Variable Sub-Region Canonical Variate Analysis for Dynamic Process Monitoring

YUPING CAO\(^{1}\), LEI YU\(^{1}\), XIAOGANG DENG\(^{1}\), AND XIAOLING ZHANG\(^{2}\)

\(^{1}\)College of Information and Control Engineering, China University of Petroleum (East China), Qingdao 266580, China
\(^{2}\)College of Shengli, China University of Petroleum (East China), Dongying 257097, China

Corresponding author: Yuping Cao (caoyyp@upc.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61403418 and Grant 21606256, in part by the Natural Science Foundation of Shandong Province under Grant ZR2016FQ21 and Grant ZR2016BQ14, in part by the Shandong Provincial Key Program of Research and Development under Grant 2018GGX101025, in part by the Fundamental Research Funds for the Central Universities under Grant 17CX02054, and in part by the Project of Shandong Province Higher Educational Science and Technology Program under Grant J18KA359.

ABSTRACT When all process variables are used to establish one data-based model, some variables have little contributions to the model. When these variables are grouped into a small local variable set to develop a data-based model, their contributions will become large. It is called local variable characteristic. Standard canonical variate analysis (CVA) based process monitoring method doesn’t consider local process variable characteristic. To solve this problem, a variable sub-region based canonical variate analysis (V-CVA) is proposed for dynamic process monitoring by enhancing the information of local variables. The proposed method combines variable sub-region division and Bayesian fusion. First, standard CVA is performed by using all process variables. Then, K-means clustering is adopted to divide process variables into sub-regions based on the mutual information of process variables and canonical variables. For each variable sub-region, CVA is conducted to compute local process monitoring statistics. Finally, local statistics are fused to build ensemble statistics based on the idea of Bayesian inference. Compared with the standard CVA method, the proposed V-CVA method can emphasize the information of local variables, and has better monitoring performance in the canonical variable feature subspace and the prediction residual subspace. Two examples demonstrate the effectiveness of the proposed method.

INDEX TERMS Process monitoring, dynamic, local information, variable sub-region, canonical variate analysis, Bayesian inference.

I. INTRODUCTION

Process monitoring is of great importance to large scale continuous operating plants, such as oil refineries, chemical plants, and so on [1]. Continuous processes often have some process characteristics, such as dynamic characteristic and local variable characteristic. Usually, process variables differ in physical meanings and values. Take the nominal state of a continuous stirred tank reactor (CSTR) for example [2], the concentration of reactant is 0.037 mol/L, the reactor liquid level is 0.6 m, the feed concentration is 1 mol/L, the coolant flow rate is 15 L/min, the feed flow rate is 100 L/min, and the feed temperature is 320 K. It can be seen that the value disparities between them are large. Process variables change according to their own laws. For instance, pressure variables and flow rate variables change quickly, while temperature variables change slowly. When all process variables are used to establish one data-based model, some variables have little contributions to the model. When these variables are grouped into a small local variable set to develop a data-based model, their contributions will become large. It is called local variable characteristic. Furthermore, due to process operating principle and closed loop control, high dimensional process variables influence each other seriously. Hence, correlation is an important variable characteristic. Process monitoring methods that consider process characteristics and variable characteristics can achieve better monitoring effects.

Process monitoring technologies have developed rapidly in recent years. Until now, three research directions have been formed in process monitoring field [3]. They are driven...
by model [4]–[7], knowledge [8], [9] and data [10]–[16], respectively. Model driven fault diagnosis approaches are very effective for electrical and electronic devices, which have precise models. However, it is difficult for engineers to build an accurate chemical process model. Knowledge driven methods fully utilize process prior knowledge. For complex chemical and oil refining processes, collecting and describing knowledge are challenging tasks. Data based approaches are more suitable for chemical and oil refining processes, since a large number of real operating data are collected every few minutes. Because process variables are high dimensional, multivariate dimensionality reduction and feature extraction methods are widely popular, such as principal component analysis (PCA) [17]–[21], independent component analysis (ICA) [22]–[25], and slow feature analysis (SFA) [26], [27].

Standard PCA, ICA and SFA are data processing methods for one variable set. Partial least square (PLS) and canonical variate analysis (CVA) can analyze the relationship between two variable sets. For PLS method, measured variables are used to form a process vector and a quality vector. Then, PLS abstracts process features according to their predictive capability for quality variables [28]–[32]. However, standard PCA/PLS models are static. They don’t consider variables’ time serial correlations, i.e. dynamic characteristic. Although dynamic PCA [17] was proposed to solve this problem, dynamic PCA mixes variable correlations and time serial correlations together to perform optimization. It is difficult to interpret dynamic relations. For CVA method, time serial data are utilized to construct a past vector and a future vector. Its optimization objective is just the correlation coefficient of the past and the future. The extracted canonical variables are uncorrelated, and have the strongest capability to predict the future. They can be seen as a kind of subspace identification state. Hence, CVA not only takes the correlations between variables into account, but also considers the correlations of variables’ time series. The main differences between dynamic PCA and CVA are the optimization objectives and feature interpretations, which make CVA outperform dynamic PCA in handling dynamic data. The optimization objective of dynamic PCA is the variances of features, while the optimization objective of CVA is the correlations of the past features and the future features. It is difficult to interpret the meanings of dynamic PCA features, while the features of CVA are state estimations of actual systems [33], [34].

To enhance the interpretability of models, variables can be divided into some groups. A variable group is called a variable sub-block or a variable sub-region. Multi-block PCA was first proposed. Westerhuis et al. [35] compared and analyzed multi-block and hierarchical PCA and PLS. By analyzing multi-block PCA and PLS, Qin et al. [36] proposed a decentralized monitoring strategy. Li et al. [37] proposed an improved PCA to detect and isolate sensor faults based on abnormal sub-region. To simplify fault diagnosis for semiconductor processes, Cherry and Qin [38] introduced multi-block fault contributions to a combined PCA monitoring index. Ge et al. [39] used the correlations of variables and principal components to group variables and developed multiple linear subspaces to monitor nonlinear processes. Ge and Song [40] constructed variable sub-blocks through different directions of principal components, and proposed a distributed PCA. Wang et al. [41] blocked variables based on Kullback-Leibler divergence and variable probability distribution, and proposed a new multi-block PCA. In order to reduce modeling complexity, Wang et al. [42] presented a variable sub-region PCA. The above variable division methods are based on PCA or PLS. Standard PCA and PLS models are static models.

Most process variables have time serial correlations, i.e. dynamic characteristic. To consider process dynamic characteristic, some dynamic process monitoring methods were proposed. Ku et al. [17] used time shift data to augment the measured data matrix and proposed a dynamic PCA based process monitoring method. In addition to data matrix augmentation, dynamic process monitoring research is also focused on SFA and CVA. Shang et al. [26] proposed a SFA based process monitoring method to monitor steady state changes and dynamic changes. For multiphase batch processes with transitions, Zhang et al. [27] presented a global preserving statistics SFA. To achieve adaptive monitoring, Shang et al. [43] proposed a recursive SFA. For closed-loop control systems, Zheng and Zhao [44] developed a performance-relevant full decomposition SFA. Compared to the slow features of SFA, the canonical variables extracted by CVA have clear meanings, i.e. system states. Standard CVA can effectively solve the dynamic characteristic problem and the multivariate correlation problem [45]. Until now, some improved approaches have been presented to solve other process problems, such as nonlinearity, non-Gaussian characteristic, process structure changes, time varying characteristic, and multiple operating conditions. Odiowe and Cao [46] took process nonlinear characteristic into account and employed kernel density estimation to compute CVA statistics’ control limits. Deng and Tian [47] proposed kernel CVA to solve the nonlinear problem. Samuel and Cao [48] presented a new kernel CVA structure to solve the matrix singular problem. In addition, a sparse CVA was proposed by Lu et al. [49] for the covariance matrix singular problem. For process non-Gaussian characteristic, Yang et al. [50] stringed CVA and ICA together and detected fault 3 of the Tennessee Eastman process effectively, which is very difficult to be detected. Jiang et al. [51] proposed a causal dependency feature representation based CVA to monitor process structure changes. For time-varying rotating machines, Li et al. [52] presented an adaptive CVA with a recursively calculated correlation matrix. A mixture CVA was proposed by Wen et al. [53] to solve the multimode problem. However, the CVA that relates to local variable characteristic has not been investigated.

