Journal of the Franklin Institute 361 (2024) 107114

Contents lists available at ScienceDirect

Journal of the Franklin Institute

journal homepage: www.elsevier.com/locate/fi

Incipient fault detection and isolation with Cauchy–Schwarz divergence: A probabilistic approach

Feiya Lv [a], Shujian Yu [b,c], Huawei Ye [d], Jinsong Zhao [a,*], Chenglin Wen [e]

[a] Department of Chemical Engineering, Tsinghua University, Beijing 100084, China
[b] Department of Computer Science, Vrije Universiteit Amsterdam, 1081 HV Amsterdam, The Netherlands
[c] Machine Learning Group, UiT - The Arctic University of Norway, 9037 Tromsø, Norway
[d] Sinopec Jiujiang Company, Jiujiang, 332004, China
[e] School of Automation, Guangdong University of Petrochemical Technology, Maoming, 525000, China

ARTICLE INFO

Keywords:
Cauchy–Schwarz divergence
Conditional Cauchy–Schwarz divergence
Fault detection
Fault isolation
Process monitoring

ABSTRACT

To monitor the dynamics and non-stationarity inherent in industrial processes, we propose a novel incipient fault detection and isolation scheme grounded in a probabilistic perspective, using the Cauchy–Schwarz (CS) divergence. Our innovation lies in the utilization of marginal CS divergence for incipient fault detection and the conditional CS divergence for fault isolation. This approach neither require prior parametric assumptions about the underlying data distribution nor depend on historical fault data, while simultaneously providing explanatory diagnostics. Beyond this, we develop a change point detection-base diagnosis technique for practical engineering applications. This online process monitoring technique guarantees timely intervention to uphold process stability and safety. We demonstrate the compelling performance, higher detection rate and lower alarm rate, of the CS divergence over prevalent divergence-based approaches, such as Kullback–Leibler divergence, Wasserstein distance and Mahalanobis distance. We also illustrate the explanatory insights offered by conditional CS divergence in fault isolation on synthetic data, benchmarks of continuous stirred-tank reactor process and continuous stirred-tank heater process and even a real-world continuous catalytic reforming process. Code of this CS divergence based fault detection and isolation is available at https://github.com/Feiya-Lv/Incipient-Fault-Detection-and-Isolation-with-Cauchy--Schwarz-Divergence.

1. Introduction

With the growing demand for security equipments and high-quality products, process monitoring is an essential task with critical applications in process safety and production maintenance [1]. As fault is formally defined as an "unpermitted deviation of at least one parameter of a process from acceptable conditions", fault detection and isolation is a heavily studied area to systematically determine whether a fault has occurred (*detection*) and identify the process variables affected by the fault (*isolation*) [2,3].

Multivariate statistical process monitoring approaches have been widely used in complex industrial and perform satisfactorily in highly correlated multi-modal variables, e.g., principal component analysis (PCA) for Gaussian distributed source signals, and independent component analysis as a non-Gaussian extension. They in general use distance-based statistics to timely detect any deviation from normal or "in-control" region that has been defined to accommodate the acceptable variations. Various improved or

* Corresponding author.
E-mail addresses: lfy61290@mail.tsinghua.edu.cn (F. Lv), s.yu3@vu.nl (S. Yu), yehw.jjsh@sinopec.com (H. Ye), jinsongzhao@tsinghua.edu.cn (J. Zhao), wencl@hdu.edu.cn (C. Wen).

https://doi.org/10.1016/j.jfranklin.2024.107114
Received 23 October 2023; Received in revised form 28 March 2024; Accepted 24 July 2024
Available online 2 August 2024

recursive statistical process monitoring methods have been proposed thereafter [4,5]. They are conceptual simplicity and effective in dealing with dynamic processes, but their performances still need improvement in terms of increased sensitivity, early detection, and minimal false alarms during normal operation. In recent years, detection of incipient faults has attracted significant attention, as it is crucial to prevent system deterioration. Incipient faults are easily masked by normal variation or noise due to their small magnitudes. They seem more likely to alter the probability distribution rather than cause obvious parameter changes. As a result, particular techniques based on specific distance measures have been applied to process control over the past few decades to detect such faults, such as *f-divergences*: Kullback–Leibler (KL) divergence [6–8], Jensen–Shannon divergence [9,10]; and *integral probability metric (IPMs)*: Wasserstein distance [11], maximum mean discrepancy [12], etc.

The divergence-based detection identifies out-of-control conditions in a process by monitoring the dissimilarity between the distribution of the current monitoring process and a reference one [6,8,11,13,14]. KL divergence, frequently employed within the PCA framework for fault detection, facilitates the comparison of distributions along principal axes, demonstrating good performance in detecting incipient faults and providing insights into fault severity [6–8]. However, computing the KL divergence in a non-parametric way can be computationally challenging, particularly in high-dimensional spaces. Given the uncertainties inherent in industrial engineering and the complexities of chemical reactions, it is unrealistic to assume that real-world process data adhere to a specific parametric distribution, such as Gaussian. Furthermore, in many cases, prior knowledge regarding the shape or form of the underlying probability density function (PDF) is limited or unavailable. This lack of information makes it difficult to rely on traditional statistical methods that assume a particular distribution for process monitoring and fault detection. Therefore, if a divergence can be efficiently estimated or calculated, even in the absence of knowledge about the true distribution or its conformity to typical parametric assumptions, the resulting fault detection and isolation framework will exhibit both versatility and resilience.

In contrast to the above-mentioned *f-divergences* measures, which become challenging to compute with an increase in data dimensions, and *IPMs*, which entail high computational complexity, the Cauchy–Schwarz (CS) divergence offers a straightforward approach to quantifying distributional dissimilarity through the measurement of the tightness of an inequality pertaining to two distributions [15,16]. Notably, it boasts a closed-form expression for mixture-of-Gaussians and imposes no prior parametric assumptions on the underlying data distribution. This flexibility of the CS divergence renders it well-suited for a range of practical applications, particularly when the underlying data distribution is unknown or deviates from simple parametric forms. Such benefits make CS divergence an enticing choice for fault detection tasks. Indeed, classical CS divergence has been employed in process monitoring—measuring the angle between two datasets after feature extraction within a kernel entropy component analysis framework [17,18]—this strategy is fundamentally different from our proposal, which emphasizes joint distribution from a probabilistic perspective.

To prevent undesired accidents and maintain system safety, fault isolation is a main function for prompt handling and mitigation of the detected deviations [19]. Prompt and accurate fault isolation is imperative to ascertain the key observation variables that are critical for fault diagnosis. However, this remains a challenging task due to the complexity of the processes involved and the propagation of faults [2,20–22]. In the field of process monitoring, the traditional contribution plots and reconstruction-based contribution methods are two typical classes tailored for fault isolation. Contribution plots are the most popular tool for identifying which variables are pushing the monitoring statistics out of their control limits without prior knowledge [23–25]. However, it assumes that the fault-free variables after fault occurs follows the same distribution as that under normal operating conditions, which may not always be the case. As an alternative strategy, reconstruction-based contribution methods rely on historical fault data to extract fault features and build models for fault reconstruction and classification [26,27]. This method, however, demands a substantial collection of historical fault data, which may be difficult to be obtained in practice. Besides, both of them may have difficulty handling unknown disturbances that are not covered by historical fault data [28]. Even for divergence-based fault diagnosis, they almost involve calculating the contribution rate of each variable being faulty [9–11,13,29], which thus has the same problem as contribution plot. Additionally, to take advantage of the aforementioned detection framework based on CS divergence, we tend to develop an isolation approach from the perspective of conditional probability distributions. Due to the physical connections across sensors and the inherent reaction mechanisms, a compromised or faulty sensor is likely to alter the dependence of other sensors on it [30,31]. Recognizing this, the variation in dependency between a faulty sensor and other remaining sensors (which can be characterized by their conditional distributions) becomes a potent means for fault isolation. Nevertheless, accurately measuring the dissimilarity between two conditional distributions is a challenging task. While measures like *f-divergences* and *IPMs* present potential solutions, they do not seamlessly adapt to this specific context. This predicament led us to propose a novel fault isolation approach leveraging the conditional CS divergence [16]. This method simplifies the task of online incipient fault diagnosis without requiring priori fault information, as we will elucidate in the upcoming sections.

Our main contributions are multi-fold:

1. **Novel scheme:** A statistical divergence-based fault detection and isolation scheme is developed from a probabilistic perspective, combining both CS divergence (for incipient fault detection) and conditional CS divergence (for fault isolation). The proposed monitoring statistics are derived from the perspective of changes in process distribution, utilizing CS divergence, and do not rely on any prior parametric assumptions regarding the underlying data distribution.
2. **Detection accuracy:** Experiments, conducted on both synthetic data and benchmarks of the continuous stirred-tank reactor (CSTR) and continuous stirred-tank heater (CSTH) processes, indicate that the CS divergence based fault detection achieves comparable or even higher detection rates than prevalent divergence based detection approaches, such as KL divergence [6], Mahalanobis distance [13], and Wasserstein distance [11]. Moreover, it maintains a reduced false alarm rate while ensuring expedited detection delays.

Abbreviation			
CCR	Continuous Catalytic Reforming	CS	Cauchy–Schwarz
CSTH	Continuous Stirred-tank Heater	CSTR	Continuous Stirred-tank Reactor
DCS	Distributed Control System	FAR	False Alarm Rate
FDD	Fault Detection Delay	FDR	Fault Detection Rate
FNR	Fault-to-noise Ratios	*IPMs*	Integral Probability Metric
KDE	Kernel Density Estimator	KL	Kullback–Leibler
PCA	Principal Component Analysis	PDF	Probability Density Function
SMD	Statistics Mahalanobis Distance	WD	Wasserstein Distance

3. **Explanatory isolation:** The fault isolation technique, grounded in conditional probability distributions, leverages conditional CS divergence to effectively identify altered dependencies between a faulty sensor and other remaining sensors. This methodology offers explanatory insights of which sensors are affected by fault propagation, thereby providing better-informed decision-making for maintaining process safety and risk assessment.
4. **Scalability:** Given the dynamics and non-stationary inherent in industrial processes, we have enhanced an online CS divergence-based process monitoring technique by incorporating change point detection. This approach has been verified on large scale monitoring sites of an actual continuous reforming (CCR) process in China.

The remainder of this paper is organized as follows. We first introduce the definition of CS divergence and conditional CS divergence in Section 2. We then describe our proposed probabilistic framework of fault detection and isolation in Section 3, includes the CS divergence based incipient fault detection, conditional CS divergence based fault isolation, and an overall dynamic process monitoring framework in practical process. Experiments on synthetic, CSTH benchmark and an actual CCR process are performed in Section 4. We conclude this work in Section 5.

Notations. Throughout this paper, scalars are denoted by lowercase letters (e.g., x), vectors appear as lowercase boldface letters (e.g., $\mathbf{x}$), and matrices are indicated by uppercase letters (e.g., X). The ith row of a matrix X is declared by the row vector $\mathbf{x}^i$, while the jth column is indicated with the column vector $\mathbf{x}_j$. Moreover, subscript indicates time (or sample) index, superscript indicates variable (or sensor) index. $p_r(\mathbf{x})$ wrt. $p_c(\mathbf{x})$ are denoted as two distributions of the reference state and the current monitoring process.

2. Background knowledge

2.1. Problem formulation

Let a multivariate system with m sensors, gather the process variables $\aleph = \{\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^m\}$, if a fault happens in sensor l $(1 \le l \le m)$[1], the faulty measurements observed along this sensor are represented by $\mathbf{x}^l$ and the direction of this fault is denoted as f_l.

Given the monitoring observations $\psi_c = \{\mathbf{x}_{c,i}\}_{i=1}^N$ and a reference state $\psi_r = \{\mathbf{x}_{r,i}\}_{i=1}^M$ of the steady process that are assumed to be independently and identically distributed (*i.i.d.*) with PDFs $p_r(\mathbf{x})$ and $p_c(\mathbf{x})$ respectively, fault detection is to determine whether the current process exhibits any deviation from its steady-state operations. Due to the physical connections across sensors and the inherent reaction mechanisms, any independent manipulation of a subset of sensors will be noticeable when considering their joint distribution [30,31]. This interdependence offers a holistic view of the system's state. Specifically, the dissimilarity between ψ_r and ψ_c can be quantified by the divergence between their respective PDFs, denote as $D(p_r(\mathbf{x}); p_c(\mathbf{x}))$. Such divergences have demonstrated strong performance in detecting incipient faults and provide insights into fault severity [11–13].

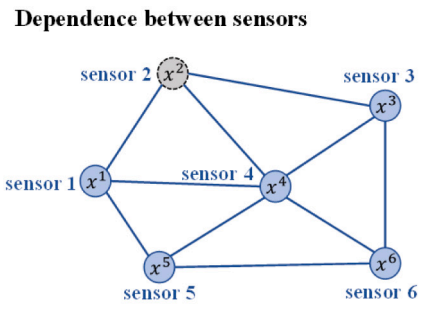
Fault isolation is to identify the most likely fault sensor or variable $\mathbf{x}^l$, the innovation of this paper lies in reformulating this task as a conditional distribution-based probabilistic problem,

$$\mathbf{x}^l = \arg\max_{\mathbf{x}^l} D(p_r(\mathbf{x}^l|X^{-l}), p_c(\mathbf{x}^l|X^{-l})), \tag{1}$$

here $X^{-l} = [\mathbf{x}^1, \dots, \mathbf{x}^{l-1}, \mathbf{x}^{l+1}, \dots, \mathbf{x}^m]^T$ is a remaining sensor stream that includes all sensors excluding the fault sensor $\mathbf{x}^l$ (e.g., $X^{-2} = [\mathbf{x}^1, \mathbf{x}^3, \mathbf{x}^4, \dots, \mathbf{x}^m]^T$), $p_c(\mathbf{x}^l|X^{-l})$ and $p_r(\mathbf{x}^l|X^{-l})$ are the conditional probability of $\mathbf{x}^l$. By leveraging the stability and dynamics of dependencies across sensors, the variable with the maximum conditional dissimilarity D can be considered as the primary contributor to the observed faulty state.

To capture the changes between probability distributions with high sensitivity, we propose the development of a monitoring statistic based on CS divergence, a probabilistic measure taken from Information Theory, for incipient fault detection, denoted as $D_{\mathrm{CS}}(p_r(\mathbf{x}); p_c(\mathbf{x}))$; and a monitoring statistic based on conditional CS divergence, from a conditional probabilistic perspective, for fault isolation, denoted as $D_{\mathrm{CS}}(p_{k-1}(\mathbf{x}^l|X^{-l}); p_k(\mathbf{x}^l|X^{-l}))$, as Fig. 1 shown.

[1] Here taking the univariate fault as example, multivariate faults can be also analyzed similarly.



