Free Probability Theory Based Event Detection for Power Grids Using IoT-enabled Measurements

Hongxia Wang, Bo Wang, Jiaxin Zhang, Chengxi Liu, and Hengrui Ma

***Abstract*—Taking the advantage of Internet of Things (IoT) enabled measurements, this paper formulates the event detection problem as an information-plus-noise model, and detects events in power systems based on free probability theory (FPT). Using big data collected from phasor measurement units (PMUs), we construct the event detection matrix to reflect both spatial and temporal characteristics of power gird states. The event detection matrix is further described as an information matrix plus a noise matrix, and the essence of event detection is to extract event information from the event detection matrix. By associating the event detection problem with FPT, the empirical spectral distributions (ESDs) related moments of the sample covariance matrix of the information matrix is computed, to distinguish events from "noises", including normal fluctuations, background noises, and measurement errors. Based on central limit theory (CLT), the alarm threshold is computed using measurements collected in normal states. Additionally, with the aid of sliding window, this paper builds an event detection architecture to reflect power grid state and detect events online. Case studies with simulated data from Anhui, China, and real PMU data from Guangdong, China, verify the effectiveness of the proposed method. Compared with other data-driven methods, the proposed method is more sensitive and has better adaptability to the normal fluctuations, background noises, and measurement errors in real PMU cases. In addition, it does not require large number of training samples as needed in the training-testing paradigm.**

***Index Terms*—Big data, event detection, empirical spectral distribution (ESD), free probability theory (FPT), information-plus-noise model, Internet of Things (IoT), phasor measurement unit (PMU).**

I. Introduction

EVENT detection is critical to the reliable operation of power grids. However, with the increasing scale and the rapidly growing integrations of new energy and electrical vehicles [1], [2], the topology and operating mechanism of modern power grids become more and more complex [3], which makes the event detection in modern power grids more difficult.

Nowadays, with the fast development of Internet of Things (IoT) [4], [5], thousands of IoT devices such as phasor measurement units (PMUs) [6], supervisory control and data acquisition (SCADA) [7], and smart meters [3], are equipped in power systems, making it easy to reflect the power grid states based on big data. Events in power systems, whether they are caused by attacks or failures, will be reflected in the relevant data changes. Therefore, it is of great significance to implement data-driven methods in IoT-enabled measurements, to conduct online state monitoring and event detection in power systems [8], [9], and to diminish or even avoid the severe consequences [10].

One typical data-driven method for event detection is the training-testing paradigm, which discovers abnormalities through two main steps: training the model using a training set and validating it using a testing set. For example, support vector machine (SVM) [11], [12], random forest [13], [14], and deep learning [15], [16] are used to classify the faults in power grids. In overall, the training-testing methods get fault mechanism by studying mapping functions between the input measurements and the event type. They are of high precision and straightforward. However, these methods require large amount of data representing events, which are difficult to access because power grids operate normally most of the time. Besides, the performance of these methods heavily relies on the training set, and the poor generalization may further limit their applications in real cases [17], [18].

Another typical method is the statistical method, which identifies anomalies through statistical characteristics of data. Wherein, in addition to those rather simple indicators like mean, maximum, and minimum values [19], some other indicators including wavelet-based method [20] and Chebyshev-based method [21] are developed to extract event features. Nevertheless, the above methods in essence analyze the data from different measurements independently, which have limited abilities to reveal the correlations between data from different devices. To deal with this problem, the principal component analysis (PCA) based methods in [22] and [23] construct the event detection matrix using measurements from various devices, and learn the warning threshold based on the measurements reflecting normal state. However, these methods assume the applied data to be linear and Gaussian

Manuscript received: April 2, 2023; revised: June 21, 2023; accepted: September 6, 2023. Date of CrossCheck: September 6, 2023. Date of online publication: November 22, 2023.

This work was supported by the National Key Research and Development Program of China (No. 2021YFB2401302).

H. Wang, B. Wang (corresponding author), J. Zhang, C. Liu, and H. Ma are with Hubei Engineering and Technology Research Center for AC/DC Intelligent Distribution Network, School of Electrical Engineering and Automation, Wuhan University, wuhan, China, and they are also with and School of Electrical Engineering and Automation, Wuhan University, Wuhan, China (e-mail: 2018282070092@whu. edu. cn; whwdwb@whu. edu. cn; 2020282070157@whu. edu.cn; liuchengxi@whu.edu.cn; henrym@whu.edu.cn).

DOI: 10.35833/MPCE.2023.000205

distributed, which is not always the case in real power systems [23], [24]. In addition, the extraction of main components will inevitably result in the loss of original information, especially when the dimension of the event detection matrix is large, which is exactly the case within IoT, because the growing integration of IoT-enabled devices will increase the data volume [4].

Instead of focusing on a few main components that the largest eigenvalues corresponded to, the spectral distribution approaches detect events by analyzing the spectral distribution of the big data matrix, where all the eigenvalues are considered and the loss of information can be avoided to some extent. For example, based on random matrix theory (RMT), [25]-[27] catch the fault signals by comparing the empirical spectral distributions (ESDs) with the standard limit spectral distributions (LSDs), where the mean spectral radius (MSR) from the complex plane is used as the event detection indicator, and the difference between ESDs and LSDs is used to represent the fault degree. However, these methods describe the measurements as independently identically distributed (IID.), and the alarm threshold (e.g., the inner circle of the circular law in RMT) is also based on the IID. assumption, which is too strict to meet in real power grids. In addition, the lack of consideration for normal fluctuations, background noises, etc. makes the MSR-based methods insensitive especially when the signal is weak.

To address the above problems, this paper takes full advantage of IoT-enabled measurements, and proposes a spectral distribution analysis method to detect events online as well as catch event signals from the multivariate statistical view without dimension reduction. Firstly, using measurements from the PMU, which is one of the IoT-enabled devices, the event detection matrix reflecting spatio-temporal characteristics is constructed and formulated by an information-plus-noise model. Wherein, the event signals are regarded as information matrix. Normal fluctuations, background noises, measurement errors and so on are regarded as noise matrix. The sum of the two is the measurement matrix. The essence of event detection is to extract event signal from the measurement matrix. Secondly, based on the spectral distribution relations among information matrix, noise matrix, and measurement matrix in information-plus-noise model provided by free probability theory (FPT), the ESD-related moments of the information matrix are taken to reflect the event signals. With the aid of FPT, the strict Gaussian and IID. conditions can be relaxed. Thirdly, using measurements collected under normal states, this paper reveals the ESD information representing normal fluctuations, background noises, measurement errors and so on. With the calculation rules given by FPT, the event is detected based on hypothesis testing, where the central limit theorem (CLT) is used to determine the alarm threshold. In addition, the sliding window is used to conduct online event detection. The proposed method can determine whether there is an event or not. To further determine the event type, it is necessary to further analyze the event using other methods.

The contributions of this paper are summarized as follows.

1) Based on the event detection matrix reflecting spatio-temporal characteristics of the IoT-enabled measurements, the event detection problem is formulated as an information-plus-noise model. Wherein, the event signal is taken as information. Normal fluctuations, background noises, and measurement errors are regarded as "noises". The essence of event detection is to reveal event signal from the measurements including the "noises" mentioned above. The consideration of "noises" makes the event detection sensitive to weak signals.

2) By leveraging FPT and sliding window, the ESD-based moments of the event information matrix are calculated to tell events from the event detection matrix online. The proposed method converts the IoT-caused multivariate problem to a univariate problem, and extracts statistical signals without dimension reduction. What is more, the spectral distribution method relaxes the strict conditions in the RMT-based method and is more adaptive to real IoT-enabled measurements.

3) The event detection is conducted based on hypothesis testing. The ESD-based moments of the noise matrix are calculated using normal state measurements, and the alarm threshold is calculated based on the CLT.

The rest of this paper is organized as follows. Section II discusses the basic principles of FPT and its application in information-plus-noise model. Section III formulates the event detection problem in power systems by information-plus-noise model and proposes the event detection architecture based on FPT. Section IV verifies the FPT-based method based on both simulated and real PMU data. Section V concludes this paper.

II. FPT and Its Application in Information-plus-noise Model

FPT is a tool for high-dimensional matrix analysis, which is mainly used in multiple-input multiple-output (MIMO) channel capacity estimation [28]. In the realm of IoT, the data from various devices and diverse locations result in high-dimensional big data within power systems. Consequently, it is feasible to extract event signals using FPT for such data configurations. This section introduces the concepts of ESD and FPT and presents the applications of FPT in information-plus-noise model.

A. ESD of Matrices

As shown in (1), for a measurement matrix $\boldsymbol{M}$ with p measurement variables and n sampling points, the ESD of its sample covariance matrix $\boldsymbol{A}$ reflects the distribution characteristics of the eigenvalues of $\boldsymbol{A}$ [27], [29].

$$\mu_{\boldsymbol{A}}(x)=\frac{1}{p}\sum_{i=1}^{p}I\left(\lambda_i(\boldsymbol{A})\le x\right) \tag{1}$$

where $\mu_{\boldsymbol{A}}(x)$ is the ESD of $\boldsymbol{A}$; x is the independent variable of the ESD function; $I(\cdot)$ is the indicator function; and $\lambda_i(\boldsymbol{A})$ is the i^{th} eigenvalue of $\boldsymbol{A}$.

