

Deviation Contribution Plots of Multivariate Statistics

Ruomu Tan¹ and Yi Cao, *Senior Member, IEEE*

Abstract—As data analytic techniques evolve and the accessibility of process measurements improves, data-driven process monitoring has enjoyed a quick development in both theoretical and application perspectives recently. Although abundant process measurements will facilitate data-driven process monitoring and lead to better monitoring indexes, it becomes difficult to identify the underlying variables that are responsible for a fault directly with the monitoring indexes as the scope of measured variables is getting broader. To restrain the scope and identify the source of fault, contribution plots are commonly used in fault diagnosis in order to quantify the influence of process variables in presence of fault. Nevertheless, as sophisticated monitoring techniques become more and more complicated, deriving corresponding contribution plots is challenging. The concept of deviation contribution plots is proposed to address this issue. By extending the original definition of contribution for linear processes, the deviation contribution is defined to quantify the contribution of deviations in originally measured variables to the deviation of monitoring indexes. The ability of the proposed deviation contribution plots to identify influential variables in monitoring algorithms based on nonlinear feature extractions is verified by both numerical simulation and the Tennessee Eastman process benchmark case study.

Index Terms—Contribution, fault diagnosis, feature extraction, multivariate statistical process monitoring.

I. INTRODUCTION

FAULT diagnosis is an indispensable step in process monitoring. When a fault has been detected, its properties, such as type, location, severity, and scope are of great interest if proper action is to be taken to sustain reliable and favorable process performance. In data-driven process monitoring, fault diagnosis further analyzes relevant information about the fault so as to characterize it and provide more specified knowledge for process operation, decision support, maintenance schedul-

ing, etc. Faults can be characterized from different aspects using various fault diagnosis algorithms. For example, data-driven classification methods have been utilized to recognize patterns of behavior in process data and the fault can be detected and categorized simultaneously. These include Fisher discriminant analysis [1], support vector machine, Gaussian mixture models [2], and their extensions [3], [4], to name a few. On the other hand, the likelihood of fault occurrence in chosen piece of equipment can be calculated by probabilistic approaches including Bayesian network [5] and Bayesian inference [6]. Causal map of process variables has also been integrated for fault isolation [7]. Another major task of fault diagnosis is to identify influential variables with significant contributions to faults detected. Performance of data-driven fault detection may be enhanced as the number of process variables increases; meanwhile, a large amount of variables will make locating and identifying the fault challenging. Hence, spotting process variables related to a fault is an essential stage in multivariate statistical fault diagnosis. As one of the most widely used approaches, contribution plots quantify the influence of process variables to the monitoring statistics concluded in the multivariate statistical fault detection step [8]. The mathematical formulation of contribution has made it easy to be combined with different multivariate statistical fault detection algorithms developed for handling various complexities, including multiway principal component analysis [9] for batch process data, independent component analysis [10] for handling non-Gaussianity, and canonical variate analysis for handling temporal correlation [11]. In general, contribution analysis can be exempt from complexity and quality issues in the data owing to its compatibility with data preprocessing and feature extraction methods applied in the fault detection stage.

Nevertheless, it has been pointed out that the extension of standard contribution plots is difficult [12] when nonlinear feature extraction approaches, such as kernel principal component analysis (PCA), are adopted. In order to accommodate kernel-based multivariate statistical methods in nonlinear process monitoring, contribution rate, which is defined as the first-order derivative of monitoring statistics with respect to original variables, is proposed and used for estimating the velocity of change in monitoring statistics caused by variation in each process variable [13], [14]. The concept of reconstruction-based contribution (RBC) has also been applied to determine the fault direction and magnitude [12] for nonlinear process monitoring with generalization to fault-specific [15] and combination with variable selection methods, such as Branch and Bound for further isolation [16]. However, neither the contri-

Manuscript received January 12, 2018; revised March 27, 2018; accepted April 20, 2018. This work was supported by the European Union's Horizon 2020 research and innovation programme under grant agreement No. 675215 - PRONTO - H2020-MSCA-ITN-2015. Paper no. TII-18-0105. (Corresponding author: Ruomu Tan.)

R. Tan was with the School of Water, Energy and Environment, Cranfield University, Cranfield MK43 0AL, U.K., and is now with the Centre for Process Systems Engineering, Imperial College London, London SW7 2AZ, U.K. (e-mail: r.tan@imperial.ac.uk).

Y. Cao is with the College of Chemical and Biological Engineering, Zhejiang University, Hangzhou 310027, China (e-mail: caoyi2018@zju.edu.cn).

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

Digital Object Identifier 10.1109/TII.2018.2841658

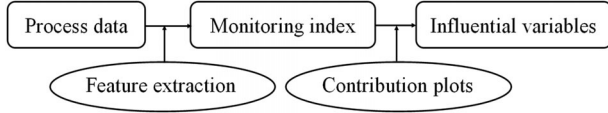


Fig. 1. General structure of the fault detection and diagnosis algorithm.

tribution rate nor RBC provides a direct quantification of the contribution of process variables to the monitoring statistics for nonlinear process monitoring as the standard contribution plots do under linear circumstance. Further challenges emerge when nonparametric and probabilistic monitoring algorithms are introduced [17], [18] as both detection and diagnosis procedures are different from those in multivariate statistical approaches. To conclude, disparity still exists in the mathematical formulation of different contribution-related definitions and there lacks the contribution definition which is derived directly from standard contribution plots. In addition, as the feature extraction approach used in monitoring statistics calculation gets more sophisticated, computation of monitoring statistics is no longer straightforward and contributions are transmitted through multiple layers of feature extraction, some of which are mathematically complicated; hence, the original definition of contribution and way of estimating it should also be modified to account for advanced feature extraction algorithms used in fault detection.

In this paper, the concept of deviation contribution is proposed for fault identification when nonlinear feature extraction methods are applied in a fault detection stage by extending the original definition of contribution. In this framework, the contribution of deviations in original measured variables to the deviation of monitoring statistics is quantified. Similarly to the contribution plots, locating and identification of the fault will be facilitated by the restrained scope of variables that have large deviation contributions to the fault in faulty scenarios when the process is assumed to be nonlinear.

The rest of this paper is organized as follows. Section II briefly revisits the multivariate statistical approaches for fault detection and contribution-based fault diagnosis. The deviation contribution, as an extension of the original definition of contribution, is proposed in Section III for contribution-based fault diagnosis of nonlinear processes. An integrated flowchart of this algorithm for offline training and online implementation is also included along with two practical issues. In Section IV, the deviation contribution for kernel PCA based process monitoring [19] is derived as an example. This result is applied to two case studies afterward: first, a numerically simulated case study; and second, the Tennessee Eastman process. Its performance is compared with the existing nonlinear contribution-based fault diagnosis approach in both cases. Following the case studies, Section V discusses the future extensions of the proposed approach. Findings of this paper are summarized in Section VI.

