Synergistic Data Analytics for Electromechanical Oscillation in Electric Power Systems

Tao Jiang[1]*, Xiao Kou[2], Guodong Liu[3]*, Haoyu Yuan[2]

[1]: *Department of Electrical Engineering, Northeast Electric Power University, Jilin, JL 132012, China*
[2]: *University of Tennessee, Knoxville, TN 37996, USA*
[3]: *Power and Energy Systems Group, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA*
[*]: Contact author: tjiang@neepu.edu.cn, liug@ornl.gov

Abstract—Accurate real-time estimation of the four electromechanical oscillation properties, i.e., dominant oscillation modes, mode shapes, participation factors, and coherent groups, is of great importance to assess and mitigate potential electromechanical oscillations in interconnected power systems. While eigenvalue analysis can realize such estimation, it requires precise linearized dynamic models and accurate parameters, which are highly difficult to obtain in practice. Data fusion-based modal estimation methods can extract the properties of electromechanical oscillations from measurement data without the power system model and parameters, but most of the time only one or two property assessments can be accomplished each time. To overcome this challenge, this paper presents a synergistic data analytics solution to characterize the dynamic behaviors of electromechanical oscillations from real-time measurement data. The proposed method uses optimized variable projection, and it is capable of estimating all four electromechanical oscillation behavior properties from measured responses. Case studies are performed using the simulated measurement data of a 16-generator 68-bus test system and the field measurements collected by the PMUs deployed in the Yunnan Power Grid. The results demonstrate that the proposed synergistic data analytics solution can achieve satisfactory performance in capturing the properties of electromechanical oscillations from measurement data and exhibit strong robustness against measurement noise when compared with existing measurement-based methods.

***Keywords*—**Electromechanical oscillation; dominant mode; mode shape; participation factor; coherent group; measurement data.

1. Introduction

Electromechanical oscillation, an intrinsic dynamic characteristic of synchronous power grids, can be a great threat to the stability of power systems [1], [2]. An undamped electromechanical oscillation mode might intensify and eventually lead to power system breakups and wide-area blackouts [3], [4]. There are many incidents of adverse electromechanical oscillation events worldwide, among which one notable event was the Western Electricity Coordinating Council (WECC) system breakup caused by an undamped inter-area oscillation mode on August 10, 1996 [5]. Sustained low-frequency oscillations were observed during these events and then resulted in power system instability incidents [1], [3], [5], [6]. Therefore, assessing the power system electromechanical oscillation in real-time is of great importance.

Mode estimation, which can reveal vital dynamic characteristics of the electromechanical oscillations in power systems, is essential for assessing and damping the oscillation [3], [4], [6]. Mathematically, the modes are the eigenvalues of a linear dynamic model [3], [4]. A traditional method to obtain the power system oscillation modes is conducting eigenvalue analysis (EA) on a linearized power system dynamic model at its given operating point [3], [4]. However, the estimated result of the EA depends on the accuracy of the linearized power system dynamic model. Maintaining an accurate dynamic model with high precision parameters is a huge challenge, especially in today's bulk power systems. As demonstrated by the WECC system breakup on August 10, 1996 and the sustained ultralow frequency oscillation event in the China Southern Power Grid (CSG) on March 28, 2016, the simulated dynamic response of the initial formulated power system dynamic models was unable to match the field measurement data, i.e., the observed oscillations cannot be reflected and observed through dynamic models [5].

To improve the model-based EA approaches and facilitate the electromechanical oscillation assessment and controls in interconnected power grids, data-driven mode estimation methods have been extensively applied to extract the dynamic oscillation properties from the real-time PMU data [6], [7], [8]. The field-measured data collected by PMUs, which are synchronized with GPS/GLONASS/BDS/Galileo clock and stamped with time labels, typically have a sampling frequency of at least 25Hz to enable the dynamic features of power systems are fully captured in real-time [6]. Due to the unique advantages of measurement-based approaches over model-based EA, various data-driven methods have been developed to estimate the electromechanical oscillation properties from field-measured responses, which have attracted widespread attention from both academia and industry [7]-[16]. Particularly, organizations and communities, such as North American SynchroPhasor Initiative (NASPI), WECC Joint Synchronized Information Subcommittee (WECC-JSIS), etc., have been funded to investigate the applications of data-driven mode estimation methods in real-world power grids. A large number of PMUs have been installed and the Wide Area Monitoring System (WAMS) has been constructed in power grids. In addition, many commercial and open-source tools, such as open-source phasor data concentrator (OpenPDC), real time dynamics monitoring system (RTDMS), synchroWAVe, etc., were also developed to estimate and monitor the dominant oscillation modes through field-measured data collected by PMUs in the power systems.

In the dominant mode estimation, given typical disturbances in real power systems, measurements are usually divided into two groups: ringdown and ambient. The ringdown data is excited by power system major disturbances such as HVDC blocking, branch/generator tripping, etc. To detect the dominant modes from the ringdown data, a Prony method was proposed in [12] and extended in [17]. However, its performance is subject to measurement noise and signal offsets [18]. To address this issue, Messina *et al*. [19] employed empirical mode decomposition (EMD) to estimate the dominant modes, but this approach suffers from the mode mixing problem. Later, a multi-band modal analysis was proposed in [20] to eliminate the mode mixing. Besides the ringdown data, the ambient data, which is excited by the random variations of small load changes, is more common in power systems. Different from the high dynamic information density embedded in the ringdown data, the ambient data only contains finite dynamic information density. Hence, the Prony and EMD are incapable of extracting the dominant modes from the ambient data. In this context, the autoregressive (AR) and autoregressive moving average (ARMA), which are estimated by the Yule-Walker [22], modified Yule-Walker [23], or extended modified Yule-Walker [24], were adopted to detect the dominant oscillation modes from ambient data. To improve the estimation quality, a bootstrap-based confidence interval mode estimation was designed by Anderson *et al* in [25]. To suppress outliers and missing data in the ARMA-based dominant mode estimation, a robust recursive least square (RRLS) based ARMA model was developed in [26] to detect the dominant modes from ambient data. Since the RRLS-based ARMA model is unable to deal with the nontypical data in the power systems, a regularized RRLS (R3LS) based autoregressive moving average exogenous (ARMAX) model was developed to consider both typical and nontypical measurement data [27]. In addition to the ARMA or ARMAX based transfer function (TF) methods, the eigensystem realization algorithm (ERA) and subspace state space system identification (N4SID) based subspace identification methods were proposed in [28] and [29]. The subspace identification methods have been confirmed to be robust against signal noise, but they yield high computational burdens. To improve the computational efficiency of the subspace identification methods in real-time, a recursive adaptive stochastic subspace identification method was proposed by Nezam *et al* [30].

Similar to the estimated damping and frequency in the dominant oscillation mode, the mode shapes, participation factors, and coherent groups of generators are also important dynamic properties of electromechanical oscillations in power systems. They can potentially aid in oscillation mitigation and control decision-making by providing critical information. For example, the mode shapes could help system operators decide the optimal generator/branch tripping plan to increase the damping of an oscillation mode. Participation factors can be used to assess the observability and controllability of each generator associated with a single oscillation mode, and then determine the best placement of power system stabilizer (PSS) to damp the corresponding oscillation. The coherent groups of generators can be used to split the entire power grid into several sustainable islands to minimize the spread of blackouts caused by the sustained electromechanical oscillations with increasing amplitudes.

In the mode shape estimation, the EMD was applied by Messina *et al* [31] to estimate the mode shape properties from the measured ringdown data. However, the lack of connection between the data-driven analysis and the eigen-solution of the system obstructs the application of the measurement-based data-driven methods in the mode shape estimation. Later, this issue was resolved by Trudnowski *et al* [32], where the theoretical connection between measurable spectral properties and the system's eigenvector properties has been derived. After that, many efforts were devoted to estimating mode shapes from the measurement data. Similar to the measurement-based mode estimation methods, the existing mode shape estimation methods are classified into two categories: ringdown and ambient. The mode shape estimation from the measurement data was first applied to the ringdown data by Hauer *et al* via the Prony. Then Hilbert-Huang Transform was presented by Senroy [33] to track the instantaneous mode shapes using the ringdown data. In [34], recursive Kalman filtering was proposed by Chaudhuri *et al* to identify the mode shapes from the measurement data under transient conditions. In [35], wavelet transform was employed by Avdaković *et al* to estimate the mode shapes through analyzing the phase angle difference of generators. In [36], the cross wavelet transform (XWT) was proposed to estimate the mode shapes from ringdown data and to estimate the mode shapes from the measured ambient data. In [37] and [38], Autoregressive Exogenous (ARX) and multichannel ARMAX based Transfer Function (TF) methods were developed to estimate the mode shapes from the measured ambient data, respectively. In [15] and [39], wavelet-based Spectral methods such as XWT and multichannel CWT were implemented to identify the mode shape. In [40] and [41], the frequency domain decomposition (FDD) was applied to identify the mode shape. In [42] and [43], channel matching was proposed to estimate the mode shapes from the ambient data. In [44], [45], and [46], the ERA and N4SID based subspace identification methods were developed for mode shape identification using ambient data. Further, the performances of TF, spectral, FDD, channel matching, and subspace identification methods were compared by Dosiek *et al*. in [47]. The result showed that the Spectral, FDD, channel matching and TF methods failed in the high damping case, while the subspace identification exhibited better performance under different ambient conditions.

In the participation factor estimation, Koopman mode was developed by Susuki and Mezic [48] to estimate the participation factors from the measured non-stationary data. In [49], a dynamic mode decomposition (DMD) method was adopted to identify the participation factors of the generators associated with the dominant modes from measurement data. In our previous works, the CWT [50], ERA [44], and SSI [45] were employed to estimate the participation factors from measurement data. Other than Koopman, DMD, CWT, ERA, and SSI, little research has been conducted on using measurement data to estimate participation factors.

In the coherent group estimation, the measurement-based coherency identification methods are roughly classified into two major categories: clustering [48]-[54] and direction cosine [44], [45], [55], [56]. In the clustering-based coherency identification, the critical dynamic characteristics representing coherency property are extracted from the measurement data using Koopman mode [48], dynamic mode decomposition [49], principal component analysis [51], and independent component analysis [52]. The hierarchical clustering [53] and spectral clustering [54] are then adopted to cluster the coherent groups of generators. The clustering methods have been proven to be reliable in the coherent group cluster using the measurement responses. However, the disadvantage is that the number of coherent groups or the distance threshold has to be specified in advance. The direction cosine-based coherent group separation methods can overcome the deficiencies of the clustering-based coherent group identification methods. In [55] and [56], the connection between coherent generators and the direction cosine was presented. The conclusion is that two generators are completely coherent if the direction cosine is 1, and they are completely incoherent if the direction cosine is close to -1 [44], [45], [50]. When using measurement data, the measurement-based direction cosine can be represented by the correlation coefficient between the measurement data from two different measurement channels, and the coherency matrix of the generators can be further formed through measurement-based direction cosine. Based on the entries in the formed coherency matrix, coherent groups of generators can be clustered. Nevertheless, a prior direction cosine threshold is needed. In addition to clustering and direction cosine, a projection pursuit approach was developed in [57] to identify the critical coherent group and the dominant coherent groups from measurement data via the optimal orthogonal projection direction. Also, the theoretical connection between the measurable coherency properties and the coherency in the Center of Inertia (COI) was provided.

Given the value of dominant modes, mode shapes, participation factors, and coherent groups as well as the intrinsic connections among the four dynamic properties, estimating these four properties in a systematic way has been accomplished using the model-based EA, and the results have been used in the electromechanical oscillation assessments and real-world power system controls [4]. Comparing with the holistic model-based EA, most measurement-based data-driven methods can only estimate one or two oscillation properties among the four dynamic properties. Although measurement-based data-driven methods have big advantages as they eliminate the need of building complex power system dynamic models, existing research that has been dedicated to estimating all of the four dynamic properties using measurement data in a holistic framework is still limited [39], [49]. In [39], a multichannel continuous wavelet transform (CWT) approach was proposed to estimate the electromechanical oscillation modes, mode shapes, and coherent groups from the measurement data. However, the proposed approach was incapable of estimating the participation factors. In addition, the proposed procedure of coherent group estimation through inversed

CWT is extremely complicated and time-consuming. In [49], A DMD method was proposed to estimate the dominant modes, mode shapes, participation factors, and coherent groups of generators from the measured responses. Nevertheless, the proposed DMD method is very sensitive to measurement noise. Moreover, the DMD-based method requires given initial modes to estimate the dominant modes from measurement data. Given the drawbacks of existing research, developing a synergistic data analytics framework for estimating the dominant mode, mode shapes, participation factors, and coherent groups from the power system field measurement data in a holistic way remains an open issue.

To bridge the gap between model-based EA and measurement-based methods on power system dynamic property estimation, this paper develops an integrated framework to estimate the dominant oscillations, mode shapes, participation factors, and coherent groups from measurement data in a systematic way. In specific, an optimized variable projection (OptVarPro) based synergistic data analytics solution is developed, which extends the functionality of the VarPro-based data-driven method from dominant mode estimation to dominant mode, mode shape, participation factor, and coherent group estimations. The VarPro method has been proven by Borden *et al* [58] to have better accuracy on mode estimation than linear measurement-based mode estimation methods, such as Prony, ERA, Matrix-Pencil. However, the performance of VarPro is subject to the initially given modes which are usually difficult to be obtained. Moreover, even though the VarPro can accurately predicate the dominant modes from measurement data, the estimations of mode shape, participation factor, and coherent generator group using VarPro are rarely investigated [59]. To solve these two issues, the OptVarPro-based synergistic data analytics solution is used in this work to determine appropriate given modes and realize the dominant mode, mode shape, participation factor, and coherent group estimations in a single framework. The accuracy and effectiveness of the proposed OptVarPro-based method are validated through simulation using a 16-generator 68-bus system and the field-measured PMU data from the Yunnan Power Grid (YPG) in China. Similar to model-based EA methods, the proposed OptVarPro-based method could obtain accurate dominant modes, mode shapes, participation factors, and coherent groups in a systematic way. In addition, the proposed OptVarPro-based data-driven method is robust to measurement noise compared with other measurement-based data-driven methods. The main contributions of this paper include:

1) The proposed OptVarPro based synergistic data analytics solution realizes the estimation of electromechanical oscillation modes, mode shapes, participate factors, and coherent groups in one holistic framework.
2) The proposed OptVarPro method overcomes the disadvantages of the VarPro on the initial mode selection and extends the VarPro-based method to further estimate mode shapes, participation factors, and coherent groups.
3) The proposed OptVarPro is robust against measurement noise on the estimations of mode, mode shape, participate factor, and coherent group when compared with the existing measurement-based methods.

The rest of this paper is organized as follows. In Section 2, an OptVarPro-based synergistic data analytics solution is developed to estimate the dominant mode, mode shape, participation factor, and coherent group from multichannel measurement data. In Section 3, the performance of the proposed OptVarPro method is evaluated in the 16-generator 68-bus as well as the YPG test systems. Finally, section 4 draws conclusions.

2. Optimized Variable Projection

In this section, the VarPro-based mode estimation is briefly introduced. Then an OptVarPro-based method is developed to obtain more reasonable initial given modes $\boldsymbol{\alpha}$. Further, the mode shapes, participation factors, and coherent groups are also estimated using the developed OptVarPro method.

2.1 Variable Projection

The VarPro method was initially developed by Golub *et al* [60], [61], and [62] to cope with the variable separation problem in nonlinear least squares. In [58], VarPro is further employed to estimate the oscillation modes from the measurement responses in the power system. The procedures of the VarPro based mode estimation can be summarized as follows [59]:

1) Gather the synchrophasor measurements $\boldsymbol{Y}=[\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_m]^{\mathrm{T}}$ from PMUs, where $\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_m \in \mathbf{R}^{1\times N}$ are the measurement responses, m is the number of the measurement channels, and N is the number of measurement samples.

2) Construct basis function matrix $\boldsymbol{f}(\boldsymbol{\lambda})=[\mathrm{e}^{\lambda_1 t}, \mathrm{e}^{\lambda_2 t}, \cdots, \mathrm{e}^{\lambda_n t}]$ where $\boldsymbol{\lambda}=[\lambda_1, \lambda_2, \cdots, \lambda_n]$ are initial given mode vector.

3) Reconstruct the measurements $\boldsymbol{Y}$ as $\hat{\boldsymbol{Y}}=(\boldsymbol{f}(\boldsymbol{\lambda})\boldsymbol{F})^{\mathrm{T}}$ using the basis function matrix $\boldsymbol{f}(\boldsymbol{\lambda})$ and Frobenius matrix $\boldsymbol{F}$.