- **Fault detection:**
  Divergence between joint distributions $p_r(\mathbf{x})$ and $p_c(\mathbf{x})$

$$\begin{cases} H_0: \forall i, p_r(x^1, x^2, \cdots, x^6) = p_c(x^1, x^2, \cdots, x^6) \\ H_1: \exists i, p_r(x^1, x^2, \cdots, x^6) \neq p_c(x^1, x^2, \cdots, x^6) \end{cases}$$

- **Fault isolation:**
  Divergence between conditional distributions $p_r(x^i|\mathbf{x}^{-i})$ and $p_c(x^i|\mathbf{x}^{-i})$

$$\begin{cases} H_0: \forall i, p_r(x^i|x^1, \cdots, x^{i-1}, x^{i+1}, \cdots, x^6) = p_c(x^i|x^1, \cdots, x^{i-1}, x^{i+1}, \cdots, x^6) \\ H_1: \exists i, p_r(x^i|x^1, \cdots, x^{i-1}, x^{i+1}, \cdots, x^6) \neq p_c(x^i|x^1, \cdots, x^{i-1}, x^{i+1}, \cdots, x^6) \end{cases}$$

Fig. 1. An illustrative example of fault detection and isolation in a probabilistic perspective. $p_r(\mathbf{x})$ wrt. $p_c(\mathbf{x})$ are two distributions of the reference state and the current monitoring process.

2.2. The Cauchy–Schwarz divergence and conditional Cauchy–Schwarz divergence

In the early 2000s, Principe et al. [15,32] suggested a way to quantify the distributional dissimilarity simply by measuring the tightness of the famed Cauchy–Schwarz (CS) inequality associated with two distributions:

$$\left| \int p_r(\mathbf{x}) p_c(\mathbf{x}) d\mathbf{x} \right|^2 \leq \int |\, p_r(\mathbf{x})\,|^2\, d\mathbf{x} \int |\, p_c(\mathbf{x})\,|^2\, d\mathbf{x}, \tag{2}$$

with equality if and only if $p_r(\mathbf{x})$ and $p_c(\mathbf{x})$ are linear dependent, a measure of the "distance" between PDFs can be defined, which was named the CS divergence, with:

$$\begin{aligned} D_{\mathrm{CS}}(p_r(\mathbf{x}); p_c(\mathbf{x})) &= -\log\left(\frac{\left| \int p_r(\mathbf{x}) p_c(\mathbf{x}) d\mathbf{x} \right|^2}{\int |\, p_r(\mathbf{x})\,|^2\, d\mathbf{x} \int |\, p_c(\mathbf{x})\,|^2\, d\mathbf{x}} \right) \\ &= -2\log\left(\int p_r(\mathbf{x}) p_c(\mathbf{x}) d\mathbf{x} \right) + \log\left(\int p_r(\mathbf{x})^2 d\mathbf{x} \right) + \log\left(\int p_c(\mathbf{x})^2 d\mathbf{x} \right). \end{aligned} \tag{3}$$

The CS divergence enjoys a few appealing properties [33]. For example, it has closed-form expression for mixture-of-Gaussians [34], a property that KL divergence does not hold [16]. Moreover, it can be simply evaluated with the kernel density estimator (KDE).

Given $\{\mathbf{x}_{r,i}\}_{i=1}^M$ and $\{\mathbf{x}_{c,i}\}_{i=1}^N$ both in $\mathbb{R}^m$, drawn *i.i.d.* from $p_r(\mathbf{x})$ and $p_c(\mathbf{x})$ respectively. Using the KDE with Gaussian kernel $G_\sigma(\cdot) = \exp(-\frac{\|\cdot\|^2}{2\sigma^2})$, Eq. (3) can be estimated as [33]:

$$\begin{aligned} \widehat{D}_{\mathrm{CS}}(p_r(\mathbf{x}); p_c(\mathbf{x})) &= \log\left(\frac{1}{M^2} \sum_{i,j=1}^{M} G_{\sqrt{2}\sigma}(\mathbf{x}_{r,i} - \mathbf{x}_{r,j}) \right) \\ &+ \log\left(\frac{1}{N^2} \sum_{i,j=1}^{N} G_{\sqrt{2}\sigma}(\mathbf{x}_{c,i} - \mathbf{x}_{c,j}) \right) - 2\log\left(\frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} G_{\sqrt{2}\sigma}(\mathbf{x}_{r,i} - \mathbf{x}_{c,j}) \right). \end{aligned} \tag{4}$$

The non-parametric estimation used in this calculation is not confined to specific parametric forms or distributions, providing flexibility for the design of fault detection and isolation framework and enabling its adaptation to a wide variety of industrial processes and scenarios.

The CS divergence for two conditional distributions $p_r(\mathbf{y}|\mathbf{x})$ and $p_c(\mathbf{y}|\mathbf{x})$, in which $\mathbf{y} \in \mathbb{R}$ (e.g., $\mathbf{x}^l$) and $\mathbf{x} \in \mathbb{R}^{m-1}$ (e.g., X^{-l}), can be expressed naturally as:

$$\begin{aligned} D_{\mathrm{CS}}(p_r(\mathbf{y}|\mathbf{x}); p_c(\mathbf{y}|\mathbf{x})) &= -2\log\left(\int_{\mathcal{X}} \int_{\mathcal{Y}} p_r(\mathbf{y}|\mathbf{x}) p_c(\mathbf{y}|\mathbf{x}) d\mathbf{x} d\mathbf{y} \right) \\ &+ \log\left(\int_{\mathcal{X}} \int_{\mathcal{Y}} p_r^2(\mathbf{y}|\mathbf{x}) d\mathbf{x} d\mathbf{y} \right) + \log\left(\int_{\mathcal{X}} \int_{\mathcal{Y}} p_c^2(\mathbf{y}|\mathbf{x}) d\mathbf{x} d\mathbf{y} \right) \\ &= -2\log\left(\int_{\mathcal{X}} \int_{\mathcal{Y}} \frac{p_r(\mathbf{x},\mathbf{y}) p_c(\mathbf{x},\mathbf{y})}{p_r(\mathbf{x}) p_c(\mathbf{x})} d\mathbf{x} d\mathbf{y} \right) \\ &+ \log\left(\int_{\mathcal{X}} \int_{\mathcal{Y}} \frac{p_r^2(\mathbf{x},\mathbf{y})}{p_r^2(\mathbf{x})} d\mathbf{x} d\mathbf{y} \right) + \log\left(\int_{\mathcal{X}} \int_{\mathcal{Y}} \frac{p_c^2(\mathbf{x},\mathbf{y})}{p_c^2(\mathbf{x})} d\mathbf{x} d\mathbf{y} \right). \end{aligned} \tag{5}$$

which contains two conditional quadratic terms (i.e., $\int_{\mathcal{X}} \int_{\mathcal{Y}} \frac{p_r^2(\mathbf{x},\mathbf{y})}{p_r^2(\mathbf{x})} d\mathbf{x} d\mathbf{y}$ and $\int_{\mathcal{X}} \int_{\mathcal{Y}} \frac{p_c^2(\mathbf{x},\mathbf{y})}{p_c^2(\mathbf{x})} d\mathbf{x} d\mathbf{y}$) and a cross term (i.e., $\int_{\mathcal{X}} \int_{\mathcal{Y}} \frac{p_r(\mathbf{x},\mathbf{y}) p_c(\mathbf{x},\mathbf{y})}{p_r(\mathbf{x}) p_c(\mathbf{x})} d\mathbf{x} d\mathbf{y}$). Similarly, Eq. (5) can be elegantly estimated by the KDE.

Given observations $\psi_r = \{(\mathbf{x}_{r,i}, \mathbf{y}_{r,i})\}_{i=1}^M$ and $\psi_c = \{(\mathbf{x}_{c,i}, \mathbf{y}_{c,i})\}_{i=1}^N$ which are sampled from distributions $p_r(\mathbf{x}, \mathbf{y})$ and $p_c(\mathbf{x}, \mathbf{y})$, respectively. Let K^r and L^r denote, respectively, the Gram matrices for variables $\mathbf{x}$ and $\mathbf{y}$ in the distribution p_r (e.g., $K_{ij}^r = \kappa(\mathbf{x}_{r,i} - \mathbf{x}_{r,j})$).

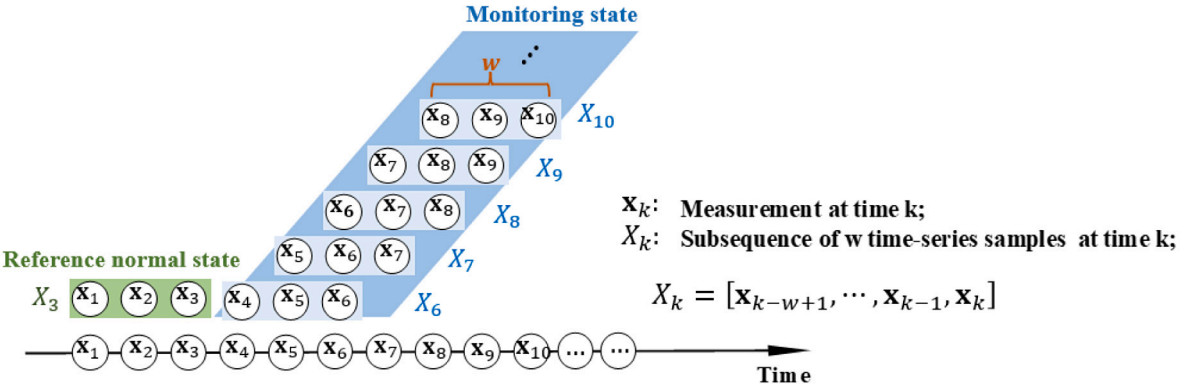


Fig. 2. An illustrative example of monitoring measurements, in which window length $w = 3$.

Similarly, let K^c and L^c denote, respectively, the Gram matrices for the variables $\mathbf{x}$ and $\mathbf{y}$ in the distribution p_c. Meanwhile, let $K^{rc} \in \mathbb{R}^{M\times N}$ (i.e., $K^{rc}_{ij} = \kappa(\mathbf{x}_{r,i} - \mathbf{x}_{c,j})$) denote the Gram matrix from distribution p_r to distribution p_c for variable $\mathbf{x}$, and $L^{rc} \in \mathbb{R}^{M\times N}$ the Gram matrix from distribution p_r to distribution p_c for variable $\mathbf{y}$. Similarly, let $K^{cr} \in \mathbb{R}^{N\times M}$ (i.e., $K^{cr}_{ij} = \kappa(\mathbf{x}_{c,i} - \mathbf{x}_{r,j})$) denote the Gram matrix from distribution p_c to distribution p_r for variable $\mathbf{x}$, and $L^{cr} \in \mathbb{R}^{N\times M}$ the Gram matrix from distribution p_c to distribution p_r for variable $\mathbf{y}$. $D_{\mathrm{CS}}(p_r(\mathbf{y}|\mathbf{x}); p_c(\mathbf{y}|\mathbf{x}))$ can be estimated by [16]:

$$\begin{aligned}\widehat{D}_{\mathrm{CS}}(p_r(\mathbf{y}|\mathbf{x}); p_c(\mathbf{y}|\mathbf{x})) \approx &\log\left(\sum_{j=1}^{M}\left(\frac{\sum_{i=1}^{M} K^r_{ji} L^r_{ji}}{(\sum_{i=1}^{M} K^r_{ji})^2}\right)\right) + \log\left(\sum_{j=1}^{N}\left(\frac{\sum_{i=1}^{N} K^c_{ji} L^c_{ji}}{(\sum_{i=1}^{N} K^c_{ji})^2}\right)\right)\\ &- \log\left(\sum_{j=1}^{M}\left(\frac{\sum_{i=1}^{N} K^{rc}_{ji} L^{rc}_{ji}}{(\sum_{i=1}^{M} K^r_{ji})(\sum_{i=1}^{N} K^{rc}_{ji})}\right)\right) - \log\left(\sum_{j=1}^{N}\left(\frac{\sum_{i=1}^{M} K^{cr}_{ji} L^{cr}_{ji}}{(\sum_{i=1}^{M} K^{cr}_{ji})(\sum_{i=1}^{N} K^c_{ji})}\right)\right),\end{aligned} \tag{6}$$

where κ refers to a Gaussian kernel with width σ and takes the form of $\kappa(x - y) = \exp\left(-\frac{\|x-y\|^2}{2\sigma^2}\right)$.

3. Incipient fault diagnosis with Cauchy–Schwarz divergence

In this section, we develop a general fault detection and isolation scheme that combines CS and conditional CS divergences. To enhance the practicality of our scheme for real-world industrial processes, we also develop a dynamic monitoring mechanism using change point detection.

3.1. The CS divergence based incipient fault detection

In multivariate process systems, deviations in internal dynamics can indicate the presence of faults. However, these faults might not always manifest themselves in individual sensors. Instead, they can cause shifts in the system's overall distribution. This subsection introduces a detection method on the basis of CS divergence to identify these faults with high sensitivity.

For incipient fault, samples belonging to normal and abnormal conditions usually overlap to a large extent [4,10]. Window-based fault detection methods can partially reduce the missed detection rate by utilizing statistical information among measurements, and alleviate data overlap [4]. Given the process observations $\aleph = \{\mathbf{x}_1, \mathbf{x}_2, \ldots\} : \mathbf{x}_i \in \mathbb{R}^m$, we construct a "time lag shift" observer X_k at time $k, k \in [w + 1, n]$,

$$\begin{aligned}X_k &= \left[\mathbf{x}_{k-w+1}, \ldots, \mathbf{x}_{k-1}, \mathbf{x}_k\right]\\ &= \begin{bmatrix} x^1_{k-w+1} & x^1_{k-w+2} & \cdots & x^1_k \\ x^2_{k-w+1} & x^2_{k-w+2} & \cdots & x^2_k \\ \vdots & \vdots & \ddots & \vdots \\ x^m_{k-w+1} & x^m_{k-w+2} & \cdots & x^m_k \end{bmatrix} \in \mathbb{R}^{m\times w},\end{aligned} \tag{7}$$

w is the length of sliding window. X_k is the moving window data that are consecutively updated along time direction. Note that X_k has the special form of a *Hankel* matrix which is widely used in system identification. Fig. 2 illustrates the time-lagged matrix X_k [35].

As process variables are jointly monitored, the data covering the measurement data under normal condition are used as reference distribution, $\{\mathbf{x}_{r-w+1}, \ldots, \mathbf{x}_{r-1}, \mathbf{x}_r\}$, and then a moving window $\{\mathbf{x}_{c-w+1}, \ldots, \mathbf{x}_{c-1}, \mathbf{x}_c\}$ with length of w is used for dissimilarity analysis to evaluate the changes of distribution structure along time direction.

The dissimilarity between two PDFs, $p_r(\mathbf{x})$ and $p_c(\mathbf{x})$, of the above observations can be quantified by CS divergence, $D_{\mathrm{CS}}(p_r(\mathbf{x}); p_c(\mathbf{x}))$ with Eq. (3). When using D_{CS} for process monitoring, its upper control threshold D_{cl} can be estimated by empirical method with a given confidence level η during the offline stage. Following the occurrence of a fault, the state may shift or oscillate concomitant with the propagation of fault. Therefore, online monitoring aims to quantify the deviation of the operating process by CS divergence. It is possible to detect whether the process is "in-control" by applying the criterion: $D_{\mathrm{CS}}(k) \geq D_{cl}$.

In our work, fault detection based on CS divergence is performed in the observation space. Considering the uncertainties inherent in industrial environments, wherein data may deviate from idealized or assumed distributions, we have adopted an adaptive kernel computation method for online kernel density estimation. This method flexibly adjusts the kernel size based on real-time changes in the data, thereby ensuring that the estimation results are more aligned with the actual situation. Indeed, it serves as a versatile and flexible framework that can integrate subspace projection or feature extraction to further enhance its detection accuracy. Additionally, with memory items recorded in sliding window, the Type-II error (i.e., fails to reject a false null-hypothesis) may be effectively avoid consequently [36].

