Data-Driven Islanding Detection Using a Principal Subspace of Voltage Angle Differences

Tin Rabuzin, *Graduate Student Member, IEEE*, and Lars Nordström, *Senior Member, IEEE*

***Abstract*—The likelihood of an unintentional power system islanding is increased in systems with significant penetration of distributed generation. To mitigate the adverse effects of islanding, a quick and reliable islanding detection method is needed. This paper first analyzes covariance matrices of a linearized power system model, and relates them to the principal component analysis of experimentally obtained covariance matrices. Additionally, a new model-independent islanding detection method is proposed that uses measurements of voltage angle differences between multiple locations in the system. The angle differences are first preprocessed to remove the effects of nonstationarity. Thereafter, a probabilistic model of principal component analysis is trained using the acquired measurements. The principal and residual spaces extracted from the measurements are used to discriminate between islanding and other events in the system. The applicability of the proposed method is demonstrated by using real measurements gathered from several locations in a transmission grid.**

***Index Terms*—Bayes methods, covariance matrices, islanding, phase angle differences, principal component analysis, wide area monitoring.**

I. Introduction

SUSTAINABLE production of electrical energy is one of the cornerstones of many countries' environmental agendas. For instance, the European Union (EU) has put forward its 2020 strategy that aims to increase the share of renewable energy sources in Europe to 20% by the year 2020. Eurostat reports that the EU is on track to reach the goals of the strategy with some member countries surpassing their respective goals [1]. In particular, Sweden generated 54% of its electric energy in 2017 from renewable energy sources. The environmental targets, in addition to the technological advancements and economic opportunities, have facilitated the integration of renewable distributed generation (DG) in electric power systems [2].

However, significant penetration levels of DG can introduce technical challenges for the system operators which, if not appropriately addressed, impact the safe and stable operation of power systems [3]. For example, DG can increase the likelihood of sustained islanded operation of a power system. A power system island is created when an area of a power system is electrically disconnected from the rest of the system, yet it continues to be energized by local energy sources. If an islanding event remains undetected, it can endanger line workers, impair the effectiveness of voltage and frequency controls, cause out-of-step reclosing when automatic reclosers operate and reduce power quality [4]. Therefore, it is important to quickly and reliably detect an islanding event to be able to take appropriate corrective actions such as disconnecting DG units in the island or, alternatively, ensure a safe transition to an islanded mode of operation. A recent event in Great Britain's power system exemplifies the importance of reliable islanding detection schemes in systems containing distributed generation. A lightning strike triggered a sequence of events leading to a loss of power supply to approximately 1 million customers [5]. One of the events which contributed to the loss of supply was the operation of islanding protection schemes installed at distributed generators, namely vector shift and rate of change of frequency (ROCOF) protection.

Literature proposes various islanding detection methods [6]. Generally, the methods can be broadly classified into local and remote. Local methods base their operating principle on a locally observed system's response to either intentionally or unintentionally introduced system perturbations. On the other hand, the advent of wide-area monitoring systems (WAMSs) has enabled the development of remote islanding detection methods that rely on the availability of time-synchronized measurements from phasor measurement units (PMUs) at different locations in the network. It is generally expected that the PMU measurements from remote locations can make islanding detection more robust. Moreover, the report in [7] states that the synchrophasor-based islanding detection can provide information to the operators which are not available from SCADA systems, it can aid decision making and assist with resynchronization and monitoring of the electrical island. Examples of such methods are based on comparison of ROCOF [8], accumulated phase angle drift [9], frequency or angle difference [10] and slip frequency and acceleration [11]. However, all of the aforementioned methods depend on threshold settings which are hard to optimize for timely and reliable islanding detection.

An alternative to setting hard thresholds can be found in the analysis of information useful for islanding detection extracted from historical PMU measurements. The abundance of measured data produced by the WAMSs has facilitated research

Manuscript received February 15, 2020; revised June 23, 2020, October 9, 2020, and January 4, 2021; accepted March 23, 2021. Date of publication March 29, 2021; date of current version August 23, 2021. This work was supported by the SamspEl program of the Swedish Energy Agency. Paper no. TSG-00231-2020. *(Corresponding author: Tin Rabuzin.)*

The authors are with the Division of Electric Power and Energy Systems, School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, 114 28 Stockholm, Sweden (e-mail: rabuzin@kth.se).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TSG.2021.3069287.

Digital Object Identifier 10.1109/TSG.2021.3069287

in data-driven applications for power systems [12]. The central data analysis tool of this paper is principal component analysis (PCA). PCA is a dimensionality reduction technique introduced nearly a century ago [13]. It is well known within various application areas such as face recognition, process monitoring, and data compression. Dimensionality reduction has also been exploited in power system applications. For example, the authors of [14] provide a linear system analysis view on measurement analysis by using PCA and demonstrate its applicability for online event detection, although the proposed approach does not provide information on which specific event occurred. Similarly, PCA along with the k-Nearest Neighbor algorithm has been used for detection and localization of power system events in real-time [15]. In [16], the authors propose to identify events based on a comparison of the data subspace of a current event obtained from PCA to the elements of a dictionary of subspaces corresponding to known past events. Other applications of PCA include voltage stability assessment [17], methods for alleviating the bad data problem [18]–[20], PMU data compression [21] and coherency identification [22].

Specifically for islanding detection, PCA and PMU measurements of frequency have been utilized in [23]. The authors of [23] show that PMU measurements of frequency during normal operation are centered around a mean point at the nominal frequency and they assume that no significant deviations of frequency can occur between two points of measurements during non-islanding events. They also show that statistical limits for frequency deviations can be obtained using PCA and measurements collected during normal operation. If frequency measurements start to deviate from each other and surpass the statistical limit, the proposed method concludes that an islanding event occurred.

To address the time-varying nature of power systems and lower the number of false alarms, the work in [24] proposes an improvement of the aforementioned method by recursive updates of the correlation matrix used by PCA. Furthermore, since frequency differences can be small in cases of balanced islanded systems, frequency-based methods suffer from a non-detection zone (NDZ). To reduce the NDZ, the same group of authors in [25] propose a refined method involving a moving window kernel PCA (KPCA) that models the behavior of nonstationary signals of voltage angle differences. The voltage angle differences are mapped to a higher dimensional space where a PCA is again applied to detect islanding. This approach requires a pre-defined kernel (a mapping function) that removes the physical interpretability of PCA. The method is computationally expensive since the KPCA model needs to be updated with each newly reported measurement. Similarly to the moving window approach of [25], the nonstationarity issue of angle differences was addressed by iteratively updating the probabilistic PCA model of angle differences in [26].

A. Scope and Contributions

The contributions of this paper are the following. First, the paper derives the covariance matrices of a linearized power system model subject to random load variations and relates those to the system's controllability Gramian. It also claims that an islanding event can be characterized as a departure of the measured voltage angle differences from the principal subspace obtained from historical PMU measurements using PCA.