4) Solve the Frobenius matrix $\boldsymbol{F}$ by minimizing the residual $\boldsymbol{r}$ satisfying $\|\boldsymbol{r}\|_2^2=\|\boldsymbol{Y}^{\mathrm{T}}-\hat{\boldsymbol{Y}}\|_2^2=\|\boldsymbol{Y}^{\mathrm{T}}-\boldsymbol{f}(\boldsymbol{\lambda})\boldsymbol{B}\|_2^2$, and the linear least-squares optimization solution of $\boldsymbol{F}$ is $\boldsymbol{F}=\boldsymbol{f}^{+}(\boldsymbol{\lambda})\boldsymbol{Y}^{\mathrm{T}}$ where $\boldsymbol{f}^{+}(\boldsymbol{\lambda})$ is Moore-Penrose pseudo-inverse of $\boldsymbol{f}(\boldsymbol{\lambda})$.

5) Achieve optimal modes $\boldsymbol{\lambda}^*$ by solving the following variable projection function

$$\min\|\boldsymbol{r}(\boldsymbol{\lambda})\|_2^2=\min\|\boldsymbol{Y}^{\mathrm{T}}-\hat{\boldsymbol{Y}}\|_2^2=\min\|\boldsymbol{Y}^{\mathrm{T}}-\boldsymbol{f}(\boldsymbol{\alpha})\boldsymbol{f}^{+}(\boldsymbol{\alpha})\boldsymbol{Y}^{\mathrm{T}}\|_2^2 \tag{1}$$

6) Calculate the frequency f_i and damping ζ_i of the ith achieved optimal mode λ_i^*

Detailed procedures of the mode estimation using VarPro can be found in [58]. Note that solving the variable projection function (1) may be computational intensive since it is subject to the initially given modes λ_is. Although a good λ_is value can significantly accelerate the solution efficiency and improve the accuracy of variable projection function (1), there is currently no effective method to directly determine what an appropriate λ_is is. Moreover, VarPro is incapable of estimating mode shapes, participation factors, and coherent groups, which also provide critical dynamic information for grid operators to make decisions. Therefore, an OptVarPro-based mode, mode shape, participation factor, and coherent group estimations are developed in the following sections.

2.2. OptVarPro based Mode Estimation

As mentioned in Section 2.1, the computational efficiency and accuracy of the VarPro-based mode estimation are subject to the initially given modes α_is in the variable projection function (1). To address this problem, an OptVarPro-based method proposed in [63] is adopted to estimate the initially given modes α_is and improve the computational efficiency of the variable projection function (1).

Assume the adjacent measured data sequences $\boldsymbol{y}_m$ and $\boldsymbol{y}_{m+1}$ are generated from a discrete-time linear dynamical system, then $\boldsymbol{y}_m$ and $\boldsymbol{y}_{m+1}$ satisfy the linear mapping $\boldsymbol{y}_{m+j+1}=\boldsymbol{O}\boldsymbol{y}_{m+j}+\boldsymbol{r}_0$ where $\boldsymbol{O}$ is a time-independent operator matrix capturing the inherent dynamics of the studied system in this matrix, and $\boldsymbol{r}_0$ is the measurement noise. The linear mapping $\boldsymbol{y}_{m+j+1}=\boldsymbol{O}\boldsymbol{y}_{m+j}+\boldsymbol{r}_0$ is a typical linear approximation, where eigenvalues of $\boldsymbol{O}$ and their corresponding eigenvectors can represent the features of the dynamic system.

Further, in the noise-free case, the measurements $\boldsymbol{Y}=[\boldsymbol{y}_1, \boldsymbol{y}_2,\cdots, \boldsymbol{y}_m]^{\mathrm{T}}$ yields the Krylov sequences by using $\boldsymbol{Y}_1=[\boldsymbol{y}_1, \boldsymbol{y}_2,\cdots, \boldsymbol{y}_m]^{\mathrm{T}}=[\boldsymbol{y}_1, \boldsymbol{O}\boldsymbol{y}_1,\cdots, \boldsymbol{O}^{m-1}y_1]^{\mathrm{T}}$ and $\boldsymbol{Y}_2=[\boldsymbol{y}_2, \boldsymbol{y}_3,\cdots, \boldsymbol{y}_m]^{\mathrm{T}}=[\boldsymbol{O}\boldsymbol{y}_1, \boldsymbol{O}^2\boldsymbol{y}_1,\cdots, \boldsymbol{O}^m\boldsymbol{y}_1]^{\mathrm{T}}$. Hence, the $\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$ satisfy $\boldsymbol{Y}_2=\boldsymbol{O}\boldsymbol{Y}_1$

If the $\boldsymbol{y}_m$ is sampled by following the trapezoidal rule, $\boldsymbol{y}_m$ and $\boldsymbol{y}_{m+1}$ should fulfill

$$\frac{\boldsymbol{y}_{m+1}-\boldsymbol{y}_m}{t_{m+1}-t_m}=\boldsymbol{O}\frac{\boldsymbol{y}_{m+1}+\boldsymbol{y}_m}{2} \tag{2}$$

Hence, $\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$ satisfy

$$\frac{\boldsymbol{Y}_2-\boldsymbol{Y}_1}{\boldsymbol{T}}=\boldsymbol{O}\frac{\boldsymbol{Y}_1+\boldsymbol{Y}_2}{2} \tag{3}$$

where $\boldsymbol{T}$=diag(t_2-t_1, … , t_m-t_{m-1}) is a matrix with m-1 sampling intervals.

Define the matrices $\boldsymbol{\Re}_1$ and $\boldsymbol{\Re}_2$ as

$$\begin{cases} \boldsymbol{\Re}_1=\dfrac{\boldsymbol{Y}_1+\boldsymbol{Y}_2}{2} \\ \boldsymbol{\Re}_2=\dfrac{\boldsymbol{Y}_2-\boldsymbol{Y}_1}{\boldsymbol{T}} \end{cases} \tag{4}$$

After taking the singular value decomposition (SVD) of the matrix $\boldsymbol{\Re}_1$, the left singular vector matrix $\boldsymbol{U}$, singular value matrix $\boldsymbol{\Sigma}$ and right singular vector matrix $\boldsymbol{V}$ are

$$\boldsymbol{\Re}_1=\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^* \tag{5}$$

By substituting (5) into (4), the matrix $\boldsymbol{\Re}_2$ can be approximated by

$$\boldsymbol{\Re}_2 \approx \boldsymbol{O}\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^* \tag{6}$$

Further, a low order state matrix $\boldsymbol{A}$ can be defined as

$$\boldsymbol{A}=\boldsymbol{U}^*\boldsymbol{O}\boldsymbol{U} \tag{7}$$

According to (6) and (7), $\boldsymbol{A}$ can be rewritten as

$$\boldsymbol{A} \approx \boldsymbol{U}^*\boldsymbol{\Re}_2\boldsymbol{V}\boldsymbol{\Sigma}^{-1} \tag{8}$$

Theoretically, the critical oscillation modes embedded in the response measurements are contained in the eigenvalues of $\boldsymbol{A}$. Hence, the eigenvalues of $\boldsymbol{A}$ are

$$\boldsymbol{A}\boldsymbol{\varphi}_i=\lambda_i\boldsymbol{\varphi}_i \tag{9}$$

where λ_i is the eigenvalue of $\boldsymbol{A}$, and $\boldsymbol{\varphi}_i$ is the eigenvector of λ_i.

By employing the eigenvalues of $\boldsymbol{A}^*$ as the initially given modes $\boldsymbol{\lambda}$ in (1), the variable projection function (1) can be solved. Also, the optimal oscillation modes $\boldsymbol{\lambda}^*$ embedded in the response measurement, which can reveal the electromechanical oscillation property of the power systems, are obtained.

Note that to suppress the measurement noise and improve the accuracy of the oscillation mode estimation, additional trivial modes are intentionally included in the solved optimal oscillation modes $\boldsymbol{\lambda}^*$. Thus, the dominant oscillation modes are mixed with the trivial modes in the results that are solved by the proposed OptVarPro method. The dominant modes can reveal some system dynamic features. However, the trivial modes are only utilized to suppress the measurement noise. To avoid false alarms and separate the dominant modes from the trivial modes, cumulative energy weights $E_{s,k}$ is defined as [59]

$$E_{s,k}=\sum_{i=1}^{k}\frac{E_i}{\sum_{i=1}^{m}E_i}\times 100\% \tag{10}$$

where t$E_{s,k}$ is the cumulative energy weight of the ranked E_is, E_i =||$\boldsymbol{F}$* i ||2 2 is the energy weight of the i-th mode of the $\boldsymbol{\lambda}^*$, $\boldsymbol{F}$* i is the ith row vector of $\boldsymbol{F}^*$.

Give a cumulative energy weight threshold E_0, if $E_{s,k}$ is greater than E_0, the first k modes can be considered as the dominant oscillation modes. Otherwise, more modes should be considered. The E_0 plays a critical role in the dominant mode separation. A larger E_0 may make the separated dominant modes mix with the trivial modes, while a smaller E_0 may miss some dominant modes. To ensure the separation accuracy, the value of E_0 in this work is set to be 85%, and its performance will be demonstrated in the case studies section.

2.3. Mode Shape Estimation

Mode shape is another important dynamic feature that can provide additional information for electromechanical oscillation mitigations. In this work, the mode shapes of the i-th dominant mode estimated by the proposed OptVarPro can be achieved from the solved $\boldsymbol{F}^*$ in (1).

$$\boldsymbol{\phi}_i=\frac{\boldsymbol{F}_i^{*\mathrm{T}}}{\left\|\boldsymbol{F}_i^*\right\|_2} \tag{11}$$

where $\boldsymbol{\phi}_i$ is the mode shape vector of the i-th dominant oscillation mode.

2.4. Participation Factor Estimation

The participation factors can be used to quantitatively evaluate the degrees of generators participating in the oscillation, locate the most energetic generators associated with the focused oscillation, and select the placement of PSS to damp the oscillation. It is defined as

$$\boldsymbol{p}_i=\left|\boldsymbol{F}_i^{*\mathrm{T}}\left\|\boldsymbol{\phi}_i\left(\lambda_i^*\right)\right\|_2\right| \tag{12}$$

where $\boldsymbol{p}_i=[p_{i1}\ p_{i2}\ \ldots\ p_{ik}]^{\mathrm{T}}$ is the participation factor vector of the generators associated with the i-th dominant oscillation mode, and p_{ik} is the participation factor of kth generator associated with the i-th dominant oscillation mode.

2.5. Coherent Groups Separation

The estimated mode shapes in subsection 2.3 can identify two relative oscillation coherent groups among the whole generators in the power system under one of the dominant oscillation modes. However, they are unable to separate the coherent groups of generators subjected to the joint multiple oscillation modes. To solve this problem, a graph cut-based coherent group separation is developed in this work. Following the general graph cut method in [64], a virtual undirected graph $\boldsymbol{G}$, which can represent the coherency among the whole generators, is defined. In the defined virtual undirected graph, the vertex of $\boldsymbol{G}$ is defined as $\boldsymbol{V}=\{1,2,\cdots,n\}$, where n is the number of generators. The edge is defined as $\boldsymbol{E}\subset\boldsymbol{V}\times\boldsymbol{V}$, where $e(i, j)\in \boldsymbol{E}$ indicates the coherency degree between generators i and j. Further, the vertex-weight and edge-weight are represented as follows

Vertex-Weights: the weight of the vertex i in $\boldsymbol{G}$ is defined as

$$w_{ij} = 1 \tag{13}$$

Edge-Weights: the weight of the edge $e(i, j)$, $i\neq j$, is defined as

$$w_{ij} = \begin{cases} r_{ij},\ r_{ij} \geq 0 \\ 0,\ r_{ij} < 0 \end{cases} \tag{14}$$

where r_{ij} is the direction cosine between generators i and j and is calculated as

$$r_{ij} = \frac{\sum_{k=1}^{m}(\varphi_{ki}-\bar{\varphi}_i)(\varphi_{kj}-\bar{\varphi}_j)}{\sqrt{\sum_{k=1}^{m}(\varphi_{ki}-\bar{\varphi}_i)^2\sum_{k=1}^{m}(\varphi_{kj}-\bar{\varphi}_j)^2}} \tag{15}$$

By employing the defined vertex weight and edge weight, the unnormalized Laplacian matrix $\boldsymbol{L}$ of $\boldsymbol{G}$ can be formed, and element $L_{i,j}$ in $\boldsymbol{L}$ is equal to

$$\boldsymbol{L}(i, j) = \begin{cases} d_i - w_{ii},\ i = j \\ -w_{ij},\ i \neq j \end{cases} \tag{16}$$

where d_i is the weighting factor of the vertex i, and is defined as

$$d_i = \sum_{j=1}^{n} w_{ij} \tag{17}$$

According to the formulated unnormalized Laplacian matrix $\boldsymbol{L}$, the normalized Laplacian matrix $\boldsymbol{L_N}$ of $\boldsymbol{G}$ is

$$\boldsymbol{L}_N = \boldsymbol{D}^{1/2}\boldsymbol{L}\boldsymbol{D}^{-1/2} \tag{18}$$

where $\boldsymbol{D}$ is a diagonal matrix, and it can be represented as $\boldsymbol{D}$=diag[d_1, d_2, $\cdots$, d_n].

For the formulated normalized Laplacian matrix $\boldsymbol{L_N}$, the graph-cut problem of the $\boldsymbol{G}$ can be solved by spectral clustering, and the coherent groups can be clustered. The detailed spectral clustering-based graph-cut solution can be found in [64]. Note that in the spectral clustering-based graph-cut method, the number of the coherent groups should be predetermined. Various criteria have been used to select the number of the coherent groups, e.g., the largest eigengap, maximum relative eigengap. In this paper, the largest eigengap of $\boldsymbol{L_N}$ is adopted to determine the number of coherent groups and to separate the coherent groups from measurement data.

2.6. Procedures of the Developed OptVarPro Solution

In this section, the flowchart of using the proposed OptVarPro method to extract electromechanical oscillation modes, mode shapes, participation factors, and coherent groups from measurement responses is shown in Fig.1. The major steps are summarized as follows:

1) Gather the response measurements from PMU, and form the measurement matrix $\boldsymbol{Y}$=[$\boldsymbol{y}_1$, $\boldsymbol{y}_2$,$\cdots$, $\boldsymbol{y}_m$]$^{\mathrm{T}}$.
2) Extract matrices $\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$ from $\boldsymbol{Y}$, and calculate matrices $\mathfrak{R}_1$ and $\mathfrak{R}_2$ via (4).
3) Take the SVD of $\mathfrak{R}_1$ to obtain matrices $\boldsymbol{U}$, $\boldsymbol{\Sigma}$, and $\boldsymbol{V}$ in (5).
4) Estimate the dynamic matrix $\boldsymbol{A}$ through (8) by using the matrices $\mathfrak{R}_2$, $\boldsymbol{U}$, $\boldsymbol{\Sigma}$, and $\boldsymbol{V}$.

5) Achieve the optimal modes **λ*** and ***F**** by solving variable projection function (1) with the eigenvalues of ***A*** used as initially given modes **λ.**

6) Distinguish the dominant oscillation modes from the achieved optimal modes **λ*** via the cumulative energy weights in (10).

