POLITECNICO DI TORINO
Repository ISTITUZIONALE

Review of Clustering Methods for Slow Coherency-Based Generator Grouping

(Article begins on next page)

25 April 2024

Review of Clustering Methods for Slow Coherency-Based Generator Grouping

Gianfranco Chicco*

Dipartimento Energia "Galileo Ferraris", Politecnico di Torino, Torino, Italy

Abstract — **Slow coherency is one of the most relevant concepts used in power systems dynamics to group generators that exhibit similar response to disturbances. Among the approaches developed for generator grouping based on slow coherency, clustering algorithms play a significant role. This paper reviews the clustering algorithms applied in model-based and data-driven approaches, highlighting the metrics used, the feature selection, the types of algorithms and the comparison among the results obtained considering simulated or measured data.**

Index Terms: **clustering, contingency screening, data-driven, distance, feature selection, generator grouping, model-based, phasor measurement unit, power system dynamics, similarity metric, slow coherency, stability analysis.**

NOMENCLATURE

COA	Centre of Angle
COI	Centre of Inertia
COIFD	Centre of Inertia Frequency Deviation
DBI	Davies–Bouldin Index
DBSCAN	Density-Based Spatial Clustering of applications with Noise
DCD	Dynamic Coherency Determination
DFT	Discrete Fourier Transform
DMD	Dynamic Mode Decomposition
DTW	Dynamic Time Warping
FCM	Fuzzy c-Means
FCMdd	Fuzzy c-Medoids
ICA	Independent Component Analysis
KPCA	Kernel Principal Component Analysis
PCA	Principal Component Analysis
PMU	Phasor Measurement Unit
SOM	Self-Organising Maps
SVC	Support Vector Clustering
SVD	Singular Value Decomposition
WSSE	Weighted Sum Squared Error

I. INTRODUCTION

The evolution of the computational frameworks for applying efficient data analytics through machine learning has a significant impact on the formulation and update of the computational codes used in many applications. The power system sector is heavily affected by this evolution, which is marking new differences between the two most used approaches:

1. The *model-based* approach, in which a suitable model is constructed by connecting a number of elements that represent the physical components with different degrees of approximation. Suitable parameters have to be associated to each element, and the solution is determined through simulations.

2. The *data-driven* approach (also indicated as *measurement-based* or *signal-based* approach), in which the computations are based on the available data gathered from the field in actual situations and are carried out by using machine learning approaches that do not need the construction of the system model nor the parameter setup, and do not depend on theoretical assumptions.

Advantages and drawbacks of these approaches are summarised in [1], together with the characterisation of different aspects that explain data consistency in terms of data characteristics, data quality, and information quality. The scientific community is contributing to the debate about the convenience of one approach or the other, by developing new computational procedures in both directions.

Concerning power system dynamics, both approaches share a major issue, given by the variability in time of the

* Corresponding author.
  E-mail: gianfranco.chicco@polito.it

data used in dynamic conditions. In particular:
- Methods based on the model-based approach are highly sensible on the accuracy of the model and parameters used to represent dynamics at different time scales. Dynamic models are available at different levels of detail, and dynamic reduction methods are available and are incorporated in the dynamic simulation tools. Modal analysis is applied, where the power system model is linearised around an equilibrium point, and the oscillation modes are extracted by calculating eigenvalues and eigenvectors from the small-signal analysis of the linearised power system model [2]. Further insights may come from the application of Koopman mode analysis, which is a nonlinear technique that requires no linearisation nor assumptions of scale separation [3, 4].
- Methods based on the data-driven approach require time series of data with appropriate resolution in time and data window length, to form a reasonable history, benchmark or comparative dataset. These methods are sensible to the consistency of the data gathered from the field. However, possible issues depend on accuracy of the measurements, sensitivity to disturbances, and in particular to the occurrence of major changes in the operational conditions of the network (e.g., changes in topology, loss of a generator, or large net power variations [5]). Following major disturbances, the dataset is typically partitioned into portions that represent the system operation before and after the change, by identifying the starting instant of the event and removing pre-event data. In particular, the methods based on similarities among time series can be adapted and used online, while the methods that carry out mode estimation or use information about the frequency spectrum (e.g., of inter-area oscillations) require longer computation time and are thus applied offline [6].

The determination of the dynamic equivalents of the generators connected to the power system power is a challenging and insightful line of research. Starting from dynamic models developed during the years, in the Seventies two main approaches emerged for the identification of similar behaviour of the generators in a multi-machine power system, especially during contingencies, and the formation of dynamic equivalents. One of these approaches is based on *modal analysis* [7] [8], while the other approach is based on the notion of *coherency* [9]. Two (or more) generators are considered to be coherent if their rotor angles follow the same evolution in time, both in normal conditions and after the occurrence of disturbances. The identification of coherency requires to set up appropriate criteria based on distance measures [10]. Many algorithms have been formulated for both modal analysis [11] and for coherency assessment, with various approaches [12, 13], among which clustering algorithms that exploit time-series similarity-based techniques. The creation of reduced dynamic equivalent models of

generators is a major asset to increase the computational speed and enable real-time applications. A discussion on methods for dynamic reduction can be found for example in [14]. Among these methods, the application of clustering techniques is an alternative to deal with this non-linear and large-scale problem in power system dynamics by using a suitable set of representative features (e.g., variables or indicators). Coherency-based generator grouping is the problem referring to power system dynamics for which most contributions that exploit clustering algorithms are available in the literature.

This paper addresses the application of clustering algorithms in the methods used to deal with the generator grouping based on coherency aspects. The application of clustering algorithms deals with the need of identifying groups inside the system that exhibit some similarity, and this similarity may be useful for specific purposes. Clustering algorithms have shown particularly promising results, and more refined variants are under continuous development.

In general terms, the main requirements of the clustering algorithms include:
• The choice of appropriate *features*.
• The incorporation of a suitable notion of *distance* for similarity assessment.
• The ability to operate with *high-dimensional* data set.
• The effectiveness of the clustering algorithm (assessed through *clustering validity* indices).
• The ability to discover possible *outliers* (when needed).
• The ability to estimate the *best number of clusters*.

In power system dynamic studies, clustering algorithms are applied to both model-based and data-driven approaches. In general, the challenging objective is to test and validate the proposed approaches as close as possible to real-time applications in actual systems. In model-based approaches, in addition to software simulation tools, examples oriented to practical validation have been explored with a hybrid analog-digital power system simulator [15], while solutions that incorporate the use of real-time data in power system models have recently emerged with the use of real-time simulators, which allow the implementation of hardware-in-the-loop solutions [6].