The presented derivation of the covariance matrices is valid only for stationary signals in the proximity of an equilibrium. The PMU measurements of angle differences are typically not stationary in their means, and thus we propose to preprocess the angle differences by using a high-pass finite impulse response (FIR) filter prior to building the PCA model of the measurements. Filtering data in such a way removes the long term dynamics observed in the angle differences which makes them stationary during normal operation. It is also demonstrated that the filtered measurements follow a normal distribution which makes inference using newly reported PMU measurements possible.

Finally, based on the first two contributions, an islanding detection method independent of the power system model is proposed that is based on the application of a probabilistic approach to PCA. The detection method uses a machine learning algorithm and historical measurements of voltage angles to extract the principal subspace of the data which is later used in real-time to detect islanding. The learning algorithm requires no human interaction nor knowledge of the system states and topology. The performance of the method is demonstrated using PMU measurements taken from a European transmission system.

II. PCA and Covariance Matrices

Consider a bulk power system with PMUs installed at a subset of the buses in a network. This section discusses the possibility of detecting an islanding event based on the PCA of measurements of voltage angle differences in a such PMU-equipped system. First, however, PCA as a mathematical tool for data analysis is introduced.

A. PCA—Classical Interpretation

PCA can be interpreted as a transformation of originally correlated set of variables into a new set of uncorrelated variables (principal component scores). In essence, PCA does that by extracting the directions of the highest variance observed in the data.

Consider a matrix of N multivariate observations, $Y \in \mathbb{R}^{N\times m}$. Each row vector $\boldsymbol{y}[k] \in \mathbb{R}^{m\times 1}$ in this matrix corresponds to one observation at time $t = k/f_r$ and f_r is the sampling rate. In the case of PMU measurements in power system applications, f_r corresponds to PMU reporting rate. Singular value decomposition (SVD) allows one to express the data matrix as:

$$Y = USV^T, \tag{1}$$

where S is a diagonal matrix of decreasingly ordered singular values, and the columns of the matrices U and V are left and right singular vectors of Y, respectively. A covariance matrix

of the measurements, assuming that the measurements are zero mean, can then be expressed as:

$$\Sigma_y = \frac{1}{N-1} Y^T Y = \frac{1}{N-1} V S^2 V^T, \tag{2}$$

since $U^T U = I$. The matrix V is now treated as the matrix of eigenvectors of Σ_y. Graphically, Σ_y can be depicted by forming an m-dimensional ellipsoid that encircles the data using the eigenvectors in V scaled by the eigenvalues in S as the semi-axes of the ellipsoid. These semi-axes indicate the directions in the space where the data has the highest variance. We refer to the space spanned by the direction of the highest variance as the principal subspace.

Typically, most of the data's variance can be captured by retaining only the directions of the highest variance, i.e., it is considered that the data lies on a low-dimensional manifold where the most information is contained. Following this principle, the data matrix can be decomposed as follows:

$$Y = \underbrace{U_r S_r}_{T} V_r^T + E, \tag{3}$$

where the subscripts denote that only $r < m$ columns were retained from the original matrices. The matrix E contains errors caused by the truncation and T is the truncated matrix of principal component scores.

B. Covariance Matrices of a State Space Model

Dynamics of a multi-machine power system can be represented by a set of nonlinear differential-algebraic equations (DAEs) of the following form:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{z}(t)), \tag{4a}$$

$$0 = \boldsymbol{g}(\boldsymbol{x}(t), \boldsymbol{z}(t)). \tag{4b}$$

DAEs in (4) can be linearized around an equilibrium to obtain the following linear DAEs

$$\Delta\dot{\boldsymbol{x}}(t) = f_x \Delta\boldsymbol{x}(t) + f_h \Delta\boldsymbol{z}(t), \tag{5a}$$

$$0 = g_x \Delta\boldsymbol{x}(t) + g_z \Delta\boldsymbol{z}(t) + g_u \Delta\boldsymbol{p}_L(t), \tag{5b}$$

where f_x, f_h, g_x, and g_z are the corresponding Jacobian matrices. An additional term $g_u \Delta\boldsymbol{p}_L(t)$ was added to the algebraic set of equations in order to be able to analyze the impact of random load variations on the system dynamics. The algebraic variables in (5) can be eliminated by Kron reduction which results in the following continuous linear time-invariant (LTI) model:

$$\Delta\dot{\boldsymbol{x}}(t) = A_c \Delta\boldsymbol{x}(t) + B_c \Delta\boldsymbol{p}_L(t), \tag{6a}$$

$$\Delta\boldsymbol{y}(t) = C_c \Delta\boldsymbol{x}(t), \tag{6b}$$

where A_c, B_c, and C_c are the matrices of the state space model. The state space model in (6) can be discretized which yields the following discrete LTI system:

$$\Delta\boldsymbol{x}[k+1] = A\Delta\boldsymbol{x}[k] + B\Delta\boldsymbol{p}_L[k], \tag{7a}$$

$$\Delta\boldsymbol{y}[k] = C\Delta\boldsymbol{x}[k] + \boldsymbol{v}[k], \tag{7b}$$

where A, B, and C are the discrete versions of matrices A_c, B_c, and C_c, respectively. We assume that the vector of random load variations is normally distributed according to $\Delta\boldsymbol{p}_L[k] \sim \mathcal{N}(0, \Sigma_u)$. A random variable is included in the output equation in (7b) to model measurement noise which we assume to be normally distributed, i.e., $\boldsymbol{v}[k] \sim \mathcal{N}(0, \Sigma_v)$. Furthermore, it is assumed that the random variables $\boldsymbol{x}[k]$, $\Delta\boldsymbol{p}_L[k]$ and $\boldsymbol{v}[k]$ are independent.

Now, we are interested in properties of the covariance matrix of the measurements Δy. First, the steady-state covariance matrix of state variables Σ_x can be obtained from the following discrete Lyapunov equation:

$$\Sigma_x = \mathbf{E}\big[\Delta\boldsymbol{x}[k+1]\Delta\boldsymbol{x}[k+1]^T\big] \tag{8a}$$

$$= A\Sigma_x A^T + B\Sigma_u B^T, \tag{8b}$$

where $\mathbf{E}[\bullet]$ is the expectation operator. Then, the steady-state covariance matrix of the measurements is obtained from:

$$\Sigma_y = \mathbf{E}\big[\Delta\boldsymbol{x}[k]\Delta\boldsymbol{x}[k]^T\big] + \mathbf{E}\big[\boldsymbol{v}[k]\boldsymbol{v}[k]^T\big] \tag{9a}$$

$$= C\,\mathbf{E}\big[\Delta\boldsymbol{x}[k-1]\Delta\boldsymbol{x}[k-1]^T\big]C^T + \Sigma_v \tag{9b}$$

$$= C\Sigma_x C^T + \Sigma_v. \tag{9c}$$

It is apparent from (8b) and (9c) that any change in the model will cause a change in the covariance matrix Σ_x and in Σ_y. However, it is necessary to identify a change in the covariance matrix which can be related only to islanding detection and not other system events.