In addition, the sample covariance matrix of the measurement matrix $\boldsymbol{M}$ is computed by (2).

$$\boldsymbol{A}=\frac{1}{n}\tilde{\boldsymbol{M}}\tilde{\boldsymbol{M}}^{\mathrm{H}} \tag{2}$$

where H represents the transpose of a matrix; and $\tilde{\boldsymbol{M}}$ is the normalization of $\boldsymbol{M}$, which is computed by:

$$\tilde{\boldsymbol{M}}_i=\left(\boldsymbol{M}_i-\mu\left(\boldsymbol{M}_i\right)\right)/\sigma\left(\boldsymbol{M}_i\right) \tag{3}$$

where $\boldsymbol{M}_i$ is the i^{th} row of $\boldsymbol{M}$; $\widetilde{(\cdot)}$ is the symbol of normalization of $\boldsymbol{M}_i$; and $\mu\left(\boldsymbol{M}_i\right)$ and $\sigma\left(\boldsymbol{M}_i\right)$ are the mean and variance of $\boldsymbol{M}_i$, respectively.

B. Basic Rules of FPT

In traditional mathematical theory, it is difficult to get the ESD of the sum or multiplication of two matrices from the ESDs of two individual matrices (except for the case when they are commutative) and vice versa. However, the FPT introduces the concept of "asymptotic freeness" [30]-[32], which is adaptive to non-commutative random variables like matrices. Based on FPT, one can get the ESD of the sum or multiplication of the two matrices when they are asymptotic free.

Given two matrices $\boldsymbol{C}_1\in\mathbb{C}^{p\times p}$ and $\boldsymbol{C}_2\in\mathbb{C}^{p\times p}$, the ESDs of them are $\mu_{\boldsymbol{C}_1}$ and $\mu_{\boldsymbol{C}_2}$, respectively, and the ESD of $\boldsymbol{C}_1+\boldsymbol{C}_2$ is $\mu_{\boldsymbol{C}_1+\boldsymbol{C}_2}$. If both $\boldsymbol{C}_1$ and $\boldsymbol{C}_2$ are asymptotic free, we have:

$$\mu_{\boldsymbol{C}_1+\boldsymbol{C}_2}=\mu_{\boldsymbol{C}_1}\boxplus\mu_{\boldsymbol{C}_2} \tag{4}$$

where $\boxplus$ represents the additive free convolution. Equation (4) indicates that the ESD of the sum of two matrices equals to additive free convolution of the ESDs of the two matrices.

Let the ESD of multiplication of $\boldsymbol{C}_1$ and $\boldsymbol{C}_2$ be $\mu_{\boldsymbol{C}_1\times\boldsymbol{C}_2}$, then $\mu_{\boldsymbol{C}_1\times\boldsymbol{C}_2}$ can be computed based on $\mu_{\boldsymbol{C}_1}$ and $\mu_{\boldsymbol{C}_2}$.

$$\mu_{\boldsymbol{C}_1\times\boldsymbol{C}_2}=\mu_{\boldsymbol{C}_1}\boxtimes\mu_{\boldsymbol{C}_2} \tag{5}$$

where $\boxtimes$ represents the multiplicative free convolution. Equation (5) indicates that the ESD of multiplication of two asymptotic free matrices equals to the multiplicative free convolution of the ESDs of two respective matrices.

In addition, based on additive free deconvolution $\boxminus$ and multiplicative free deconvolution $\oslash$, we have (6) and (7), respectively.

$$\mu_{\boldsymbol{C}_1}=\mu_{\boldsymbol{C}_1+\boldsymbol{C}_2}\boxminus\mu_{\boldsymbol{C}_2} \tag{6}$$

$$\mu_{\boldsymbol{C}_1}=\mu_{\boldsymbol{C}_1\times\boldsymbol{C}_2}\oslash\mu_{\boldsymbol{C}_2} \tag{7}$$

C. Application of FPT in Information-plus-noise Model

As shown in (8), the standard information-plus-noise model consists of three parts: ① the measurement matrix $\boldsymbol{Y}$; ② the information matrix $\boldsymbol{X}$; and ③ the noise matrix $\boldsymbol{S}$.

$$\boldsymbol{Y}=\boldsymbol{X}+\boldsymbol{S} \tag{8}$$

The sample covariance matrix of each part in (8) is denoted as $\boldsymbol{W}_n=\frac{1}{n}\tilde{\boldsymbol{Y}}\tilde{\boldsymbol{Y}}^{\mathrm{H}}$, $\boldsymbol{R}_n=\frac{1}{n}\tilde{\boldsymbol{X}}\tilde{\boldsymbol{X}}^{\mathrm{H}}$, and $\boldsymbol{\Gamma}_n=\frac{1}{n}\tilde{\boldsymbol{S}}\tilde{\boldsymbol{S}}^{\mathrm{H}}$, respectively; and the ESDs of $\boldsymbol{W}_n$, $\boldsymbol{R}_n$, and $\boldsymbol{\Gamma}_n$ are denoted as $\mu_{\boldsymbol{W}}$, $\mu_{\boldsymbol{R}}$, and $\mu_{\boldsymbol{\Gamma}}$, respectively.

Based on the FPT rules mentioned above, [31] reveals the ESD relation between the elements in the information-plus-noise model, as shown in:

$$\mu_{\boldsymbol{W}}\oslash\mu_c=\left(\mu_{\boldsymbol{R}}\oslash\mu_c\right)\boxplus\mu_{\boldsymbol{\Gamma}} \tag{9}$$

where μ_c is the ESD of Wishart matrix [27], [33]. The density of μ_c $\left(\mathrm{d}\mu_c/\mathrm{d}x\right)$ is described as Marchenko-Pastur law (M-P law), based on which μ_c can be determined once the dimensions of the matrices in the information-plus-noise model are determined.

Then, we have (10), from which one can get event information $\mu_{\boldsymbol{R}}$ from the measurement matrix and the noise matrix.

$$\mu_{\boldsymbol{R}}=\left(\left(\mu_{\boldsymbol{W}}\oslash\mu_c\right)\boxminus\mu_{\boldsymbol{\Gamma}}\right)\boxtimes\mu_c \tag{10}$$

D. Computation of ESD-based Moments of Information Matrix Based on FPT

The ESD of a matrix can be uniquely expressed by its moments, so one can compute the moments of $\mu_{\boldsymbol{R}}$ to get signal about the information matrix. As shown in (11), the "moment-cumulant formula" describes the relation between the moments and cumulants of an ESD μ, based on which the computation of the moments of $\mu_{\boldsymbol{R}}$ can be calculated.

$$m_k^{\mu}=\sum_{i\le k}\alpha_i^{\mu}\cdot coef_{k-i}\left(\left(1+m_1^{\mu}z+m_2^{\mu}z^2+\ldots\right)^i\right) \tag{11}$$

where $coef_i(\cdot)$ is the coefficient of z^i; m_k^{μ} is the k^{th} moment of μ; and α_i^{μ} is the i^{th} cumulant of μ.

Based on the iterative operation, one can compute the first k cumulants from the first k moments, and vice versa [31].

$$coef_{k-i}\left(\left(1+m_1^{\mu}z+m_2^{\mu}z^2+\ldots\right)^i\right)=M_i(k-i) \tag{12}$$

$\boldsymbol{M}_i(i=1,2,\ldots,k)$ is computed as:

$$\begin{cases}\boldsymbol{M}_1=\boldsymbol{m}\\ \boldsymbol{M}_2=\boldsymbol{m}\star\boldsymbol{m}\\ \quad\vdots\\ \boldsymbol{M}_k=\star_k\boldsymbol{m}\end{cases} \tag{13}$$

$$\boldsymbol{m}=\begin{bmatrix}1 & m_1^{\mu} & m_2^{\mu} & \ldots & m_k^{\mu}\end{bmatrix} \tag{14}$$

where $\star_k$ represents the k-fold convolution; and $\boldsymbol{m}$ is a vector composed by the first k moments of μ.

The expression of $\mu_{\boldsymbol{R}}$ in (10) includes a multiplicative free deconvolution between $\mu_{\boldsymbol{W}}$ and μ_c, an additive free deconvolution between $\mu_{\boldsymbol{W}}\oslash\mu_c$ and $\mu_{\boldsymbol{\Gamma}}$, and a multiplicative free convolution between $\left(\mu_{\boldsymbol{W}}\oslash\mu_c\right)\boxminus\mu_{\boldsymbol{\Gamma}}$ and μ_c. As depicted in (15) and (16), the cumulants of addictive free (de)convolution between $\mu_{\boldsymbol{A}}$ and $\mu_{\boldsymbol{B}}$ can be computed based on the addition or subtraction between cumulants as:

$$\alpha_k^{\mu_{\boldsymbol{A}}\boxplus\mu_{\boldsymbol{B}}}=\alpha_k^{\mu_{\boldsymbol{A}}}+\alpha_k^{\mu_{\boldsymbol{B}}} \tag{15}$$

$$\alpha_k^{\mu_{\boldsymbol{A}}\boxminus\mu_{\boldsymbol{B}}}=\alpha_k^{\mu_{\boldsymbol{A}}}-\alpha_k^{\mu_{\boldsymbol{B}}} \tag{16}$$

Also, the moments of multiplicative free (de)convolution of μ and μ_c can be computed based on their moments:

$$cm_k^{\mu}=\sum_{i\le k}cm_i^{\mu\oslash\mu_c}\cdot coef_{k-i}\left(\left(1+cm_1^{\mu}z+cm_2^{\mu}z^2+\ldots\right)^i\right) \tag{17}$$

$$cm_k^{\mu\boxtimes\mu_c}=\sum_{i\le k}cm_i^{\mu}\cdot coef_{k-i}\left(\left(1+cm_1^{\mu}z+cm_2^{\mu}z^2+\ldots\right)^i\right) \tag{18}$$

where c is the ratio of rows to columns of the matrices in the information-plus-noise model.