The aim of this paper is to illustrate and discuss why, in which phase of the analysis and how the clustering algorithms are used for solving the generator grouping on the basis of their dynamic behaviour. A selected set of references is considered, including recent references together with a number of historical references, without the intention to provide an exhaustive list of literature contributions.

The next sections of the paper are organised as follows. Section II discusses the notions of distance and metrics used in the clustering algorithms for analysing power system dynamics. Section III recalls the notion of slow coherency and its use for generator grouping. Section IV addresses in more details the feature selection stage. Section V deals

with the clustering algorithms used. Section 6 illustrates the comparisons carried out among the results obtained from literature contributions. The last section contains the concluding remarks.

II. METRICS USED IN THE CLUSTERING ALGORITHMS

For the purpose of the contents presented in this paper, the input dataset is composed of M time series with T data points each, corresponding to a given feature. For $m = 1, \ldots, M$, the time series of the individual feature ψ_m is represented as

$$\psi_m = \{\psi_{mt}, t = 1, \ldots, T\}. \tag{1}$$

The metrics used are based on determining the distance between two time series. Since the dataset is discrete, discrete versions of the metrics are used. Without loss of generality, the distances are indicated in the sequel by considering two feature vectors. The same formulations are applicable to calculate the distance between a feature vector and a centroid, by replacing one of the feature vectors with the centroid.

In a clustering procedure, the choice of the metric to calculate the distances between time series is crucial [16]. In particular, four types of metrics can be identified:

1 Metrics based on the calculation of the distance for points located at the *same instant*: these metrics are relevant if it is important to maintain the identity of the individual time instants, calculating distances between time series on the vertical axis. The most classical metrics of this type are based on the Minkowski distances, defined for an integer parameter $p > 0$:

$$d_M^{(p)}\left(\psi_i, \psi_j\right) = \left(\sum_{t=1}^{T} \left|\psi_{it} - \psi_{jt}\right|^p\right)^{1/p}, \tag{2}$$

where in particular the distance for $p = 2$ is the Euclidean distance:

$$d_M^{(2)}\left(\psi_i, \psi_j\right) = d_E\left(\psi_i, \psi_j\right) = \sqrt{\sum_{t=1}^{T} \left|\psi_{it} - \psi_{jt}\right|^2}. \tag{3}$$

When a weight factor w_{ijt} is introduced, which depends on the features, time, or combinations of them, the weighted versions of the Minkowski distances become:

$$d_{Mw}^{(p)}\left(\psi_i, \psi_j\right) = \left(\sum_{t=1}^{T} w_{ijt} \left|\psi_{it} - \psi_{jt}\right|^p\right)^{1/p}. \tag{4}$$

In particular, $d_{Mw}^{(2)}\left(\psi_i, \psi_j\right)$ is the weighted sum squared error (WSSE).

2 Metrics based on the *proximity between points* in the time series. These distances are used if comparing the time series without the constraint of maintaining the correspondence between the instants of time is acceptable. The distance obtained is a single value that depends on the overall evolution of the trajectories. In this way, time series with similar shape but (slight) shift in time can be characterised by a small distance. Among these metrics, the *Hausdorff distance* is given

by the largest distance from a point belonging to the time series ψ_i to the closest point belonging to the other time series ψ_j. Let us first define the minimum distance between a point $\psi_{it} \in \psi_i$ for a given value of t and the points $\psi_{jq} \in \psi_j$, $q = 1, \ldots, T$, as follows:

$$d\left(\psi_{it}, \psi_j\right) = \min_{\psi_{jq} \in \psi_j} \left\{d_E\left(\psi_{it}, \psi_{jq}\right)\right\}. \tag{5}$$

Likewise,

$$d\left(\psi_{jt}, \psi_i\right) = \min_{\psi_{iq} \in \psi_i} \left\{d_E\left(\psi_{ij}, \psi_{iq}\right)\right\}. \tag{6}$$

The Hausdorff distance is then computed as:

$$d_H\left(\psi_i, \psi_j\right) = $$
$$= \max\left\{\min_{\psi_{it} \in \psi_i} \left\{d\left(\psi_{it}, \psi_j\right)\right\}, \min_{\psi_{jt} \in \psi_j} \left\{d\left(\psi_{jt}, \psi_i\right)\right\}\right\} \tag{7}$$

Moreover, the *Fréchet distance* is the maximum pairwise distance between the points located on the two time series [17]. It is also known as the person-dog metric, being based on the concept of considering a person who walks a dog on a leash. Both the person and the dog follow a separate trajectory but cannot move backwards (this is a conceptual difference with respect to the Hausdorff distance). The Fréchet distance is the length of the shortest leash needed to follow the entire trajectories.

The *Dynamic time warping* (DTW) distance is another possibility to determining the distance between two time series. In DTW, the differences between two time series are determined by changing the connections between the corresponding points in the time series. DTW creates a warping path of minimum local distances and computes the final distance as the mean distance along this path. The length of the warping path may be higher than the length of the time series.

The need for considering a distance notion that does not strictly refer to the same time instants is recognised as well in [18], where an *integrated weighted distance* is used. This distance is based on the definition of three distance functions: (i) "horizontal absolute value"; (ii) weighted distance of "rate of change at adjacent time points"; (iii) weighted distance of "coefficient of variation". The values obtained for each distance function are then subject to min-max normalisation. Finally, the weighted distance matrix is formed by weighting the matrices that contain the entries of the normalised distance functions, by using entropy-based weights.

3 Metrics that consider the *global behaviour* of the time series. These metrics are based on determining the similarity between the time series on the [0,1] scale, then defining the distances as the complement to unity of the similarity. The *correlation* coefficient is the most classical notion of similarity. Since the correlation coefficient ρ varies from −1 to 1, it is

needed to reconduct it to the [0,1] range by calculating $(\rho + 1)/2$. In addition, the *cosine similarity* derives from the definition of the Euclidean internal product and is the ratio between the internal product and the product of the magnitudes. The cosine distance is then calculated as:

$$d_C\left(\psi_i, \psi_j\right) = 1 - \frac{\sum_{t=1}^{T} \psi_{it}\psi_{jt}}{\sqrt{\sum_{t=1}^{T}\psi_{it}^2}\sqrt{\sum_{t=1}^{T}\psi_{jt}^2}} . \qquad (8)$$