Processes usually have local variable characteristics. For the above CVA methods, all process variables are used to compose a past vector, no matter how different they are. Consequently, some variables may have little contributions
to the process monitoring model. Faults related to those variables become difficult to be detected. To solve this problem, we present a variable sub-region canonical variate analysis (V-CVA) based dynamic process monitoring approach to enhance the information of local variables. Firstly, the proposed V-CVA performs regular CVA on measured variables. Then, divide process variables into sub-regions based on K-means clustering and mutual information of measured variables and canonical variables. Perform CVA in each sub-region, and calculate local monitoring statistics. Lastly, develop ensemble statistics on the basis of local statistics and the idea of Bayesian inference. Our contribution is proposing a dynamic process monitoring method based on variable sub-region canonical variate analysis. For the proposed method, we design a variable blocking strategy based on the mutual information of process variables and canonical variables. The blocking strategy makes use of canonical variable information, and takes process dynamic characteristic into account. We design above variable blocking strategy to improve standard CVA, and make it consider local process variable characteristic and emphasize the information of local variables. We also develop CVA ensemble monitoring statistics. The ensemble monitoring statistics are fusions of local sub-region monitoring results. They are global and sensitive monitoring indexes. The proposed V-CVA is capable of considering process dynamic characteristic, making the information of local variables stand out, and achieving enhanced monitoring performance in the canonical variable feature subspace and the prediction residual subspace. Simulation results confirm the effectiveness of the proposed V-CVA based dynamic process monitoring approach.

The remainder of this contribution is organized as follows. In section II, the CVA based process monitoring method is introduced briefly. Section III presents the proposed V-CVA for dynamic process monitoring. Two case studies are given in section IV to compare the performance of the proposed V-CVA with that of the standard CVA. The concluding remarks are summarized in section V.

II. PROCESS MONITORING BASED ON CANONICAL VARIATE ANALYSIS

A. CANONICAL VARIATE ANALYSIS

Canonical variate analysis is a multivariate statistical analysis method for two sets of variables. Variable dimensionality can be reduced by maximizing a correlation statistic during the CVA processing procedure. When applied to dynamic systems, CVA can extract process states by analyzing the correlations between the past data and future data.

Consider the following linear dynamic system,

$$
x(k + 1) = Ax(k) + Bu(k) + w(k)$$
$$y(k) = Cx(k) + Du(k) + v(k)$$

where $x(k) \in \mathbb{R}^{1 \times 1}$, $y(k) \in \mathbb{R}^{m \times 1}$, and $u(k) \in \mathbb{R}^{n \times 1}$ are the state vector, measured output vector, and measured input vector, respectively. $w(k) \in \mathbb{R}^{1 \times 1}$ and $v(k) \in \mathbb{R}^{m \times 1}$ are the process noise vector and measurement noise vector, respectively. They are independent with one another and obey the zero-mean normal distributions.

Construct a process vector $\tilde{p}(k) = [y(k)T, u(k)T]T$. According to the subspace identification algorithm for the state space model, the state vector $x(k)$ is the linear combination of the past data,

$$x(k) = A_I^T p(k)$$

where $p(k) = [\tilde{p}(k-1)T, \tilde{p}(k-2)T, \ldots, \tilde{p}(k-h)T]T$ is a past vector, $h$ is the number of lags. $A_i = [\alpha_1, \alpha_2, \ldots, \alpha_s]$ is a projection matrix composed of $s$ projection vectors $[33, 34]$. Those projection vectors can be searched by maximizing a correlation $\rho$,

$$\max_{\alpha, \beta} \frac{\alpha^{T}C_{pf}\beta}{\sqrt{\alpha^{T}C_{pp}\alpha}\sqrt{\beta^{T}C_{pf}\beta}}$$

s.t. $\text{var}(x) = \alpha^{T}p) = 1$, $\text{var}(d) = \beta^{T}f = 1$

where $x$ and $d$ are called canonical variables. $\alpha$ and $\beta$ denote projection vectors. $f(k) = [y(k)T, y(k+1)T, \ldots, y(k+l-1)T]T$ is a future vector, $l$ indicates the number of lags. The covariance matrix of $p$ and $f$ are denoted by $C_{pp}$ and $C_{pf}$, respectively. Their cross covariance matrix is denoted by $C_{pf}$. The projection vectors $\alpha$ and $\beta$ can be searched one pair by one pair. They can also be computed by singular value decomposition at once. Use the first $s$ projection vectors $\alpha$ to compose the projection matrix $A_s$. Then, low dimensional canonical variable features can be obtained on the basis of (2).

B. PROCESS MONITORING

Because CVA takes process time serial correlations into account, canonical variables can be utilized as process dynamic features to monitor process operating conditions. At a particular time instant $k$, the $T_2^1$, $T_2^2$ and SPE statistics are established as follows to monitor different subspaces,

$$T_2^1(k) = \frac{p(k)^T A_s A_s^T p(k)}{\sqrt{\text{var}(x)}}$$
$$T_2^2(k) = \frac{p(k)^T A_s A_s^T p(k)}{\sqrt{\text{var}(d)}}$$
$$\text{SPE}(k) = \frac{p(k)^T (I - A_s A_s^T) (I - A_s A_s^T)^T p(k)}{\sqrt{\text{var}(x)}}$$

where $A_s = [\alpha_{s+1}, \alpha_{s+2}, \ldots, \alpha_{h(m+n)}]$ contains the last $h(m+n) - s$ projection vectors of $\alpha$ [33]. The $T_2^2$ statistic is a measure for the low dimensional canonical variable subspace. The $T_2^2$ statistic is applied to monitor the variations outside the canonical variable subspace. The SPE statistic is a measure for the prediction residual space.

Standard CVA based process monitoring method utilizes all process variables to establish a whole model at once, and does not consider local process variable characteristic. The $T_2^s$, $T_2^r$ and SPE statistics in (4) only present global process operating conditions. Local process variable variations may be blanketed. When all process variables are put into one group to model, some process variables may attract few attentions, and may be assigned small contributions to monitoring statistics. For standard CVA, the canonical variable $x_j, j = 1, 2, \ldots, s$, is a linear combination of process
variables. When a process variable becomes abnormal, its weighted item changes obviously. However, that weighted item is added to other weighted items to compute the canonical variable. Consequently, the change of the canonical variable \( x_i \) may not be obvious. It is difficult for traditional process monitoring methods to detect thus faults effectively. How to consider the information of local variables and enlarge some variables’ contributions to monitoring statistics is a challenging problem.

