Density-based Unsupervised Learning Approach for Generators Coherency Evaluation in Complex Domain

Mohammad Hossein Rezaeian Koochi [a], Mohammad Hasan Hemmatpour [b,*]

[a] Department of Electrical Engineering, Shahid Bahonar University of Kerman, Kerman, Iran

[b] Department of Electrical Engineering, Faculty of Engineering, Jahrom University, Jahrom, Iran

Email: m.h.hematpour@jahromu.ac.ir

Corresponding author: Mobile: +989173918471, Tel.:+987154372252

Abstract: In measurement-based approaches, generators coherency can be determined on the basis of excited modes, which are obtained in the form of a vector of complex values. In order to find coherent generators in the frequency domain, which is suitable for applications such as wide-area control of power systems, these complex vectors should be clustered on the basis of their similarities. To do that, in this paper a new density-based unsupervised learning approach is proposed for clustering complex vectors, which is crucial for finding coherent generators in online applications. The proposed coherency evaluation approach is straightforward and more practical according to the reality of power systems since (i) eliminating the complexities in previous studies, it evaluates the coherency from the excited modes point of view using the complex correlation of frequency spectrums, (ii) it uses a new density-based learning approach with only one parameter setting making it more suitable for clustering generators. These features have been demonstrated on the well-known 16-machine, 68-bus system.

Keywords: clustering, coherency detection, complex vector, data mining, frequency domain, power system dynamic signature

1. Introduction

The deployment of state-of-the-art synchrophasor measurement technology has provided accessibility to a huge volume of data gathered throughout a bulk power system. The information hidden in these data can open new horizons in power systems operating and planning. However, these data are raw and must be turned into useful information. In this context, the use of pattern recognition and data mining for analyzing the raw data is getting more attention from researchers in different fields of power system studies, such as load forecasting [1], fault locating [2], security of systems [3], stability analysis [4], etc.

In power system dynamic studies, system analysis for determining the stability of the system after fault occurrence is very important. Besides, there is other information, such as coherency in the system, that should be analyzed online since it is a useful requirement for wide-area control systems. According to the concept of coherency, generators that reveal approximately the same dynamic responses to disturbances are called coherent and are placed in the same group [5]. In post-disturbance controlling of power systems, by knowing the groups of coherent generators, control actions are triggered to mitigate the post-disturbance severe conditions and make the system move toward a stable or an islanded situation. It has been recently examined that DERs in active distribution networks can reveal coherent participation in local and inter-area modes, which means that active distribution networks massively penetrated with DERs should participate in post-disturbance control actions as well [6]. Basically, the concept of coherency in a power system is a reflection of modes excited across the system. Therefore, extracting these modes and analyzing them can help to find coherent generators and eliminate many complexities related to previous coherency evaluation methods.

Up to now, various techniques have been proposed for coherency assessment in power systems, which can be divided into two categories: model-based and measurement-based methods. In model-based methods, the linearized model of the power system is used for detecting the degree of coherency [7 - 10]. While methods in this category use concepts such as singular perturbation [aa11], epsilon decomposition [12], etc, the most well-known approach is the slow-coherency analysis in which the linear representation of the system is employed in order to separate its slow dynamics from fast ones [13]. [14]. However, in model-based methods, system configuration changes or disturbance severity and location are not usually considered. Moreover, linearizing the model of a large system with hundreds of generators is a difficult task. The methods in the second category use various techniques to analyze the patterns in the post-disturbance variations of the parameters measured across the system. Therefore, detailed models and accurate parameters of the controllers and other types of equipment are not required in these methods. Accordingly, measurement-based coherency evaluation has received attention in recent decades. Examples of these methods are those that use the wavelet transform [15], Hilbert-Huang transforms [16], cosine similarity evaluation [17], projection pursuit [18], Koopman mode analysis [19], etc. to find coherent generators. Refs. [20] and [21] reviews comprehensively these approaches.

Machine learning techniques are mainly divided into supervised, unsupervised, and reinforcement learning categories [22, 23]. Unsupervised learning approaches, which are also called clustering type of data mining techniques, are used to discover hidden patterns in the unlabeled data in order to partition the dataset into a number of clusters. In unsupervised learning approaches, the algorithm is allowed to act on the available information or features without guidance. In other words, the task of the machine is to group unsorted datasets according to similarities, patterns, and differences without any prior training of data. On the contrary, supervised learning methods, also known as methods used to build classifiers, use a set of data with already known group labels for training the classifier. In fact, supervised learning methods try to find the rules that determine the relationships between the already observed data and their associated known labels and then use these rules to predict the label of the next observation. Among supervised learning techniques, neural networks, decision trees, and Bayesian networks are the most well-known. Reinforcement learning is the third branch of machine learning techniques, which is concerned with the problem of finding suitable actions to take in a given situation through a process of trial and error in order to maximize a reward. In these types of algorithms, there is a sequence of states and actions in which the learning algorithm interacts with its environment.

In the literature related to coherency-based power system partitioning, the use of both supervised and unsupervised learning approaches is proposed. Supervised learning methods have been used to build classifiers for predicting the post-disturbance generator grouping or system islanding strategies [24 – 28]. Whereas, unsupervised learning approaches have been proposed to be used for online generators or bus clustering [29]. Examples of unsupervised learning approaches are k-means clustering (KM) [30, 31], fuzzy c-means (FCM) [32] and fuzzy c-medoids (FCMd) [33, 34] clustering, subtractive clustering [35], partition around medoid clustering (PAM) [36], Monte Carlo consensus clustering [37], typicality-based analysis [38] [39], affinity propagation [40] and support vector clustering [41] algorithms, which have been evaluated so far. Among the recent works, [42] has tried to develop KM clustering by adding a new concept based on clustering slopes. In another recent work, Affinity propagation clustering is proposed in which a prior assumption on the number of clusters is not required [40].

There are several key problems with many of the abovementioned unsupervised learning techniques in relation to their capabilities to discover clusters of coherent generators, which are as follows; these techniques can not work well for highly unbalanced data sets (i.e. cases where both very small and big clusters exist at the same time in the dataset), they are highly dependent on random operations, they require the number of clusters to be defined in advance, and they have dependency on the settings of two or more parameters. It should be noted that not all the above-mentioned clustering techniques suffer from all the above shortcomings. For example, by using affinity propagation proposed in [40] the need to have a prior assumption on the number of clusters is solved, but it is still required to set various parameters for this algorithm.

Among clustering techniques, a density-based approach such as density-based spatial clustering of applications with noise (DBSCAN) [43] is a highly suitable one for grouping the generators in a power system. This is because not only no random selection operation is used in the DBSCAN algorithm, but no prior assumption on the number of clusters is required, which is vital for this study since the number of groups of coherent generators and their members is not fixed and may change for different disturbance characteristics and system conditions. Also, DBSCAN has the ability to find clusters with a very low number of members. This is important since, for example in unstable conditions, generators in a power system may be divided into groups with different sizes, i.e. a very small group containing only one generator may exist along with other medium and large groups. However, DBSCAN is introduced for real vectors. Therefore, in this paper, a new density-based learning approach for clustering complex vectors, called density-based clustering for complex vectors (DBCXV), is introduced, which is suitable for problems such as coherent generator grouping in power system analysis. This method is a developed version of DBSCAN in which new rules are added to make additional concepts for evaluating the similarities of complex vectors in 2D spaces, i.e. complex planes.

Aiming to address the abovementioned challenges, in this paper, coherency evaluation is performed in frequency domain to achieve a practical solution for discovering groups of coherent generators in terms of excited modes. This method is based on complex values which contain all the necessary information required for coherency assessment. It is clear that two generators may reveal the same oscillatory frequency with the same magnitude but with a phase angle difference of 180°. Accordingly, it is said that these two generators are out of phase (or they are not coherent). This suggests that relying only on the magnitude of oscillatory modes is not sufficient for coherency evaluation. Therefore, in this paper, a new approach is proposed to discover coherency among generators by analyzing complex values. Fig. 1 shows the conceptual illustration of the steps of the proposed approach. Note that, unlike methodologies in which disturbances are considered to be distributed over the whole system as a white noise applied on loads in all buses [44], in the proposed approach coherency evaluation is performed on the basis of parameters' variations observed as the result of a single disturbance occurrence. In this regard, for each generator a low-frequency spectrum, or in other words a complex vector consisting of low-frequency modes, will exist. In order to cluster the complex vectors associated with generators and in turn find coherent generators, the proposed DBCXV will be used here as an alternative to cope with the challenging problems described in this section regarding the clustering of coherent generators. To be specific, the proposed approach converts the problem of discovering coherent generators into clustering the complex vectors associated with generators using an innovative technique. Note that the proposed methodology is an unsupervised learning one since neither training data nor training step is required. In summary, from both mathematical and technical points of view, the methodology proposed for coherency evaluation has the following features.

Fig. 1. Conceptual illustration of the proposed approach

- It is simple, since unlike many complicated measurement-based coherency evaluation methods, it uses frequency components, which can be simply extracted from generators' speed signals.
- It is fast, since using DFT for extracting frequencies has a low computational burden, in addition, unlike algorithms such as KM and FCM, the proposed DBCXV algorithm is not an iteration-based technique.
- It is accurate in terms of clustering coherent generators since the proposed DBCXV algorithm can accurately cluster the complex vectors.
- The proposed DBCXV algorithm is highly suitable for this problem since its solution is deterministic, no prior assumption on the number of clusters is required for the algorithm, it has fewer parameters to be set, and it has the capability to find generators that lose synchronization in unstable cases. It should be noted that, in a power network, some generators are located at the corners of the network and therefore if a fault happens near them they probably swing against the

rest of the generators (this means that this generator individually forms a group as it is non-coherent with the rest of the generators). The same situation happens to a generator that loses its synchronism with the rest of the network. In such cases, the proposed algorithm is able to find these single-generator clusters by detecting outliers (which are called noises in the approach) in the dataset.

- Since this methodology makes use of data measured by phasor measurement units (PMUs) and is based on excited modes, it is applicable for measurement-based applications, particularly wide-area control systems, which need accurate groups of coherent generators to be available online for tuning controllers.

2. Proposed Approach for Determining Groups of Coherent Generators

In this section, the proposed coherency evaluation in the complex domain and the proposed density-based learning technique for the mode-based grouping of generators are discussed.