In the calculation of the cosine similarity, if the vectors are changed by subtracting the vector means, the resulting version is called *centred cosine similarity*, which corresponds to the Pearson correlation coefficient. Another way to calculate correlations by excluding the mean value is to take only non-zero F harmonic components of the spectrum computed from the Discrete Fourier Transform (DFT), in which the zero-order harmonic component, proportional to the mean value, is not considered [19]. In this case, the features ζ_m for $m = 1,\ldots, M$ are expressed as a complex number, which can be written with the real part $\zeta_m^{Re} = \left\{\zeta_{mv}^{Re}\right\}$ and the imaginary part $\zeta_m^{Im} = \left\{\zeta_{mv}^{Im}\right\}$ for $v = 1, \ldots, F$. The correlation coefficient between ζ_i^{Re} and ζ_j^{Re} is expressed as c_{ij}^{Re} (and with c_{ij}^{Im} considering the imaginary parts). The density function δ_{Sm}, based on the dissimilarity between two entries expressed in an exponential form and depending on the user-defined parameter r that represents the neighbouring radius, is formulated as

$$\delta_{Sm} = \sum_{m=1}^{M} e^{-\left(\frac{\left(1-c_{im}^{Re}\right)^2\left(1-c_{im}^{Im}\right)^2}{(r/2)^2}\right)} . \qquad (9)$$

The *Jaccard similarity* between two sets is generally defined as the ratio of the intersection over the union of the two sets. For two datasets formed by the same number of discrete points, the Jaccard distance is:

$$d_J\left(\psi_i, \psi_j\right) = 1 - \frac{\sum_{t=1}^{T}\min\left\{\psi_{it}, \psi_{jt}\right\}}{\sum_{t=1}^{T}\max\left\{\psi_{it}, \psi_{jt}\right\}} . \qquad (10)$$

4) Metrics that exploit *global information* on the dataset. In general, when it is acceptable to compare two time series without taking into account the relationships between the instants of time, by observing only the global behaviour, it is possible to define metrics based on the probabilistic distribution of the data belonging to the time series. Both probability density functions and cumulative distribution functions can be constructed. In these cases, the connection with time is completely lost, and only the distribution of the amplitudes is relevant. This kind of metrics have not been used in the cases discussed in this paper.

III. Slow Coherency And Generator Grouping

A. Slow coherency

In large power systems, a distinction is made between slow dynamics (due to low frequency oscillations between groups of coherent generators) are fast dynamics (with oscillations at higher frequencies that occur among the generators of the same group). *Slow coherency* is then considered when the oscillatory frequency varies in an indicative range from 0.1 to 0.8 Hz [6], while the transient dynamics of the system are not considered by the slow coherency theory. In [20] slow coherency is associated to the weak coupling among coherent areas, due to low connection or high impedance of the lines among coherent areas, or transmission lines with heavy load. Moreover, in the present systems with distributed generator and load dynamics, the causes of disturbances are increasing with respect to the classical power system supplied by large synchronous generators [21, 22]. Hence, the use of a model-based approach that incorporates all the variables needed becomes more and more challenging. This increases the interest towards exploiting measurement-based (data-driven) approaches.

The rotor angle gives a direct information on the machine dynamics [23]. The speed deviation of generators represents the energy absorbed or delivered by the generator [24]. Particularly relevant are the quantities that can be measured by using data taken from PMUs. A PMU installed at the generator terminals can provide positive sequence voltage measurements. However, the rotor angle cannot be measured directly nor estimated accurately with a PMU [6]. Hence, the rotor angles can be used in model-based approaches, in which they can be calculated by simulations, while for a data-driven approach it could be needed to find other variables.

The calculation of dynamic equivalents based on coherency by considering a model-based approach, in which the relevant variables are the voltage phase angles at the generator terminal nodes or at the generator internal nodes, is described in [25]. In this respect, two generator nodes are defined as coherent if the angular difference of their voltages does not exceed a certain tolerance over a given time interval. The corresponding automatic formation of dynamic equivalents is addressed in [26].

The identification of coherent generators enables the grouping of coherent terminal buses, which is at the basis of numerical procedures for power system model reduction in the model-based approach. In these procedures, after the determination of the coherent generator groups, the network is partitioned into an internal area (in which detailed modelling is desired) and in an external area (the rest of the network, where resorting to reduced modelling is of interest) [14, 27]. In the external area, each coherent generator group is aggregated by maintaining an equivalent generator, and the nodes in that area are reduced by using suitable techniques to form equivalent circuits [28].

B. Slow coherency-based clustering algorithms for generator grouping

Following the above concepts, the application of clustering algorithms has become a viable way to form groups of generators characterised by slow coherency. In [25], a greedy clustering algorithm (i.e., an algorithm that makes direct choices at each step, following a predefined strategy) is applied to the terminal voltage angles (voltage swing curves). Starting from a reference generator, another generator is compared with it by calculating whether all terminal voltage angle variations in a given time interval do not differ more than a certain threshold from the terminal voltage angle variations at the reference generator. If this happens, the new generator is added to the same cluster of the reference generator, otherwise it becomes the reference generator of a new cluster. The process continues until all generators have been compared with the reference generators of the existing clusters. A successive check is made on the coherency of the generator internal bus voltages, looking for possible reallocations among the clusters. This contribution is based on the simulation of the power system by using a simplified linear model. With the evolution of the computational speed and resources, more contributions emerged to carry out calculations considering non-linear models or data-driven machine learning approaches.

Table I shows the features, the notions of distance, the clustering algorithms adopted, and additional notes on contributions presented in a number of selected publications. The main concepts that emerge from these applications are discussed in the next sections.

IV. FEATURE SELECTION

The feature selection stage has the objective to form a matrix with dimensions (M, H), where M is the number of generators, and H is the number of features. A plurality of features has been used in the literature. The various contributions have used the features directly coming from actual or simulated data or have created customised features by including specific knowledge about slow coherency.

The rotor angle and speed of the generators are the most used features, especially in the model-based approach in which all variables can be assumed to be available. In these approaches, the solution method for computing the variables of interest is based on simulations, assuming full observability by the measurements that could be provided by PMUs [29, 30]. In addition, the time window with which the data are available is long enough to represent the slow oscillations of the generators, and the data resolution inside the time window is appropriate. In general, PMUs cannot record the rotor angles and speeds of the synchronous generators. In some solutions the data on rotor angles and speeds are reconstructed from theoretical calculations based on electrical measurements [31, 32]. In [6] it is assumed that the generator terminal frequency is available from PMUs installed at each generator terminal

node. Higher frequency components are suppressed by using digital low-pass filtering with 0.8 Hz band. The time window is different in the case of normal operation and of the occurrence of transient phenomena due to topology variations or contingencies. In the presence of transient events, pre-event data are separated from post-event data. For the definition of the features to be used in the clustering algorithm, the distances among pairs of generators are calculated by using the cosine distance and the Minkowski distance. Then, the generator distance feature is formed by computing the weighted combination of the squared normalised distance matrices formed by the cosine and Minkowski metrics, taking user-defined weights.

