Multivariate Detection of Power System Disturbances Based on Fourth Order Moment and Singular Value Decomposition

Lianfang Cai, Nina F. Thornhill, *Senior Member, IEEE*, and Bikash C. Pal, *Fellow, IEEE*

***Abstract*—This paper presents a new method to detect power system disturbances in a multivariate context, which is based on Fourth Order Moment and multivariate analysis implemented as Singular Value Decomposition. The motivation for this development is that power systems are increasingly affected by various disturbances and there is a requirement for the analysis of measurements to detect these disturbances. The application results on the measurements of an actual power system in Europe illustrate that the proposed multivariate detection method achieves enhanced detection reliability and sensitivity.**

***Index Terms*—Disturbance, fourth order moment, multivariate, power system, singular value decomposition.**

I. Introduction

POWER system operation and control are becoming more and more complex because of the high penetration of renewable generation and increasing electricity consumption. Meanwhile, a variety of power system disturbances pose an increasingly severe threat to system security [1]. In this challenging context, detecting these disturbances effectively plays a crucial role in improving the system security and stability.

The power system disturbances generally result from the loss of a synchronous generator or tie-line, the penetration of heavy or light loads, the intermittency of renewable generation and sudden changes in the operation conditions. They usually take the form of deviations from the previous and subsequent trends [2], [3]. Detection of power system disturbances is a difficult task because of system complexity, diversity of operating conditions and interference from noise. It places high demands on reliable, sensitive and real-time implementation and is of concern in power system monitoring and control, as indicated by present research [4]–[11].

Manuscript received April 6, 2016; revised September 8, 2016; accepted November 12, 2016. Date of publication March 2, 2017; date of current version October 18, 2017. This work was supported by the U.K. Engineering and Physical Sciences Research Council under Grant "EP/L014343/1" "Stability and Control of Power Networks with Energy Storage (STABLE-NET)." Paper no. TPWRS-00536-2016.

L. Cai and N. F. Thornhill are with the Centre for Process Systems Engineering, Imperial College London, London SW7 2AZ, U.K. (e-mail: l.cai@imperial.ac.uk; n.thornhill@imperial.ac.uk).

B. C. Pal is with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, U.K. (e-mail: b.pal@imperial.ac.uk).

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

Digital Object Identifier 10.1109/TPWRS.2016.2633321

Advanced measuring devices, such as Phasor Measurement Units (PMUs), provide abundant measurements for the development of the data-driven disturbance detection. The existing data-driven approaches can be classified into three main categories according to the applications: (1) for the protection of power systems, e.g., the hidden Markov model based method [4] and the wavelet coefficient energy based method [5]; (2) for the assessment of power quality (mainly focusing on alternate voltage), e.g., the power quality state estimation based method [6] and the wavelet packet and Tsallis entropy based method [7]; (3) for the wide-area monitoring of power systems, typically, the principal component analysis based methods [8], [9]. Usually, the first two categories of methods take a univariate approach to deal with electrical variables separately, whereas the third category of methods use multivariate analysis to handle all the electrical variables together. This work focuses on the latter.

A common limitation of the present methods is that the non-Gaussian information in the electrical measurements has not been much explored previously. It has been reported in [12], [13] that the recorded measurements usually have a non-Gaussian distribution due to system nonlinearity and the non-Gaussian information is important for system monitoring. Usually, the non-Gaussian information needs high order (order greater than two) analysis. As indicated in [14], fourth order moment (FOM) contains significant non-Gaussian information. The motivation of this work is to explore such information by FOM for disturbance detection. A further motivation is to combine FOM with multivariate analysis to meet the requirement of the wide-area monitoring.

Against this background, this paper proposes a new multivariate detection method based on a combination of high order statistical analysis known as FOM and multivariate analysis known as Singular Value Decomposition (SVD). Firstly, the non-Gaussianity of each electrical variable is quantified through the robust estimation of negentropy (NE) which is widely used for evaluating the extent of non-Gaussianity [13]. Then, the non-Gaussian information in the measurements of each variable is explored by FOM to provide fourth order statistics. Furthermore, the obtained fourth order statistics are dealt with using SVD. Through this multivariate analysis, the presence of the same disturbance in the measurements of different variables can be jointly explored to allow a more satisfactory detection. A case study involving measurements from an actual power

system in Europe (called European power system here) is used to demonstrate the effectiveness of the proposed method.

This paper is organized as follows. Section II presents the univariate detection method based on FOM. Then, the multivariate detection method based on a combination of FOM and SVD is detailed in Section III. The application results and analysis on the case study are provided in Section IV, while our conclusions are drawn in Section V.

II. Univariate Detection Based on FOM

In this section, a univariate detection method based on FOM, referred to as UD-FOM, is presented to lay the foundation for the subsequent multivariate detection. UD-FOM consists of three parts with the details as follows.

A. Robust Estimation of NE

The symbols $x_1, x_2, \cdots, x_m$ denote m measured electrical variables. The non-Gaussianity of each variable can be quantified using NE. For the *i*-th variable x_i, its NE can be expressed as follows:

$$NE_i = H_G(\nu) - H_i({x_i}^*) \tag{1}$$

where $H_G(\nu) = -\int_{-\infty}^{\infty} p_G(\nu)\log p_G(\nu)d\nu$ denotes the entropy of the Gaussian variable ν with zero mean and unit variance, $H_i({x_i}^*) = -\int_{-\infty}^{\infty} p_i({x_i}^*)\log p_i({x_i}^*)d{x_i}^*$ denotes the entropy of ${x_i}^*$, while ${x_i}^* = \frac{x_i - E(x_i)}{std(x_i)}$ is the normalized version of x_i with $E(\cdot)$ and $\mathrm{std}(\cdot)$ respectively denoting the expectation operator and the standard deviation operator, $p_i({x_i}^*)$ and $p_G(\nu)$ denote the probability density functions of ${x_i}^*$ and ν, and $\log(\cdot)$ denotes the natural logarithmic function. $NE_i >= 0$, and $NE_i = 0$ only when x_i has a Gaussian distribution. In practice, it is difficult to obtain an exact zero value for NE_i even if x_i has a Gaussian distribution, because NE_i needs to be estimated based on the measurements of x_i and the number of measurements in the dataset may exert some effect. Thus, a small threshold value slightly larger than zero is supposed to be preset. Here, the value 10^{-3} suggested in [15] is also taken as the threshold since it provides a good tradeoff between the theoretical significance and practical implementation.

From (1), it can be seen that the calculation of NE_i involves the probability density estimation and an integration operation that are difficult and burdensome to implement. Practically, the entropy $H_G(\nu)$ can be proven a constant:

$$H_G(\nu) = \log\sqrt{2\pi} + \int_{-\infty}^{\infty} \frac{\nu^2}{2\sqrt{2\pi}} e^{-\frac{\nu^2}{2}} d\nu = \log\sqrt{2\pi} + \frac{1}{2} \tag{2}$$

Thus, to estimate NE_i, only the estimation of $H_i({x_i}^*)$ needs to be considered. This can be achieved by the following robust entropy estimator that has been demonstrated to have the excellent performance even under the effect of outliers [16]:

$$\hat{H}_i\left({x_i}^{*(1)}, {x_i}^{*(2)}, \cdots, {x_i}^{*(N)}\right) = \frac{1}{N - \left\lfloor\sqrt{N}\right\rfloor} \sum_{j=1}^{N-\left\lfloor\sqrt{N}\right\rfloor} \log\left(\frac{N+1}{\left\lfloor\sqrt{N}\right\rfloor}\left({x_i}^{*\left(j+\left\lfloor\sqrt{N}\right\rfloor\right)} - {x_i}^{*(j)}\right)\right) \tag{3}$$

where ${x_i}^{*(1)}, {x_i}^{*(2)}, \cdots, {x_i}^{*(N)}$ denote the N realizations of ${x_i}^*$ arranged in the ascending order, and $\lfloor\cdot\rfloor$ denotes the round down operator. Substituting (2) and (3) into (1), the robust estimation of NE_i can be obtained. Then, the non-Gaussian information in the measurements of x_i can be explored in the following part.