It can be observed that, by setting Σ_u to be an identity matrix, the solution of (8b) yields the controllability Gramian of the system. The decreasingly ordered eigenvalues of the controllability Gramian and the associated eigenvectors describe the directions in the space of the system's states where the system is most controllable which is then by extension also reflected by Σ_x and Σ_y. This property of the controllability Gramian is utilized, for instance, in the field of model order reduction as shown in [27] where PCA is used to define the most controllable subspaces of the state space.

Power system models exhibit the property of high correlations between the measurements of the system's frequencies. Furthermore, the controllability Gramian similarly shows that it is not "easy" to control the generator's speeds in all directions of the space by using $\Delta\boldsymbol{p}_L$ as an input, i.e., the frequency tends to be equal in all parts of the system. This property of the Gramian and Σ_y, although not stated explicitly, was used for islanding detection in [23]. A similar phenomenon can be observed in the case of voltage angle differences. There are principal directions of movements of the voltage angle differences and thus the deviation of the angle differences from the dominant directions can signify islanding. It will be shown in Section V that this is indeed the case. However, conversely to the case of frequency measurements, there might exist more than one principal direction of angle differences.

C. Gaussian Characteristics of Angle Differences

One of the premises for successful inference by using PCA of PMU measurements is that the measurements are normally distributed. It was shown in [23] that the histograms of frequency measurements follow the Gaussian distribution, while that was not the case for the measurements of the phase

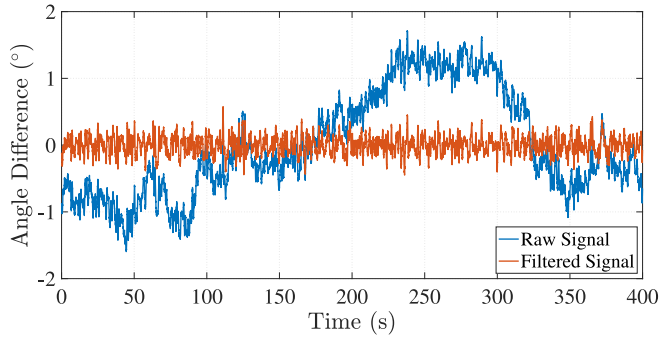


Fig. 1. The effect of FIR filtering of the angle differences.

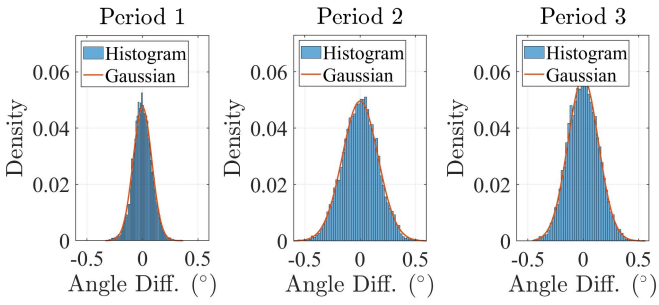


Fig. 2. Histograms of preprocessed voltage angle differences.

angle differences. We claim that the non-Gaussian probability distribution of voltage angle differences arises due to the nonstationary behavior of random loads in the system, i.e., that the random load is comprised of slowly varying equilibrium changes and high frequency Gaussian fluctuations. We also claim that, if the angle differences are detrended, they do indeed follow a Gaussian distribution.

To support the aforementioned claims, real PMU measurements of voltage angle difference between two buses in a system were analyzed over three different periods of time. The measurements were preprocessed by a high-pass FIR filter with a cut-off frequency at 0.1Hz. The length of each considered period corresponds to the time of 400s. A raw and a filtered signal of one of the periods are shown in Fig. 1. Histograms and the fitted pdfs for the measurements in each of the periods are shown in Fig. 2. It can be observed that, if the slow dynamics contained in the measurement signals are removed, the filtered measurements do obey Gaussian distribution. Therefore, as it will be shown in Section V, PCA can be used to perform statistical inference on the newly reported PMU measurements.

III. Probabilistic PCA

A manual selection of r in (3), i.e., the number of principal directions of data, can be challenging when there are no clear dominant directions of variance as can be the case with voltage angle differences. For this reason, we propose to model the PMU measurements of angle differences using the probabilistic PCA (PPCA) model introduced in [28]. The probabilistic treatment permits one to automatically determine the dimensionality of a dataset using the data itself.

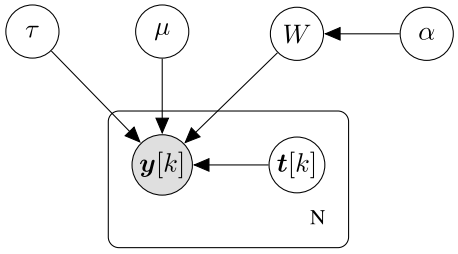


Fig. 3. A graphical representation of the probabilistic model of PCA.

A. Model Description

The graphical representation of the relationships between the random variables of the PPCA model is given in Fig. 3.

Similarly as in the case of classical PCA, $\mathbf{y}[k]$ are observations of an m-dimensional random variable at time k (i.e., a PMU measurement sample), and $\boldsymbol{t}[k]$ is a q-dimensional latent (unobserved) random variable, i.e., a principal component score. PPCA defines the following latent variable model:

$$\boldsymbol{y}[k] = W\boldsymbol{t}[k] + \boldsymbol{\mu} + \boldsymbol{\epsilon}[k], \tag{10}$$

where $W \in \mathbb{R}^{m\times r}$ is a transformation matrix that relates the two random variables. The m-dimensional vector $\boldsymbol{\mu} \sim \mathcal{N}(\boldsymbol{\mu}_\mu, \beta^{-1}I_m)$ models a nonzero mean of the measurements and $\boldsymbol{\epsilon}[k] \sim \mathcal{N}(\mathbf{0}, \tau^{-1}I_m)$ is a noise variable. The latent variable is also assumed to be Gaussian with the following prior:

$$p(\boldsymbol{t}[k]) = \mathcal{N}(\mathbf{0}, I_m). \tag{11}$$

Similarly as in classical PCA, the elements of the principal component score (the latent variable in case of PPCA) are assumed to be uncorrelated since the covariance matrix in (11) is diagonal.