II. PRELIMINARY

Fig. 1 summarizes the general procedure of multivariate fault detection and contribution-based fault diagnosis. As shown in

(1), multivariate statistical feature extraction approaches are applied to normal operation dataset in training; d -dimensional representative features $\mathbf{z} \in \mathbb{R}^d$ of the process are projected from m -dimensional measured variables $\mathbf{x} \in \mathbb{R}^m$ using predefined feature extraction method \mathbf{f} . Monitoring index I is commonly a quadratic function of representative features \mathbf{z} . The upper control limit of I , i.e., I_{UCL} , is calculated based on its probability density function with a given confidence level α . In particular, kernel density estimation is applied for estimating I_{UCL} when \mathbf{f} is assumed to be nonlinear [20]

$$\mathbf{z} = \mathbf{f}(\mathbf{x}); I = \mathbf{z}^T \mathbf{z}$$

$$I_{UCL} = \arg P(I < I_0) = \alpha. \quad (1)$$

The online monitoring procedure is described by (2). The monitoring index (e.g., I^*) of new sample \mathbf{x}^* is calculated and compared with its control limit in online monitoring and/or algorithm validation. A fault will be detected if one or more monitoring statistics exceed their thresholds, namely $I^* > I_{UCL}$

$$\mathbf{z}^* = \mathbf{f}(\mathbf{x}^*); I^* = \mathbf{z}^{*T} \mathbf{z}^*. \quad (2)$$

Different multivariate statistical fault detection algorithms based on feature extraction approaches share similar structures except for the selection of feature extraction method \mathbf{f} . For a linear approach, e.g., the PCA, $\mathbf{z}^* = \mathbf{P}\mathbf{x}^*$, Chiang *et al.* proposed the following definition of variable contribution [21]:

$$C_{i,j} = z_j P_{j,i} x_i \quad (3)$$

$$C_{i,I} = \sum_{j=1}^d C_{i,j} \quad (4)$$

where $C_{i,j}$ represents the contribution of variable x_i to z_j^2 , i.e., the monitoring index with only one principal component z_j ; $P_{j,i}$ is the loading coefficient obtained by the PCA when \mathbf{x} is assumed to be normalized beforehand; and $C_{i,I}$ is the contribution of x_i to the monitoring index I .

The influential variables are identified if they have the largest contributions, namely $C_{i,I}$, to the monitoring index I . Moreover, (5) holds, which indicates that the variable contributions are the breakdown of monitoring index value

$$I = \mathbf{z}^T \mathbf{z} = \mathbf{z}^T \mathbf{P} \mathbf{x} = \sum_{i=1}^m \sum_{j=1}^d z_j P_{j,i} x_i = \sum_{i=1}^m C_{i,I}. \quad (5)$$

However, in order to enhance the fault detection performance in complex processes with nonlinearity, temporal correlation, and multiple operating modes, feature extraction method \mathbf{f} is supposed to be comprehensive and sophisticated. Owing to the rapid development of feature extraction algorithms in data analytics, the selection of \mathbf{f} is flexible. In the meantime, corresponding contribution-based fault diagnosis gets more complex along with \mathbf{f} : for example, $C_{i,I}$ cannot be calculated directly when \mathbf{f} is nonlinear since (3) and (4) are no longer valid. Therefore, contribution-based fault diagnosis requires a more generic definition of contribution that accounts for the intricacy and variability of feature extraction algorithms.

III. DEVIATION-BASED CONTRIBUTION PLOTS

A. Deviation Contributions

When the feature extraction approach \mathbf{f} is nonlinear, (3) does not apply. Instead \mathbf{x}_{new} , the new faulty sample is diagnosed against a reference point under normal condition \mathbf{x}_{ref} by estimating the contribution of difference between \mathbf{x}_{new} and \mathbf{x}_{ref} . First, (6) is obtained by applying the first-order Taylor expansion to $\mathbf{f}(\mathbf{x}_{\text{ref}})$ around \mathbf{x}_{new}

$$\mathbf{f}(\mathbf{x}_{\text{ref}}) \approx \mathbf{f}(\mathbf{x}_{\text{new}}) + \sum_{i=1}^m \frac{\partial \mathbf{f}}{\partial x_i} \bigg|_{\mathbf{x}=\mathbf{x}_{\text{new}}} (x_{\text{ref},i} - x_{\text{new},i}). \quad (6)$$

Equivalently

$$\mathbf{f}(\mathbf{x}_{\text{new}}) \approx \mathbf{f}(\mathbf{x}_{\text{ref}}) + \sum_{i=1}^m \frac{\partial \mathbf{f}}{\partial x_i} \bigg|_{\mathbf{x}=\mathbf{x}_{\text{new}}} (x_{\text{new},i} - x_{\text{ref},i}). \quad (7)$$

Hence, the deviation of \mathbf{z} is $\mathbf{dz} = \mathbf{z}_{\text{new}} - \mathbf{z}_{\text{ref}}$, where $\mathbf{dz} = \sum_{i=1}^m \frac{\partial \mathbf{f}}{\partial x_i} \bigg|_{\mathbf{x}=\mathbf{x}_{\text{new}}} (x_{\text{new},i} - x_{\text{ref},i})$. For j th component z_j , by ignoring the high-order deviation

$$\begin{aligned} z_{\text{new},j}^2 &= z_{\text{ref},j}^2 + 2z_{\text{new},j}\mathbf{dz}_j - (\mathbf{dz}_j)^2 \\ &\approx z_{\text{ref},j}^2 + 2z_{\text{new},j}\mathbf{dz}_j \\ &= z_{\text{ref},j}^2 + 2z_{\text{new},j} \sum_{i=1}^m \frac{\partial f_j}{\partial x_i} \bigg|_{\mathbf{x}=\mathbf{x}_{\text{new}}} (x_{\text{new},i} - x_{\text{ref},i}). \end{aligned} \quad (8)$$

Define $\mathcal{C}_{i,j,\text{ref}}$ and $\mathcal{C}_{i,j,\text{new}}$ as contributions of $x_{\text{new},i}$ to $z_{\text{ref},i}$ and $z_{\text{new},i}$, respectively, and $\mathbf{dC}_{i,j}$ as the deviation contribution

$$z_{\text{new},j}^2 = \sum_{i=1}^m \mathcal{C}_{i,j,\text{new}} = \sum_{i=1}^m \mathcal{C}_{i,j,\text{ref}} + \sum_{i=1}^m \mathbf{dC}_{i,j}. \quad (9)$$