Blocking variables is a common way to consider the information of local variables. References [35], [36], [38] used process prior knowledge to divide variables into meaningful groups. References [37], [39], [42] grouped variables according to the correlations between variables and principal components. Ge et al. [40] divided variables based on their contributions to the selected principal components. Wang et al. [41] applied the Kullback-Leibler divergence of variable probability distribution to block variables. Inspired by the above methods and integrating with CVA characteristics, we design a variable blocking strategy based on the correlations between process variables and canonical variables. Compared with references [35], [36], [38], the proposed strategy does not need process prior knowledge. References [37], [39], [40], [42] are based on PCA. PCA models are static, which do not take process dynamics into account. The proposed strategy is on the basis of CVA, which considers process dynamic characteristic.

### III. VARIABLE SUB-REGION CANONICAL VARIATE ANALYSIS BASED PROCESS MONITORING

To enhance local process variable information, we propose a new dynamic process monitoring method based on V-CVA. First, perform standard CVA on measured variables. Then, divide process variables into sub-regions. Establish a local CVA model and compute local process monitoring statistics for each sub-region. Finally, integrate sub-region statistics to develop ensemble process monitoring statistics. The proposed method utilizes local process monitoring statistics to describe local process variations intensively.

#### A. VARIABLE SUB-REGION DIVISION

The aim of the sub-region division is to heighten the impact of local process variables. The guideline of sub-region division is to make the process variables within a sub-region have strong similarity, and make the process variables from different sub-regions have strong difference. The correlation between one process variable and the canonical variable subspace is different from that of another process variable. Therefore, we adopt correlation intensity as a division index. For the selection of the number of sub-regions, there is no acknowledged and unified method. If the number of variable sub-region is too small, the number of variables in a sub-region may be too large. It is unfavorable to make the information of local variables outstanding. As the number of variable sub-regions increases, the number of the local statistical models will increase. In addition, the selection of the number of variable sub-regions should also take the number of process variables into account. If the number of process variables is not large, a large variable sub-region number is not appropriate. For a multizone tubular reactor for the production of low-density polyethylene, Westerhuis et al. [35] divided variables into two blocks. Li et al. [37] divided variables into principal component related variables and other variables. Cherry and Qin [38] divided correlations into strong, common and weak, qualitatively. In the same light, we partition process variables into three sub-regions: Sub-region 1, strong correlation variables; Sub-region 2, common correlation variables; Sub-region 3, weak correlation variables.

We utilize mutual information to measure the correlation intensity between a process variable and a canonical variable. Mutual information can analyze one random variable’s information that is contained in another random variable. The mutual information of the \( i \) th element of the past vector \( p_i \) and the canonical variable \( x_j \) can be defined as

\[
I_{ev}(p_i, x_j) = \varepsilon \left\{ \lg \frac{f_{p_ix_j}(p_i, x_j)}{f_{p_i}(p_i)f_{x_j}(x_j)} \right\}
\]

where \( i = 1, 2, \cdots, h(m + n) \), \( j = 1, 2, \cdots, s \). \( f_{p_i}(\cdot) \) and \( f_{x_j}(\cdot) \) indicate marginal distribution functions of \( p_i \) and \( x_j \), respectively. \( f_{p_ix_j}(\cdot, \cdot) \) is the joint probability density function of \( p_i \) and \( x_j \). \( \varepsilon \left\{ \cdot \right\} \) denotes ensemble expectation. It can be seen that mutual information can reflect the common information owed by two variables. Mutual information is non-negative. Therefore, using mutual information to describe the correlation between an element in the past vector and a canonical variable is appropriate.

Assume both \( f_{p_i}(\cdot) \) and \( f_{x_j}(\cdot) \) are zero-mean Gaussian distribution function,

\[
f_{p_i}(p_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp \left( -\frac{p_i^2}{2\sigma_i^2} \right)
\]

\( f_{p_ix_j}(\cdot, \cdot) \) is a zero-mean jointly Gaussian distribution function

\[
f_{p_ix_j}(p_i, x_j) = \frac{1}{\sqrt{2\pi |\Sigma|}} \exp \left( -\frac{1}{2} X^T \Sigma^{-1} X \right)
\]

where \( X = \begin{bmatrix} p_i \end{bmatrix} \), \( \Sigma \) is the covariance matrix. It can be derived that

\[
I_{ev}(p_i, x_j) = -\frac{1}{2} \lg \left( 1 - \rho_{ij}^2 \right)
\]

where \( \rho_{ij} \) is a correlation coefficient between \( p_i \) and \( x_j \) [54].

The metric \( I_{ev}(p_i, x_j) \) presents the mutual information of one element of the past vector and one canonical variable. On this basis, the mutual information of the \( i \) th element \( p_i \) of the past vector and the canonical variable subspace is defined as

\[
I_{ev}(p_i) = \sum_{j=1}^{s} I_{ev}(p_i, x_j)
\]
where \( i = 1, 2, \cdots, h(m+n) \). Because the elements of the past vector \( p(k) \) are from \( h \) sample instants, \( h \) elements of \( p(k) \) are from the same process variable. Compute a mutual information summation \( S_{\tau} \),

\[
S_{\tau} = \sum_{q=1}^{h} I_{es} (p_{\tau+(m+n)(q-1)})
\]

where \( \tau = 1, 2, \cdots, m+n \). The mutual information summation \( S_{\tau} \) is corresponding to the \( \tau \) th element of the process vector \( \tilde{p}(k) \). Define \( i' = i \mod (m+n) \), and construct a mutual information summation vector \( S' \),

\[
S'_i = \begin{cases} 
S_{m+n}, & \text{if } i' = 0, \\
S_i, & \text{otherwise.}
\end{cases}
\]

where \( i = 1, 2, \cdots, h(m+n) \). The elements of the mutual information summation vector \( S' \) are corresponding to those of the past vector \( p(k) \). Use mutual information proportion to weight mutual information \( I_{es} \), and obtain weighted mutual information \( I_{esw} \)

\[
I_{esw} (p_i) = \frac{I_{es} (p_i)}{S'_i} I_{es} (p_i) = \frac{[I_{es} (p_i)]^2}{S'_i} \quad (12)
\]

\( I_{esw} (p_i) \) reflects the common information of the \( i \) th element of the past vector and the canonical variable subspace. The larger the variable \( I_{esw} (p_i) \) is, the more their common information is, and the stronger their correlation is. Therefore, we perform variable sub-region division by using weighted mutual information \( I_{esw} (p_i) \).

The elements of the past vector \( p(k) \) are clustered into three sub-sets based on their mutual information \( I_{esw} (p_i) \). K-means clustering is a widely used clustering algorithm. We adopt the K-means clustering algorithm to obtain three mutual information groups. Assume the maximum and minimum of the \( b \) th clustering group are \( I_{esw,\text{max}}(b) \) and \( I_{esw,\text{min}}(b) \), respectively, \( b = 1, 2, 3 \). Two mutual information segment points can be yield by

\[
\xi_1 = \frac{I_{esw,\text{min}}(1) + I_{esw,\text{max}}(2)}{2},
\]

\[
\xi_2 = \frac{I_{esw,\text{min}}(2) + I_{esw,\text{max}}(3)}{2} \quad (13)
\]

According to the segment points, three element sub-sets can be obtained as follows

\[ G_1 = \{ p_i | \xi_1 < I_{esw} (p_i) \leq I_{esw,\text{max}}(1), \ i = 1, 2, \cdots, h(m+n) \} \]

\[ G_2 = \{ p_i | \xi_2 < I_{esw} (p_i) < \xi_1, \ i = 1, 2, \cdots, h(m+n) \} \]

\[ G_3 = \{ p_i | I_{esw,\text{min}}(3) \leq I_{esw} (p_i) \leq \xi_2, \ i = 1, 2, \cdots, h(m+n) \} \]

(14)

For the past vector \( p(k) \), if \( h \) elements corresponding to one process variable are divided into different element sub-sets, all of them are adjusted into one element sub-set according to the latest element. Assume \( y_i \) denotes the \( i \) th output variable, take it for example to describe the element adjustment process. \( y_i(k-1), y_i(k-2), \cdots, y_i(k-h) \) are \( h \) elements of the past vector \( p(k) \). They are corresponding to the same output variable \( y_i \). If \( y_i(k-1) \) belongs to \( G_1 \), and the other \( h-1 \) elements belong to \( G_2 \) or \( G_3 \), we put \( h \) elements into the sub-set \( G_1 \). The adjusted element sub-sets are denoted by \( G'_1, G'_2 \) and \( G'_3 \).

