IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 11, 2017, accepted December 5, 2017, date of publication December 15, 2017, date of current version March 19, 2018.

Digital Object Identifier 10.1109/ACCESS.2017.2784097

Joint-Individual Monitoring of Parallel-Running Batch Processes Based on MCCA

YANG WANG[1], QINGCHAO JIANG[2], (Member, IEEE), BINBIN LI[1], AND LIZHI CUI[3]

[1]School of Electric Engineering, Shanghai Dianji University, Shanghai 200240, China
[2]Key Laboratory of Advanced Control and Optimization for Chemical Processes, Ministry of Education, East China University of Science and Technology, Shanghai 200237, China
[3]Key Laboratory of Control Engineering of Henan Province, Henan Polytechnic University, Jiaozuo 454003, China

Corresponding author: Qingchao Jiang (qchjiang@ecust.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61603138, in part by the Shanghai Pujiang Program under Grant 17PJD009, in part by the Fundamental Research Funds for the Central Universities under Grant 222201717006 and Grant 222201714027, in part by the Programme of Introducing Talents of Discipline to Universities (the 111 Project) under Grant B17017, and in part by the Open Funding through the Key Laboratory of Control Engineering of Henan Province under Grant KG2016-10.

ABSTRACT A modern production plant may consist of several parallel-running batch processes, and the monitoring of such processes is imperative. This paper proposes a multiset canonical correlation analysis (MCCA)-based joint-individual monitoring scheme for parallel-running batch processes, which considers the individual feature of each batch process and the joint features shared by all batch processes. First, four-way batch process data are unfolded into two-way time-slice data. Second, MCCA is performed at each time instant to extract joint features throughout all running batch processes. Then, for each batch process, the measurements are projected onto a joint feature subspace and its orthogonal complement subspace that contains the individual features of the batch process. Finally, monitoring statistics are constructed to examine the joint and individual features. The proposed monitoring scheme is applied on a numerical example and the simulated parallel-running batch-fed penicillin fermentation processes. Monitoring results show the efficiency of the proposed approach.

INDEX TERMS Parallel-running batch processes, fault detection, multiset canonical correlation analysis, process monitoring.

I. INTRODUCTION

Batch processes play an important role in the chemical, biological, pharmaceutical, and semi-conductor industries [1]–[3]. Nowadays, a production plant may consist of several parallel-running batch processes. These batch processes have intense interactions and communicate with each other, which construct a cyber-physical system (CPS). With the increasing demand in plant safety and product quality, fault detection and diagnosis for this kind of CPS are imperative [4], [5]. Meanwhile, due to the rapid advancement of data collecting, transmitting, and storing techniques, data-driven process monitoring methods have become popular, among which multivariate statistical process monitoring methods are of particular interest [6]–[13].

Numerous multivariate statistical batch process monitoring methods have been developed. In [14], the characteristics of batch processes were summarized, and multivariate statistical process control methods were developed for online monitoring of batch processes. In [15], multiway principal component analysis (MPCA) was proposed to extract information by projecting the data onto low-dimensional spaces defined by the latent variables. In [16], multi-way partial least squares (MPLS) was developed to extract information from process measurement variable trajectories that is more relevant to the final quality variables of the product. The MPCA and MPLS methods establish the basis of batch process monitoring, and several extensions are proposed to address various process characteristics such as uneven length, multiple phases, or multiple operation modes [17]–[20]. Although numerous successful applications are reported, these monitoring methods are limited on an individual batch process. However, a modern production plant may consist of several parallel-running batch processes, and the monitoring of which is imperative.

Aside from PCA and PLS, canonical correlation analysis (CCA) is an alternative classical multivariate analysis

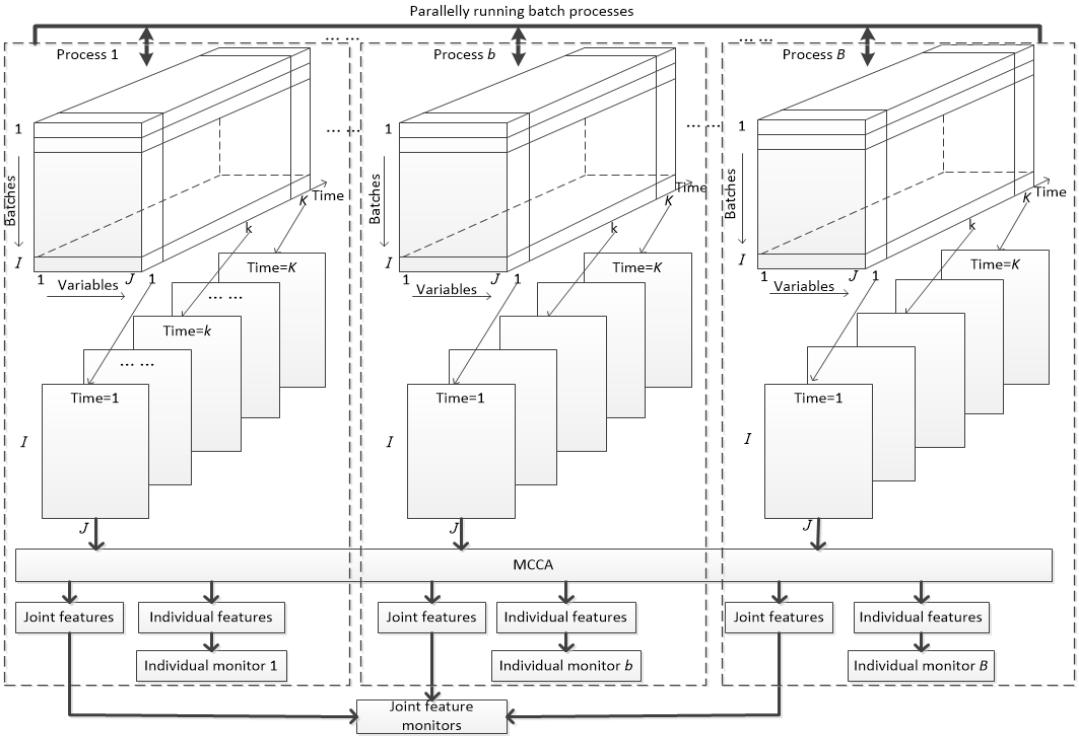

FIGURE 1. Illustration of the data unfolding and the proposed monitoring scheme.

method, which aims to characterize the relationship between two sets of variables. The first applications of CCA on process monitoring have been reported [21], [22]. Recently, a residual generation-based CCA fault detection scheme was proposed, in which CCA is used to characterize the relationship between input and output [23]. Shortly afterwards, a genetic algorithm regularized CCA-based distributed monitoring scheme was proposed, in which CCA is employed to characterize the relationship between different subsystems [24]. The efficiency is theoretically analyzed and experimentally tested. However, conventional CCA efficiently characterizes the relationship between two sets of variables but is limited in handling more than two sets of variables. Recently, a multiset CCA (MCCA)-based joint-individual monitoring scheme for multi-unit processes was proposed, which splits each operation unit into joint feature subspace and individual feature subspace [25]. However, this monitoring scheme was developed for continuous processes that have two-way data and have relatively constant correlations. For batch processes, the correlations may change from time to time at a running cycle, and the usage of MCCA for batch process monitoring is under investigation.

This study develops a MCCA-based joint-individual monitoring scheme to achieve efficient monitoring of parallel-running batch processes. The contributions of the current work can be summarized as follows:

(i) A four-way to two-way data unfolding method is introduced for dealing with the data of parallel-running batch processes. A set of parallel-running batch process data generally has four dimensions, namely, variable dimension, time dimension, batch dimension, and process dimension, as illustrated in Fig. 1. The history data collected from the process include four-way data. First, the four-way data are unfolded along the process wise, and then B sets of three-way data are obtained. For each process, the three-way data are unfolded as time-slice data, as illustrated in Fig. 1.