Based on (11) and (15)-(18), the moments of $\mu_{\boldsymbol{R}}$ can be computed from Algorithm 1. Algorithm 1 depicts the calculation details of moments of $\mu_{\boldsymbol{R}}$, where *moment*(·) denotes the calculation of moments through cumulant; and *cumulant*(·) denotes the computation of cumulants through moments. The higher-order moment provides a better fit to the ESD, but it also demands a more extensive computational effort. Considering the tradeoff between computation cost and accuracy, this paper takes the third moment of the ESD of $\mu_{\boldsymbol{R}}$ as the event detection indicator.

Algorithm 1: calculation of moments of $\mu_{\boldsymbol{R}}$

Input: the measurement matrix $\boldsymbol{M}$ and the first three moments of the noise matrix calculated from Algorithm 2 ($m_k^{\mu_r}$)

Step 1: compute moments of $v=\mu_{\boldsymbol{W}} \oslash \mu_c$ by (17)

1) $\{\gamma_k\}=cumulant\left(\left\{cm_k^{\mu_W}\right\}\right), k=1,2,3$

2) $\{m_k^v\}=\{\gamma_k/c\}, k=1,2,3$

Step 2: compute moments of $\rho=v \ominus \mu_{\boldsymbol{I}}$ by (11) and (16)

1) $\{\alpha_k^v\}=cumulant\left(\{m_k^v\}\right), k=1,2,3$

2) $\{\alpha_k^{\mu_I}\}=cumulant\left(\{m_k^{\mu_I}\}\right), k=1,2,3$

3) $\{\alpha_k^\rho\}=\{\alpha_k^v-\alpha_k^{\mu_I}\}, k=1,2,3$

4) $\{m_k^\rho\}=moment\left(\{\alpha_k^\rho\}\right), k=1,2,3$

Step 3: compute moments of $\rho \otimes \mu_c$ by (18)

1) $\{\alpha_k^{\mu_R}\}=moment\left(\left\{cm_k^\rho\right\}\right), k=1,2,3$

2) $\{m_k^{\mu_R}\}=\{\alpha_k^{\mu_R}/c\}, k=1,2,3$

Output: the first three moments of information matrix ($m_k^{\mu_R}$)

Algorithm 2: calculation of moments of noise matrix and alarm threshold

Step 1: compute moments of noise matrix $m_k^{\mu_r}$

1) For each sampling moment t, calculate the k^{th} moment of the noise matrix $m_{k_t}^{\mu_r}$ formed by the historical data collected under normal states, where $m_{k_t}^{\mu_r}$ is the k^{th} moment of $m_k^{\mu_r}$ calculated at moment t

2) Calculate the mean of the k^{th} moment of the noise matrix $m_k^{\mu_r}=mean(m_{k_t}^{\mu_r})$

Step 2: compute alarm threshold $m_3^{\mu_R^0}$

1) At each sampling moment t, calculate the moments of the event matrix $m_{k_t}^{\mu_R}$ using Algorithm 1, where the measurement matrix is constructed using the historical measurements from $t-n+1$ to t, and the moments of the noise matrix $m_k^{\mu_r}$ are from ***Step 1***. Repeat this step until all the moments are considered

2) Calculate the mean and standard deviation of the third moment of the event matrix, depicted as $mean(m_{3_t}^{\mu_R})$ and $dev(m_{3_t}^{\mu_R})$, respectively

3) Based on $\alpha=0.05$, the alarm threshold in standard Gaussian distribution is 1.65, so the alarm threshold is calculated according to $m_3^{\mu_R^0}=1.65\cdot dev(m_{3_t}^{\mu_R})+mean(m_{3_t}^{\mu_R})$

III. Event Detection Architecture Based on FPT

In this section, the event detection matrix is firstly constructed using the PMU measurements. Secondly, the event detection problem is formulated as an information-plus-noise model. Thirdly, the alarm threshold is calculated based on the CLT considering normal fluctuations, background noises, and measurement errors. Lastly, the online event detection is proposed by combining FPT with sliding window.

A. PMU Big Data and Event Detection Matrix

PMU [6], [34], [35] is one of the typical IoT-enabled devices. PMUs generate data streams with accurate time stamps and high resolutions. In this paper, by leveraging IoT-enabled measurements from PMUs, the event detection matrix $\boldsymbol{Y}$ with the size of $p\times n$ is constructed to reflect both spatial and temporal characteristics. As shown in (19), there are p PMUs at different locations (spatial characteristics), and n denotes the length of the time series (temporal characteristics). Besides, the elements in each column have the same time stamp.

$$\boldsymbol{Y}=\begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & & \vdots \\ y_{p1} & y_{p2} & \cdots & y_{pn} \end{bmatrix} \tag{19}$$

The ratio of rows to columns is c ($c=p/n$). The elements in $\boldsymbol{Y}$ can be nodal voltage, branch current, active power and so on [19], [26], [27], depending on the type of the event that needs to be detected. For example, the events associated with active power disturbances primarily lead to transient in frequency and voltage magnitude, thus this kind of events can be detected using event detection matrix constructed by active power, and frequency or voltage magnitude. However, those related to reactive power affect the voltage magnitude (locally) only [19]. In this situation, instead of using frequency and active power, the voltage magnitude should be taken as the elements of event detection matrix.

B. Information-plus-noise Model of Event Detection Problem

The event detection matrix constructed in (19) reflects power grid state and contains event signals [25]-[27], [29], so it is feasible to conduct event detection based on it. However, the matrix in (19) is actually a measurement matrix, which means that in addition to the event signals, normal fluctuations, background noises, and measurement errors are also included. These factors always exist and can be regarded as "noises", because they are mixed with event signals and can deteriorate the detection results or even totally cover the event signals, especially when the event signals are weak. However, the previous research work doesn't consider the "noises" and monitor the power grid state by directly extracting event signals from $\boldsymbol{Y}$ [25]-[27], [29], so it is necessary to reconsider the event detection based on PMU measurement data, and detect events after eliminating the effects of the "noises" mentioned above.

As shown in (8), the event detection matrix in (19) can be described as an information-plus-noise model, which consists of two parts.

1) The information matrix $\boldsymbol{X}$, which indicates whether an event occurs or not, and the severity of the event, etc. For instance, [27] abstracts the event as an step signal, and the magnitude of the step signal is used to describe the event severity.

2) The noise matrix $\boldsymbol{S}$, which contains the normal fluctuations, measurement errors, background noises, etc.

From the discussions above, we can conclude that the essence of event detection in power grids is to extract signals of $\boldsymbol{X}$ from the measurement matrix $\boldsymbol{Y}$, which is exactly what

this paper uses FPT to do. The logic is as follows. As discussed in Section III, ESD is a common tool to reflect properties of a specific matrix. Based on the ESD of the measurement matrix, studies have been done to reflect power grid states [25], [26], detect faults [27], [29], and so on. Thus, ESD can also be used to reflect event signals from $\boldsymbol{X}$. FPT reflects the ESD-based relation between the elements in information-plus-noise model, based on which ESD of the information matrix $\boldsymbol{X}$ can be extracted from the ESDs of measurement matrix and noise matrix using (10). Furthermore, the ESD-based moments can uniquely describe a specific matrix, and Algorithm 1 describes the computation of ESD-based moments of information matrix based on FPT, so this paper takes the ESD-based moments as the event indicators, and detects events using Algorithm 1.

C. Alarm Threshold Considering "Noises" in Power Systems

Algorithm 1 describes the method to get ESD-based moments of the information matrix from those of the measurement matrix and noise matrix, where the moments and cumulants of $\mu_{\boldsymbol{W}}$ and $\mu_{\boldsymbol{\Gamma}}$ are needed. The moments and cumulants of $\mu_{\boldsymbol{W}}$ are easy to calculate based on the PMU measurements. However, there is no direct way to get the moments and cumulants of $\mu_{\boldsymbol{\Gamma}}$, which represent the information of the noise matrix. As mentioned in Section III-B, the noise matrix includes normal fluctuations, measurement errors, background noises, and so on. The measurements collected in normal operation states contain all the information above. Thus, this paper uses the measurements in normal operation states to calculate moments and cumulants of $\mu_{\boldsymbol{\Gamma}}$.