2.1. Coherency Evaluation in the Complex Domain

From a measurement-based point of view, two generators revealing highly similar rotor angle or speed variations are called highly coherent. The degree of these similarities can be obtained by calculating the correlation coefficient (CC) between post-disturbance variations in speed or rotor angle signal of each pair of generators. It has been described in [45] that the variations in the speed of generators in a power system are a reflection of the energy absorbed or delivered by the generators. On the other hand, according to Parseval's theorem, the sum of the squares of a function is equal to the sum of the squares of its Fourier transform as follows [46].

$$\sum_{k=0}^{N-1} \left| x(k) \right|^2 = \frac{1}{N} \sum_{f=0}^{N-1} \left| X(f) \right|^2 \tag{1}$$

where $X(f)$ is the f^{th} component in the discrete Fourier transform of the finite-duration signal x and N is the number of samples. The practical interpretation derived from this theorem is that the energy of a signal is the same as the energy calculated from the frequency components that exist in the signal waveform. Authors in [47] showed that the Euclidean distance between two signals in the time domain is the same as their Euclidean distance in the frequency domain, meaning that for coherency evaluation on the basis of excited modes, it can be possible to easily move to the frequency domain where these excited modes are represented by complex values. Since in this paper the interest is on inter-area and local modes, which are lower than 2 Hz, and considering the approach as a frequency-weighted one, filters can be added in real applications to filter the higher frequencies. To extract the sinusoidal contents of speed signals, DFT is used as in (2):

$$F_i(f) = \sum_{k=0}^{N-1} \omega_i(k)\, e^{-\left(j\frac{2\pi fk}{N}\right)} \qquad f = 0,1,\ldots,N-1 \tag{2}$$

where

$$\omega_i(k) = \frac{\theta_i(k) - \theta_i(k-1)}{\Delta t} \tag{3}$$

In the above formula, ω_i (k) and θ_i (k) are the speed and rotor angle of generator i at time instant k, and, F_i (f) represents the Fourier transform of the speed signal. Δt is the time interval between two consecutive samples. To be more clear, F_i (f) denotes the f^{th} component in the frequency spectrum obtained for the speed of generator i or in other words the f^{th} mode excited by generator i.

One of the efficient ways for assessing the degree of similarity between two vectors of real values such as speed signals of generators is to use the Pearson correlation coefficient. In this regard, for assessing the similarity between the speed signals of generators i and j, the Pearson correlation coefficient takes the form defined in (4).

$$c_{i,j} = \frac{\sum_{k=1}^{N} \omega_i(k)\omega_j(k)}{\sqrt{\sum_{k=1}^{N} \omega_i^2(k) \sum_{k=1}^{N} \omega_j^2(k)}} \tag{4}$$

In case the similarity between two vectors with complex values, such as frequency spectrums of generators i and j, is of interest, the CC is then proposed to be obtained using (5). In this paper, the CC calculated between vectors with complex elements is denoted as $cx_{i,j}$.

$$cx_{i,j} = \frac{\sum_{f=1}^{N_f} F_i(f) F_j^*(f)}{\sqrt{\sum_{f=1}^{N_f} |F_i(f)|^2 \sum_{f=1}^{N_f} |F_j(f)|^2}} \quad (5)$$

In (5), * is the conjugate operator and | . | represents the magnitude of frequency components. Moreover, N_f is the number of low frequencies (lower than 2 Hz) in the frequency spectrums. Note that for two signals with real values, $c_{i,j}$ obtained from (4) takes a real value within the range [-1 , 1]. In this regard, for two similar signals, $c_{i,j}$ = 1, while for two negatively correlated signals, its value will be equal to -1. However, since the frequency spectrum is a vector whose elements are obtained from (2), the CC between each pair of these complex vectors (frequency spectrums) obtained from (5) will be a complex value as well. The position of *cx* in the complex plane is somewhere within a circle with a radius equal to 1 and centered on the point 0 + j0 in the complex plane. Clearly, if a pair of generators *i* and *j* have completely similar frequency spectrums, $cx_{i,j}$ will take the complex value 1 + j0, whereas if generator *i* oscillates exactly against generator *j* (meaning that a phase angle difference of 180 degrees exists), $cx_{i,j}$ takes the value -1 + j0. Fig. 2 shows an example of values of *cx* between generator *i* and two other generators *j* and *k* plotted in the complex plane. Note that, the more the location of $cx_{i,j}$ is close to the point 1 + j0, the more generator *i* is coherent with generator *j*. For example, from Fig 2, it can be concluded that the coherency between generators *j* and *i* is more than the coherency between generator *k* and generator *i*.

Fig. 2. Definition of coherency in the complex plane

2.2. Proposed Density-Based Learning Algorithm for Generators Clustering

2.2.1. Proposed Definition of Clusters in the Complex Domain

DBSCAN, which is a clustering algorithm simpler than k-means and fuzzy c-means algorithms, is proposed in [48] for generators grouping due to the following features:

- A prior assumption on the number of clusters is not required by DBSCAN. This is a very important feature since from a measurement point of view groups of coherent generators are not fixed for different disturbances.

- DBSCAN needs fewer parameters to be set compared to conventional clustering algorithms.
- No random selection operation is used by DBSCAN, and therefore its solution is deterministic.

DBSCAN works by detecting dense clusters, which is a concept different from the one used in KM and FCM algorithms. One shortcoming of KM and FCM, which are also named as centroid-based clustering algorithms, is that they have no notion of outliers, meaning that these algorithms assign all data points to clusters even with low degrees of membership. On the contrary, the density-based algorithm identifies places where points are clustered and where they are separated by vacant or sparse regions. Accordingly, this concept allows for identifying outliers (which are called noises in this paper) as well.

However, DBSCAN is introduced for vectors with real elements. As stated in section 2.1, an efficient coherency evaluation can be obtained by analyzing the frequency components that exist in the speed signal of generators. Since the value of $cx_{i,j}$ obtained from (5) is a complex value, using DBSCAN to group the generators on the basis of frequency spectrums is not feasible in this way. Therefore, in this paper, a new density-based learning approach derived from DBSCAN, called DBCXV, is introduced for clustering the vectors with complex values. This algorithm not only can be used for generators grouping in power system studies but it can be applied to other scientific fields where clustering of complex vectors is needed.

It should be noted that, in the proposed DBCXV algorithm, concepts including core data points, noises, definite clusters, and non-definite clusters are utilized. Most of these concepts have been previously defined in the DBSCAN algorithm [48]. However, in this paper, some of them are re-defined to be applicable to the proposed DBCXV algorithm. Below is the definition of these concepts:

- *Core point*: In a complex plane, a data point p_r is a core data point if at least *Pmin* data points including p_r are within distance α of it.
- *Being directly reachable*: In a complex plane, if data point q_r is within distance α of data point p_r, it is said that data point q_r is directly reachable from data point p_r.
- *Noise*: In a complex plane, a data point p_r is a noise if it is not directly reachable from any core data point.
- *Being reachable*: A data point q_r is reachable from data point p_r if there is a path p_{r1}, ..., p_{rn}

with $p_{r1} = p_r$ and $p_{rn} = q_r$, where each $p_{r(i+1)}$ is directly reachable from p_{ri}.

- *Non-definite cluster*: In a complex plane, a non-definite cluster is a set of at least *Pmin* data points so that each two of them are directly reachable, and additionally none of them are the members of a definite cluster.
- *Definite cluster*: In a complex plane, a definite cluster is a non-definite cluster that contains a core point located exactly at point 1 + j0.

According to the last two definitions, any core data point can form a non-definite cluster if none of the data points that are directly reachable from it are members of a definite cluster. In addition, it is clear that a definite cluster forms around point 1 + j0.

Fig. 3 presents an example of data points in the complex plane to provide a better understanding of the above definitions. In this example, nine data points are assumed in the complex plane. In Fig. 3, dashed circles are drawn to show a neighborhood with radius $\alpha = 0.2$ around each data point. By assuming *Pmin* set on 2 (meaning that a cluster can be formed by at least two data points), it can be seen that there is a data point d_2 in the neighborhood of data point d_1, which means that data points d_1 and d_2 form a definite cluster shown by *D* in the figure. There is also one noise (data point d_9) in this example since there is no other data point within its neighborhood as can be seen from the figure. Moreover, there are two non-definite clusters in Fig. 3 named A and B. It should be noted that, as stated in the definitions, the data points in each of the clusters are reachable.

Fig. 3. Example of noise, definite cluster, and non-definite cluster in the complex plane

2.2.2. Proposed DBCXV Algorithm

In order to group the generators of a power system on the basis of post-disturbance excited modes, the frequency spectrum of the speed signal of the generators should be extracted. Then, the similarity between each two of the frequency spectrums is calculated using (5), and a correlation matrix, i.e. the matrix *CM* defined in (6), is formed. It should be noted that *CM* is a matrix with dimension $N_G \times N_G$ where N_G is the number of generators. Moreover, in (6) the i^{th} and j^{th} element of *CM* is the correlation between the frequency spectrums of i^{th} and j^{th} generators, i.e. $cx_{i,j}$. It is also clear that $cx_{i,i} = 1$ and

$cx_{i,j} = cx^{*}_{j,i}$.

$$CM = \begin{bmatrix} cx_{1,1} & \cdots & cx_{1,i} & \cdots & cx_{1,N_G} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ cx_{i,1} & \cdots & cx_{i,i} & \cdots & cx_{i,N_G} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ cx_{N_G,1} & \cdots & cx_{N_G,i} & \cdots & cx_{N_G,N_G} \end{bmatrix} \tag{6}$$

By looking at (6), it can be concluded that the elements in the i^{th} row indicate how generators oscillate from the view point of generator *i*. In other words, by analyzing the i^{th} row, it is possible to find the generators that may form a definite cluster with generators *i* and the other possible non-definite clusters and noises. Fig. 3 is an example that shows the elements of row d_1 of a *CM* matrix obtained for a system with nine generators d_1 to d_9 (or it can be said that that this figure shows how other generators d_2 to d_9 oscillate from the viewpoint of generator d_1). By looking at Fig. 3 it can be concluded that generators d_1 and d_2 oscillate in phase to a high extent (are coherent), whereas the generators in clusters A or B swing against generator d_1. However, further analysis is still required to figure out whether clusters A and B oscillate against each other or not. The reason is that, as it is stated above, the location of data points plotted in Fig. 3 are from the viewpoint of generator d_1. Thus, these locations in the complex plane in Fig. 3 are valid only with respect to generator d_1, and if another row of matrix *CM* is considered, the locations of data points will be different from those plotted in Fig 3.