(ii) Considering the correlation may change from time to time, the MCCA-based joint-individual feature extracting and modeling is performed at each time instant to characterize the relations among the parallel-running batch processes. For the k-th time instant of all batch processes, MCCA is performed to extract the joint features throughout the entire plant. Then, the measurements of each batch process are projected onto a joint feature subspace, which contains the features related to the entire plant and an individual subspace, which consists of individual features that show only the local batch process.

(iii) The superiority of the proposed monitoring scheme is theoretically analyzed within the multivariate statistical framework of hypothesis testing, which enhances the data-driven batch process monitoring theory basis. The efficiency of the proposed fault detection scheme is illustrated by case studies on a numerical example and simulated parallel-running batch-fed penicillin fermentation (FBPF) processes.

The remainder of this article is organized as follows: In Section II, the basics of MCCA are briefly reviewed, and the monitoring problem of parallel-running batch processes is formulated. In Section III, the proposed MCCA-based monitoring scheme for parallel-running batch processes is presented in detail. Some characteristics of the proposed monitoring scheme are also discussed. Then, in Section IV,

application examples on a numerical example and the parallel-running simulated FBPF benchmark processes are provided. Finally, in Section V, conclusions are drawn.

II. PRELIMINARIES AND PROBLEM FORMULATION

In this study, $\mathbb{R}^n$ denotes the n-dimensional Euclidean space, $\mathbb{R}^{n\times m}$ denotes the set of all $n \times m$ matrices, $diag(\ldots)$ denotes a diagonal matrix, $E(\cdot)$ denotes the expectation operator, $\boldsymbol{\Sigma}_{(\cdot)}$ denotes the covariance. $\chi^2(m)$ denotes the chi-square distribution with m degrees of freedom, $rank(\boldsymbol{A})$ denotes the rank of a matrix $\boldsymbol{A}$, and $prob(A)$ denotes the probability of the event A. The other notations are standard.

A. MCCA BASICS

Given two sets of random vectors $\boldsymbol{x}_1 \in \mathbb{R}^{m_1}$ and $\boldsymbol{x}_2 \in \mathbb{R}^{m_2}$ with

$$\begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{bmatrix} \sim N\left(\begin{bmatrix} E(\boldsymbol{x}_1) \\ E(\boldsymbol{x}_2) \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_{\boldsymbol{x}_1} & \boldsymbol{\Sigma}_{\boldsymbol{x}_1\boldsymbol{x}_2} \\ \boldsymbol{\Sigma}_{\boldsymbol{x}_2\boldsymbol{x}_1} & \boldsymbol{\Sigma}_{\boldsymbol{x}_2} \end{bmatrix}\right),$$

CCA finds canonical correlation vectors $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ to maximize the correlation between $\boldsymbol{w}_1^T\boldsymbol{x}_1$ and $\boldsymbol{w}_2^T\boldsymbol{x}_2$, i.e., [25], [26],

$$\begin{aligned}(\boldsymbol{w}_1, \boldsymbol{w}_2) &= \underset{(\boldsymbol{w}_1,\boldsymbol{w}_2)}{\arg\max}\, \rho_{(\boldsymbol{w}_1^T\boldsymbol{x}_1)(\boldsymbol{w}_2^T\boldsymbol{x}_2)} \\ &= \underset{(\boldsymbol{w}_1,\boldsymbol{w}_2)}{\arg\max} \frac{\boldsymbol{w}_1^T\boldsymbol{\Sigma}_{\boldsymbol{x}_1\boldsymbol{x}_2}\boldsymbol{w}_2}{\left(\boldsymbol{w}_1^T\boldsymbol{\Sigma}_{\boldsymbol{x}_1}\boldsymbol{w}_1\right)^{1/2}\left(\boldsymbol{w}_2^T\boldsymbol{\Sigma}_{\boldsymbol{x}_2}\boldsymbol{w}_2\right)^{1/2}}. \end{aligned} \tag{1}$$

The solution can be obtained by performing singular value decomposition on a matrix $\boldsymbol{K}$ as

$$\boldsymbol{K} = \boldsymbol{\Sigma}_{\boldsymbol{x}_1}^{-1/2}\boldsymbol{\Sigma}_{\boldsymbol{x}_1\boldsymbol{x}_2}\boldsymbol{\Sigma}_{\boldsymbol{x}_2}^{-1/2} = \boldsymbol{L}\boldsymbol{\Sigma}\boldsymbol{V}^T, \tag{2}$$

where

$$\boldsymbol{\Sigma} = \begin{bmatrix} diag(\sigma_1, \ldots, \sigma_r) & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \in \mathbb{R}^{m_1\times m_2}$$

and $r = rank(\boldsymbol{\Sigma}_{\boldsymbol{x}_1\boldsymbol{x}_2})$. The canonical correlation vectors can be obtained as [26]

$$\boldsymbol{w}_1 = \boldsymbol{\Sigma}_{\boldsymbol{x}_1}^{-1/2}\boldsymbol{L} \in \mathbb{R}^{m_1\times m_1}, \tag{3}$$

$$\boldsymbol{w}_2 = \boldsymbol{\Sigma}_{\boldsymbol{x}_2}^{-1/2}\boldsymbol{V} \in \mathbb{R}^{m_2\times m_2}. \tag{4}$$

MCCA extends the CCA to more than two set forms. Given B sets of variables with $\boldsymbol{x}_b \in \mathbb{R}^{m_b}$, $b = 1, 2, \ldots, B$, MCCA takes multiple stages to find canonical vectors that maximize the correlation among the canonical variables [25], [27], [28]. Five objective functions for MCCA have been introduced [27], among which the maximum variance method, given its simplicity, is employed in the current study. The method has been demonstrated to be equivalent to solving the following optimization problem [25], [29]:

$$\begin{aligned}(\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_B) = \underset{\boldsymbol{w}_1,\boldsymbol{w}_2,\ldots,\boldsymbol{w}_B}{\arg\max}\ \rho = \sum_{a\neq b}^{B} \boldsymbol{w}_a^T\boldsymbol{x}_a\boldsymbol{x}_b^T\boldsymbol{w}_b \\ s.t.\ \frac{1}{B}\sum_{b=1}^{B} \boldsymbol{w}_b^T\boldsymbol{x}_b\boldsymbol{x}_b^T\boldsymbol{w}_b = 1. \end{aligned} \tag{5}$$

Then, we can obtain the multiset canonical variables as

$$\boldsymbol{z}_b = \boldsymbol{w}_b^T\boldsymbol{x}_b\ (b = 1, \ldots, B), \tag{6}$$

where $\boldsymbol{w}_b = [\,\boldsymbol{w}_b^{(1)}\ \ldots\ \boldsymbol{w}_b^{(D)}\,]^T$ with D denoting the number of extracted joint features. More detailed discussions on the MCCA are provided in [25], [27], and [28].

B. PROBLEM FORMULATION AND MOTIVATION

For a production process that has parallel-running batch processes, the process data include a four-way matrix as $\boldsymbol{X}$ $(I \times J \times K \times B)$, where I is the number of batches, J is the number of measured variables in each process, K is the number of measurements in each operation cycle, and B is the number of batch processes. For process monitoring, the four-way data are first arranged as B sets of three-way data as $\boldsymbol{X}_b$ $(I \times J \times K)$ $(b = 1, \ldots, B)$. Each three-way data can be further unfolded as two-way data, which can be achieved by several unfolding methods. To achieve online monitoring, which means that the process data should be tested in each time instant, the three-way data are unfolded along the time dimension, and then K time-slice data matrices are obtained, as illustrated in Fig. 1. The k-th time slice for the b-th process can be denoted as $\boldsymbol{X}_{b,k}$ $(I \times J)$ $(k = 1, \ldots, K)$. Let the measurement of the b-th process at the k-th time instant of the i-th batch be $\boldsymbol{x}_{b,k,i} = \left[x_{b,k,i,1}, x_{b,k,i,2}, \ldots, x_{b,k,i,J}\right]^T$.