With these definitions, the conditional distribution of the observed variable can be written as:

$$p(\mathbf{y}[k]|\boldsymbol{t}[k], \boldsymbol{\mu}, \tau) = \mathcal{N}\Big(W\boldsymbol{t}[k] + \boldsymbol{\mu}, \tau^{-1}I_m\Big). \tag{12}$$

Applying Bayes' rule and integrating the observed variable results in the marginal distribution of the observed variable:

$$p(\mathbf{y}[k]) = \mathcal{N}\big(\mu, \Sigma_y\big), \tag{13}$$

where $\Sigma_y = WW^T + \tau^{-1}I_m$. Also, we can obtain the posterior distribution of the latent and noise variables:

$$p(\boldsymbol{t}[k]|\mathbf{y}[k]) = \mathcal{N}\Big(M^{-1}W^T(\mathbf{y}[k] - \boldsymbol{\mu}), \tau^{-1}M^{-1}\Big), \tag{14}$$

$$p(\boldsymbol{\epsilon}[k]|\mathbf{y}[k]) = \mathcal{N}\Big(\Big(I_m - WM^{-1}W^T\Big)(\mathbf{y}[k] - \boldsymbol{\mu}), \Sigma_\epsilon\Big), \tag{15}$$

where $M = W^TW + I_r$ and $\Sigma_\epsilon = \tau^{-1}WM^{-1}W^T$. It can be observed here that in contrast to the deterministic PCA, which deterministically projects the data points into a lower-dimensional latent and residual spaces, PPCA provides both the expected value of the projection and its uncertainty.

The automatic determination of dimensionality is achieved by including a prior distribution over the columns of matrix W:

$$p(W|\boldsymbol{\alpha}) = \prod_{i=1}^{r}\Big(\frac{\alpha_i}{2\pi}\Big)e^{-0.5\alpha_i\mathbf{w}_i^T\mathbf{w}_i}, \tag{16}$$

where $\boldsymbol{w}_i$ are the columns of W. The random r-dimensional Gamma-distributed random variable $\boldsymbol{\alpha} = [\alpha_1 \ \cdots \ \alpha_r]$ governs the norm of the columns of W and, in turn, also the dimensionality of the principal subspace.

In summary, this subsection presented the assumed prior distributions for all of the random variables comprising the PPCA model. These distributions can be used to incorporate known information prior to the training of the model. Furthermore, once the PPCA model is trained, it is possible to obtain posterior distributions of the unobserved (latent and noise) variables given new observations, e.g., PMU measurements in the application of islanding detection. These posterior distributions can then be used to quantify the how well the model fits to the new measurements.

B. Learning Algorithm

The probabilistic treatment of PCA makes it possible to apply machine learning algorithms to obtain the model's parameters from measurements. Therefore, in this subsection, we introduce an algorithm for learning the model's parameters originally proposed in [29].

Integrations over the distribution of the parameters of the PPCA model introduced in the previous subsection are analytically intractable due to the model's complexity. In order to circumvent this problem, variational inference methods can be employed [30]. These methods introduce a new simplified probability distribution of the unobserved variables (latent variables and model's parameters):

$$q(\psi) = q(T)q(W)q(\boldsymbol{\alpha})q(\boldsymbol{\mu})q(\tau), \tag{17}$$

where $\psi = \{T, W, \boldsymbol{\alpha}, \boldsymbol{\mu}, \tau\}$ is a set of model's parameters and X is a matrix of latent variables. The factorized distribution $q(\psi)$ serves as an approximation of the true posterior distribution, which simplifies the estimation of parameters described later in (19). Nonetheless, the distribution q can be chosen to be sufficiently close to the original one as shown in [29]. Using $q(\psi)$, the log marginal likelihood can be expanded as:

$$\ln p(Y) = \ln \int p(Y, \psi) d\psi = \ln \int q(\psi) \frac{p(Y, \psi)}{q(\psi)} d\psi \tag{18a}$$

$$\geq \int q(\psi) \ln \frac{p(Y, \psi)}{q(\psi)} d\psi = \mathcal{L}(q). \tag{18b}$$

The inequality in (18b) defines a lower bound on the log data likelihood, $\mathcal{L}(q)$. Increasing $\mathcal{L}(q)$ corresponds to increasing the fit of the PPCA model to the data. It can be shown that the lower bound $\mathcal{L}(q)$ is maximized by iteratively updating model's parameters according to:

$$\ln q_i(\psi_i) = \mathbf{E}_{\psi_i}\big[\ln p(Y, \psi)\big] + \text{const}, \tag{19}$$

where $q_i(\psi_i)$ are elements of the product in (17). Moreover, each iteration of (19) strictly increases the lower bound $\mathcal{L}(q)$ which can serve as an indication of convergence.

C. Monitoring Measures

To be able to quantify how well the newly observed measurements fit the trained PPCA model two monitoring measures need to be introduced. First, a new measurement is projected into the latent and noise variable spaces by taking the expectations of the distributions in (14) and (15), which are denoted by $\mathbf{E}[\boldsymbol{t}[k]]$ and $\mathbf{E}[\boldsymbol{\epsilon}[k]]$, respectively.

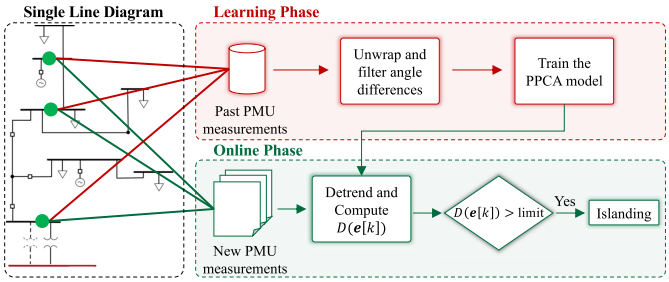


Fig. 4. Proposed islanding detection method.

The following two squared Mahalanobis distances (SMDs) are defined:

$$D(\boldsymbol{t}[k]) = \mathbf{E}[\boldsymbol{t}[k]]^T \mathbf{E}[\boldsymbol{t}[k]], \tag{20}$$

$$D(\boldsymbol{\epsilon}[k]) = \tau^{-1} \mathbf{E}[\boldsymbol{\epsilon}[k]]^T \mathbf{E}[\boldsymbol{\epsilon}[k]]. \tag{21}$$

Using these two distances, one can effectively measure how closely the newly observed variable obeys the trained PPCA model in both the latent (principal) and residual space. A probability distribution of the two measures can be derived and the confidence limits can be defined as shown in [31]. The significance level for calculating the confidence limits for the distances is set to 99.9% in this paper.

IV. Proposed Islanding Detection Method

By utilizing the concept of principal directions of voltage angle differences discussed in Section II and the PPCA model presented in Section III that can be used to identify these principal directions, this section proposes a new islanding detection method. We claim that the proposed method is able to distinguish between islanding and other types of events, and, thereby, addresses the potential weaknesses of methods in [23] and [25].

