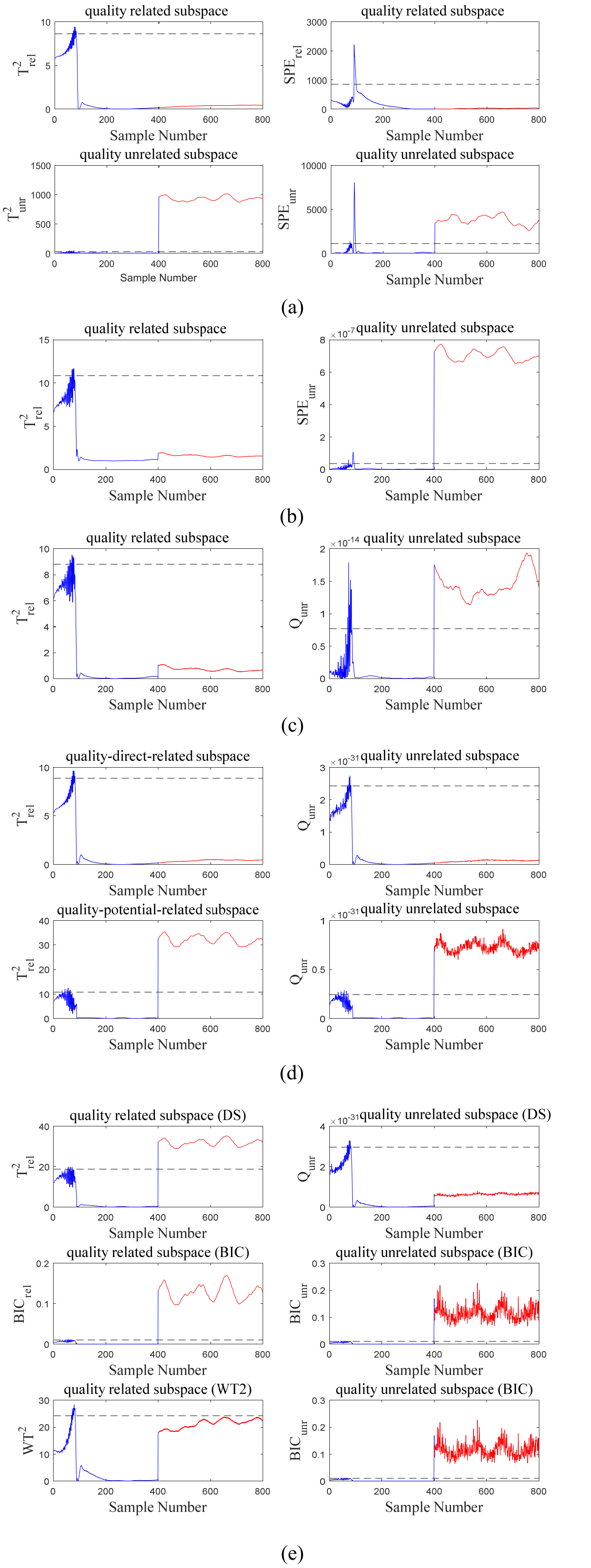


FIGURE 9. The monitoring results of MI-KPCA, KLS, DSFE, DSDSFE, BIDSFE and WDSFE. (a)MI –KPCA. (b) KPLS. (c) KLS. (d) DSFE. (e) DSDSFE, BIDSFE and WDSFE.