The alarming threshold can be decided based on the moments of $\mu_{\boldsymbol{\Gamma}}$ using CLT, and events are detected based on hypothesis testing. When the power grid is in normal operation state, the measurements matrix will only contain the "noises" mentioned above, which means that the measurement matrix and the noise matrix have the same distributions. Because of the subtraction process mentioned in Algorithm 1, the moments $\mu_{\boldsymbol{R}}$ will be really small and follow the Gaussian distribution according to CLT. When there is an event, the information matrix will contain event signals, and the difference between $\mu_{\boldsymbol{W}}$ and $\mu_{\boldsymbol{\Gamma}}$ will make the moments of $\mu_{\boldsymbol{R}}$ relatively large and will follow the Gaussian distribution with very low probability. Consequently, this paper accomplishes event detection by assessing the probability of the moments of $\mu_{\boldsymbol{R}}$ adhering to the Gaussian distribution, using hypothesis testing theory, which is depicted in (20).

$$\begin{cases} \mathcal{H}_0: m_k^{\mu_R} < m_k^{\mu_R^0} \\ \mathcal{H}_1: m_k^{\mu_R} \geq m_k^{\mu_R^0} \end{cases} \tag{20}$$

where $m_k^{\mu_R^0}$ denotes the alarm threshold; $\mathcal{H}_0$ denotes the null hypothesis corresponding to the normal operation states, where $m_k^{\mu_R} < m_k^{\mu_R^0}$; and $\mathcal{H}_1$ corresponds to the situation where $m_k^{\mu_R} \geq m_k^{\mu_R^0}$ and null hypothesis is rejected, which means that there is an event. In hypothesis testing theory, the significance level α determines whether your sample evidence is strong enough to suggest that an effect exists in the entire population of the null hypothesis $\mathcal{H}_0$. In this paper, α is set to be 0.05, and from the normalized Gaussian distribution, we can see that if the normalized event indicator is more than 1.65, we have less than 5% probability that the power grid operates in normal operation state, so the null hypothesis is rejected and $\mathcal{H}_1$ is accepted, that is, there is an event. Details of calculations of moments of noise matrix ($m_k^{\mu_\Gamma}$) and alarm threshold ($m_3^{\mu_R^0}$) are shown in Algorithm 2. In this paper, considering the trade-off among sensitivity, robustness, and the computation, we have $k=1,2,3$, and the third moment of the information matrix is used to detect the event. Details of order of moment and cumulant selections are shown in Section IV-D.

D. Real-time Event Detection Architecture Based on Sliding Window

To conduct real-time event detection, the sliding window is used to observe the power grid state. As shown in (19), the length of the window is n. In each window, there are $n-1$ columns of historical samples and 1 column of current samples. As time goes on, we have 1 column of historical data moving out and 1 column of new data moving in at each sampling moment.

Based on sliding window, the overall framework of event detection based on FTP using IOT-enabled measurements is depicted in Fig. 1, which mainly includes three steps.

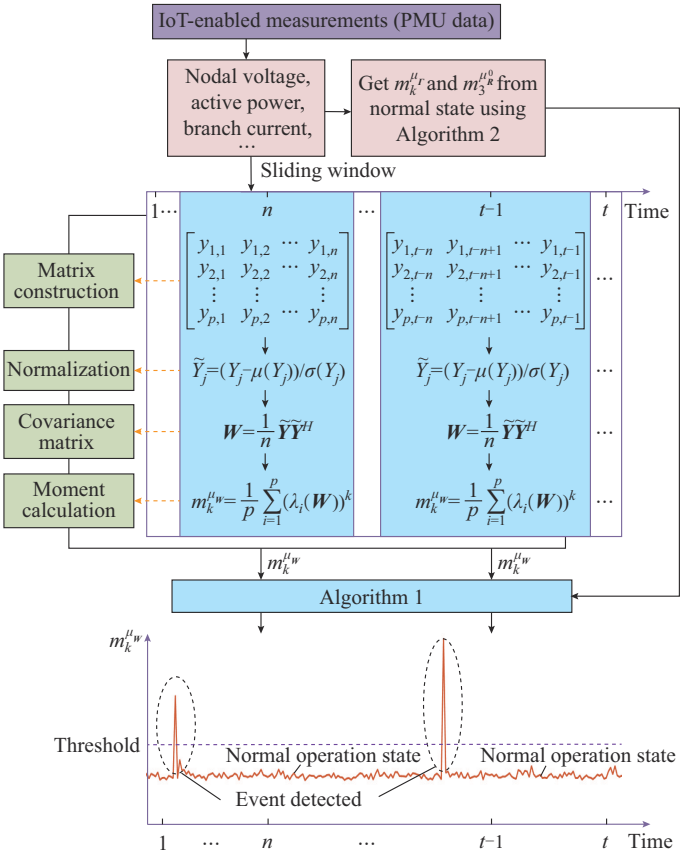

Fig. 1. Event detection based on FPT using IoT-enabled measurements.

Step 1: using the historical measurements, construct the

noise matrix $\boldsymbol{S}$, calculate the first three moments of $\boldsymbol{\Gamma}(m_k^{\mu_r})$, and calculate the alarm threshold ($m_3^{\mu_R^0}$) using Algorithm 2.

Step 2: using the PMU measurements collected at current time and the historical measurements, construct the event detection matrix $\boldsymbol{Y}$, and calculate the first three moments of $\boldsymbol{W}$ (depicted as $m_k^{\mu_W}$).

Step 3: taking $m_k^{\mu_W}$ and $m_k^{\mu_r}$ as inputs, compute $m_k^{\mu_R}$ using "moment-cumulant formula" in Algorithm 1, and monitor the power grid state based on the alarm threshold. When the measurements of PMUs are updated, repeat *Step 2* and *Step 3*.

IV. Case Studies

Based on MATLAB R2016b, the FPT-based event detection method is validated in three cases: ① its effectiveness towards simulated data; ② its effectiveness towards real PMU data; and ③ its advantages over other data-driven methods. At the end of this section, the way to decide orders of moments and cumulants of the event detection indicator is introduced.

A. Case 1: Case Study with Simulated Data

In this subsection, a 500 kV power grid covering 26 provinces is built based on power system analysis software package (PSASP), according to the real power grid topology of China. To focus on the operation state of Anhui Province, data streams from 40 PMUs are collected and analyzed. The 40 PMUs are deployed at 40 buses in Anhui, China. The sampling rate is 100 Hz, and the simulation duration is 5 s, so we have 500 sampling points in total. During the simulation, the load at bus 15 experiences an step increase of 50 MW at 1.01 s, and stays at 50 MW during the following 399 sampling points. Event settings for case 1 are depicted in Table I. In this case, both voltage magnitude and frequency at each bus are collected to construct the event detection matrix and verify the adaptability of the proposed method to different kinds of PMU data in power systems. In addition, the parameter c in (19) is set to be 0.8, so the width of the window is 50, and the indicators in Fig. 2(b) and Fig. 3(b) have their values from the 50th sampling moment. The test results based on voltage magnitude and frequency will be analyzed, respectively.

TABLE I
Event Settings for Case 1

Bus No.	Sampling time	Event
15	$t = 0$ - 1.01 s	Unchanged
	$t = 1.01$ s	$\Delta P = 0 \rightarrow 50$ MW (linearly increase)
	$t = 1.01$ - 5 s	The same as that when $t = 1.01$ s
Others	$t = 0$ - 5 s	Unchanged

1) Event Detection Based on Voltage Magnitude Measurements

Figure 2 depicts the event detection based on voltage magnitude measurements. Based on the voltage measurements collected under normal operation state, the first three moments of sample covariance matrix of noise matrix in (8) and the alarm threshold are calculated. Then, the FPT-based indicator is used to monitor the operation state during the whole simulation. From Fig. 2, the following two conclusions can be obtained.

1) The FPT-based indicator can catch the starting point of the event based on voltage magnitude measurements. The alarm threshold is 4.25. As shown in Fig. 2(b), the indicator hits 4.28 when $t = 1.02$ s, which is exactly the next sampling moment after the event. It indicates the ability of the indicator to sensitively catch the event signal. In addition, the indicator keeps alarming during $t = 1.02$ - 2.20 s, which corresponds to the drastic changing trend during that period in Fig. 2(a).The 1.18 s duration alarm will make the event detection result more credible.

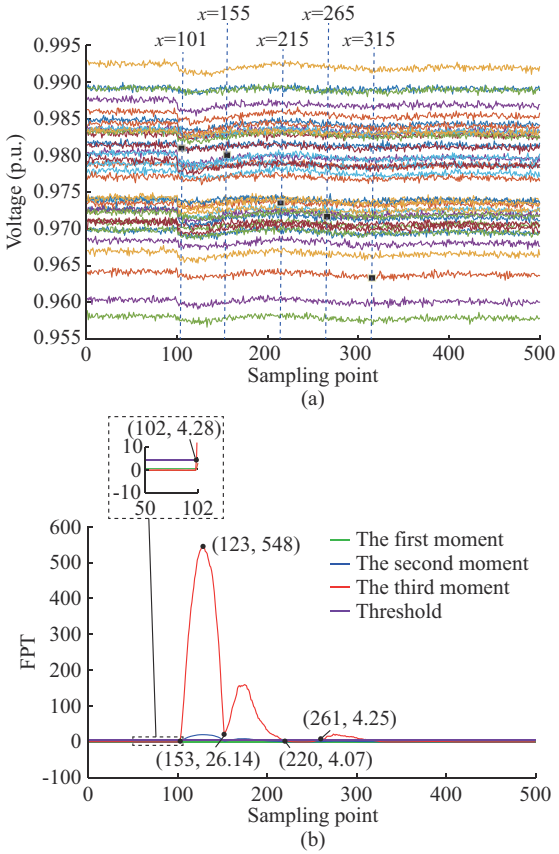

Fig. 2. Event detection based on voltage magnitude measurements. (a) Voltage magnitude. (b) Event detection results.