Therefore, the individual deviation contributions of x_i to z_j are

$$\mathbf{dC}_{i,j} = 2z_{\text{new},j} \frac{\partial f_j}{\partial x_i} \bigg|_{\mathbf{x}=\mathbf{x}_{\text{new}}} (x_{\text{new},i} - x_{\text{ref},i}). \quad (10)$$

For the new monitoring index I_{new} at \mathbf{x}_{new} , (11) can be obtained using (1), (4), and (9)

$$\begin{aligned} I_{\text{new}} &= I_{\text{ref}} + \mathbf{dI} = \sum_{j=1}^d z_{\text{new},j}^2 \\ &= \sum_{j=1}^d \sum_{i=1}^m \mathcal{C}_{i,j,\text{ref}} + \sum_{j=1}^d \sum_{i=1}^m \mathbf{dC}_{i,j} \\ &= I_{\text{ref}} + \sum_{j=1}^d \sum_{i=1}^m \mathbf{dC}_{i,j} = I_{\text{ref}} + \sum_{i=1}^m \mathbf{dC}_{i,I} \end{aligned} \quad (11)$$

for $i = 1, 2, \dots, m$.

The deviation contribution of $x_{\text{new},i}$ to I_{new} is, therefore, defined as

$$\mathbf{dC}_{i,I} = \sum_{j=1}^d \mathbf{dC}_{i,j} = \left[2\mathbf{z}_{\text{new}}^T \frac{\partial \mathbf{f}}{\partial x_i} \bigg|_{\mathbf{x}=\mathbf{x}_{\text{new}}} \mathbf{dx}_i \right] \quad (12)$$

where $\mathbf{dx}_i = x_{\text{new},i} - x_{\text{ref},i}$.

Additional conclusion of deviation contributions can be drawn by adopting the contribution of original variables \mathbf{x} to extracted features \mathbf{z} , i.e., \mathbf{dx} and \mathbf{dz} , defined in [8] and assuming (9) for multilayer feature extraction approaches. When a feature extraction approach with k layers, namely $\mathbf{f}^{(1)}, \mathbf{f}^{(2)}, \dots, \mathbf{f}^{(k)}$ is applied, the deviation contribution of intermediate feature $\mathbf{dy}_i^{(l-1)}$ of $(l-1)$ th layer to intermediate feature $\mathbf{dy}_j^{(l)}$ of l th layer is denoted as $\mathbf{dC}_{y_i^{(l-1)}, y_j^{(l)}}$

$$\mathbf{dC}_{y_i^{(l-1)}, y_j^{(l)}} = \frac{\partial f_j^{(l)}}{\partial y_i^{(l-1)}} \bigg|_{\mathbf{y}^{(l-1)}=\mathbf{y}_{\text{new}}^{(l-1)}} \mathbf{dy}_i^{(l-1)}. \quad (13)$$

The deviation contribution is transmitted between two layers based on the chain rule, as shown in the following equation:

$$\begin{aligned} \mathbf{dy}_k^{(l+1)} &= \sum_{j=1}^{r_l} \frac{\partial f_k^{(l+1)}}{\partial y_j^{(l)}} \mathbf{dy}_j^{(l)} \\ &= \sum_{j=1}^{r_l} \left[\frac{\partial f_k^{(l+1)}}{\partial y_j^{(l)}} \sum_{i=1}^{r_{l-1}} \frac{\partial f_j^{(l)}}{\partial y_i^{(l-1)}} \mathbf{dy}_i^{(l-1)} \right] \\ &= \sum_{i=1}^{r_{l-1}} \frac{\partial f_k^{(l+1)}}{\partial y_i^{(l-1)}} \bigg|_{\mathbf{y}^{(l-1)}=\mathbf{y}_{\text{new}}^{(l-1)}} \mathbf{dy}_i^{(l-1)} \end{aligned} \quad (14)$$

and the contribution of $\mathbf{dy}_i^{(l-1)}$ in $\mathbf{dy}_k^{(l+1)}$ is

$$\mathbf{dC}_{y_i^{(l-1)}, y_k^{(l+1)}} = \frac{\partial f_k^{(l+1)}}{\partial y_i^{(l-1)}} \bigg|_{\mathbf{y}^{(l-1)}=\mathbf{y}_{\text{new}}^{(l-1)}} \mathbf{dy}_i^{(l-1)}. \quad (15)$$

By recursion, the contribution of deviation in original measured variables \mathbf{dx}_i to variation in monitoring index \mathbf{dI} is

$$\mathbf{dC}_{i,I} = \sum_{i_1=1}^{r_1} \dots \sum_{i_L=1}^{d_1} \frac{\partial I}{\partial z_{i_L}} \dots \frac{\partial y_{i_1}^{(1)}}{\partial x_i} \mathbf{dx}_i = \frac{\partial I}{\partial x_i} \bigg|_{\mathbf{x}=\mathbf{x}_{\text{new}}} \mathbf{dx}_i. \quad (16)$$

When $\mathbf{f}^{(1)}, \dots, \mathbf{f}^{(k)}$ are all linear, the result is equivalent to the contribution propagation framework proposed for PCA-enhanced canonical variate analysis based linear dynamic process monitoring [22].

B. Algorithm Summary

Fig. 2 shows the integrated scheme of offline training and online monitoring of feature extraction based fault detection and deviation contribution plots based fault diagnosis. Red-dashed boxes highlight the stages where deviation contribution is involved. The parameters of feature extraction method and control limits of monitoring index are obtained by offline training using historical data collected from normal operating conditions. In succession, unlike the ordinary contribution plots, functional

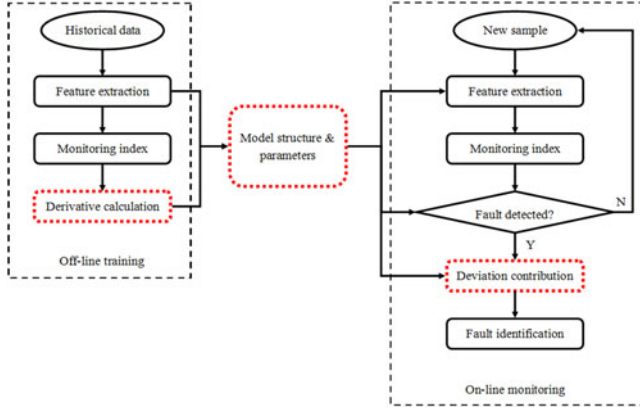


Fig. 2. Integrated monitoring flowchart with feature extraction and deviation contribution plots.

structure and coefficients used for deviation contribution calculation are also learned in the training stage. When implemented online, monitoring indexes are calculated using the new measurement point retrieved from the process and compared with the control limits. If the new sample is labeled as faulty, deviation contributions of all measured variables in this sample are calculated and influential variables are identified for further diagnosis and characterization. Online implementation will be time efficient owing to the fact that all the functional structures and parameters have been acquired in offline training stage and online fault detection stage.