3.2. Fault isolation by conditional CS divergence

Fault isolation is performed to identify the variables which are responsible for the changes of distribution structure, i.e., changes in probabilistic dependencies. With the occurrence of fault, the conditional dependence between the faulty variable and other variables will differ from that under normal states. Thus, it is feasible to identify faulty variables by investigating whether there is a deviant change in their conditional distributions. Considering the power of CS divergence in distribution monitoring, we aim to develop a fault isolation strategy based on conditional CS divergence.

Given the above mentioned multivariate process $\aleph$, let us re-denote the "time lag shift" observer X_k at time k as

$$X_k \triangleq [\mathbf{x}^1, \ldots, \mathbf{x}^l, \ldots, \mathbf{x}^m]_k^T \in \mathbb{R}^{m \times w}, \tag{8}$$

$\mathbf{x}^l$ $(1 \leq l \leq m)$ is the lth variable consists of w observations within a sliding window. To better describe fault characteristics, we isolate each variable $\mathbf{x}^l$ individually during the monitoring phase to observe its specific contribution and impact on the overall monitoring status. In other words, we examine how the variable $\mathbf{x}^l$ depends on the remaining sensor streams X^{-l} to get a comprehensive understanding of the system's behavior. This dependencies across sensors can be captured by conditional probability distributions, denoted as $p(\mathbf{x}^l|X^{-l})$. Thus, we can utilize the dissimilarity between two conditional distributions, $p_c(\mathbf{x}^l|X^{-l})$ and $p_r(\mathbf{x}^l|X^{-l})$, for fault isolation. The logic behind this is that the steady-state process should contain similar internal dynamics, in the sense that an isolation model trained on reference state should perform satisfactory on monitoring phase as well. If the stability and dynamism of these dependencies across sensors deviate, the variable responsible for the greatest deviation is deemed as the fault variable we seek to identify. That is, fault isolation is the problem of identifying the most likely fault sensor or variable $\mathbf{x}^l$ given its remaining sensor streams X^{-l}:

$$\mathbf{x}^l = \arg\max_{\mathbf{x}^l} D_{\mathrm{CS}}(p_r(\mathbf{x}^l|X^{-l}); p_c(\mathbf{x}^l|X^{-l})). \tag{9}$$

The variable with the maximum D_{CS} can be considered as the primary contributor to the observed faulty state.

To identify the specific variable or sensor which is in disrepair or affected by the fault, we propose the normalized total conditional divergence (nTCD) to estimate the ratio of conditional divergence,

$$T(l) = \frac{D_{\mathrm{CS}}(p_r(\mathbf{x}^l|X^{-l}); p_c(\mathbf{x}^l|X^{-l}))}{\sum_{l=1}^{m} D_{\mathrm{CS}}(p_r(\mathbf{x}^l|X^{-l}); p_c(\mathbf{x}^l|X^{-l}))}. \tag{10}$$

Variables with the maximum deviation ratio can be considered as the main sensor related to the fault, recorded as $\mathbf{x}^{l*}$. In Eq. (10), each dimension of $D_{\mathrm{CS}}(p_r(\mathbf{x}^l|X^{-l}); p_c(\mathbf{x}^l|X^{-l}))$ should be mean centered and normalized with reference mean μ_{CS} and standard deviation σ_{CS} of the training set,

$$\hat{D}_{\mathrm{CS}}(p_r(\mathbf{x}^l|X^{-l}); p_c(\mathbf{x}^l|X^{-l})) = \frac{D_{\mathrm{CS}}(p_r(\mathbf{x}^l|X^{-l}); p_c(\mathbf{x}^l|X^{-l})) - \mu_{\mathrm{CS}}}{\sigma_{\mathrm{CS}}^2}. \tag{11}$$

By monitoring the multivariate data with a sliding window sequentially, fault isolation resorts to the following hypothesis test for identification of a fault variable $\mathbf{x}^l$ on X^{-l},

$$\begin{cases} \mathcal{H}_0 : \forall l, p_c(\mathbf{x}^l|X^{-l}) = p_r(\mathbf{x}^l|X^{-l}), \\ \mathcal{H}_1 : \exists l, p_c(\mathbf{x}^l|X^{-l}) \neq p_r(\mathbf{x}^l|X^{-l}). \end{cases} \tag{12}$$

Eq. (12) relies on the dependencies between $\mathbf{x}^l$ and X^{-l}. This fault isolation strategy enables us to individually calculate the anomaly of each sensor and dynamically assess its contribution to the current state. Importantly, it operates within the observation space, eliminating the need for prior fault information or historical samples. This flexibility allows the method to adapt seamlessly to various process scenarios without the need for extensive preprocessing or assumption fitting. By leveraging relative discrimination instead of fixed control limits or thresholds, we can precisely identify specific sensors that have been compromised or affected. This brings explanatory benefits to fault isolation, as operators need to understand the reasoning behind the algorithm's decisions to trust its results.

3.3. Dynamic monitoring using change point detection

For actual industrial process, faults may arise or develop due to various dynamic changes. If corresponding regulation can be implemented during the fault symptom stage, it will effectively prevent faults from occurring. This subsection discusses a more flexible process monitoring strategy using change point detection.

Given two phased observations $\{\mathbf{x}_{k-w_c}, \dots \mathbf{x}_{k-2}, \mathbf{x}_{k-1}\}$ and $\{\mathbf{x}_{k+1}, \dots \mathbf{x}_{k+w_c-2}, \mathbf{x}_{k+w_c-1}\}$ before and after time index k, written as $\mathbf{x}_{k-w_c:k-1}$ and $\mathbf{x}_{k:k+w_c-1}$ correspondingly, if the divergence between their pdfs $p(\mathbf{x}_{k-w_c:k-1})$ and $p(\mathbf{x}_{k:k+w_c-1})$ reaches a peak or exceeds a threshold δ:

$$D(p(\mathbf{x}_{k-w_c:k-1}); p(\mathbf{x}_{k:k+w_c-1})) \geq \delta, \tag{13}$$

fault or change in dynamic characteristics of processes may occurs at time k. Here, w_c is the length of retrospective interval. The threshold δ is a minimum peak height within the interval $[k - w_c, k + (w_c - 1)]$, and it can be customized by industry-specific knowledge. In our experiment, we utilize the average divergence value calculated over the retrospective interval. Specifically, when the divergence value D surpasses δ within its respective traceable interval, we classify that point as a change point. Change point detection gives support to the hypothesis of process changes or fault occurrence through comparing interval integral distributions [35]. The higher $D(p(\mathbf{x}_{k-w_c:k-1}); p(\mathbf{x}_{k:k+(w_c-1)}))$ is, the more likely the monitoring point is an abrupt change of distribution structure along time direction.

For process monitoring, we aim to assess whether the conditional distribution of each sensor has shifted based on the other sensors. If the conditional CS divergence reaches a peak, an alarm will be triggered,

$$\sum_{l=1}^{m} D_{\mathrm{CS}}(p_{k-1}(\mathbf{x}^l|X^{-l}); p_{k+w_c-1}(\mathbf{x}^l|X^{-l})) \geq \delta, \tag{14}$$

this alarm time is regarded as c_k, and the alarm time set is formed as $C = \{c_1, c_2, \dots, c_{k-1}, c_k, \dots\}$. Note that, the alarm is true only if there exists a real fault at step k^* such that $k \in [k^* - \tau, k^* + \tau]$, in which τ is the acceptable maximum detection delay, $w_c \leq \tau$. With process changes alarmed, this developed process monitoring strategy allows operators to detect incipient faults earlier. Besides, to localize which sensor is compromised or affected, isolation should be analyzed through Eq. (10) within the interval $[c_{k-1}, c_k]$ of two adjacent change alarm points c_{k-1} and c_k., the alarm variable is recorded as $\mathbf{x}^{s*}$.

This online process monitoring method using change point detection based on conditional CS divergence tends to give more robust and accurate detection since the retrospective interval w_c can be set to a value smaller than the fault incubation period. Actually, it can quantitatively and flexibly describe the correlation and even variable transmission relationships between sensors, which is more suitable for the supervision of processes in practical engineering that may undergo steady-state switching due to operating conditions, control loops, etc. Algorithm 1 summarizes the offline modeling and online monitoring of our proposed fault detection and isolation framework.

4. Applications to process monitoring

In the subsequent sections, we conduct experiments on a simulation dataset (subSection 4.1), benchmark datasets from the CSTR process (subSection 4.2) and CSTH process (subSection 4.3), as well as a real-world CCR process (subSection 4.4), to showcase the exceptional capabilities of our fault detection and isolation method based on CS divergence in process monitoring. We also verified the performance improvement of CS divergence based detection when integrated with latent projection.

For quantitative comparison, three generally used metrics, namely fault detection rate (FDR), false alarm rate (FAR) and fault detection delay (FDD) are employed for performance evaluation [1,37–39]. FDR is the probability of event where an alarm is raised when a fault really occurs,

$$\mathrm{FDR} = \mathrm{prob}(D \geq D_{\mathrm{cl}} | \mathrm{fault} \neq 0), \tag{15}$$

where D and D_{cl} are respectively the divergence index and its corresponding control limit. By contrast, FAR is the percentage of samples under normal state but are identified as faults,

$$\mathrm{FAR} = \mathrm{prob}(D \geq D_{\mathrm{cl}} | \mathrm{fault} = 0). \tag{16}$$

Usually, a higher FDR and a lower FAR is expected in fault detection methods. Additionally, we look forward to a smaller FDD because it reflects the sensitivity of the algorithm,

$$\mathrm{FDD} = \mathrm{Num}(D\langle D_{\mathrm{cl}} | \mathrm{fault} = 0 \rightarrow D \geq D_{\mathrm{cl}} | \mathrm{fault} \neq 0). \tag{17}$$

4.1. Numerical simulation

Motivated by [4,13], we consider a simulation example generated by the following multivariate process:

$$\mathbf{x} = A\mathbf{s} + \mathbf{e}, \quad \mathbf{s} = [s_1, s_2, s_3, s_4]^T, \tag{18}$$

Algorithm 1 The fault detection and isolation framework with CS divergence.

Input: The offline process observations $\{\mathbf{x}_1, \mathbf{x}_2, \cdots\}$ and online measurements $\{\mathbf{x}_{\text{test},1}, \mathbf{x}_{\text{test},2}, \cdots\}$; The sliding window size w; The retrospective reaction periods w_c.
Output: Alarm or not, alarm variable $\mathbf{x}^{s*}$; Fault or not, faulty variable $\mathbf{x}^{l*}$.
Offline modeling:
Construct the reference matrix X_{ref};
for $k = w + 1$ to n **do**
  Construct X_k and obtain $D_{\text{CS}}(p_k(\mathbf{x}); p_{\text{ref}}(\mathbf{x}))$ with Eq. (3);
  for $l = 1$ to m **do**
    Compute $D_{\text{CS}}(p_{k-1}(\mathbf{x}^l|X^{-l}); p_k(\mathbf{x}^l|X^{-l}))$ with Eq. (5);
  end for
end for
Determine the control limit D_{cl} at the significance level η.
Calculate reference mean μ_{CS}, standard deviation σ_{CS}.
return D_{cl}; μ_{CS}; σ_{CS}.
Online monitoring:
Construct the reference matrix $X_{\text{test,ref}}$;
while End of process not reached **do**
  Construct $X_{\text{test},k}$ and obtain $D_{\text{test,CS}}(p_k(\mathbf{x}); p_{\text{ref}}(\mathbf{x}))$ with Eq. (3);
  for $l = 1$ to m **do**
    Compute $D_{\text{test,CS}}(p_{k-w_c+1}(\mathbf{x}^l|X^{-l}); p_k(\mathbf{x}^l|X^{-l}))$;
    Compute $D_{\text{test,CS}}(p_{k-1}(\mathbf{x}^l|X^{-l}); p_k(\mathbf{x}^l|X^{-l}))$;
  end for
  if $D_{\text{test,CS}}(p_{k-w_c+1}(\mathbf{x}^l|X^{-l}); p_k(\mathbf{x}^l|X^{-l})) \geq \delta$ **then**
    Alarm the change position, recorded as c_k;
    for $l = 1$ to m **do**
      $\sum_{k=c_{k-1}}^{c_k} D_{\text{test,CS}}(p_{k-1}(\mathbf{x}^l|X^{-l}); p_k(\mathbf{x}^l|X^{-l}))$;
    end for
    Variables with the maximum deviation ratio, $\mathbf{x}^{s*}$;
  end if
  if $D_{\text{test,CS}}(p_k(\mathbf{x}); p_{\text{ref}}(\mathbf{x})) \geq D_{\text{cl}}$ **then**
    Alarm the occurrence of fault;
    for $l = 1$ to m **do**
      Normalize the conditional divergence by μ_{CS} and σ_{CS}, and obtain the nTCD $T(l)$ by Eq. (10);
    end for
    Variables with the maximum deviation ratio, $\mathbf{x}^{l*}$.
  end if
  $k = k + 1$; Go back to Step 13.
end while
return Decisions: Alarm or not, alarm variable $\mathbf{x}^{s*}$; Fault or not, faulty variable $\mathbf{x}^{l*}$.

where $\mathbf{x}$ represents the process measurements that result from a linear combination of data sources following a multivariate Gaussian distribution $\mathbf{s} \sim \mathcal{N}(\boldsymbol{\mu}, I)$ with $\boldsymbol{\mu} = [1.2, 0.7, 2.3, 1.7]^T$, A is the coefficient matrix, $A = \begin{bmatrix} 0.55 & 0.72 & 0.60 & 0.54 \\ 0.42 & 0.65 & 0.44 & 0.89 \\ 0.96 & 0.38 & 0.79 & 0.53 \\ 0.57 & 0.93 & 0.07 & 0.09 \\ 0.02 & 0.83 & 0.78 & 0.87 \\ 0.98 & 0.80 & 0.46 & 0.78 \end{bmatrix}$. Here, a fault of sensor bias is considered: $\hat{\mathbf{x}} = \mathbf{x} + \mathbf{f}$, $\hat{\mathbf{x}}$ is the measurement under fault, and $\mathbf{x}$ denotes the fault-free portion. A total of 700 samples were generated under steady-state process conditions, while an additional 300 samples were generated that involved introducing a ramp fault on x_4, as $\hat{x}_4 = x_4 + 0.01(k - 700), k = 701, \ldots, 1000$ [13].

We compare our proposed method, denoted as CS, with state-of-the-art divergence-based detection approaches: KL divergence (KL) [6], Statistics Mahalanobis Distance (SMD) [13] and Wasserstein Distance (WD) [11]. To ensure fairness, we follow the same settings as in [11,13], $w = 100$ and $\eta = 0.01$. In KL and WD, we select 3 principal components to capture more than 95% of the variance. To robustly evaluate the performance, we examined the system's detection capabilities under various conditions corrupted by different noise distributions. Specifically, we considered Gaussian noises $\mathbf{e} \sim \mathcal{N}(0, \sum_{\mathbf{e}})$ with $\sum_{\mathbf{e}} = diag\{0.031, 0.023, 0.192, 0.173, 0.150, 0.113\}$, Gamma noises $\mathbf{e} \sim \Gamma(2, 1)$, and Exponential noises $\mathbf{e} \sim \exp(-1.2)$. To ensure a comprehensive evaluation, we conducted 100 experiments for each noise scenario related to the fault, considering various input biases. Table 1 has summarized the average detection performance. Under Gaussian noise, our method's FDRs in the observation space is comparable to that of the KL divergence method after PCA projection, but with a FAR of 0. Under Gamma noise, our method's detection rate is second only to that of the WD distance in the principal component space, also exhibiting a reduced FAR. Under exponential noise, our method shows improvement in both FDRs and FARs when compared to other methods. Overall, our method has demonstrated promising detection performance across all three noise distributions, achieving a relatively high FDR and lower

Table 1
The FDRs (%) of different methods for the numerical simulations.