2) The FPT indicator based on voltage magnitude measurements can reflect power grid state in almost real time. During $t = 0$ - 1 s, the third moment stays steady at nearly 0 (as shown in the zoomed part in Fig. 2(b)), which correctly reflects the stable state of the power grid during that period. During $t = 1.02$ - 5.00 s, there are three peaks in total. The first peak reflects the voltage drop during $t = 1.02$ - 1.54 s, which is caused by the step increase of the load at $t = 1.01$ s. The second peak corresponds to the voltage recovery during $t = 1.55$ - 2.14 s. The third peak reflects the gentle decrease of the voltage magnitude during $t = 2.65$ - 3.14 s. What is more, the indicator shows a stable trend when $t = 2.15$ - 2.64 s,

which corresponds to the relatively steady trend of the voltage magnitude during that period. After $t=3.15$ s, the indicator returns to nearly 0, which correctly reflects the stable state of the voltage magnitude, because the voltage magnitude reverts to a state similar to the condition prior to the occurrence of the event.

2) Event Detection Based on Frequency Measurements

Figure 3 depicts the event detection results based on frequency measurements. Similar to the voltage magnitude measurements, the first three moments of the sample covariance matrix of the noise matrix and the alarm threshold are calculated based on the measurements collected under normal operation state. From Fig. 3, two conclusions can be obtained.

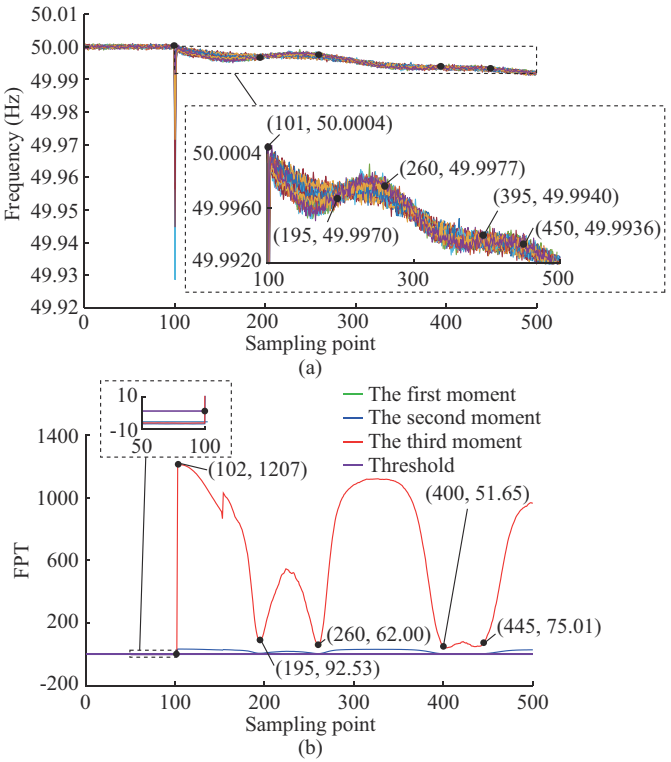

Fig. 3. Event detection based on frequency measurements. (a) Frequency. (b) Event detection results.

1) The FPT-based indictor can catch the event signal based on frequency measurements. The alarm threshold is 4.25 in this case. As shown in Fig. 3(b), the indicator hits 1208 when $t=1.02$ s, which is exactly the nearest sampling point after the event occurs. In addition, the indicator keeps alarming since $t=1.02$ s, which correctly reflects the event during $t=1.01$ - 5.00 s.

2) The FPT indicator based on frequency measurements can reflect power grid state in almost real time. During t=0 - 1 s, the indicator stays steady, which reflects the stable state of the frequency during that period. Then, during $t=$ 1.02 - 5.00 s, the third moment calculated fluctuates severely and keeps alarming, which correctly reflects the frequency decreasing trend during that period. To be specific, when $t=$ 1.02 s, this indicator surges to about 1200, which corresponds to the slump of the frequency at that moment. Besides, the two peaks during $t=1.95$ - 2.06 s and $t=2.06$ - 4.00 s, respectively, reflect the increasing and decreasing trends. When $t=4.00$ - 4.45 s, the third moment of ESD shows a relatively steady trend and stays nearly 50, which corresponds to the relatively stable trend of the frequency during that period. After $t=4.45$ s, the increasing trend of the indicator reflects the decreasing trend of the frequency.

3) More Discussions About Simulated Data Case

From the event detection results based on voltage magnitude and frequency, more conclusions can be summarized as follows.

1) The proposed method is valid for all the data characterizing power grid state, as long as they have corresponding changes when the event occurs. For example, in this subsection, both voltage magnitude and frequency can catch the event and reflect the power grid state. The main reason is that, the event causes the changes of both voltage and frequency measurements, and the indicators catch the event signal based on the changes of the two types of data.

2) The type of data used to build the indicator to monitor the power grid state can be determined by the type of the event and the sensitivity expected. For the same event, the indicators can react differently because of different changing degrees and trends of the corresponding type of data used. To be specific, in case 1, the frequency-based indicator hits up to 1208 at $t=1.02$ s, while the voltage magnitude based indicator reaches up to 548 until $t=1.03$ s. In addition, the frequency-based indicator alarms during $t=1.02$ - 5.00 s, while the voltage magnitude based indicator alarms from 1.02 s to 2.06 s. Also, the accuracies of the two measurements are 100% and 41%, respectively. To conclude, for the active power step increase event, both frequency measurements and voltage magnitudes are effective in detecting the events. However, the frequency measurements are better in this case. For other events related to reactive power, only voltage magnitude will be affected and should be used to monitor the power grid state.

B. Case 2: Case Study with Real PMU Data

In this subsection, the current measurements of a 10 kV distribution network with 11 lines in Guangdong, China, are provided to validate the effectiveness of the proposed method. For each line, there is a PMU monitoring each phase, so we have 33 PMUs in total. According to $c=0.8$ (the same as case 1), the width of the sliding window is set to be 41. The sampling frequency is 100 Hz, and 500 sampling points including the event occurring moment are selected for testing. To be specific, a single-phase grounding fault occurs at the 101st moment and continues for the following 399 sampling points. Details of the current measurements collected are shown in Fig. 4(a). Based on current measurements collected in normal operation state, the alarm threshold is computed to be 93.3. Then, based on the event detection architecture in Fig. 1, the power grid state is monitored and the results are depicted in Fig. 4(b), from which the following two conclusions can be obtained.

1) The proposed method can detect the real-world event. The single-phase grounding fault occurs at the 101st moment.

The ideal performance is to detect the event at the nearest moment after the event occurs (i.e., the 102nd moment). As shown in Fig. 4(b), the event indicator reaches the alarm threshold at the 107th sampling point, delayed by only 0.05 s. In addition, the indicator keeps alarming until the 150th moment, which correctly reflects the increasing and decreasing trends during this period.

2) The proposed method can reflect the real-world power grid state. Although the indicator does not necessarily hit the alarm threshold because of the severity of the event, the changing trend of the measurements and the power grid state can be obtained. For example, from the 41st moment to the 106th moment, although the indictor keeps below the alarm threshold, the first peak during $t = 0.65 - 1.02$ s correctly reflects the first current increase shown in Fig. 4(a). The increasing trend during $t = 1.02 - 1.07$ s reflects the event from $t = 1.01$ s. In addition, between the 151st and the 359th moments, the indicator steadily runs at about −51, which is corresponding to the relatively steady trend of the current measurements during that period. After that, the overall downward and upward trends in Fig. 4(a) are reflected by the two peaks at the 360th and 452nd moments in Fig. 4(b), respectively. Compared with the peak at the 106th moment, these two peaks are not that severe and do not trigger the alarm, because the current changes are not strong enough.

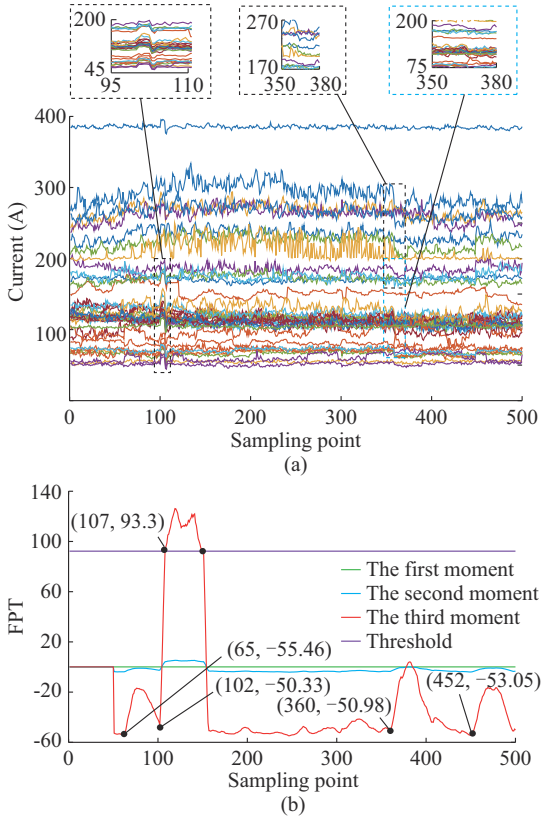

Fig. 4. Event detection results of real-world data. (a) Current. (b) Event detection results.