C. Practical Issues

1) *Reference Point Selection*: The proposed approach applies to both linear and nonlinear process monitoring, and its domain of application is specified by properly selecting \mathbf{x}_{ref} . When linear feature extraction approach is applied to data after centering, $\mathbf{x}_{\text{ref}} = 0$ is selected, correspondingly, $I_{\text{ref}} = 0$; the deviation contribution becomes equivalent to the standard contribution defined by (3) and (4).

The selection of reference point for deviation calculation will influence the accuracy of approximating the deviation contribution and the diagnosis result thereafter when nonlinearity exists. The most intuitive choice will be the mean point of all training samples, which is a linear combination of these normal points; however, when the unknown normal operating space is nonconvex, the mean point may not be normal at all, hence inappropriate to be taken as a reference point. To ensure the reference point is normal, it has to be selected from these training samples. On the other hand, the accuracy of Taylor expansion depends on the distance of deviation. The smaller the deviation is, the more accurate the approximation will be. Therefore, nearest neighbor of \mathbf{x}_{new} in the normal dataset is selected as the reference point, as shown in the following equation:

$$\mathbf{x}_{\text{ref}} = \arg \min \|\mathbf{x} - \mathbf{x}_{\text{new}}\|_2, \quad \mathbf{x} \in \mathbf{X}_{\text{normal}} \quad (17)$$

where $\mathbf{X}_{\text{normal}}$ denotes the set containing all normal samples used for training.

When an outlier with large deviation in faulty variable is observed and considered as a fault signature, even the deviation

between the outlier and its nearest neighbor in normal dataset will be large toward the direction of faulty variable according to (17). The faulty variable can be identified since its contribution calculated by (12) will, therefore, be significant.

2) *Online Derivative Computation*: Fig. 2 indicates that the representative features and their derivatives should be updated online using both training data and new sample when applied to nonlinear feature extraction layers. When the analytical formulations of monitoring index I and its explicit derivative is available, the functional structure can be computed in advance in offline training stage and the value of derivatives can be obtained by substituting the new sample. However, numerical approaches for derivative calculation should be considered when the feature extraction layer is nonparametric or the computation of $\frac{\partial I}{\partial x_i}$ is difficult, such as automatic differentiation [23] and complex-step differentiation [24].

The applicability of different derivative computation approaches will differ with respect to different feature extraction methods. In this paper, only the fault diagnosis performance of the analytical solution will be tested in nonlinear monitoring algorithm.

IV. NUMERICAL EXAMPLES

As an example of nonlinear feature extraction approach, the deviation contribution based fault diagnosis for kernel PCA is investigated in this section. Corresponding fault identification performance is compared with the performance obtained by contribution rate of a numerical simulation and the Tennessee Eastman process benchmark case study.

A. Deviation Contribution Plots for Kernel PCA Based Process Monitoring

The general formulation of feature extraction using the kernel PCA model is stated as follows:

$$\mathbf{z} = \mathbf{v}_1 \cdot \Phi(\mathbf{x}) = \alpha K_c; \quad \mathbf{e} = \mathbf{v}_2 \cdot \Phi(\mathbf{x}) = \beta K_c;$$

$$K(i, j) = k(\mathbf{x}(i), \mathbf{x}(j)); \quad K_c = K - KE - EK + EKE \quad (18)$$

where

- $\mathbf{x} \in \mathbb{R}^m$ original measured variables;
- $\Phi(\mathbf{x}) \in \mathbb{R}^{v_h}$ projection of \mathbf{x} in higher dimensional nonlinear feature space;
- $\mathbf{z} \in \mathbb{R}^{d_1}, \mathbf{e} \in \mathbb{R}^{d_2}$ projected kernel principal components;
- $[\alpha; \beta]$ projection matrix comprising of eigenvectors obtained by eigenvalue decomposition of K ;
- k kernel function used for computing kernel matrix K ;
- $E_{ij} = \frac{1}{n}$ for centering of K matrix; n is the number of sample points.

The most widely used kernel function is radial basis function (RBF), as shown in the following equation:

$$k(\mathbf{x}(t_1), \mathbf{x}(t_2)) = \exp \left(-\frac{1}{c} \sum_{i=1}^m (x_i(t_1) - x_i(t_2))^2 \right) \quad (19)$$

where c is the bandwidth of the kernel.

Based on (12), deviation contributions of i th measured variable x_i to T^2 and Q statistics at time t^* , if assuming RBF as the

TABLE I
FAULT SPECIFICATION OF NUMERICAL SIMULATION

	Type	Description	Influential variable
Fault 1	Model mismatch	$x_3 = -1.1t^3 + 3.2t^2 + e_3$	x_3
Fault 2	Constant bias	$x_1 = t + 0.2 + e_1$	x_1
Fault 3	Random variation	$e_2 \sim N(0, 0.05)$	x_2

kernel function, are presented as follows:

$$dC_{i,T^2}(t^*) = 2 \sum_{j=1}^{d_1} z_j(t^*) \left[- \sum_{t=1}^n \frac{2\alpha_{t^*,t}}{c} A_i(t^*, t) - \frac{\sum_{\tau=1}^n A_i(t^*, \tau)}{n} \sum_{t=1}^n \alpha_{t^*,t} \right] (x_i(t^*) - x_{\text{ref},i}) \quad (20)$$

$$dC_{i,Q}(t^*) = 2 \sum_{j=1}^{d_2} e_j(t^*) \left[- \sum_{t=1}^n \frac{2\beta_{t^*,t}}{c} A_i(t^*, t) - \frac{\sum_{\tau=1}^n A_i(t^*, \tau)}{n} \sum_{t=1}^n \beta_{t^*,t} \right] (x_i(t^*) - x_{\text{ref},i}) \quad (21)$$

where

$$A_i(t^*, t) = k(\mathbf{x}(t^*), \mathbf{x}(t)) \cdot (x_i(t^*) - x_i(t)).$$

B. Numerical Simulation

1) *Simulation Model*: Equation (22) shows the process model of a numerical simulation, adopted from [25]

$$\begin{cases} x_1 = t + e_1 \\ x_2 = t^2 - 3t + e_2 \\ x_3 = -t^3 + 3t^2 + e_3 \end{cases} \quad (22)$$

where $t \in [0.01, 1]$ and e_i are the independent Gaussian noises with distribution $N(0, 0.01)$ for $i = 1, 2, 3$.