Process variable sub-sets can be obtained according to the adjusted element sub-sets. Assume the adjusted element sub-set \( G'_a \) is corresponding to \( m_a \) process variables, and \( \sum m_a = m+n \). Use the above \( m_a \) process variables to construct a sub-region vector \( g_a = [g_{a,1}, g_{a,2}, \cdots, g_{a,m_a}]^T \), \( a = 1, 2, 3 \). Then, for sub-region \( a \), its local past vector is \( p'_a(k) = [g_{a}(k-1)^T, g_{a}(k-2)^T, \cdots, g_{a}(k-h_1)^T]^T \). The local past vector \( p'_a(k) \) has a strong correlation with the canonical variable subspace. The local past vector \( p'_a(k) \) has a weak correlation with the canonical variable subspace. The correlation intensity between \( p'_a(k) \) and the canonical variable subspace is common and transitional. The variable sub-region division is based on the correlations between process variables and the canonical variable subspace. The canonical variable subspace is obtained by global canonical variable extraction. So, global canonical variable extraction can influence variable sub-region division.

The flowchart of the variable sub-region division is shown in Fig.1. The procedure is listed as follows:

1) Perform CVA on process variables and obtain canonical variables;
2) Calculate the mutual information of elements and canonical variables \( I_{esw} (p_i, x_j) \) according to (8);
3) Compute the weighted mutual information of elements and the canonical subspace \( I_{esw} \) according to (12);
4) Perform K-means clustering on \( I_{esw} \) and determine segment points according to (13);
5) Divide element sub-sets according to (14) and adjust them;
6) Divide variable sub-regions according to \( G'_1, G'_2 \) and \( G'_3 \), and construct sub-region vectors.

**B. MODELING AND MONITORING OF SUB-REGIONS**

After yielding three sub-regions, perform CVA in each sub-region, and compute local monitoring statistics. For sub-region \( a \), its local future vector is defined by \( f'_a(k) = [g_{a}(k)^T, g_{a}(k+1)^T, \cdots, g_{a}(k+l-1)^T]^T, a = 1, 2, 3 \). Conduct CVA on the local past vector \( p'_a(k) \) and future vector \( f'_a(k) \) to obtain a local canonical vector

\[
x'_a(k) = (A'_{s,a})^T f'_a(k) \quad (15)
\]

where \( A'_{s,a} \) is the local projection matrix of reserved canonical variables. Because three local past vectors may have different dimensionalities, the corresponding local canonical vectors may also have different dimensionalities. Local \( T'_s \),

VOLUME 8, 2020
 monitoring statistics are constructed by using the weighted combination of sub-region failure probabilities [20].

At time instant $k$, for the local past vector $p'_a(k)$ of sub-region $a$, the failure posterior probabilities and normal posterior probabilities of $T^2_a$, $T^2_{slim,a}$ and $SPE_a$ statistics are defined as

$$P_{T^2_a}(p'_a(k)|F) = \exp\left(-\frac{T^2_{slim,a}(k)}{T^2_{a}(k)}\right)$$

$$P_{T^2_a}(p'_a(k)|N) = \exp\left(-\frac{T^2_{a}(k)}{T^2_{slim,a}}\right)$$

$$P_{T^2_a}(p'_a(k)|F) = \exp\left(-\frac{T^2_{slim,a}(k)}{T^2_{r,a}(k)}\right)$$

$$P_{SPE_a}(p'_a(k)|F) = \exp\left(-\frac{SPE_{lim,a}}{SPE_a(k)}\right)$$

where $T^2_{slim,a}$, $T^2_{r,a}$ and $SPE_{lim,a}$ denote the control limits of the $T^2_a$, $T^2_{r,a}$ and $SPE_a$ statistics, respectively. Assume the confidence level of each statistic is $\alpha'$. Then, the failure probabilities of $T^2_a$, $T^2_{r,a}$ and $SPE_a$ statistics can be computed by

$$P_{T^2_a}(F|p'_a(k)) = \frac{P_{T^2_a}(p'_a(k)|F) P(F)}{P_{T^2_a}(p'_a(k)|N) P(N) + P_{T^2_a}(p'_a(k)|F) P(F)}$$

$$P_{T^2_a}(F|p'_a(k)) = \frac{P_{T^2_a}(p'_a(k)|F) P(F)}{P_{T^2_a}(p'_a(k)|N) P(N) + P_{T^2_a}(p'_a(k)|F) P(F)}$$

$$P_{SPE_a}(F|p'_a(k)) = \frac{P_{SPE_a}(p'_a(k)|F) P(F)}{P_{SPE_a}(p'_a(k)|N) P(N) + P_{SPE_a}(p'_a(k)|F) P(F)}$$

where failure prior probability $P(F)$ is equal to the confidence level $\alpha'$, and normal prior probability $P(N) = 1 - \alpha'$. Finally, construct ensemble statistics based on sub-region information

$$T^2_{slim,a}(k) = \frac{1}{3} \sum_{a' = 1}^{3} \omega_{T^2_{slim,a}} P_{T^2_a}(F|p'_a(k))$$

$$T^2_{r,a}(k) = \frac{1}{3} \sum_{a' = 1}^{3} \omega_{T^2_{r,a}} P_{T^2_a}(F|p'_a(k))$$

$$SPE_a(k) = \frac{1}{3} \sum_{a' = 1}^{3} \omega_{SPE,a} P_{SPE}(F|p'_a(k))$$

where $\omega_{T^2_{slim,a}}$, $\omega_{T^2_{r,a}}$ and $\omega_{SPE,a}$ denote the weight coefficients of $T^2_{slim,a}$, $T^2_{r,a}$ and $SPE_a$ statistics, respectively.
where the weights of sub-region $a$ are defined as

$$
\omega_{T^2,a} = \exp \left( \frac{T^2_a(k) - T^2_{slim,a}}{T^2_{slim,a}} \right), \quad \omega_{SPE,a} = \exp \left( \frac{SPE_a(k) - SPE_{lim,a}}{SPE_{lim,a}} \right), \quad a = 1, 2, 3
$$

The above weights can highlight the sub-regions which have strong correlations with process faults. When a process fault occurs, the local monitoring statistics of the relevant sub-region should violate their control limits. The relevant sub-region should be assigned a bigger weight. On the contrary, the sub-region which has a weak correlation with the process fault should have a small weight. Hence, on one side, ensemble statistics $T^2$, $T^2_{slim}$ and $SPE$ are global monitoring statistics, which integrate all process variable information. On the other side, ensemble statistics can make the information of local variables stand out, and enlarge fault variables’ contributions.