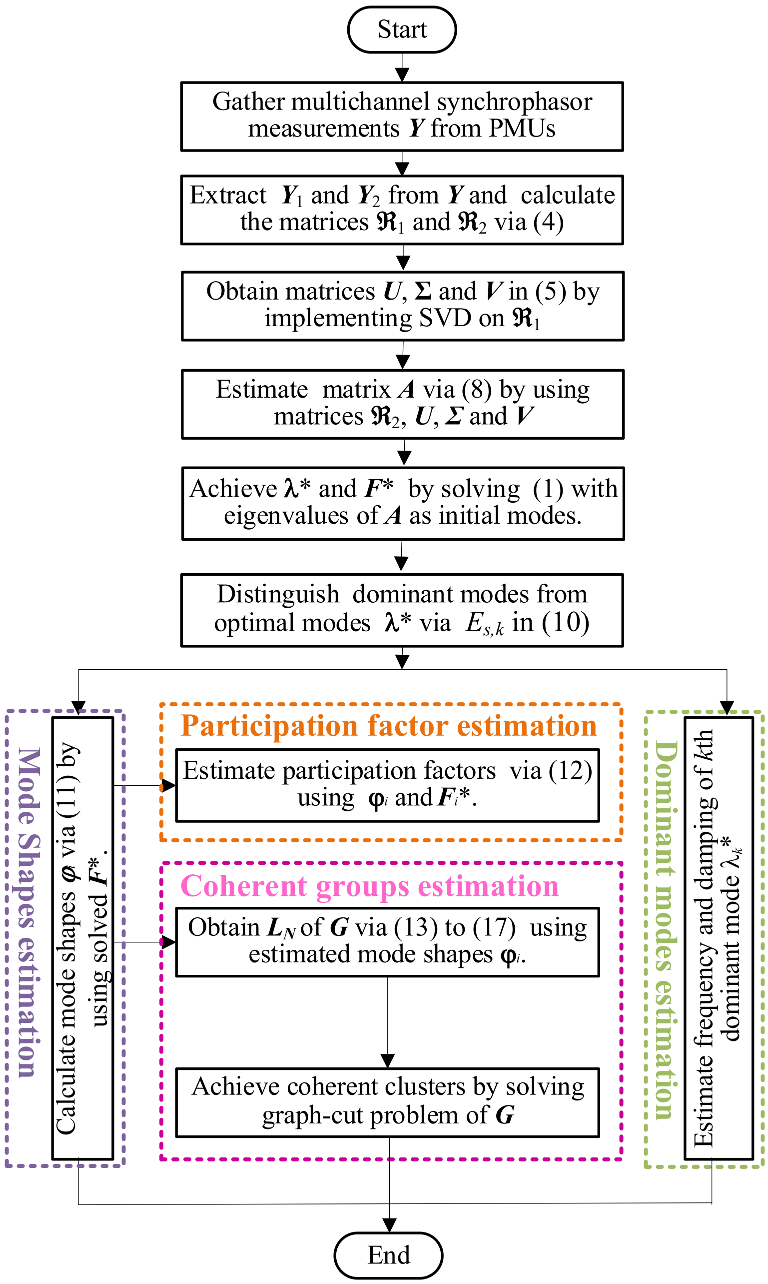


Fig. 1 Flowchart of the proposed OptVarPro Solution.

7) Calculate the mode shapes associated with the distinguished dominant oscillation modes through (11).

8) Estimate the participation factors associated with the distinguished dominant oscillation modes via (12).

9) Calculate the direction cosine r_{ij} between the generators i and j via (15) by using the estimated mode shapes.

10) Form the unnormalized Laplacian matrix $\boldsymbol{L}$ of the virtual undirected graph $\boldsymbol{G}$ via (16) by using the defined vertex-weights and edge-weights.

11) Normalize the Laplacian matrix $\boldsymbol{L}$ to obtain the normalized Laplacian matrix $\boldsymbol{L}_N$ via (18).

12) Solve the graph-cut problem of the $\boldsymbol{G}$ with the normalized $\boldsymbol{L}_N$ by using the spectral clustering method.

13) Achieve coherent clusters.

3. Case studies

In this section, the performance of the proposed OptVarPro-based solution in estimating the dominant oscillation modes, mode shapes, participation factors, and coherent groups is evaluated through the simulated measurements of a 16-generator 68-bus test system and the real measurements from the PMUs in Yunnan Power Grid (YPG).

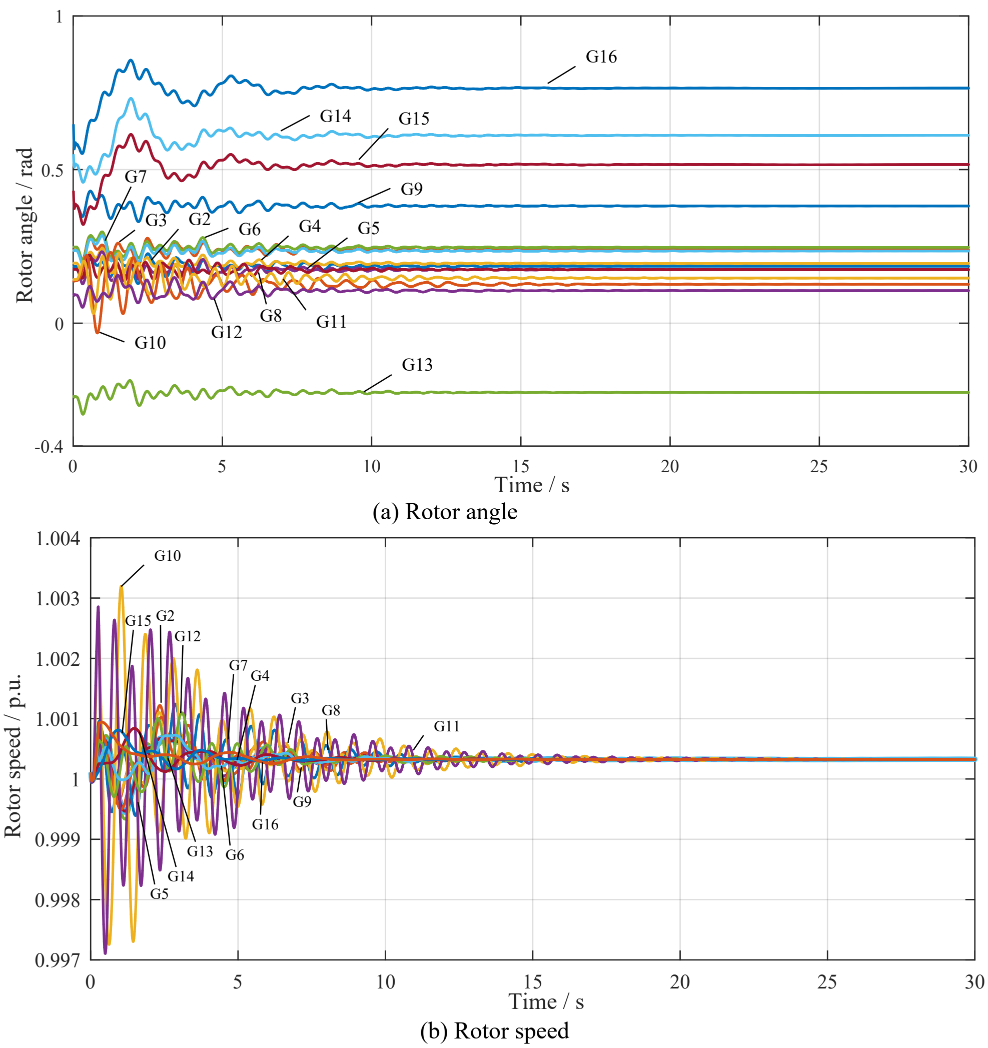


Fig. 2 Swing rotor angles and speed of generators in the 16-generator 68-bus test system.

3.1. 16-generator 68-bus Test System

In the 16-generator 68-bus test system, generators G1-G9, G10-G13, G14, G15, and G16 are located in five different areas, respectively. Detailed parameters of the system are available in [4]. To mimic the response collected from PMUs, a 3-phase fault that occurred at the end of branch 46-49 is added. Then branch 46-49 is tripped after 0.12s to clear this fault. The swing rotor angles and speed of the generators during the disturbance are illustrated in Fig.2.

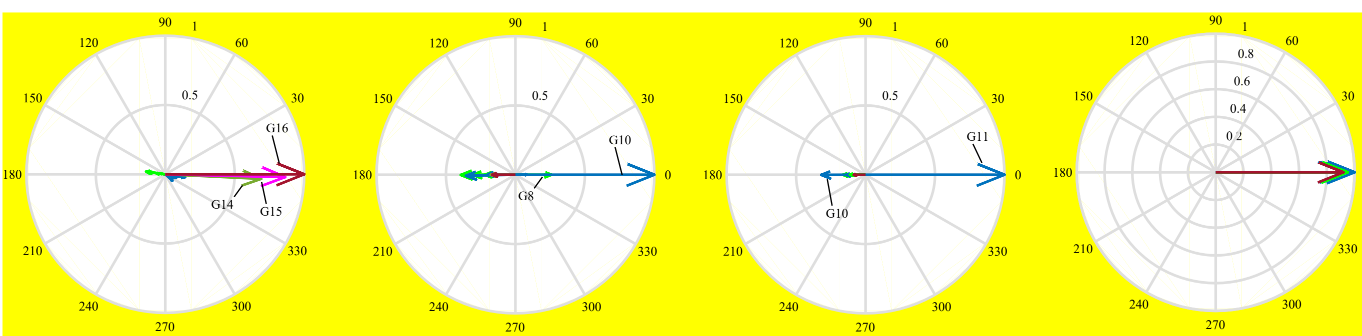
3.1.1 Electromechanical oscillation properties estimated by EA

Although there are 15 natural oscillation modes in this test system, the previous works have demonstrated that four dominant modes whose oscillation frequencies are around 0.30Hz, 1.15Hz, 1.60Hz, and 2.10Hz, may be excited by the branch 46-49 tripping [39], [59]. To fully evaluate the performance of the proposed OptVarPro-based data analytics solution, the results of EA are used as the baseline in this work. Table 1 lists these four dominant modes, and the corresponding mode shapes, participation factors, and coherent groups estimated by EA in detail. The estimated mode shapes, participation factors, and coherent groups are also plotted in Figs. 3, 4, and 5.

Table 1 Modes, mode shapes, participation factors, and coherent groups estimated by EA

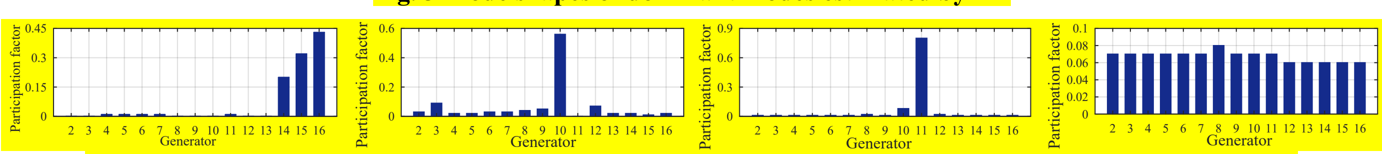
	Mode 1			Mode 2			Mode 3			Mode 4		
	f=0.29Hz, ζ=19.49%			f=1.14Hz, ζ=3.51%			f=1.59Hz, ζ=2.52%			f=2.10Hz, ζ=1.90%		
	Mode shape		PF	Mode shape		PF	Mode shape		PF	Mode shape		PF
	Mag.	Ang./°		Mag.	Ang./°		Mag.	Ang./°		Mag.	Ang./°	
G2	0.06	174.12	0.00	0.22	179.92	0.03	0.11	-179.90	0.01	0.93	0.00	0.07
G3	0.07	173.26	0.00	0.39	-179.81	0.09	0.11	-179.91	0.01	0.93	0.00	0.07
G4	0.12	171.86	0.01	0.17	-179.99	0.02	0.10	-179.92	0.01	0.94	0.00	0.07
G5	0.14	171.85	0.01	0.21	179.98	0.02	0.08	-179.94	0.01	0.93	0.00	0.07
G6	0.13	171.82	0.01	0.22	179.98	0.03	0.09	-179.93	0.01	0.93	0.00	0.07
G7	0.13	171.83	0.01	0.21	179.98	0.03	0.09	-179.93	0.01	0.93	0.00	0.07
G8	0.03	172.31	0.00	0.26	-0.11	0.04	0.16	-179.89	0.02	1.00	0.00	0.08
G9	0.11	171.36	0.00	0.31	-180.00	0.05	0.09	-179.94	0.01	0.94	0.00	0.07
G10	0.11	-8.69	0.00	1.00	0.00	0.56	0.32	-179.87	0.08	0.93	0.00	0.07
G11	0.14	-8.25	0.01	0.08	0.61	0.00	1.00	0.00	0.80	0.95	0.00	0.07
G12	0.08	-19.57	0.00	0.36	-178.85	0.07	0.14	-179.71	0.02	0.92	0.00	0.06
G13	0.07	-42.11	0.00	0.20	178.78	0.02	0.11	179.18	0.01	0.92	0.00	0.06
G14	0.69	-2.74	0.20	0.18	-180.00	0.02	0.09	-179.93	0.01	0.92	0.00	0.06
G15	0.86	-1.35	0.32	0.16	-179.93	0.01	0.08	-179.94	0.01	0.92	0.00	0.06
G16	1.00	0.00	0.43	0.17	179.94	0.02	0.08	179.99	0.01	0.92	0.00	0.06
Coherent groups	C1={G2、G3、G4、G5、G6、G7、G9}, C2={G8、G10、G11}, C3={G12、G13、G14、G15、G16}											

In Table 1, four dominant modes, which are 0.29Hz with 19.49% damping, 1.14Hz with 3.51% damping, 1.59Hz with 2.52% damping, and 2.10Hz with 1.90% damping, are estimated. From the mode shapes in Table 1 and Fig. 3, it could be seen that mode 1 is an inter-area oscillation mode, i.e., generators G2-G9 in area 1 swinging against generators G10-G16 in the rest areas, and generators G14, G15, and G16 have the highest degree of observability in this oscillation. Mode 2 is representative of generators G8, G10, and G11 swinging against generators G2, G3, G4, G5, G6, G7, G9, G12, G13, G14, G15, and G16. This mode has strong observability at generator G10. Similarly, Table 1 and Fig. 3(c) show generator G11 oscillates against the rest generators in mode 3, and generators G10 and G11 have significant observability in this single mode. Lastly, Table 1 and Fig. 3(d) illustrate mode 4 has a unique shape with generators G2-G16 swinging against generator G1.



(a) Mode 1 (b) Mode 2 (c) Mode 3 (d) Mode 4

Fig. 3 Mode shapes of dominant modes estimated by EA



(a) Mode 1 (b) Mode 2 (c) Mode 3 (d) Mode 4

Fig. 4 Participation factors of dominant modes estimated by EA

Table 1 and Fig.4 further show the estimated participation factors of generators estimated by EA. It is observed that generators G14, G15, and G16 have relatively higher participation factors in mode 1, and the largest participation factors of modes 2, 3, and 4 are associated with G10, G11, and G8, respectively.

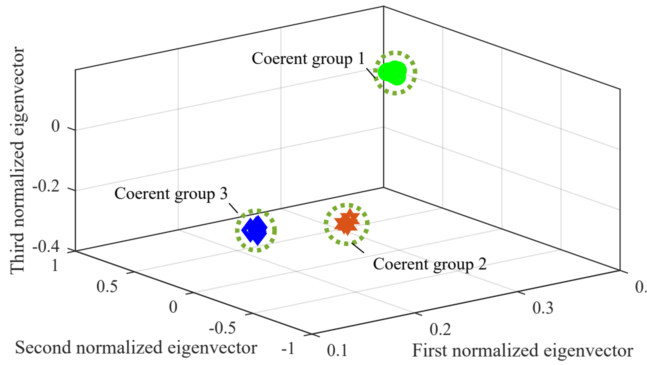


Fig.5 Coherent groups of generators separated by EA

From the estimated mode shapes in Table 1 and Fig.3, the coherent groups of the generators are separated by using the proposed spectral clustering-based graph-cut solution discussed in section 2.5. The results are listed in Table 1 and clustered in Fig. 5. Three coherent groups, including C1={G2、G3、G4、G5、G6、G7、G9}, C2={G8、G10、G11} and C3={G12、G13、G14、G15、G16}, are separated by EA.

Table 2 Estimated modes and corresponding energy

Mode	f /Hz	ζ /%	E_i	$E_{s,k}$/%
Mode 1	**0.30**	**19.58**	**0.019**	**54.92**
Mode 2	**1.14**	**3.54**	**0.006**	**72.24**
Mode 3	**1.60**	**2.51**	**0.002**	**79.17**
Mode 4	**2.09**	**1.89**	**0.003**	**87.47**
Mode 5	0.43	13.49	0.001	91.96
Mode 6	0.57	9.67	0.001	94.94
Mode 7	0.94	4.37	0.001	96.83
Mode 8	1.22	3.29	0	98.06
Mode 9	1.02	4.64	0	99.19
Mode 10	0.71	9.94	0	99.75
Mode 11	0.87	4.66	0	99.98
Mode 12	1.37	4.22	0	100.0

3.1.2 Dominant oscillation mode estimation

The mimicked generator rotor angles and speed in Fig.2 are used as the input data to the proposed OptVarPro-based data analytics method. Then the optimal $\boldsymbol{\lambda}^*$ and $\boldsymbol{F}^*$ in (1) are solved. The oscillation energy and cumulative energy weight of each λ^* are calculated, and the results are listed in Table 2. According to the predefined threshold of cumulative energy weights, four modes with frequencies at 0.30 Hz, 1.14 Hz, 1.60 Hz, and 2.09 Hz, are considered as the dominant oscillation modes embedded in the mimicked response measurements.