As stated in Section 2.2.1, in a complex plane definite clusters are formed around point 1 + j0. Accordingly, another matrix called *B* is extracted from matrix *CM* whose elements are defined as below.

$$B(i,j) = \begin{cases} 1 & \text{if} \quad |1 - CM(i,j)| \leq \alpha \\ 0 & \text{if} \quad |1 - CM(i,j)| > \alpha \end{cases} \tag{7}$$

Matrix *B* is very helpful in finding the definite clusters which are the clusters of generators in the final solution. In order to achieve the clustering solution, the algorithm defined below is proposed.

Step 1: For a power system with N_G generators, following a disturbance occurrence, first, create N_G frequency spectrums for all N_G generators from their speed signal using (2) and then form the matrix *B* defined in (7).

Step 2: select a row number randomly (for example row i) and store the row number i in a set called S.

Step 3: for each row number in S find the unity values in the corresponding row in matrix B and expand S by storing the column number of unity values in the selected row.

Step 4: repeat step 3 until no new row number is added to S.

Step 5: if the number of row numbers stored in S is greater than or equal to Pmin, label all the numbers in S as the first cluster of generators. Otherwise, label them as noise.

Step 6: remove all rows and columns from matrix B corresponding to the row numbers stored in S and then clear all the numbers of S.

Step 7: repeat steps 2 to 6 for the rest of the rows remaining in matrix B until matrix B becomes empty.

It is worthy noting that, similar to most of tolerance-based coherency evaluation techniques (either model-based or measurement-based), there is a threshold (tolerance) defined in the proposed methodology below which two generators are treated as coherent and vice versa. This threshold is the parameter α defined in Section 2.2.1. In fact, the proposed clustering methodology assigns two generators to the same group as coherent generators, if the value of dissimilarity between them (expressed in (7)) is less than α. This threshold is user-defined like all other tolerance-based techniques in the literature. It means that coherency is deemed to be a relative term, and it is the user who decides at which level of similarity, two generators are considered coherent.

For a better understanding of the proposed approach, an example with a small number of generators is presented here. Fig. 4 shows the speed signals of five generators following a disturbance occurrence. From Fig. 4 it can be seen that generators 1 and 2 (shown by dashed line) have oscillations almost against the rest of the generators. Accordingly, it can be said that generators 1 and 2 are coherent and form a cluster while the other cluster includes generators 3, 4, and 5.

Fig. 4. Speed variations of generators (example)

Matrices CM and B obtained for this example are shown in (8) and (9) respectively. As stated, the elements in the i^{th} row imply how the rest of generators oscillate with respect to generator i. To confirm that, the elements of each row of matrix *CM* are plotted in complex planes as shown in Fig. 5. It can be clearly seen in all sub-figures of Fig. 5 that generators 1 and 2 are coherent and oscillates against the cluster including generators 3, 4 and 5.

$$CM = \begin{bmatrix} 1 & 0.89+j0.16 & -0.25+j0.35 & -0.37+j0.25 & -0.21+j0.37 \\ 0.89-j0.16 & 1 & 0.13+j0.37 & 0.00+j0.32 & 0.16+j0.37 \\ -0.25-j0.35 & 0.13-j0.37 & 1 & 0.98+j0.14 & 0.99-j0.03 \\ -0.37-j0.25 & 0.00-j0.32 & 0.98-j0.14 & 1 & 0.97-j0.17 \\ -0.21-j0.37 & 0.16-j0.37 & 0.99+j0.03 & 0.97+j0.17 & 1 \end{bmatrix} \tag{8}$$

$$B = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix} \tag{9}$$

Fig. 5. Correlation values with respect to (a) generator 1, (b) generator 2, (c) generator 3, (d) generator 4, and (e) generator 5.

3. Simulation Results

In this paper, the efficacy of the proposed coherency evaluation approach as well as the proposed DBCXV algorithm is assessed using two test cases applied to the 16-machine, 68-bus test system (shown in Fig. 6), which is a well-known system and is widely used for power system dynamics related studies in the literature. The details of this system can be found in [49]. In this paper, all simulations are carried out within the MATLAB environment. Particularly, the Power System Toolbox (PST) [50] is used for simulating disturbances. Moreover, it should be noted that, in this study it is assumed that the system is fully observable through a wide-area measurement system, meaning that by using voltage phasors measured across the system rotor angle and speeds of generators can be obtained [51, 52].

Fig. 6. Line diagram of the test system

There is one note about the use of DBCXV that should be under attention. As stated in Section 1, it is desired to use a clustering algorithm with lowest number of parameters to be set. DBCXV intrinsically has two parameters, i.e. *Pmin* and α. However, according to the nature of coherency detection problem, the value of one of these parameters can be easily set. In doing so, firstly note that in a power system, at least two generators can form a coherent group in some parts of the system. Moreover, if an unstable condition occurs after the disturbance, a generator may gradually lose its synchronization and form an individual group, or in other words form a noise. Thus, to make the algorithm capable of distinguishing between a group of coherent generators with two members and an unstable generator, *Pmin* should be essentially set to 2. Therefore, the only remained parameter that should be set is α, which has been set at 0.2, in this study.

3.1. Case #1

In this case, a three-phase fault is applied on the line connecting buses 52 and 42 and near to bus 52 and is cleared after 31 ms. Furthermore, a loading factor of 1.0217 is applied to the normal load of the system detailed in [49]. This case is simulated for 10 seconds to let the variations appropriately damped. Fig. 7 shows the speed variations of the generators of the system. As it can be seen, among the 16 generators, three of them, i.e. generators 9, 14, and 16 (dashed lines), reveal different oscillations. Note that in these simulations, firstly, DFT is applied to the speed signals to extract complex vectors of modes. Then, frequency components less than 2 Hz as the inter-area and local modes have been selected to be used in the clustering algorithm.

Fig. 7. Speed variations of generators (Case 1)

Fig. 8 depicts the elements of matrix B formed by using (7). The diagonal elements of matrix B are the core points. From this figure it can be seen that a large group of generators are forming a cluster meaning that a large part of the system is oscillating against the rest. Another focus in Fig. 8 shows that generators 9 and 14 are found to be coherent. However, these generators are in two apart sides of the network as can be seen in Fig. 9, although they reveal, to some extent, similar oscillations. Therefore, they should be placed in separate areas.

Fig. 8. Elements of matrix B (Case 1)

Fig. 9. Groups of generators discovered in Case 1

Table 1 presents the final generators grouping scheme obtained in this case by using the proposed method. This grouping scheme is also graphically depicted in Fig. 9. As stated in Introduction section, one of the main advantages of DBCXV is that it can easily find a small cluster along with large clusters in the data set. In this case, the DBCXV algorithm has found a large cluster containing 13 generators (which is a large one compared to the others) and three single-generator clusters, which confirms its capability to find both small and large clusters in the dataset.

Table 1. Generator grouping scheme obtained in Case 1

In addition, an examination on the impact of various values of α on the generators clustering is also carried out. To do so, the proposed algorithm is applied with α values of 0.1 and 0.3. The additional results are presented in Figs. 10 and 11. By looking at these figures, it can be concluded that, as it is expected, the number of clusters is reduced as the value of α increased to 0.3, and vice versa. Note that the groups of generators obtained for $\alpha = 0.3$ is the same as the one obtained for $\alpha = 0.2$, meaning that still a value of α higher than 0.3 is required to assign generators 9, 14, or 16 to the large area in Fig. 9.

Fig. 10. Generators clustering with $\alpha = 0.1$ (Case 1)

Fig. 11. Generators clustering with $\alpha = 0.3$ (Case 1)

3.2. Case #2

In this case, which is more complicated than case 1, a fault similar to that of case 1 is applied, except that it is applied on the line connecting buses 4 and 5 and near to bus 4 and is cleared after 28 ms. Moreover, in this case the loading factor is considered to be 1.0414. Fig. 12 shows the speed variations of the generators of the system. This case is aimed to show that in a power system generators grouping schemes can be different for different disturbances occurring at different parts of the system. Elements of matrix B obtained in this case have been shown in Fig. 13. The low number of unity values in B indicates that more clusters but with small sizes are expected to be found compared to Case 1.

Fig. 12. Speed variations of generators (Case 2)

Fig. 13. Elements of matrix B (Case 2)

Table 2 details the clusters with their members in the final solution obtained in this case. Moreover, a graphical presentation of areas is shown in Fig. 14. To show that how generators in each area are coherent, the speed variations of generators in each area is depicted in Fig. 15. It should be noted that the last sub-figure contains those generators that have not been assigned to any of the other clusters. Fig. 15 clearly shows that generators in each cluster reveal approximately the same oscillations which proves the capability of the proposed approach. As it can be seen, in this case larger number of areas have been detected compared to Case 1, which is due to the excitation of several local modes. However, as will be shown in Section 3.3.1, when local modes are damped some of these areas will combine to form a larger area.

Table 2. Generator grouping scheme obtained in Case 2

Fig. 14. Groups of generators discovered in Case 2

Fig. 15. Speed variations of generators (Case 2)

Similar to Case 1, another examination is also carried out with α set to 0.1 and 0.3 to investigate the impact of α on the clustering solution. The new clustering solutions are presented in Figs. 16 and 17. It can be seen that, as it is expected, the number of clusters reduces if α is set to a lower value.

Fig. 16. Generators clustering with $\alpha = 0.1$ (Case 2)

Fig. 17. Generators clustering with $\alpha = 0.3$ (Case 2)

3.3. Discussion

In this section, the impact of data resolution and observation time frame on the solution of the proposed clustering algorithm as well as the performance of DFT will be discussed, and a comparison with traditional clustering algorithms will be provided.

3.3.1. Data Resolution and Observation Time Frame

The number of frequencies in the frequency spectrum obtained from applying DFT to a waveform is equal to the number of samples of the waveform, as DFT transforms the N discrete time samples to the same number of frequency samples. It means that, in order to obtain more low-frequency samples and in turn more accuracy of results, it is necessary to use higher sampling rates of the speed signals. In a wide-area measurement system, PMUs provide voltage or speed samples with various rates to be used by the control center. According to the IEEE Standard for Synchrophasor Data Transfer for Power Systems, in a 60 Hz system, PMUs have to provide these data with rates up to 60 sample/s, while higher rates such as 120 sample/s have also been encouraged by the standard. In the simulations presented in Sections 3.1 and 3.2, it is assumed that the generator speed data is sampled with the frequency of 100 Hz (1.66 sample/cycle). However, if a higher rate, for example 2 or more samples per cycle, is available, a frequency spectrum with more samples could be obtained, which can lead to a more accurate coherency evaluation.