Considering the variations of voltage angle $\Delta\theta_{mt}$, radian frequency $\Delta\omega_{mt}$ and voltage magnitude with respect to equilibrium conditions of the power system [33], for the generator $m = 1,\ldots, M$ at discrete points in time $t = 1,\ldots, T$, the root-mean square coherency criteria [34] can be written, concerning voltage angle and radian frequency, as:

$$\alpha_{ij} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(\left|\Delta\theta_{it} - \Delta\theta_{jt}\right|^2 + \left|\Delta\omega_{it} - \Delta\omega_{jt}\right|^2\right)}. \quad (11)$$

and for voltage magnitudes, as:

$$\beta_{ij} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left|\Delta V_{it} - \Delta V_{jt}\right|^2}. \quad (12)$$

On these bases, it is possible to form the coherency matrices $\mathbf{A} = \{\alpha_{ij}\}$ for angles and $\mathbf{B} = \{\beta_{ij}\}$ for voltages [35].

When the evolution in time of the rotor angles and speeds are available, a classical way to address dynamic problems is to convert the variables with respect to the Centre of Inertia (COI). In this case, the COI is conceptually considered as an infinite bus and is taken as the reference for rotor angle and speed variations. By considering the constant of inertia H_m, the rotor angle δ_{mt} and the rotor speed ω_{mt} for the generator $m = 1,\ldots, M$ at time t, the corresponding quantities reported to the COI are:

$$\delta_{mt}^{(COI)} = \delta_{mt} - \frac{\sum_{m=1}^{M} H_m \delta_{mt}}{\sum_{m=1}^{M} H_m}, \quad (13)$$

$$\omega_{mt}^{(COI)} = \omega_{mt} - \frac{\sum_{m=1}^{M} H_m \omega_{mt}}{\sum_{m=1}^{M} H_m}, \quad (14)$$

Since the evolution in time of the rotor angle and speed has a different starting point after the occurrence of an event, rotor angles and speeds have been further elaborated in [21] by calculating the differences with respect to the first instant of interest ($t = 1$):

$$\hat{\delta}_{mt}^{(COI)} = \delta_{mt}^{(COI)} - \delta_{m1}^{(COI)}, \quad (15)$$

$$\hat{\omega}_{mt}^{(COI)} = \omega_{mt}^{(COI)} - \omega_{m1}^{(COI)}. \quad (16)$$

The evolution of these new variables starts from the same point and allow easier application of Jaccard similarity principles (10), based on the definition of areas calculated on the basis of the trajectories to be compared. This pre-processing step is rarely carried out. It seems promising for comparing the evolution of the time series that have different initial values, e.g., when the distances considered do not include explicit information about correlations. In addition, the initial difference has a negative impact on methods that use Euclidean distances between the time series [36]. However, if the initial difference is caused by far location of the generators, while the type of generator indicates similar behaviour, it has to be verified whether the connection point is relevant for establishing reduced dynamic models.

The centre of inertia frequency deviation (COIFD) is defined in [37] by considering the vector that contains the average frequency deviation for the group of M_g generators that belong to a coherent group g:

$$\Delta f_g^{(COIFD)} = \frac{\sum_{m=1}^{M_g} \Delta f_{gm}}{M_g}. \tag{17}$$

Considering voltages, a relevant quantity is the voltage angle $\theta^{(COA)}$ of the Centre of Angle (COA), calculated by considering the constant of inertia H_m and the voltage angle θ_m of the generator $m = 1,\ldots,M$:

$$\theta^{(COA)} = \frac{\sum_{m=1}^{M} H_m \theta_m}{\sum_{m=1}^{M} H_m}. \tag{18}$$

Once defined, the COA is taken as the angle reference [15]. The difference between the voltage phase angle θ_m of the mth generator and the angle $\theta^{(COA)}$ gives the angle variation at each individual generator $m = 1,\ldots,M$, denoted as:

$$\delta_m = \theta_m - \theta^{(COA)}. \tag{19}$$

The principal component analysis (PCA) has been used in [38] to reduce the number of features from a dynamic model in which the rotor angles and speeds of the generators are taken as input variables. Results have been presented for simulations carried out on a system with 244 generators. The application of PCA to identify the first three principal components of generator speed and node voltage angles was suggested in [39], where some results were provided on a 16-machine 68-bus test system in the form of scattered data plots, without executing a clustering algorithm.

In [40] the extraction of the principal components of the rotor angles is enhanced by using the projection pursuit theory, with which optimal projection directions are identified and the signs of the entries of the projection direction vector are considered to assess the coherency of the generators. In particular, the generators for which the same combination of signs is found in all dominant projection directions are grouped together.

A dynamic coupling-based criterion taken from [41] is used in [30] to define the following similarity function between generators:

$$\xi_{ij} = \left(\frac{1}{H_i} + \frac{1}{H_j} \right) E_i E_j B_{ij} \cos(\delta_{ij0}), \tag{20}$$

where, for the generators i and j, H_i and H_j are the inertia constants, E_i and E_j are the internal node voltages, δ_{ij0} is the relative voltage angle, and B_{ij} is the imaginary part of the bus admittance entry. The similarity values are then included in the matrix $\Xi = \{\xi_{ij}\}$ and are used to form the diagonal matrix $\mathbf{G}_\xi$ with entries $g_{im} = \sum_{m=1}^{M} \xi_{im}$. A normalised Laplacian matrix is then formed as the positive semi-definite matrix

$$\mathbf{L}_\xi = \mathbf{G}_\xi^{-1/2} \left(\Xi - \mathbf{G}_\xi \right) \mathbf{G}_\xi^{-1/2}. \tag{21}$$

and its entries are used to calculate a matrix of Euclidean distances. This matrix is then processed with a method based on Kernel Principal Component Analysis (KPCA), to obtain the features embedded into a Euclidean space, which are used in the clustering algorithm. The embedding strategy used has the advantage to allow the incorporation of non-Euclidean distance measures in the clustering procedure. KPCA is also used in [21] to create a similarity matrix by using the eight outcomes obtained from the calculation of four indices (based on Euclidean distance, Fréchet distance, cosine similarity and Jaccard similarity) on the variables (15) and (16), taking the correlations among these indices as weights. Spectral independent component analysis (ICA) is applied to form a set of features starting from the generator speeds and node voltage angles in the data-driven technique presented in [42], without following with a clustering algorithm. ICA has been used to overcome the drawbacks of PCA, which linear-type decomposition is considered to be insufficient for grouping the coherent generators on the basis of the measured data [42].