The ability of fault detection of kernel PCA in this case has already been testified by the previous works [26], [27]. Both deviation contribution plots and contribution rate are adopted successively in order to identify influential variables in faulty scenarios and their performances are compared. A set of 500 samples from the model specified by (22) with no fault is generated in advance for offline training. For validation, three types of faults have been introduced to this simulated system separately, as specified in Table I; for each fault, 500 samples are collected as the validation set.

The deviation contributions of x_1, x_2 , and x_3 in validation datasets are calculated via analytical derivative of \mathbf{z} and \mathbf{e} . Taking T^2 statistics as an example, index of diagnosed faulty variable in faulty conditions detected by T^2 is obtained as follows:

$$i_{\text{diagnosed}} = \arg \max_i dC_{x_i, T^2} \quad (23)$$

and the diagnosis result is claimed to be correct if $i_{\text{diagnosed}}$ matches the influential variable of corresponding validation set referring to Table I.

TABLE II
DIAGNOSIS RATES OF DEVIATION CONTRIBUTION AND CONTRIBUTION RATE

	Deviation contribution		Contribution rate	
Fault	T^2	Q	T^2	Q
1	0.3425	0.3205	0.2188	0.1798
2	0.9058	0.8896	0.5991	0.4890
3	0.7531	0.7058	0.7726	0.7121

TABLE III
COMPARISON OF RECONSTRUCTED MONITORING INDEXES

		Deviation contribution		Contribution rate	
	Fault	Corr	RMSE	Corr	RMSE
T^2	1	0.9057	3.033	0.7511	286.0
	2	0.8914	26.70	0.7871	333.9
	3	0.9572	3.508	0.8241	311.9
Q	1	0.9739	0.0347	0.8620	2.546
	2	0.8937	0.1880	0.7617	2.861
	3	0.9823	0.0366	0.8507	2.410

In order to eliminate the influence of random error, 100 times of Monte Carlo simulation have been conducted; the correct diagnosis rate is $DR = \frac{\text{num}(\text{Correct Diagnosis})}{\text{num}(\text{Faulty Samples})}$ for each run and the averaged diagnosis rate over 100 times Monte Carlo simulation is used for diagnosis performance evaluation in different faulty scenarios.

2) *Results and Discussion*: Table II shows the correct fault identification rates using the proposed deviation contribution plots and contribution rate algorithm over 100 simulations.

It can be seen from Table I that the deviation contribution plots obtain a better diagnosis rate of influential variables for Fault 2 comparing to the performance of contribution rate. It may be justified by the following reasoning: the deviation in process variable is significant if there is constant bias fault in it; hence, including the variable deviation along with the derivative will achieve a more accurate estimation of its contribution. In the meantime, although Fault 1 is difficult to both detect [27] and diagnose, the proposed deviation-based contribution plots can also improve fault diagnosis performance. Moreover, deviation contribution plots reflect the breakdown of monitoring statistics directly and can be used to reconstruct the original monitoring statistics along with the reference points. For a new sample \mathbf{x}_{new} at time t , (24) holds for its reconstructed monitoring index \hat{I}_{new} and the deviation contributions of $x_{\text{new},j}$

$$\hat{I}_{\text{new}} = I_{\text{ref}} + dI \approx I_{\text{ref}} + \sum_{j=1}^n C_{\mathbf{x}_{\text{new},j}, I} \cdot \quad (24)$$

To further illustrate, two measures of similarity between \hat{I}_{new} and I_{new} are calculated: the linear correlation and the root-mean-squared-error (RMSE), as shown in the following equations:

$$\text{Corr}(\hat{I}_{\text{new}}, I_{\text{new}}) = \frac{\sum_{t=1}^N (\hat{I}_{\text{new}}(t) - \mu_{\hat{I}})(I_{\text{new}}(t) - \mu_I)}{\sigma_{\hat{I}} \sigma_I} \quad (25)$$

TABLE IV
MEASURED PROCESS VARIABLES FOR FAULT DIAGNOSIS

No.	Variable name	No.	Variable name
XMEAS(1)	A feed (stream 1)	XMEAS(12)	Separator level
XMEAS(2)	D feed (stream 2)	XMEAS(13)	Separator pressure
XMEAS(3)	E feed (stream 3)	XMEAS(14)	Separator underflow
XMEAS(4)	A and C feed	XMEAS(15)	Stripper level
XMEAS(5)	Recycle flow	XMEAS(16)	Stripper pressure
XMEAS(6)	Reactor feed rate	XMEAS(17)	Stripper underflow
XMEAS(7)	Reactor pressure	XMEAS(18)	Stripper temperature
XMEAS(8)	Reactor level	XMEAS(19)	Stripper steam flow
XMEAS(9)	Reactor temperature	XMEAS(20)	Compressor work
XMEAS(10)	Purge rate	XMEAS(21)	Reactor water temperature
XMEAS(11)	Separator temperature	XMEAS(22)	Separator water temperature

TABLE V
FAULT SPECIFICATION IN TEP CASE STUDY

	Process Variable	Type
IDV(11)	Reactor cooling water inlet	Random variation
IDV(13)	Reaction kinetics	Slow drift

$$RMSE(\hat{I}_{new}, I_{new}) = \sqrt{\frac{1}{N} \sum_{t=1}^N (\hat{I}_{new}(t) - I_{new}(t))^2} \quad (26)$$

where $(\mu_{\hat{I}}, \sigma_{\hat{I}})$ and (μ_I, σ_I) are the mean and standard deviation pairs of \hat{I}_{new} and I_{new} , respectively; N is the number of faulty samples being diagnosed.

Table III compares the averaged performances of reconstructed monitoring indexes obtained by deviation contribution and contribution rate. The reconstructed T^2 and Q values by (24) using the deviation contribution plots are very similar to their true values, comparing to those obtained by the contribution rate. It can be concluded that, similar to the original contribution defined by (5), the deviation contribution plots directly represent the breakdown of deviation in monitoring statistics with respect to individual process variables.

C. Tennessee Eastman Process

1) *Process Description*: As the most widely accepted benchmark case study for validating process monitoring algorithms, Tennessee Eastman process is a well-built simulation model of a real chemical plant with 41 measured variables and 11 manipulated variables. The configuration of Tennessee Eastman process (TEP) used in this case study is adopted from [28]. Table IV shows the 22 process variables collected in this case study for fault detection and diagnosis.