In summary, the procedure of the proposed V-CVA method is shown in Fig.2. The variables with a similar correlation level are put into one sub-region according to the mutual information of process variables and canonical variables. The sub-region division method is simple. No prior process knowledge is needed during the process of sub-region division. Developing a CVA model in a small sub-region can be useful for emphasizing the information of local variables and describing small process variations, compared to the modeling of all process variables. Hence, sub-region monitoring statistics can reflect the local variations delicately. Ensemble monitoring statistics are constructed by using the idea of Bayesian inference. Ensemble monitoring statistics are global. And they emphasize the information of local variables according to data based sub-region weights.

**IV. SIMULATION RESULTS**

We use a fourth-order dynamic system and a continuous stirred tank reactor to compare the performance of the proposed V-CVA based process monitoring method with that of the traditional CVA based method.

**A. MONITORING FOR THE FOURTH-ORDER DYNAMIC SYSTEM**

Consider the following dynamic system

$$
x(k + 1) = Ax(k) + Bu(k) + w(k)
y(k) = Cx(k) + Du(k) + v(k)
$$

where

$$
x = [x_1, x_2, \ldots, x_4]^T, \quad u = [u_1, u_2]^T,
y = [y_1, y_2, \ldots, y_8]^T, \quad D = 0,
A = \begin{bmatrix} 0.118 & -0.191 & 0.287 & 0.624 \\
-0.847 & 0.264 & 0.943 & 0.438 \\
-0.333 & 0.514 & -0.217 & -0.112 \\
0 & 0.129 & -0.443 & 0.321 \end{bmatrix},
B = \begin{bmatrix} 1 & 2 \\
3 & -4 \\
-2 & 1 \\
1 & 3 \end{bmatrix},
C = \begin{bmatrix} 0.076 & 0.372 & 0.415 & -0.689 \\
-0.409 & 0.272 & 0.217 & 0.514 \\
0.627 & 0.348 & 0.156 & 0.947 \\
0.332 & -0.286 & 0.513 & 0.711 \end{bmatrix}
$$

$u_1 \in [1.4, 1.5]$ and $u_2 \in [0.9, 0.95]$, and they obey uniform distributions. The elements of $w$ and $v$ obey the zero-mean normal distributions with the variance 0.01. A normal operating condition and two failure operating conditions were simulated to produce 500 samples for each condition. Fault 1 is a unit step change in variable $y_3$. Fault 2 is a ramp change in variable $y_2$. The ramp rate is 0.0051. Both of them occurred at the 201st sample time.

For traditional CVA, the number of lags is $h = l = 2$, and the number of retained canonical variables is $k = 7$. The proposed V-CVA has the same number of lags that the traditional CVA has. Process variables and the canonical variable order of each sub-region are listed in Table 1. The confidence level of all statistics is set as $\alpha' = 0.01$.

Fault 1 is a unit step change in variable $y_3$. The results obtained under this failure status are shown in Fig. 3 and 4. For the CVA method, the fault detection time of the $T^2$, $T^2_{slim}$ and $SPE$ statistics is 204, 201, and 203 sample time, respectively. The corresponding fault detection rates are
TABLE 1. Process variables and the canonical variable order of each sub-region.

| Sub-region No. | Process variables | Canonical variable order |
|----------------|-------------------|-------------------------|
| 1              | u₁, y₃            | 2                       |
| 2              | u₂              | 1                       |
| 3              | y₁, y₂, y₃, y₄, y₅, y₆, y₇ | 10                      |

FIGURE 3. CVA based monitoring results under Fault 1 operating condition.

FIGURE 4. V-CVA based monitoring results under Fault 1 operating condition.

TABLE 2. Process parameters of CSTR.

| Parameter               | Value      | Parameter               | Value      | Parameter               | Value      |
|-------------------------|------------|-------------------------|------------|-------------------------|------------|
| Aᵣ (dm³)                | 16.666     | Cₐ (mol/L)              | 0.037      | Cₛ (mol/L)              | 1          |
| E/R (K)                 | 8750       | hₗ (m)                  | 0.6        | k₀ (min⁻¹)              | 7.2×10⁻³   |
| Qₑ (L/min)              | 100        | Qₑ (L/min)              | 15         | Tₑ (K)                  | 402.35     |
| Tₑ (K)                  | 320        | Tₑ (K)                  | 345.44     | Tₛ (K)                  | 300        |
| UAₑ (min·K)             | 5×10⁴      | Vₑ (L)                  | 10         | ρₑCₑ (U/L·K)            | 239        |
| ρₑCₑ (U/L·K)            | 4175       | -ΔH (J/mol)             | 5×10⁴      |                         |            |

84.51%, 100% and 89.56%, respectively. While for the V-CVA method, all of the $T_{s,e}^2$, $T_{re}^2$, and $SPE_e$ statistics detect the fault at 201 sample time. The corresponding fault detection rates are 100%, 100% and 99.66%, respectively. The $T_{s}^2$ and $T_{re}^2$ statistics monitor the variations of retained canonical variables. The $SPE$ and $SPE_e$ statistics reflect the variations of prediction error. Hence, the proposed V-CVA method is superior to the CVA method in monitoring the canonical variable feature subspace and the prediction residual subspace. Fig. 5 shows the monitoring results of all sub-regions. Only the statistics of sub-region 3 signal alarm. Table 1 shows that fault root cause $y_3$ belongs to sub-region 3. It can be seen that the monitoring results are in accordance with the fault mechanism. For the V-CVA method, locally modeling makes the information of local variables stand out, and makes statistics be sensitive to local abnormal variations. But for CVA method, local abnormal variations are submerged by other information to a certain extent.

Fault 2 is a ramp change in variable $y_2$, the ramp rate is 0.0051. Monitoring results are shown in Fig. 6 and 7. For
FIGURE 5. V-CVA sub-region monitoring results under Fault 1 operating condition.

FIGURE 6. CVA based monitoring results under Fault 2 operating condition.

the CVA method, the fault detection time of the $T^2_s$, $T^2_r$ and $SPE$ statistics is 320, 258, and 290 sample time, respectively. For the V-CVA method, the fault detection time of the $T^2_s$, $T^2_r$ and $SPE_e$ statistics is 252, 357 and 276 sample time, respectively. It can be seen that the proposed V-CVA method signals alarm $258 - 252 = 6$ sample time ahead. Fig. 8 shows the monitoring results of each sub-region. It can be seen that only the statistics of sub-region 3 violate their control limits. It is because that fault variable $y_2$ belongs to sub-region 3. Hence, the V-CVA method can enhance the information of local variables and improve monitoring performance.