The flowchart of the islanding detection method is shown in Fig. 4. It consists of a learning (offline) phase and an online phase. In the learning phase, historical PMU measurements of voltage angles are first unwrapped and the angle differences are calculated with respect to a selected reference angle. Before proceeding, such angle differences are filtered to remove the effects of nonstationarity discussed in Section II-C. These preprocessed measurements are then used to train the PPCA model using the learning algorithm presented in Section III-B. Once the PPCA model is trained, it is ready to be used in the online phase. The training of the model can continue in parallel with the online operation, and the online phase can be updated, for instance at regular intervals, with the results from the offline phase.

In the online phase, PMU measurements are sequentially processed. First, the voltage angle measurements are unwrapped, which can be done as proposed in [32], and the angle differences are calculated. These differences also need to be detrended to be able to use the trained PPCA model. The detrended data is then used to estimate the means of the latent

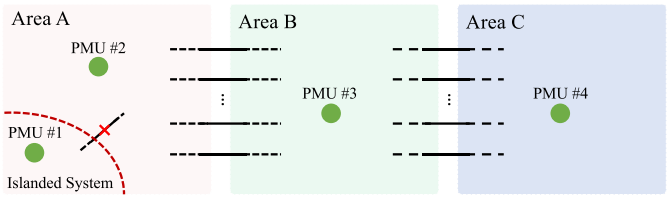


Fig. 5. Overview of the PMU locations in the power system.

and residual variables using (14) and (15). Finally, SMD of the residual variable, computed using (21), can be compared with the confidence limit. If the distance is above the limit, it is concluded that angle differences have departed from the principal subspace and an islanding event has happened. Similarly, the SMD of the latent variable from (20) can be used to detect other nonislanding types of events.

A. Signal Filtering

As mentioned, angle differences are filtered in both the learning and the online phase. In the learning phase, the signals are filtered using a finite impulse response (FIR) high-pass filter with a cut-off frequency at 0.1Hz. This filter is of linear phase which ensures that the filtered angle differences are not distorted. Since the filtering delay in the learning phase does not degrade the method's performance, filter's order was selected to be 2129 resulting in a group delay of 21.28s.

The high-pass filter from the learning phase cannot be used in the online phase due to this large group delay. Therefore, in the online phase the signals are instead detrended by subtracting a moving average from the signal. The moving average is computed over a window of 500 samples of angle differences which corresponds to a group delay of 4.98s. It should be noted that the length of the moving average filter might imply a nonnegligible filtering delay. However, since the subtraction of the moving average does not affect high frequencies notably, these are reflected in the output with essentially no delay.

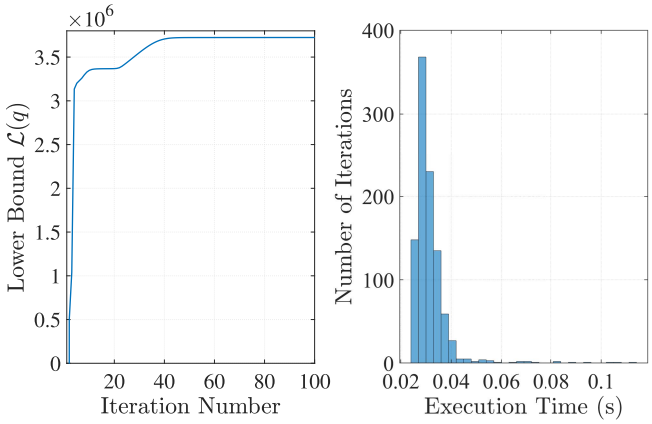
V. Case Studies

This section demonstrates the performance of the proposed islanding detection method by applying it to real PMU measurements taken from a European transmission system which were reported at a rate of 50 samples per second. The abstracted diagram in Fig. 5 shows the locations of the PMUs in the system. The reference location with respect to which the positive-sequence angle differences were computed was taken to be at PMU #4.

The following subsection presents results from the learning phase of the method followed by the results of the method's performance during two events, namely, frequency oscillations and islanding. Additionally, the results of the methods proposed in [23] and [24], respectively referred to as PCA and recursive PCA (RPCA), are shown for the same two events.

A. Model Learning

The historical data which was used in the learning phase of the case study is comprised of $N = 177964$ observations corresponding to 59.3 hours of measurements. These measurements were taken from time periods spread over a year and 4 months of system's normal operation. Given that each observation is a vector containing angle differences from $d = 4$ different locations, the number of processed measurement samples in the learning phase is 533892. Since the large amount of data can imply a high computational burden of the learning phase, the case study evaluates not only the performance of the method but also the feasibility in terms of computational burden.



(a) Lower bound during training (b) Histogram of execution times

Fig. 6. Convergence of the learning phase reflected in the monotonic increase of the lower bound and the histogram of iterations' execution times.

As previously mentioned, the convergence of the learning algorithm can be tracked via the lower bound on the log data likelihood introduced in (18b), i.e., the training of the model is interrupted once a sufficiently small change in the lower bound between two consequtive iterations is observed. Fig. 6(a) shows the evaluation of the lower bound at each iteration of the learning algorithm. It can be observed that the bound is monotonically increasing and already past the 50th iteration there is no large change in the bound which then indicates model convergence. The training was carried out on a PC equipped with an Intel Core i7-7600U 2.8 GHz CPU and 32 GB of RAM. The execution of 100 training iterations took 4.01s. Furthermore, a histogram of execution times of 1000 iterations is shown in Fig. 6(b). Most of the iterations were executed faster than 0.06s.

In the following subsections, previously unseen measurements from two events and the trained model discussed in this section are utilized to validate the detection method.

B. Frequency Oscillations

As stated in Section I, an islanding detection method needs to be stable, i.e., it must not operate for nonislanding events such as frequency oscillations. Therefore, this event is used in the case study to illustrate the stability of the proposed detection method. Fig. 7 shows PMU measurements of frequency at three different locations in the system, and Fig. 8 shows angle differences with respect to an angle at one of the locations. For the sake of figure clarity, the angle differences were centered

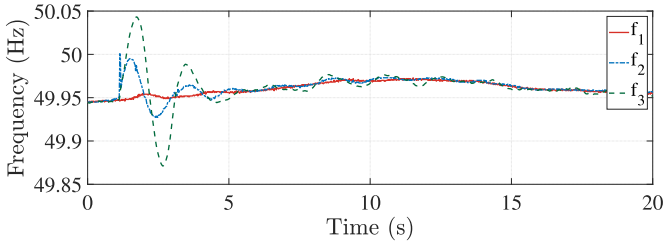


Fig. 7. PMU frequency measurements during an oscillation event.

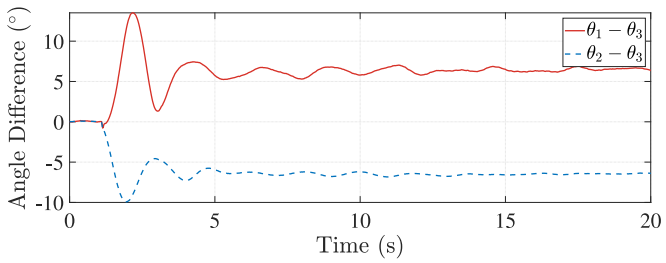


Fig. 8. PMU measurements of angle differences during an oscillation event.