B. Non-Gaussian Information Exploration Using FOM

The non-Gaussian information in the electrical measurements usually requires high order analysis. According to [14], FOM contains significant non-Gaussian information. It is thus taken to explore such high order information here. For the variable x_i, its FOM can be written as (4) [14]:

$$FOM_i = E\left(x_i(k)\, x_i(k-\tau_1)\, x_i(k-\tau_2)\, x_i(k-\tau_3)\right) \tag{4}$$

where k is a time variable denoting the sampling time point, $x_i(k)$ denotes the measurement of x_i at the *k*-th sampling time point, while τ_1, τ_2 and τ_3 denote the time lags, and $\tau_3 >= \tau_2 >= \tau_1 >= 0$.

In statistical analysis, it is often assumed that the signal is stationary or ergodic [15]. Only in this condition, FOM_i can be estimated when a number of samples are obtained. However, this is applicable for the off-line condition. For the on-line situation, FOM_i needs to be calculated approximately as the instantaneous term $x_i(k)x_i(k-\tau_1)x_i(k-\tau_2)x_i(k-\tau_3)$. The main requirement here is that the statistical nature of the underlying physical system that generates the measurements must not change erratically at over a timescale encompassed by the maximum delay τ_3. This can be more easily met by assigning the maximum delay a smaller value, as was done in [17]. Afterwards, the following fourth order statistics can be taken for the evaluation of the non-Gaussian information in the original measurements:

$$c_i(k) = x_i(k)\, x_i(k-\tau_1)\, x_i(k-\tau_2)\, x_i(k-\tau_3) \tag{5}$$

Regarding the use of (5), the choices of τ_1, τ_2 and τ_3 need to be considered. Generally, they should match with the sampling rate of the measurements and the dynamics of the disturbances. Since the sampling rate relies on the measuring instruments and the disturbances are unpredictable before detection, there is no universal criterion for determining the time lags. However, the absolute differences between the time lags should not be large from the detection perspective, as stated in [17], and this literature suggested to set them as one. Accordingly, $\tau_1 = 1$, $\tau_2 = 2$ and $\tau_3 = 3$ can be a good choice, which can not only be in accordance with the above discussion, but also can reflect the cumulative effect of the present measurement and three most recent measurements on the non-Gaussian information.

C. Monitoring Statistic and Detection Threshold

To conduct the univariate detection, a monitoring statistic for the i -th variable x_i can be constructed using c_i as:

$$UD_i(k) = |c_i(k)|, \quad i = 1, 2, \cdots, m \tag{6}$$

where $|\cdot|$ denotes the absolute value operator.

Besides, a global characterization of the group of variables with regard to the disturbances can be given by a system-wide monitoring statistic as shown below:

$$UD(k) = \frac{1}{m}\sum_{i=1}^{m} |c_i(k)| \tag{7}$$

After the construction of the monitoring statistics, the related detection thresholds need to be determined for judging whether a disturbance occurs or not. As no prior knowledge is available with regard to the distribution of UD_i and UD, their detection thresholds $UD_{i,\alpha}$ and UD_α with the confidence level α can be determined as follows. Based on the training dataset $\{x_i(k)\}_{k=1}^{N}$ historically measured on x_i under the ambient condition with no disturbance, the time-series values $\{c_i(k)\}_{k=\tau_3+1}^{N}$ of c_i are calculated using (5). Then, the time-series values $\{UD_i(k)\}_{k=\tau_3+1}^{N}$ of UD_i are calculated using (6), while the time-series values $\{UD(k)\}_{k=\tau_3+1}^{N}$ of UD are calculated using (7). Finally, $(1-\alpha)(N-\tau_3)$ is rounded towards the nearest integer δ, and the δ-th highest value of $\{UD_i(k)\}_{k=\tau_3+1}^{N}$ is adopted as the detection threshold for UD_i, while the δ-th highest value of $\{UD(k)\}_{k=\tau_3+1}^{N}$ is adopted as the detection threshold for UD.

On completion of the above modelling procedure, the on-line disturbance detection can be performed. For the present measurement $x_i(p)$ collected online from x_i, where $k = p$ represents the present sampling time point, the present value $c_i(p)$ of c_i can be calculated using (5). Then, the present value $UD_i(p)$ of UD_i can be calculated using (6) and compared with $UD_{i,\alpha}$, while the present value $UD(p)$ of UD can be calculated using (7) and compared with UD_α. If $UD_{i,\alpha}$ is exceeded consecutively for a number of sampling time points, a disturbance is then detected in the variable x_i, while if UD_α is exceeded consecutively for a number of sampling time points, a disturbance is then detected at the system-wide level.

Thus, UD-FOM which can explore the non-Gaussian information in the measurements has been developed. However, it performs the disturbance detection in a univariate analysis manner and cannot take the correlation between variables into account. In the next section, UD-FOM is extended to a multivariate detection method by means of SVD, for further using correlation information to detect disturbances.

III. Multivariate Detection Based on FOM and SVD

To explore correlation information in the measurements besides the non-Gaussian information, a multivariate detection method based on FOM and SVD, referred to as MD-FOMSVD, is proposed here. In the following, MD-FOMSVD is presented in detail by four parts.

A. Multivariate Extension Using SVD

SVD is a multivariate statistical analysis technique, which can factorize any data matrix into a product of three other matrices with specific properties as follows:

$$\boldsymbol{C} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^{\mathrm{T}} = \sum_{j=1}^{m} \boldsymbol{u}_j s_j \boldsymbol{v}_j^{\mathrm{T}} \tag{8}$$

where the columns $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_m$ of $\boldsymbol{U}$ are orthonormal basis functions for the columns of $\boldsymbol{C}$, while the rows $\boldsymbol{v}_1^{\mathrm{T}}, \boldsymbol{v}_2^{\mathrm{T}}, \cdots, \boldsymbol{v}_m^{\mathrm{T}}$ of $\boldsymbol{V}^{\mathrm{T}}$ are orthonormal basis functions for the rows of $\boldsymbol{C}$, $(\cdot)^{\mathrm{T}}$ denotes the transpose operator, and $\boldsymbol{S}$ is a diagonal matrix with diagonal elements $s_1, s_2, \cdots, s_m$ as the descending-order singular values (SVs) of $\boldsymbol{C}$. The square of a SV is directly associated with the total variance of $\boldsymbol{C}$ along the direction defined by the corresponding row of $\boldsymbol{V}^{\mathrm{T}}$. This means that the first few rows of $\boldsymbol{V}^{\mathrm{T}}$ are the directions capturing the largest variance proportion of the rows of $\boldsymbol{C}$.

The objective of applying SVD is to further identify principal features from $c_1, c_2, \cdots, c_m$, where principal features are those capturing the largest proportion of the variances of $c_1, c_2, \cdots, c_m$. This usually occurs when a feature is present in many of the correlated variables.