C. Comparisons with Other Data-driven Methods

To verify the superiority of the FPT-based method over other data-driven methods, this subsection compares the FPT-based method with the statistical MSR-based method in [26] and the training-testing CNN-based method in [16]. The MSR-based method maps the eigenvalues of the covariance matrix of (19) to the complex plane and detects the event by comparing the mean value of the eigenvalues with the limit value under IID. condition. The CNN-based method constructs the feature extraction network and maps the measurements to the event type. In this paper, the CNN-based method is used to do a binary classification problem, where normal and abnormal operation states are the two classes need to be distinguished. The comparisons include the adaptability to the simulated data, the adaptability to the real PMU data, and the sensitivity under different signal-to-noise ratios (SNRs).

1) Comparisons Based on Simulated Data

In this part, the measurements from different events are used for comparisons of the FPT-based method with the MSR-based method and the CNN-based method. As shown in Fig. 5, the performances of MSR-based and CNN-based methods are shown based on frequency measurements from the event in case 1.

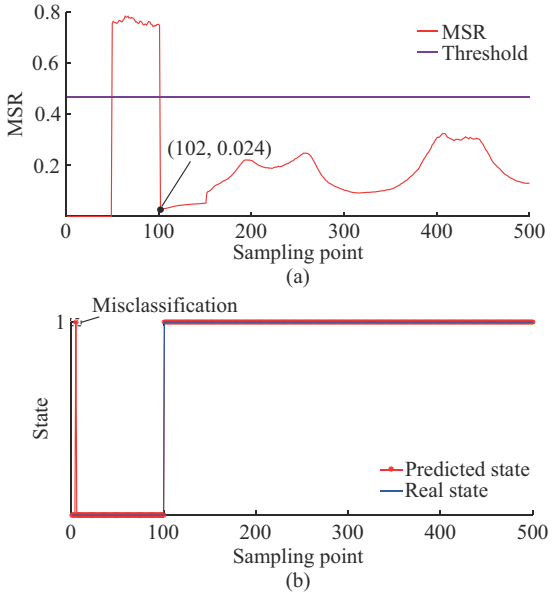

Fig. 5. Performances of different methods based on simulated data. (a) MSR-based method. (b) CNN-based method.

Figure 5(a) depicts the performance of the MSR-based method. For this method, the same sliding window width is set as in FPT-based method. It finds that the MSR-based method can also effectively detect the event and reflect the power grid state. According to [26], the alarm threshold of this method is 0.469. The MSR indicator drops below the threshold at the 102nd sampling point, which is exactly the next sampling point the event occurs. In addition, MSR runs steadily above the threshold during 0.5-1.01 s, which corresponds to the steady stage of the power grid during that period. Then, MSR keeps below the threshold during 1.02-5 s, which correctly reflects the event during that period. To conclude, the MSR-based method shows a comparable performance as the FPT-based method for the simulated data.

Figure 5(b) depicts the performance of the CNN-based method, where 0 represents the normal operation state and 1 denotes the abnormal operation state with event. For this method, the parameter τ in (7) in [16] is set to be 5, the size of images is 20×20, and *ROCOF_net* in [16] is the network used to predict the power grid state. The training set is collected by setting other load changing events and collecting the frequency measurements, and the testing set is collected by the frequency measurements from case 1. To evaluate the power grid state from the first sampling point of the testing set, this paper concatenates another 403 sampling points collected in normal operation state before the testing set. As shown in Fig. 5(b), for the first 101 sampling points, the power grid is depicted as normal in most sampling points, and only 1 point is wrongly classified as abnormal. When the fault is set during 1.01-5.00 s, all the samples are rightly described as abnormal states with events. To conclude, the CNN-based method can effectively detect the event in this case.

From the above results shown in Figs. 3, 5, and 6, we can see that the three methods perform equally well in case 1, because all of them detect the event at the 102nd sampling point and keep alarming for the whole event occurring period.

To further compare the three methods, this paper set other three events, as shown in Table II, to test the performances of the three methods. In this part, the following two indicators are used to evaluate the performances of the event detection methods considered.

TABLE II
PERFORMANCE COMPARISONS OF DIFFERENT DATA-DRIVEN METHODS

Event	Measurement	Method	Performance	
			DM	AR (%)
Load changing in case 1	Voltage	FPT-based	0	56.40
		MSR-based	16	24.20
		CNN-based	70	67.00
Single-phase grounding fault	Current	FPT-based	11	50.30
		MSR-based	32	33.10
		CNN-based	83	50.74
One generator is cut	Frequency	FPT-based	0	100.00
		MSR-based	0	100.00
		CNN-based	0	97.82
One transformer is cut	Voltage	FPT-based	0	100.00
		MSR-based	0	98.64
		CNN-based	0	98.66

1) Delayed moment (DM): DM represents the number of sampling points during the delayed time between the detection of an event and the occurence of that event. DM describes the sensitivity of the event detection method. The smaller this indicator is, the more sensitive the event detection method is.

2) Accuracy ratio (AR): AR denotes the ratio that the method correctly describes the power grid state. The larger the indicator is, the better the method is at detecting events.

Based on the above two indicators, Table II shown the performances of the three methods for different events and measurements. From Table II, the FPT-based method performs best in DM for all the events. As for AR, it shows worse performance than the CNN-based method in the first event in Table II, which is caused by the stable state of the voltage measurements after $t=3.15$ s, as shown in Fig. 2(a). In the single-phase grounding fault, the FPT-based method achieves 50.30% in AR, which is slightly lower than 50.74% of the CNN-based method. This inferiority is really slight and can be caused by statistical errors, so we can say that the two methods have comparable performance in this event when it comes to the AR. In addition, the FPT-based method has much lower DM than the CNN-based method in this event (11 v. s. 83), which means that the FPT-based method is much more sensitive than the CNN-based method. Overall, the FPT-based method is the most sensitive and accurate method compared with the MSR-based and CNN-based methods.

2) Comparisons Based on Real PMU Data

This paper also compares the adaptability of the above methods towards the real PMU data mentioned in case 2. As shown in Fig. 6 and Fig. 4(b), the conclusions are as follows.

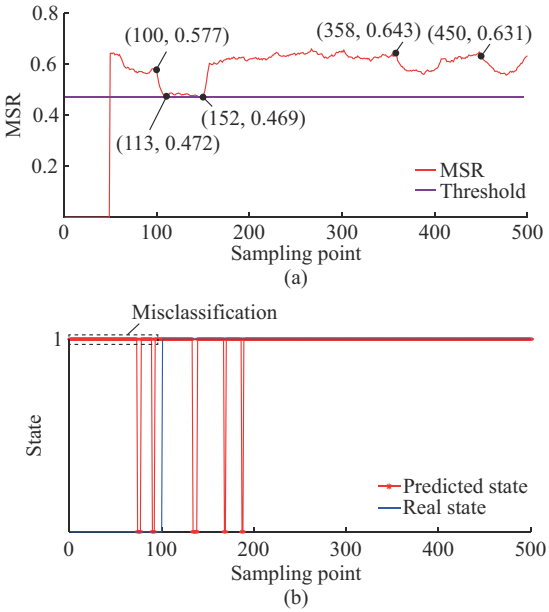

Fig. 6. Comparison of different methods based on real PMU data. (a) MSR-based method. (b) CNN-based method.

1) The MSR-based method is effective in reflecting the operation trend of the power grid, but has limited sensitivity in detecting events. As shown in Fig. 6(a), the MSR indicator shows similar changing trend as the FPT-based method. For example, MSR shows downward trends from $t=1.01$ s, $t=3.58$ s, and $t=4.50$ s, which corresponds to the upward trends of the FPT-based method when $t=1.02$ s, $t=3.60$ s, and $t=4.52$ s, respectively, in Fig. 4(b). However, MSR triggers the threshold only when $t=1.13$ s and $t=1.52$ s, which means that MSR has limited sensitivity. We may need to further lower the threshold for MSR to make it adaptable to the real PMU data.

2) For the CNN-based method, its effectiveness towards the real-world cases is limited. To collect the training set, a distribution network model is built according to the real PMU distribution network in case 2, and other single-phase grounding faults are set to collect the current measurements. Then, the CNN-based method is tested using the real PMU data from case 2. From Fig. 6(b), we can find that the CNN-based method fails to detect the event, because it classifies almost all the sampling points as abnormal operation states. The reason is that the method in [16] constructs the images based on the differences between measurements. However, the real-world measurements are always with relatively severe fluctuations, thus the differences between the measurements are relatively large and the CNN-based method mistakenly classifies the large differences as abnormal operation states. What needs to be mentioned is that, during the training stage, some other real-world current measurements are also used to train the CNN-based method. However, because of the small number of the real-word training samples, the CNN-based method still performs poorly and shows poor generalization performance. From this case, we can conclude that the FPT-based method has better adaptability over the CNN-based method for the real-world PMU data.