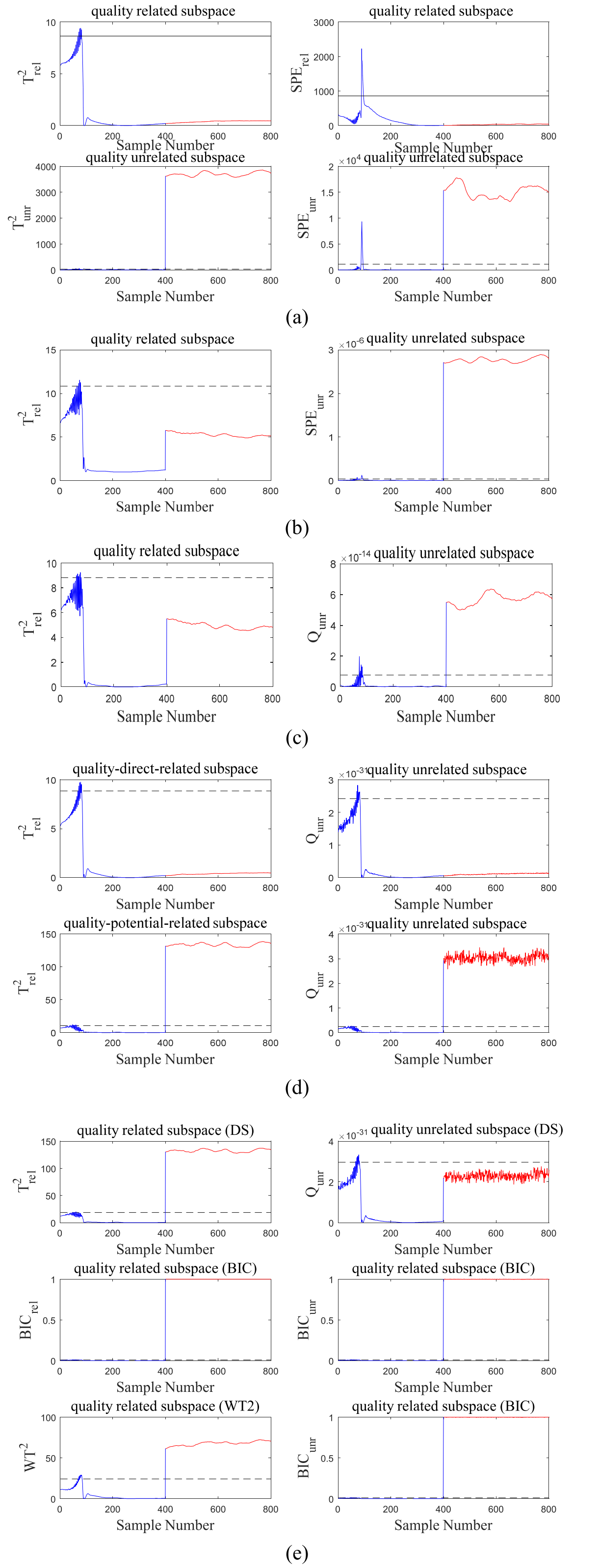


FIGURE 10. The monitoring results of MI-KPCA, KLS, DSFE, DSDSFE, BIDSFE and WDSFE. (a) MI –KPCA. (b) KPLS. (c) KLS. (d) DSFE. (e) DSDSFE, BIDSFE and WDSFE.

described as

$$rate = \frac{y_f - y_n}{y_n} = 0.81\% \tag{47}$$

Other than the fault scenario 1, the fault is a quality related fault. The monitoring results of MI-KPCA, KPLS, KLS, DSFE, DSDSFE, BIDSFE and WDSFE for the fault scenario are shown in Fig. 10.

In Fig. 10, DSFE, DSDSFE, BIDSFE and WDSFE can all get satisfactory monitoring results. For MI-KPCA, the aeration rate is partitioned into the quality unrelated variable block. Thus, MI-KPCA considers that the changes in the aeration rate have no effect on the quality. Although the amplitude of the fault has increased, the information of the quality weakly related variables is still obscured by the information of quality strongly related variable. Hence, the results of KPLS and KLS are still wrong. DSFE takes into account both the process variables classification and post processing technology, which can overcome the difficulties on classification of quality weakly related variables and avoid the information of quality weakly related variables being submerged. In this case, the three integration methods can all obtain the correct monitoring results directly.

To sum up, WDSFE can give reasonable explanation for the faults on the quality weakly related variables. No matter whether the faults on the quality weakly related variables affect the significant fluctuation of the quality, the results of WDSFE are always trustworthy.