Thus, MD-FOMSVD starts from c_i generated by (5). That is, based on the training dataset $\{x_i(k)\}_{k=1}^{N}$, the time-series values $\{c_i(k)\}_{k=\tau_3+1}^{N}$ of c_i are calculated using (5) and treated as the elements of the i-th row $\boldsymbol{c}_i^{\mathrm{T}}$ of $\boldsymbol{C}$, and then $\boldsymbol{C}$ is factorized with SVD as the following equation:

$$\begin{aligned}\boldsymbol{C} &= \begin{bmatrix} \boldsymbol{c}_1^{\mathrm{T}} \\ \boldsymbol{c}_2^{\mathrm{T}} \\ \vdots \\ \boldsymbol{c}_m^{\mathrm{T}} \end{bmatrix} = \begin{bmatrix} c_1(\tau_3+1) & c_1(\tau_3+2) & \cdots & c_1(N) \\ c_2(\tau_3+1) & c_2(\tau_3+2) & \cdots & c_2(N) \\ \vdots & \vdots & \cdots & \vdots \\ c_m(\tau_3+1) & c_m(\tau_3+2) & \cdots & c_m(N) \end{bmatrix} \\ &= \sum_{j=1}^{m} \begin{bmatrix} u_{1,j} \\ u_{2,j} \\ \vdots \\ u_{m,j} \end{bmatrix} s_j \boldsymbol{v}_j^{\mathrm{T}} \end{aligned} \tag{9}$$

where $m < N - \tau_3$, and $u_{r,j}$ denotes the r-th row and the j-th column element of $\boldsymbol{U}$.

Inspired by the statement in the work [3], two important issues are now discussed, regarding the implementation of SVD. Firstly, power system disturbances have a larger impact on principal features that account for most of the variance of the rows of $\boldsymbol{C}$. Those principal features of interest are captured by the first few basis functions $\boldsymbol{v}_1^{\mathrm{T}}, \boldsymbol{v}_2^{\mathrm{T}}, \cdots, \boldsymbol{v}_{m'}^{\mathrm{T}}$, where $m' < m$, whereas the remaining basis functions should capture the details of comparatively less significance to the description of disturbances. Accordingly, the determination of m' should be performed to remove the basis functions $\boldsymbol{v}_{m'+1}^{\mathrm{T}}, \boldsymbol{v}_{m'+2}^{\mathrm{T}}, \cdots, \boldsymbol{v}_m^{\mathrm{T}}$ as well as the related SVs $s_{m'+1}, s_{m'+2}, \cdots, s_m$ and the column vectors $\boldsymbol{u}_{m'+1}, \boldsymbol{u}_{m'+2}, \cdots, \boldsymbol{u}_m$. How to choose an appropriate m' is discussed later in Section III-D. After this selection step

called *Selection* Γ here, the data matrix $\hat{\boldsymbol{C}}$ is formed by the retained $\boldsymbol{u}_j, s_j, \boldsymbol{v}_j^{\mathrm{T}}, j = 1, 2, \cdots, m'$ as:

$$\hat{\boldsymbol{C}} = \begin{bmatrix} \hat{\boldsymbol{c}}_1{}^{\mathrm{T}} \\ \hat{\boldsymbol{c}}_2{}^{\mathrm{T}} \\ \vdots \\ \hat{\boldsymbol{c}}_m{}^{\mathrm{T}} \end{bmatrix} = \boldsymbol{U}_{:,1:m'} \boldsymbol{S}_{1:m'} \boldsymbol{V}_{1:m',:}^{\mathrm{T}} = \sum_{j=1}^{m'} \begin{bmatrix} u_{1,j} \\ u_{2,j} \\ \vdots \\ u_{m,j} \end{bmatrix} s_j \boldsymbol{v}_j^{\mathrm{T}} \tag{10}$$

where $\boldsymbol{U}_{:,1:m'}$ denotes the matrix consisting of the first m' columns of $\boldsymbol{U}$, while $\boldsymbol{V}_{1:m',:}^{\mathrm{T}}$ denotes the matrix consisting of the first m' rows of $\boldsymbol{V}^{\mathrm{T}}$, and $\boldsymbol{S}_{1:m'}$ denotes the diagonal matrix with diagonal elements as the first m' diagonal elements of $\boldsymbol{S}$.

Secondly, some of the basis functions retained in *Selection* Γ may show little similarity to the behavior of the rows of $\boldsymbol{C}$ because the corresponding elements in $\boldsymbol{U}_{:,1:m'}$ are very small. This can lead to the degraded detection sensitivity, as discussed in [3]. Therefore, another selection step is needed to remove from each row $\hat{\boldsymbol{c}}_r{}^{\mathrm{T}}$ of $\hat{\boldsymbol{C}}$ the term $u_{r,j_r} s_{j_r} \boldsymbol{v}_{j_r}^{\mathrm{T}}$, $j_r \in \{1, 2, \cdots, m'\}$ that is not relevant to the corresponding row $\boldsymbol{c}_r^{\mathrm{T}}$ of $\boldsymbol{C}$, for r from 1 to m. The term $u_{r,j_r} s_{j_r} \boldsymbol{v}_{j_r}^{\mathrm{T}}$, $j_r \in \{1, 2, \cdots, m'\}$ can be removed by setting the related element u_{r,j_r} of $\boldsymbol{U}_{:,1:m'}$ to zero. How to determine whether the term $u_{r,j_r} s_{j_r} \boldsymbol{v}_{j_r}^{\mathrm{T}}$, $j_r \in \{1, 2, \cdots, m'\}$ is relevant to the row $\boldsymbol{c}_r^{\mathrm{T}}$ or not is also discussed later in Section III-D. The adjusted $\boldsymbol{U}_{:,1:m'}$ is denoted as $\tilde{\boldsymbol{U}}_{:,1:m'}$. After this selection step called *Selection* $\Gamma\Gamma$ here, the data matrix $\hat{\boldsymbol{C}}$ is further adjusted to $\tilde{\boldsymbol{C}}$ as (11):

$$\begin{aligned} \tilde{\boldsymbol{C}} &= \begin{bmatrix} \tilde{c}_1(\tau_3+1) & \tilde{c}_1(\tau_3+2) & \cdots & \tilde{c}_1(N) \\ \tilde{c}_2(\tau_3+1) & \tilde{c}_2(\tau_3+2) & \cdots & \tilde{c}_2(N) \\ \vdots & \vdots & \cdots & \vdots \\ \tilde{c}_m(\tau_3+1) & \tilde{c}_m(\tau_3+2) & \cdots & \tilde{c}_m(N) \end{bmatrix} \\ &= \tilde{\boldsymbol{U}}_{:,1:m'} \boldsymbol{S}_{1:m'} \boldsymbol{V}_{1:m',:}^{\mathrm{T}} \end{aligned} \tag{11}$$

B. Monitoring Statistic and Detection Threshold

To detect disturbances in the context of multivariate analysis, a monitoring statistic for the i-th electrical variable x_i can be built by the variable $\tilde{c}_i$ as (12):

$$MD_i(k) = |\tilde{c}_i(k)|, \quad i = 1, 2, \cdots, m \tag{12}$$

Moreover, a system-wide monitoring statistic providing a global characterization of the group of variables with respect to the disturbances can be built as (13):

$$MD(k) = \frac{1}{m} \sum\nolimits_{i=1}^{m} |\tilde{c}_i(k)| \tag{13}$$

In order to discover whether a disturbance occurs or not, the detection thresholds of MD_i and MD with the confidence level α need to be determined. The determination procedure is similar with that of UD-FOM. Specifically, based on the training dataset, the time-series values $\{\tilde{c}_i(k)\}_{k=\tau_3+1}^{N}$ of $\tilde{c}_i$ are calculated using (11). Then, the time-series values $\{MD_i(k)\}_{k=\tau_3+1}^{N}$ of MD_i are calculated using (12), while the time-series values $\{MD(k)\}_{k=\tau_3+1}^{N}$ of MD are calculated using (13). Finally, the δ-th highest value of $\{MD_i(k)\}_{k=\tau_3+1}^{N}$ is taken as the detection threshold $MD_{i,\alpha}$ for MD_i, and the δ-th highest value of $\{MD(k)\}_{k=\tau_3+1}^{N}$ is taken as the detection threshold MD_α for MD.