3) Comparisons Under Different SNRs

The frequency measurements in case 1 are used to compare the sensitivity of the three methods under different SNRs. The event detection matrix is constructed by $\boldsymbol{Y}=\boldsymbol{Y}+m\boldsymbol{E}$, where m is the index of SNR; and $\boldsymbol{E}$ is the noise matrix meeting $\boldsymbol{E}\sim\mathcal{N}(0,1)$. This paper sets m to be 0.0005 and 0.0010, respectively, to mimic the increased noise and the decreased SNR. The performances of the three methods are shown in Figs. 7-9, respectively. The comparisons between the three methods are shown in Table III, from which we can conclude that although the abilities of the three methods to detect anomalies deteriorate with the decrease of SNR, the FPT-based method has the best sensitivity. More details are as follows.

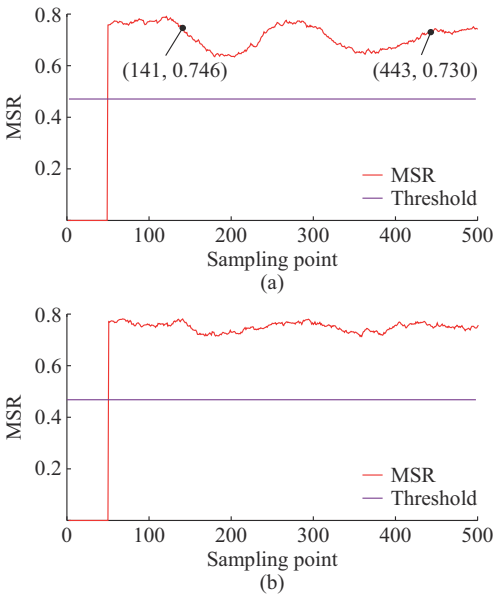

Fig. 7. Performance of MSR-based method under different SNRs. (a) $m=0.0005$. (b) $m=0.001$.

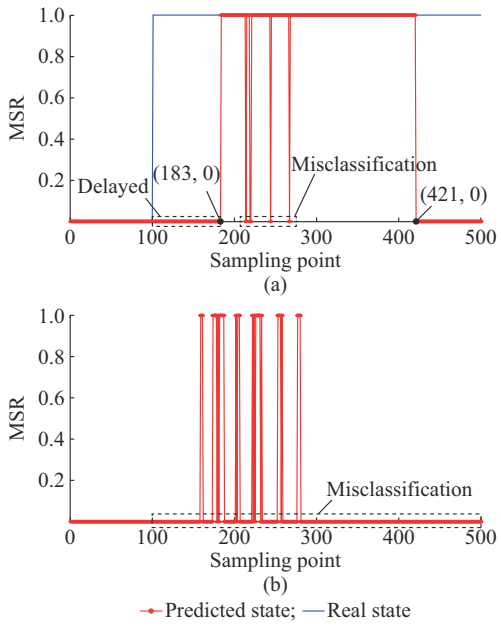

Fig. 8. Performance of CNN-based method under different SNRs. (a) $m=0.0005$. (b) $m=0.001$.

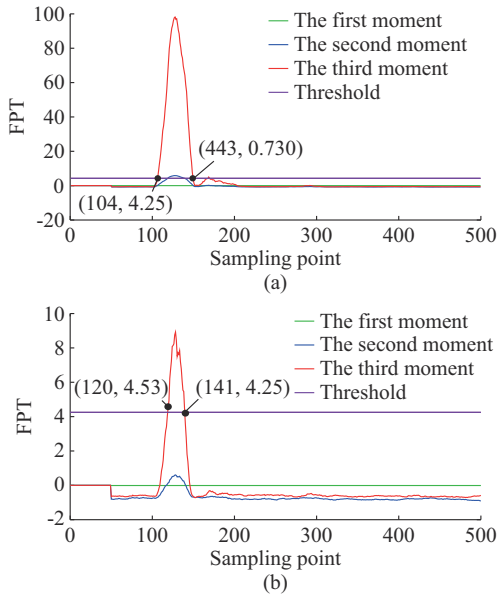

Fig. 9. Performance of FPT-based method under different SNRs. (a) $m=0.0005$. (b) $m=0.001$.

TABLE III
COMPARISONS OF DIFFERENT DATA-DRIVEN METHODS UNDER DIFFERENT SNRS

SNR	Method	Performance	
		DM	AR (%)
0.0005	FPT-based	2	29.4
	MSR-based		
	CNN-based	81	66.6
0.0010	FPT-based	18	24.4
	MSR-based		
	CNN-based	74	22.4

1) The MSR-based method is the least sensitive among the three methods. As shown in Fig. 7, the indicator calculated from this method stays above the threshold under two SNRs, which means that the MSR-based method is not sensitive enough to trigger the alarm threshold and detect the event when the noise is not strong.

2) The CNN-based method can detect the event when the SNR is 0.0005, because it classifies the power grid state as abnormal during 1.83-4.21 s. However, when the SNR increases to 0.001, this method fails to detect the event because only 12 sampling points are classified as abnormal and they are discontinuous, so it is difficult to distinguish the event from the noise.

3) The FPT-based method is the most sensitive and effective method to detect events. As shown in Fig. 9 and Table III, this method triggers the alarm threshold with the least DM under two SNRs. In addition, it keeps alarming for a while in both cases.

4) Although the FPT-based method shows lower AR compared with the CNN-based method when the SNR is 0.0005, it has smaller DM. In addition, the FPT-based method works when SNR increases to 0.001, while the CNN-based method does not. Overall, the FPT-based method outperforms the other two data-driven methods for low SNRs.

4) More Comparisons Between CNN-based Method and FPT-based Method

This paper uses extra training sets for the CNN-based method. However, it shows worse performance than the FPT-based method in most cases, especially when it comes to sensitivity. What is more, although the CNN-based method can classify the events, it requires large amount of training sets for each type of event, which is unavailable in real power systems, and the simulated data lead to poor generalization. While for the FPT-based method, only the data collected from the normal operation state are needed, which is more applicable in real power systems.

5) Orders of Moments and Cumulants of Event Detection Indicator

Based on the simulations above, this part describes how the orders of moments and cumulants of the event detection indicator are decided, from the perspectives of sensitivity and robustness.

Figures 2(b), 3(b), and 4(b) describe the event detection results by the first, second, and third moments in three different cases, respectively. In the above three figures, we can see that the first moment is unable to detect the event because the data are normalized during the calculation.

For the ideal simulated data from Figs. 2(a) and 3(a), both the second and third moments can detect the event, because both of them keep steady when the power grid operates normally, and show increasing trend when there is an event. The difference is that the third moment is much more sensitive than the second one, which can make it easier to catch the event signals.

When it comes to the real-world case, as shown in Fig. 4(b), although the third moment fluctuates more because of its better sensitiveness, it does not hit the threshold and can distinguish between the normal operation state and the event, so its robustness is acceptable. In addition, the second order moment is so robust that it may fail to detect some data stream changing. For example, for the decreasing trend at around 3.6 s, the second moment stays steady and does not catch the event signal while the third moment successfully catch it.

To conclude, the third moment can meet the requirements of both sensitivity and robustness. This paper does not use higher order more than the third order, because the higher the order, the more sensitive the indicator, which may not fit the real-world case. In addition, the higher order will take more time to do the calculation.

V. Conclusion

IoT enables the real-time big data collection and online state monitoring possible in power systems. In this paper, a data-driven event detection architecture is proposed based on FPT. The event detection matrix constructed by IoT-enabled measurements is described as an information-plus-noise model, where normal fluctuations of measurements, background noises, and measurement errors are regarded as "noises", and the event signal is taken as the information. Then, using FPT as the tool, the ESD-based moments of the event information matrix are extracted from the information-plus-noise model, where the noise matrix is constructed using the measurements collected under normal state, and the alarm threshold is determined based on the CLT. Besides, the online event detection is conducted based on the sliding window. Based on IoT-enabled measurements from PMUs, case studies prove that our method is effective in detecting the events using voltage magnitude, frequency, and current. In addition, because of consideration of the "noises", the proposed method can reduce the influences of the "noises" mentioned above, so it has better sensitivity and is more adaptive to the real PMU data.

The work in this paper can help handle the challenges of power systems under IoT. In the future, data streams from different IoT-enabled devices will be considered to provide a more universal solution for power system event detection.

References

[1] Z. Zhang, Z. Chen, Q. Zhao *et al*., "Situation awareness and sensitivity analysis for absorption of grid-connected renewable energy power generation integrating robust optimization and radial basis function neural network," *Journal of Modern Power Systems and Clean Energy*, vol. 10, no. 6, pp. 1795-1803, Nov. 2023.

[2] W. Wang, L. Liu, J. Liu *et al*., "Energy management and optimization of vehicle-to-grid systems for wind power integration," *CSEE Journal of Power and Energy Systems*, vol. 7, no. 1, pp. 172-180, Jan. 2021.

[3] R. Morello, C. de Capua, G. Fulco *et al*., "A smart power meter to monitor energy flow in smart grids: the role of advanced sensing and IoT in the electric grid of the future," *IEEE Sensors Journal*, vol. 17, no. 23, pp. 7828-7837, Dec. 2017.

[4] G. Bedi, G. K. Venayagamoorthy, R. Singh *et al*., "Review of Internet of Things (IoT) in electric power and energy systems," *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 847-870, Apr. 2018.