V. CONCLUSION

In this paper, WDSFE is presented for quality weakly related fault detection. Firstly, DSFE is used to improve the sensitivity of quality weakly related information and overcome the impact on the wrong classification of quality weakly related variables. In the first step, process variables are divided into the quality strongly related variable block and the quality weakly related variable block based on MIC. In the second step, KLS is performed on the two variable blocks, respectively. Both the process variables classification and post processing technology are considered by DSFE. In addition, a weighted T^2 statistic based on the correlation is proposed for integrating the statistics of two quality related subspaces, which is an effective way to improve monitoring accuracy for quality weakly related fault detection. The *BIC* statistic is proposed for integrating the statistics of two quality unrelated subspaces. Two integrated statistics can provide satisfactory monitoring results directly. The simulation results of the literature example and industry example show that WDSFE can achieve the most satisfied monitoring results for quality weakly related fault detection.

REFERENCES

[1] C. Tong, X. Shi, and T. Lan, "Statistical process monitoring based on orthogonal multi-manifold projections and a novel variable contribution analysis," *ISA Trans.*, vol. 65, pp. 407–417, Nov. 2016.
[2] J. Zhu, Y. Yao, D. Li, and F. Gao, "Monitoring big process data of industrial plants with multiple operating modes based on Hadoop," *J. Taiwan Inst. Chem. Eng.*, vol. 91, pp. 10–21, Oct. 2018.
[3] Z. Lv, X. Yan, and Q. Jiang, "Batch process monitoring based on multiple-phase online sorting principal component analysis," *ISA Trans.*, vol. 64, pp. 342–352, Sep. 2016.
[4] C. Tong, T. Lan, and X. Shi, "Double-layer ensemble monitoring of non-Gaussian processes using modified independent component analysis," *ISA Trans.*, vol. 68, pp. 181–188, May 2017.
[5] B. Song and H. Shi, "Temporal-spatial global locality projections for multimode process monitoring," *IEEE Access*, vol. 6, pp. 9740–9749, 2018.
[6] Y.-J. Liu, T. Chen, and Y. Yao, "Nonlinear process monitoring and fault isolation using extended maximum variance unfolding," *J. Process Control*, vol. 24, no. 6, pp. 880–891, 2014.
[7] R. Raveendran, H. Kodamana, and B. Huang, "Process monitoring using a generalized probabilistic linear latent variable model," *Automatica*, vol. 96, pp. 73–83, Oct. 2018.
[8] Q. Lu, B. Jiang, R. B. Gopaluni, P. D. Loewen, and R. D. Braatz, "Sparse canonical variate analysis approach for process monitoring," *J. Process Control*, vol. 71, pp. 90–102, Nov. 2018.
[9] Q. Jiang and X. Yan, "Parallel PCA–KPCA for nonlinear process monitoring," *Control Eng. Pract.*, vol. 80, pp. 17–25, 2018.
[10] X. Chen, J. Wang, and J. Zhou, "Process monitoring based on multivariate causality analysis and probability inference," *IEEE Access*, vol. 6, pp. 6360–6369, 2018.
[11] Z. Ge, "Distributed predictive modeling framework for prediction and diagnosis of key performance index in plant-wide processes," *J. Process Control*, vol. 65, pp. 107–117, May 2018.
[12] K. Zhang, Y. A. W. Shardt, Z. Chen, X. Yang, S. X. Ding, and K. Peng, "A KPI-based process monitoring and fault detection framework for large-scale processes," *ISA Trans.*, vol. 68, pp. 276–286, May 2017.
[13] C. H. Zhao and Y. X. Sun, "Multispace total projection to latent structures and its application to online process monitoring," *IEEE Trans. Control Syst. Technol.*, vol. 22, no. 3, pp. 868–883, May 2014.