Ten similarity indices are calculated in [32], constructed on the basis of characteristic indicators defined for the angle trajectories in the COI coordinate, such as the dissimilarity of amplitude deviation, location deviation, corner deviation, swing direction, and correlation of the trajectories. Then, multicriteria decision-making principles are used to synthesise a single similarity matrix Γ to send to the clustering algorithm. For this purpose, starting from the entries of the qth index normalised similarity matrix Γ_q, the similarity matrix Γ is obtained as

$$\Gamma = \sum_{q=1}^{Q} c_q \Gamma_q, \tag{22}$$

Where c_q is the normalised weight coefficient determined from the combination of two coefficients (the Gini coefficient and the Kendall rank correlation coefficient), such that $\sum_{q=1}^{Q} c_q = 1$. In particular, the Gini

coefficient measures the inequality among the indices, while the Kendall rank correlation coefficient measures the correlations. The exploitation of multiple indicators is also addressed in [18], where a multi-indicator panel data is constructed by considering, for each generator $m = 1,…, M$, the amplitude of the voltage at the generator terminal, the increment of the rotor kinetic energy (with respect to the initial instant t_0), and the voltage phase angle variation $\Delta\delta_{mt}$ at time t with respect to the voltage phase angle $ä_{mt_0}$ at the initial instant t_0:

$$\Delta\delta_{mt} = \delta_{mt} - \delta_{mt_0}. \qquad (23)$$

Another aspect included in [18] is the use of a monotonically decreasing function of time to weight the variables considered, in such a way that the importance of the variable decreases during time.

In some contributions, the Fourier coefficients of the generator speed deviations are determined through the Discrete Fourier transform (DFT) [15]. The rationale of using the DFT is that in normal operating conditions the rotor speed deviations of the generators are null, and after a disturbance there are oscillations that introduce non-zero frequency components (in amplitude and phase) of the rotor speed deviations with respect to the fundamental speed, from which it is possible to recognise the coherent behaviour of some groups of generators. In [15], the features constructed are the DFT coefficients (in amplitude and phase) of the angle variation at each individual generator with respect to the COA. Positive or negative variations of the phase spectrum components determined from the DFT are taken to identify possible generator groups, without explicitly indicating a clustering algorithm. Another use of the evolution of the voltage phase angle variations is to consider the correlations between the real parts and the imaginary parts of the frequencies, found from the DFT, of the velocity variation of the voltage phasors [19]. These correlations are calculated individually, and the cosine distances obtained from these correlations are then summed up in the Euclidean way.

The Discrete Cosine Stockwell Transform is used in [22] to transform the data taken from generator speed and frequency into data in the time-frequency domain. The features are prepared by defining a feature matrix with five indicators: cumulative energy of spectral difference, relative spectral flatness, relative spectral flux, relative spectral coherence, and relative linear correlation. The feature matrix is then subject to min-max normalisation before sending it to the clustering procedure. Wavelet-based coefficients have been used as features in [36], where the dominant frequency components of the generator rotor angles are identified, and the difference in their instantaneous phase is included in a phase difference matrix, which is used as input data for clustering. Features defined in the Taylor-Fourier subspace [43] are used in [44], with the possibility of constructing band-pass filters with different window lengths.

Simulated samples of the generator frequency, extracted with time window 0.25 s and sampling time 100 ms, are used in [45] to form the input data matrix. The data matrix dimension is then reduced by using the Singular Value Decomposition (SVD), and the reduced data are then sent to the clustering process.

The definition of the coherent generator groups could depend on the specific event considered, even though the identification of an overall coherency could be useful. For this purpose, in [34] multiple disturbances and weighted scenarios are considered, calculating the overall coherency matrix as the expected value of the coherency matrices formed with the coefficients α_{ij} from Equation (11) or β_{ij} from Equation (12), determined by applying the same set of disturbances to randomly selected operating conditions. Likewise, in [46] an equivalent rotor speed deviation is determined by considering multiple events and calculating the weighted sum of the generator rotor speed deviations, assuming the active power of the generator as the weighting factor.

Further coherency indices have been defined by using specific data, such as:
- the average angular speed deviation of generators i and j [47]:

$$\chi_{\omega ij} = f_0 \max\{|\Delta\omega_i(t) - \Delta\omega_j(t)|\}; \qquad (24)$$

- the similarity measure between generators i and j [29, 48]:

$$\chi_{Mij} = 1 - \frac{d_M^{(1)}(\boldsymbol{\psi}_i, \boldsymbol{\psi}_j)}{\max_{i,j}\left\{\max\left\{d_M^{(1)}(\boldsymbol{\psi}_i, \boldsymbol{\psi}_j)\right\}\right\}}. \qquad (25)$$

A key point in the identification of the data to be used to form the features is the possible incorporation of information on the network topology. In principle, in the analysis of a network-based problem, a mere comparison among time series, without considering additional information referring to the network connections, could provide groupings in which some generators located into different portions of the network could be included in the same cluster. However, the nature itself of the power system dynamics tends to create groups in which similar types of generators located in a zone close to the large disturbance tend to have a coherent response, while the same type of generators located far from the area tend to have another response. These concepts are discussed in [49] with an algebraic characterisation of coherency, indicating that a generator with small inertia located far from a disturbance could be coherent with a generator with larger inertia but closer to the disturbance. This result also clarifies that using only the electrical distance for determining the coherency is not sufficient. These aspects are also discussed in [50], indicating that for each fault the coherent generators are not always electrically close. Hence, even though the role of the network can remain "hidden", it is significant also when only the time series are analysed. This aspect may be helpful to justify the use of data-driven approaches for

Table 1. Characteristics of Selected Clustering Algorithms for Slow Coherency-based Generator Grouping.