In the other hand, in this study, a time frame of 10 seconds has been used for coherency evaluation, which is suitable for capturing frequencies within the range of 0.1 – 2 Hz. It can be seen from Fig. 18 that there are oscillations in the speed signals that are damped in the first four seconds. Therefore, if the start point of the time frame is where these modes are damped, a grouping solution is obtained in which modes that have frequencies at around 1 Hz and above are not considered. Fig. 18 shows the grouping solutions obtained by the proposed approach using two different time frames. As it can be seen, considering a time frame starting at t = 4s, a different grouping scheme is obtained in which generators 1 to 8 forms a large group.

Fig. 18. Coherent generators obtained in two different time frames (case 2)

3.3.2. Performance evaluation of DFT and Prony analysis

As described in Section 2.1, the proposed algorithm is a measurement-based one in which excited modes extracted from generators' speed signals are used for coherency evaluation. To extract these modes, DFT has been used in this paper. DFT in its original description has some limitations such as frequency leakage, which affects the estimations on the magnitude and phase angle of oscillations. However, it should be mentioned that the aim has not been to find the exact value of magnitude and phase angle of each excited mode and a small error can be neglected, since if two generators reveal the same excited modes in their oscillations, the frequency spectrum corresponding to them are then approximately similar (considering the error in the estimation made by DFT). It should be noted that coherency is a relative term meaning that two generators are coherent if they reveal relatively similar dynamic responses. In fact, the reason for using DFT instead of other signal processing techniques such as wavelet or Hilbert

Huang transforms, or even other DFT-based phasor estimation techniques, is to make the procedure fast and simple as much as possible while keeping the accuracy within proper range.

In the literature, Prony analysis is proposed for extracting the characteristics of an oscillatory decaying signal, which are the magnitude, phase angle, frequency, and damping factor of the signal. The accuracy of estimating these characteristics using Prony analysis highly depends on the appropriate selection of fitting parameters, which varies for different oscillatory signals [53]. In the other hand, with Prony analysis there is no prior knowledge on the frequency of the oscillatory signal, meaning that it is the Prony analysis responsible for finding the frequency of the damping oscillations. Accordingly, for each oscillatory signal an extra task of finding proper values for the fitting parameters is required.

However, in the proposed methodology, damping characteristics of the modes are not required, as the similarity of two signals is evaluated only in terms of the similarity of magnitudes and phase angles of the same oscillatory modes. This means that the aim is to look for the same frequencies in all generators, since it is clear that coherent generators reveal the same low frequency modes. If calculation of frequency is integrated with the calculation of magnitude and phase angle of the oscillatory signals, a small change in the phase angle may affect the estimation of the frequency as well. Therefore, making these comparisons is not highly reliable if Prony analysis is used to estimate the magnitude and phase angle of oscillatory signals.

In the following example, the speed signals of generators 2 and 3 in Case 1 are examined. As shown in Fig. 19 these two generators reveal highly similar speed variations and therefore they are highly coherent. Prony analysis is applied to these two signals using three common methods, i.e. classic, least square (LS), and total least square (TLS) [54]. Figs, 20 to 22 present magnitudes and phase angles of the low-frequency modes extracted by Prony analysis.

Fig. 19. Speed signals of generators 2 and 3 in Case 1

Fig. 20. Modes obtained by TLS Prony analysis, (a) magnitude, and (b) phase angle

Fig. 21. Modes obtained by LS Prony analysis, (a) magnitude, and (b) phase angle

Fig. 22. Modes obtained by classic Prony analysis, (a) magnitude, and (b) phase angle

As it can be seen from Figs. 20 to 22, Prony analysis is not always capable of extracting the characteristics of the signals accurately. These figures clearly show dissimilarity between magnitudes and phase angles of the two generators 2 and 3, although the speed of these generators are similar as shown in Fig. 19.

However, by applying DFT to the above speed signals (which is done in the proposed methodology) the magnitudes and phase angles of low frequency oscillations (below 1 Hz) will be obtained as shown in Fig. 23. It can be seen that more frequencies below 1 Hz are examined by DFT. More importantly, Fig. 23 shows a very high similarity of magnitudes as well as phase angles compared to Prony analysis results shown in Figs. 20 to 22, demonstrating the better performance of DFT over Prony analysis in terms of same-frequency-based comparison applications.

Fig. 23. Modes obtained by DFT, (a) magnitude, and (b) phase angle

3.3.2. Performance Comparison with Other Clustering Methods

In the literature, several clustering algorithms have been proposed for clustering the generators of a power system. As stated in Section 1, DBCXV has advantages over other methods which make it more suitable for coherency evaluation. To show the efficacy of DBCXV, a performance comparison between the results of DBCXV and the results of applying KM, FCM and PAM has been provided here. In doing so, all KM, FCM, and PAM algorithms have been coded in MATLAB environment and have been applied to both cases simulated in Sections 3.1 and 3.2. It has to be noted that, KM, FCM and PAM are iteration-based algorithms and use random selection in their first iteration. This means that their solution is not always correct nor fixed and therefore it is advised to run them multiple times to find the best solution. Moreover, KM, FCM and PAM algorithms need to have a prior assumption on the number of clusters. Besides, although there have been extra techniques, such as subtractive clustering algorithm, for proper selection of cluster centers in the first iteration of these algorithms as well as proper determination of number of clusters, here the focus is on the clustering algorithms themselves and no preprocessing step is considered.

Table 3 details the results of 5000 runs carried out for both cases using KM, FCM and PAM algorithms. The number of clusters in cases 1 and 2 has been set on 3 and 9 respectively for all three algorithms since these algorithms need a prior assumption on the number of clusters. As it can be seen, in case #1 all algorithms could find the correct solution in less than 40 % of runs. This get worse in case #2 where the best performance is achieved by the PAM algorithm (only 2 % of runs led to the correct solution). This comparison demonstrates that using traditional clustering algorithms for the problem of online measurement-based coherency evaluation could lead to erroneous results, while the proposed DBCXV algorithm has successfully overcome the disadvantages of these algorithms and has effectively found the correct solution.

Table 3. Comparison between the results obtained by other clustering methods

4. Conclusion

In this paper, a new approach for coherency evaluation in power systems was presented. In the proposed approach, low-frequency modes including inter-area and local modes were selected to form a complex vector for each generator. Then, the correlation coefficients between each pair of complex vectors were considered as data points in the complex plane. In addition, a new density-based unsupervised learning approach was introduced in the paper for clustering the complex vectors of generators. The new clustering approach, which is called DBCXV, not only is suitable for the problem of the paper, but it can be applied to other problems where clustering of complex vectors is required. The results of applying the proposed approach on two cases showed that it can evaluate the coherency between generators and discover clusters of coherent generators fast and accurately, which makes it suitable for online applications.

References

[1] Qiu, X. Ren, Y. Suganthan, P. N., et al. "Empirical mode decomposition based ensemble deep learning for load demand time series forecasting", *Applied Soft Computing*, **54**, pp. 246-255 (2017).

[2] Reche, E.A. de Sousa, J.V. Coury, D.V., et al. "Data mining-based method to reduce multiple estimation for fault location in radial distribution systems", *IEEE Transactions on Smart Grid,* **10**(4), pp. 3612-3619 (2018).

[3] Ahmed, S. Lee, Y. Hyun, S.H., et al. "Unsupervised machine learning-based detection of covert data integrity assault in smart grid networks utilizing isolation forest", *IEEE Transactions on Information Forensics and Security*, **14**(10), pp. 2765-2777 (2019).

[4] Wang, B. Fang, B. Wang, Y., et al. "Power system transient stability assessment based on big data and the core vector machine", *IEEE Transactions on Smart Grid,* **7**(5), pp. 2561-2570 (2016).

[5] Koochi, M.H.R. Esmaeili, S. and Dehghanian, P., February. "Coherency detection and network partitioning supported by wide area measurement system", *2018 IEEE Texas power and energy conference (TPEC)*, pp. 1-6 (2018).

[6] Singh, P. Yadav, R. Pradhan, A.K., et al. "Fundamental factors influencing bus coherency in distribution networks with distributed energy resources", *International Journal of Electrical Power & Energy Systems*, **147**, pp.108778 (2023).

[7] Ourari, M.L. Dessaint, L.A. and Do, V.Q. "Generating units aggregation for dynamic equivalent of large power systems", In IEEE Power Engineering Society General Meeting,. pp. 1535-1541 (2004).

[8] Miah, A. M. “Dynamic reduction of power systems using a simple equivalent”, in Proc. IEEE Power Syst. Conf. Expo., Atlanta, GA, USA, , pp. 1410–1417 (2006).

[9] Yuan, Z. Xia, T. Zhang, Y., et al. “Inter-area oscillation analysis using wide area voltage angle measurements from FNET”, *IEEE PES General Meeting*, Minneapolis, MN, USA, pp. 1-7 (2010).

[10] Mehrzad, A. Darmiani, M. Mousavi, Y., et al. “An Efficient Rapid Method for Generators Coherency Identification in Large Power Systems”, in *IEEE Open Access Journal of Power and Energy*, **9**, pp. 151-160 (2022).

[11] Fritzsch J. and Jacquod, P. “Long Wavelength Coherency in Well Connected Electric Power Networks” in *IEEE Access*, **10**, pp. 19986-19996 (2022),

[12] De Tuglie, E., Iannone, S.M., Torelli, F. “A coherency recognition based on structural decomposition procedure”, *IEEE Trans. Power System*, **23**, pp. 555–563 (2008)

[13] Tyuryukanov, I. Popov, M. van der Meijden, M.A., et al. “Slow Coherency Identification and Power System Dynamic Model Reduction by Using Orthogonal Structure of Electromechanical Eigenvectors”, *IEEE Transactions on Power Systems*, **36**(2), pp. 1482-1492 (2021).

[14] Amini Badr, A. Safari, A. Najafi Ravadanegh, S. “Segmentation of interconnected power systems considering microgrids and the uncertainty of renewable energy sources”, *IET Gener. Transm. Distrib.* **17**, pp. 3814–3827 (2023).

[15] Avdakovic, S. Becirovic, E. Nuhanovic, A., et al. “Generator coherency using the wavelet phase difference approach”, *IEEE Trans. Power System*, **29**(1), pp. 271–278(2014).