[14] Q. Liu, Q. Zhu, S. J. Qin, and T. Chai, "Dynamic concurrent kernel CCA for strip-thickness relevant fault diagnosis of continuous annealing processes," *J. Process Control*, vol. 67, pp. 12–22, Jul. 2018.
[15] G. Wang, J. Li, C. Sun, and J. Jiao, "Least squares and contribution plot based approach for quality-related process monitoring," *IEEE Access*, vol. 6, pp. 54158–54166, 2018.
[16] J. Zhu, Z. Ge, and Z. Song, "Multimode process data modeling: A Dirichlet process mixture model based Bayesian robust factor analyzer approach," *Chemometrics Intell. Lab. Syst.*, vol. 142, pp. 231–244, Mar. 2015.
[17] Z. Lv and X. Yan, "Hierarchical support vector data description for batch process monitoring," *Ind., Eng. Chem. Res.*, vol. 55, no. 34, pp. 9205–9214, 2016.
[18] Z. Ge and X. Chen, "Supervised linear dynamic system model for quality related fault detection in dynamic processes," *J. Process Control*, vol. 44, pp. 224–235, Aug. 2016.
[19] B. Song and H. Shi, "Fault detection and classification using quality-supervised double-layer method," *IEEE Trans. Ind. Electron.*, vol. 65, no. 10, pp. 8163–8172, Oct. 2018.
[20] J. Huang and X. Yan, "Relevant and independent multi-block approach for plant-wide process and quality-related monitoring based on KPCA and SVDD," *ISA Trans.*, vol. 73, pp. 257–267, Feb. 2018.
[21] J. Dong, Q. Wang, M. Wang, and K. Peng, "Data-driven quality monitoring techniques for distributed parameter systems with application to hot-rolled strip laminar cooling process," *IEEE Access*, vol. 6, pp. 16646–16654, 2018.
[22] B. Song, X. Zhou, H. Shi, and Y. Tao, "Performance indicator oriented concurrent subspace process monitoring method," *IEEE Trans. Ind. Electron.*, to be published, doi: 10.1109/TIE.2018.2868316.
[23] Q. Li, F. Pan, Z. Zhao, and J. Yu, "Process modeling and monitoring with incomplete data based on robust probabilistic partial least square method," *IEEE Access*, vol. 6, pp. 10160–10168, 2018.
[24] J. Jiao, N. Zhao, G. Wang, and S. Yin, "A nonlinear quality-related fault detection approach based on modified kernel partial least squares," *ISA Trans.*, vol. 66, pp. 275–283, Jan. 2017.
[25] Z. Yan, T.-H. Kuang, and Y. Yao, "Multivariate fault isolation of batch processes via variable selection in partial least squares discriminant analysis," *ISA Trans.*, vol. 70, pp. 389–399, Sep. 2017.
[26] K. Zhang, H. Hao, Z. Chen, S. X. Ding, and K. Peng, "A comparison and evaluation of key performance indicator-based multivariate statistics process monitoring approaches," *J. Process Control*, vol. 33, pp. 112–126, Sep. 2015.
[27] Y. Zheng, Z. Liu, W. Yang, B. Tao, and Y. Wang, "Parallel projection to latent structures for quality-relevant process monitoring," *J. Taiwan Inst. Chem. Eng.*, vol. 80, pp. 76–84, Nov. 2017.
[28] Q. Zhu, Q. Liu, and S. J. Qin, "Concurrent quality and process monitoring with canonical correlation analysis," *J. Process Control*, vol. 60, pp. 95–103, Dec. 2017.

[29] K. Peng, K. Zhang, B. You, and J. Dong, "Quality-related prediction and monitoring of multi-mode processes using multiple PLS with application to an industrial hot strip mill," *Neurocomputing*, vol. 168, pp. 1094–1103, Nov. 2015.

[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, 2010.

[31] 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, 2010.

[32] 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, 2013.

[33] 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, Mar. 2015.

[34] G. Wang, H. Luo, and K. Peng, "Quality-related fault detection using linear and nonlinear principal component regression," *J. Franklin Inst.*, vol. 353, pp. 2159–2177, Apr. 2016.