Reference	Features	Distance	Algorithm	Notes
Agrawal & Thukaram (2013) [29]	Generator rotor measurement (angle or speed), simulated	Similarity measure χ_{Mij} between generators i and j	SVC	Initial number of clusters not required. Specific clustering validity ratio for SVC and statistical assessment.
Alsafih & Dunn (2010) [46]	Equivalent rotor speed deviation, simulated	Euclidean	Hierarchical (average linkage)	Multiple events are weighted to define the equivalent rotor speed deviation.
Babaei et al. (2019) [30]	Components of a distance matrix calculated from the simulated generator rotor angles and voltages	Euclidean, based on an embedding strategy to incorporate non-Euclidean distance measures	SVC	Initial number of clusters not required. Total number of iterations reduced. Cluster validity measure defined based on the cluster compactness and separation between clusters.
Barocio et al. (2019) [45]	Generator frequency samples, simulated, reduced by SVD	Euclidean	kmeans	New concept based on clustering slopes. Simple global measure for stability condition.
Chen et al. (2019) [18]	Weighted distance matrix from a multi-indicator panel of simulated data (generator terminal voltage amplitude, increment of the rotor kinetic energy, and voltage phase angle variation with respect to the first instant, weighted with respect to time and indicators)	Weighted distance function with three indicators (horizontal absolute value, rate of change at adjacent time points, and coefficient of variation between two generators)	Hierarchical (Ward linkage criterion)	DBI for determining the number of clusters.
Joo et al. (2001) [52]	Modal responses of the generators from a linearised system model, simulated	Euclidean	kmeans	Calculation of centroids for each cluster.
Kamwa et al. (2007) [34]	Coherency matrix, calculated from simulated phase angle and radian frequency variation signals at the relevant nodes	Euclidean	Fuzzy c-medoids	Medoids are real time series.
Khalil & Iravani (2016) [37]	Simulated frequency deviations, with respect to the system rated frequency, at generator and non-generator nodes	Coherency coefficient (cosine similarity)	Threshold-based DCD	Initial number of clusters not required. Non-generator buses are associated with generator groups, forming electrical areas. Suitable for online applications.
Lin et al. (2018) [21]	Similarity matrix formed through KPCA, based on four indices defined from generator rotor angles and speeds, reported to the COI	Four indices: - Euclidean distance - Fréchet distance - Cosine similarity - Jaccard similarity Correlations among the indices are used as weights in KPCA	Affinity propagation	Initial number of clusters not required.
Lin et al. (2018) [32]	Generator rotor angles and speeds, recorded or simulated	Ten indices, integrated through a decision-making method based on the Gini and Kendall rank correlation coefficients	Spectral	Initial number of clusters not required. Average silhouette index used for determining the best number of clusters.

the creation of the generator groups, in which network data and modelling of generators and their control systems are not needed [51]. Refined versions of these methods will be useful for online applications.

V. CLUSTERING ALGORITHMS

A. Clustering algorithms for coherency-based generator or node grouping

In general, a clustering algorithm is applied starting from the matrix of features with dimensions (M, H), to form an output vector of dimensions $(M, 1)$ in which each entry contains the number of the cluster to which the generator $m = 1,…, M$ has been assigned.

Different clustering algorithms have been used in the literature. Most papers have implemented classical versions of known clustering algorithms, including new algorithms that have become available during time. In general, the authors have tried to get benefits from the advances of

Table 1 (continued). Characteristics of selected clustering algorithms for slow coherency-based generator grouping.

Reference	Features	Distance	Algorithm	Notes
Mei *et al.* (2008) [53]	Simulated generator rotor angle speeds	WSSE, considering the distance from each cluster element to the centroid, with inertia of the generator as the weight factor	Hierarchical (Ward linkage criterion)	
Naglic *et al.* (2020) [6]	Simulated frequency values at generator nodes, used to form a generator distance (weighted combination of the squared normalised distance matrices with cosine and Minkowski metrics, and user-defined weights)	Max-product algorithm	Affinity propagation	Initial number of clusters not required. Hardware in-the-loop implementation.
Rezaeian *et al.* (2018) [19]	Simulated frequency components existing in the angular velocity variation of voltage phasors, in the range of interarea and local oscillation modes	Euclidean sum of the correlations between the real parts and the imaginary parts of the features	Subtractive clustering	Initial number of clusters not required. Coherency assessed in two dimensions (real and imaginary parts).
Singh & Fozdar (2019) [36]	Rotor angles of the generators, used to form a phase difference matrix based on the outcomes of the complex wavelet transform	Euclidean	DBSCAN	Determines the instantaneous phase difference between rotor angles using the complex wavelet transform.
Tyuryukanov *et al.* (2021) [54]	Normalised eigenvectors of a matrix obtained from the second-order electromechanical model with neglected damping, simulated	Normalised graph cuts	Spectral clustering and classical coherency	Initial number of clusters not required. Model improvement for the second-order model of aggregated generators, to reduce the stiffening effect.
Wilfert *et al.* (2001) [38]	Simulated time responses of the rotor motions	Membership of the generators with their clusters, based on normalised principal components	Self-organising maps	
Yadav *et al.* (2019) [22]	Min-max normalised feature matrix with five indicators: cumulative energy of spectral difference, relative spectral flatness, relative spectral flux, relative spectral coherence, and relative linear correlation	Least mean	Mean shift spectral clustering	Useful for data-driven application, independent of changes in system structure and operating conditions.

the state of the art, tying to discover clustering algorithms most suitable for the structure of the problem with data available as time series. In a few cases (e.g., [37] and [54]), the proposed versions include come characteristics of the technical problem in the solution technique.

A classical distinguishing aspect among the clustering algorithms is the need for setting up the number of clusters as an input. In this respect, with reference to the selected contributions shown in Table I:

- The number of clusters has to be preliminarily defined in many classical methods, such as kmeans [55], hierarchical clustering [18, 46, 53], and fuzzy c-medoids [34].
- Methods that do not require the preliminary definition of the number of clusters include non-parametric clustering methods, such as support vector clustering [29, 30], the mean shift spectral clustering [22], the affinity propagation clustering versions [6, 21], the spectral clustering versions [32, 54], the threshold-based clustering [37], the subtractive clustering [19],

and the density-based spatial clustering of applications with noise (DBSCAN) [36].

Another typical distinction among clustering methods is based on their ability to isolate uncommon behaviours as *outliers*. The following cases can be identified, where the first two cases represent extreme situations:

a) Methods that tend to isolate the outliers into small groups and to merge all the other cases into one or a few clusters. These methods have not been exploited for slow coherency studies.

b) Methods that tend to create relatively uniform groups, in which the outliers are aggregated inside one of the groups.

c) Methods in which the outliers are identified and singled out during the clustering process.

The situation b) may occur for clustering methods such as *kmeans*, *k-medoids* and their fuzzy logic-based versions, which tend to create a group that includes the

outliers. The classical kmeans clustering algorithm [56] is used in [52] to determine the groups of generators and calculate the centroids on the basis of an input vector that contains the modal responses of the generators taken from a linearised system model. Furthermore, kmeans is used in [45] in conjunction with the SVD, taking int account the acceleration and deceleration of the synchronous machines. The kmeans method is probabilistic and needs the initialisation of the centroids based on random numbers (or other specific criteria), so that multiple executions should be carried out to obtain a statistically significant analysis of the results. The *k-harmonic means* clustering used in [57] and [58] has been formulated to be non-sensitive to the initialisation of the centroids. The *fuzzy c-means* (FCM) algorithm [59] has been applied in [60]. The interest towards a fuzzy logic-based approach is the possibility of defining the degree of membership with which one generator belongs to different clusters. This is appropriate for interconnected power systems in which there are connections among the nodes that result in stronger or weaker coupling among the generators [34]. However, also FCM depends on some randomness in the initialisation. The subtractive clustering proposed in [19] on the basis of the method formulated in [61] to form the cluster centres can be used a pre-processing step to avoid the dependence on random initialisation. Medoid-based clustering algorithms (e.g., k-medoids and fuzzy c-medoids) are preferable to their centroid-based counterparts (kmeans and FCM), because the centroid is a fictitious output and does not represent any of the actual generators or nodes [34]. Conversely, the medoid gives by definition one of the actual generators or nodes per group, which can be taken as the representative element of the group. For this reason, the *fuzzy c-medoids* (FCMdd) algorithm developed in [62] has been used in [34], where the medoids are initialised by using the sequential PMU placement technique described in [35]. The concept of partitioning around medoids has also been used in [63], where the initial medoids are determined taking into account also information on the nodes to which the generators are connected and on the control systems of the generators.