B. PROCESS MONITORING FOR THE CSTR

The non-isothermal CSTR with a first-order irreversible reaction $A_r \rightarrow B_r$ and a cooling jacket to remove heat was described by [2, 20]

\[
\begin{align*}
\frac{dC_A}{dt} &= -k_0e^{-\frac{E}{RT}} C_A + \frac{Q_F C_{AF} - Q_F C_A}{A_d h_f} \\
\frac{dT}{dt} &= k_0e^{-\frac{E}{RT}} C_A(-\Delta H) + \frac{(Q_F T_F - QT)}{\rho_d C_p} + \frac{A_p h_f}{UA_C(T_C - T)} \\
\frac{dT_C}{dt} &= \frac{Q_C(T_{CF} - T_C)}{\rho C_p C_v} + \frac{UA_C(T - T_C)}{\rho C_p C_v} \\
\frac{dh_f}{dt} &= \frac{Q_F - Q}{A_d}
\end{align*}
\]

(27)

where $A_d$ is the cross section area of reactor; $A_C$ is the heat transfer area; $C_A$ is the concentration of reactant $A_r$; $C_{AF}$ is the feed concentration; $C_P$ is the mass specific heat of
reaction ; $C_{PC}$ is the mass specific heat of coolant ; $E$ is the activation energy; $h_l$ is the reactor liquid level; $k_0$ is the frequency coefficient; $Q$ is the outlet flow rate; $Q_F$ is the feed flow rate; $Q_C$ is the coolant flow rate; $R$ is the gas constant; $T$ is the reactor temperature; $T_F$ is the feed temperature; $T_C$ is the coolant outlet temperature; $T_{CF}$ is the coolant inlet temperature; $U$ is the overall heat transfer coefficient; $V_C$ is the jacket volume; $\rho_d$ is the reactant density; $\rho_C$ is the coolant density; $\Delta H$ is the reaction enthalpy. The CSTR parameters are listed in Table 2. The schematic diagram of CSTR is shown in Fig.9. Cascade PID control is used to control the reactor liquid level and temperature.

4 output variables $[C_A T T_C h_l]^T$ and 6 input variables $[Q_F Q C_A F T_F Q_C T_{CF}]^T$ are used to monitor process operating conditions. A normal operating condition and 6 failure operating conditions were simulated to produce 1000 samples for each condition. A list of process faults is given in Table 3. All faults occurred at the 301th sample time. Faults 1-3 are step changes. Faults 4-6 are ramp changes.
For traditional CVA, parameters $h = l = 3$, and $k = 2$. The weighted mutual information $I_{esw}$ of the elements of the past vector and the canonical variable subspace is shown in Fig. 10. It can be seen that different elements have different mutual information. Hence, different variables should have different contributions to the canonical variable subspace. It is necessary to make the information of local variables stand out. For the proposed V-CVA, process variables and the canonical variable order of each sub-region are listed in Table 4. The confidence level of all statistics is set as $\alpha = 0.01$.

Fault 1 is a bias in the coolant outlet temperature sensor, with a bias $+1.15\text{K}$ for $T_C$. The monitoring results for CVA and V-CVA methods are shown in Fig. 11 and 12, respectively. The $T^2_s$ statistic of the CVA method signals alarm at 302 sample time, and its fault detection rate is 92.09%. While the $T^2_{re}$ statistic of the V-CVA method detects the fault at 301 sample time, and its fault detection rate is 100%. Both of the CVA $T^2_s$ statistic and V-CVA $T^2_{re}$ statistic signal the fault at 301 sample time, and both of their fault detection rates are 100%. Both of the CVA $SPE$ statistic and V-CVA $SPE_e$ statistic signal the fault at 301 sample time. But the fault detection rate of the CVA $SPE$ statistic is 97.70%.

While, the fault detection rate of the V-CVA $T^2_{re}$ statistic is 100%. It can be seen that the fault detection performance is improved. The fault root cause is coolant outlet temperature $T_C$. Coolant outlet temperature changes led to reactor temperature $T$ changes. The closed-loop controller regulated the coolant flow rate $Q_C$. Those process variables belong to sub-region 2. Fig. 13 shows the sub-region monitoring results. It can be seen that three statistics of sub-region 2 violate their control limits obviously. The monitoring results are in accordance with the fault mechanism. Hence, the CVA method can blanket the information of local variables, and the proposed V-CVA can highlight the information of local variables and improve monitoring performance effectively.

Fault 6 is the heat exchanger fouling simulated as a ramp change in the heat transfer coefficient. The ramp rate for $UA_C$ is $-15 \text{(J/(min\textbullet K)/min)}$. The monitoring results for CVA and V-CVA methods are illustrated in Fig. 14 and 15, respectively. For the CVA method, the $T^2_s$, $T^2_r$ and $SPE$ statistics detect the fault at 441, 403 and 416 sample time, respectively.

### Table 3. Fault Settings for the CSTR.

| No. | Operating condition                  | Description and size                  |
|-----|-------------------------------------|--------------------------------------|
| 1   | Bias in coolant outlet temperature sensor | The bias for $T_C$ was $+1.15\text{K}$ |
| 2   | Step change in feed flow rate       | The steps change for $Q_f$ was $+0.2\text{(L/min)}$ |
| 3   | Bias in reactor temperature sensor  | The bias for $T$ was $+0.3\text{K}$   |
| 4   | Ramp change in feed concentration   | The ramp rate for $C_{in}$ was $+2.1\times10^{-4}\text{((mol/L)/min)}$ |
| 5   | Ramp change in catalyst activity    | The ramp rate for $E/R$ was $+2.7\text{ (K/min)}$ |
| 6   | Heat exchanger fouling              | The ramp rate for $UA_C$ was $-15 \text{(J/(min\textbullet K)/min)}$ |

### Table 4. Process variables and the canonical variable order of each sub-region.

| Sub-region No. | Process variables No. | Canonical variable order |
|----------------|-----------------------|--------------------------|
| 1              | 1,4,5,6,7,10          | 2                        |
| 2              | 2,3,9                 | 4                        |
| 3              | 8                     | 2                        |

### Table 5. Fault Detection Sample Time.

| Fault No. | CVA  | V-CVA  |
|-----------|------|--------|
|           | $T^2_s$ | $T^2_r$ | $SPE$ | $T^2_{re}$ | $SPE_e$ |
| 1         | 302   | 301    | 301   | 301   | 301   |
| 2         | 309   | 301    | 324   | 304   | 314   | 304   |
| 3         | 301   | 301    | 301   | 301   | 301   | 301   |
| 4         | 384   | 359    | 372   | 346   | 513   | 340   |
| 5         | 365   | 353    | 351   | 320   | 460   | 318   |
| 6         | 441   | 403    | 416   | 386   | 521   | 372   |
Their fault detection rates are 79.86%, 83.31% and 83.74%, respectively. While for the V-CVA method, the $T^2_s$, $T^2_r$ and $SPE$ statistics signal alarm at 386, 521 and 372 sample time, respectively. Their fault detection rates are 87.63%, 61.44% and 91.37%, respectively. Compared to the CVA method, the proposed V-CVA method detects the fault 403-372 = 31 samples ahead. It can be seen that the monitoring performance of the feature subspace and the prediction residual subspace...
is improved. The sub-region monitoring results are shown in Fig. 16. It can be seen that sub-region 2 is sensitive to the fault. Local variable fault information is enhanced in this sub-region, and global monitoring performance is improved.

The fault detection sample time of all faults is compared in Table 5. It can be seen that for ramp changes such as faults 4, 5, and 6, the proposed V-CVA method detects faults earlier than the CVA method. Take fault 4 for example,
the fault detection sample time of the V-CVA and CVA methods are 340 and 359, respectively. The proposed V-CVA method signals alarm 19 sample time ahead, compared to the CVA method. The fault detection rates of all faults are shown in Table 6. For faults 1, 2, 4, 5 and 6, the fault detection CV A method can detect ramp changes early. Two examples have been shown that the V-CVA method can make the information of local variables stand out, and raise the fault detection rates of the canonical variable feature subspace and the prediction residual subspace. The V-CVA based process monitoring method can detect ramp changes early. Two examples have been used to demonstrate the effectiveness of the proposed V-CVA for dynamic process monitoring. It should be noted that the proposed V-CVA method does not consider process nonlinearity and the dissimilarity between the past and future features [55]. The related studies for the above problems are worth further investigating.