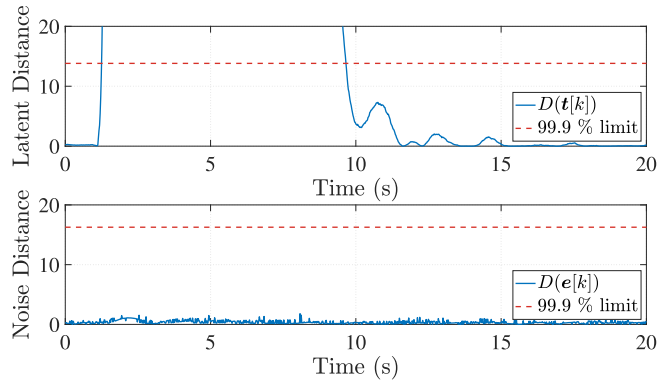


Fig. 9. Squared Mahalanobis distances of latent and noise variables during a frequency oscillations event.

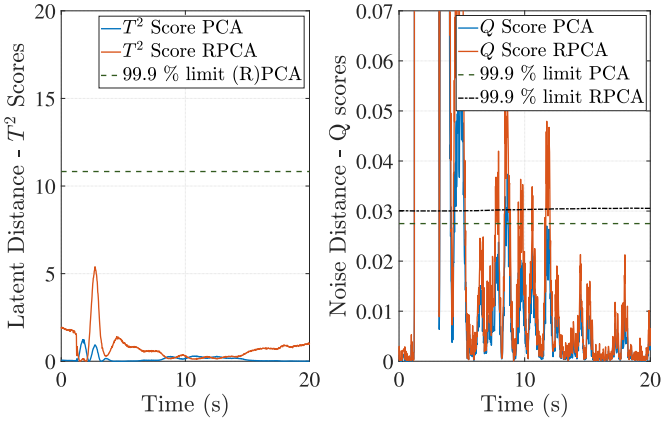


Fig. 10. PCA and RPCA results during the frequency oscillation event.

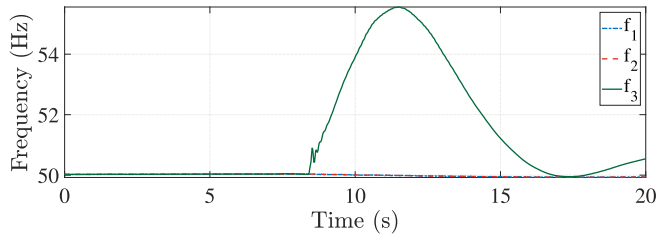


Fig. 11. PMU frequency measurements during an islanding event.

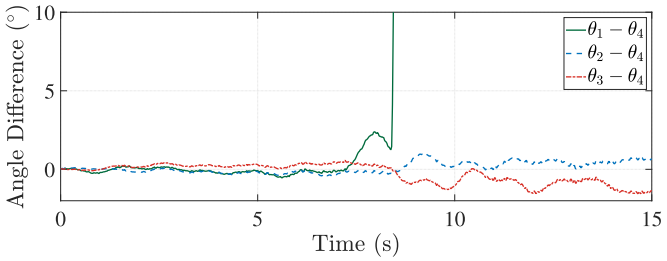


Fig. 12. PMU measurements of angle differences during an islanding event.

by subtracting the measurements at $t = 0$s from measurements at other time instants. The measurements were taken during a period of frequency oscillations that were caused by an unknown event at approximately $t = 1.1$s.

The PPCA model, previously trained with historical PMU measurements of filtered voltage angle differences, was used to obtain the SMDs of latent and noise variables that are shown in Fig. 9. As can be seen, the SMD of the latent variable surpasses the confidence limit at the start of the oscillation event, while that is not the case for the SMD of the noise variable. This signifies that an event occurred which caused the angle differences to behave as modeled by the principal components of the trained PPCA model. Therefore, the method is stable during the frequency oscillation event.

Similarly, the T^2 and Q scores, i.e., the equivalents to the SMDs of latend and noise variables respectively, of the PCA and RPCA methods are shown in Fig. 10 along with their respective statistical limits. It can be observed that the Q score of both PCA and RPCA methods surpasses the statistical limit during the first part of the frequency oscillations which would result in a false detection of an islanding event.

C. System Islanding

Another case that demonstrates the performance of the method is an islanding event. Fig. 11 and Fig. 12 show PMU measurements of frequency and voltage angle differences, respectively. Again, the angle differences were centered with respect to the initial value at the start of the observation period.

At approximately $t = 8.4$s, an islanding event occurs. Both the frequency and the angle difference at the islanded, i.e., at the location of PMU #1 in Fig. 5, start to diverge from the ones at the other two locations. The same PPCA model as in the preceding case study was used to obtain the SMDs of the latent and noise variables shown in Fig. 13. Similarly, the T^2 and Q scores of the PCA and RPCA methods are shown in Fig. 14. It can be observed that, upon the inception of the islanding event, both of the monitoring measures of the proposed method start growing and exceed their respective limits at $t = 9.12$s, i.e., the method operated in 720ms. It can similarly be observed that the PCA and RPCA methods operate for the islanding event within 60ms after the islanding event, disregarding the proposed time delay of 500ms suggested by [23].

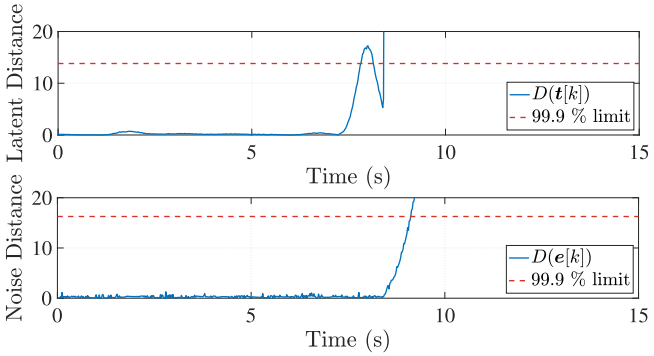


Fig. 13. SMDs of the latent and noise variables during an islanding event.