No.		KL	KL		SMD	WD		CS
			K_z	K_e		W_z	W_e	
$\mathbf{e} \sim \mathcal{N}(0, \sum_{\mathbf{e}})$	FDR	77.55	71.67	**83.11**	61.55	70.0	79.67	**82.33**
	FAR	0.07	**0**	5.84	1.0	1.0	1.0	**0**
	FDD	67	85	**51**	115	90	61	**53**
$\mathbf{e} \sim \Gamma(2, 1)$	FDR	56.78	55.56	35.67	60.44	**65.44**	62.56	**63.78**
	FAR	0.53	**0**	**0**	1.0	1.0	1.0	**0.4**
	FDD	130	133	193	119	104	112	**109**
$\mathbf{e} \sim \exp(-1.2)$	FDR	**75.56**	71.89	69.11	54.11	54.78	70.44	**72.01**
	FAR	24.10	0.53	10.29	1.0	1.0	1.0	**0**
	FDD	**73**	84	93	138	136	89	**84**

To robustly evaluate the performance, this fault scenario undergoes 100 trials with varying input disturbances. $*_z$ and $*_e$ denote monitoring statistics in principal space and residual space, respectively. $w = 100$, $\eta = 1\%$.

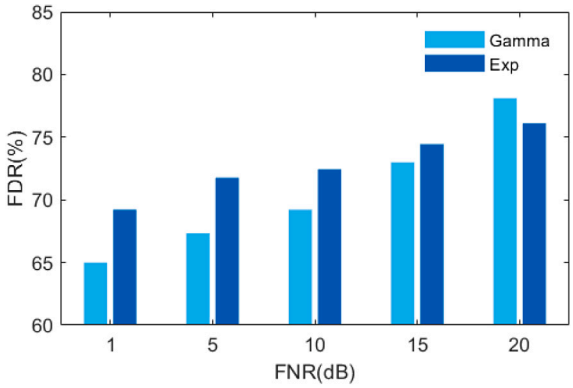


Fig. 3. Detection performance of the proposed CS divergence-based fault detection method along with varying FNR values.

FAR. This represents a favorable balance between FDR and FAR. Additionally, our proposed CS divergence exhibits shorter detection delays, with the FDD of 53 mins, 109 mins and 84 mins. The shorter detection delay facilitates prompt implementation of effective regulatory measures, ensuring a swift restoration of the process to its normal state. However, it is worth noting that due to the slow-varying characteristics of this incipient fault, it can be masked by higher noise levels during its early stages; meanwhile, the use of sliding windows inevitably causes detection delays, thereby decreasing the detection rate. Then, we calculated the detection accuracy across different fault-to-noise ratios (FNR) to evaluate the robustness of our algorithm against noises. The results, as depicted in Fig. 3, demonstrate that our approach remains effective in detecting faults even when the FNR is low. This indicates the algorithm's resilience to noise and its ability to accurately identify faults in challenging conditions. Besides, the FDRs increase concurrent with the escalation of FNR, which underscores the algorithm's adaptability in handling different noise levels.

Fig. 4 illustrates the conditional divergence scores, notably, x_4 exhibits a significant deviation from its steady state, indicating that it is the main variable affected by the fault. This observation aligns with the actual cause, where $\hat{x}_4 = x_4 + f$. Besides, deviations observed within the interval $[427, 700]$, without any alarms in the detection results, can be attributed to system dynamics or operational deviations. This suggests that certain information can be discerned from the conditional distribution, which differs from the overall distribution. Consequently, real-time conditional monitoring becomes crucial as it facilitates timely regulation in response to such deviations.

Moreover, we plot the monitoring results using change point detection in Fig. 5. In Fig. 5(a), the alarm set is $[324, 427, 732]$, which forms 4 stages. While the first stage displays a right-skewed state and the second stage shows a left-skewed state, the variables remain relatively stable under normal conditions. The third stage exhibits an overall upward shift, as seen in Fig. 4, the upper quartile of x_4 increases, resulting in right skewness under normal conditions. However, in the fourth stage, there is a significant upward shift in x_4, indicating a faulty state during this stage.

It is worth noting that the proposed fault detection scheme, based on CS divergence, does not employ any projections of latent variables, such as PCA. This suggests that CS divergence is a more versatile distance metric suited for detecting incipient faults. Of course, PCA projection can be incorporated into this scheme to further enhance its detection capabilities, a claim we will validate in subsequent experiments.

4.2. Application to CSTR process

This section discusses the effectiveness of the proposed method in a closed-loop continuous stirred-tank reactor (CSTR) process, which is a common chemical system that designed especially for simulating incipient faults [40–43]. The dynamic characteristics

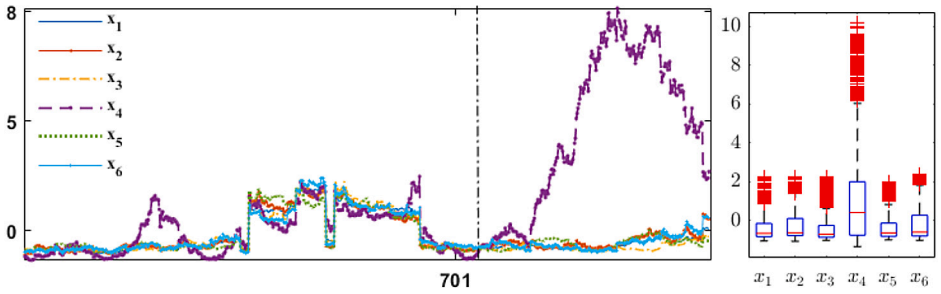
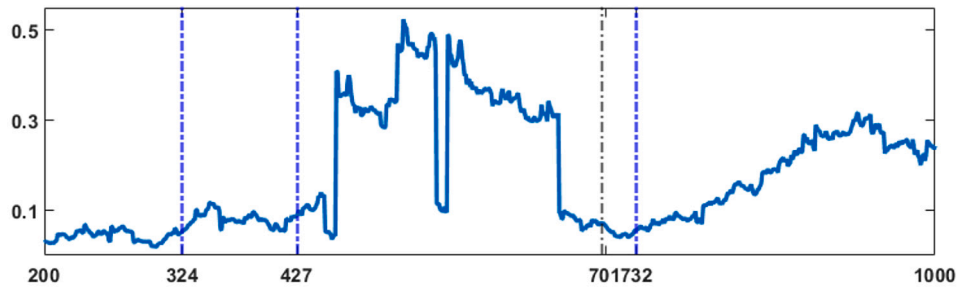
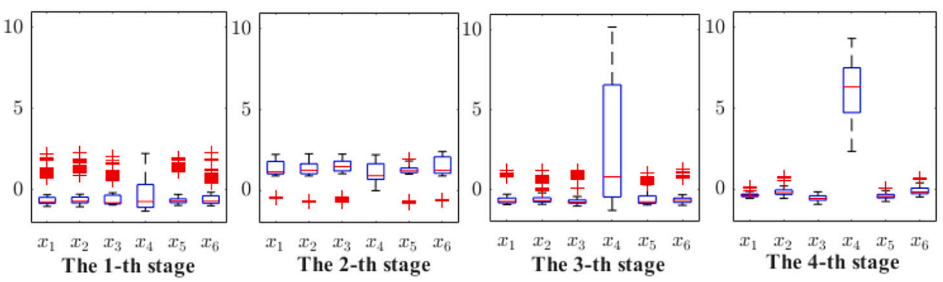


Fig. 4. The conditional divergences for fault isolation in the simulation example.



(a) Conditional CS divergence based change-point detection.



(b) Distributions for each state split by change points.

Fig. 5. The monitoring results of the simulation example. (a) Conditional CS divergence based change point detection. (b) Distributions for each state split by change points.

of CSTR process is described by 3 ordinary differential equations which are mass and energy balances around the system,

$$\frac{dC}{dt} = \frac{Q}{V}(C_i - C) - akC + v_1, \tag{19}$$

$$\frac{dT}{dt} = \frac{Q}{V}(T_i - T) - a\frac{(\Delta H_r)kC}{\rho C_p} - b\frac{UA}{\rho C_p V}(T_i - T) + v_2, \tag{20}$$

$$\frac{dT_c}{dt} = \frac{Q_c}{V_c}(T_{ci} - T_c) + b\frac{UA}{\rho_c C_{pc} V_c}(T_c - T) + v_3, \tag{21}$$

where the system's inputs are $\mathbf{x} = [C_i, T_i, T_{ci}]^T$, the outputs are $\mathbf{y} = [C, T, T_c, Q_c]^T$, v_i is process noise, and k is an Arrhenius-type rate constant with $k = k_0 e^{(\frac{-E}{RT})}$. A schematic diagram of the CSTR and feedback control system is shown in Fig. 6: the tank temperature T is maintained using a cooling jacket [41,42]. Details of parameters are described in Table 2. It is worth noting that, input disturbances can bring out system dynamics, such that measurements are temporally correlated, non-Gaussian-distributed and noisy due to the process non-linearity. The sampling interval is 1 min, 2000 points under normal conditions are collected as training data.[2]

[2] We use the Simulink model developed by Pilario [41], available at http://www.mathworks.com/matlabcentral/fileexchange/66189-feedback-controlled-cstr-process-for-fault-simulation.

Table 2
Parameter description in CSTR process.

Parameter	Description	Value
Q	Inlet flow rate	100.0 L/min
V	Tank volume	150.0 L
V_c	Jacket volume	10.0 L
ΔH_r	Chemical reaction heat	-2.0×10^5 cal/mol
UA	Heat transfer coefficient	7.0×10^5 cal/min/K
k_0	Pre-exponential factor to k	7.2×10^{10} min^{-1}
E/R	Activation energy and gas constant	1.0×10^4 K
ρ, ρ_c	Fluid density	1000 g/L
C_p, C_{pc}	Fluid heat capacity	1.0 cal/g/K
v_1, v_2, v_3	Gaussian noise	$N(0, 0.01)$

Table 3
Incipient fault scenarios in CSTR process.

Fault ID	Description	Value of f	Type
f_1	$Q = Q_0 + \delta$	5	Additive bias
f_2	$a = a_0 exp(-\delta t)$	0.0005	Multiplicative
f_3	$b = b_0 exp(-\delta t)$	0.001	Multiplicative
f_4	f_1, f_2 simultaneous		Multiplicative
f_5	$C_i = C_{i,0} + \delta t$	0.001	Additive
f_6	$T_c = T_{c,0} + \delta t$	0.05	Additive
f_7	$C = C_0 + \delta t$	0.001	Additive
f_8	$T = T_0 + \delta t$	0.05	Additive
f_9	$T_i = T_{i,0} + \delta t$	0.05	Additive
f_{10}	$Q_c = Q_{c,0} + \delta t$	−0.1	Additive

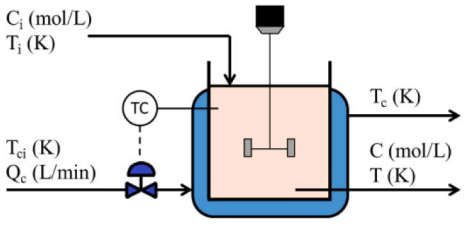


Fig. 6. Schematic diagram of the CSTR process.

We generated 100 faulty datasets for each fault scenario, as detailed in Table 3, differing in the random seeds for process noise, measurement noise, and input disturbances. Each testing set comprises 1600 samples, with faults imposed starting from the 201st sample. We averaged performance metrics for monitoring across all trials, and the comprehensive results, including FDR, FAR, and FDD, are presented in Table 4. It is clear that, although most methods achieve a FAR of 0, their FDD are not exceptionally early, suggesting there is room for refining strategies to enhance incipient fault detection. Incorporating latent projection, represented as CS_z and CS_e, has significantly boosted our proposed method by minimizing its detection delay. Early detection of small process deviations, prior to serious failure of the overall process, is crucial to preventing undesired consequences. Notably, our proposed method consistently performs well across all datasets, and even stands out compared to other methods, highlighting the efficacy of CS divergence in differentiating faulty from normal states. One of the challenges in diagnosing incipient faults lies in detecting tiny faults with power levels close to or even lower than the surrounding noise. To intensely evaluate the detection performance of CS divergence in such scenarios, we conducted an assessment for f_8 under different noise conditions of signals v_1, v_2, and v_3. Band-limited white noise was used in this assessment, and the resulting detection outcomes are graphically presented in Fig. 7. As evident from the figure, with the reduction of noise, the FDRs of our method increase accordingly, demonstrating effective fault detection across different noise power levels. Except for the noise level of 1e-1, where the detection rate is relatively low due to the overwhelming noise obscuring the fault, but the overall stable detection rate convincingly illustrates the robustness of the proposed algorithm in mitigating the adverse effects of noise. In particular, CS_e after PCA projection (shown in blue) exhibits the best noise robustness, maintaining high detection accuracy even under challenging noise conditions.

Subsequently, we assess the efficacy of the proposed isolation method, the isolation results for the mentioned fault scenarios in Table 3 were calculated and box-plotted in Fig. 8. For each fault scenario, the calculation highlights a specific variable with a higher/wider interval or various outliers than the others; this corresponds to the actual faulty one that was set. Specifically, f_4 to f_{10} were designed as unvariable faults due to additive settings, and the identified variables exhibited a higher/wider interval. Conversely, f_2, f_3, and f_4 were multiplicative faults characterized by changes in coefficients, these faults manifested as outliers across a broader range of variables as the system evolved. Take f_1 for example, Q is subjected to a sensor bias fault of 5 L/min, which aligns with the 201st sample. Given that the bias 5 L is notably small compared to its initial value of 330.9 L, f_1 qualifies

Table 4
The fault detection performance (%) in the CSTR process.