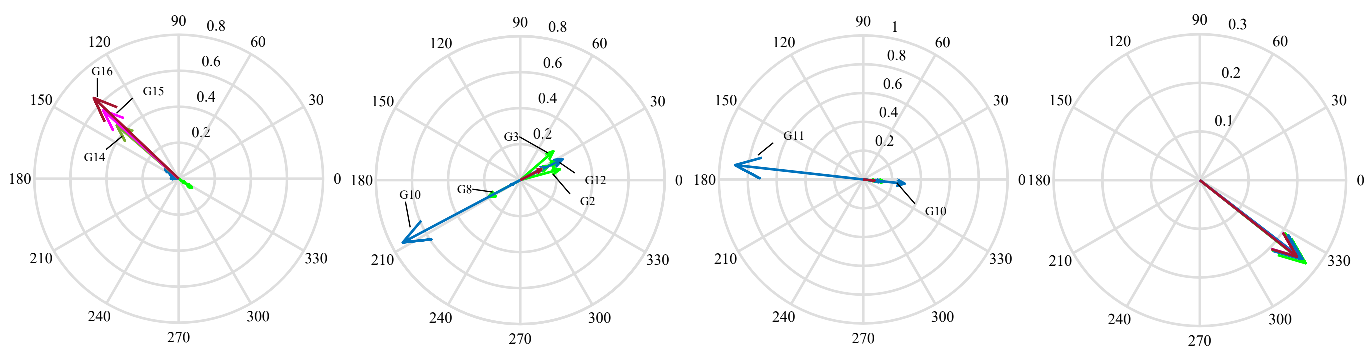
Table 3 Comparison of the dominant oscillation modes estimated by different methods

	Mode 1				Mode 2				Mode 3				Mode 4			
	frequency		damping		frequency		damping		frequency		damping		frequency		damping	
	f/Hz	Error /%	ζ/%	Error /%	f/Hz	Erro r/%	ζ /%	Error /%	f/Hz	Error /%	ζ/%	Error /%	f/Hz	Error /%	ζ/%	Error /%
OptVarPro	0.30	3.45	19.58	0.46	1.14	0.00	3.54	0.85	1.60	0.63	2.51	0.40	2.09	0.48	1.89	0.53
DMD	0.24	17.24	40.89	109.80	1.13	0.88	4.50	28.21	1.64	3.14	5.28	109.52	2.10	0.00	2.49	31.05
POD	0.29	0.00	-	-	1.14	0.00	-	-	1.60	0.63	-	-	2.09	0.48	-	-
KMD	0.31	6.90	7.35	62.29	1.15	0.88	9.43	168.66	1.58	0.63	10.72	325.40	2.05	2.38	12.54	560.00

Table 3 compares the dominant oscillation modes estimated by OptVarPro, DMD, POD, and KMD. Moreover, the results from EA in Tabel 1 are served as the baseline. It is observed from Table 3 that the maximum estimation errors of OptVarPro, DMD, POD, and KMD are 3.54%, 109.80%, 0.63%, and 560.00%, respectively. The POD exhibits high accuracy in oscillation frequency estimation but fails in the damping estimation. The DMD and KMD can estimate the damping and frequency of the dominant modes from measurements, but the results from DMD and KMD are not very accurate compared to the baseline. Therefore, the accuracy of the proposed OptVarPro-based data analytics solution in dominant oscillation mode estimation is validated.

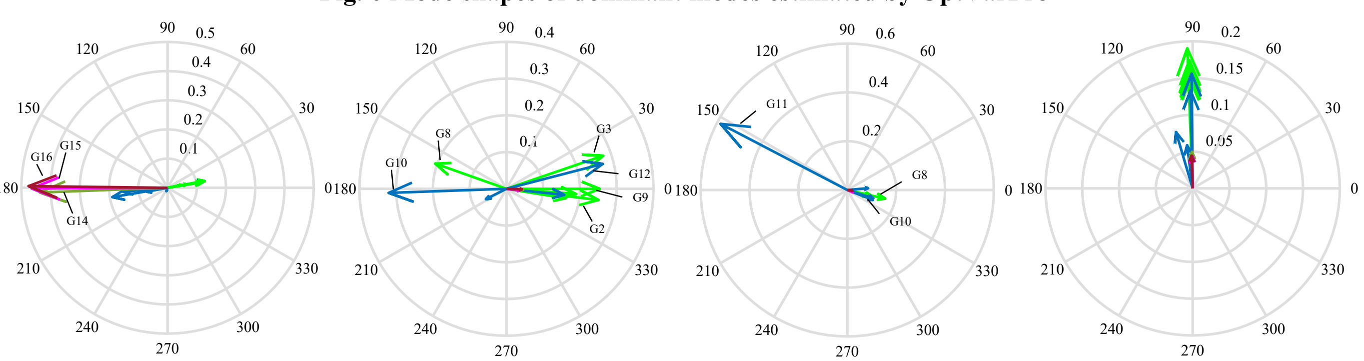
3.1.3 Mode shape estimation

After determining the dominant modes, the corresponding mode shapes of all the generators are estimated via (11) using the solved ***F****. The results are shown in Fig. 6, and the detailed mode shapes containing magnitudes and angles are listed in Table 4. From Fig. 6(a) and Table 4, it could be seen that mode 1 is an inter-area oscillation mode, i.e., generators G2-G9 in area 1 swinging against generators G10-G16 in the rest areas. Further, generators G14, G15, and G16 have the highest degree of observability in this oscillation. The mode shapes in Table 4 and Fig. 6(b) reveals that mode 2 is representative of generators G8, G10, and G11 swinging against generators G2, G3, G4, G5, G6, G7, G9, G12, G13, G14, G15, and G16 in the test system. This mode exhibits strong observability at generator G10. Similarly, Table 4 and Fig. 6(c) show generator 11 oscillates against the rest generators in mode 3, and generators G10 and G11 have significant observability in this single mode. Lastly, Table 4 and Fig. 6(d) illustrate mode 4 has a unique shape with generators G2-G16 swinging against generator G1.



(a) Mode shapes of mode 1 (b) Mode shapes of mode 2 (c) Mode shapes of mode 3 (d) Mode shapes of mode 4

Fig. 6 Mode shapes of dominant modes estimated by OptVarPro



(a) Mode shapes of mode 1 (b) Mode shapes of mode 2 (c) Mode shapes of mode 3 (d) Mode shapes of mode 4

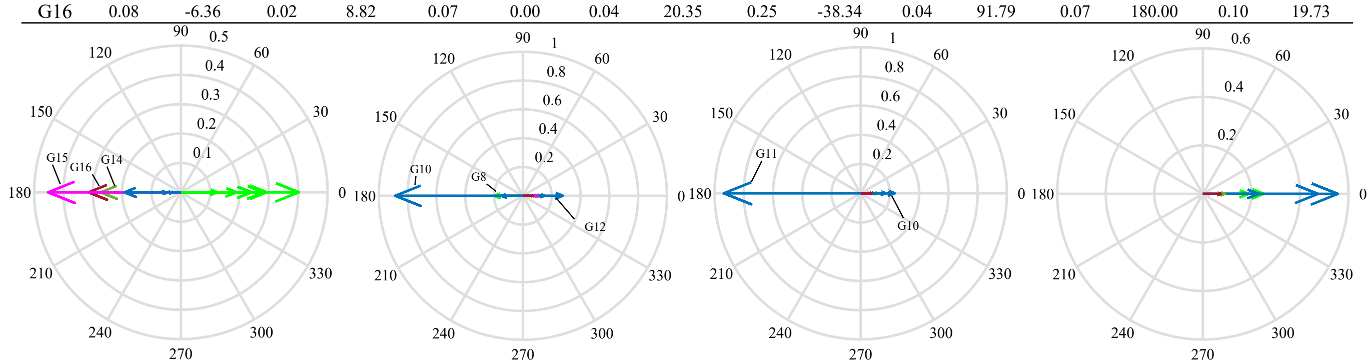
Fig. 7 Mode shapes of dominant modes estimated by DMD

The mode shapes of all the generators associated with the detected four dominant modes are also estimated by DMD, POD, and KMD. The estimated mode shapes are plotted in Figs. 7, 8, and 9, and the complete mode shape information is listed in Table 4. By comparing these graphs and values with the results estimated by OptVarPro in Fig. 3, Fig. 6, and Table 1, it is observed that the coherent generators and the generators with strong observability in these four modes are all the same. Even though the magnitudes of the estimated mode shapes are slightly different, the accuracy of the proposed OptVarPro in the mode shape estimation is proven.

Table 4 Comparison of the mode shapes estimated by different methods

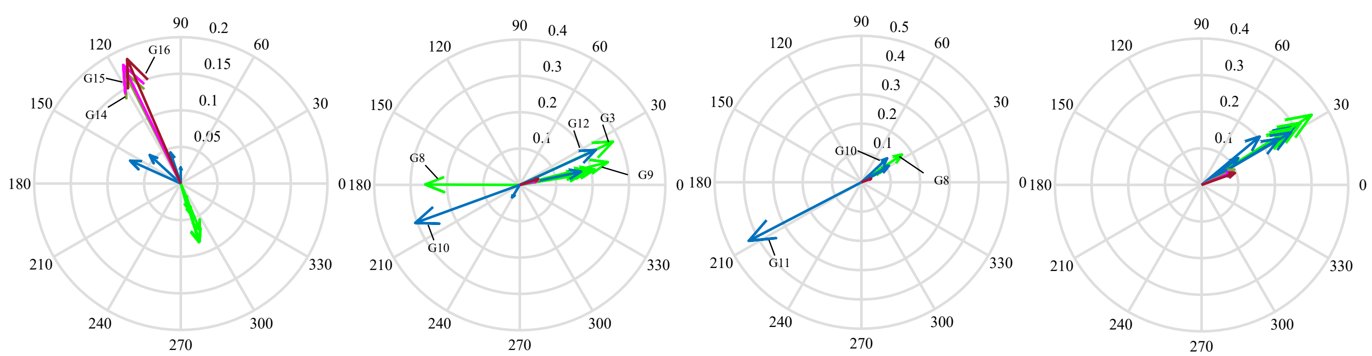
Gen.	Mode 1								Mode 2							
	OptVarPro		DMD		POD		KMD		OptVarPro		DMD		POD		KMD	
	Mag.	Ang./°	Mag.	Ang./°	Mag.	Ang./°	Mag.	Ang./°	Mag.	Ang./°	Mag.	Ang./°	Mag.	Ang./°	Mag.	Ang./°
G2	0.04	-38.03	0.05	11.77	0.19	0.00	0.03	-74.98	0.23	15.26	0.25	-6.93	0.12	0.00	0.21	11.60
G3	0.04	-34.22	0.07	10.43	0.24	0.00	0.04	-72.48	0.24	41.08	0.28	18.88	0.08	0.00	0.28	24.93
G4	0.08	-34.97	0.13	10.31	0.24	0.00	0.08	-72.12	0.13	27.95	0.16	-5.85	0.05	0.00	0.16	11.42
G5	0.09	-35.22	0.13	10.66	0.24	0.00	0.08	-72.78	0.16	26.53	0.16	-9.02	0.09	0.00	0.17	9.46
G6	0.08	-35.50	0.13	10.27	0.29	0.00	0.08	-72.01	0.17	26.37	0.19	-4.34	0.05	0.00	0.19	11.79
G7	0.08	-35.48	0.13	10.28	0.28	0.00	0.08	-72.01	0.17	26.32	0.19	-4.15	0.05	0.00	0.19	11.96
G8	0.02	-35.10	0.04	11.70	0.12	0.00	0.02	-70.34	0.20	-152.02	0.21	160.48	0.20	180.00	0.26	179.76
G9	0.07	-34.64	0.12	8.97	0.40	0.00	0.07	-67.96	0.24	27.26	0.25	-0.03	0.03	0.00	0.25	14.50
G10	0.07	146.42	0.01	-108.62	0.07	180.00	0.02	90.28	0.74	-151.98	0.32	-178.01	0.89	180.00	0.31	-159.98
G11	0.09	145.11	0.07	-169.32	0.05	180.00	0.04	108.37	0.06	-153.42	0.06	-153.39	0.15	180.00	0.04	-121.63
G12	0.05	155.11	0.15	-174.95	0.03	180.00	0.06	137.43	0.26	26.23	0.27	14.50	0.15	0.00	0.23	24.70
G13	0.04	179.24	0.19	-170.43	0.31	180.00	0.08	155.64	0.15	28.65	0.16	-6.02	0.28	0.00	0.17	12.21
G14	0.45	139.15	0.44	-177.88	0.27	180.00	0.17	116.80	0.14	27.38	0.05	-3.55	0.01	0.00	0.05	15.09
G15	0.56	137.74	0.46	-179.74	0.45	180.00	0.18	115.85	0.13	27.17	0.04	-8.05	0.10	0.00	0.05	13.39
G16	0.65	136.35	0.48	179.31	0.19	180.00	0.19	113.27	0.13	27.45	0.04	-1.92	0.07	0	0.05	18.01

Gen.	Mode 3								Mode 4							
	OptVarPro		DMD		POD		KMD		OptVarPro		DMD		POD		KMD	
	Mag.	Ang./°	Mag.	Ang./°	Mag.	Ang./°	Mag.	Ang./°	Mag.	Ang./°	Mag.	Ang./°	Mag.	Ang./°	Mag.	Ang./°
G2	0.10	-7.23	0.09	-14.72	0.04	0.00	0.09	32.38	0.26	-38.48	0.15	91.67	0.25	0.00	0.28	32.08
G3	0.09	-5.34	0.10	-13.38	0.03	0.00	0.10	32.84	0.26	-37.84	0.16	91.80	0.19	0.00	0.30	32.16
G4	0.09	-7.52	0.10	-14.41	0.05	0.00	0.10	32.77	0.26	-38.29	0.17	91.45	0.03	0.00	0.32	32.12
G5	0.08	-6.02	0.08	-19.21	0.05	0.00	0.07	29.42	0.26	-38.22	0.15	90.95	0.01	0.00	0.27	31.89
G6	0.08	-6.56	0.09	-16.58	0.03	0.00	0.08	31.10	0.26	-38.29	0.16	91.24	0.00	0.00	0.30	32.01
G7	0.08	-6.44	0.09	-16.44	0.03	0.00	0.09	31.20	0.26	-38.27	0.17	91.26	0.01	0.00	0.30	32.01
G8	0.15	-6.45	0.16	-12.90	0.11	0.00	0.17	34.39	0.28	-38.21	0.19	92.18	0.08	0.00	0.36	32.51
G9	0.08	-6.42	0.08	-13.88	0.02	180.00	0.08	32.36	0.26	-38.26	0.16	91.58	0.01	180.00	0.29	32.19
G10	0.29	-5.77	0.09	5.03	0.23	0.00	0.12	42.88	0.26	-38.28	0.06	97.29	0.23	0.00	0.12	36.21
G11	0.89	173.60	0.59	152.42	0.94	180.00	0.43	-152.48	0.27	-37.61	0.08	106.45	0.23	0.00	0.21	39.85
G12	0.12	-6.85	0.11	-17.57	0.09	0.00	0.11	31.14	0.25	-38.26	0.13	91.27	0.47	0.00	0.25	30.91
G13	0.10	-5.92	0.11	-21.16	0.17	0.00	0.10	28.81	0.25	-38.26	0.16	90.52	0.56	0.00	0.28	30.43
G14	0.08	-6.96	0.03	-19.38	0.02	180.00	0.03	23.80	0.25	-38.23	0.05	90.86	0.09	180.00	0.10	25.16
G15	0.07	-6.99	0.02	-21.70	0.06	0.00	0.03	20.14	0.25	-38.27	0.04	90.56	0.03	0.00	0.09	22.44
G16	0.08	-6.36	0.02	8.82	0.07	0.00	0.04	20.35	0.25	-38.34	0.04	91.79	0.07	180.00	0.10	19.73