After the above modelling procedure, the on-line detection can then be conducted. For the present measurement $x_i(p)$ of x_i, the present value $c_i(p)$ of c_i is calculated using (5). Then, the present value $\tilde{c}_i(p)$ of $\tilde{c}_i$ can be obtained using (14):

$$\begin{aligned} &\left[\tilde{c}_1(p)\ \tilde{c}_2(p)\ \cdots\ \tilde{c}_m(p)\right]^{\mathrm{T}} = \\ &\tilde{\boldsymbol{U}}_{:,1:m'} \left(\boldsymbol{U}_{:,1:m'}{}^{\mathrm{T}} \left[c_1(p)\ c_2(p)\ \cdots\ c_m(p)\right]^{\mathrm{T}}\right) \end{aligned} \tag{14}$$

Based on $\tilde{c}_i(p)$ calculated presently, the present value $MD_i(p)$ of MD_i can be calculated using (12), while the present value $MD(p)$ of MD can be calculated using (13). If $MD_{i,\alpha}$ is exceeded consecutively for a number of sampling time points, a disturbance is then detected in the variable x_i, while if MD_α is exceeded consecutively for a number of sampling time points, a disturbance is then detected at the system-wide level.

Thus, MD-FOMSVD which is the multivariate extension of UD-FOM has been developed. In comparison to UD-FOM which deals with the individual variable separately without considering the important correlation information, MD-FOMSVD excavates such information in a multivariate analysis manner by using SVD.

C. Procedure for MD-FOMSVD

The procedure for MD-FOMSVD is now summarized below, which consists of the off-line modelling based on the training dataset and the on-line detection.

1) The off-line modelling

① The historically recorded measurements of x_i are taken to form the training dataset $\{x_i(k)\}_{k=1}^{N}$.

② The non-Gaussianity of x_i is quantified by implementing the robust NE estimation of (1)–(3).

③ The time-series values $\{c_i(k)\}_{k=\tau_3+1}^{N}$ of c_i are calculated using (5).

④ The data matrix $\boldsymbol{C}$ is built and SVD is performed on $\boldsymbol{C}$ to obtain the matrices $\boldsymbol{U}$, $\boldsymbol{S}$ and $\boldsymbol{V}^{\mathrm{T}}$ according to (9).

⑤ The *Selection* Γ is used to obtain the matrices $\boldsymbol{U}_{:,1:m'}$, $\boldsymbol{S}_{1:m'}$ and $\boldsymbol{V}_{1:m',:}^{\mathrm{T}}$, and the matrix $\hat{\boldsymbol{C}}$ is calculated using (10).

⑥ The *Selection* $\Gamma\Gamma$ is used to obtain the matrix $\tilde{\boldsymbol{U}}_{:,1:m'}$, and the time-series values $\{\tilde{c}_i(k)\}_{k=\tau_3+1}^{N}$ are calculated using (11).

⑦ The time-series values $\{MD_i(k)\}_{k=\tau_3+1}^{N}$ are calculated using (12), while the time-series values $\{MD(k)\}_{k=\tau_3+1}^{N}$ are calculated using (13), and the detection thresholds $MD_{i,\alpha}$ and MD_α are determined.

2) The on-line detection

① The present measurement $x_i(p)$ of x_i is taken to calculate the present value $c_i(p)$ of c_i using (5).

② The present value $c_i(p)$ of c_i obtained in ① is used in (14) to calculate the present value $\tilde{c}_i(p)$ of $\tilde{c}_i$.

③ The present value $\tilde{c}_i(p)$ of $\tilde{c}_i$ obtained in ② is used in (12) and (13) to calculate the present value $MD_i(p)$ of MD_i and the present value $MD(p)$ of MD, respectively.

④ If $MD_{i,\alpha}$ is exceeded consecutively for a number of sampling time points, an alarm is given for x_i. If MD_α is exceeded consecutively for a number of sampling time points, an alarm is given at the system-wide level.

In the above procedure, the time-series values $\{c_i(k)\}_{k=\tau_3+1}^{N}$ calculated from the training dataset $\{x_i(k)\}_{k=1}^{N}$ and the present value $c_i(p)$ calculated from the present measurement $x_i(p)$ are normalized with the mean and variance of $\{c_i(k)\}_{k=\tau_3+1}^{N}$.

D. Parameter Settings for MD-FOMSVD

As stated above, MD-FOMSVD involves the *Selection* Γ and the *Selection* $\Gamma\Gamma$. The strategies for these two selection steps are now described in detail below.

1) Selection Γ*:* This selection step is to retain the first m' basis functions $\boldsymbol{v}_1^{\mathrm{T}}, \boldsymbol{v}_2^{\mathrm{T}}, \cdots, \boldsymbol{v}_{m'}^{\mathrm{T}}$ from all the m basis functions. It is performed on a multivariate level that can be quantitatively reflected by the m diagonal entries $s_1, s_2, \cdots, s_m$ of $\boldsymbol{S}$. More specifically, the squares of these m SVs are drawn up in a chart with respect to the SV numbers $1, 2, \cdots, m$, and the number after which the squares of the SVs are small enough to be neglected is chosen as the initial value of m' and is denoted as m'_{ini}. This can be easily understood as only larger SVs account for the main variance of the rows of $\boldsymbol{C}$ and give an obviously large trend in the former part of the chart. In addition to the intuitive chart, the percentage of $\sum_{j=1}^{m'} s_j^2$ in $\sum_{j=1}^{m} s_j^2$ is calculated in (15):

$$\eta = \frac{\sum_{j=1}^{m'} s_j^2}{\sum_{j=1}^{m} s_j^2} \times 100\% \tag{15}$$

where η is the cumulative percentage variance (CPV) criterion widely applied in multivariate statistical monitoring [12], [18], s_j^2 directly relates to the variance of the j-th row of $\boldsymbol{C}$ along the direction of the basis function $\boldsymbol{v}_j^{\mathrm{T}}$. Accordingly, $\sum_{j=1}^{m} s_j^2$ relates to the total variance of $\boldsymbol{C}$ along all the basis functions. Usually, $\eta \geq 90\%$ is sufficient to signify that most variance of the original data is well captured [12], [19]. Thus, m'_{ini} needs to be further evaluated by (15). If m'_{ini} makes the condition $\eta \geq 90\%$ satisfied, it can then be used as an appropriate value for m'; otherwise, $m'_{\mathrm{ini}}+1, m'_{\mathrm{ini}}+2, \cdots, m$ need to be evaluated in sequence until the condition is met.