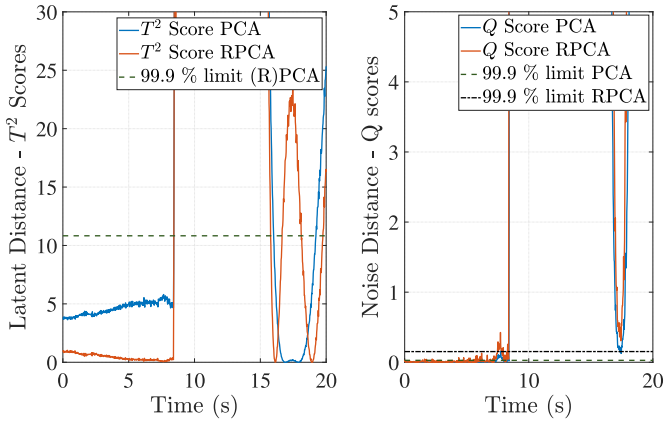


Fig. 14. PCA and RPCA results during the islanding event.

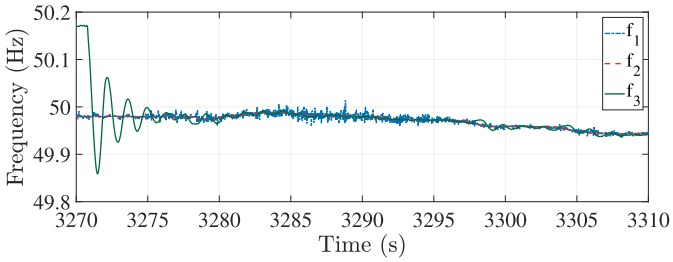


Fig. 15. PMU frequency measurements during the resynchronization event.

Both SMDs of the proposed method exceed the limits during approximately 55 minutes after which the island was resynchronized to the rest of the system. Fig. 15 and Fig. 16 show measurements of frequency and centered angle differences during the resynchronization event which started at $t = 3270.76$s. Fig. 17 shows the monitoring measures during the resynchronization event. The SMD of the noise variable falls below the confidence limit at $t = 3271.58$s which indicates detection of the resynchronization in 820ms.

VI. Discussion

The preceding section demonstrated a successful application of the proposed method for islanding detection. Additionally, it should be pointed out that the method remained stable during the frequency oscillation event. Conversely, we demonstrated that, since the frequency measurements shown in Fig. 7 were not equal at all three PMU locations, the methods based on

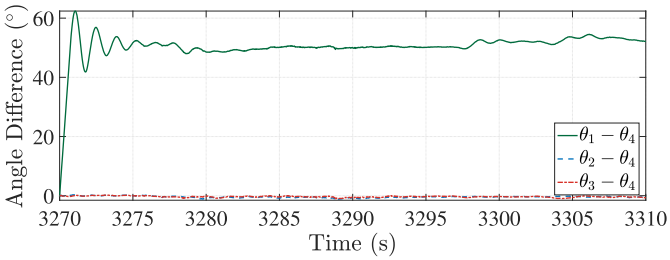


Fig. 16. PMU measurements of angle differences during the resynchronization event.

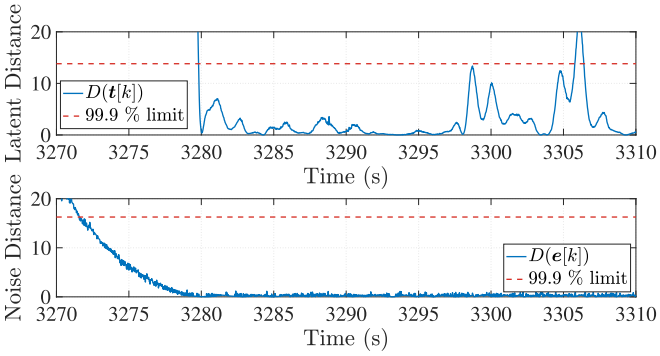


Fig. 17. SMDs of latent and noise variables during the resynchronization event.

frequency differences, namely those proposed in [24] or [23], can erroneously classify this event as islanding.

It should also be noted that the proposed method is limited by the existence of its non-detection zone. The analyzed islanding event is an example where the generation and load in the island were not balanced, and consequently, this was reflected in the apparent divergence of both frequency and angle differences. Conversely, in a case of a perfect balance in the island, the divergence would not be as apparent. However, the high sensitivity of voltage angle differences with respect to the power imbalance greatly reduces the possibility for an unsuccessful islanding detection.

Furthermore, care has to be taken when selecting the measurement locations. The proposed method is based on the existence of the principle subspace in the prefiltered measurements of angle differences. If the locations are chosen in such a way that they are uncorrelated, e.g., when the covariance matrix Σ_y is diagonal, the principal subspace would be of an effective dimensionality equal to m, i.e., the dimension of the measurements. This would mean that the measurements can "move" in all directions and that the method would not be able to discriminate between different types of events in the system. Therefore, one has to ensure, while selecting the measurement locations, that the angle differences are sufficiently correlated, which also means that the measurements have to be sufficiently electrically close. In other words, the more electrically close the measurements are, the better the method will perform.

Another potential pitfall in the application of the proposed method is imperfect detrending. While the PPCA model was trained by using the measurements processed by an FIR filter, the same filter cannot be applied in real-time due to the filter delay which is the reason why we resorted to the moving

average filter. Certainly, the frequency response of the moving average filter is not as good as the one of the FIR filter. The poorer performance of the moving average detrending might jeopardize the stability of the method, however, this was not observed in the cases studied in the previous section. In relation to the filtering delay, it should also be noted that the performance of the proposed method might be affected by delays introduced by the communication infrastructure.

VII. Conclusion

This paper analyzed the covariance matrices of a linearized power system model and showed that detrending ensures normality of the voltage angle differences and permits the use of the proposed inference procedures. Building on this analysis, a data-driven method for islanding detection using probabilistic PCA of detrended measurements of angle differences was proposed. The advantage of the method is that, by using only the measurements of the angle differences, the non-detection zone is reduced compared to methods based on frequency differences. Furthermore, the need for user-defined settings of the thresholds is removed through the employed machine learning algorithm. The method was applied to the real historical and also previously unseen PMU measurements from two different events which illustrated that the SMD of the noise variable can be used for timely and reliable identification of islanding events and that the SMD of the latent variable indicates the occurrence of nonislanding events.

It should also be noted that, conversely to other similar methods which are based on the adaptive (moving window) PCA models, the proposed method in this paper performs all of the computationally expensive operations in the offline phase as shown in the Fig. 4. In the online phase, the method only performs the computationally inexpensive operations needed to compute SMDs of the two variables. In other words, the method is suitable for real-time application.

References

[1] *Smarter, Greener, More Inclusive? Indicators to Support the Europe 2020 Strategy*, Eurostat, Luxembourg City, Luxembourg, 2019.

[2] G. Pepermans, J. Driesen, D. Haeseldonckx, R. Belmans, and W. D'haeseleer, "Distributed generation: Definition, benefits and issues," *Energy Policy*, vol. 33, no. 6, pp. 787–798, 2005.

[3] N. Jenkins, J. Ekanayake, and G. Strbac, *Distributed Generation* (Energy Engineering Series). London, U.K.: Inst. Eng. Technol., 2010, pp. 12–18.