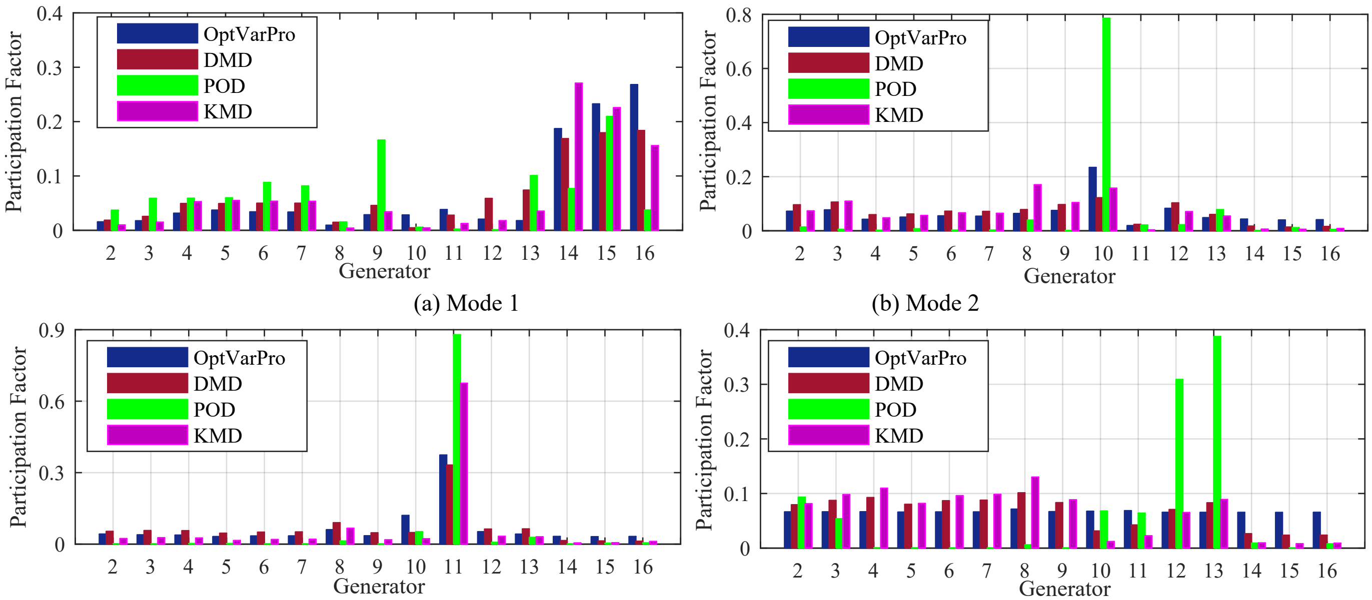
(a) Mode shapes of mode 1 (b) Mode shapes of mode 2 (c) Mode shapes of mode 3 (d) Mode shapes of mode 4

Fig. 8 Mode shapes of dominant modes estimated by POD



(a) Mode shapes of mode 1 (b) Mode shapes of mode 2 (c) Mode shapes of mode 3 (d) Mode shapes of mode 4

Fig. 9 Mode shapes of dominant modes estimated by KMD



(a) Mode 1 (b) Mode 2

(a) Mode 3 (b) Mode 4

Fig. 10 Comparison of participation factors estimated by different methods

3.1.4 Participation factor estimation

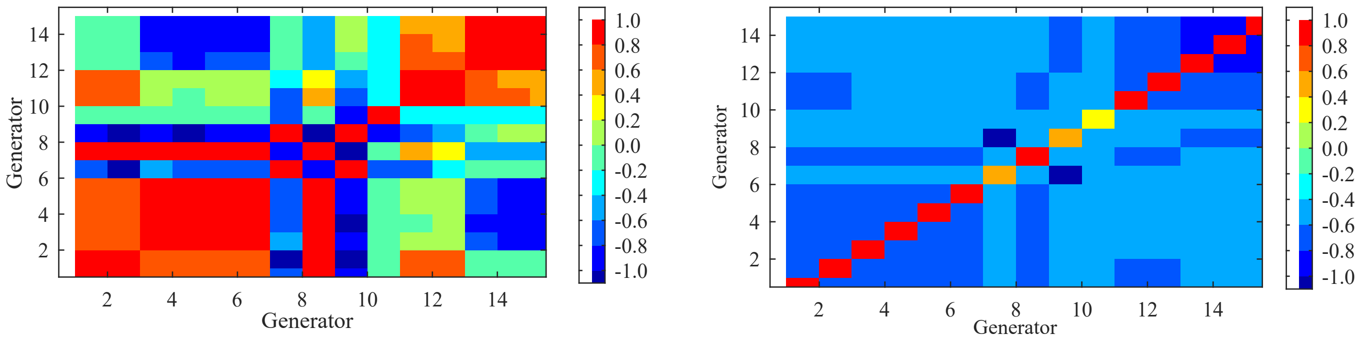
Following the procedures of the proposed OptVarPro-based data analytics solution in participation factor estimation, the achieved optimal modes $\boldsymbol{\alpha}^*$ and $\boldsymbol{B}^*$ are adopted to calculate the participation factors of all the generators associated with the dominant modes 1, 2, 3, and 4 via (12). Fig.10 shows the generator participation factors estimated by different approaches. In OptVarPro, generators G14, G15, and G16 have relatively higher participation factors in mode 1, and the largest participation factors in modes 2, 3, and 4 are associated with the generators G10, G11, and G8, respectively. Table 5 lists the dominant generators in the participation factors estimated by different approaches. By comparing the results in Tables 1 and 5, it is observed that the dominant generators identified by the proposed OptVarPro in these four modes are consistent with the dominant generators determined by EA, DMD, POD, and KMD. Still, the values of the participation factors estimated by these methods may be different, but the accuracy of the proposed OptVarPro method in estimating the participation factors of oscillation modes is proved.

Table 5 Comparison of dominant generators estimated by different methods

Method	Mode 1	Mode 2	Mode 3	Mode 4
OptVarPro	G14, G15, G16	G10	G11	G8
DMD	G14, G15, G16	G10	G11	G8
POD	G9、G15	G10	G11	G12, G13
KMD	G14, G15, G16	G8、G10	G11	G8

3.1.5 Coherent group estimation

According to Step 9) in subsection 2.6, the direction cosines among generators are calculated using the solved mode shapes and the results are shown in Fig. 11 (a). The vertex-weights and edge-weights of the virtual undirected graph $\boldsymbol{G}$ are determined via (13) and (14). By using the determined vertex-weights and edge-weights of $\boldsymbol{G}$, the normalized Laplacian matrix $\boldsymbol{L}_N$ of $\boldsymbol{G}$ is formed through (18) and illustrated in Fig.11 (b). The eigenvalues of $\boldsymbol{L}_N$ are calculated and shown in Fig. 12. From Fig. 12, the largest eigengap of $\boldsymbol{L}_N$ appears between the 4th eigenvalue and 5th eigenvalue, and the number of the coherent groups is hereby determined as 4.



(a) direction cosine matrix (b) normalized Laplacian matrix

Fig.11 Direction cosine and normalized Laplacian matrices of the 16-generator 68-bus test system

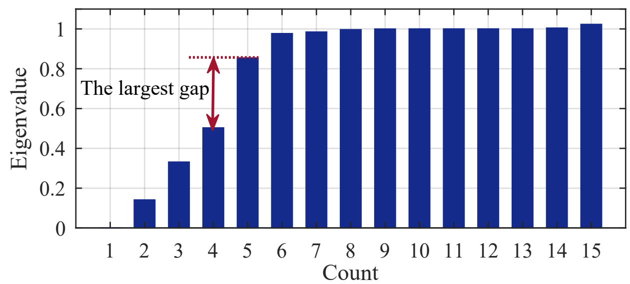
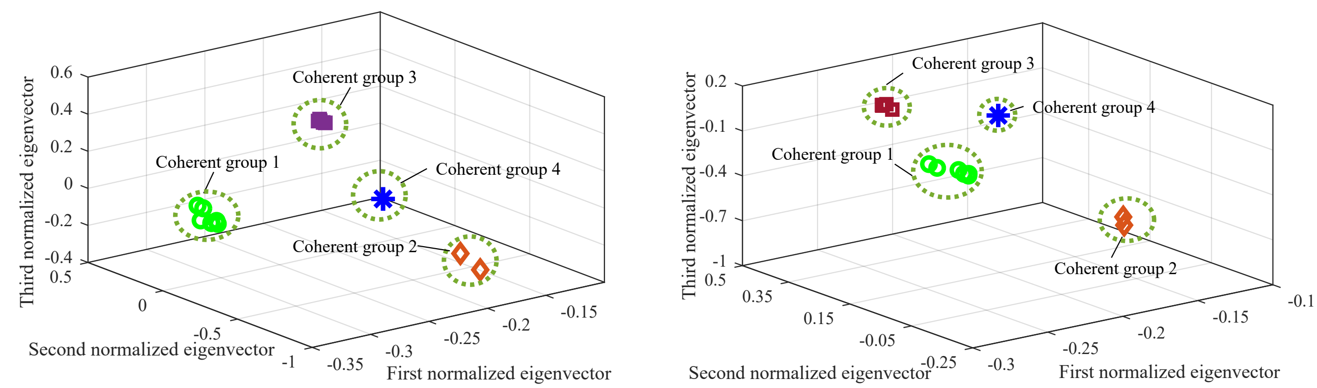
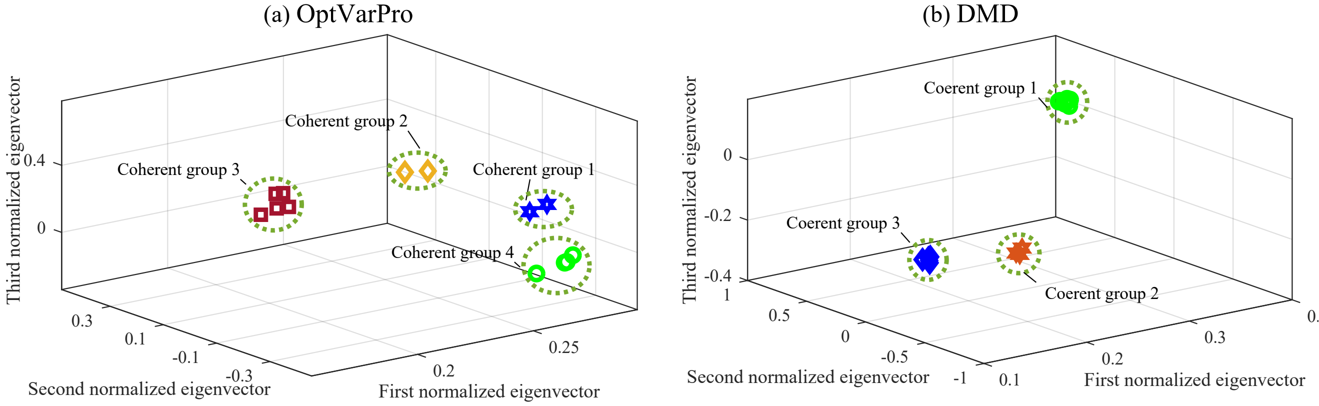


Fig.12 Eigenvalues of L_N for the 16-generator 68-bus test system



(a) OptVarPro (b) DMD



(c) POD (d) KMD

Fig.13 Coherent groups of generators separated by OptVarPro, DMD, POD, and KMD

Following the determined number of coherent groups, the coherent groups of generators are clustered using the spectral clustering method to solve the graph-cut problem of the $\boldsymbol{G}$. Fig. 13(a) shows the separation results of the coherent groups by the proposed method, and the details are listed in Table 6. As can be seen in Fig. 13(a) and Table 6, all the generators that are subject to the implemented disturbance are clustered into four coherent groups, where generators G2, G3, G4, G5, G6, G7, and G9 form the first coherent group C1, generators G8, and G10 form the second coherent group C2, generators G12, G13, G14, G15, and G16 form the third coherent group C3, and generator G11 alone forms the fourth coherent group C4.

The coherent groups clustered via the proposed graph cut method using the estimated mode shapes of DMD, POD, and KMD are also compared in Fig. 13. Detailed results are summarized in Table 6. It is observed that DMD and KMD can separate all the generators into four coherent groups and the generators in each group are identical to the coherent groups separated by OptVarPro. POD clusters all the generators into four coherent groups, and the generators composing coherent groups C2 and C3 are the same as the generators in the coherent groups C2 and C3 that are estimated by OptVarPro. However, in POD, the coherent group C1 is formed by generators G2, G3, and G11, and the coherent group C4 is formed by generators G4, G5, G6, G7, and G9. Further, by comparing Table 1 and Fig.5 with Table 6 and Fig.13, it is observed that the coherent group C2={G8、G10、G11}estimated by EA is further divided by into two coherent groups C2={G8、G10 } and C4={G11} by the proposed OptVarPro.

Table 6 Comparison of coherent groups separated by different methods

method	Coherent groups
OptVarPro	C1={G2、G3、G4、G5、G6、G7、G9}, C2={G8、G10}, C3={G12、G13、G14、G15、G16}, C4={G11}
DMD	C1={G2、G3、G4、G5、G6、G7、G9}, C2={G8、G10}, C3={G12、G13、G14、G15、G16}, C4={G11}
POD	C1={G2、G3、G11}, C2={G8、G10}, C3={G12、G13、G14、G15、G16}, C4={G4、G5、G6、G7、G9}
KMD	C1={G2、G3、G4、G5、G6、G7、G9}, C2={G8、G10}, C3={G12、G13、G14、G15、G16}, C4={G11}

From Fig. 13 and Table 6, the coherent groups separated by OptVarPro, DMD, POD, and KMD are slightly different from each other due to the various mode shapes estimated by different methods. Nevertheless, the critical generators in each coherent group that separated by OptVarPro, DMD, POD, and KMD are always the same. This indicates that the proposed OptVarPro method can precisely cluster the coherent groups of generators from measurement data.

3.1.6 Discussion on the threshold of relative energy weight

As mentioned earlier, the preset threshold of relative energy weight E_0 is a critical parameter to separate the dominant oscillation modes from trivial ones in the results of α^*. A larger E_0 may mix the trivial modes into the dominant modes while a smaller E_0 may miss the dominant modes. To obtain the appropriate E_0, the selection of E_0 is discussed in this subsection in the following four scenarios:

Scenario 1: a 3-phase fault occurs at the end of branch 1-30, then branch 1-30 is tripped after 0.13s.

Scenario 2: a 3-phase fault occurs at the end of branch 3-4, then branch 3-4 is tripped after 0.16s.

Scenario 3: 2.57 p.u. load at bus 1 is shed.

Scenario 4: 0.5 p.u. load disturbance at bus 47 is excited between 0.1s and 2s.

Scenario 5: the Gaussian white noise with signal-to-noise ratios (SNR) of 85dB, 65dB, 45dB, and 25dB, are intentionally added to the simulated data in Fig. 2, respectively.

(1) Scenario 1

Following the procedures of the proposed approach, the simulated swing rotor angles and the speed of the generators in the 16-generator 68-bus test system during the disturbance in scenario 1 are used as the input to the proposed OptVarPro method. Table 7 lists the solved $\boldsymbol{\alpha}^*$ and their corresponding energy and relate energy calculated from $\boldsymbol{B}^*$. It is observed from Table 7 that when the threshold of relative energy weight E_0 is set to 85%, four modes (i.e., 0.35 Hz mode, 1.15 Hz mode, 1.59 Hz mode, and 2.09 Hz mode) are considered as the dominant modes excited by this disturbance.

Table 7 Estimated modes and corresponding energy in scenario 1

Mode	f /Hz	ζ /%	E_i	$E_{s,k}$/%
Mode 1	**2.09**	**1.88**	**0.066 7**	**38.96**
Mode 2	**0.35**	**16.6**	**0.045 3**	**65.43**
Mode 3	**1.15**	**3.57**	**0.026 0**	**80.6**
Mode 4	**1.59**	**2.53**	**0.012 9**	**88.15**
Mode 5	0.55	10.8	0.010 6	94.34
Mode 6	0.94	4.37	0.004 6	97.02
Mode 7	1.02	4.63	0.001 9	98.14
Mode 8	1.12	5.39	0.001 5	99.04
Mode 9	1.22	3.18	0.000 8	99.51
Mode 10	0.87	4.56	0.000 8	99.96
Mode 11	1.38	3.54	0.000 1	100.0

Fig. 14 shows the calculated PSD of the frequencies via FFT using the same simulated data as input. As can be seen, there are four PSD peaks at 0.34Hz, 1.15Hz, 1.60Hz, and 2.09Hz in the spectral. Hence, four modes with frequencies at 0.35 Hz, 1.15 Hz, 1.59 Hz, and 2.09 Hz are excited in this disturbance. By comparing the detected dominant modes in Table 7 and Fig. 14, it is concluded that setting the threshold E_0 to 85% can accurately separate dominant modes from trivial ones in the results of solved $\boldsymbol{\alpha}^*$.