Other clustering algorithms provide more balanced solutions, in which some outliers can be recognised, while the other generators are located into an appropriate number of clusters. Specific notes are indicated below.

DBSCAN [64] is a density-based method that defines the neighbourhood with a maximum radius and creates groups with a minimum number of points in the neighbourhood. DBSCAN is a deterministic algorithm, in which the number of groups is not fixed a priori. In [36], DBSCAN is applied by considering a phase difference matrix determined from the results of the complex wavelet transform (also considered in [65]). Because of these thresholds, some generators could not be included in any group and are considered as outliers.

Self-organising maps (SOM) [66] have been used in [38] with the use of normalised principal components, and [67] with PCA-based features. The maximum number of clusters is given by the dimension of the map. For a two-dimensional SOM, the output has to be post-processed to identify the groups (e.g., with the kmeans algorithm). A one-dimensional self-organising map can be used if the number of clusters is relatively small (e.g., lower than 6 or 7) to avoid post-processing.

The *affinity propagation* clustering [68] requires in input a similarity matrix, and there is no need to specify the number of groups a priori. It has been used in [21] with a similarity matrix formed through KPCA, composed of four indices, whose correlations have been used as weights in KPCA. It has also been used in [6] to select the post-event measurements and identify groups of slow coherent generators after the occurrence of a disturbance, by using a max-product algorithm. Outliers are included into specific clusters.

In *subtractive* clustering [19], the cluster centres are selected on the basis of the density function δ_{Sm} from Equation (9), where the index m is extended to all the system nodes. The first cluster centre (centroid) is chosen on the basis of the highest density. The effect of the first centroid is then removed from the other nodes through a subtraction process. The procedure is continued to find the next centroids, until the stop criterion based on the ratio between the density at the current step and the density at the first step is satisfied.

Hierarchical clustering [69] is another typical algorithm, which can be used in the agglomerative or divisive version. In the agglomerative case, initially all nodes are taken as independent clusters, a distance metric is used, and the clusters with the lowest distance are progressively merged, until the user-defined number of clusters is reached. In the divisive version there is initially a single cluster, which is partitioned into two clusters by considering the relevant metric, then successive partitions are created until reaching the user-defined number of clusters. All contributions that use the hierarchical clustering have chosen the agglomerative version. The distances are defined by using the linkage criterion, with different solutions. The single linkage (based on the distance between the closest components of pairs of clusters) is used in [70], the average linkage (based on the average distances among the components of pairs of clusters) in [46], the complete linkage (based on the distance between the farthest components of pairs of clusters) in [71], which also considers unstable rotor angle trajectories, and the centroid linkage (based on the distances among the centroids) in [44]. The Ward linkage criterion [72] is used in [53], with the customised option of using the total inertia of the generators during the update of the dissimilarity matrix. In [73] the distance used to decide which pairs of clusters have to be merged is replaced by the dissimilarity coefficient constructed on the basis of ten normalised indicators.

Support vector clustering (SVC) [74] is a non-parametric clustering method that does not need to define the number of clusters a priori. It is able to find suitable clusters of arbitrary shape by determining boundaries of any type. The SVC procedure includes two steps, namely, (i) the mapping of the data points into a higher-dimensional vector space, with the identification of a number of support vectors that depend on a user-defined parameter, and (ii) the formation of the clusters by considering the support vectors to define the cluster boundaries. In [29] the SVC is applied by considering a non-Euclidean distance measure χ_{Mij} between the generators *i* and *j*. The SVC is also applied in [55], where an embedding strategy is developed for making it possible to deal with non-Euclidean distance measures.

The *spectral clustering* algorithm [75] has gained interest in recent years in power system applications [76], also because it is possible to include user-defined similarities [77]. For this purpose, it is applied in [21] to a feature set that includes ten similarity indices. The spectral clustering is based on the definition of nodes in a connected graph, in which the branches are assigned suitable weights. For slow coherency studies, the generators are associated with the nodes, and the dissimilarity among the time series referring to appropriate variables referring to the dynamic behaviour (e.g., rotor angles and speeds) are considered as weights. The average silhouette index is used to assess the most appropriate number of clusters. The mean shift spectral clustering is used in [22] in a data-driven approach in which inertia is not used in the input dataset. Because of this, the method is suitable to deal with the intermittent variations in the renewable energy sources and is not affected by the changes of inertia due to the diffusion of converter-interfaced generators.

The threshold-based *dynamic coherency determination* (DCD) procedure presented in [37] includes two steps. In the first step the generators are partitioned into groups by imposing a minimum threshold on the cosine similarity, and the COIFD is calculated for each group. In the second step the threshold is relaxed to merge pairs of groups, recalculating the COIFD for the new groups as the weighted mean of the COIFD values of the two groups, using the number of generators as the weighting factor. Then, non-generator buses are associated to the groups formed on the basis of the cosine similarity between each bus and the groups formed, assuming full observability at the non-generator nodes. The DCD is extended in [78] by taking into account the impact of wind generators, highlighting how the coherent groups can change because of the presence of wind power plants.

B. New findings for enhancing model-based approaches

Close relations between the notion of slow coherency and the use of normalised graph cuts have been shown in [54]. Starting from the second-order electromechanical

model with neglected damping, the matrix that contains the synchronising torque coefficients between the generators is simplified by neglecting the components associated with the transfer conductances. In this way, the resulting matrix denoted as $\mathbf{K}^B$ has the structure of a negated graph Laplacian matrix. Moreover, the scaled inertias of the generators are considered as weight factors and are included in a diagonal matrix called $\mathbf{M}$. By using these two matrices, it is shown that the solution of minimising the normalised graph cuts with relaxed constraints is equal to the negated sum of the K eigenvalues of $\mathbf{M}^{-1}\mathbf{K}^B$ with smallest magnitude, where K is the number of groups formed, and the groups are determined on the basis of the corresponding eigenvectors. On these bases, a customised spectral clustering algorithm is developed, together with a refined algorithm for aggregating the generators belonging to the same group. The results shown indicate further improvements in the accuracy of model-based approaches based on the calculation of eigenvectors for dynamic model reduction, leading to renewing the interest towards these approaches.