The proposed deviation contribution is applied to two fault scenarios [IDV(11) and IDV(13)], where kernel PCA has satisfactory fault detection performance, as described in Table V.

A training dataset of 500 samples is acquired under normal operating condition. Two faulty datasets of 1000 sample points with the fault occurring from Sample 200 are collected for IDV(11) and IDV(13), respectively. Kernel PCA is used for feature extraction and monitoring statistics calculation. Deviation contribution and contribution rate are applied successively

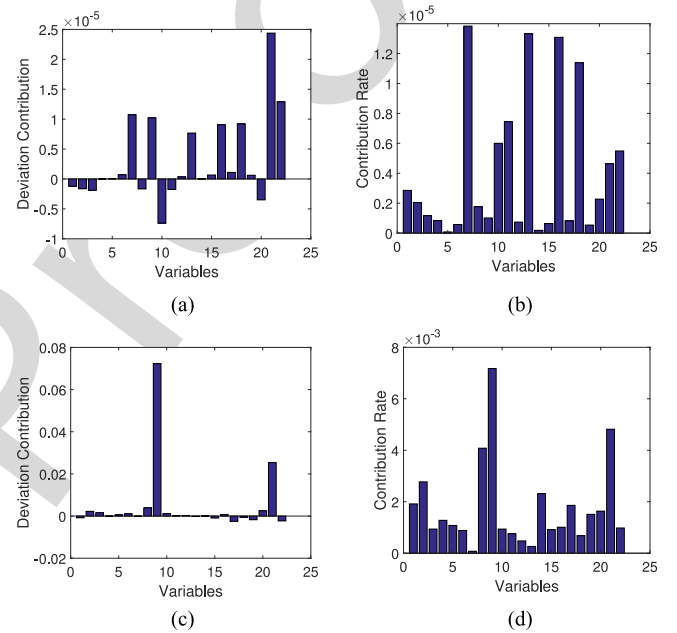


Fig. 3. Contribution plots at sample number 500: IDV(11). (a) Deviation contributions to T^2 . (b) Contribution rates to T^2 . (c) Deviation contributions to Q . (d) Contribution rates to Q .

for fault diagnosis. Diagnosis performances and reconstruction of monitoring statistics of these two approaches are compared.

2) *Results and Discussion*: Similar to the standard contribution plots, deviation contribution plots at certain time stamp can also be visualized as bar plots, as shown in Figs. 3 and 4.

For the reactor cooling water inlet fault [IDV(11)], Fig. 3(c) indicates the most influential variables are XMEAS(9): reactor temperature and XMEAS(21): reactor water temperature. These two variables are most directly related to the random variations in the inlet flow rate of reactor cooling water. While Fig. 3(d) also identifies the same results with small distraction from XMEAS(8). For T^2 statistics, due to the fact that T^2 -based fault detection performance is not as good as Q , the diagnosis performance also decays. Nevertheless, Fig. 3(a) can still provide some insights, such as XMEAS(21) has the largest contribution, whereas the message from Fig. 3(b) is relatively vague with XMEAS(7), XMEAS(13), XMEAS(16), and XMEAS(18), all having similar magnitudes of contribution.

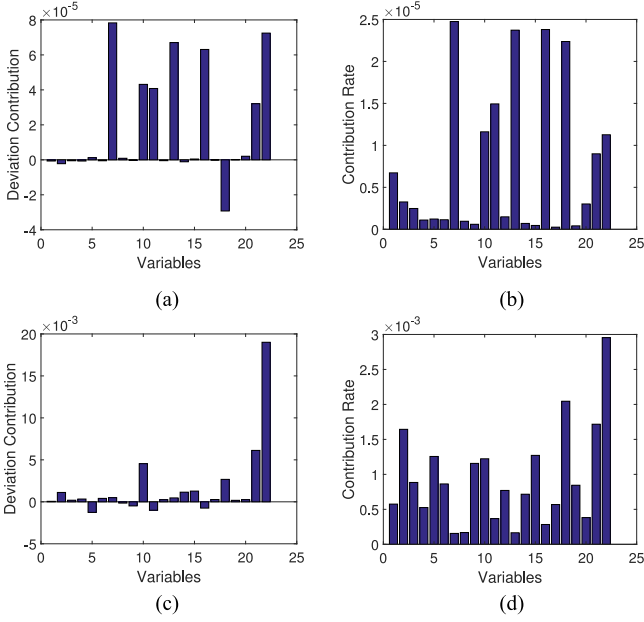


Fig. 4. Contribution plots at sample number 500: IDV(13). (a) Deviation contributions to T^2 . (b) Contribution rates to T^2 . (c) Deviation contributions to Q . (d) Contribution rates to Q .

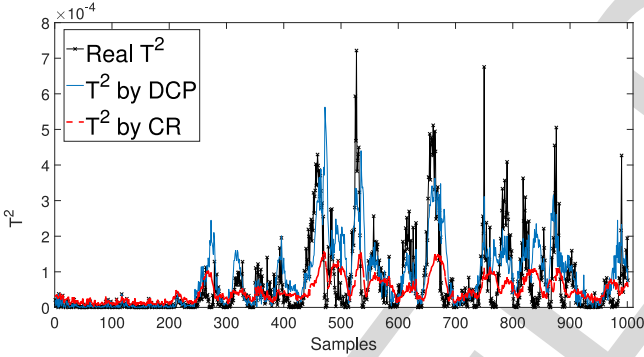


Fig. 5. Reconstructed T^2 statistics for IDV(11).

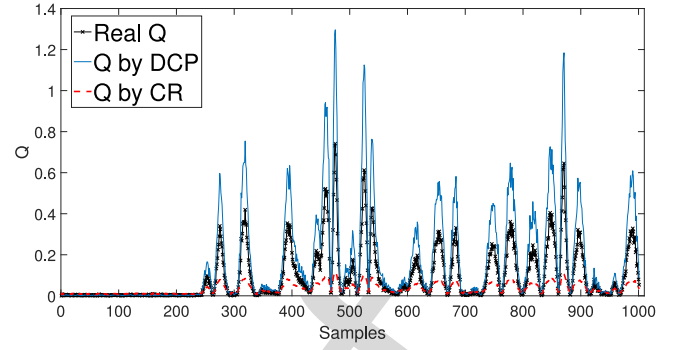


Fig. 6. Reconstructed Q statistics for IDV(11).