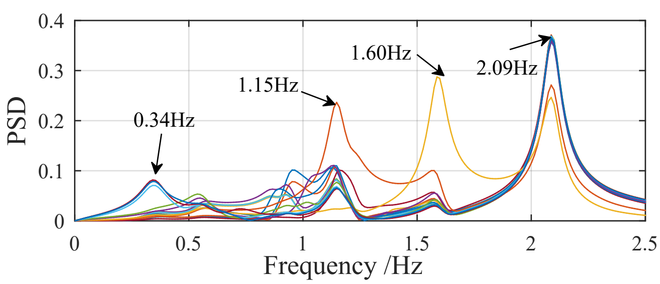


Fig.14 PSD of frequencies in scenario 1

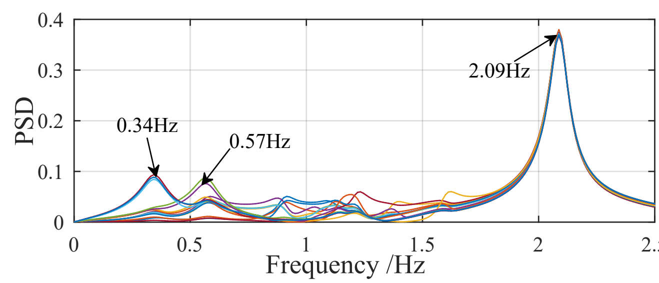


Fig.15 PSD of frequencies in scenario 2

(2) Scenario 2

The swing rotor angles and the speed of all the generators in scenario 2 are used as input to the proposed OptVarPro method. The estimated nine modes and their related energy are summarized in Table 8. When the predetermined threshold of relative energy weight E_0 is set to 85%, it is observed that the dominant modes are 0.35Hz with 16.4% damping, 0.57Hz with 10% damping, and 2.09Hz with 1.87% damping. The rest are considered as trivial modes.

Table 8 Estimated modes and corresponding energy in scenario 2

Mode	f /Hz	ζ /%	E_i	$E_{s,k}$/%
Mode 1	**2.09**	**1.87**	**0.301 8**	**42.03**
Mode 2	**0.35**	**16.4**	**0.284 5**	**81.64**
Mode 3	**0.57**	**10.0**	**0.113 7**	**97.48**
Mode 4	0.90	4.54	0.006 0	98.31
Mode 5	1.22	3.37	0.005 0	99.01
Mode 6	1.14	3.93	0.004 1	99.57
Mode 7	1.05	5.28	0.001 3	99.75
Mode 8	1.60	2.58	0.001 1	99.90
Mode 9	1.38	2.99	0.000 7	100.0

Furthermore, Fig.15 shows the calculated PSD of the simulated swing rotor angles and the speed of the generators in the 16-generator 68-bus test system during the disturbance in scenario 2 using FFT. By comparing the calculated relative energy in Table 8 and the estimated PSD in Fig.15, three dominant modes, i.e., 0.35Hz mode, 0.57Hz mode, and 2.09Hz mode, are excited by this disturbance. This demonstrates that the preset E_0 can distinguish the dominant modes from the raw estimated modes.

(3) Scenario 3

The swing rotor angles and the speed of all the generators in scenario 3 are used as the input of the proposed OptVarPro method. The estimated nine modes and their related energy are summarized in Table 9. When the predetermined threshold of relative energy weight E_0 is set to 85%, it is observed that 0.34Hz with 16.3% damping mode, 0.54Hz with 10.7% damping mode, 0.71Hz with 5.65% damping mode, and 2.08Hz with 1.89% damping mode can be viewed as the dominant modes, and the rest are considered as the trivial modes.

Table 9 Estimated modes and corresponding energy in scenario 3

Mode	f /Hz	ζ /%	E_i	$E_{s,k}$/%
Mode 1	**0.34**	**16.3**	**0.2588**	**27.78**
Mode 2	**2.08**	**1.89**	**0.2226**	**51.67**
Mode 3	**0.54**	**10.7**	**0.1968**	**72.80**
Mode 4	**0.71**	**5.65**	**0.1760**	**91.69**
Mode 5	1.21	3.36	0.0412	96.11
Mode 6	1.14	3.74	0.0189	98.14
Mode 7	1.05	5.09	0.0081	99.01
Mode 8	0.94	5.24	0.0046	99.50
Mode 9	1.60	3.35	0.0046	100.0

Fig. 16 shows the calculated PSD of the simulated swing rotor angles and the speed of the generators in the 16-generator 68-bus test system during the disturbance in scenario 3 using FFT. By comparing the calculated relative energy in Table 9 and the estimated PSD in Fig. 16, four dominant modes, i.e., 0.34Hz mode, 0.54Hz mode, 0.71Hz, and 2.08 Hz mode, are excited by this disturbance. Once again, this demonstrates that the preset E_0 can distinguish the dominant modes from the raw estimated modes.

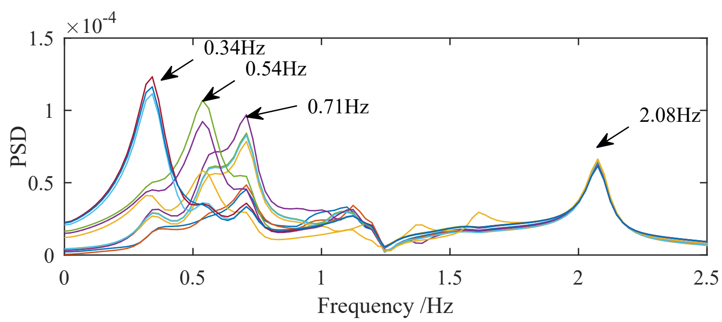


Fig. 16 PSD of frequencies in scenario 3

(4) Scenario 4

In this scenario, by following the same procedure as above, the modes and the estimated energy are shown in Table 10. According to the threshold of the relative energy weight E_0=85%, it is indicated from Table 10 that, two dominant modes, 0.33 Hz with 18.0% damping mode and 2.07 Hz with 2.24% damping mode, are identified as the dominant modes from the estimated modes.

The calculated PSD of the frequencies via the FFT using the same simulated data as input is illustrated in Fig. 17. Obviously, there are two PSD peaks associated with 0.33 Hz and 2.07 Hz in the results of FFT spectral analysis. Hence, two dominant modes with frequencies at 0.33 Hz and 2.07 Hz are excited in this disturbance. By comparing the detected dominant modes in Table 10 and Fig. 17, it is concluded that setting the threshold E_0 to 85% can accurately separate the dominant modes from the trivial ones in the results of solved $\boldsymbol{\alpha}^*$.

Table 10 Estimated modes and corresponding energy in scenario 4

Mode	f /Hz	ζ /%	E_i	$E_{s,k}$/%
Mode 1	**0.33**	**18.0**	**0.1184**	**67.73**
Mode 2	**2.07**	**2.24**	**0.0343**	**87.36**
Mode 3	1.59	2.59	0.0116	93.99
Mode 4	1.11	4.24	0.0037	96.11
Mode 5	0.94	4.66	0.0029	97.77
Mode 6	0.88	5.96	0.0022	99.03
Mode 7	0.58	10.8	0.0008	99.49
Mode 8	1.02	4.76	0.0006	99.83
Mode 9	1.38	2.20	0.0003	100.0
Mode 10	0.33	18.0	0.1184	67.73

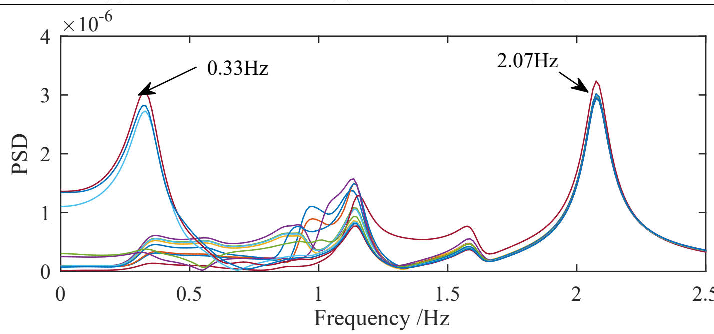


Fig. 17 PSD of frequencies in scenario 4

(5) Scenario 5

In this scenario, the Gaussian white noise with SNR=85dB, 65dB, 45dB, and 25dB is intentionally added to simulated swing rotor angles and the speed of all the generators in Fig. 2, respectively. Table 11 lists the modes and their corresponding energy and corresponding energy estimated from the simulated measurement data in Fig. 2 with different noise levels.

Table 11 Estimated modes and corresponding energy in scenario 5

Mode	85dB		65dB		45dB		25dB	
	E_i	$E_{s,k}$/ %	E_i	$E_{s,k}$/ %	E_i	$E_{s,k}$/ %	E_i	$E_{s,k}$/ %
Mode 1	**0.019**	**55.36**	**0.019**	**55.42**	**0.019**	**54.99**	**0.020**	**54.32**
Mode 2	**0.006**	**71.68**	**0.006**	**71.50**	**0.006**	**71.66**	**0.006**	**70.20**
Mode 3	**0.003**	**80.08**	**0.003**	**79.92**	**0.003**	**79.94**	**0.003**	**78.00**
Mode 4	**0.002**	**87.11**	**0.002**	**86.97**	**0.002**	**86.78**	**0.003**	**85.24**
Mode 5	0.002	91.58	0.002	91.45	0.002	91.19	0.002	89.98
Mode 6	0.001	94.57	0.001	94.44	0.001	94.17	0.002	94.72
Mode 7	0.001	96.41	0.001	96.29	0.001	96.17	0.001	98.06
Mode 8	0.000	97.67	0.000	97.58	0.000	97.46	0.001	99.45
Mode 9	0.000	98.74	0.000	98.67	0.000	98.70	0.000	99.98
Mode 10	0.000	99.29	0.000	99.23	0.000	99.34	0.000	99.99
Mode 11	0.000	99.77	0.000	99.78	0.000	99.81	0.000	100.0
Mode 12	0.000	100.0	0.000	100.0	0.000	100.0	-	-

From Table 11, even though the SNR is decreased from 85 dB to 25 dB, the determined relative energy threshold E_0=85% still exhibits strong robustness on differentiating dominant modes and trivial modes.

From the above discussions, it is concluded that setting the value of E_0 to 85% can accurately differentiate dominant modes and trivial ones in the proposed OptVarPro method.

3.1.7 Robustness of proposed OptVarPro method

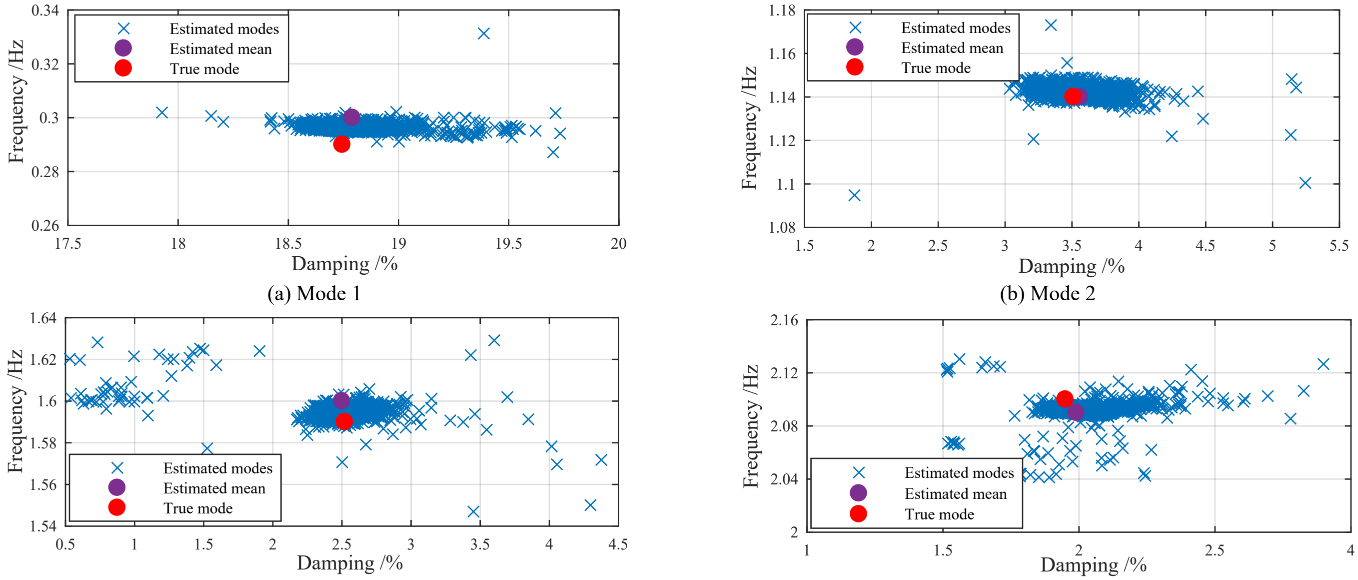
One advantage of OptVarPro over DMD, POD, and KMD is that OptVarPro is robust against measurement noise. In this subsection, the Gaussian white noise with SNR=85dB, 65dB, 45dB, 25dB, 10dB, 5dB, and 1dB is intentionally added to the simulated data in Fig. 2 to evaluate the robustness of the proposed OptVarPro against measurement noise.

Table 12 Comparison of the dominant modes estimated by different methods under different SNRs

SNR	Mode	OptVarPro				DMD				POD				KMD			
		frequency		damping		frequency		damping		frequency		damping		frequency		damping	
		f/Hz	Error /%	ζ%	Error /%	f/Hz	Error r/%	ζ/%	Error /%	f/Hz	Error /%	ζ%	Error /%	f/Hz	Error /%	ζ%	Error /%
85dB	Mode 1	0.30	0.00	19.6	0.10	0.28	16.67	86.0	110.3	0.29	0.00	-	-	0.31	0.00	8.10	10.20
	Mode 2	1.14	0.00	3.48	1.69	1.16	2.65	20.6	357.8	1.14	0.00	-	-	1.15	0.00	23.3	147.1
	Mode 3	1.60	0.00	2.52	0.40	1.60	2.44	12.5	136.7	1.60	0.00	-	-	1.58	0.00	15.6	45.52
	Mode 4	2.09	0.00	1.90	0.53	2.09	0.48	8.82	254.2	2.09	0.00	-	-	2.08	1.46	10.4	17.07
65dB	Mode 1	0.30	0.00	19.6	0.10	0.41	70.83	79.1	93.45	0.29	0.00	-	-	0.24	22.58	4.64	36.87
	Mode 2	1.14	0.00	3.54	0.00	1.07	5.31	55.5	1133	1.14	0.00	-	-	1.14	0.87	10.4	10.29
	Mode 3	1.60	0.00	2.51	0.00	1.40	14.63	-	-	1.60	0.00	-	-	1.66	5.06	14.4	34.33
	Mode 4	2.09	0.00	1.89	0.00	1.99	5.24	-	-	2.09	0.00	-	-	2.07	0.98	12.3	1.91
45dB	Mode 1	0.30	0.00	19.7	0.61	0.43	79.17	-	-	0.29	0.00	-	-	0.33	6.45	7.71	4.89
	Mode 2	1.14	0.00	3.55	0.28	-	-	-	-	1.14	0.00	-	-	1.15	0.00	5.95	36.8
	Mode 3	1.60	0.00	2.53	0.80	-	-	-	-	1.60	0.00	-	-	1.65	4.43	7.13	33.5
	Mode 4	2.09	0.00	1.90	0.53	-	-	-	-	2.09	0.00	-	-	2.18	6.34	7.82	37.7
25dB	Mode 1	0.30	0.00	19.1	2.45	-	-	-	-	0.31	6.90	-	-	0.38	22.58	8.21	11.7
	Mode 2	1.14	0.00	3.56	0.56	-	-	-	-	1.14	0.00	-	-	1.26	9.57	7.31	22.5
	Mode 3	1.60	0.00	2.41	3.98	-	-	-	-	1.59	0.63	-	-	1.62	2.53	10.9	1.68
	Mode 4	2.09	0.00	1.84	2.65	-	-	-	-	2.09	0.00	-	-	2.02	1.46	5.80	53.8
15dB	Mode 1	0.30	0.00	19.58	0.00	-	-	-	-	0.35	20.69	-	-	0.29	6.45	7.95	8.16
	Mode 2	1.14	0.00	3.39	4.24	-	-	-	-	1.12	1.75	-	-	1.15	0.00	6.42	31.92
	Mode 3	1.60	0.00	2.50	0.40	-	-	-	-	1.59	0.63	-	-	1.60	1.27	4.59	57.18
	Mode 4	2.09	0.00	1.89	0.00	-	-	-	-	-	-	-	-	2.10	2.44	4.19	66.59
10dB	Mode 1	0.29	0.00	19.76	0.92	-	-	-	-	0.35	20.69	-	-	0.30	3.23	5.11	30.48
	Mode 2	1.14	0.00	3.63	2.54	-	-	-	-	1.12	1.75	-	-	1.13	1.74	6.32	32.98
	Mode 3	1.60	0.00	2.52	0.40	-	-	-	-	1.59	0.63	-	-	1.59	0.63	6.96	35.07
	Mode 4	2.09	0.00	1.91	1.06	-	-	-	-	-	-	-	-	2.09	1.95	7.80	37.80
5dB	Mode 1	0.30	0.00	19.16	2.15	-	-	-	-	0.35	20.69	-	-	0.31	0.00	6.87	6.53
	Mode 2	1.14	0.00	3.49	1.41	-	-	-	-	1.12	1.75	-	-	1.14	0.87	6.61	29.90
	Mode 3	1.60	0.00	2.49	0.80	-	-	-	-	1.59	0.63	-	-	1.59	0.63	8.85	17.44
	Mode 4	2.09	0.00	1.89	0.00	-	-	-	-	-	-	-	-	2.09	1.95	7.80	37.80
1dB	Mode 1	0.28	6.67	21.95	12.10	-	-	-	-	0.31	6.90	-	-	0.29	6.45	6.01	18.23
	Mode 2	1.05	7.89	2.54	28.25	-	-	-	-	1.15	0.88	-	-	1.13	1.74	6.03	36.06
	Mode 3	1.59	0.63	1.41	43.82	-	-	-	-	1.60	0.00	-	-	1.60	1.27	5.35	50.09
	Mode 4	2.03	2.87	1.21	35.98	-	-	-	-	-	-	-	-	2.09	1.95	7.15	42.98