No.	KL		SMD			WD		CS		
	K_z	K_e		MD_z	MD_e	W_z	W_e		CS_z	CS_e
f_1	91.50	97.69	98.05	97.93	96.07	96.02	97.64	97.31	95.17	**98.19**
	0	33.33	8.82	**0**	1.63	**0**	**0**	**0**	**0**	**0**
	71	33	27	28	55	51	34	38	67	**26**
f_2	69.92	78.36	76.86	75.21	75.36	71.36	76.64	69.93	69.50	**85.71**
	0	**0**	8.82	23.53	**0**	**0**	**0**	**0**	**0**	**0**
	15	25	**10**	133	346	382	235	419	33	**33**
f_3	61.35	**97.0**	94.28	87.28	96.35	69.35	89.57	85.78	70.71	93.93
	0	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
	485	**43**	57	86	52	427	147	200	360	**19**
f_4	67.85	95.43	90.07	90.21	**95.93**	73.71	87.57	81.21	76.64	90.96
	0	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
	437	65	110	137	**58**	344	175	237	326	129
f_5	93.29	91.29	94.85	95.43	85.93	94.91	93.29	94.64	**94.93**	93.29
	0	**0**	**0**	0.98	**0**	**0**	**0**	**0**	**0**	**0**
	95	106	73	65	198	72	95	76	**72**	**8**
f_6	95.93	88.36	96.29	**97.29**	89.50	97.28	89.14	96.50	96.57	89.71
	0	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
	58	107	53	**39**	146	39	151	50	49	145
f_7	86.29	40.57	85.93	83.07	29.79	89.43	50.86	**92.36**	91.71	61.50
	0	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
	193	**16**	64	153	566	124	558	**82**	83	44
f_8	80.0	**97.78**	94.78	93.21	98.86	88.43	92.79	93.07	89.21	94.71
	0	**0**	**0**	**0**	6.86	**0**	**0**	**0**	**0**	**0**
	185	**32**	73	92	17	163	102	**11**	147	75
f_9	84.43	98.5	3.64	93.64	**98.86**	89.57	93.86	92.64	89.29	95.29
	0	**0**	**0**	**0**	**0**	20.59	**0**	**0**	**0**	**0**
	213	22	419	86	**17**	146	87	104	148	67
f_{10}	79.29	**95.5**	90.42	93.36	95.00	87.79	94.36	93.57	88.50	94.36
	0	**0**	0.98	6.86	36.27	**0**	**0**	**0**	**0**	**0**
	215	**64**	135	93	71	165	78	91	**45**	78
Ave.FDR	80.99	88.05	82.52	**90.67**	86.17	85.79	86.57	89.72	86.22	**89.77**
Ave.FAR	**0**	3.33	2.16	3.14	4.48	2.06	**0**	**0**	**0**	**0**
Ave.FDD	197	51	102	91	152	191	166	130	133	**62**

In each fault scenario, 100 trials were conducted, with variations introduced through different random seeds for process noise, measurement noise, and input disturbances. $w = 100$, $\eta = 1\%$.

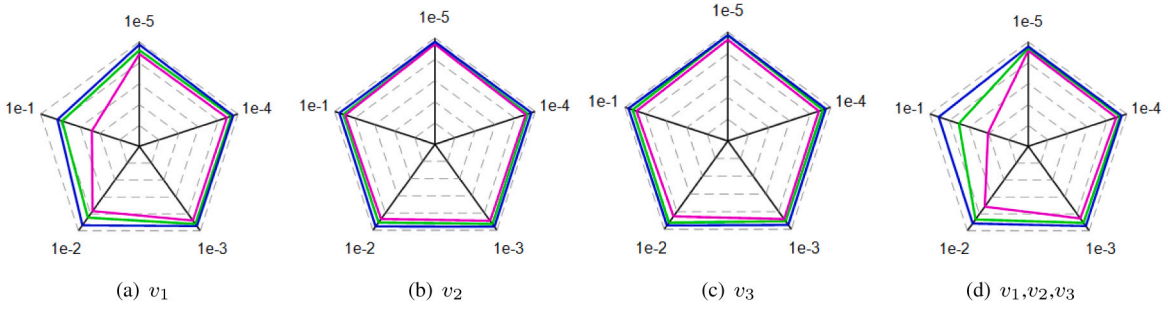


Fig. 7. Detection performance of the proposed CS divergence-based fault detection method for f_8 along with different noise power in the CSTR process. (a) v_1. (b) v_2.(c) v_3. (d) v_1, v_2, v_3. The radar coordinates represent the noise power, while the green color represents the FDRs of CS, red stands for the FDRs of CS_z, and blue indicates the FDRs of CS_e. The closer to the edge, the higher the FDR is. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

as an incipient fault. Fig. 9 clearly indicates that the dimensions of Q, C_i, T_i are the most affected by the fault. This observation is consistent with the dependence between Q and $\{C_i, T_i\}$ as detailed in Eqs. (19) and (20). The box-plot to the right validates this, showcasing a significant number of outliers for these variables, which are predominantly of larger values, thus causing a rightward skew in the distribution. To provide a more detailed explanation, we specifically illustrate the monitoring result based on conditional CS divergence for f_1 in Fig. 10. From Fig. 10(a), the alarm points are segmented into 8 phased stages. The dependencies largely remain consistent during the time intervals $[200, 303]$ (ascending) and $[302, 524]$ (descending), which aligns with Fig. 10(b). In the

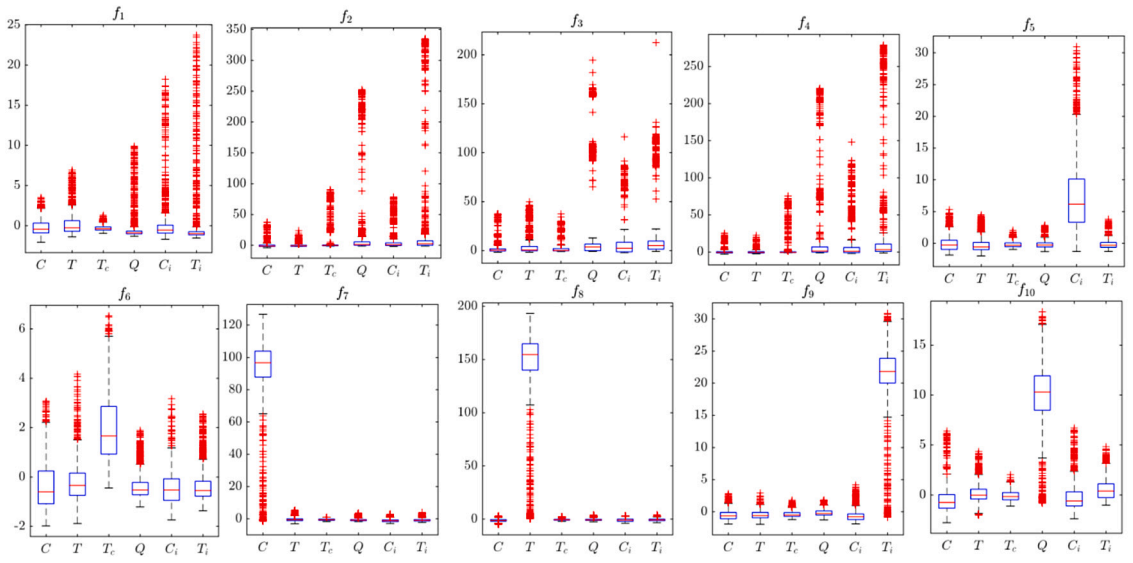


Fig. 8. Fault isolation results for the mentioned fault scenarios in the CSTR process.

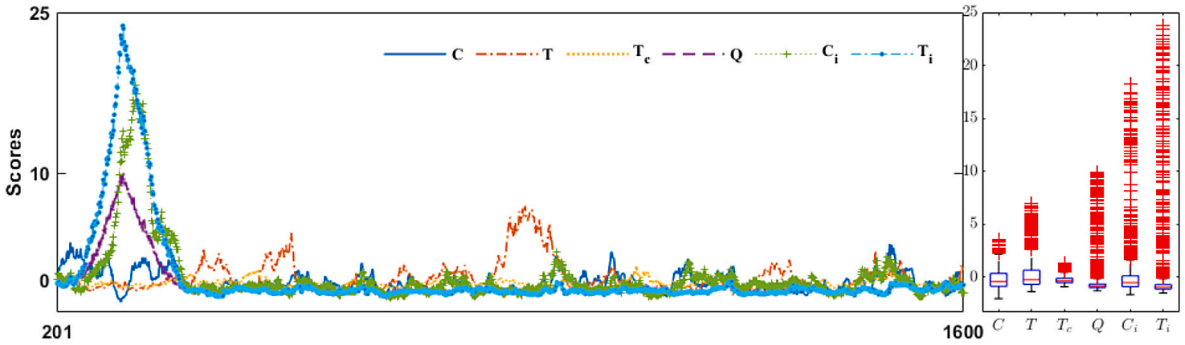


Fig. 9. Fault isolation results of f_1 in the CSTR process.

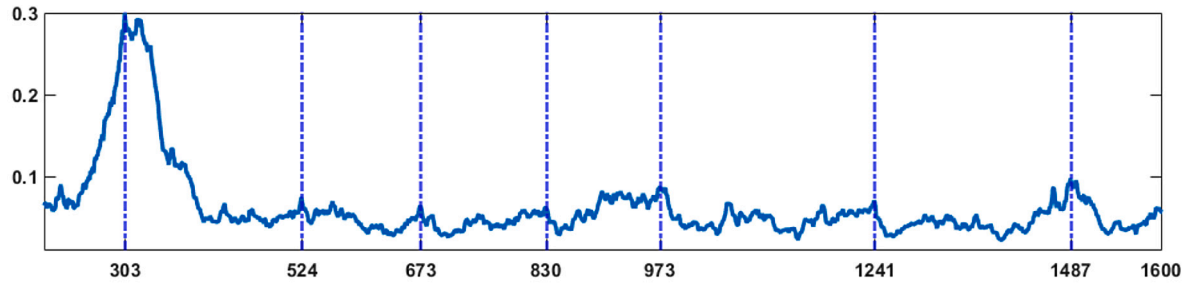
Table 5
Incipient fault scenarios in the cascade-controlled CSTR process.

Fault ID	Description	Primary fault variables	Secondary fault variables
cf_1	Cooling water heat transfer and scaling	T_c, F_c	T, C
cf_2	Increase in feed flow rate	F_i, F	C, F_c, T, T_c

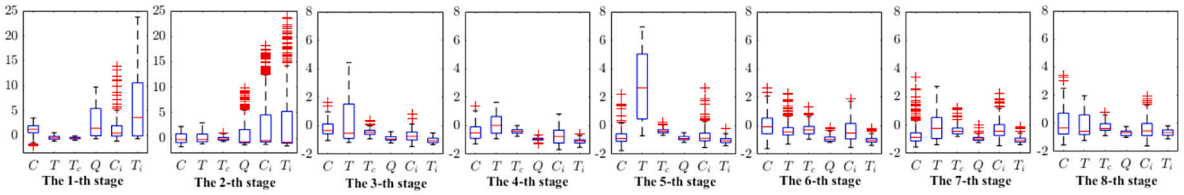
third stage, the 2-th dimension, which is the controlled object T, has a significant impact due to shifts in its dependence structure with Q. Consequently, our method enhances the explanatory of dynamic monitoring and isolation.

Additionally, in order to test the algorithm's compatibility with actual data even in the presence of multiplicative faults, we further extended its application to a 4-state CSTR process dynamic simulation under 2 cascade control loops. Two multiplicative faults, as shown in Table 5, are simulated[3]. Each testing set comprises 1000 samples, with faults imposed starting from the 201st sample. Additive white noise is present in all output measurements. The isolation results, summarized in Table 6, are almost entirely consistent with the actual situation. This demonstrates the ability of the proposed method to accurately identify fault variables without causing false alarms or omissions. Specifically, we present the isolation and identification results for cf_2. cf_2 denotes an increase in feed flow F_i, due to the liquid level control loop, concurrently induces an increase in discharge flow F. This, in turn, leads to an increase in outlet concentration C and a rise in reaction heat, ultimately resulting in an increase of the cooling water flow rate F_c. Indeed, while the changes in F_i and F are significant, the variations in C and F_c are slowly and slightly above their normal values. Fig. 11 displays the monitoring results of cf_2 based on conditional CS divergence, F_i and F exhibit significant increases

[3] We use the Simulink model accessible at https://www.mathworks.com/matlabcentral/fileexchange/65091-cascade-controlled-cstr-for-fault-simulation.



(a) Conditional CS based change point detection of f_1.



(b) Distributions for each state split by change points.

Fig. 10. The monitoring results based on change point detection of f_1 in the CSTR process. (a) Conditional CS based change point detection of f_1. (b) Distributions for each state split by change points.

Table 6
The fault isolation results in the cascade-controlled CSTR process.

cf_1	s_1	s_2	s_3	s_4		
	T_c, F_c T_{ci}, T	T_c, F_c T, C	T_c, F_c C, T	T_c, F_c C, T		
cf_2	s_1	s_2	s_3	s_4	s_5	s_6
	F_i, F C, F_c, T_i	F_i, F C, F_c, C_i	F_i, F C, F_c, T	F_i, F F_c, T_c, C	F_i, F C, F_c, C_i	F_i, F F_c, C, T_{ci}

Under each fault, the first line represents the primary fault variables, the second line indicates the secondary fault variables.

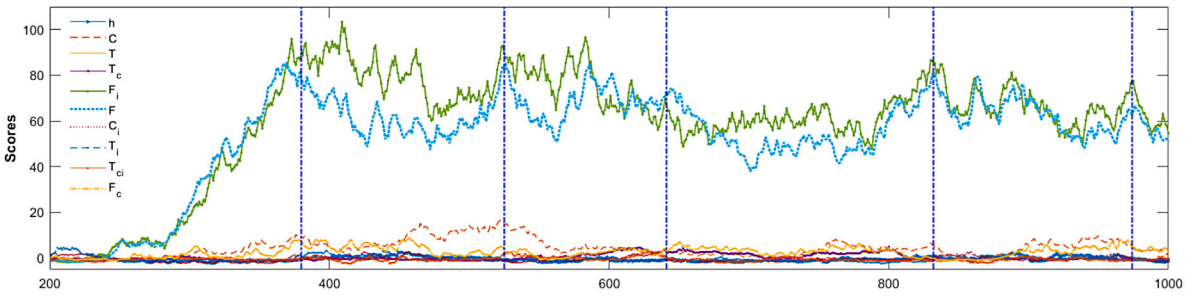
across all stages, highlighting their status as the primary fault variables. Meanwhile, Fig. 12 presents the variations of each variable in relation to their respective thresholds, which are calculated using the comparison algorithm, SPE contribution and IG methods[4]. In the SPE contribution method, the contributions of nearly all variables surpassed their thresholds, except T_i. Likewise, almost all variables, excluding C_i and T_i, exceeded their thresholds in the IG method. Both the SPE contribution and IG methods are suffering from smearing effect, the smearing out of fault variables' residuals over other non-fault process variables in the latent space [23]. Such issue may be particularly detrimental in the field of chemical engineering, where the operations are often characterized by highly interconnected equipment and numerous control-loops.

Our conditional CS divergence approach effectively isolates $\mathbf{x}^l$ from the overall distribution during the identification process. Through our experiments conducted on two distinct CSTR systems, we have demonstrated its robust performance across the given fault scenarios. By utilizing $\mathbf{x}^l|X^{-l}$ instead of $\mathbf{x}^l|X$, we are able to mitigate the influence of fault variables on other variables, ensuring accurate isolation (see Table 7).

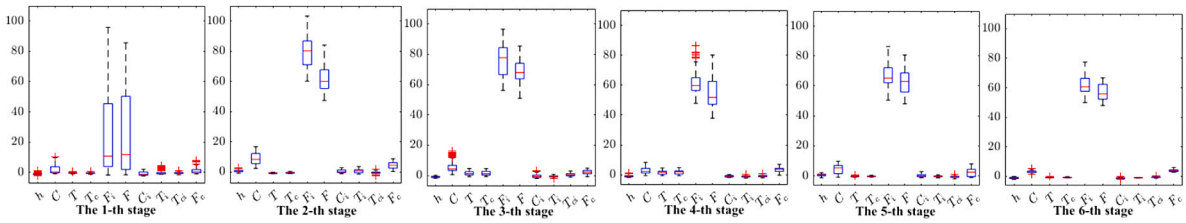
4.3. Application to CSTH process

In this part, we investigate a benchmark of a closed-loop continuous stirred-tank heater (CSTH) process [47]. The CSTH process represents a typical chemical mixing vessel and is equipped with three sensors, tank level L, steam valve F and outflow temperature T, refer to the schematic diagram in Fig. 13. As depicted by dashed lines in Fig. 13, two closed loops with three PI controllers are

[4] Note that, since traditional fault isolation methods can be employed in conjunction with DNNs for fault isolation, we use long short-term memory-autoencoder (LSTM-AE) as the baseline network model for comparison algorithms owing to its demonstrated effectiveness in detecting faults, as evidenced in previous studies [44–46]. The hidden dimension of the LSTM-AE used in our experiment is 32 and the bottleneck of latent dimension is 16. Interested readers are referred to our previous work [44] for more details.