[16] Senroy, N. “Generator coherency using the Hilbert–Huang transform”, *IEEE Trans. Power System*, **23**(4), pp. 1701–1708 (2008).

[17] Khalil, A.M., Iravani, R. ‘A dynamic coherency identification method based on frequency deviation signals’, *IEEE Trans. Power Syst.*, 2016, **31**, pp. 1779–1787

[18] Jiang, T., Jia, H., Yuan, H., *et al.*: ‘Projection pursuit: a general methodology of wide-area coherency detection in bulk power grid’, *IEEE Trans. Power Syst.*, 2016, **31**, pp. 2776–2786

[19] Jafarzadeh, S. Genc, I. and Nehorai, A. ‘‘Real-time transient stability prediction and coherency identification in power systems using Koopman mode analysis”, *Electric Power Systems Research*, **201**, pp.107565 (2021).

[20] Koochi, M. H. R. Esmaeili, S. and Ledwich, G. “Taxonomy of coherency detection and coherency-based methods for generators grouping and power system partitioning”, *IET Generation, Transmission & Distribution*, **13**(12), pp. 2597-2610 (2019).

[21] Chamorro, H.R. Gomez- Diaz, E.O. Paternina, M.R., et al. “Power system coherency recognition and islanding: Practical limits and future perspectives”, *IET Energy Systems Integration*, **5**(1), pp.1-14 (2023).

[22] Jordan, M. I., & Mitchell, T. M. “Machine learning: Trends, perspectives, and prospects” *Science*, **349**(6245), pp. 255-260 (2015).

[23] Ray, S. “A quick review of machine learning algorithms” *2019 International conference on machine learning, big data, cloud and parallel computing (COMITCon),* pp. 35-39, (2019).

[24] Guo T. and Milanović, J. V. “Online identification of power system dynamic signature using PMU measurements and data mining” *IEEE Transactions on Power Systems,* **31**(3), pp. 1760-1768 (2015).

[25] Koochi, M.H.R. Dehghanian, P. Esmaeili, S., et al. “A synchrophasor-based decision tree approach for identification of most coherent generating units”, *IECON 2018-44th Annual Conference of the IEEE Industrial Electronics Society*, pp. 71-76 (2018).

[26] Koochi, M. H. R. Esmaeili, S. and Fadaeinedjad R. “New phasor-based approach for online and fast prediction of generators grouping using decision tree”, *IET Generation, Transmission & Distribution,* **11**(6), pp. 1566-1574 (2017).

[27] Liu, F. Yin, Y. and Li, B. “A novel probabilistic framework with interpretability for generator coherency identification”, *International Journal of Electrical Power & Energy Systems*, **4**(1), pp. 13-17 (2022).

[28] Sadeghi, M. Akbari, H. Mousavi, S., et al. “Coherency-Based Supervised Learning Approach for Determining Optimal Post-Disturbance System Separation Strategies”, *IEEE Access*, **11**, pp. 5894-5907 (2023)

[29] Rezaeian Koochi, M.H. Hemmatpour, M.H. and Dehghanian, P. “Clustering in Power Systems Using Innovative Machine Learning/Deep Learning Methods”, *Application of Machine Learning and Deep Learning Methods to Power System Problems*, pp.151-176 (2021).

[30] Davodi, M. Modares, H. Reihani, E., et al. “Coherency approach by hybrid PSO, K-Means clustering method in power system”, *2008 IEEE 2nd International Power and Energy Conference*, pp. 1203-1207 (2008).

[31] Tang, K. and Venayagamoorthy, G.K. “Online coherency analysis of synchronous generators in a power system”, *ISGT 2014*, pp. 1-5 (2014).

[32] Wang, S.C. and Huang, P.H. “Fuzzy c-means clustering for power system coherency”, IEEE International Conference on Systems, Man and Cybernetics, **3**, pp. 2850-2855 (2005).

[33] Kamwa, I., Pradhan, A.K. and Joós, G. “Automatic segmentation of large power systems into fuzzy coherent areas for dynamic vulnerability assessment”, *IEEE Transactions on Power Systems*, **22**(4), pp. 1974-1985 (2007).

[34] Kamwa, I. Pradhan, A.K. and Joós, G. “Fuzzy partitioning of a real power system for dynamic vulnerability assessment”, *IEEE Transactions on Power Systems*, **24**(3), pp. 1356-1365 (2009).

[35] Rezaeian, M.H., Esmaeili, S. and Fadaeinedjad, R. “Generator coherency and network partitioning for dynamic equivalencing using subtractive clustering algorithm”, *IEEE Systems Journal*, **12**(4), pp. 3085-3095 (2017).

[36] Pyo, G.C., Park, J.W. and Moon, S.I. “Coherency identification of generators using a PAM algorithm for dynamic reduction of power systems”, *Energies*, **5**(11), pp. 4417-4429 (2012).

[37] De Caro, F. Pepiciello, A. Milano, F., et al. “Measurement-based coherency detection through Monte Carlo Consensus Clustering”, *Electric Power Systems Research*, **216**, pp.109075 **(**2023).

[38] Lugnani, L., Paternina, M.R.A., Dotta, D., et al. “Power System Coherency Detection From Wide-Area Measurements by Typicality-Based Data Analysis”, *IEEE Transactions on Power Systems*, **37**(1), pp. 388-401 (2022).

[39] Lugnani, L., Dotta, D., Paternina, M.R., et al. “Real-time Coherency Identification using a Window-Size-Based Recursive Typicality Data Analysis” *2022 International Conference on Smart Grid Synchronized Measurements and Analytics (SGSMA)*, Split, Croatia, pp. 1-6 (2022)

[40] Ortiz-Bejar, J., Paternina, M.R.A., Zamora-Mendez, A., et al. “Power system coherency assessment by the affinity propagation algorithm and distance correlation” *Sustainable Energy, Grids and Networks*, **30**, pp.100658 (2022)

[41] Babaei, M., Muyeen, S.M. and Islam, S. “Identification of coherent generators by support vector clustering with an embedding strategy”, *IEEE Access*, **7**, pp. 105420-105431 (2019)

[42] Barocio, E., Korba, P., Sattinger, W., et al. “Online coherency identification and stability condition for large interconnected power systems using an unsupervised data mining technique”, *IET Generation, Transmission & Distribution*, **13**(15), pp.3323-3333 (2019).

[43] Ester, M., Kriegel, H.P., Sander, J., et al. “A density-based algorithm for discovering clusters in large spatial databases with noise”, **96**(34), pp. 226-231(1996)

[44] Vahidnia, A., Ledwich, G., Palmer, E., et al. “Generator coherency and area detection in large power systems”, *IET generation, transmission & distribution*, **6**(9), pp. 874-883 (2012).

[45] Jonsson, M. Begovic, M. and Daalder, J. “A new method suitable for real-time generator coherency determination”, *IEEE Transactions on Power Systems*, **19**(3), pp. 1473-1482 (2004).

[46] Oppenheim, A. V. and Schafer, R. W. “Digital signal processing(Book)”, *Research supported by the Massachusetts Institute of Technology, Bell Telephone Laboratories, and Guggenheim Foundation. Englewood Cliffs, N. J., Prentice-Hall, Inc.* (1975).

[47] Agrawal, R. Faloutsos, C. and Swami A. “Efficient similarity search in sequence databases”, 4th International Conference, FODO'93 Chicago, Illinois, USA, October 13–15, 1993 Proceedings 4 (pp. 69-84). Springer Berlin Heidelberg (1993).

[48] Znidi, F., Davarikia, H., Arani, M., et al. “Coherency detection and network partitioning based on hierarchical DBSCAN”, *2020 IEEE Texas Power and Energy Conference (TPEC)*, pp. 1-5, IEEE (2020).

[49] Rogers, G. “Power system oscillations”, Springer Science & Business Media (2012).

[50] Power System Toolbox (PST), available online at https://sites.ecse.rpi.edu/~chowj/PST_2020_ Aug_10.zip

[51] Tripathy, P. Srivastava, S.C. and Singh, S.N. “A divide-by-difference-filter based algorithm for estimation of generator rotor angle utilizing synchrophasor measurements”, *IEEE Transactions on Instrumentation and Measurement*, **59**(6), pp. 1562–1570 (2010).

[52] Yan, J. Liu, C.C. and Vaidya, U. “PMU-based monitoring of rotor angle dynamics”, *IEEE Transactions on Power System*, **26**(4), pp. 2125–2133 (2011)

[53] Lei, H., Huang, Y., Jiang, L., et al. “Research on Fourier and Prony analysis algorithm of inter-harmonic in power system”, *Energy Reports*, ***8***, pp.728-737 (2022)

[54] Fernández Rodríguez, A., de Santiago Rodrigo, L., López Guillén, E., et al. “Coding Prony’s method in MATLAB and applying it to biomedical signal filtering”, *BMC bioinformatics*, **19**, pp.1-14 (2018)

Table List

Table 1. Generator grouping scheme obtained in Case 1

Group No.	**Generators**
1	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15
2	9
3	14
4	16

Table 2. Generator grouping scheme obtained in Case 2

Group No.	**Gens.**	**Group No.**	**Gens.**	**Group No.**	**Gens.**
1	4, 5, 6, 7,8	4	2, 3	7	10
2	11,12	5	1	8	13
3	15,16	6	9	9	14

Table 3. Comparison between the results obtained by other clustering methods

Case #1			**Case #2**		
Algorithm	**Total No. of runs**	**No. of runs with the correct solution**	**Algorithm**	**Total No. of runs**	**No. of runs with the correct solution**
KM	5000	2000	KM	5000	16
FCM	5000	803	FCM	5000	0
PAM	5000	422	PAM	5000	103