2) Selection $\Gamma\Gamma$*:* After an appropriate value is chosen for m', this selection step is to determine whether the term $u_{r,j_r} s_{j_r} \boldsymbol{v}_{j_r}^{\mathrm{T}}$, $j_r \in \{1, 2, \cdots, m'\}$ is relevant to the r-th row $\boldsymbol{c}_r^{\mathrm{T}} = \sum_{j=1}^{m} u_{r,j} s_j \boldsymbol{v}_j^{\mathrm{T}}$ of $\boldsymbol{C}$ in (9) or not, and to remove the terms that have little relevance to $\boldsymbol{c}_r^{\mathrm{T}}$. Different from the *Selection* Γ performed on a multivariate level, the *Selection* $\Gamma\Gamma$ is carried out on an individual level for the selection of relevant terms. Specifically, the following index is constructed to evaluate the m' terms $u_{r,1} s_1 \boldsymbol{v}_1^{\mathrm{T}}, u_{r,2} s_2 \boldsymbol{v}_2^{\mathrm{T}}, \cdots, u_{r,m'} s_{m'} \boldsymbol{v}_{m'}^{\mathrm{T}}$:

$$\mu_{r,j_r} = \frac{u_{r,j_r}^2}{\sum_{j=1}^{m'} u_{r,j}^2} \times 100\%,\ j_r \in \{1, 2, \cdots, m'\} \tag{16}$$

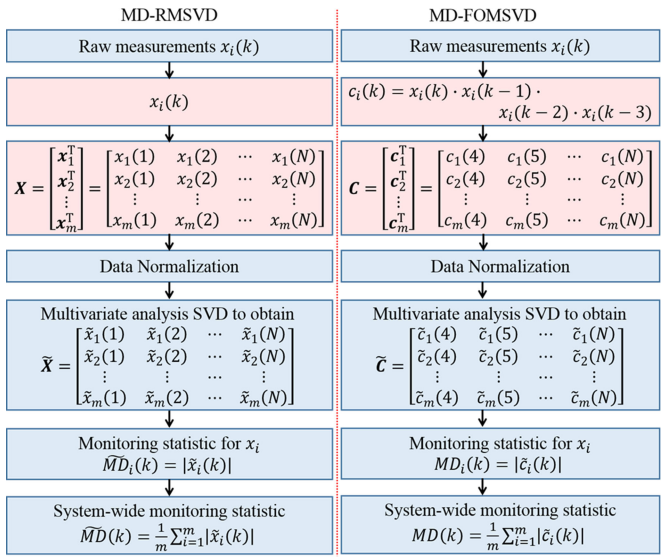

Fig. 1. The schematic diagram describing MD-RMSVD and MD-FOMSVD

where μ_{r,j_r} measures the relevance of $s_{j_r} \boldsymbol{v}_{j_r}^{\mathrm{T}}$ to $\boldsymbol{c}_r^{\mathrm{T}}$. If μ_{r,j_r} is far smaller than the average level $\frac{1}{m'} \times 100\%$, typically, smaller than $\frac{1}{\mathrm{m}m'} \times 100\%$, $u_{r,j_r} s_{j_r} \boldsymbol{v}_{j_r}^{\mathrm{T}}$ is removed since $s_{j_r} \boldsymbol{v}_{j_r}^{\mathrm{T}}$ has little relevance to $\boldsymbol{c}_r^{\mathrm{T}}$. As a parallel with the well-known principal component analysis (PCA) [9], u_{r,j_r} corresponds to the *r*-th element of the j_r-th loading vector in PCA, and $s_{j_r} \boldsymbol{v}_{j_r}^{\mathrm{T}}$ corresponds to the j_r-th score. In this sense, this step only retains the scores making obvious contribution to the rows of $\boldsymbol{C}$ and removes the ones making little contribution.

IV. Case Study

In this section, UD-FOM and MD-FOMSVD are evaluated using the measurements from the European power system. Besides, a multivariate detection method based on Raw Measurements (RM) and SVD, referred to as MD-RMSVD here, is taken for the comparison with MD-FOMSVD. Referring to Fig. 1, the difference between the two multivariate detection methods is that MD-FOMSVD conducts SVD on fourth order statistics obtained by FOM whereas MD-RMSVD directly conducts SVD on raw measurements.

The electrical variables for monitoring are described in Table I, and 128000 measurements are recorded on each variable with the 10 Hz sampling frequency that are plotted in Fig. 2. The first 8000 measurements are taken to form training data since there is nothing abnormal in them to worry an operator in the control room and they reflect the characteristic of ambient operation, whereas the remaining measurements are taken to form testing data since they contain a disturbance that is highlighted in the rectangles. The observations from the supplier of the data are that this disturbance may be caused by a switching operation and the staircase effect in the restoration of the ambient operation towards the right hand side of the rectangles is due to the automated response of transformer tap changer to the occurring disturbance. After this disturbance, there is no other obvious abnormality in the subsequent operation condition. The

TABLE I
THE VARIABLES FOR MONITORING AND THE ESTIMATES OF NE

Variable	Description	Estimate of NE
x_1	Voltage amplitude in substation 1	$NE_1 = 0.0872$
x_2	Current amplitude in substation 1	$NE_2 = 0.0911$
x_3	Active power in substation 1	$NE_3 = 0.0872$
x_4	Apparent power in substation 1	$NE_4 = 0.0894$
x_5	Reactive power in substation 1	$NE_5 = 0.3675$
x_6	Voltage amplitude in substation 2	$NE_6 = 0.3421$
x_7	Current amplitude in substation 2	$NE_7 = 0.0900$
x_8	Active power in substation 2	$NE_8 = 0.0875$
x_9	Apparent power in substation 2	$NE_9 = 0.0882$
x_{10}	Reactive power in substation 2	$NE_{10} = 0.2403$

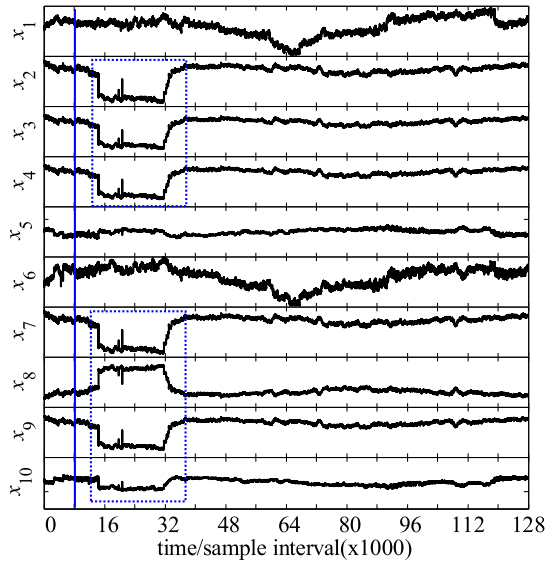

Fig. 2. The electrical measurements.

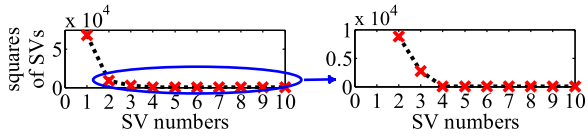

Fig. 3. The squares of SVs with respect to the SV numbers for MD-FOMSVD.

expectation here is that the disturbance in the rectangles be effectively detected and very few continuous alarms be triggered when there is no obvious disturbance occurring.