The parallel-running batch processes may share a common power system, pneumatic system, material supplying system, or reaction environment. Therefore, correlations among variables from different batch processes generally exist. T^2 test is a likelihood-ratio test and has been widely used for fault detection [30]. According to the Neyman-Pearson lemma, given a certain false alarm rate (FAR), the T^2 test provides the best fault detectability if no prior fault information is available [30]. Then, to monitor a local process, the following two T^2 tests using different variables can be performed.

(i) T_L^2: T^2 test on only the measured variables in the b-th process, that is,

$$\begin{aligned} T_{L,b}^2(k) &= \left[x_{b,k,i,1}, x_{b,k,i,2}, \ldots, x_{b,k,i,J}\right]^T \\ &\quad \times \boldsymbol{\Sigma}_{b,k}^{-1}\left[x_{b,k,i,1}, x_{b,k,i,2}, \ldots, x_{b,k,i,J}\right] \\ &\sim \chi^2\left(m_{L,b,k}\right), \end{aligned} \tag{7}$$

where $m_{L,b,k} = rank(\boldsymbol{\Sigma}_{b,k})$ and $\boldsymbol{\Sigma}_{b,k}$ is the covariance matrix of the $\boldsymbol{x}_{b,k,i}$. This T^2 test can detect a fault that affects only the local process but may ignore the correlation with other processes.

(ii) T_G^2: T^2 test on all measured variables from all processes. Then, $\boldsymbol{x}_G(k) = [x_{1,k,i,1}, x_{1,k,i,2}, \ldots, x_{1,k,i,J}, x_{2,k,i,1}, x_{2,k,i,2}, \ldots, x_{2,k,i,J}, \ldots, x_{B,k,i,1}, x_{B,k,i,2}, \ldots, x_{B,k,i,J}]^T$. and

$$T_G^2(k) = \boldsymbol{x}_G^T(k)\boldsymbol{\Sigma}_{xG}^{-1}(k)\,\boldsymbol{x}_G \sim \chi^2\left(m_{xG,k}\right), \tag{8}$$

where $m_{xG,k} = rank(\boldsymbol{\Sigma}_{xG}(k))$ and $\boldsymbol{\Sigma}_{xG}(k)$ is the covariance matrix of $\boldsymbol{x}_G(k)$. This T^2 test considers the correlation among processes. However, involving all measured variables

may introduce non-beneficial information into the monitoring, which will relax the control limit and degrade the monitoring performance [6], [24].

III. MCCA-BASED MONITORING FOR PARALLEL-RUNNING BATCH PROCESSES

A. MCCA-BASED FAULT DETECTION

In a continuous process, the correlation is generally kept constant. However, the correlation in a batch process may change time to time. Therefore, the correlation should be analyzed at each time instant. Following the offline modeling and online monitoring procedures, the MCCA-based monitoring scheme is established as follows.

1) OFFLINE MODELING

Step 1: Data preparation

The four-way data $\boldsymbol{X}(I \times J \times K \times B)$ are collected under normal operating condition, and the four-way data are unfolded into two-way time-slice data as illustrated in Fig. 1. First, the data are arranged along the process-wise as $\boldsymbol{X}_b(I \times J \times K)(b = 1, \ldots, B)$. Second, the three-way data are unfolded into time-slice data as $\boldsymbol{X}_{b,k}(I \times J)(k = 1, \ldots, K)$ for each process. Then at each time instant, it is appropriate to assume that the process data are Gaussian distributed with a large number of running cycles [17]. The mean and variance of measurements are calculated at each time slice, and the data at each time instant are mean–variance normalized. It is worth mentioning that since uneven length batches can generally be synchronized through various methods [14], [31], the current study assumes that the batches are synchronized to be with even length.

Step 2: MCCA joint feature extraction at each time instant k

On the basis of MCCA, the canonical vectors for the b-th batch process are obtained as $\boldsymbol{w}_b(k) = [\boldsymbol{w}_b^{(1)}(k) \ldots \boldsymbol{w}_b^{(D)}(k)]^T$ using the time-slice data $\boldsymbol{X}_{b,k}(I \times J)$. These canonical vectors are supposed to be responsible for the joint features throughout the entire process. Then, the d-th canonical variable of the b-th batch process is obtained as

$$z_b^{(d)}(k) = \left(\boldsymbol{w}_b^{(d)}(k)\right)^T \boldsymbol{x}_b(k), \quad (d = 1, 2, \ldots, D; b = 1, 2, \ldots, B). \tag{9}$$

The canonical variables related to the d-th joint feature from all operation units can be obtained as $\boldsymbol{z}^{(d)}(k) = [z_1^{(d)}(k) \ldots z_B^{(d)}(k)]^T (d = 1, 2, \ldots, D)$.

Step 3: Joint feature statistic establishment at each time instant k

For the d-th joint feature, the T^2 statistic $T_{J,d}^2(k)$ can be calculated as

$$T_{J,d}^2(k) = \left(\boldsymbol{z}^{(d)}(k)\right)^T \left(\frac{\boldsymbol{Z}^{(d)}(k)\left(\boldsymbol{Z}^{(d)}(k)\right)^T}{N-1}\right)^{-1} \boldsymbol{z}^{(d)}(k), \tag{10}$$

where $\boldsymbol{Z}^{(d)}(k)$ denotes the d-th joint feature of the training data and N denotes the number of historical training data samples. The corresponding threshold can be calculated as

$$T_{th,J,d}^2(k) = \frac{B(N^2-1)}{N(N-B)} F_\alpha(B, N-B), \tag{11}$$

where $F_\alpha(B, N-B)$ denotes the F distribution with B and N-B degrees of freedom and significant level α.

Step 4: Individual feature extraction at each time instant k

The individual feature in each batch process is obtained by projecting the measurements onto the orthogonal complement subspace of canonical vectors, which can be achieved through the QR decomposition on $\boldsymbol{w}_b(k)$ as

$$\boldsymbol{w}_b(k) = \left[\boldsymbol{Q}_{1,b}(k)\, \boldsymbol{Q}_{2,b}(k)\right] \begin{bmatrix} \boldsymbol{R}_{1,b}(k) \\ 0 \end{bmatrix} = \boldsymbol{Q}_{1,b}(k)\, \boldsymbol{R}_{1,b}(k). \tag{12}$$

where $\boldsymbol{Q}_{2,,b}(k)$ consists of the orthogonal projecting vectors of the individual feature subspace.

Step 5: Individual feature statistic establishment at each time instant k

The T^2 test for the individual features in the b-th batch process can be established as

$$T_{I,b}^2(k) = \boldsymbol{x}_b^T(k)\, \boldsymbol{Q}_{2,b}(k) \times \left(\frac{\boldsymbol{Q}_{2,b}^T(k)\, \boldsymbol{X}_b(k)\, \boldsymbol{X}_b^T(k)\, \boldsymbol{Q}_{2,b}(k)}{N-1}\right)^{-1} \times \boldsymbol{Q}_{2,b}^T(k)\, \boldsymbol{x}_b(k), \tag{13}$$

where $\boldsymbol{X}_b(k)$ denotes the historical training data of the b-th operation unit. The threshold can be obtained as

$$T_{th,I,b}^2(k) = \frac{(m_b - D)(N^2-1)}{N(N-m_b+D)} F_\alpha(m_b - D, N - m_b + D). \tag{14}$$

Step 6: Steps 1 to 5 are repeated to establish the local fault detector for each local process at each time instant.