[4] M. Bollen and F. Hassan, *Integration of Distributed Generation in the Power System* (Power Engineering). Hoboken, NJ, USA: Wiley-IEEE Press, 2011, pp. 340–342.

[5] "Technical report on the events of 9 August 2019," Nat. Grid ESO, Warwick, U.K., Rep., 2019. [Online]. Available: https://www.nationalgrideso.com/document/152346/download

[6] P. Mahat, Z. Chen, and B. Bak-Jensen, "Review of islanding detection methods for distributed generation," in *Proc. 3rd Int. Conf. Electr. Utility Deregulation Restructuring Power Technol.*, Apr. 2008, pp. 2743–2748.

[7] "Using synchrophasor data during system islanding events and blackstart restoration," NASPI Control Room Solutions Task Team, North Amer. Synchrophasor Initiative, Rep. NASPI-2015-TR-013, 2015.

[8] C. G. Bright, "COROCOF: Comparison of rate of change of frequency protection. A solution to the detection of loss of mains," in *Proc. 7th Int. Conf. Develop. Power Syst. Prot. (IEE)*, Apr. 2001, pp. 70–73.

[9] C. An, G. Millar, G. J. Lloyd, A. Dyśko, G. M. Burt, and F. Malone, "Experience with accumulated phase angle drift measurement for islanding detection," in *Proc. 11th IET Int. Conf. Develop. Power Syst. Prot. (DPSP)*, Apr. 2012, pp. 1–6.

[10] Z. Lin *et al.*, "Application of wide area measurement systems to islanding detection of bulk power systems," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 2006–2015, May 2013.

[11] J. Mulhausen, J. Schaefer, M. Mynam, A. Guzmán, and M. Donolo, "Anti-islanding today, successful islanding in the future," in *Proc. 63rd Annu. Conf. Protective Relay Eng.*, Mar. 2010, pp. 1–8.

[12] R. Arghandeh and Y. Zhou, *Big Data Application in Power Systems*. Amsterdam, The Netherlands: Elsevier Sci., 2017.

[13] K. Pearson, "LIII. On lines and planes of closest fit to systems of points in space," *London Edinburgh Dublin Philosoph. Mag. J. Sci.*, vol. 2, no. 11, pp. 559–572, Nov. 1901.

[14] 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.

[15] L. Cai, N. F. Thornhill, S. Kuenzel, and B. C. Pal, "Wide-area monitoring of power systems using principal component analysis and *k*-nearest neighbor analysis," *IEEE Trans. Power Syst.*, vol. 33, no. 5, pp. 4913–4923, Sep. 2018.

[16] W. Li, M. Wang, and J. H. Chow, "Real-time event identification through low-dimensional subspace characterization of high-dimensional synchrophasor data," *IEEE Trans. Power Syst.*, vol. 33, no. 5, pp. 4937–4947, Sep. 2018.

[17] J. M. Lim and C. L. DeMarco, "SVD-based voltage stability assessment from phasor measurement unit data," *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp. 2557–2565, Jul. 2016.

[18] P. Gao, M. Wang, S. G. Ghiocel, J. H. Chow, B. Fardanesh, and G. Stefopoulos, "Missing data recovery by exploiting low-dimensionality in power system synchrophasor measurements," *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 1006–1013, Mar. 2016.

[19] K. Mahapatra, N. R. Chaudhuri, R. G. Kavasseri, and S. M. Brahma, "Online analytical characterization of outliers in synchrophasor measurements: A singular value perturbation viewpoint," *IEEE Trans. Power Syst.*, vol. 33, no. 4, pp. 3863–3874, Jul. 2018.

[20] M. Brown, M. Biswal, S. Brahma, S. J. Ranade, and H. Cao, "Characterizing and quantifying noise in PMU data," in *Proc. IEEE Power Energy Soc. Gen. Meeting (PESGM)*, Jul. 2016, pp. 1–5.

[21] P. H. Gadde, M. Biswal, S. Brahma, and H. Cao, "Efficient compression of PMU data in WAMS," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2406–2413, Sep. 2016.

[22] K. K. Anaparthi, B. Chaudhuri, N. F. Thornhill, and B. C. Pal, "Coherency identification in power systems through principal component analysis," *IEEE Trans. Power Syst.*, vol. 20, no. 3, pp. 1658–1660, Aug. 2005.

[23] X. Liu, D. M. Laverty, R. J. Best, K. Li, D. J. Morrow, and S. McLoone, "Principal component analysis of wide-area phasor measurements for islanding detection—A geometric view," *IEEE Trans. Power Del.*, vol. 30, no. 2, pp. 976–985, Apr. 2015.

[24] Y. Guo, K. Li, D. M. Laverty, and Y. Xue, "Synchrophasor-based islanding detection for distributed generation systems using systematic principal component analysis approaches," *IEEE Trans. Power Del.*, vol. 30, no. 6, pp. 2544–2552, Dec. 2015.

[25] X. Liu, J. M. Kennedy, D. M. Laverty, D. J. Morrow, and S. McLoone, "Wide-area phase-angle measurements for islanding detection—An adaptive nonlinear approach," *IEEE Trans. Power Del.*, vol. 31, no. 4, pp. 1901–1911, Aug. 2016.

[26] T. Rabuzin, J. Lavenius, N. Taylor, and L. Nordström, "Bayesian detection of islanding events using voltage angle measurements," in *Proc. IEEE Int. Conf. Commun. Control Comput. Technol. Smart Grids (SmartGridComm)*, Oct. 2018, pp. 1–6.

[27] B. Moore, "Principal component analysis in linear systems: Controllability, observability, and model reduction," *IEEE Trans. Autom. Control*, vol. 26, no. 1, pp. 17–32, Feb. 1981.

[28] M. E. Tipping and C. M. Bishop, "Probabilistic principal component analysis," *J. Royal Stat. Soc. B, Stat. Methodol.*, vol. 61, no. 3, pp. 611–622, 1999.

[29] C. Bishop, "Variational principal components," in *Proc. 9th Int. Conf. Artif. Neural Netw. (ICANN)*, vol. 1, Jan. 1999, pp. 509–514.

[30] M. I. Jordan, Z. Ghahramani, T. S. Jaakkola, and L. K. Saul, "An introduction to variational methods for graphical models," *Mach. Learn.*, vol. 37, no. 2, pp. 183–233, Nov. 1999.

[31] D. Kim and I.-B. Lee, "Process monitoring based on probabilistic PCA," *Chemometr. Intell. Lab. Syst.*, vol. 67, no. 2, pp. 109–123, 2003.

[32] V. Venkatasubramanian, "Real-time strategies for unwrapping of synchrophasor phase angles," *IEEE Trans. Power Syst.*, vol. 31, no. 6, pp. 5033–5041, Nov. 2016.