Figure List

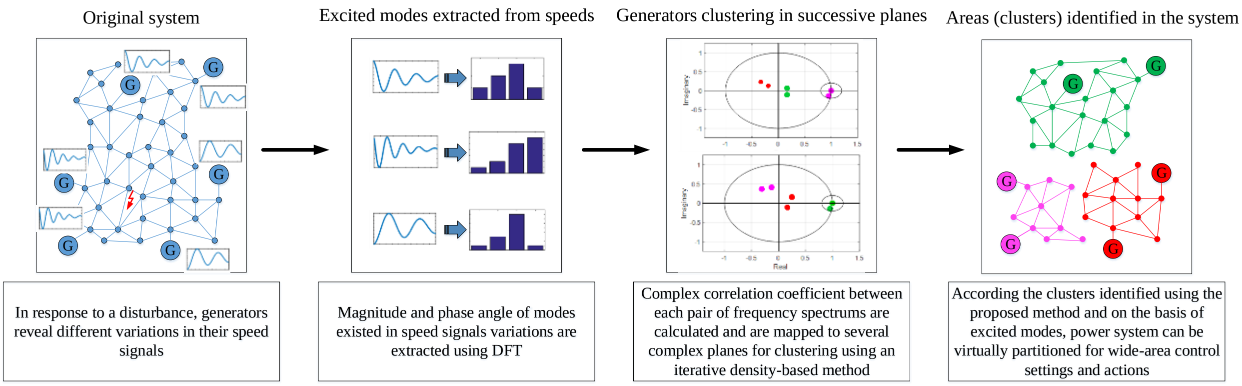


Fig. 1. Conceptual illustration of the proposed approach

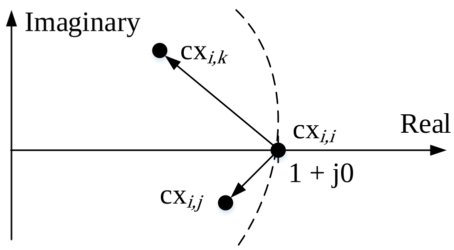


Fig. 2. Definition of coherency in the complex plane

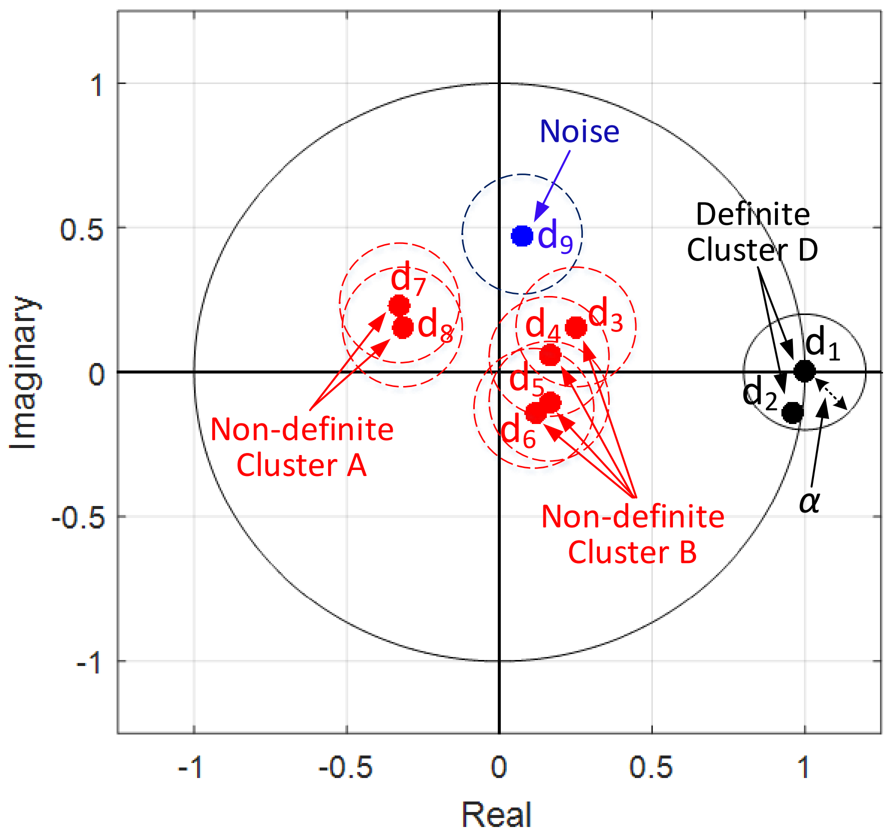


Fig. 3. Example of noise, definite cluster, and non-definite cluster in the complex plane

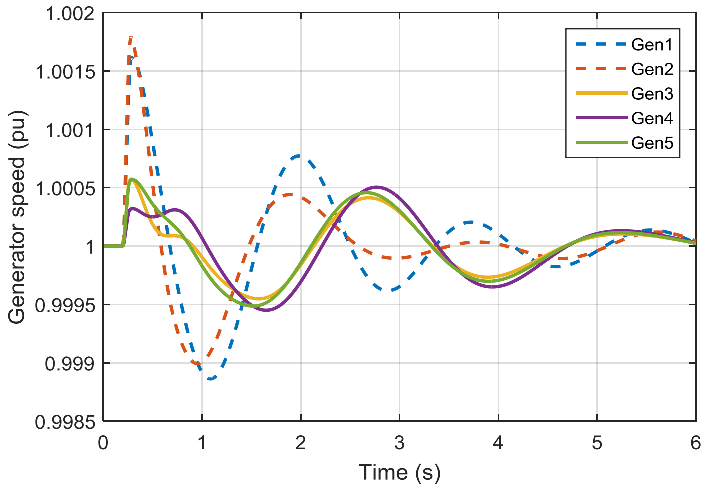


Fig. 4. Speed variations of generators (example)

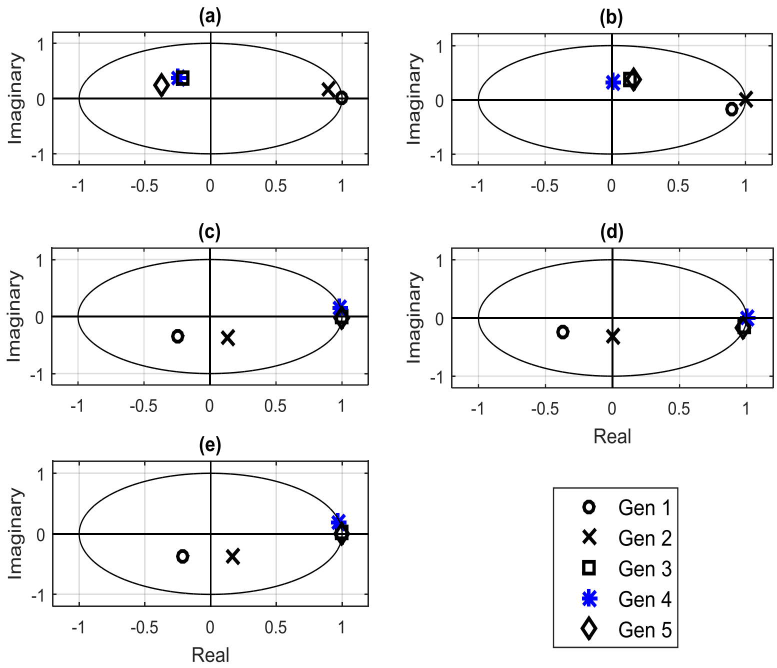


Fig. 5. Correlation values with respect to (a) generator 1, (b) generator 2, (c) generator 3, (d) generator 4, and (e) generator 5.

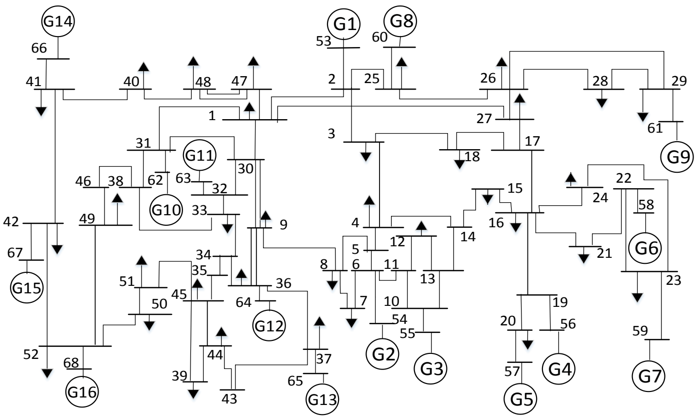


Fig. 6. Line diagram of the test system

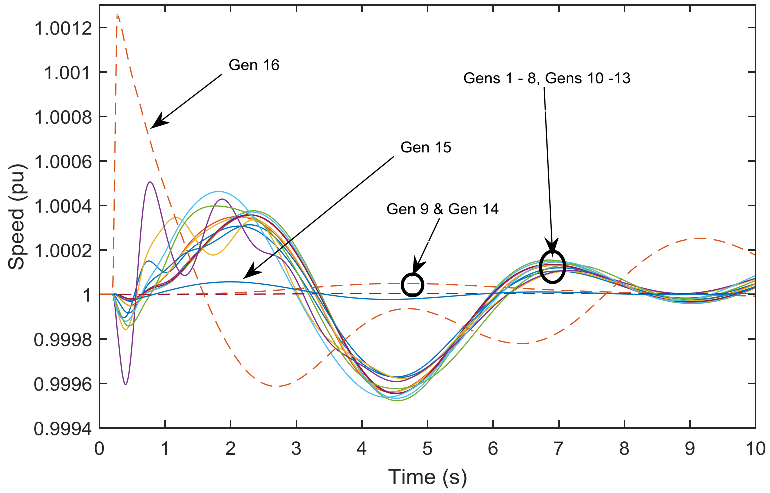


Fig. 7. Speed variations of generators (Case 1)

Generator no.	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12	G13	G14	G15	G16
G1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	0
G2	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	0
G3	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	0
G4	1	1	1	1	1	1	1	1	0	0	0	1	0	0	0	0
G5	1	1	1	1	1	1	1	1	0	0	0	1	0	0	0	0
G6	1	1	1	1	1	1	1	1	0	0	0	1	0	0	0	0
G7	1	1	1	1	1	1	1	1	0	1	0	1	0	0	0	0
G8	1	1	1	1	1	1	1	1	0	1	1	1	0	0	1	0
G9	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0
G10	1	1	1	0	0	0	1	1	0	1	1	1	1	0	0	0
G11	1	1	1	0	0	0	0	1	0	1	1	1	1	0	0	0
G12	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	0
G13	1	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0
G14	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0
G15	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0
G16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

Fig. 8. Elements of matrix *B* (Case 1)