2) ONLINE MONITORING

Step 7: Online sample scaling

An online measurement at the time instant k is mean–variance scaled. For an online measurement of the b-th process at time instant k, the measurement is mean–variance scaled using the previously obtained mean and variance.

Step 8: Statistics calculation

The scaled measurements are projected onto the joint-feature subspace and individual feature subspace of each batch process. The values of corresponding statistics are calculated.

Step 9: Status determination

The existence of a fault is determined, and the type of the fault is identified based on the detection logic as

$$\begin{cases} T_{J,d}^2(k) \le T_{th,J,d}^2(k) \text{ and } T_{I,b}^2(k) \le T_{th,I,b}^2(k) \Rightarrow \textit{fault-free} \\ T_{J,d}^2(k) > T_{th,J,d}^2(k) \Rightarrow \textit{fault in the d-th joint feature} \\ T_{I,b}^2(k) > T_{th,I,b}^2(k) \Rightarrow \textit{fault in the b-th subprocess.} \end{cases} \tag{15}$$

If a fault is detected by an individual feature statistic, the fault generally affects the corresponding sub-process and the fault location can be identified. Otherwise, if a fault is detected by a joint feature statistic, the fault may affect several processes and spread to the entire process.

B. CHARACTERISTICS OF THE PROPOSED MONITORING SCHEME

Non-detection rate (NDR) and FAR are two important indices used to evaluate the monitoring performance of a proposed method. On the basis of the properties of the NDR, the following propositions on the fault detection performance of the proposed joint-individual monitoring scheme can be derived:

Proposition* 1**: To detect a local fault that affects only the individual features of the b-th batch process, the T^2 test on the individual features, i.e., $T^2_{I,b}$, will provide better monitoring performance than the $T^2_{L,b}$ and T^2_G. The proof is provided in ***Appendix A.

***Proposition* 2**: For a fault that affects only the joint features, the inclusion of the individual features in the T^2 test is not necessary. The derivation is similar to that of ***Proposition* 1**.

Proposition* 3**: For a fault that affects only the joint feature in the b-th batch process, involving canonical variables from other batch processes may provide better monitoring performance than examining the joint feature individually. The proof is provided in ***Appendix B.

It is noted that if the data are unfolded along the batch dimension, considering that a batch run may have multiple phases and the correlation may change from phase to phase in a batch cycle, some phase partition algorithms such as clustering-based algorithms may be needed to deal with the multiple operating conditions [32]–[34]. However, the proposed MCCA-based monitoring scheme unfolds the data along the time dimension, and then the underlying assumption is the correlations among variables and processes keep unchanged at a time instant k in different batch running cycles. Then it is appropriate to assume that the process data at each time slice are Gaussian distributed with a large number of normal running cycles [17].

IV. APPLICATION EXAMPLES

A. APPLICATION ON A NUMERICAL EXAMPLE

In this study, a CPS that consists of three batch processes is employed to test the monitoring performance of the proposed monitoring scheme, which is as follows:

$$\begin{aligned} x_{1,1}(k) &= 50 + k + \varepsilon_1 \\ x_{1,2}(k) &= t_1 - 0.5 * k + \varepsilon_2 \\ x_{2,1}(k) &= 50 + k + \varepsilon_3 \\ x_{2,2}(k) &= t_2 - 0.5 * k + \varepsilon_4 \\ x_{3,1}(k) &= 50 + k + \varepsilon_5 \\ x_{3,2}(k) &= t_3 - 0.5 * k + \varepsilon_6, \end{aligned} \tag{16}$$

in which

$$\begin{bmatrix} t_1 & t_2 & t_3 \end{bmatrix}^T \sim N\left(\begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^T, \begin{bmatrix} 1 & 0.5 & 0.5 \\ 0.5 & 1 & 0.5 \\ 0.5 & 0.5 & 1 \end{bmatrix} \right)$$

and $\varepsilon_1, \ldots, \varepsilon_6$ are Gaussian-distributed independent noise with zero mean and variance of 0.01. Since the source variables t_1, t_2 and t_3 are correlated, the three processes are related. Each batch run consists of 30 samples. To establish the MCCA-based model, process data of 200 batches under normal operating conditions are collected. Two faults are constructed to test the monitoring performance of the proposed monitoring scheme as follows:

Fault 1: A ramp change with an amplitude of $0.05*(k-10)$ is added to $x_{1,1}$ from the 11th point to the 25th point.

Fault 2: A ramp change with an amplitude of $0.05*(k-10)$ is added to $x_{1,2}$ from the 11th point to the 25th point.

Fault 1 is a fault that affects only the individual feature of batch process 1. The monitoring results for fault 1 using $T^2_{J,1}$ (the T^2 test on the joint feature from only process 1), T^2_J (the T^2 test on the joint feature with all processes involved), $T^2_{I,1}$ (the T^2 test on the individual feature from process 1), $T^2_{L,1}$ (the T^2 test on the variables from process 1), and T^2_G (the T^2 test on variables from all processes) are provided in Fig. 2. As shown in Fig. 2, the fault is not detected by the statistics that examine the joint features ($T^2_{J,1}$ and T^2_J). The fault is detected by $T^2_{I,1}$, $T^2_{L,1}$, and T^2_G, among which $T^2_{I,1}$ has the fewest ND points (highlighted by the ellipses and arrows). The monitoring results are consistent with the fault property, i.e., only the individual feature is affected by the fault. $T^2_{L,1}$ and T^2_G have more ND points and worse monitoring performance because they introduce monitoring redundancy, as discussed by ***Proposition* 1**.

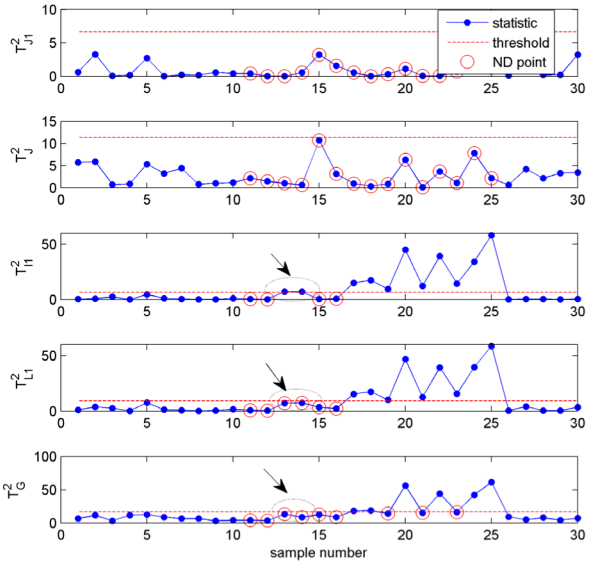

FIGURE 2. Monitoring results for the numerical example fault 1.

Fault 2 is a fault that affects the joint feature of batch process 1, which is related to the other processes.

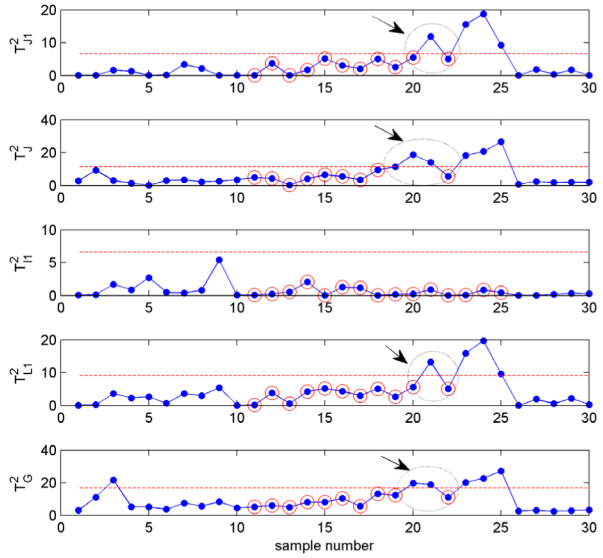

FIGURE 3. Monitoring results for the numerical example fault 2.