V. CONCLUSION

A variable sub-region canonical variate analysis method has been proposed for dynamic process monitoring. Our motivation for V-CVA is based on the observation that the standard CVA method can blanket the information of local variables, which can influence its fault detection performance. It has been shown that the V-CVA method can make the information of local variables stand out, and raise the fault detection rates of the canonical variable feature subspace and the prediction residual subspace. The V-CVA based process monitoring method can detect ramp changes early. Two examples have been used to demonstrate the effectiveness of the proposed V-CVA for dynamic process monitoring. It should be noted that the proposed V-CVA method does not consider process nonlinearity and the dissimilarity between the past and future features [55]. The related studies for the above problems are worth further investigating.

REFERENCES

[1] L. H. Chiang, R. D. Braatz, and E. L. Russell, Fault Detection and Diagnosis in Industrial Systems. London, U.K.: Springer-Verlag, 2001, pp. 35–98.
[2] M. C. Johannesmeyer, A. Singhal, and D. E. Seborg, “Pattern matching in historical data,” AIChE J., vol. 48, no. 9, pp. 2022–2038, 2006.
[3] Z. Gao, C. Cecati, and S. X. Ding, “A survey of fault diagnosis and fault-tolerant techniques—Part I: Fault diagnosis with model-based and signal-based approaches,” IEEE Trans. Ind. Electron., vol. 62, no. 6, pp. 3757–3767, Jun. 2015.
[4] C. Liu, B. Jiang, R. J. Patton, and K. Zhang, “Integrated fault-tolerant control for close formation flight,” IEEE Trans. Aero. Electron. Syst., to be published.
[5] Q. Liu, Z. Wang, X. He, and D. H. Zhou, “Event-triggered resilient filtering with measurement quantization and random sensor failures: Monotonicity and convergence,” Automatica, vol. 94, pp. 458–464, Aug. 2018.
[6] Y. Li, H. R. Karimi, M. Zhong, S. X. Ding, and S. Liu, “Fault detection for linear discrete time-varying systems with multiplicative noise: The finite-horizon case,” IEEE Trans. Circuits Syst. I, Reg. Papers, vol. 65, no. 10, pp. 3492–3505, Oct. 2018.
[7] S. Liu, B. Jiang, Z. Mao, and S. X. Ding, “Adaptive backstepping based fault-tolerant control for high-speed trains with actuator faults,” Int. J. Control, Autom. Syst., vol. 17, no. 6, pp. 1408–1420, May 2019.
[8] C. G. Siontorou, F. A. Batzias, and V. Tskakiri, “A knowledge-based approach to online fault diagnosis of FET biosensors,” IEEE Trans. Instrum. Meas., vol. 59, no. 9, pp. 2345–2364, Sep. 2010.
[9] Y. Zhao, T. Li, X. Zhang, and C. Zhang, “Artificial intelligence-based fault detection and diagnosis methods for building energy systems: Advantages, challenges and the future,” Renew. Sustain. Energy Rev., vol. 109, pp. 85–101, Jul. 2019.
[10] N. Md Nor, C. R. Che Hassan, and M. A. Hussain, “A review of data-driven fault detection and diagnosis methods: Applications in chemical process systems,” Rev. Chem. Eng., vol. 0, no. 0, Jan. 2019.
[11] M. Friglavudin, F. Khan, S. Imitiaz, and S. Ahmed, “A bibliometric review and analysis of data-driven fault detection and diagnosis methods for process systems,” Ind. Eng. Chem. Res., vol. 57, no. 32, pp. 10719–10735, Jul. 2018.
[12] Z. Ge, Z. Song, S. X. Ding, and B. Huang, “Data mining and analytics in the process industry: The role of machine learning,” IEEE Access, vol. 5, pp. 20590–20616, 2017.
[13] X. Deng, X. Xie, and H. Luo, “A review on basic data-driven approaches for industrial process monitoring,” IEEE Trans. Ind. Electron., vol. 61, no. 11, pp. 6418–6428, Nov. 2014.
[14] S. J. Qin, “Survey on data-driven industrial process monitoring and diagnosis,” Ann. Rev. Control, vol. 36, no. 2, pp. 220–234, 2012.
[15] Y. Zhu, D. Liu, G. Chen, H. Jia, and H. Yu, “Mathematical modeling for active and dynamic diagnosis of crop diseases based on Bayesian networks and incremental learning,” Math. Comput. Model., vol. 58, nos. 3–4, pp. 514–523, Aug. 2013.
[16] B. Luo, H. Wang, H. Liu, B. Li, and F. Peng, “Early fault detection of machine tools based on deep learning and dynamic identification,” IEEE Trans. Ind. Electron., vol. 66, no. 1, pp. 509–518, Jan. 2019.
[17] W. Ku, R. H. Storer, and C. Georgakis, “Disturbance detection and isolation by dynamic principal component analysis,” Chemometrics Intell. Lab. Syst., vol. 30, no. 1, pp. 179–196, Nov. 1995.
[18] J.-M. Lee, C. Yoo, S. W. Choi, P. A. Vanrolleghem, and I.-B. Lee, “Non-linear process monitoring using kernel principal component analysis,” Chem. Eng. Sci., vol. 59, no. 1, pp. 223–234, Jan. 2004.
[19] Q. Jiang, B. Huang, and X. Yan, “GMM and optimal principal components-based Bayesian method for multimode fault diagnosis,” Comput. Chem. Eng., vol. 84, pp. 338–349, Jan. 2016.
[20] X. Deng, X. Tian, S. Chen, and C. J. Harris, “Fault discriminant enhanced kernel principal component analysis incorporating prior fault information for monitoring nonlinear processes,” Chemometrics Intell. Lab. Syst., vol. 162, pp. 21–34, Mar. 2017.
[21] X. Deng and J. Deng, “Incipient fault detection for chemical processes using two-dimensional weighted SLKPCA,” Ind. Eng. Chem. Res., vol. 58, no. 6, pp. 2280–2295, Jul. 2019.
[22] C. K. Yoo, J.-M. Lee, P. A. Vanrolleghem, and I.-B. Lee, “On-line monitoring of batch processes using multiway independent component analysis,” Chemometric Intell. Lab. Syst., vol. 71, no. 2, pp. 151–163, May 2004.
[23] G. Stefatos and A. Ben Hamza, “Dynamic independent component analysis approach for fault detection and diagnosis,” Expert Syst. Appl., vol. 37, no. 12, pp. 8606–8617, Dec. 2010.
[24] C. Bo, X. Qiao, G. Zhang, Y. Bai, and S. Zhang, “An integrated method of independent component analysis and support vector machines for industry distillation process monitoring,” J. Process Control, vol. 20, no. 10, pp. 1133–1140, Dec. 2010.
[25] L. Cai, X. Tian, and S. Chen, “Monitoring nonlinear and non-Gaussian processes using Gaussian mixture model-based weighted kernel independent component analysis,” IEEE Trans. Neural Netw. Learn. Syst., vol. 28, no. 1, pp. 122–135, Jan. 2017.
[26] C. Zhang, F. Yang, X. Gao, X. Huang, J. A. K. Suykens, and D. Huang, “Concurrent monitoring of operating condition deviations and process dynamics anomalies with slow feature analysis,” AIFChE J., vol. 61, no. 11, pp. 3666–3682, Jun. 2015.
[27] H. Zhang, X. Tian, X. Deng, and Y. Cao, “Multiphase batch process with transitions monitoring based on global preserving statistics slow feature analysis,” Neurocomputing, vol. 293, pp. 64–86, Jun. 2018.
[28] L. H. Chiang, E. L. Russell, and R. D. Braatz, “Fault diagnosis in chemical processes using Fisher discriminant analysis, discriminant partial least squares, and principal component analysis,” Chemometric Intell. Lab. Syst., vol. 50, no. 2, pp. 243–252, Mar. 2000.
J. C. Gunther, J. S. Conner, and D. E. Seborg, “Process monitoring and quality variable prediction utilizing PLS in industrial fed-batch cell culture,” J. Process Control, vol. 19, no. 5, pp. 914–921, May 2009.