A. The Off-Line Modelling

The non-Gaussianity of each variable in Table I is quantified by estimating the NE based on the training data and the results are also listed in Table I. The table shows that each variable has a non-Gaussian distribution, because the related estimate of NE is far larger than the threshold 10^{-3}. Then, such non-Gaussian information is explored using the fourth order statistic of (5).

For MD-FOMSVD, the *Selection* Γ is used to determine the number m' of retained basis functions. Firstly, the relationship between the squares of the SVs $s_1, s_2, \cdots, s_{10}$ and the SV numbers $1, 2, \cdots, 10$ is plotted in Fig. 3. From Fig. 3, s_1 is much larger than s_2 and s_3, while s_2 and s_3 are still much larger than the remaining SVs. Thus, s_2 and s_3 are not negligible. Another reason for retaining s_2 and s_3 is due to the multivariate correlation. As shown in Fig. 2, there exist three groups of variables $(x_1, x_6; x_5, x_{10}; x_2, x_3, x_4, x_7, x_8, x_9)$ and the variables in each group are strongly correlated. Actually, the number of the retained SVs directly relates with the number of such groups. If only s_1 is chosen, it cannot adequately reflect the characteristics of the whole groups of variables. Accordingly, the initial value m'_{ini} of m' is set to 3. Then, substituting $m' = m'_{\text{ini}}$ into (15), the CPV index η is calculated as 99.87% much larger than 90%. This means $m' = m'_{\text{ini}}$ could be taken for the use in MD-FOMSVD.

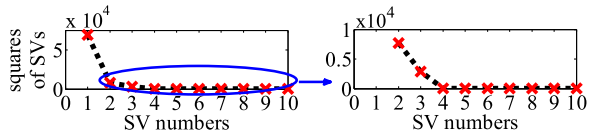

Fig. 4. The squares of SVs with respect to the SV numbers for MD-RMSVD.

TABLE II
THE CALCULATION RESULTS OF $\mu_{r,j_r}, j_r \in \{1,2,3\}$ (%) FOR r FROM 1 TO 10 FOR MD-FOMSVD AND MD-RMSVD

	MD-FOMSVD			MD-RMSVD		
r	$\mu_{r,1}$	$\mu_{r,2}$	$\mu_{r,3}$	$\mu_{r,1}$	$\mu_{r,2}$	$\mu_{r,3}$
1	9.81	0.06	90.13	9.30	0.14	90.56
2	68.16	24.51	7.33	66.52	28.81	4.67
3	69.86	18.40	11.74	68.40	22.23	9.37
4	64.74	24.22	11.04	63.25	29.02	7.73
5	20.23	69.51	10.26	24.29	65.63	10.08
6	48.78	22.58	28.64	46.78	31.83	21.39
7	72.54	19.71	7.75	71.06	23.19	5.75
8	69.68	18.44	11.88	68.26	22.31	9.43
9	64.93	23.73	11.34	63.40	28.53	8.07
10	10.64	87.11	2.25	12.01	80.46	7.53

Similarly, for MD-RMSVD, the *Selection* Γ is also used to determine the number m' of retained basis functions. The relationship between the squares of the SVs and the SV numbers is plotted in Fig. 4. From Fig. 4, s_1, s_2 and s_3 are also much larger than the remaining SVs. Thus, the initial value m'_{ini} of m' is also set to 3. Substituting $m' = m'_{\text{ini}}$ into (15), the CPV index η is calculated as 99.99% also much larger than 90%, meaning $m' = m'_{\text{ini}}$ is a proper choice for MD-RMSVD.

For MD-FOMSVD, after the determination of m', the *Selection* $\Gamma\Gamma$ is used to determine whether the term $u_{r,j_r} s_{j_r} \boldsymbol{v}_{j_r}^{\mathrm{T}}$, $j_r \in \{1,2,3\}$ is relevant to the row $\boldsymbol{c}_r^{\mathrm{T}}$ or not. For r from 1 to 10, μ_{r,j_r} is calculated by (16) and the results in Table II show that $\mu_{1,2}$ (0.06%) and $\mu_{10,3}$ (2.25%) are smaller than the level $\frac{1}{\mathrm{m}m'} \times 100\% = 3.33\%$. This means that $s_2 \boldsymbol{v}_2^{\mathrm{T}}$ has little relevance to $\boldsymbol{c}_1^{\mathrm{T}}$, while $s_3 \boldsymbol{v}_3^{\mathrm{T}}$ has little relevance to $\boldsymbol{c}_{10}^{\mathrm{T}}$. Accordingly, $u_{1,2} s_2 \boldsymbol{v}_2^{\mathrm{T}}$ and $u_{10,3} s_3 \boldsymbol{v}_3^{\mathrm{T}}$ are removed by setting $u_{1,2}$ and $u_{10,3}$ to zero, and the matrices $\boldsymbol{U}_{:,1:m'}$ and $\tilde{\boldsymbol{U}}_{:,1:m'}$ for the on-line use in (14) are then determined.

Similarly, for MD-RMSVD, after the determination of m', the *Selection* $\Gamma\Gamma$ is also used to determine whether the term

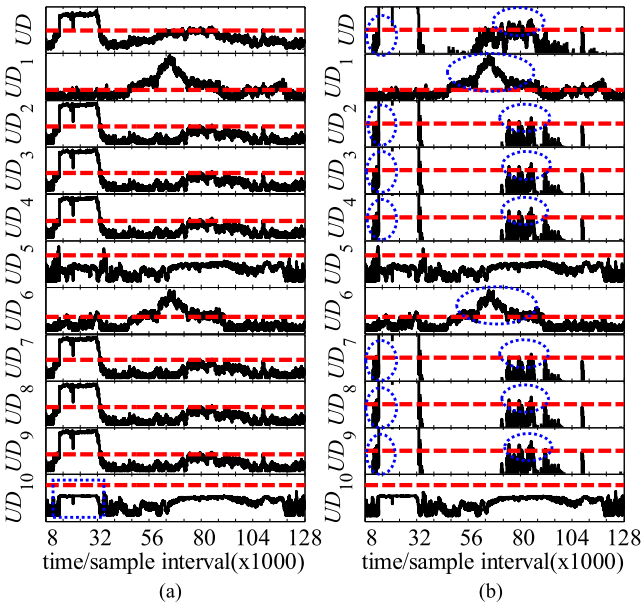

Fig. 5. The detection results of UD-FOM on the testing data, showing the time-series values of monitoring statistics (solid lines) and the detection thresholds (dashed lines). (a) The complete charts. (b) The charts with UD, UD_2-UD_4, UD_7-UD_9 zoomed in.

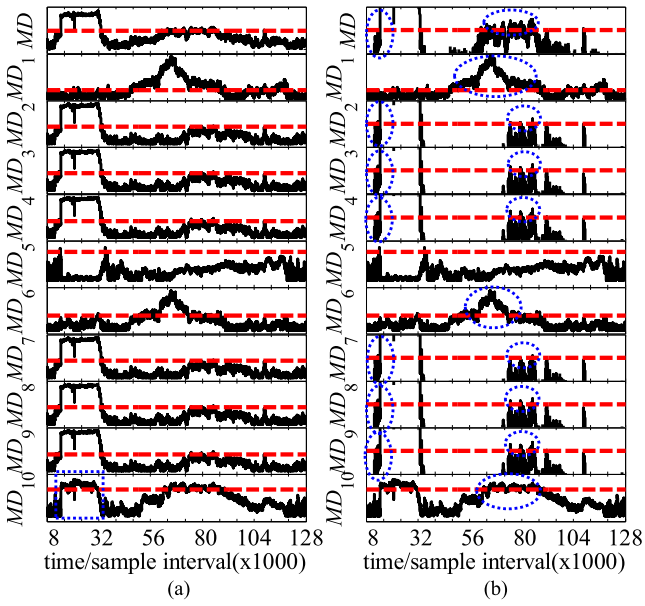

Fig. 6. The detection results of MD-FOMSVD on the testing data, showing the time-series values of monitoring statistics (solid lines) and the detection thresholds (dashed lines). (a) The complete charts. (b) The charts with MD, MD_2-MD_4, MD_7-MD_9 zoomed in.