The monitoring results for fault 2 using $T^2_{J,1}$,T^2_J,$T^2_{I,1}$, $T^2_{L,1}$, and T^2_G are provided in Fig. 3. The individual feature of process 1 is not affected as no fault point is detected. T^2_J has the fewest ND points and performs the best in detecting the fault (highlighted by the ellipses and arrows). The monitoring results support the points discussed by ***Propositions* 2 *and* 3**. For better comparison, Monte Carlo tests of 1000 times are carried out and the average NDRs for the two faults are presented in Fig. 4. From the Fig.4, it can be seen that $T^2_{I,1}$ performs the best for the fault 1 and the T^2_J performs the best for the fault 2 (with lowest NDR and highlighted by the ellipses and arrows).

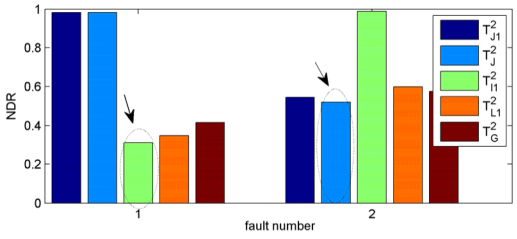

FIGURE 4. Monte Carlo test results for the two faults in the numerical example.

B. APPLICATION ON THE SIMULATED FBPF PROCESS

The FBPF industrial process is a well-known benchmark batch process for testing batch process monitoring schemes [35], [36]. A modular simulator for the fermentation industrial process is developed by the monitoring and control group of the Illinois Institute of Technology (http://mypages.chee.iit.edu/~cinar). A simplified flow diagram of the penicillin fermentation is presented in ***Appendix C*** Fig. 7. In the current work, 16 measured variables are employed for monitoring, which are listed in ***Appendix C*** Table 1. Initial conditions and set points are listed in ***Appendix C*** Table 2. Three batch processes are

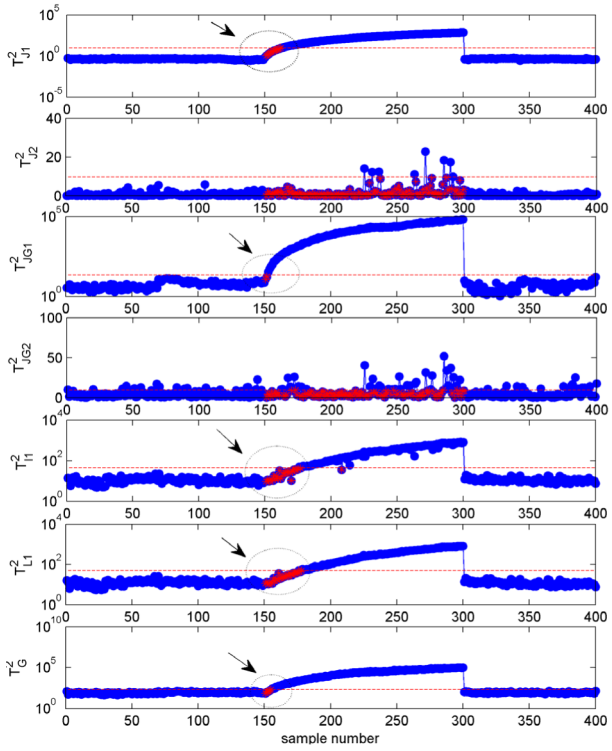

FIGURE 5. Monitoring results for the FBPF fault 1.

simulated simultaneously, and the correlations among batch processes are generated by adding the same Gaussian distributed signal to the initial conditions of aeration rate. The sampling time is 1 h, and the overall duration of each batch is 400 h, including batch and fed-batch stages. Process data of 100 batches are collected. Two joint features are extracted in each batch process, and the MCCA-based monitors are established.

Two different faults are simulated to test the monitoring performance of the proposed monitoring scheme as follows.

Fault 1: A ramp fault of magnitude 20% is introduced into the aeration rate (variable 1) from the 150th to the 300th hour of batch process 1. Since the same Gaussian distributed signal is added to the initial conditions of aeration rate, the fault 1 is a fault that affects the joint feature.

Fault 2: A ramp fault of magnitude 5% is introduced into the agitator power (variable 2) from the 150th to the 300th hour of batch process 1. Since the agitator power runs independently, the fault 2 does not affect the joint feature at the beginning of the fault.

The monitoring results for the fault 1 are presented in Fig. 5. T^2_{J1} denotes the T^2 test on the first joint feature, T^2_{J2} denotes the T^2 test on the second joint feature, T^2_{JG1} denotes the T^2 test on the first joint feature with joint features from the other processes included, T^2_{JG2} denotes the T^2 test on the second joint feature with joint features from the other processes included, T^2_{I1} denotes the T^2 test on the individual features of process 1, T^2_{L1} denotes the T^2 test on the measured variables of process 1, and T^2_G denotes the T^2 test on all measured variables from all processes. At the early stage of

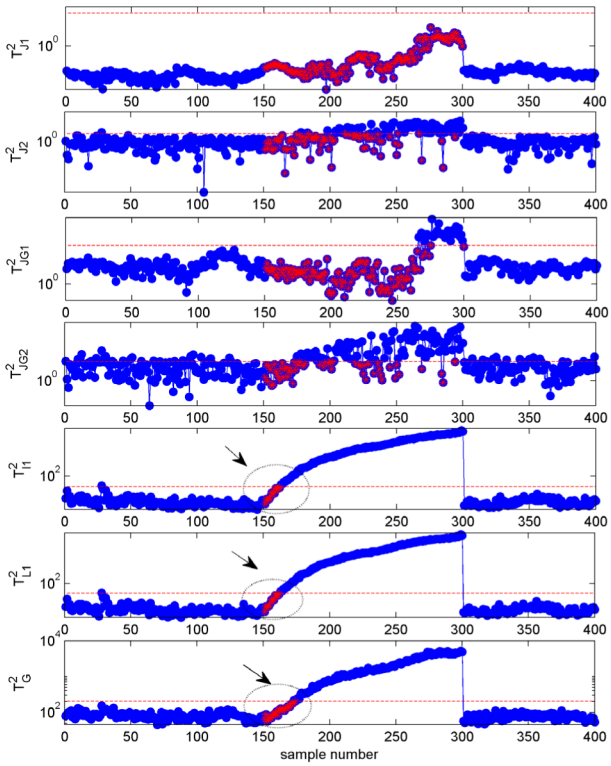

FIGURE 6. Monitoring results for the FBPF fault 2.

the fault, the T^2_{JG1} detected the fault first and then the other statistics, which means that the first joint feature is affected first. The second feature is not affected, and almost no fault point is detected. T^2_{I1} has not detected the fault at the early stage; however, as time goes on, the individual features are affected and the fault is revealed. The results are consistent with the fault property, i.e., the fault affects the joint feature first.

The monitoring results for the fault 2 are presented in Fig. 6. The fault is first detected by T^2_{I1} and then by T^2_{JG2}. That means at the beginning of the fault, the joint features are not affected by the fault. The monitoring results are consistent with the fault property. The efficiency of the proposed joint-individual monitoring scheme is verified.

V. CONCLUSIONS

In this study, a MCCA-based joint-individual monitoring scheme for parallel-running batch processes is proposed. First, the four-way process data are unfolded into two-way time-slice data. Second, MCCA is performed to extract the joint features throughout the entire process. Then, for each batch process, the individual features are obtained by projecting the measurement to the orthogonal complement space of the joint features. Finally, monitoring statistics are constructed to examine the joint and individual features. The efficiency of the proposed monitoring scheme is shown through application examples on a numerical example and the FBPF benchmark process.