[5] H. Jahangir, S. Lakshminarayana, C. Maple *et al*., "A deep-learning-based solution for securing the power grid against load altering threats by IoT-enabled devices," *IEEE Internet of Things Journal*, vol. 10, no. 12, pp. 10687-10697, Jan. 2023.

[6] I. Kosen, C. Huang, Z. Chen *et al*., "UPS: unified PMU-data storage system to enhance T+D PMU data usability," *IEEE Transactions on*

Smart Grid, vol. 11, no. 1, pp. 739-748, Jan. 2020.
[7] A. G. Bruce, "Reliability analysis of electric utility SCADA systems," *IEEE Transactions on Power Systems*, vol. 13, no. 3, pp. 844-849, Aug. 1998.
[8] Z. Li, J. Wan, P. Wang *et al*., "A novel fault section locating method based on distance matching degree in distribution network," *Protection and Control of Modern Power Systems*, vol. 6, no. 1, p. 20, Dec. 2021.
[9] S. Admasie, S. B. A. Bukhari, T. Gush *et al*., "Intelligent islanding detection of multi-distributed generation using artificial neural network based on intrinsic mode function feature," *Journal of Modern Power Systems and Clean Energy*, vol. 8, no. 3, pp. 511-520, May 2020.
[10] C. Wang, H. He, and T. Zhu, "Impacts of the blackout on August 14 on the renovation of power market in China," in *Proceedings of ESMO 2006 – 2006 IEEE 11th International Conference on Transmission & Distribution Construction*, *Operation and Live-line Maintenance*, Albuquerque, USA, Oct. 2006, pp.1-5.
[11] G. Zhou, X. Zhang, M. Han *et al*., "Single-ended fault detection scheme using support vector machine for multi-terminal direct current systems based on modular multilevel converter," *Journal of Modern Power Systems and Clean Energy*, vol. 11, no. 3, pp. 990-1000, May 2023.
[12] M. D. Borrás, J. C. Bravo, and J. C. Montaño, "Disturbance ratio for optimal multi-event classification in power distribution networks," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 5, pp. 3117-3124, May 2016.
[13] D. Feng, Z. Deng, T. Wang *et al*., "Identification of disturbance sources based on random forest model," in *Proceedings of 2018 International Conference on Power System Technology* (*POWERCON*), Guangzhou, China, Nov. 2018, pp. 3370-3375.
[14] T. Wu, Y. Zhang, and X. Tang, "Online detection of events with low-quality synchrophasor measurements based on *i*Forest," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 1, pp. 168-178, Jan. 2021.
[15] S. Zhang, Y. Wang, M. Liu *et al*., "Data-based line trip fault prediction in power systems using LSTM networks and SVM," *IEEE Access*, vol. 6, pp. 7675-7686, Dec. 2017.
[16] W. Wang, H. Yin, C. Chen *et al*., "Frequency disturbance event detection based on synchrophasors and deep learning," *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 3593-3605, Jul. 2020.
[17] X. Liu, S. Lin, J. Fang *et al*., "Is extreme learning machine feasible? A theoretical assessment (part I)," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 1, pp. 7-20, Jul. 2015.
[18] Y. Kanagawa and T. Kaneko, "Rogue-gym: a new challenge for generalization in reinforcement learning," in *Proceedings of 2019 IEEE Conference on Games* (*CoG*), London, UK, Aug. 2019, pp. 1-8.
[19] S. S. Negi, N. Kishor, K. Uhlen *et al*., "Event detection and its signal characterization in PMU data stream," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 6, pp. 3108-3118, Dec. 2017.
[20] D. I. Kim, T. Y. Chun, S. H. Yoon *et al*., "Wavelet-based event detection method using PMU data," *IEEE Transactions on Smart Grid*, vol. 8, no. 3, pp. 1154-1162, May 2017.
[21] P. Wang, H. Wang, P. Hart *et al*., "Application of Chebyshev's inequality in online anomaly detection driven by streaming PMU data," in *Proceedings of 2020 IEEE General Meeting*, Montreal, Canada, Aug. 2020, pp. 1-5.
[22] L. Xie, Y. Chen, and P. R. Kumar, "Dimensionality reduction of synchrophasor data for early event detection: linearized analysis," *IEEE Transactions on Power Systems*, vol. 29, no. 6, pp. 2784-2794, Nov. 2014.
[23] X. Liu, D. M. Laverty, R. J. Best *et al*., "Principal component analysis of wide-area phasor measurements for islanding detection – a geometric view," *IEEE Transactions on Power Delivery*, vol. 30, no. 2, pp. 976-985, Apr. 2015.
[24] X. Liu, J. Kennedy, D. Laverty *et al*., "Wide area phase angle measurements for islanding detection – an adaptive nonlinear approach," in *Proceedings of 2017 IEEE General Meeting*, Chicago, USA, Jul. 2017, pp. 1-10.
[25] X. He, L. Chu, R. C. Qiu *et al*., "A novel data-driven situation awareness approach for future grids – using large random matrices for big data modeling," *IEEE Access*, vol. 6, pp. 13855-13865, Mar. 2018.
[26] X. He, Q. Ai, R. Qiu *et al*., "A big data architecture design for smart grids based on random matrix theory," *IEEE Transactions on Smart Grid*, vol. 8, no. 2, pp. 674-686, Mar. 2017.
[27] F. Yang, M. Wei, Z. Ling *et al*., "Brown measure based spectral distribution analysis for spatial-temporal localization of cascading events in power grids," *IEEE Transactions on Smart Grid*, vol. 12, no. 2, pp. 1805-1820, Mar. 2021.
[28] A. Skupch, D. Seethaler, and F. Hlawatsch, "Free probability based capacity calculation for MIMO channels with transmit or receive correlation," in *Proceedings of 2005 International Conference on Wireless Networks*, *Communications and Mobile Computing*, Maui, USA, Jul. 2005, pp. 1041-1046.
[29] H. Wang, B. Wang, P. Luo *et al*., "State evaluation based on feature identification of measurement data: for resilient power system," *CSEE Journal of Power and Energy Systems*, vol. 8, no. 4, pp. 983-992, Jul. 2022.
[30] O. Ryan and M. Debbah, "Free deconvolution for signal processing applications," in *Proceedings of 2007 IEEE International Symposium on Information Theory*, Nice, France, Jun. 2007, pp. 1846-1850.
[31] Ø. Ryan and M. Debbah. (2007, Feb.). Multiplicative free convolution and information-plus-noise type matrices. [Online]. Available: http://arxiv.org/pdf/math/070234221.pdf
[32] V. Dan, "The analogues of entropy and of Fisher's information measure in free probability theory, I," *Communications in Mathematical Physics*, vol. 155, no. 1, pp. 71-92, Jan. 1993.
[33] B. Wang, H. Wang, L. Zhang *et al*., "A data-driven method to detect and localize the single-phase grounding fault in distribution network based on synchronized phasor measurement," *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, pp. 1-13, Aug. 2019.
[34] Y. W. Law, M. Palaniswami, G. Kounga *et al*., "WAKE: key management scheme for wide-area measurement systems in smart grid," *IEEE Communications Magazine*, vol. 51, no. 1, pp. 34-41, Jan. 2013.
[35] L. Zhu and J. Lin, "Learning spatiotemporal correlations for missing noisy PMU data correction in smart grid," *IEEE Internet of Things Journal*, vol. 8, no. 9, pp. 7589-7599, Nov. 2021.

Hongxia Wang received the B.S. degree in electrical engineering and automation from Hefei University of Technology, Hefei, China, in 2018, and the M.S. degree in School of Electrical Engineering and Automation, Wuhan University, Wuhan, China, in 2020. She is currently pursuing the Ph.D. degree in School of Electrical Engineering and Automation, Wuhan University. Her research interests include big data and data fusion in power systems.

Bo Wang received the Ph.D. degree in computer science from Wuhan University, Wuhan, China, in 2006. He did the Post-doctoral Research with the School of Electrical Engineering and Automation, Wuhan University, from 2007 to 2009. He is currently a Professor with the School of Electrical Engineering and Automation, Wuhan University. His research interests include power system online assessment, big data, and integrated energy systems.

Jiaxing Zhang received M.Sc. degree from the School of Electrical Engineering and Automation, Wuhan University, Wuhan, China, in 2022. She is currently pursuing the Ph.D. degree in the School of Electrical Engineering and Automation, Wuhan University. Her current research interests include distribution network online assessment and digitization of distribution network.

Chengxi Liu received the B.S. and M.S. degrees in electrical engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2005 and 2007, respectively, and the Ph.D. degree from Aalborg University, Aalborg, Denmark, in 2013. He worked as System Analyst with Energinet.dk (the Danish TSO), Fredericia, Danmark, until 2016, and a Research Associate with the University of Tennessee, Knoxville, USA, from 2016 to 2018. He was an Associate Professor with the Department of Energy Technology, Aalborg University. He is currently a Professor with the School of Electrical Engineering and Automation, Wuhan University, Wuhan, China. His research interests include power system stability and analysis, simulation methods of power systems, machine learning applied on power systems.

Hengrui Ma received the Ph.D. degree in power system and automation from the School of Electrical Engineering, Wuhan University, Wuhan, China, in 2018. His research interests include integrated energy systems and energy storage systems.