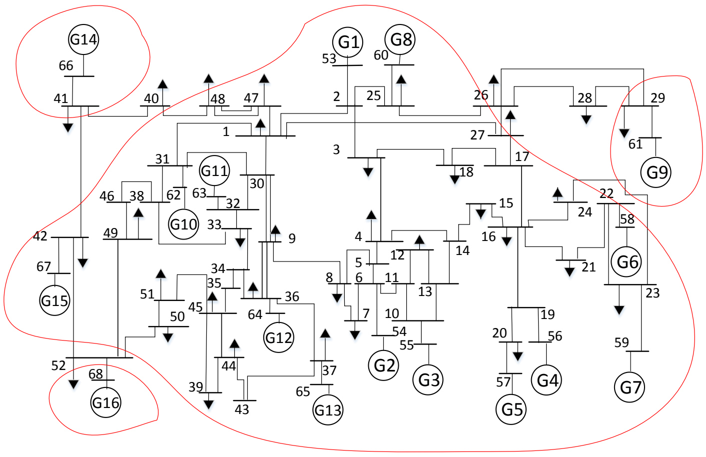


Fig. 9. Groups of generators discovered in Case 1

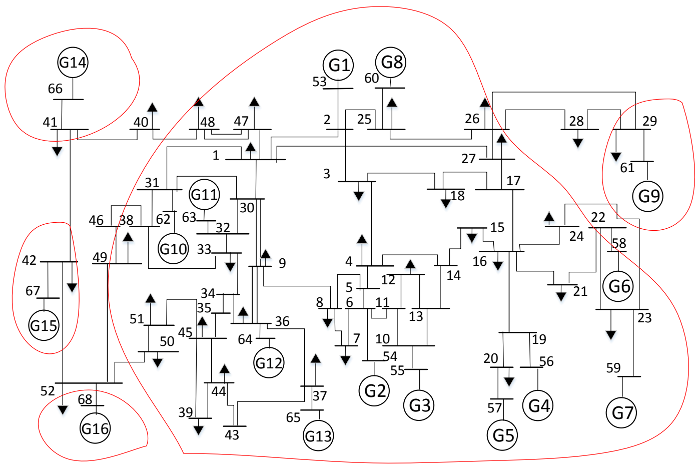


Fig. 10. Generators clustering with α = 0.1 (Case 1)

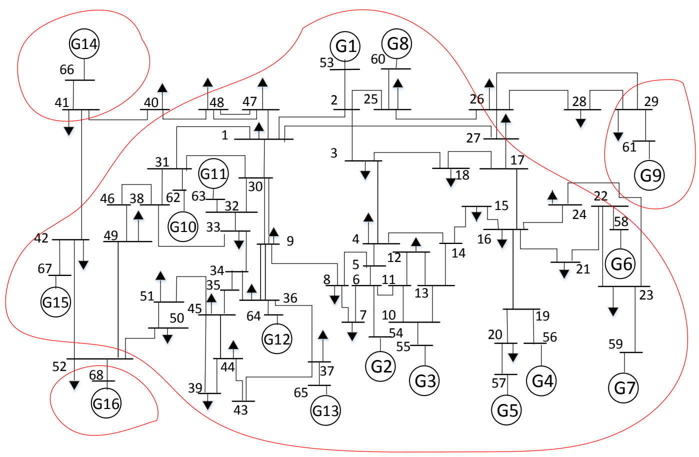


Fig. 11. Generators clustering with α = 0.3 (Case 1)

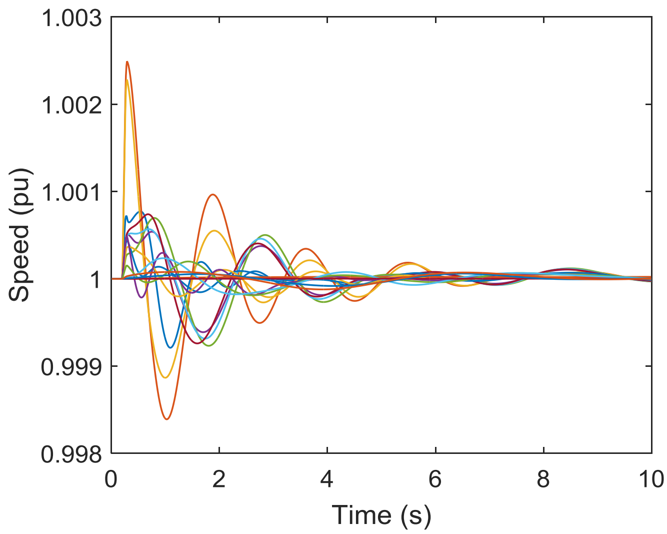


Fig. 12. Speed variations of generators (Case 2)

Generator no.	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12	G13	G14	G15	G16
G1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
G2	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
G3	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
G4	0	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0
G5	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0
G6	0	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0
G7	0	0	0	1	0	1	1	1	0	0	0	0	0	0	0	0
G8	0	0	0	1	0	1	1	1	0	0	0	0	0	0	0	0
G9	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
G10	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
G11	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0
G12	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0
G13	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
G14	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
G15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
G16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1

Generator no.

Fig. **13**. Elements of matrix *B* (Case 2)

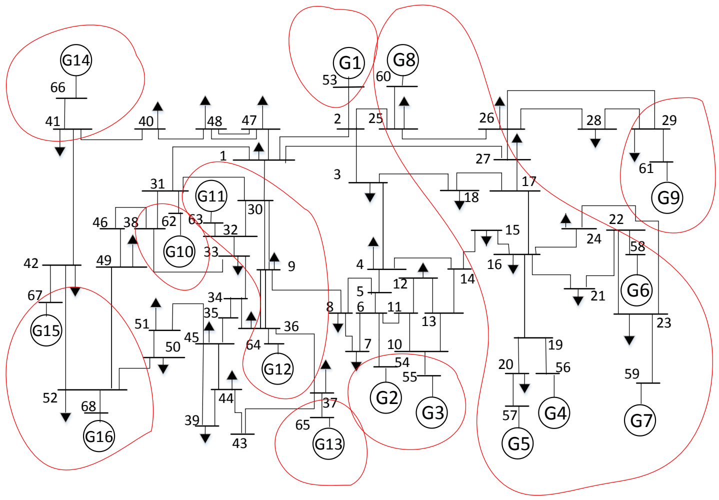


Fig. 14. Groups of generators discovered in Case 2

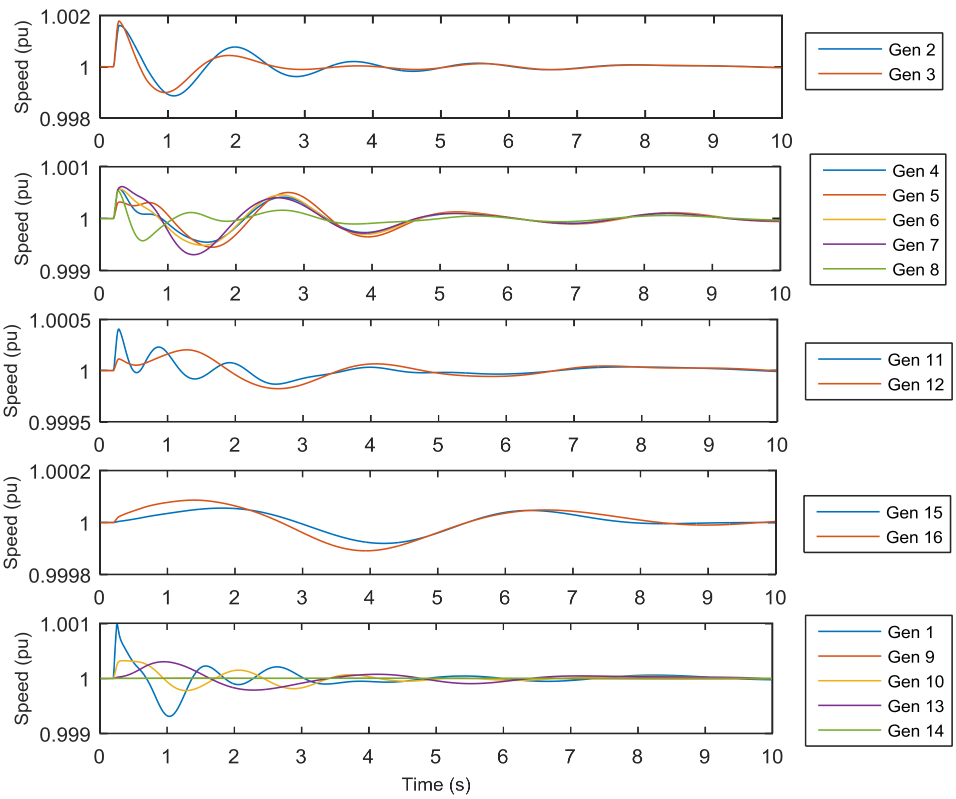


Fig. 15. Speed variations of generators (Case 2)

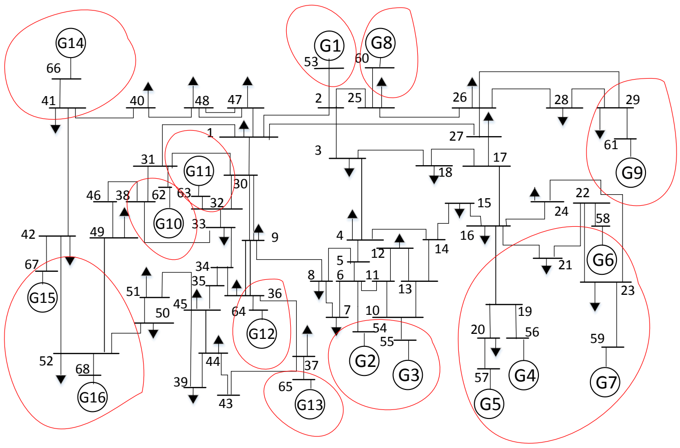


Fig. 16. Generators clustering with α = 0.1 (Case 2)

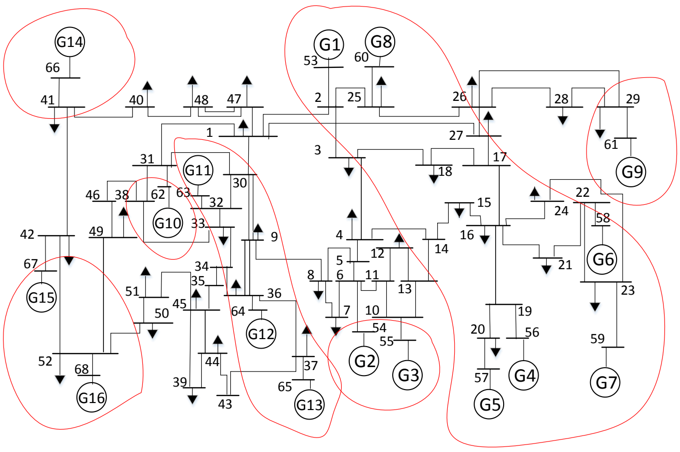


Fig. 17. Generators clustering with α = 0.3 (Case 2)