It is worth mentioning that this work provided preliminary studies on the fault detection of parallel-running batch processes. Fault isolation and diagnosis can be the subjects of future work. Also, extensions to more complex batch processes such as nonlinear, uneven length or multiple phase batch processes are under investigation.

APPENDIX A
PROOF OF PROPOSITION 1

In the following proofs, we ignore the mark of time instant k when no ambiguity is observed. The following properties of NDR are employed here [24].

Property 1: Given the same degree of freedom m, the NDR is a monotonically decreasing function of the non-central parameter $v \in (0, +\infty)$.

Property 2: Given the same non-central parameter v, the NDR is a monotonically increasing function with the degree of freedom m.

Assume that a fault model of the measurement of the b-th process $\boldsymbol{x}_b$ can be expressed as

$$\boldsymbol{x}_{b,f} = \boldsymbol{x}_{b,N} + \Theta f. \tag{A1}$$

For a local fault that affects only the individual feature of the process, we can obtain $(\Theta f)^T \boldsymbol{Q}_{1,b} = 0$, $(\Theta f)^T \boldsymbol{Q}_{2,b} \neq 0$. The NDR for the $T^2_{I,b}$ test on only the individual features for the b-th operation unit is

$$NDR\left(T^2_{I,b}\right) = F_{\chi^2}\left(T^2_{th,I,b}; m_b - D, v_{I,b}\right), \tag{A2}$$

where $F_{\chi^2}\left(T^2_{th,I,b}; m_b - D, v_{I,b}\right)$ is the cumulative distribution function of the non-central chi-squared distribution with $m_b - D$ degree of freedom and non-central parameter $v_{I,b}$.

$$\begin{aligned} v_{I,b} &= \left(\boldsymbol{Q}^T_{2,b}(\Theta f)\right)^T \boldsymbol{\Sigma}^{-1}_{I,b}\left(\boldsymbol{Q}^T_{2,b}(\Theta f)\right) \\ &= (\Theta f)^T\left(\boldsymbol{Q}_{2,b}\boldsymbol{\Sigma}^{-1}_{I,b}\boldsymbol{Q}^T_{2,b}\right)(\Theta f) \\ &= (\Theta f)^T\left(\boldsymbol{Q}_{2,b}\left(\frac{\boldsymbol{Q}^T_{2,b}\boldsymbol{X}_b\boldsymbol{X}^T_b\boldsymbol{Q}_{2,b}}{N-1}\right)^{-1}\boldsymbol{Q}^T_{2,b}\right)(\Theta f) \\ &= (\Theta f)^T\left(\boldsymbol{Q}_{2,b}\left(\boldsymbol{Q}^T_{2,b}\boldsymbol{\Sigma}_{x_b}\boldsymbol{Q}_{2,b}\right)^{-1}\boldsymbol{Q}^T_{2,b}\right)(\Theta f). \end{aligned} \tag{A3}$$

The T^2 test on all variables in the b-th batch process is

$$T^2_{L,b} = \boldsymbol{x}^T_b\boldsymbol{\Sigma}^{-1}_{x_b}\boldsymbol{x}_b. \tag{A4}$$

The NDR for the $T^2_{L,b}$ is

$$NDR\left(T^2_{L,b}\right) = F_{\chi^2}\left(T^2_{th,b}; m_b, v_b\right), \tag{A5}$$

where

$$\begin{aligned} v_{L,b} &= (\Theta f)^T\boldsymbol{\Sigma}^{-1}_{x_b}(\Theta f) \\ &= (\Theta f)^T\boldsymbol{\Sigma}^{-1}_{x_b}(\Theta f) \\ &= (\Theta f)^T\left(\frac{\boldsymbol{X}_b\boldsymbol{X}^T_b}{N-1}\right)^{-1}(\Theta f) \\ &= (\Theta f)^T\left(\left[\boldsymbol{Q}_{1,b}\boldsymbol{Q}_{2,b}\right]\right. \end{aligned}$$

$$\times \left(\frac{\left[\boldsymbol{Q}_{1,b}\boldsymbol{Q}_{2,b}\right]^T \boldsymbol{X}_b \boldsymbol{X}_b^T \left[\boldsymbol{Q}_{1,b}\boldsymbol{Q}_{2,b}\right]}{N-1} \right)^{-1}$$
$$\times \left[\boldsymbol{Q}_{1,b}\boldsymbol{Q}_{2,b}\right]^T \Big) (\Theta f)$$
$$= \left[\mathbf{0}\,(\Theta f)^T \boldsymbol{Q}_{2,b}\right] \left(\begin{bmatrix} \boldsymbol{Q}_{1,b}^T \\ \boldsymbol{Q}_{2,b}^T \end{bmatrix} \boldsymbol{\Sigma}_{x_b} \left[\boldsymbol{Q}_{1,b}\boldsymbol{Q}_{2,b}\right] \right)^{-1}$$
$$\times \begin{bmatrix} \mathbf{0} \\ \boldsymbol{Q}_{2,b}^T (\Theta f) \end{bmatrix}$$
$$= \left[\mathbf{0}\,(\Theta f)^T \boldsymbol{Q}_{2,b}\right] \begin{pmatrix} \boldsymbol{Q}_{1,b}^T \boldsymbol{\Sigma}_{x_b} \boldsymbol{Q}_{1,b} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{Q}_{2,b}^T \boldsymbol{\Sigma}_{x_b} \boldsymbol{Q}_{2,b} \end{pmatrix}^{-1}$$
$$\times \begin{bmatrix} \mathbf{0} \\ \boldsymbol{Q}_{2,b}^T (\Theta f) \end{bmatrix}$$
$$= (\Theta f)^T \left(\boldsymbol{Q}_{2,b} \left(\boldsymbol{Q}_{2,b}^T \boldsymbol{\Sigma}_{x_b} \boldsymbol{Q}_{2,b} \right)^{-1} \boldsymbol{Q}_{2,b}^T \right) (\Theta f). \quad \text{(A6)}$$

Evidently, the non-central parameters are $v_{L,b} = v_{I,b}$. The T^2 test on all variables in the entire process is

$$T_G^2 = \boldsymbol{x}_G^T \boldsymbol{\Sigma}_{x_G}^{-1} \boldsymbol{x}_G. \quad \text{(A7)}$$

The NDR for the T_G^2 is

$$NDR\left(T_G^2\right) = F_{\chi^2}\left(T_{th,G}^2; m_G, v_G\right). \quad \text{(A8)}$$

Thus, $v_{L,b} = v_{I,b} = v_G$. The degree of freedom is $m_G \geq m_b \geq m_b - D$. According to ***Property 2***, the NDR of the $T_{I,b}^2$ will be smaller than that of the $T_{L,b}^2$ and T_G^2 for the batch process.