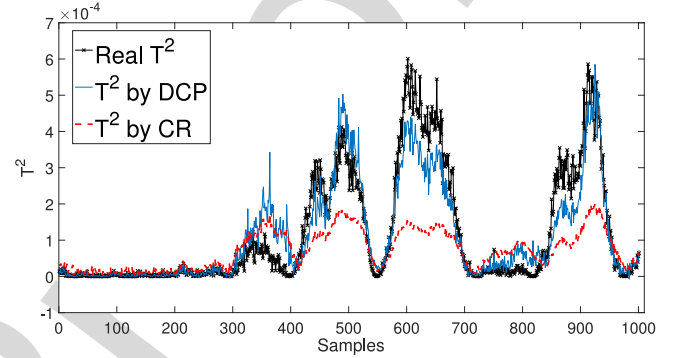


Fig. 7. Reconstructed T^2 statistics for IDV(13).

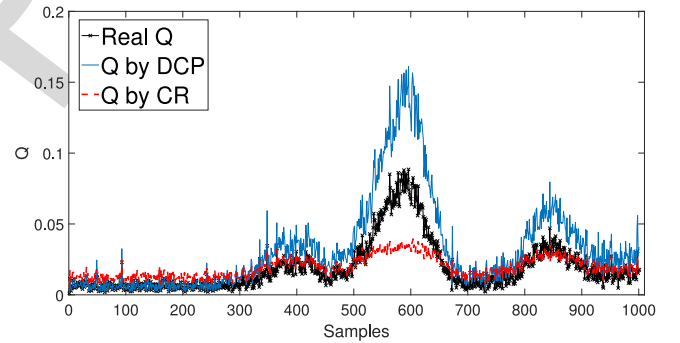


Fig. 8. Reconstructed Q statistics for IDV(13).

As for the slow drift in reaction kinetics [IDV(13)], Fig. 4(a) indicates that XMEAS(7), XMEAS(13), and XMEAS(16) are influential, which are pressures in different units. Since the reactions in this plant are turning gaseous components into liquids [29], the pressures in multiple units will be influenced directly if the reaction kinetics decay. While Fig. 4(b) indicates similar subgroup of influential variables, it behaves similar comparing to Fig. 3(b) even though the fault scenarios are totally different. It may, therefore, impair the credibility of diagnosis result of contribution rate in IDV(13). In the meantime, Fig. 4(a)–(d) all indicate XMEAS(22), which is the separator water temperature, also has large contributions in IDV(13). Further investigation is needed into the process so as to justify this observation.

The comparison of original monitoring statistics values and the reconstructed values obtained by both deviation contribution and contribution rates is shown in Figs. 5–8. The proposed deviation contribution plots still maintain a reasonable level of similarity with the real monitoring statistics while

reconstruction performance of contribution rate has improved. The reason behind is that the datasets have been normalized beforehand; therefore, the magnitude of deviations in variables may not be determined in their contributions.

In summary, it can be seen from both case studies that the deviation contribution has the following features:

- 1) The implementation and visualization are straightforward and the format of results is the same as standard contribution plots which facilitates the interpretation.
- 2) The fault diagnosis performance can be improved by applying deviation contribution in different nonlinear processes.
- 3) It also yields the good reconstructions of monitoring statistics; hence the deviation contribution may be a good quantification of variable contributions to monitoring statistics, according to the original definition.

V. DISCUSSIONS

There is great potential in improving the diagnosis performance of proposed deviation-based contribution. One of the advantages is its compatibility with both linear and nonlinear/nonparametric feature extraction algorithms which can handle extra data complexities before diagnosing the fault. More sophisticated reference point selection strategies are of interest in context of data quality. Moreover, the influential variables identified by their deviation contributions are only preliminary information for fault diagnosis. Thus, it is reasonable to consider synthesizing this framework with other diagnosis methods that incorporate additional process knowledge (e.g., causal dependency between variables) to further depict the fault and draw more comprehensive conclusions about it.

The formulation of the deviation-based contribution, especially the derivative part using Taylor series expansion, may also be updated with respect to *a priori* process knowledge, such as the importance of individual variables and their correlations. Similar idea has been presented and proven to be beneficial in neural network training through back propagation [30]. Although back propagation is based on the sensitivity of weights, whereas the contribution analysis requires the sensitivity of input (measurement) variables, the proposed deviation contribution analysis can also be carried out backward by following a similar idea to the back propagation, hence worth for the future research.

VI. CONCLUSION

The deviation-based contribution plots proposed in this paper extended the concept of contribution plots in fault diagnosis of linear processes to the situation of nonlinear feature extraction methods. It reflects the contribution of the deviation in original measured variables to the deviation in final monitoring statistics without involving intermediate features in computation. By considering the deviation in both original process variables and monitoring statistics, this framework can estimate the contributions of process variables to present fault even if the approach of projecting these variables for monitoring statistics computation is complicated. Practical issues when implementing this framework were also discussed. As an example, its capability in identifying process variables responsible for specific fault occurrence has been validated on monitoring algorithm with nonlinear feature extraction approach.