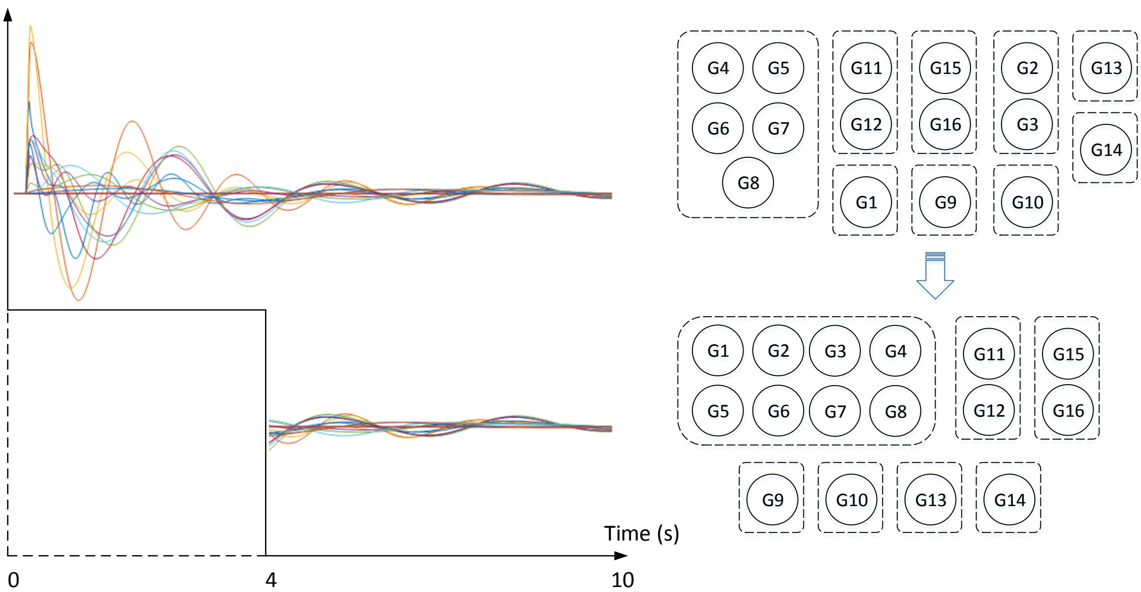


Fig. 18. Coherent generators obtained in two different time frames (case 2)

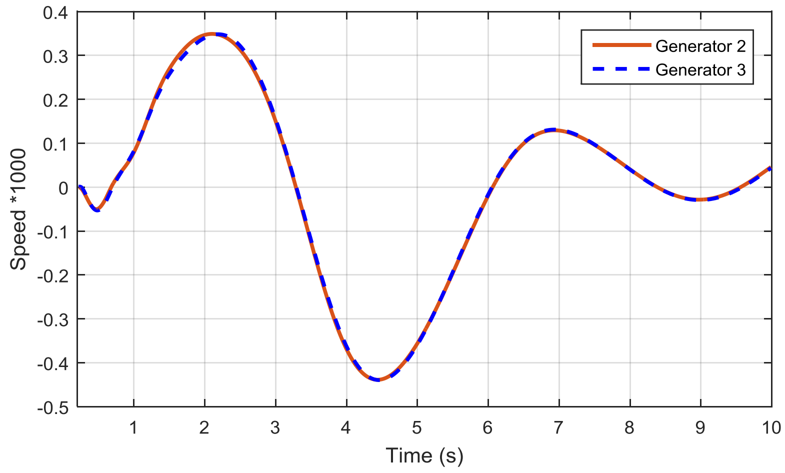
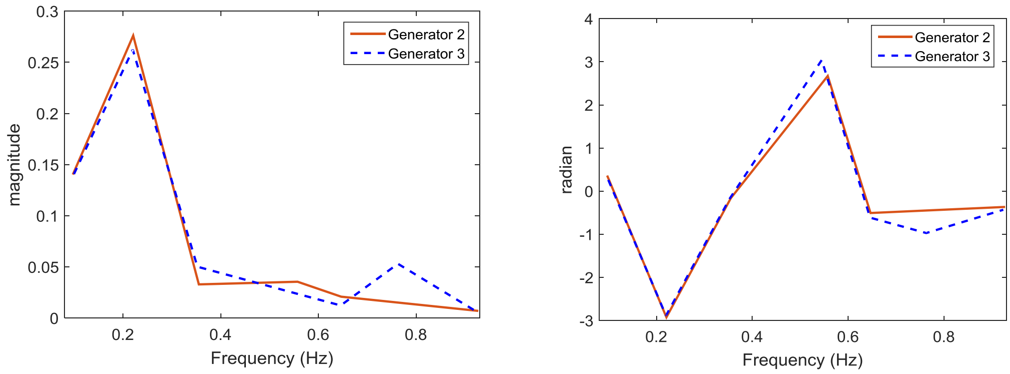
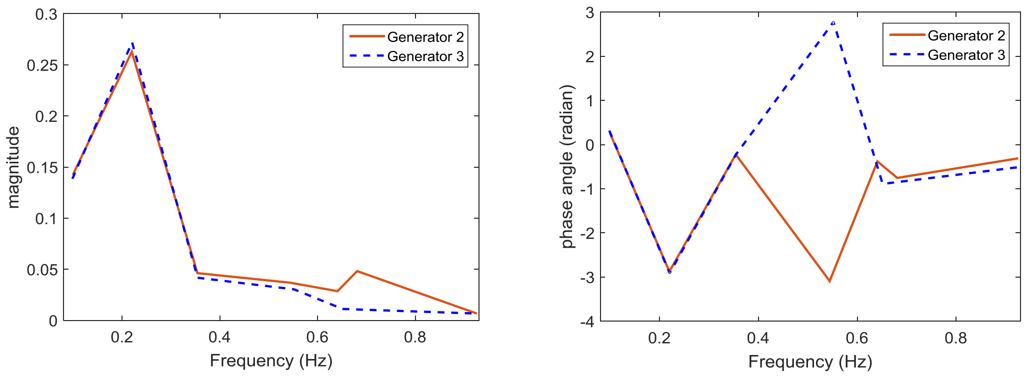


Fig. 19. Speed signals of generators 2 and 3 in Case 1



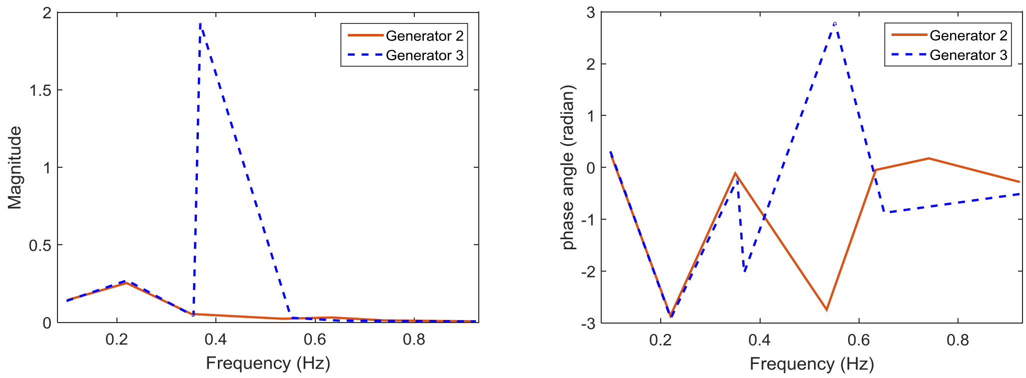
(a) (b)

Fig. 20. Modes obtained by TLS Prony analysis, (a) magnitude, and (b) phase angle



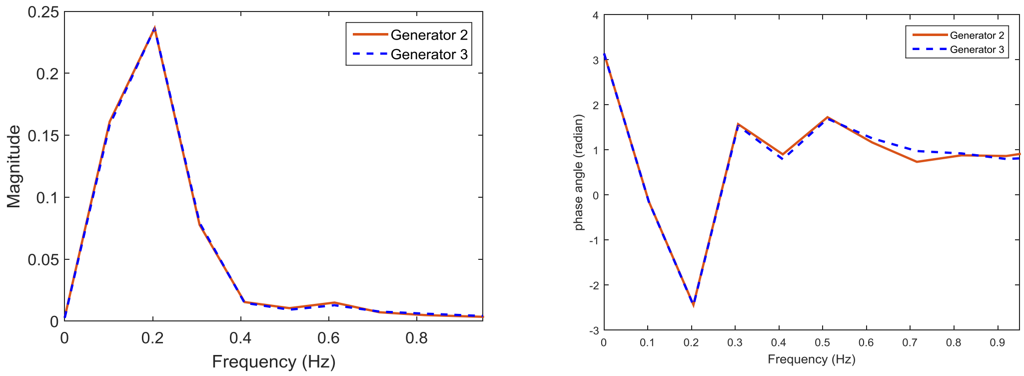
(a) (b)

Fig. 21. Modes obtained by LS Prony analysis, (a) magnitude, and (b) phase angle



(a) (b)

Fig. 22. Modes obtained by classic Prony analysis, (a) magnitude, and (b) phase angle



(a) (b)

Fig. 23. Modes obtained by DFT, (a) magnitude, and (b) phase angle

Brief technical biography

Mohammad Hossein Rezaeian Koochi the M.S. degree in Electrical Engineering from Iran University of Science and Technology (IUST), Tehran, Iran, in 2012, and the Ph. D. degree in Electrical Engineering from the Shahid Bahonar University of Kerman, Kerman, Iran, in 2017. In 2022, he was a Postdoctoral Research Fellow at the Queensland University of Technology, Brisbane, Australia. His research interests include energy management systems, smart grids, wide area measurement and control systems, power system dynamics and control, and pattern recognition and data mining in power systems.

Mohammad Hasan Hemmatpour received B.S., M.S., and Ph.D. degrees in electrical engineering from the Shahid Bahonar University of Kerman, Kerman, Iran, as an honor student in 2009, 2012, and 2016, respectively. He is an Assistant Professor at the Electrical Engineering Department of Jahrom University, Jahrom, Iran. His research interests include the operation, control, and analysis of microgrids and distribution systems, optimization techniques, and renewable energy modeling.