Table 12 lists the dominant modes estimated from the simulation data in Fig. 2 with different noise levels using OptVarPro, DMD, POD, and KMD. From Table 12, it is observed that DMD is sensitive to measurement noise. The damping of mode 1 estimated by DMD is increased from 40.89% at noise-free to 86.0% at SNR=85dB, and it is unable to detect dominant modes when the SNR is decreased to 45dB. POD exhibits robust performance against noise from SNR=85dB to SNR=25dB, but it is incapable of conducting damping estimation. The OptVarPro and KMD can extract all the dominant modes from the simulation data in Fig. 2 with different levels of noise, but the modes estimated by KMD are highly sensitive to the noise level. Table 12 compares estimation errors of the four methods using the noise-free case in Table 3 as the baseline. From SNB=85dB to SNB=5dB, it is observed that the maximum estimation errors of OptVarPro, DMD, POD, and KMD are 4.24% at the SNB=15dB, 1133% at the SNB=65dB, 20.69% at the SNB=15dB, and 66.59% at SNB=15dB, respectively. The estimation accuracy of the proposed OptVarPro falls at SNB=1dB where the maximum estimation error of the OptVarPro is 43.82%. Thus, it is concluded that the proposed OptVarPro method has strong robustness against measurement noise compared with DMD, POD, and KMD.

Further, a Monte Carlo simulation is conducted where measurement noise is intentionally injected into the mimicked data in Fig. 2. The total number of samples is 2,000. Within each sample, an independent set of measurement noise is added to generate new mimicked generator rotor angles and speed. The estimated dominant modes from OptVarPro and their means are plotted in Fig. 18. Also, it is assumed that the results estimated by EA are the true modes. From Fig.18, the means of the estimated dominant modes are very close to the true modes. In Table 13, the means and variances of the estimated dominant modes in the Monte Carlo simulation are given. It is seen that variances of the estimated dominant modes are all less than 9.7×10^{-5}. According to the comparison in Fig.18 and the calculated variances in Table 13, it is concluded that the proposed OptVarPro exhibits high accuracy and strong robustness in estimating the dominant mode from measurement data.



(a) Mode 1 (b) Mode 2

(c) Mode 3 (d) Mode 4

Fig. 18 Dominant modes estimated from the Monte Carlo simulation by using OptVarPro

Table 13 Means and variances of the dominant modes estimated from Monte Carlo simulations by using OptVarPro

Mode 1				Mode 2				Mode 3				Mode 4			
frequency		damping		frequency		damping		frequency		damping		frequency		damping	
Mean /Hz	Variance	Mean /%	Variance	Mean /Hz	Variance	Mean /%	Variance	Mean /Hz	Variance	Mean /%	Variance	Mean /Hz	Variance	Mean /%	Variance
0.30	1.97×10^{-5}	19.58	9.65×10^{-5}	1.14	8.41×10^{-6}	3.56	$4.8.4\times10^{-6}$	1.60	1.46×10^{-5}	2.50	8.22×10^{-6}	2.09	5.60×10^{-5}	1.98	9.44×10^{-6}

3.2. Yunnan Power Grid

In this subsection, the PMU field measurements in the Yunnan Power Grid are utilized to assess the practicability of the proposed approach in real-world electric power system applications.

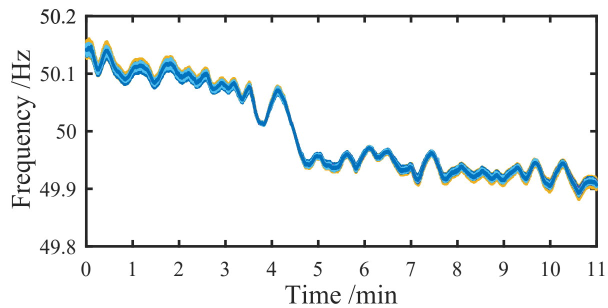


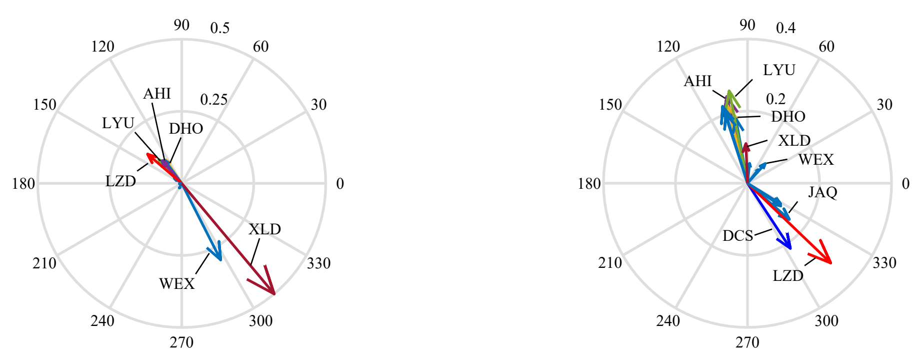
Fig. 19 Field measurement frequencies in the YPG

Yunnan Power Grid has plenty of synchronous hydro generators installed, and the electromechanical oscillation is a great threat to the grid stability. The results from the model-based EA method reveal that there are two typical electromechanical oscillation modes, Southwest-North mode, and Northwest-Northeast mode. The Southwest-North mode locates in the range of 0.52Hz~0.62Hz with 9%~15% damping, and the Northwest-Northeast mode ranges from 0.58Hz~0.65Hz with 15%~25% damping. The former mode is representative of generators in Southwest YPG swinging against generators in Northwest YPG and Northeast YPG. The latter mode represents the generators in Northwest YPG oscillating against the generators in Northeast YPG. A notable oscillation event in the YPG on March 8, 2016 has confirmed that it is essential to estimate the electromechanical oscillation characteristics from field measurement data in real-time.

Table 14 Mode shapes and participation factors estimated by EA and OptVarPro in the YPG

Bus	Southwest-North mode						Northwest-Northeast mode					
	EA			OptVarPro			EA			OptVarPro		
	f=0.59 Hz, ζ=11.28%			f=0.61 Hz, ζ=10.32%			f=0.62 Hz, ζ=23.91%			f=0.65 Hz, ζ=22.07%		
	Mode shapes		Participation factor	Ang./°	Mag.	Participation factor	Mag.	Ang./°	Participation factor	Ang./°	Mag.	Participation factor
	Mag.	Ang./°		Mag.	Ang./°		Mag.	Ang./°		Mag.	Ang./°	
AHI	0.12	-135.84	0.10	0.11	134.48	0.03	0.19	-146.62	0.17	0.25	103.55	0.12
DHO	0.09	-130.57	0.06	0.10	124.16	0.02	0.15	-143.47	0.11	0.22	104.57	0.09
GYY	0.02	-66.44	0.00	0.03	154.09	0.00	0.06	-114.76	0.02	0.05	51.78	0.00
HUE	0.09	-134.53	0.06	0.10	133.95	0.02	0.13	-144.35	0.08	0.18	100.16	0.06
JAQ	0.06	148.27	0.02	0.10	126.67	0.02	0.05	84.43	0.01	0.15	-40.83	0.05
LYU	0.12	-136.58	0.11	0.11	133.90	0.03	0.20	-144.77	0.18	0.26	101.24	0.13
LDI	0.09	-133.91	0.06	0.10	132.51	0.02	0.14	-144.10	0.09	0.22	107.95	0.10
LKK	0.06	-150.62	0.02	0.03	143.41	0.00	0.06	-177.06	0.01	0.06	83.95	0.01
MWE	0.09	-135.64	0.06	0.11	131.94	0.03	0.14	-145.40	0.09	0.20	107.32	0.08
DCS	0.07	132.63	0.03	0.10	130.28	0.02	0.09	67.87	0.04	0.22	-56.65	0.09
LZD	0.13	131.92	0.08	0.16	139.42	0.05	0.18	71.05	0.15	0.32	-43.77	0.20
DDO	0.02	127.91	0.00	0.02	-116.74	0.00	0.03	85.59	0.01	0.11	-31.86	0.02
YWA	0.01	125.17	0.00	0.01	-168.13	0.00	0.03	87.33	0.01	0.11	-34.37	0.02
WEX	0.13	-20.89	0.09	0.30	-63.00	0.20	0.03	-70.74	0.00	0.07	47.73	0.01
XLD	0.23	-10.78	0.30	0.50	-50.07	0.55	0.08	-53.03	0.02	0.11	92.37	0.02

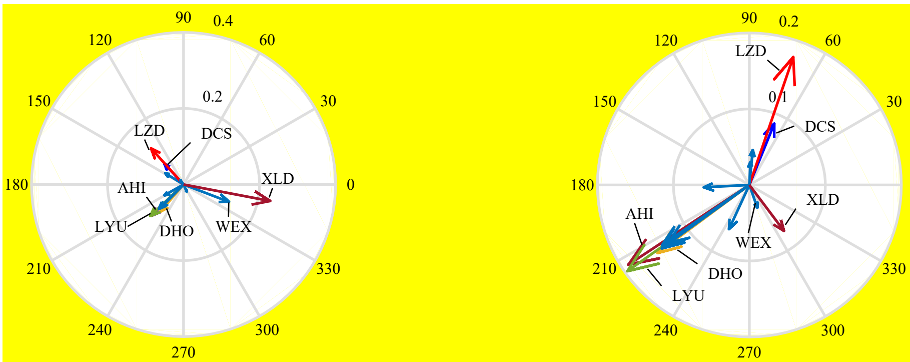
Fig. 19 shows the measured frequencies gathered from the PMUs deployed at buses AHI, DHO, GYY, HUE, JAQ, LYU, LDI, LKK, MWE, DCS, LZD, DDO, YWA, WEX, and XLD in the YPG system. This measurement data is used as the input to the proposed OptVarPro approach. Then two dominant modes, i.e., 0.61 Hz mode with 10.32% damping and 0.65 Hz mode with 22.07% damping, are detected. For both oscillation modes, the corresponding mode shapes are estimated by the proposed OptVarPro approach. The results are shown in Table 14 and Fig. 20(a). As can be seen, the estimated 0.61 Hz mode is primarily a Southwest-North mode with the generators connected to buses AHI, DHO, GYY, HUE, JAQ, LYU, LDI, LKK, MWE, DCS, LZD, DDO, and YWA oscillating against the generators connected to buses WEX and XLD. This mode has higher observability at the generators connected to buses WEX and XLD. The mode shapes associated with 0.65 Hz mode in Table 14 and Fig. 20(b) reveal that the generators connected to buses AHI, DHO, GYY, HUE, LYU, LDI, LKK, and MWE swing in phase with the generators connected to buses WEX, and XLD while oscillate against the generators connected to buses JAQ, DCS, LZD, DDO, and YWA. From the estimated mode shapes of the 0.65Hz mode, it is concluded that the detected 0.65Hz mode is the Northwest-Northeast mode.



(a) Mode shapes of Southwest-North mode (b) Mode shapes of Northwest-Northeast mode

Fig. 20 Mode shapes estimated from field measurements in the YPG

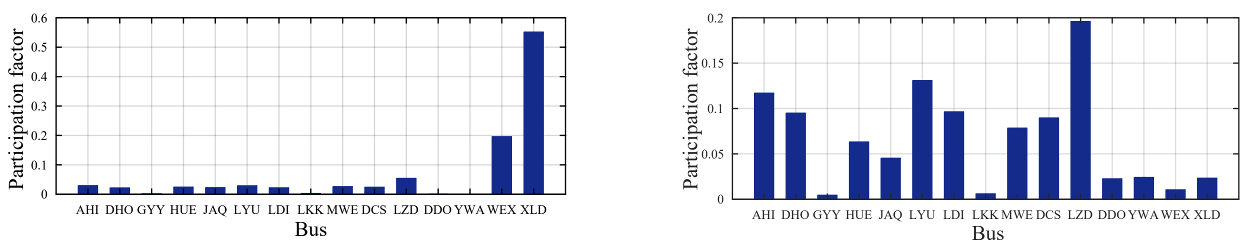
To evaluate the accuracy of the proposed OptVarPro method in estimating the mode shapes from the field-measurement data in real power systems, the mode shapes of the YPG are further estimated by the model-based EA method. The estimated results are listed in Table 14 and shown in Fig. 21. Figs. 20, 21, and Table 14 compare the mode shapes associated with the Southwest-North and Northwest-Northeast modes estimated by the proposed OptVarPro and EA. It is observed that although the magnitudes and angles of the estimated mode shapes are different in measurement-based OptVarPro and model-based EA, the buses connected with generators swinging in phase and behaving higher observability in the detected modes are the same. This proves the accuracy of the proposed OptVarPro method in real-world mode shape estimation.



(a) Mode shapes of Southwest-North mode (b) Mode shapes of Northwest-Northeast mode

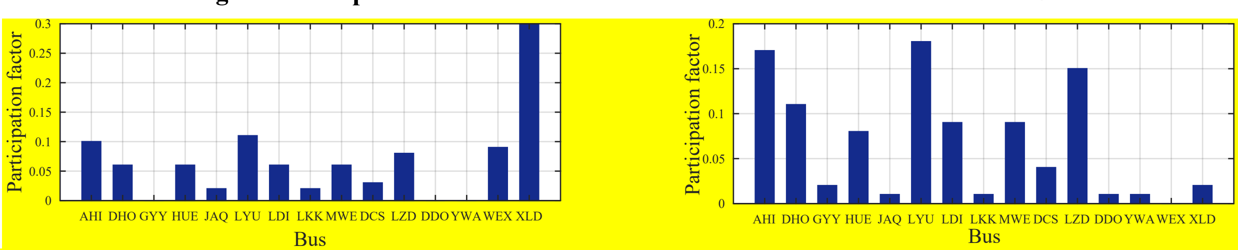
Fig. 21 Mode shapes estimated by EA in the YPG

Table 14 and Fig. 22 show the participation factors of the generators connected to the buses estimated by the OptVarPro. From Table 14 and Fig. 22, it is observed that the generators connected to bus XLD have the highest participation in Southwest-North mode, and the generators connected to buses AHI, DHO, LYU, LDI, MWE, DCS, and LZD all significantly participate in this single oscillation mode. Also, Table 14 and Fig.23 demonstrate the participation factors estimated by EA. By comparing the results of the participation factors estimated by OptVarPro and EA in Table 14 and Figs.22 and 23, it is obvious that the generators connected to the buses with dominant participation in both Southwest-North and Northwest-Northeast modes are the same. This validates the accuracy of the proposed OptVarPro method in estimating the participation factors in the YPG.