D. Zhou, G. Li, and S. J. Qin, “Total projection to latent structures for process monitoring,” AIChE J., 2010, vol. 56, no. 1, pp. 168–178.

Q. Jia and Y. Zhang, “Quality-related fault detection approach based on dynamic kernel partial least squares,” Chem. Eng. Res. Des., vol. 106, pp. 242–252, Feb. 2016.

C. Tong, T. Lan, H. Yu, and X. Peng, “Distributed partial least squares based residual generation for statistical process monitoring,” J. Process Control, vol. 75, pp. 77–85, Mar. 2019.

E. L. Russell, L. H. Chiang, and R. D. Braatz, “Fault detection in industrial processes using canonical variate analysis and dynamic principal component analysis,” Chemometric Intell. Lab. Syst., vol. 51, no. 1, pp. 81–93, May 2000.

B. C. Juricek, D. E. Seborg, and W. E. Lariumore, “Fault detection using canonical variate analysis,” Ind. Eng. Chem. Res., vol. 43, no. 2, pp. 458–474, Jan. 2004.

J. A. Westerhuis, T. Kourt, and J. F. MacGregor, “Analysis of multiblock and hierarchical PCA and PLS models,” J. Chemometrics, vol. 12, no. 5, pp. 301–321, Sep. 1998.

S. J. Qin, S. Valle, and M. J. Piovos, “On unifying multiblock analysis with application to decentralized process monitoring,” J. Chemometrics, vol. 15, no. 9, pp. 715–742, 2001.

Y. Li, Z. Xie, and D. H. Zhou, “Fault detection and isolation based on abnormal sub-regions using the improved PCA,” J. Chem. Eng. Jpn., vol. 37, no. 4, pp. 514–522, 2004.

G. A. Cherry and S. J. Qin, “Multiblock principal component analysis based on a combined index for semiconductor fault detection and diagnosis,” IEEE Trans. Semicond. Manuf., vol. 19, no. 2, pp. 159–172, May 2006.

Z. Ge, M. Zhang, and Z. Song, “Nonlinear process monitoring based on linear subspace and Bayesian inference,” J. Process Control, vol. 20, no. 5, pp. 676–688, Jun. 2010.

Z. Ge and Z. Song, “Distributed PCA model for plant-wide process monitoring,” Ind. Eng. Chem. Res., vol. 52, no. 5, pp. 1947–1957, 2013.

B. Wang, Q. Jiang, and X. Yan, “Fault detection and identification using a kullback-leibler divergence based multi-block principal component analysis and Bayesian inference,” Korean J. Chem. Eng., vol. 31, no. 6, pp. 930–943, Apr. 2014.

L. Wang, X. G. Deng, Y. Xu, and N. Zhong, “Fault detection method based on variable sub-region PCA,” CIESC J., vol. 67, no. 10, pp. 4300–4308, 2016.

C. Shang, F. Yang, B. Huang, and D. Huang, “Recursive slow feature analysis for adaptive monitoring of industrial processes,” IEEE Trans. Ind. Electron., vol. 65, no. 11, pp. 8895–8905, Nov. 2018.

J. Zheng and C. Zhao, “Online monitoring of performance variations and process dynamic anomalies with performance-relevant full decomposition of slow feature analysis,” J. Process Control, vol. 80, pp. 89–102, Aug. 2019.

C. Ruiz-Cárceles, L. Lao, Y. Cao, and D. Mba, “Canonical variate analysis for performance degradation under faulty conditions,” Control Eng. Pract., vol. 54, pp. 70–80, Sep. 2016.

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, Oct. 2009.

X. Deng and X. Tian, “Nonlinear process fault diagnosis based on kernel canonical variate analysis,” Control Decis., vol. 21, no. 10, pp. 1109–1113, 2006.

R. T. Samuel and Y. Cao, “Kernel canonical variate analysis for nonlinear dynamic process monitoring,” IFAC-PapersOnLine, vol. 48, no. 8, pp. 605–610, 2015.

Q. Lu, B. Jiang, R. B. Gopalan, 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.

Y. Yang, Y. Chen, X. Chen, and X. Liu, “Multivariate industrial process monitoring based on the integration method of canonical variate analysis and independent component analysis,” Chemometric Intell. Lab. Syst., vol. 116, pp. 94–101, Jul. 2012.

B. Jiang, X. Zhu, D. Huang, and R. D. Braatz, “Canonical variate analysis-based monitoring of process correlation structure using causal feature representation,” J. Process Control, vol. 32, pp. 109–116, Aug. 2015.

X. Li, Y. Yang, I. Bennett, and D. Mba, “Condition monitoring of rotating machines under time-varying conditions based on adaptive canonical variate analysis,” Mech. Syst. Signal Process., vol. 131, pp. 348–363, Sep. 2019.

Q. Wen, Z. Ge, and Z. Song, “Multimode dynamic process monitoring based on mixture canonical variate analysis model,” Ind. Eng. Chem. Res., vol. 54, no. 5, pp. 1605–1614, Jan. 2015.

K. Kumatani, T. Gehrig, U. Mayer, E. Stoimenov, J. McDonough, and M. Wölfel, “Adaptive beamforming with a minimum mutual information criterion,” IEEE Trans. Audio, Speech, Language Process., vol. 15, no. 8, pp. 2527–2541, Nov. 2007.

K. E. S. Pilario and Y. Cao, “Canonical variate dissimilarity analysis for process incipient fault detection,” IEEE Trans. Ind. Informat., vol. 14, no. 12, pp. 5308–5315, Dec. 2018.

YUPING CAO received the B.S. degree in engineering and the Ph.D. degree in control theory and control engineering from the China University of Petroleum (East China), in 2004 and 2010, respectively. She is currently a Lecturer with the College of Information and Control Engineering, China University of Petroleum (East China). Her research interests include process fault diagnosis and prediction.

LEI YU received the B.S. degree in engineering from the Qingdao University of Science and Technology, in 2016. She is currently pursuing the master’s degree with the College of Information and Control Engineering, China University of Petroleum (East China). Her research interest focuses on process fault diagnosis.

XIAOGANG DENG received the B.S. degree in engineering and the Ph.D. degree in chemical engineering and technology from the China University of Petroleum (East China), in 2002 and 2008, respectively. He is currently an Associate Professor with the College of Information and Control Engineering, China University of Petroleum (East China). His research interests include process modeling and simulation, fault diagnosis of industrial processes, and controller performance monitoring.

XIAOLING ZHANG received the B.S. degree in engineering and the master’s degree in control theory and control engineering from the China University of Petroleum (East China), in 2005 and 2008, respectively. She is currently a Lecturer with the College of Shenchl, China University of Petroleum (East China). Her research interest focuses on process fault diagnosis.