(a) Conditional CS divergence based change-point detection.



(b) Distributions for each state split by change points.

Fig. 11. The monitoring results of cf_2 in the cascade-controlled CSTR process. (a) Conditional CS divergence based change point detection. (b) Distributions for each state split by change points. Variables F_i, F were correctly identified as primary fault variables.

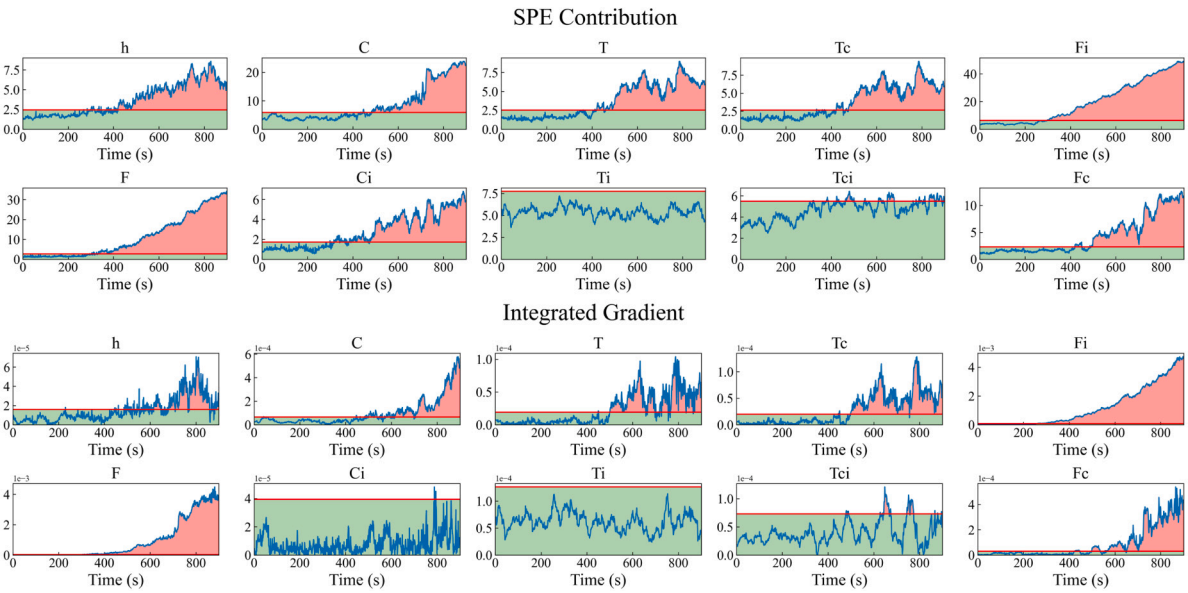


Fig. 12. The monitoring results of cf_2 by SPE contribution and IG methods in the cascade-controlled CSTR process. Variables F_i, F, C, F_c, T, T_c were correctly identified, while others are falsely recognized due to the smearing effect.

employed to regulate L and T. The system monitors five variables, namely $\mathbf{x} = [C_L, C_F, L, T, F]^T$, where C_L is the cold water flow and C_F is the cold water valve. It is worth mentioning that all the measurements are electrical signals within the range of 4-20 mA[5].

[5] The simulation code can be accessed at http://www.ps.ic.ac.uk/~nina/CSTHSimulation/index.htm.

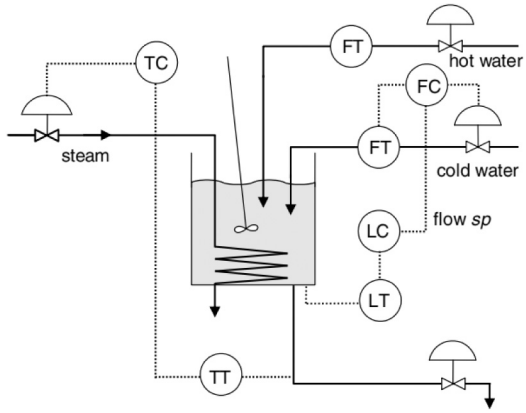


Fig. 13. Schematic diagram of the CSTH process.

Table 7
Incipient fault scenarios in the CSTH process.

No.	Description	Type
f_1	Hot water temperature	Slow increase, 50°C to 70°C
f_2	Cold water valve	Sticking

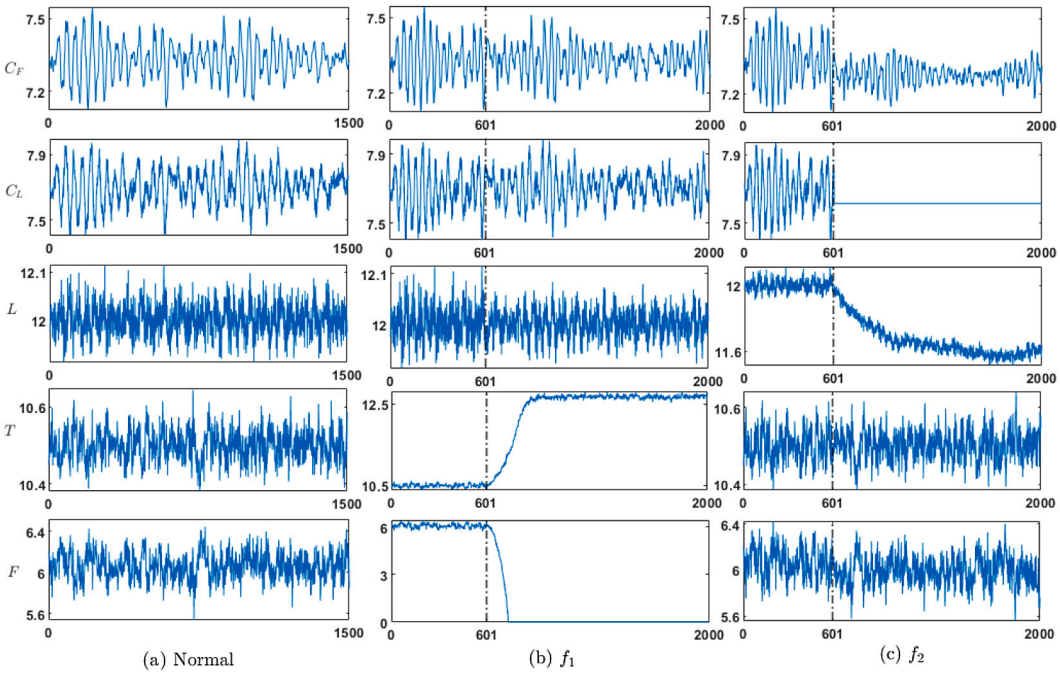
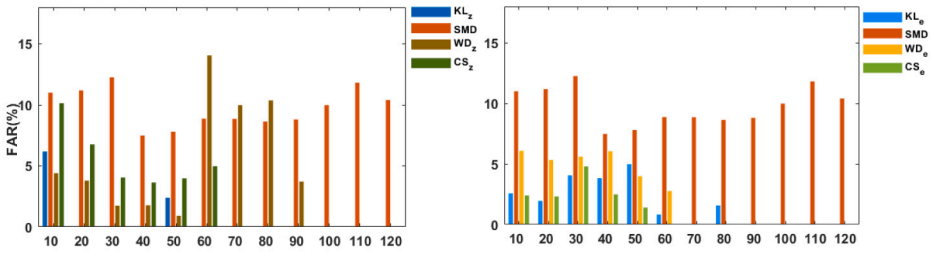


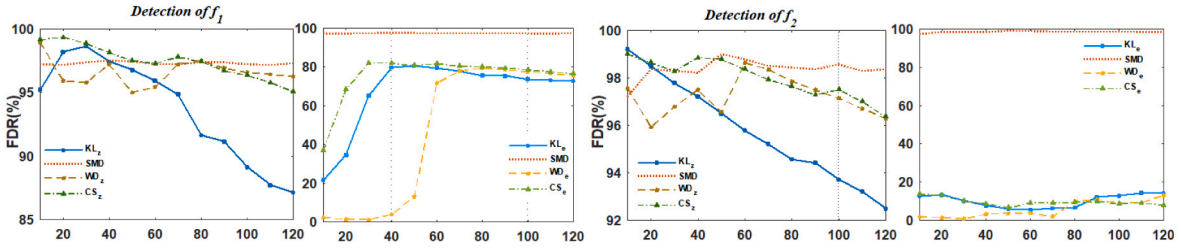
Fig. 14. The variables of CSTH (a) Normal, (b)f_1, (c) f_2.

Two scenarios of faults are investigated here, denoted as f_1 and f_2, as Table 7. f_1 stands for a slow increase in the hot water temperature, gradually rising from 50 °C to 70 °C over a period of 300 s. This gradual increase in temperature leads to a corresponding rise in the outlet temperature T, while the control loop causes the steam valve F to decrease slowly. f_2 represents a momentary sticking of the cold water valve, resulting in faulty changes in the level L and the cold water flow C_L. Both testing sets consist of 2000 samples, with the faults imposed starting from sample 601. Fig. 14 shows the variation of each variable over time under normal state and two fault conditions, each variable only fluctuates steadily in a small range around its own stationary state under normal conditions, as shown in Fig. 14(a).

We conducted an analysis on the effects of the window size w, considering the potential violation of the local stationary or smoothness assumption with a large window size and the unreliable distribution estimation with a small window size due to limited



(a) FAR(%) of f_1 and f_2.



(b) FDR(%) of f_1 and f_2.

Fig. 15. Detection results related to different window size w in the CSTH process. (a) FAR (%) of f_1 and f_2. (b) FDR (%) of f_1 and f_2.

Table 8
The fault detection results (%) in the CSTH process.

No.		KL		SMD			WD		VAE	OSAVA	CS		
		K_z	K_e		MD_z	MD_e	W_z	W_e				CS_z	CS_e
f_1	FDR	94.86	77.86	97.21	97.43	96.79	97.21	78.0	92.43	94.31	96.93	**97.79**	80.50
	FAR	0	0	8.85	20.9	3.01	9.98	**0**	**0**	**0**	**0**	**0**	**0**
	FDD	57	165	38	34	44	38	165	68	51	41	**30**	157
f_2	FDR	95.21	6.29	98.50	99.43	10.64	98.36	2.0	94.24	97.22	97.79	**97.93**	9.0
	FAR	0	0	8.85	20.9	3.01	9.98	**0**	**0**	**0**	**0**	**0**	**0**
	FDD	66	108	17	7	130	22	124	52	25	30	**28**	102

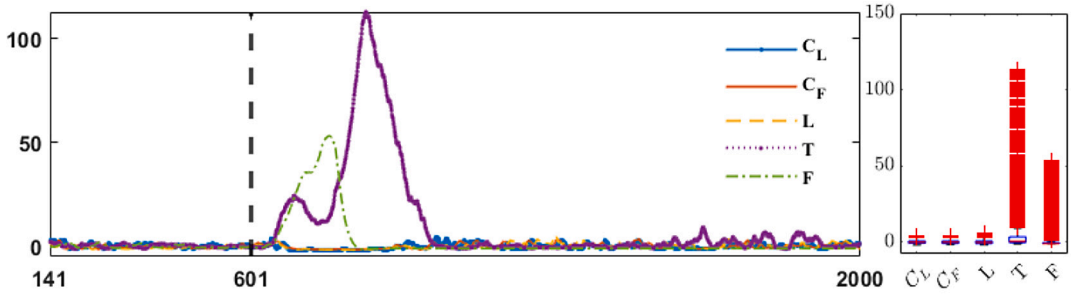
$w = 70$, $\eta = 5\%$.

samples. Fig. 15 presents the detection results by varying w from 50 to 150 in steps of ten. Note that the FARs of f_1 and f_2 are same because the first 600 normal samples in both testing sets are identical, with unique noise and disturbances. According to Fig. 15(b), as the window length increases, the FDRs of KL, WD and CS show a decreasing trend in the principal space while maintaining comparable performance in the residual space. On the other hand, their FARs gradually decrease towards 0 with the increase of w. To strike a good trade-off between FDR and FAR, $w = 70$ is optimal for CS and SMD-based detection, meanwhile $w = 50$ and $w = 100$ are more suitable for KL-based, WD-based detection, respectively.

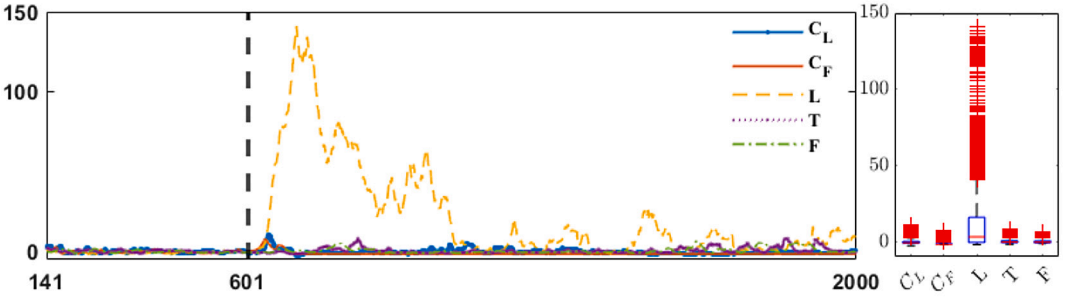
Table 8 provides a comparison of the detection performances. All methods displayed good performance, except SMD, which had higher FARs. In terms of FDR, CS divergence achieved consistent performance with KL, SMD and WD, with FDRs of 96.93% for f_1 and 97.79% for f_2. Additionally, PCA-CS, which integrates the CS divergence-based detection with PCA projections, achieved optimal detection performance, with the highest FDRs and the lowest FARs. The PCA projection also effectively reduced the detection delay, enabling the detection of f_1 in just 30 s and f_2 in 28 s. It is worth highlighting that the CS-based detection method outperformed the deep learning-based methods, such as the variational autoencoder (VAE) [48] and orthogonal self-attentive VAE [44]. This could be attributed to the fact that only 5 variables were used in this benchmark, which may be insufficient to fully demonstrate the superior representation ability of deep networks. Deep learning-based process monitoring methods generally perform well in high-dimensional data with multimodal variables.

In Fig. 16, we present the isolation results. From Fig. 16(a), T and F are related to fault f_1. The affected variable F exhibits almost identical trends to the root cause variable T. Variables gradually return to a steady state after fault occurred, indicating the creation of a new operating status. In Fig. 16(b), L is isolated as a faulty variable affected by f_2. Although f_2 is caused by the sticking of C_F, C_F itself was not identified as faulty because its valve position and flow remained within the normal range. Overall, our proposed method provides a clear isolation of the faulty variables in the CSTH process, making the results interpretable.