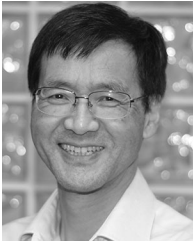
REFERENCES

- [1] 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," *Chemometrics Intell. Lab. Syst.*, vol. 50, no. 2, pp. 243–252, 2000.
- [2] J. Yu, "A particle filter driven dynamic Gaussian Mixture Model approach for complex process monitoring and fault diagnosis," *J. Process Control*, vol. 22, no. 4, pp. 778–788, 2012.
- [3] Z. Ge, S. Zhong, and Y. Zhang, "Semisupervised kernel learning for FDA model and its application for fault classification in industrial processes," *IEEE Trans. Ind. Informat.*, vol. 12, no. 4, pp. 1403–1411, Aug. 2016.
- [4] B. Yang, R. Liu, and X. Chen, "Fault diagnosis for wind turbine generator bearing via sparse representation and shift-invariant K-SVD," *IEEE Trans. Ind. Informat.*, vol. 13, no. 3, pp. 1321–1331, Jun. 2017.
- [5] Z. Yongli, H. Limin, and L. Jinling, "Bayesian networks-based approach for power systems fault diagnosis," *IEEE Trans. Power Del.*, vol. 21, no. 2, pp. 634–639, Apr. 2006.
- [6] J. Yu and S. J. Qin, "Multiway Gaussian Mixture Model based multi-phase batch process monitoring," *Ind. Eng. Chem. Res.*, vol. 48, no. 18, pp. 8585–8594, 2009.
- [7] L. H. Chiang and R. D. Braatz, "Process monitoring using causal map and multivariate statistics: fault detection and identification," *Chemometrics Intell. Lab. Syst.*, vol. 65, no. 2, pp. 159–178, 2003.
- [8] P. Miller, R. E. Swanson, and C. E. Heckler, "Contribution plots: A missing link in multivariate quality control," *Appl. Math. Comput. Sci.*, vol. 8, no. 4, pp. 775–792, 1998.
- [9] T. Kourtí, P. Nomikos, and J. F. MacGregor, "Analysis, monitoring and fault diagnosis of batch processes using multiblock and multiway PLS," *J. Process Control*, vol. 5, no. 4, pp. 277–284, 1995.
- [10] J.-M. Lee, C. Yoo, and I.-B. Lee, "Statistical process monitoring with independent component analysis," *J. Process Control*, vol. 14, no. 5, pp. 467–485, 2004.
- [11] 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, 2015.
- [12] C. F. Alcalá and S. J. Qin, "Reconstruction-based contribution for process monitoring with Kernel principal component analysis," *Ind. Eng. Chem. Res.*, vol. 49, no. 17, pp. 7849–7857, 2010.
- [13] J.-H. Cho, J.-M. Lee, S. W. Choi, D. Lee, and I.-B. Lee, "Fault identification for process monitoring using Kernel principal component analysis," *Chem. Eng. Sci.*, vol. 60, no. 1, pp. 279–288, 2005.
- [14] Y. Zhang, L. Zhang, and R. Lu, "Fault identification of nonlinear processes," *Ind. Eng. Chem. Res.*, vol. 52, no. 34, pp. 12072–12081, 2013.
- [15] G. Li, C. F. Alcalá, S. J. Qin, and D. Zhou, "Generalized reconstruction-based contributions for output-relevant fault diagnosis with application to the Tennessee Eastman process," *IEEE Trans. Control Syst. Technol.*, vol. 19, no. 5, pp. 1114–1127, Sep. 2011.
- [16] V. Kariwala, P.-E. Odiwei, Y. Cao, and T. Chen, "A Branch and Bound method for isolation of faulty variables through missing variable analysis," *J. Process Control*, vol. 20, no. 10, pp. 1198–1206, 2010.
- [17] T. Fuse and K. Kamiya, "Statistical anomaly detection in human dynamics monitoring using a hierarchical Dirichlet process hidden Markov model," *IEEE Trans. Intell. Transport. Syst.*, vol. 18, no. 11, pp. 3083–3092, Nov. 2017.
- [18] J. Zeng, S. Luo, J. Cai, U. Kruger, and L. Xie, "Nonparametric density estimation of hierarchical probabilistic graph models for assumption-free monitoring," *Ind. Eng. Chem. Res.*, vol. 56, no. 5, pp. 1278–1287, 2017.
- [19] R. T. Samuel and Y. Cao, "Nonlinear process fault detection and identification using kernel PCA and Kernel density estimation," *Syst. Sci. Control Eng.*, vol. 4, no. 1, pp. 165–174, 2016.
- [20] P. Odiwei and Y. Cao, "Nonlinear dynamic process monitoring using Canonical variate analysis and Kernel density estimations," *Comput. Aid. Chem. Eng.*, vol. 27, pp. 1557–1562, 2009.
- [21] L. H. Chiang, E. L. Russell, and R. D. Braatz, *Fault Detection and Diagnosis in Industrial Systems*. New York, NY, USA: Springer, 2000.
- [22] R. Tan and Y. Cao, "Contribution plots based fault diagnosis of a multiphase flow facility with PCA-enhanced Canonical Variate Analysis," in *Proc. 23rd Int. Conf. Autom. Comput.*, 2017, pp. 1–6.
- [23] R. D. Neidinger, "Introduction to automatic differentiation and MATLAB object-oriented programming," *SIAM Rev.*, vol. 52, no. 3, pp. 545–563, 2010.
- [24] J. R. Martins, P. Sturdza, and J. J. Alonso, "The complex-step derivative approximation," *ACM Trans. Math. Softw.*, vol. 29, no. 3, pp. 245–262, 2003.
- [25] D. Dong and T. J. McAvoy, "Nonlinear Principal Component Analysis-based on principal curves and neural networks," *Comput. Chem. Eng.*, vol. 20, no. 1, pp. 65–78, 1996.
- [26] J.-M. Lee, C. Yoo, S. W. Choi, P. A. Vanrolleghem, and I.-B. Lee, "Nonlinear process monitoring using Kernel principal component analysis," *Chem. Eng. Sci.*, vol. 59, no. 1, pp. 223–234, 2004.
- [27] S. W. Choi, C. Lee, J.-M. Lee, J. H. Park, and I.-B. Lee, "Fault detection and identification of nonlinear processes based on Kernel PCA," *Chemometrics Intell. Lab. Syst.*, vol. 75, no. 1, pp. 55–67, 2005.
- [28] S. Yin, S. X. Ding, A. Haghani, H. Hao, and P. Zhang, "A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark Tennessee Eastman Process," *J. Process Cont.*, vol. 22, no. 9, pp. 1567–1581, 2012.
- [29] J. J. Downs and E. F. Vogel, "A plant-wide industrial process control problem," *Comput. Chem. Eng.*, vol. 17, no. 3, pp. 245–255, 1993.
- [30] X. Yuan, B. Huang, C. Yang, and W. Gui, "Deep learning based feature representation and its application for soft sensor modeling with variable-wise weighted SAE," *IEEE Trans. Ind. Informat.*, p. 1, 2018, doi: [10.1109/TH.2018.2809730](https://doi.org/10.1109/TH.2018.2809730).



Ruomu Tan received the B.Eng. degree in automation from Zhejiang University, Hangzhou, China, in 2013, and the M.Sc. degree in process control from the University of Alberta, Edmonton, AB, Canada, in 2015. She is currently working toward the Ph.D. degree in chemical engineering at Imperial College London, U.K.

She is also a Marie Curie Early Stage Researcher of Horizon 2020 EID-ITN Project PRONTO. Her research interests include data-driven nonlinear process monitoring, multivariate statistical analysis, and their application to process industries.



Yi Cao (M'96–SM'10) received the M.Sc. degree in control engineering from Zhejiang University, Hangzhou, China, in 1985, and the Ph.D. degree in engineering from the University of Exeter, Exeter, U.K., in 1996.

He is currently a Professor with the College of Chemical and Biological Engineering, Zhejiang University. His research interests include advanced process control, including plant-wide control, nonlinear system identification, nonlinear model predictive control, and process monitoring.

IEEE Proof