APPENDIX B
PROOF OF PROPOSITION 3

Assume that a fault affects only a joint feature in the b-th process, i.e., $(\Theta f)^T \boldsymbol{Q}_{1,b} \neq 0$, $(\Theta f)^T \boldsymbol{Q}_{2,b} = 0$. The T^2 test on the canonical variables in the b-th batch process is

$$T_{J,d,b}^2 = \boldsymbol{x}_b^T \boldsymbol{w}_b^{(d)} \left(\boldsymbol{\Sigma}_{Z_{b,d}}\right)^{-1} \left(\boldsymbol{w}_b^{(d)}\right)^T \boldsymbol{x}_b^T, \quad \text{(A9)}$$

and the non-central parameter for the NDR is

$$v_{J,d,b} = (\Theta f)^T \left(\boldsymbol{w}_b^{(d)} \left(\boldsymbol{\Sigma}_{z_{b,d}}\right)^{-1} \left(\boldsymbol{w}_b^{(d)}\right)^T \right) (\Theta f)$$
$$= (\Theta f)^T \left(\boldsymbol{w}_b^{(d)} \left(\boldsymbol{\Sigma}_{z_{b,d}}\right)^{-1} \left(\boldsymbol{w}_b^{(d)}\right)^T \right) (\Theta f). \quad \text{(A10)}$$

The non-central parameter for NDR of the T^2 test that involves canonical variables from all operation units is

$$v_J = \left[(\Theta f)^T \boldsymbol{w}_b^{(d)} \; \mathbf{0}^T \right] \begin{pmatrix} \boldsymbol{\Sigma}_{z_{b,d}} & \boldsymbol{\Sigma}_{z_{b,d},z_{\sim b,d}} \\ \boldsymbol{\Sigma}_{z_{\sim b,d},z_{b,d}} & \boldsymbol{\Sigma}_{z_{\sim b,d}} \end{pmatrix}^{-1}$$
$$\times \begin{bmatrix} \left(\boldsymbol{w}_b^{(d)}\right)^T (\Theta f) \\ \mathbf{0} \end{bmatrix}$$
$$= (\Theta f)^T \boldsymbol{w}_b^{(d)} \left[\boldsymbol{\Sigma}_{z_{b,d}} - \boldsymbol{\Sigma}_{z_{b,d},z_{\sim b,d}} \boldsymbol{\Sigma}_{z_{\sim b,d}}^{-1} \boldsymbol{\Sigma}_{z_{\sim b,d},z_{b,d}} \right]^{-1}$$
$$\times \left(\boldsymbol{w}_b^{(d)}\right)^T (\Theta f). \quad \text{(A11)}$$

To compare the non-central parameters $v_{J,d,b}$ and v_J, the difference between them can be derived as

$$v_J - v_{J,d,b} = (\Theta f)^T \boldsymbol{w}_b^{(d)}$$
$$\times \left[\left(\boldsymbol{\Sigma}_{z_{b,d}} - \boldsymbol{\Sigma}_{z_{b,d},z_{\sim b,d}} \boldsymbol{\Sigma}_{z_{\sim b,d}}^{-1} \boldsymbol{\Sigma}_{z_{\sim b,d},z_{b,d}} \right)^{-1} - \left(\boldsymbol{\Sigma}_{z_{b,d}}\right)^{-1} \right] \left(\boldsymbol{w}_b^{(d)}\right)^T (\Theta f). \quad \text{(A12)}$$

$[(\boldsymbol{\Sigma}_{z_{b,d}} - \boldsymbol{\Sigma}_{z_{b,d},z_{\sim b,d}} \boldsymbol{\Sigma}_{z_{\sim b,d}}^{-1} \boldsymbol{\Sigma}_{z_{\sim b,d},z_{b,s}})^{-1} - .(\boldsymbol{\Sigma}_{z_{b,d}})^{-1}]$ is positive semidefinite, and the $v_J - v_{J,d,b}$ is a quadratic form. Therefore, we can have $v_J - v_{J,d,b} \geq 0$ and then $v_J \geq v_{J,d,b}$. Apparently, involving the canonical variables from other processes may increase the non-central parameter. According to ***Property 1*** of the NDR, the monitoring performance may be improved.

APPENDIX C
INTRODUCTION TO THE FBPF PROCESS

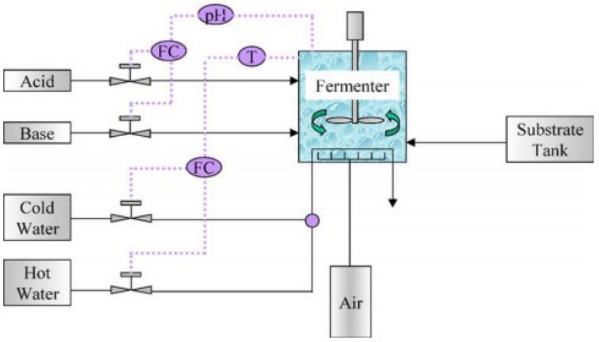

FIGURE 7. Flowsheet of the penicillin cultivation process [35], [36].

TABLE 1. Measured variables of the fed-batch penicillin fermentation process [35], [36].

Variable no.	Variable description
1	Aeration rate (L/h)
2	Agitator power (W)
3	Substrate feed rate (L/h)
4	Substrate feed temperature (K)
5	Substrate concentration (g/L)
6	Dissolved oxygen concentration (g/L)
7	Biomass concentration (g/L)
8	Penicillin concentration (g/L)
9	Culture volume (L)
10	Carbon dioxide concentration (g/L)
11	pH
12	Bioreactor temperature (K)
13	Generated heat (kcal)
14	Acid flow rate (L/h)
15	Base flow rate (L/h)
16	Cooling water flow rate (L/h)

TABLE 2. Initial conditions and set points to simulate the penicillin cultivation process [35], [36].

Initial conditions	Range
Substrate concentration (g/L)	15-17
Dissolved oxygen concentration (mmol/L)	1-1.2
Biomass concentration (g/L)	0.1
Penicillin concentration (g/L)	0
Culture volume (L)	101-103
Carbon dioxide concentration (mmol/L)	0.5-1
Hydrogen ion concentration: $[H^+]$ (mol/L)	10^{-5}
Bioreactor temperature (K)	297-299
Generated heat (kcal)	0
Set points	
Aeration rate (g/L)	8-9
Agitator power (W)	29-31
Substrate feed flow rate (L/h)	0.039-0.045
Substrate feed temperature (K)	295-296
Bioreactor temperature (K)	297-298
PH	4.95-5.05

ACKNOWLEDGMENT

The authors are grateful to the Associate Editor and anonymous reviewers for their constructive comments and suggestions based on which the paper has been greatly improved.