The monitoring results for f_1 and f_2 are depicted in Figs. 17 and 18, respectively. For instance, a large number of outliers appear in both T and F during the 1-th stage, while deviations and drift become evident in the 2-th stage in Fig. 17. Table 9 presents the corresponding isolation results. The first row indicates the variable most affected by the current state; the second row

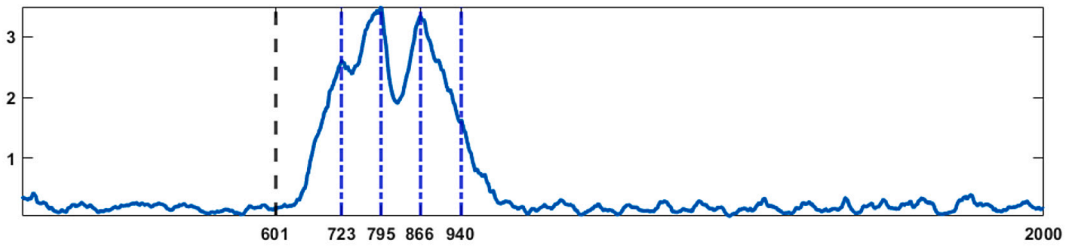


(a) Fault isolation results of f_1.

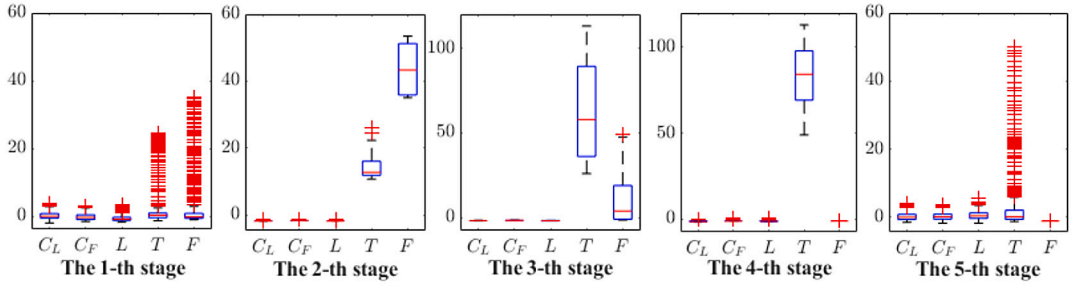


(b) Fault isolation results of f_2.

Fig. 16. Fault isolation results of f_1 and f_2 in the CSTH process. (a) Fault isolation results of f_1. (b) Fault isolation results of f_2.



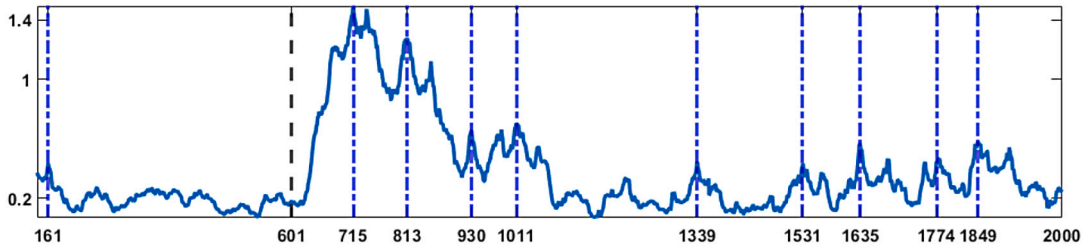
(a) Conditional CS divergence based change point detection of f_1.



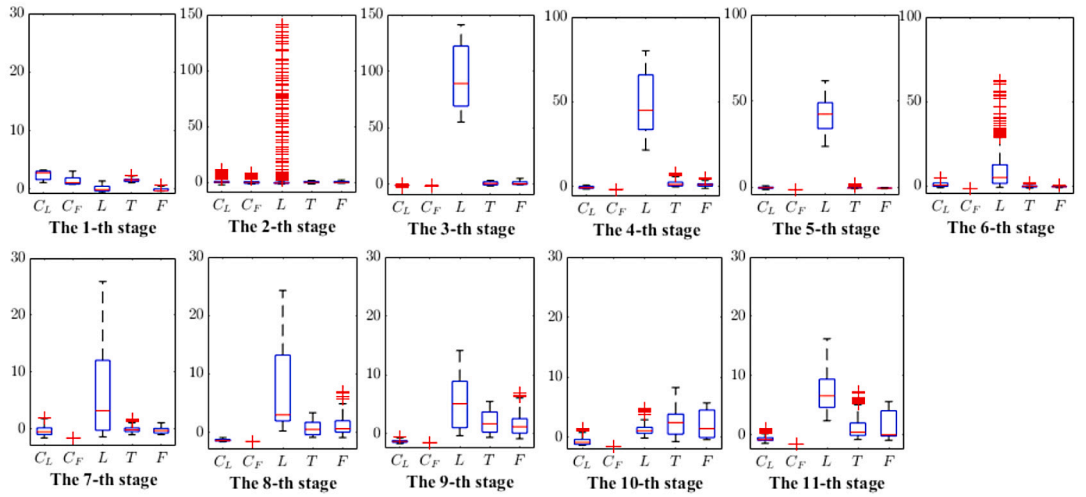
(b) Distributions for each state split by change points.

Fig. 17. The monitoring results based on change point detection of f_1 in the CSTH process. (a) Conditional CS divergence based change point detection of f_1. (b) Distributions for each state split by change points.

lists the secondary identified variables, and so forth. Table 9 offers valuable guidance for regulation. By integrating this with process knowledge, one can effectively adjust the pertinent variables.



(a) Conditional CS divergence based change-point detection of f_2.



(b) Distributions for each state split by change points.

Fig. 18. The monitoring results based on change point detection of f_2 in the CSTH process. (a) Conditional CS divergence based change point detection of f_2. (b) Distributions for each state split by change points.

Table 9
The dynamic variable isolation results for CSTH process.

f_1	s_1	s_2	s_3	s_4	s_5						
	F	F	T	T	T						
	T	T	F	C_F	C_F						
	C_F	C_F	C_F	C_L	C_L						
f_2	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}	s_{11}
	C_F	L	L	L	L	L	L	F	F	F	F
	C_L	C_F	F	F	C_L	C_L	F	L	T	T	L
	T	F	T	T	T	F	T	T	L	L	T

The first three variables identified as the principal components to the current state; Their importance decrease by row.

4.4. Application to an actual petrochemical production

In this section, we evaluate the proposed method using real-world data from a continuous catalytic reforming (CCR) process in a petrochemical enterprise in China. The CCR process involves the regeneration of catalysts. We focus on analyzing the top temperature of one reaction tower as an example. This temperature is influenced by various factors, including the feed amount, liquid level, and feed amount of its adjacent reaction tower, as well as the liquid level and return flow of its corresponding storage tank, among others. Please refer to Fig. 19 for a visual representation of these relationships. Due to the intricate physical or topological connections involving multiple monitoring variables, we used our previously proposed causality-gated time series Transformer (CGTST) model[6] select 7 most relevant variables for subsequent modeling.

[6] CGTST evaluates the causal strength of each variable's contribution to the prediction through its causal gating mechanism, making it well-suited for causal discovery within industrial big data. For a more comprehensive understanding, interested readers are encouraged to refer to our previous work [49].

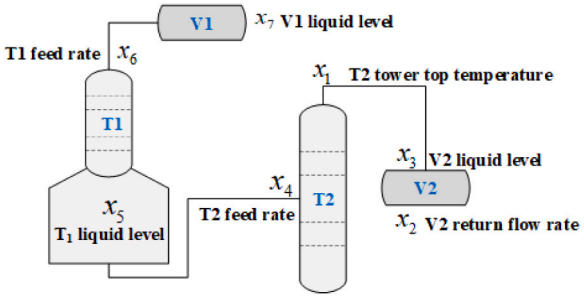


Fig. 19. Schematic diagram of a reaction tower in actual CCR process.

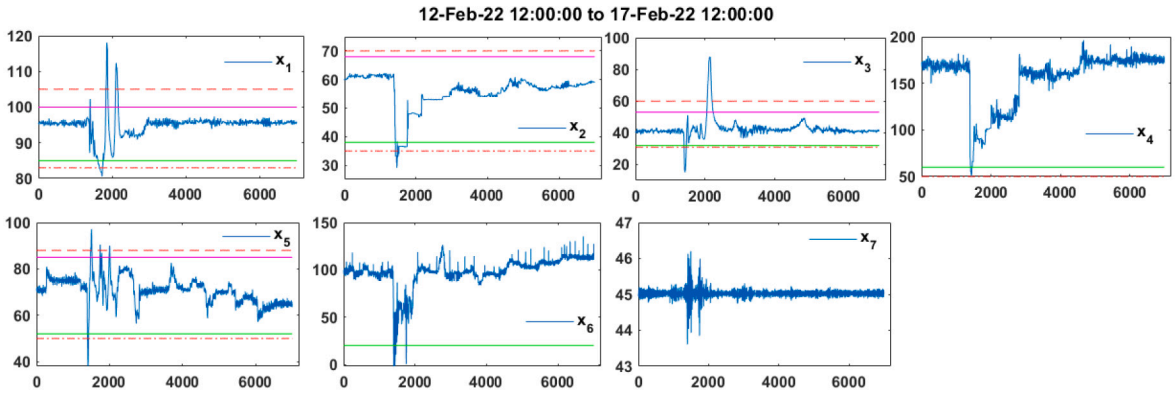
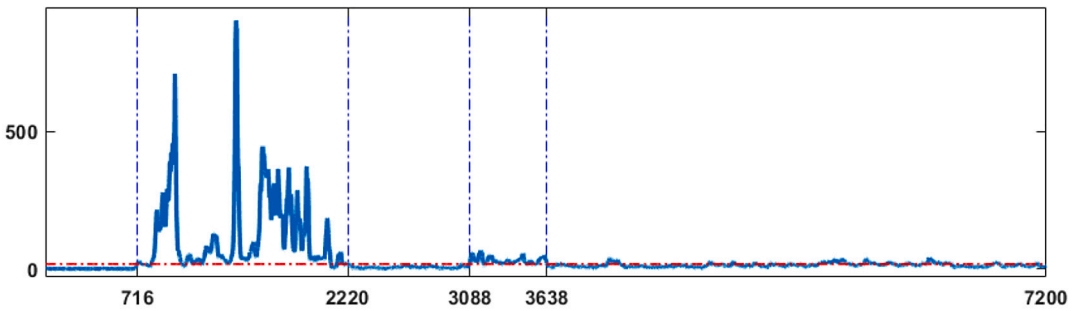
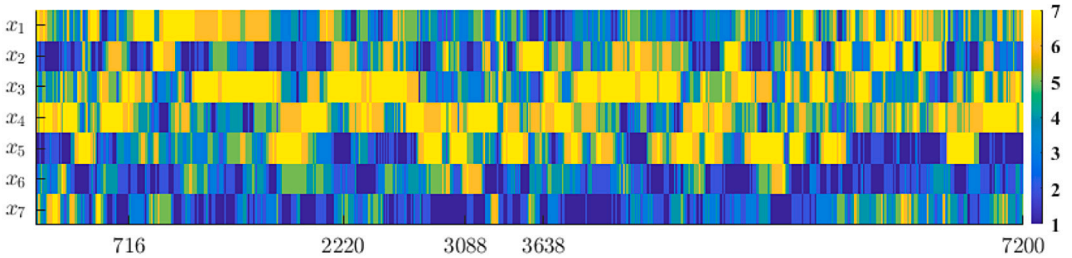


Fig. 20. The variables of the CCR process.

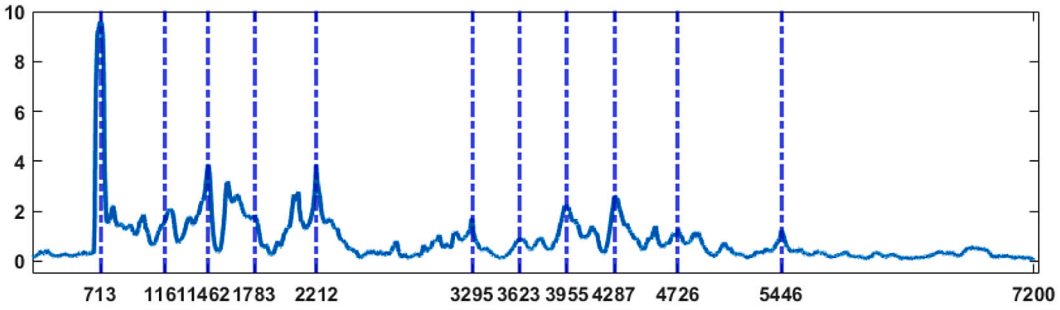


(a) The detection static for CCR process.

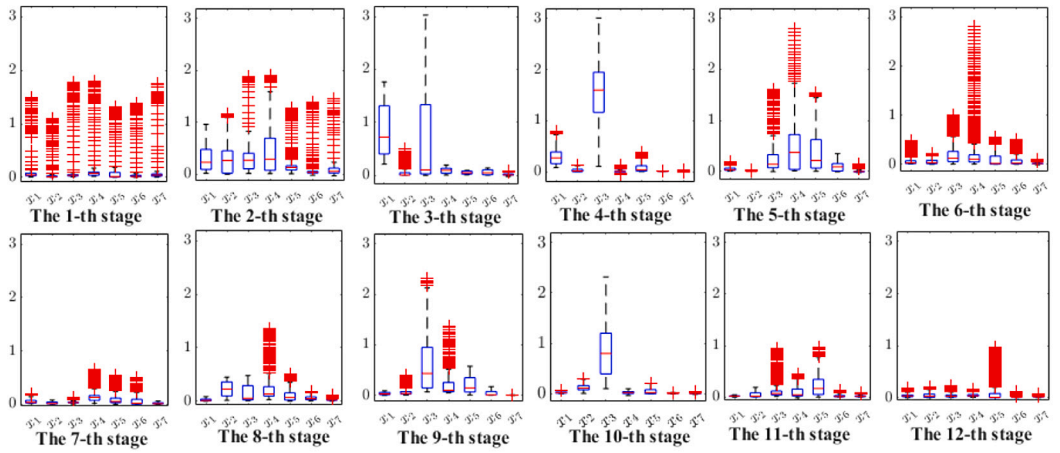


(b) The conditional CS divergence for fault isolation.

Fig. 21. Detection and isolation results for the CCR process. (a) The detection static for CCR process. (b) The conditional CS divergence for fault isolation.



(a) Conditional CS divergence based change point detection.



(b) Distributions for each stage split by change points.

Fig. 22. The monitoring results based on change point detection for the CCR process. (a) Conditional CS divergence based change point detection. (b) Distributions for each stage split by change points.

Fig. 20 displays the process measurements of a certain period in February 2022. Within the approximate range of $[675, 1515]$, the damage to a speed governor safety barrier in the recycle hydrogen compressor caused these variables to exhibit abnormal behavior after the 1515 time point, exceeding the distributed control system (DCS) control limits and triggering an alarm. Using the methodology presented in this paper, the detection results are shown in Fig. 21(a), faults are concentrated in the intervals of $[716, 2220]$ and $[3088, 3638]$. These intervals are longer than the actual DCS alarm interval of $[1515, 2220]$ due to the reliance of the DCS on univariate alarms according to its thresholds, which may not be sufficient to detect potential changes or evolving patterns that could indicate the onset of abnormalities or performance degradation. Variables in the intervals of $[716, 1516]$ and $[3088, 3638]$ may exhibit abnormal trends without triggering their DCS alarm thresholds. Indeed, the interval [716, 1516] represents the fault-tolerant range for the damage to the speed governor safety barrier, and the fault within the interval [3088, 3638] is caused by the addition of an extra process pipeline in the aromatics process. Here, the FDD, 41 mins, is acceptable compared to the sliding window length of 30 mins. The isolation results in Fig. 21(b) show that variables x_1 and x_3 have higher brightness in the interval of $[716, 2220]$, indicating their association with the fault. This aligns with the actual operation situation.