[35] K. Peng, K. Zhang, J. Dong, and B. You, "Quality-relevant fault detection and diagnosis for hot strip mill process with multi-specification and multi-batch measurements," *J. Franklin Inst.*, vol. 352, no. 3, pp. 987–1006, 2015.

[36] J. Huang and X. Yan, "Quality relevant and independent two block monitoring based on mutual information and KPCA," *IEEE Trans. Ind. Electron.*, vol. 64, no. 8, pp. 6518–6527, Aug. 2017.

[37] K. Peng, K. Zhang, and G. Li, "Quality-related process monitoring based on total kernel PLS model and its industrial application," *Math. Problems Eng.*, vol. 2013, Oct. 2013, Art. no. 707953.

[38] Y. Zhang, R. Sun, and Y. Fan, "Fault diagnosis of nonlinear process based on KCPLS reconstruction," *Chemometrics Intell. Lab. Syst.*, vol. 140, pp. 49–60, Jan. 2015.

[39] Q. Jia and Y. Zhang, "Quality-related fault detection approach based on dynamic kernel partial least squares," *Chem. Eng. Res. Design*, vol. 106, pp. 242–252, Feb. 2016.

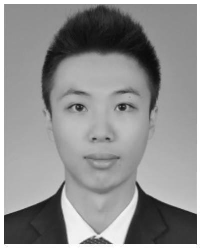
[40] G. Wang and J. Jiao, "A kernel least squares based approach for nonlinear quality-related fault detection," *IEEE Trans. Ind. Electron.*, vol. 64, no. 4, pp. 3195–3204, Apr. 2017.

[41] G. Wang, J. Jiao, and S. Yin, "A kernel direct decomposition-based monitoring approach for nonlinear quality-related fault detection," *IEEE Trans. Ind. Informat.*, vol. 13, no. 4, pp. 1565–1573, Aug. 2017.

[42] D. N. Reshef *et al.*, "Detecting novel associations in large data sets," *Science*, vol. 334, no. 6062, pp. 1518–1524, 2011.

[43] C. Jisheng, W. Yunxin, G. Hong, W. Tao, and Q. Peng, "Evaluation on light sources for electric power emergency recovery system based on Grille method and maximum information coefficient method," *Energy Procedia*, vol. 145, pp. 181–186, Jul. 2018.

[44] A. Giantomassi, F. Ferracuti, S. Iarlori, G. Ippoliti, and S. Longhi, "Electric motor fault detection and diagnosis by kernel density estimation and Kullback–Leibler divergence based on stator current measurements," *IEEE Trans. Ind. Electron.*, vol. 62, no. 3, pp. 1770–1780, Mar. 2015.



BO ZHAO received the B.S. degree in automation from the East China University of Science and Technology, Shanghai, China, in 2015, where he is currently pursuing the Ph.D. degree in control theory and control engineering. His research interests include statistical data analysis, process monitoring, and fault diagnosis.

HONGBO SHI received the M.S. degree and the Ph.D. degree in control theory and control engineering from the East China University of Science and Technology, China, in 1989 and 2000, respectively. He was the 2003 Shu Guang Scholar of Shanghai. He is currently a Professor with the East China University of Science and Technology. His research interests include modeling of industrial process and advanced control technology, theory and methods of integrated automation systems, condition monitoring, and fault diagnosis of industrial process.

BING SONG received the B.S. degree in automation and the Ph.D. degree in control theory and control engineering from the East China University of Science and Technology, Shanghai, China, in 2012 and 2017, respectively. He is currently a Postdoctoral Researcher with the Department of Automation, East China University of Science and Technology. His research interests include feature extraction, fault detection, fault diagnosis, and multimode process monitoring.

SHUAI TAN received the B.S. degree in automation and the Ph.D. degree in control theory and control engineering from Northeastern University, China, in 2005 and 2012, respectively. She is currently an Associate Professor with the East China University of Science and Technology, Shanghai, China. Her research interests include operation state evaluation for complex industrial process, fault monitoring and diagnosis, and machine learning of image information.