REFERENCES

[1] D. Bonvin, "Control and optimization of batch processes," *IEEE Control Syst.*, vol. 26, no. 6, pp. 34–45, Dec. 2014.
[2] J. Lu, Z. Cao, and F. Gao, "Batch process control-overview and outlook," *Acta Autom. Sinica*, vol. 43, no. 7, pp. 933–943, 2017.
[3] S. Yin, S. X. Ding, A. H. Abandan Sari, and H. Y. Hao, "Data-driven monitoring for stochastic systems and its application on batch process," *Int. J. Syst. Sci.*, vol. 44, no. 7, pp. 1366–1376, Jul. 2013.
[4] Q. Jiang, B. Huang, S. X. Ding, and X. Yan, "Bayesian fault diagnosis with asynchronous measurements and its application in networked distributed monitoring," *IEEE Trans. Ind. Electron.*, vol. 63, no. 10, pp. 6316–6324, Oct. 2016.
[5] Y. Jiang and S. Yin, "Recursive total principle component regression based fault detection and its application to vehicular cyber-physical systems," *IEEE Trans. Ind. Informat.*, to be published, doi: 10.1109/TII.2017.2752709
[6] Q. Jiang and B. Huang, "Distributed monitoring for large-scale processes based on multivariate statistical analysis and Bayesian method," *J. Process Control*, vol. 46, pp. 75–83, Oct. 2016.
[7] Z. Ge and J. Chen, "Plant-wide industrial process monitoring: A distributed modeling framework," *IEEE Trans. Ind. Informat.*, vol. 12, no. 1, pp. 310–321, Feb. 2017.
[8] S. Yin, S. X. Ding, 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.
[9] S. X. Ding, "Data-driven design of monitoring and diagnosis systems for dynamic processes: A review of subspace technique based schemes and some recent results," *J. Process Control*, vol. 24, no. 2, pp. 431–449, 2014.
[10] Y. Wang, Q. Jiang, and J. Fu, "Data-driven optimized distributed dynamic PCA for efficient monitoring of large-scale dynamic processes," *IEEE Access*, vol. 5, pp. 18325–18333, 2017.
[11] 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.
[12] C. Sankavaram, A. Kodali, K. R. Pattipati, and S. Singh, "Incremental classifiers for data-driven fault diagnosis applied to automotive systems," *IEEE Access*, vol. 3, pp. 407–419, 2015.
[13] S. Yin, X. Xie, and W. Sun, "A nonlinear process monitoring approach with locally weighted learning of available data," *IEEE Trans. Ind. Electron.*, vol. 64, no. 2, pp. 1507–1516, Feb. 2017.
[14] P. Nomikos and J. F. MacGregor, "Multivariate SPC charts for monitoring batch processes," *Technometrics*, vol. 37, no. 1, pp. 41–59, 1995.
[15] P. Nomikos and J. F. MacGregor, "Monitoring batch processes using multiway principal component analysis," *AIChE J.*, vol. 40, pp. 1361–1375, 1994.
[16] P. Nomikos and J. F. MacGregor, "Multi-way partial least squares in monitoring batch processes," *Chemom. Intell. Lab. Syst.*, vol. 30, pp. 97–108, Nov. 1995.
[17] Y. Qin, C. Zhao, and F. Gao, "An iterative two-step sequential phase partition (ITSPP) method for batch process modeling and online monitoring," *AIChE J.*, vol. 62, no. 7, pp. 2358–2373, 2016.
[18] C. Zhao, "Concurrent phase partition and between-mode statistical analysis for multimode and multiphase batch process monitoring," *AIChE J.*, vol. 60, no. 2, pp. 559–573, 2014.
[19] C. Zhao, S. Mo, F. Gao, N. Lu, and Y. Yuan, "Statistical analysis and online monitoring for handling multiphase batch processes with varying durations," *J. Process Control*, vol. 21, no. 6, pp. 817–829, 2011.
[20] H. Zhang, X. Tian, and X. Deng, "Batch process monitoring based on multiway global preserving kernel slow feature analysis," *IEEE Access*, vol. 5, pp. 2696–2710, 2017.
[21] L. H. Chiang, E. Russell, and R. D. Braatz, *Fault Detection and Diagnosis in Industrial Systems*. London, U.K.: Springer-Verlag, 2001.
[22] B. Jiang, D. Huang, X. Zhu, F. Yang, and R. D. Braatz, "Canonical variate analysis-based contributions for fault identification," *J. Process Control*, vol. 26, pp. 17–25, Feb. 2015.
[23] Z. Chen, S. X. Ding, K. Zhang, Z. Li, and Z. Hu, "Canonical correlation analysis-based fault detection methods with application to alumina evaporation process," *Control Eng. Pract.*, vol. 46, pp. 51–58, Jan. 2016.
[24] Q. Jiang, S. X. Ding, Y. Wang, and X. Yan, "Data-driven distributed local fault detection for large-scale processes based on the GA-regularized canonical correlation analysis," *IEEE Trans. Ind. Electron.*, vol. 64, no. 10, pp. 8148–8157, Oct. 2017.
[25] Y. Wang, Q. Jiang, X. Yan, and J. Fu, "Joint-individual monitoring of large-scale chemical processes with multiple interconnected operation units incorporating multiset CCA," *Chemom. Intell. Lab. Syst.*, vol. 166, pp. 14–22, Jul. 2017.
[26] W. Härdle and L. Simar, *Applied Multivariate Statistical Analysis*, 4th ed. Berlin, Germany: Springer, 2007.
[27] J. R. Kettenring, "Canonical analysis of several sets of variables," *Biometrika*, vol. 58, no. 3, pp. 433–451, 1971.
[28] Y.-O. Li, T. Adalı, W. Wang, and V. D. Calhoun, "Joint blind source separation by multiset canonical correlation analysis," *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 3918–3929, Oct. 2009.
[29] J. Vía, I. Santamaría, and J. Pérez, "Canonical correlation analysis (CCA) algorithms for multiple data sets: Application to blind SIMO equalization," in *Proc. 13th Eur. Signal Process. Conf.*, 2005, pp. 1–4.
[30] S. X. Ding, *Data-Driven Design of Fault Diagnosis and Fault-Tolerant Control Systems*. London, U.K.: Springer, 2014.
[31] J. M. González-Martínez, O. E. de Noord, and A. Ferrer, "Multisynchro: A novel approach for batch synchronization in scenarios of multiple asynchronisms," *J. Chemom.*, vol. 28, no. 5, pp. 462–475, 2014.
[32] N. Lu, F. Gao, and F. Wang, "Sub-PCA modeling and on-line monitoring strategy for batch processes," *AIChE J.*, vol. 50, no. 1, pp. 255–259, 2004.
[33] C. Zhao, F. Wang, N. Lu, and M. Jia, "Stage-based soft-transition multiple PCA modeling and on-line monitoring strategy for batch processes," *J. Process Control*, vol. 17, no. 9, pp. 728–741, Oct. 2007.
[34] Y. Yao and F. Gao, "Phase and transition based batch process modeling and online monitoring," *J. Process Control*, vol. 19, no. 5, pp. 816–826, 2009.
[35] G. Birol, C. Ündey, and A. Çinar, "A modular simulation package for fed-batch fermentation: Penicillin production," *Comput. Chem. Eng.*, vol. 26, no. 11, pp. 1553–1565, 2002.
[36] J.-M. Lee, C. K. Yoo, and I.-B. Lee, "Enhanced process monitoring of fed-batch penicillin cultivation using time-varying and multivariate statistical analysis," *J. Biotechnol.*, vol. 110, no. 2, pp. 119–136, 2004.

YANG WANG received the B.E. degree from the Department of Automation, Shenyang University of Technology, Shenyang, China, in 2006, the M.E. degree in measurement and control technology and instrument from Shanghai Maritime University, Shanghai, China, in 2008, and the Ph.D. degree in control science and engineering from Shanghai University, Shanghai, in 2017. She is currently a Lecturer with Shanghai Dianji University, Shanghai. Her research interests include data-driven fault detection and diagnosis, process monitoring, and wireless sensor networks.

QINGCHAO JIANG (M'17) received the B.E. and Ph.D. degrees from the Department of Automation, East China University of Science and Technology, Shanghai, China, in 2010 and 2015, respectively. In 2015, he was a Post-Doctoral Fellow with the Department of Chemical and Materials Engineering, University of Alberta, Canada. From 2015 to 2016, he was a Humboldt Research Fellow with the Institute for Automatic Control and Complex Systems, University of Duisburg-Essen, Germany. He is currently an Associate Professor with the East China University of Science and Technology, Shanghai, China. His research interests include data mining and analysis, soft sensing, multivariate statistical process monitoring, and Bayesian fault diagnosis.

BINBIN LI received the B.E. degree in mechanical design, manufacture and automatization from Hunan University, Changsha, China, in 2003, and the Ph.D. degree in measurement and control technology and instrument from Sichuan University, Chendu, China, in 2008. He is currently a Lecturer with Shanghai Dianji University, Shanghai, China. His research interests include embedding systems, mechatronics system, and signal processing theory.

LIZHI CUI received the Ph.D. degree in control science and engineering from East China University of Science and Technology, Shanghai, China, in 2015. From 2012 to 2014, he was a Co-Cultured Doctoral Fellow with the School of Information Technologies, The University of Sydney, Australia. Since 2006, he has been a Lecturer with the School of Electrical Engineering and Automation, Henan Polytechnic University, China. His research interests include signal processing, machine learning, and control engineering.