(a) Participation factors of Southwest-North mode (b) Participation factors of Northwest-Northeast mode

Fig. 22 Participation factors estimated from field measurements in the YPG



(a) Participation factors of Southwest-North mode (b) Participation factors of Northwest-Northeast mode

Fig. 23 Participation factors estimated by EA in the YPG

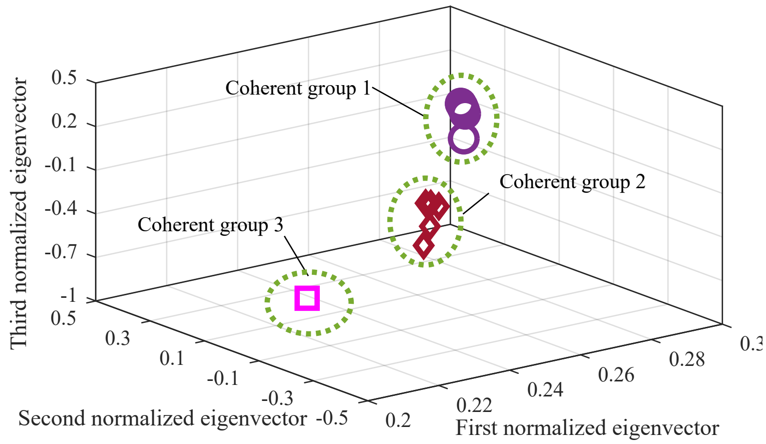


Fig. 24 Coherent groups estimated from field measurements in the YPG

The separated coherent groups of buses calculated from the proposed method are shown in Fig. 24, and the detailed results are listed in Table 15. As can be seen, the three coherent groups of buses are excited by both detected oscillation modes, where buses AHI, DHQ, GYY, HUE, LYU, LDI, LKK, and MWE form coherent group 1, buses DCS, LZD, DDO, YWA swing in phase with buses WEX, and XLD to form coherent group 2, and bus JAQ alone forms the coherent group 3.

Table 15 Coherent groups separated by OptVarPro in the YPG

method	Coherent groups
Group 1	AHI, DHQ, GYY, HUE, LYU, LDI, LKK, MWE
Group 2	DCS, LZD, DDO, YWA, WEX, XLD
Group 3	JAQ

From case studies, it is confirmed that the proposed OptVarPro-based data analytics solution haves excellent performance in estimating the dominant modes, mode shapes, participation factors, and coherent groups using real-world data.

4. Conclusions

This paper develops an OptVarPro-based data analytics solution to extract electromechanical oscillation features from measurement data in electric power systems. The performance of the proposed approach is validated using both the simulated data of a 16-generator 68-bus test system and the field PMU measurements from the YPG system. Major conclusions are drawn as follows:

1) The proposed OptVarPro-based data analytics solution can accurately extract dominant modes, mode shapes, participation factors, and coherent groups from measurement data.

2) The proposed OptVarPro-based data analytics solution can reveal the actual electromechanical oscillation features of power systems without detailed dynamic models and precise parameters.

3) The proposed OptVarPro-based data analytics solution overcomes the shortcoming of VarPro-based methods on initial modes determination and extends the capability of VarPro from mode estimation to estimation of modes, mode shapes, participation factors, and coherent groups in power systems.

4) The proposed OptVarPro-based data analytics solution can maintain accuracy under various levels of measurement noise when compared with DMD, POD, and KMD.

5) The proposed OptVarPro-based data analytics solution realizes data-driven electromechanical oscillation

assessment and offers comprehensive dynamic features for grid operators to evaluate and control the dynamic stability of electric power systems.

Acknowledgment

The authors would like to acknowledge the support in part by the National Natural Science Foundation of China (No. 51877033) and International Clean Energy Talent Programme (iCET) of China Scholarship Council.

References.

[1] IEEE/CIGRE Joint Task Force on Stability Terms and Definitions. Definition and classification of power system stability. *IEEE Trans. Power Sys*. 2004; 19(3): 1387-1401.

[2] Hatziargyriou N., Milanović J. V., Rahmann C., *et al*. Definition and Classification of Power System Stability Revisited & Extended. I *IEEE Trans. Power Sys*. early access.

[3] Kundur P. Power system stability and control. New York, NY, USA: McGraw-Hill 1994.

[4] Rogers G. Power systems oscillations. Norwell, MA: Kluwer 2000.

[5] Kosterev D., Taylor C., Mittelstadt W. Model validation for the August 10, 1996 WSCC system outage. *IEEE Trans. Power Sys*. 1999; 14(3): 967–979.

[6] Crow M., Gibbard M., Messina A., *et al*. Identification of electromechanical modes in power systems. IEEE Technical Report, Task Force on Identification of Electromechanical modes, 2012.

[7] Liu Z., Yao W., Wen J., *et al*. Effect analysis of generator governor system and its frequency mode on inter-area oscillations in power systems. *Int. J. Elect. Power Energy Syst*. 2018; 96: 1-10.

[8] Jiang T., Mu Y.F., Jia H.J., *et al*. A novel dominant mode estimation method for analyzing inter-area oscillation in China southern power grid. *IEEE Trans. Smart Grid*. 2016; 7(5): 2549-2560

[9] Rueda J., Juárez C., Erlich I. Wavelet-based analysis of power system low-frequency electromechanical oscillations. *IEEE Trans. Power Syst*. 2011; 26(3): 1733-1743.

[10] Turunen J., Thambirajah J., Larsson M., Pal B., Thornhill N. Comparison of three electromechanical oscillation damping estimation methods. *IEEE Trans. Power Syst*. 2011; 26(4): 2398–2407.

[11] Hamed K., Vaithianathan V. Modal analysis of ambient PMU measurements using orthogonal wavelet bases. *IEEE Trans. Smart Grid*. 2015; 6(6): 2954–2963.

[12] Hauer J., Demeure C., Scharf L. Initial results in Prony analysis of power system response signals. *IEEE Trans. Power Syst*. 1990; 5(1): 80–89.

[13] Jiang T., Yuan H.Y., Jia H.J., *et al*. Stochastic subspace identification-based approach for tracking inter-area oscillatory modes in bulk power system utilizing synchrophasor measurements. *IET Gen. Transm. Distrib*. 2015; 9(15): 2409-2418.

[14] Jiang T., Li X., Yuan H.Y., *et al*. Estimating Electromechanical Oscillation Modes from Synchrophasor Measurements in Bulk Power Grids Using Fast Iterative Stochastic Subspace Identification. *IET Gen. Transm. Distrib*. 2018; 12(10): 2347-2358.

[15] Jiang T., Yuan H.Y., Li G.Q., *et al*. Spatial-temporal decomposition approach for systematically tracking dominant modes, mode shapes and coherent groups in power systems. *IET Gen. Transm. Distrib*. 2017; 11(8): 1889–1900.

[16] Jiang T., Li X., Bai L.Q., Li F.X. Synchrophasor Measurement-based Modal Analysis in Power Grids. *North American Power Symposium* (NAPS), Wichita, Kansas, US, Oct 2019; 1-5.

[17] Pierre D., Trudnowski D., Hauer J. Identifying reduced-order models for large nonlinear systems with arbitrary initial conditions and multiple outputs using Prony signal analysis. *Proc.* 1990 *American Control Conf*, San Diego, CA, May 1990; 1: 149–154.

[18] Zhou N., Pierre J., Trudnowski D. A stepwise regression method for estimating dominant electromechanical modes. IEEE Trans. Power Syst. 2012; 27(2): 1051–1059.

[19] Messina A., Vittal V., Ruiz-Vega D., Enriquez-Harper G. Interpretation and Visualization of Wide-Area PMU Measurements Using Hilbert Analysis. *IEEE Trans. Power Syst*. 2006; 21(4): 1763-1771.

[20] Kamwa I., Pradhan A., Joós G. Robust detection and analysis of power system oscillations using the Teager-Kaiser energy operator. *IEEE Trans. Power Syst*. 2011; 26(1): 323–333.

[21] Peng J., Nair N. Enhancing Kalman filter for tracking ringdown electromechanical oscillations. *IEEE Trans. Power Syst*. 2012; 27(2): 1042-1050.

[22] Pierre J., Trudnowski D., Donnelly M. Initial results in electromechanical mode identification from ambient data. *IEEE Trans. Power Syst*. 1997; 12(3): 1245–1251.

[23] Wies R., Pierre J., Trudnowski D. Use of ARMA block processing for estimating stationary low-frequency electromechanical modes of power systems. *IEEE Trans. Power Syst*. 2003; 18(1): 167–173.

[24] Wies R. Estimating Low-Frequency Electromechanical Modes of Power System Using Ambient Data. Ph.D. dissertation, Dept. Electrical Eng, Univ of Wyoming, Laramie, WY 1999.

[25] Anderson M., Zhou N., Pierre J., *et al*. Bootstrap-based confidence interval estimates for electromechanical modes from multiple output analysis of measured ambient data. *IEEE Trans. Power Sys*. 2005; 20(2): 943-950.

[26] Zhou N., Pierre J., Trudnowski D., *et al*. Robust RLS methods for online estimation of power system electromechanical modes. *IEEE Trans. Power Syst*. 2007; 22(3): 1240-1249.

[27] Zhou N., Trudnowski D., Pierre J., *et al*. Electromechanical mode online estimation using regularized robust RLS methods. *IEEE Trans. Power Syst*. 2008; 23(4): 1670-1680.

[28] Jiang T., Bai L.Q, Jia H.J., *et al*. Synchrophasor measurement-based correlation approach for dominant mode identification in bulk power systems. *IET Gen. Transm. Distrib*. 2016; 10(11): 2710-2719.

[29] Jiang T., Bai L., Li X., *et al*. Estimating Electromechanical Oscillation Modes from Synchrophasor Measurements in China Southern Power Grid. *Elec. Power Syst. Res*. 2018; 161: 212-223.

[30] Nezam Sarmadi S., Venkatasubramanian V. Electromechanical mode estimation using recursive adaptive stochastic subspace identification. *IEEE Trans. Power Syst*. 2014; 29(1): 349-358.

[31] Messina A., Vittal V. Extraction of dynamic patterns from wide-area measurements using empirical orthogonal functions. *IEEE Trans. Power Syst*. 2007; 22(2): 682-692.

[32] Trudnowski D. Estimating electromechanical mode shape from synchrophasor measurements. *IEEE Trans. Power Syst*. 2008; 23(3): 1188-1195.

[33] Senroy N. Generator coherency using the Hilbert-Huang transform. *IEEE Trans. Power Syst*. 2008; 23(4): 1701-1708.

[34] Chauduri N., Chauduri B. Damping and relative mode-shape estimation in near real-time through phasor approach. *IEEE Trans. Power Syst*. 2011; 26(1): 364-373.
[35] Avdaković S., Bećirović E., Nuhanović A., *et al*. Generator coherency using the wavelet phase difference approach. *IEEE Trans. Power Syst*. 2014; 29(1): 271-278.
[36] Pan X.P., Shang F., Yuan S. Wavelet-based electromechanical mode shape online identification from ambient data. *Proc.* 2011 4*th Int Conf Electric Utility Deregulation and Restructuring and Power Technologies* (DRPT), Jul 2011; 1-7.
[37] Zhou N., Dosiek L., Trudnowski D., Pierre J. Electromechanical mode shape estimation based on transfer function identification using PMU measurements. *Proc.* 2009 *IEEE PES General Meeting*, Calgary, AB, Canada, Jul 2009.
[38] Dosiek L., Pierre J. Estimating electromechanical modes and mode shapes using the multichannel ARMAX model. *IEEE Trans. Power Syst*. 2013; 28(2): 1950-1959.
[39] Li X., Cui H., Jiang T., Xu Y., Jia H., Li F. Multichannel continuous wavelet transform approach to estimate electromechanical oscillation modes, mode shapes and coherent groups from synchrophasors in bulk power grids. *Int. J. Elect. Power Energy Syst*. 2018; 96: 222-237.
[40] Liu G., Venkatasubramanian V. Oscillation monitoring from ambient PMU measurements by frequency domain decomposition. *Proc IEEE Symp Circuits and Systems*, Seattle, WA, May 2008.
[41] Khalilinia H., Zhang L., Venkatasubramanian V. Fast frequency domain decomposition for ambient oscillation monitoring. *IEEE Trans. Power Del*. 2015; 30(3): 1631-1633.
[42] Dosiek L., Trudnowski D., Pierre J. New algorithms for mode shape estimation using measured data. *Proc* 2008 *IEEE PES General Meeting*, Pittsburgh, PA, Jul 2008.
[43] Dosiek L., Pierre J., Trudnowski D., Zhou N. A channel matching approach for estimating electromechanical mode shape and coherence. *Proc.* 2009 *IEEE PES General Meeting*, Calgary, AB, Canada, Jul 2009.
[44] Li X., Jiang T., Yuan H. Y., *et al*. An eigensystem realization algorithm based data-driven approach for extracting electromechanical oscillation dynamic patterns from synchrophasor measurements in bulk power grids. *Int. J. Elect. Power Energy Syst*. 2020; 116: 105549.
[45] Jiang T., Yuan H.Y., Zhao J.B., *et al*. Stochastic subspace identification based data-driven approach for monitoring electromechanical dynamics from phasor measurement units. *IET Gen. Transm. Distrib*. 2020; 14(19): 3983-3991.
[46] Ni J., Shen C., Liu F. Estimating the electromechanical oscillation characteristics of power system based on measured ambient data utilizing stochastic subspace method. *Proc.* 2011 *IEEE Power and Energy Society General Meeting*, Jul 2011; 1-7.
[47] Dosiek L., Zhou N., Pierre J., *et al*. Mode estimation algorithms under ambient conditions: A comparative review. *IEEE Trans. Power Syst*. 2013; 28(2): 779-787.
[48] Susuki Y., Mezić I. Nonlinear Koopman modes and coherency identification of coupled swing dynamics. *IEEE Trans. Power Syst*. 2011; 26(4): 1894-1904.
[49] Barocio E., Pal B., Thornhill N., Messina A. A dynamic mode decomposition framework for global power system oscillation analysis. *IEEE Trans. Power Syst*. 2015; 30(6): 2902-2912.
[50] Liu C., Jiang T., Zhao H., *et al*. Holistic data-driven framework for estimating electromechanical dynamic patterns from synchrophasor measurements in bulk power grids. *IET Energy Systems Integration* 2020; 2(4): 344-354.
[51] Anaparthi K., Chaudhuri B., Thornhill N., *et al*. Coherency identification in power systems through principal component analysis. *IEEE Trans. Power Syst*. 2005; 20(3): 1658-1660.
[52] Ariff M., Pal B. Coherency identification in interconnected power system-an independent component analysis approach. *IEEE Trans. Power Syst*. 2013; 28(3): 1747-1755.
[53] Lin Z., Wen F., Zhao J., *et al*. Controlled islanding schemes for interconnected power systems based on coherent generator group identification and wide-area measurements. *Journal of Modern Power Systems and Clean Energy* 2016; 4(3): 440-453.
[54] Lin Z., Wen F., Ding Y., *et al*. Data-driven coherency identification for generators based on spectral clustering. *IEEE Transactions on Industrial Informatics* 2018; 14(3): 1275-1285.
[55] Vahidnia A., Ledwich G., Palmer E., Ghosh A. Generator coherency and area detection in large power systems. *IET Gener. Transm. Distrib*. 2012; 6(9): 874-883.
[56] Aghamohammadi M., Tabandeh S. A new approach for online coherency identification in power systems based on correlation characteristics of generators rotor oscillations. *Int. J. Elect. Power Energy Syst*. 2016; 83: 470–484.
[57] Jiang T., Jia H.J., Yuan H.Y., Zhou N., Li F.X. Projection pursuit: A general methodology of wide-area coherency detection in bulk power grid. *IEEE Trans. Power Syst*. 2016; 31(4): 2776-2786.
[58] Borden A., Lesieutre B. Variable projection method for power system modal identification. *IEEE Trans. Power Syst*. 2014; 29(6): 2613-2620.
[59] Li X., Yu Y., Jiang T., *et al*. Estimating dominant oscillation characteristics from measurement responses in power systems utilizing optimized variable projection. *Transactions of China Electrotechnical Society*, DOI: 10.19595/j.cnki.1000-6753.tces.201498 (*early access*).
[60] Golub G., Pereyra V. Separable nonlinear least squares: the variable projection method and its applications. *Speech Communication* 2003; 45(1): 63-87.
[61] Golub G., Pereyra V. The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate. *SIAM J. Numer. Anal*. 1973; 10(2): 413-432.
[62] Golub G., Pereyra V. Separable nonlinear least squares: The variable projection method and its applications. *Inverse. Probl*. 2003; 9(2003): R1-R26.
[63] Askham T., Kutz J. Variable projection methods for an optimized dynamic mode decomposition. *SIAM J. Appl. Dyn. Syst*. 2018; 17(1): 380-416.
[64] Jiang T., Bai L.Q, Jia H.J., *et al*. Spectral clustering-based partitioning of volt/VAR control areas in bulk power systems. *IET Gener. Transm. Distrb*. 2017; 11(5): 1126-1133.