$u_{r,j_r}\, s_{j_r}\, \boldsymbol{v}_{j_r}^{\mathrm{T}}$, $j_r \in \{1,2,3\}$ is relevant to the row $\boldsymbol{x}_r^{\mathrm{T}}$ or not. For r from 1 to 10, μ_{r,j_r} is also calculated by (16) and the results in Table II show that $\mu_{1,2}$ (0.14%) is smaller than the level $\frac{1}{\mathrm{m}m'} \times 100\% = 3.33\%$. This suggests that $s_2 \boldsymbol{v}_2^{\mathrm{T}}$ has little relevance to $\boldsymbol{x}_1^{\mathrm{T}}$. Thus, $u_{1,2} s_2 \boldsymbol{v}_2^{\mathrm{T}}$ is removed by setting $u_{1,2}$ to zero, and the matrices $\boldsymbol{U}_{:,1:m'}$ and $\tilde{\boldsymbol{U}}_{:,1:m'}$ for the on-line use are obtained.

B. The On-Line Detection

1) Comparison of Detection Sensitivity: The detection results of UD-FOM and MD-FOMSVD on the testing data are shown in Figs. 5–6. The detection thresholds are calculated with the confidence level $\alpha = 99\%$. In Figs. 5–6, the charts on the left side are plotted completely, while the charts on the right side show UD, UD_2-UD_4, UD_7-UD_9, MD, MD_2-MD_4, and MD_7-MD_9 zoomed in so that excursions above the detection thresholds can be seen more clearly. Observing from Figs. 5–6, when the disturbance in the rectangles of Fig. 2 occurs, both UD of UD-FOM and MD of MD-FOMSVD react sharply and give the definite indication. This makes good significance for application, because UD and MD are system-wide monitoring statistics and the practical implementation of a detection method in a control room usually takes the form of a traffic light with green or red indicators for the overall system state. Furthermore, the disturbance effect on each variable can be checked by the monitoring statistics UD_1, $UD_2, \cdots, UD_{10}$ of UD-FOM and MD_1, $MD_2, \cdots, MD_{10}$ of MD-FOMSVD. For $i = 2, 3, 4, 7, 8, 9$, the disturbance is well detected in x_i by the related UD_i and MD_i. This is the expected result, since the prominent effect of the disturbance on $x_2, x_3, x_4, x_7, x_8, x_9$ can be seen in their measurements where a distinct jump happens. In contrast, the disturbance is not detected in x_1, x_5, x_6 by the related UD_1, UD_5, UD_6 and MD_1, MD_5, MD_6, which is in line with the visual inspection that x_1, x_5, x_6 are not affected by the disturbance. Moreover, as exhibited in the rectangles of Figs. 5–6, the disturbance in x_{10} is not detected by UD_{10}, but it is well detected by MD_{10}.

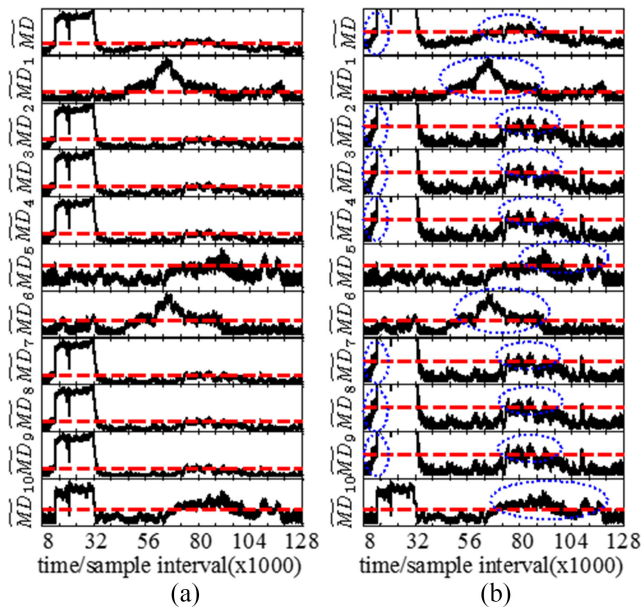

Fig. 7. The detection results of MD-RMSVD on the testing data, showing the time-series values of monitoring statistics (solid lines) and the detection thresholds (dashed lines). (a) The complete charts. (b) The charts with $\widetilde{MD}$, $\widetilde{MD}_2$-$\widetilde{MD}_4$, $\widetilde{MD}_7$-$\widetilde{MD}_9$ zoomed in.

Fig. 7 shows the detection results of MD-RMSVD on the testing data. Again, the charts on the left side are plotted com-

TABLE III
THE NUMBER OF ALARMS WHEN THE SYSTEM IS IN THE AMBIENT CONDITION

UD-FOM		MD-FOMSVD		MD-RMSVD	
UD	5211	MD	7023	$\widetilde{MD}$	16387
UD_1	52273	MD_1	52552	$\widetilde{MD}_1$	51928
UD_2	109	MD_2	75	$\widetilde{MD}_2$	7104
UD_3	204	MD_3	192	$\widetilde{MD}_3$	7684
UD_4	842	MD_4	683	$\widetilde{MD}_4$	8982
UD_5	108	MD_5	19	$\widetilde{MD}_5$	10210
UD_6	30955	MD_6	21694	$\widetilde{MD}_6$	30808
UD_7	37	MD_7	39	$\widetilde{MD}_7$	6138
UD_8	264	MD_8	205	$\widetilde{MD}_8$	7743
UD_9	799	MD_9	691	$\widetilde{MD}_9$	9023
UD_{10}	0	MD_{10}	20368	$\widetilde{MD}_{10}$	31599

pletely, while the charts on the right side show $\widetilde{MD}$, $\widetilde{MD}_2$-$\widetilde{MD}_4$, $\widetilde{MD}_7$-$\widetilde{MD}_9$ zoomed in. It can be seen that the disturbance is detected in x_i by $\widetilde{MD}_i$ of MD-RMSVD and MD_i of MD-FOMSVD for $i = 2, 3, 4, 7, 8, 9, 10$. Meanwhile, the disturbance is also detected by $\widetilde{MD}$ and MD at the system-wide level. Thus, both MD-RMSVD and MD-FOMSVD achieve effective detection of the disturbance in the rectangles of Fig. 2.

2) Comparison of Detection Reliability: Table III lists the number of alarms when the system is in the ambient condition. The fewer alarms a method gives in the ambient condition, the better reliability it has. It can be seen from Table III as well as inside the ellipses of Figs. 5–7 that MD-RMSVD generally gives more false alarms than MD-FOMSVD and UD-FOM. Here, a point worth mentioning is that the three methods all give alarms for x_1 and x_6. This is because the measurements of x_1 and x_6 have significant deviations after about the 48000th sampling time point, as can be observed from Fig. 2. There are measurements of voltage at two substations. The supplier of the data did not highlight any event during this time, however the results suggest that all three methods have detected an unusual voltage deviation.