The monitoring results are presented in Fig. 22(a), with changes alarmed at time points [713, 1161, 1462, 1783, 2212, 3295, 3623, 3955, 4287, 4726, 5446]. The retrospective interval w_c is chosen as 120 mins, falling within the actual reliability range. Notably, the alarm delay of 38 mins here is shorter than the previously mentioned FDD (41 mins), highlighting the necessity and effectiveness of change point monitoring. In Fig. 22(b), a significant number of outliers are observed in the 1-th stage, but no obvious changes in their mean values, forming an upward abrupt step change; the mean values of x_1, x_2, x_3, x_4 increased in the 2-th stage, indicating the importance of further monitoring or regulation of these variables, yielding improved physical interpretation. This enhanced online process monitoring strategy facilitates informed decision-making at each stage of the process, enabling operators to make timely and accurate adjustments that ultimately lead to better process control and optimization. It is important to note that the feasibility validation of our algorithm was carried out in the raw observation space, without performing any data preprocessing or assumption fitting. This underscores its capability to handle uncertainties inherent in real-world industrial process data.

5. Conclusions

Given that divergence-based detection illustrates the dissimilarity between the monitored process and its reference normal state, we have introduced an innovative fault detection and isolation scheme grounded in a probabilistic perspective, offering enhanced sensitivity. It leverages monitoring statistics to jointly detect the occurrence of incipient faults using Cauchy–Schwarz divergence and isolate the specific compromised or faulty sensors through conditional Cauchy–Schwarz divergence. Our monitoring statistics, on the basis of Cauchy–Schwarz divergence, boast the advantage as they do not rely on any prior parametric assumptions about the underlying data distribution or historical fault data. Experiments conducted on both synthetic datasets and benchmark datasets from the continuous stirred-tank reactor and heater processes have demonstrated the compelling performance of our method. We also developed a change point detection-based fault diagnosis approach that utilizes conditional Cauchy–Schwarz divergence. It allows operators to adjust sensors during different phases, demarcated by alarm points, to preemptively address potential faults. The approach was validated on a real-world dataset from the continuous catalytic reforming process, confirming its applicability and effectiveness.

Our work offers a holistic solution for fault detection and isolation, it exhibits a higher detection rate, a lower false alarm rate, and earlier fault detection. Importantly, the conditional probabilistic analysis provides explanatory insights of which sensors are affected by fault propagation, making it a valuable monitoring tool for industrial engineering. With this established groundwork, researchers can further concentrate and expand their explorations into multivariate fault isolation, fault estimation, as well as traceability within fault diagnosis. We look forward to providing more valuable insights and practical solutions for enhancing the reliability and safety of industrial processes.

CRediT authorship contribution statement

Feiya Lv: Formal analysis, Funding acquisition, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Shujian Yu:** Investigation, Methodology, Software, Supervision, Writing – review & editing. **Huawei Ye:** Data curation, Project administration. **Jinsong Zhao:** Funding acquisition, Investigation, Methodology, Project administration, Supervision, Writing – review & editing. **Chenglin Wen:** Funding acquisition, Investigation, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant 62003004, 62125307 and U213207.

References

[1] S. Yin, S.X. Ding, X. Xie, et al., A review on basic data-driven approaches for industrial process monitoring, IEEE Trans. Ind. Electron. 61 (11) (2014) 6418–6428.
[2] L.H. Chiang, E.L. Russell, R.D. Braatz, Fault detection and diagnosis in industrial systems, Springer Science & Business Media, 2000.
[3] R. Isermann, Model-based fault-detection and diagnosis-status and applications, Annu. Rev. Control 29 (1) (2005) 71–85, http://dx.doi.org/10.1016/j.arcontrol.2004.12.002.
[4] J. Shang, M. Chen, H. Ji, D. Zhou, Recursive transformed component statistical analysis for incipient fault detection, Automatica 80 (2017) 313–327, http://dx.doi.org/10.1016/j.automatica.2017.02.028.
[5] R.R.A. Harinarayan, S.M. Shalinie, XFDDC: eXplainable fault detection diagnosis and correction framework for chemical process systems, Process. Saf. Environ. Prot. 165 (2022) 463–474, http://dx.doi.org/10.1016/j.psep.2022.07.019.
[6] J. Zeng, U. Kruger, J. Geluk, X. Wang, L. Xie, Detecting abnormal situations using the Kullback–Leibler divergence, Automatica 50 (11) (2014) 2777–2786, http://dx.doi.org/10.1016/j.automatica.2014.09.005.
[7] H. Safaeipour, M. Forouzanfar, A. Casavola, A survey and classification of incipient fault diagnosis approaches, J. Process Control 97 (2021) 1–16, http://dx.doi.org/10.1016/j.jprocont.2020.11.005.
[8] Y. Li, W. Cao, W. Hu, Y. Xiong, M. Wu, Incipient fault detection for geological drilling processes using multivariate generalized Gaussian distributions and Kullback–Leibler divergence, Control Eng. Pract. 117 (2021) 104937, http://dx.doi.org/10.1016/j.conengprac.2021.104937.
[9] X. Zhang, C. Delpha, D. Diallo, Performance of Jensen Shannon divergence in incipient fault detection and estimation, IEEE Int. Conf. Acoust. Speech and Signal Process. (ICASSP). (2019) 2742–2746, http://dx.doi.org/10.1109/ICASSP.2019.8682206.
[10] X. Zhang, C. Delpha, D. Diallo, Incipient fault detection and estimation based on Jensen–Shannon divergence in a data-driven approach, Signal Process. 169 (2020) 107410, http://dx.doi.org/10.1016/j.sigpro.2019.107410.
[11] C. Lu, J. Zeng, S. Luo, J. Cai, Detection and isolation of incipiently developing fault using Wasserstein distance, Processes 10 (6) (2022) 1081, http://dx.doi.org/10.3390/pr10061081.
[12] Y. Li, Y. Song, L. Jia, S. Gao, Q. Li, M. Qiu, Intelligent fault diagnosis by fusing domain adversarial training and maximum mean discrepancy via ensemble learning, IEEE Trans. Ind. Inform. 17 (4) (2021) 2833–2841, http://dx.doi.org/10.1109/TII.2020.3008010.
[13] H. Ji, Statistics mahalanobis distance for incipient sensor fault detection and diagnosis, Chem. Eng. Sci. 230 (2021) 116233, http://dx.doi.org/10.1016/j.ces.2020.116233.
[14] N. Sadeghzadeh-Nokhodberiz, N. Meskin, Protocol-based particle filtering and divergence estimation, IEEE Syst. J. 15 (3) (2021) 4537–4544, http://dx.doi.org/10.1109/JSYST.2020.3002907.
[15] J.C. Principe, D. Xu, J. Fisher, S. Haykin, Information theoretic learning, Unsupervised Adapt. Filter. 1 (2000) 265–319.

[16] S. Yu, H. Li, S. Løkse, R. Jenssen, J.C. Príncipe, The conditional Cauchy-Schwarz divergence with applications to time-series data and sequential decision making, (ISSN: 1568-4946) 2023, pp. 179–192, http://dx.doi.org/10.48550/arXiv.2301.08970, ArXiv preprint.
[17] R. Jenssen, Kernel entropy component analysis, IEEE Trans. Pattern Anal. Mach. Intell. 32 (5) (2010) 847–860, http://dx.doi.org/10.1109/TPAMI.2009.100.
[18] Y. Qi, Y. Wang, C. Lu, L. Wang, Improved batch process monitoring and diagnosis based on multiphase KECA, IFAC-Papers OnLine 51 (18) (2018) 827–832, http://dx.doi.org/10.1016/j.ifacol.2018.09.255.
[19] C. Hu, J. Luo, X. Kong, X. Feng, Novel fault subspace extraction methods for the reconstruction-based fault diagnosis, J. Process Control 105 (2021) 129–140.
[20] L. Feng, C. Zhao, Fault description based attribute transfer for zero-sample industrial fault diagnosis, IEEE Trans. Ind. Inform. 17 (3) (2021) 1852–1862, http://dx.doi.org/10.1109/TII.2020.2988208.
[21] A. Abid, M.T. Khan, J. Iqbal, A review on fault detection and diagnosis techniques: basics and beyond, Artif. Intell. Rev. 54 (2021) 3639–3664, http://dx.doi.org/10.1007/s10462-020-09934-2.
[22] S. Duan, K. Zhu, P. Song, C. Zhao, Dynamic causal modeling for nonstationary industrial process performance degradation analysis and fault prognosis, J. Process Control 129 (2023) 103050, http://dx.doi.org/10.1016/j.jprocont.2023.103050.
[23] J.A. Westerhuis, S.P. Gurden, A.K. Smilde, Generalized contribution plots in multivariate statistical process monitoring, Chemometr. Intell. Lab. Syst. 51 (1) (2000) 95–114, http://dx.doi.org/10.1016/S0169-7439(00)00062-9.
[24] L.M. Elshenawy, T.A. Mahmoud, C. Chakour, Simultaneous fault detection and diagnosis using adaptive principal component analysis and multivariate contribution analysis, Ind. Eng. Chem. Res. 59 (2020) 20798–20815, http://dx.doi.org/10.1021/acs.iecr.0c04572.
[25] B. Xiao, Y. Li, B. Sun, C. Yang, K. Huang, H. Zhu, Decentralized PCA modeling based on relevance and redundancy variable selection and its application to large-scale dynamic process monitoring, Process. Saf. Environ. Prot. 151 (2021) 85–100, http://dx.doi.org/10.1016/j.psep.2021.04.043.
[26] C.F. Alcala, S.J. Qin, Reconstruction-based contribution for process monitoring, Automatica 45 (7) (2009) 1593–1600, http://dx.doi.org/10.1016/j.automatica.2009.02.027.
[27] Q. Liu, B. Song, X. Ding, S.J. Qin, Fault diagnosis of dynamic processes with reconstruction and magnitude profile estimation for an industrial application, Control Eng. Pract. 121 (2022) 105008, http://dx.doi.org/10.1016/j.conengprac.2021.105008.
[28] C. Zhao, F. Gao, A sparse dissimilarity analysis algorithm for incipient fault isolation with no priori fault information, Control Eng. Pract. 65 (2017) 70–82, http://dx.doi.org/10.1016/j.conengprac.2017.05.005.
[29] J. Yang, C. Delpha, An incipient fault diagnosis methodology using local Mahalanobis distance: Fault isolation and fault severity estimation, Signal Process. 200 (2022) 108657, http://dx.doi.org/10.1016/j.sigpro.2022.108657.
[30] A. Basu, H. Shioya, C. Park, Statistical inference: the minimum distance approach, CRC Press, 2011.
[31] S. Kulinski, S. Bagchi, D.I. Inouye, Feature shift detection: Localizing which features have shifted via conditional distribution tests, Adv. Neural Inf. Process. Syst. 33 (2020) 19523–19533, http://dx.doi.org/10.48550/arXiv.2107.06929.
[32] J.C. Principe, D. Xu, Q. Zhao, J.W. Fisher, Learning from examples with information theoretic criteria, J. VLSI Signal Process. Syst. Signal, Image and Video Technol. 26 (1) (2000) 61–77.
[33] J. Robert, P.J. C, E. Deniz, E. Torbjørn, The Cauchy–Schwarz divergence and parzen windowing: Connections to graph theory and mercer kernels, J. Franklin Inst. 343 (6) (2006) 614–629, http://dx.doi.org/10.1016/j.jfranklin.2006.03.018.
[34] K. Kampa, E. Hasanbelliu, J.C. Principe, Closed-form Cauchy-Schwarz PDF divergence for mixture of Gaussians, Int. Jt. Conf. Neural Netw. (2011) 2578–2585.
[35] T. De Ryck, M. De Vos, A. Bertrand, Change point detection in time series data using autoencoders with a time-invariant representation, IEEE Trans. Signal Process. 69 (2021) 3513–3524.
[36] F. Lv, S. Yu, C. Wen, J.C. Principe, Interpretable fault detection using projections of mutual information matrix, J. Franklin Inst. 358 (7) (2021) 4028–4057, http://dx.doi.org/10.1016/j.jfranklin.2021.02.016.
[37] S. Yin, H. Gao, O. Kaynak, Data-driven control and process monitoring for industrial applications, Part I, IEEE Trans. Ind. Electron. 61 (11) (2014) 6356–6359, http://dx.doi.org/10.1109/TIE.2014.2312885.
[38] S.X. Ding, Model-Based Fault Diagnosis Techniques—Design Schemes, Algorithms and Tools, second ed., Springer Verlag, 2013.
[39] Z.W. Chen, K. Zhang, Y.A. Shardt, S.X. Ding, et al., Comparison of two basic statistics for fault detection and process monitoring, IFAC-Papers on Line 50 (1) (2017) 14776–14781.
[40] D. Zhou, Y. Y., Modern Fault Diagnosis and Fault Tolerant Control, Beijing, Tsing Hua University Publishing House, 2000.
[41] K.E.S. Pilario, Y. Cao, Canonical variate dissimilarity analysis for process incipient fault detection, IEEE Trans. Ind. Inform. 14 (12) (2018) 5308–5315.
[42] K.E.S. Pilario, Y. Cao, M. Shafiee, Mixed kernel canonical variate dissimilarity analysis for incipient fault monitoring in nonlinear dynamic processes, Comput. Chem. Eng. 123 (2019) 143–154.
[43] Q. Jiang, D.S. X., Y. Wang, X. Yan, Data-driven distributed local fault detection for large-scale processes based on the GA-regularized canonical correlation analysis, IEEE Trans. Ind. Electron. 64 (10) (2017) 8148–8157.
[44] X. Bi, J. Zhao, A novel orthogonal self-attentive variational autoencoder method for interpretable chemical process fault detection and identification, Process. Saf. Environ. Prot. 156 (2021) 581–597, http://dx.doi.org/10.1016/j.psep.2021.10.036.
[45] D. Wu, X. Bi, J. Zhao, ProTopormer: Toward understandable fault diagnosis combining process topology for chemical processes, Ind. Eng. Chem. Res. 62 (21) (2023) 8350–8361, http://dx.doi.org/10.1021/acs.iecr.3c00206.
[46] D. Li, L. Li, X. Li, Z. Ke, Q. Hu, Smoothed LSTM-AE: A spatio-temporal deep model for multiple time-series missing imputation, Neurocomputing 411 (2020) 351–363, http://dx.doi.org/10.1016/j.neucom.2020.05.033.
[47] N. Thornhill, S. Patwardhan, S. Shah, A continuous stirred tank heater simulation model with applications, J. Process Control 18 (2008) 347–360, http://dx.doi.org/10.1016/j.jprocont.2007.07.006.
[48] F. Cheng, Q.P. He, J. Zhao, A novel process monitoring approach based on variational recurrent autoencoder, Comput. Chem. Eng. 129 (2019) 106515, http://dx.doi.org/10.1016/j.compchemeng.2019.106515.
[49] X. Bi, D. Wu, D. Xie, H. Ye, J. Zhao, Large-scale chemical process causal discovery from big data with transformer-based deep learning, Process. Saf. Environ. Prot. 173 (2023) 163–177, http://dx.doi.org/10.1016/j.psep.2023.03.017.