According to the above results, the detection sensitivity of MD-FOMSVD is more satisfactory than that of UD-FOM. This is due to the reason that MD-FOMSVD further conducts the multivariate analysis using SVD on the basis of UD-FOM to jointly explore the presence of the same disturbance in different variables. In comparison to MD-RMSVD, MD-FOMSVD behaves more reliably under the ambient condition, which can be attributed to the use of fourth order statistics obtained by FOM instead of directly using raw measurements.

Besides the above potential use as an on-line method running in the background for automated detection of system disturbances, a further potential use of MD-FOMSVD is to configure and tune the event-driven data recorders that are applied in transmission systems for storing measurements following a disturbance. In practice, it can be a challenge for operators to identify suitable sets of measurements from normal ambient operation. Once the engineering effort has been made to train MD-FOMSVD with a reliable dataset from a known episode of ambient operation, the role of MD-FOMSVD would be to provide the reassurance that other episodes also do not have any disturbance and therefore are suitable for the calibration of event recorders.

V. CONCLUSION

A new method MD-FOMSVD has been proposed to detect power system disturbances. The contribution is two-fold. Firstly, an effective scheme to obtain the fourth order statistics that can account for the significant non-Gaussian information in the measurements of each electrical variable has been developed based on FOM. Secondly, a multivariate analysis strategy that can jointly explore the presence of the same disturbance in different electrical variables has been presented by handling the obtained fourth order statistics using SVD. Thus, a unified framework combining the high order analysis with the multivariate analysis has been built to allow the full play of their respective advantages. The application results on the measurements of the European power system have shown that MD-FOMSVD achieves improved performance in terms of the detection reliability and sensitivity.

ACKNOWLEDGMENT

The authors would like to thank Dr. Mats Larsson of ABB Corporate Research Center, Baden-Dättwil, Switzerland, for providing data and power system insights to support this paper.

REFERENCES

[1] O. Samuelsson, M. Hemmingsson, A. H. Nielsen, K. O. H. Pederson, and J. Ramaussen, "Monitoring of power system events at transmission and distribution level," *IEEE Trans. Power Syst.*, vol. 21, no. 2, pp. 1007–1008, May 2006.

[2] I. M. Cecílio, J. R. Ottewill, J. Pretlove, and N. F. Thornhill, "Nearest neighbors method for detecting transient disturbances in process and electromechanical systems," *J. Process Control*, vol. 24, no. 9, pp. 1382–1393, Sep. 2014.

[3] I. M. Cecílio, J. R. Ottewill, H. Fretheim, and N. F. Thornhill, "Multivariate detection of transient disturbances for uni- and multirate systems," *IEEE Trans. Control Syst. Technol.*, vol. 23, no. 4, pp. 1477–1493, Jul. 2015.

[4] Q. Huang, L. Shao, and N. Li, "Dynamic detection of transmission line outages using hidden Markov models," *IEEE Trans. Power Syst.*, vol. 31, no. 3, pp. 2026–2033, May 2016.

[5] F. B. Costa, "Fault-induced transient detection based on real-time analysis of the wavelet coefficient energy," *IEEE Trans. Power Del.*, vol. 29, no. 1, pp. 140–153, Feb. 2014.

[6] A. Farzanehrafat and N. R. Watson, "Power quality state estimator for smart distribution grids," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2183–2191, Aug. 2013.

[7] Z. Liu, Q. Hu, Y. Cui, and Q. Zhang, "A new detection approach of transient disturbances combining wavelet packet and Tsallis entropy," *Neurocomputing*, vol. 142, pp. 393–407, Oct. 2014.

[8] L. Xie, Y. Chen, and P. R. Kumar, "Dimensionality reduction of synchrophasor data for early event detection: Linearized analysis," *IEEE Trans. Power Syst.*, vol. 29, no. 6, pp. 2784–2794, Nov. 2014.

[9] E. Barocio, B. C. Pal, D. Fabozzi, and N. F. Thornhill, "Detection and visualization of power system disturbances using principal component analysis," in *Proc. 2013 IREP Symp. Bulk Power Syst. Dyn. Control IX*, Rethymnon, Greece, Aug. 2013, pp. 1–10.

[10] E. Barocio, B. C. Pal, N. F. Thornhill, and A. R. Messina, "A dynamic mode decomposition framework for global power system oscillation analysis," *IEEE Trans. Power Syst.*, vol. 30, no. 6, pp. 2902–2912, Nov. 2015.

[11] J. Turunen *et al.*, "Comparison of three electromechanical oscillation damping estimation methods," *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 2398–2407, Nov. 2011.

[12] L. Cai, X. Tian, and S. Chen, "A process monitoring method based on noisy independent component analysis," *Neurocomputing*, vol. 127, pp. 231–246, Mar. 2014.
[13] L. Cai and X. Tian, "A new fault detection method for non-Gaussian process based on robust independent component analysis," *Process Safety Environ. Protection*, vol. 92, no. 6, pp. 645–658, Nov. 2014.
[14] B. Porat and B. Friedlander, "Direction finding algorithms based on high-order statistics," *IEEE Trans. Signal Process.*, vol. 39, no. 9, pp. 2016–2024, Sep. 1991.
[15] M. A. A. S. Choudhury, S. L. Shah, and N. F. Thornhill, "Diagnosis of poor control-loop performance using higher-order statistics," *Automatica*, vol. 40, no. 10, pp. 1719–1728, Oct. 2004.
[16] E. G. Learned-Miller and J. W. Fisher III, "ICA using spacings estimates of entropy," *J. Mach. Learn. Res.*, vol. 4, pp. 1271–1295, Dec. 2003.
[17] Y. Wang, J. Fan, and Y. Yao, "Online monitoring of multivariate processes using higher-order cumulants analysis," *Ind. Eng. Chem. Res.*, vol. 53, no. 11, pp. 4328–4338, Feb. 2014.
[18] W. Zhu *et al.*, "A novel KICA–PCA fault detection model for condition process of hydroelectric generating unit," *Measurement*, vol. 58, pp. 197–206, Dec. 2014.
[19] G. Li, S. Qin, and D. Zhou, "A new method of dynamic latent-variable modeling for process monitoring," *IEEE Trans. Ind. Electron.*, vol. 61, no. 11, pp. 6438–6445, Nov. 2014.

Lianfang Cai received the B.Eng. and Ph.D. degrees from the China University of Petroleum, Qingdao, China, in 2009 and 2014, respectively. He is currently a Postdoctoral Research Associate with Imperial College London, London, U.K. His research interests include data-driven power system monitoring, modeling of power systems with energy storage, and multivariate statistics.

Nina F. Thornhill (SM'93) received the B.A. degree in physics from Oxford University, Oxford, U.K., in 1976, the M.Sc. degree from Imperial College London, London, U.K., and the Ph.D. degree from University College London, London.

She is a Professor in the Department of Chemical Engineering, Imperial College London, where she holds the ABB Chair of Process Automation.

Bikash C. Pal (M'00–SM'02–F'13) is a Professor of Power Systems at Imperial College London. He is research active in power system stability, control and computation. He has graduated 15 Ph.D.s and published 60 technical papers in IEEE Transactions and IET journals. He has co-authored two books and two awards winning IEEE Task Force/Working Group reports. He is the Editor-in-Chief of the IEEE TRANSACTIONS ON SUSTAINABLE ENERGY and Fellow of IEEE for his contribution to power system stability and control.