Probabilistic clustering is a challenging perspective that is emerging to deal with the uncertain nature of the power output from wind generators. The identification of common characteristics for wind generators has been handled in [79] by using probabilistic clustering based on SVC, and in [80] by applying a practical four-machine wind turbine clustering method. In the presence of the sources of uncertainty in the power system operation, probabilistic methods associated with clustering techniques may provide more insights on the creation of the generator groups, also with the possible identification of critical generators [81]. The computational burden of the probabilistic power system dynamics analysis is still high, making the computations suitable for offline analysis. In addition, the power system models used have to be very detailed. The probabilistic assessment of multiple operational scenarios may provide more detailed information on the creation of the generator groups, in line with future developments towards probabilistic stability studies [82].

VI. COMPARISONS AMONG CLUSTERING ALGORITHMS FOR GENERATOR GROUPING

Wide comparisons among the solution methods are not generally easy to be carried out. For clustering applications, the network to be used should be large enough and include enough generators to create a significant number of groups. Some IEEE test systems (e.g., the IEEE 68-bus 16-machine power system [83]) provide significant benchmarks, needed in model-based approaches to reproduce the results on the same network, using the same type of model of the synchronous machines, and with the testing carried out on the same contingency. In some cases, the authors have reimplemented the algorithms proposed in the selected reference methods. For data-driven approaches, the situation may become more complex because of data

property issues. Data management by large international organisations opens the possibility of carrying out extended applications to real cases [13].

Considering the selected methods reported in Table 1, different types of comparison have been carried out, as summarised below by considering the progressive year of publication of the related papers.

The earliest example considered here is the successful comparison reported in [34] for the use of FCMdd with respect to the FCM used in [60], also confirmed in [84] for a real system with 160 generators. For the hierarchical clustering described in [53] there is just a qualitative comparison with real data. Another hierarchical clustering algorithm proposed in [46] provides the same results of the calculation of eigenvalues and eigenvectors for the linearised model on a 16-machine system. The authors conclude that their proposed method has the advantage of its applicability to a wide-area measurement system, also considering the effects of the occurrence of different events on the clustering results. A more extended comparison is provided in [29], where the proposed SVC method is compared with kmeans [52], FCM [48], hierarchical (average and Ward linkages) [53, 70]. The comparison is carried out by using the cluster validity indicators Davies–Bouldin index (DBI) [85], silhouette coefficient [86] and Xie-Beni index [87], as well as the minimum average absolute angle difference over all the coherent groups. The results consistently show that the SVC method provides the best results, followed by the hierarchical method with average linkage criterion. The DCD method introduced in [37] exhibits a lower computation burden as compared to other measurements-based methods [42] by using the same system and disturbance.

The spectral clustering method presented in [32] is compared with three methods by using the average silhouette indicator [86], where the compared methods are executed with given parameters. Positive results are reported for the proposed method (with overall better results) and for the methods reported in [40] and [39]. Sensitivity analyses would be needed to perform more detailed verification. In [21] the affinity propagation method is compared with four methods, for three of which ([39, 42], and [40]) and for the proposed method the results are considered acceptable. The proposed method is then considered more appropriate by the authors because of the possibility of using it for online monitoring. The subtractive clustering proposed in [19] is compared with the re-implemented versions of kmeans and FCMdd from [34] on a 16-machine test system. The authors conclude that appropriate setting of the parameters of the subtractive clustering can provide better results.

The discussion in [36] is focused on indicating why DBSCAN is better than kmeans. The weighted clustering of multi-indicator panel data proposed in [18] is compared with other methods such as max-min, kmeans, coherence clustering based on feature extraction of wavelet transform and PCA, and provides better results by considering

the DBI indicator. In [30], the proposed SVC method is compared with FCM [60] and spectral clustering [77] by using the average silhouette indicator [86] and leads to better results, with the advantage that can identify in an automatic way the suitable number of clusters.

More extensive comparisons are carried out in some papers. The clustering algorithm presented in [55] is compared with PCA [39], the determination of Koopman modes [3] and the dynamic mode decomposition (DMD) introduced in [45]. The PCA is unable to provide information on damping, modes and participation factors as given by the Koopman mode and DMD. The proposed clustering method based on SVD and kmeans is able to reduce the dimension of the dataset and requires shorter time windows with respect to Koopman mode and DMD for capturing slow electromechanical modes (about 0.1 Hz). In addition, the proposed method gives information about the slope of the signal. The mean shift spectral clustering from [22] is compared with FCMdd [34], PCA [39], ICA [42], DCD [37], and COI-based affinity propagation [21]. The comparison is conducted with main reference to the behaviour of the methods in intermittency cases due to renewable energy sources. In this case, the information needed for PCA, ICA and COI-based methods are difficult to be obtained or updated in the presence of intermittency, while for FCMdd and DCD the authors indicate the presence of incorrect groupings.

In [6] the results of the proposed affinity propagation method are compared with the reimplemented DCD from [37], with tests carried out for a time interval of 86.6 s with multiple perturbations. The results are presented at different times, showing when the same coherent groups are found from the two methods, and discussing the higher adaptability of the proposed method. In [54] the comparison is conducted more on the modelling side, taking the slow coherency grouping algorithm [88] as the reference. Concerning the generator grouping into areas, some comparisons are carried out with the results shown in [20] and [89], providing some indications concerning the partitions in a small number of groups for a given disturbance.

VII. Concluding Remarks

Slow coherency concepts and related applications implemented through clustering algorithms have been reviewed on the basis of historical contributions and recent developments. The topics addressed in this paper have been dedicated to the bulk power system, excluding distribution systems and microgrids, in which the context of dynamic analysis is evolving with the increasing presence of converter-interfaced generation, storage and participation of the demand-side. Different options have been followed by the authors concerning the choice of the features and the clustering algorithms.

In general, the comparisons shown are still too limited to conclude that any clustering method can be more

appropriate than others. Only extended applications to consolidated benchmarks could provide more hints. In any case, it is not expected to find that one clustering algorithm could be consistently superior to others across all test systems and for many sets of disturbances. The definition of specific benchmarks agreed among the scientific community and the definition of shared test procedures are important directions for carrying out further activity, with the possible development of more refined algorithms. The benchmarks themselves need to be upgraded to follow the current evolution of the power systems, where the integration of generation from renewable energy (mainly wind systems) with fluctuating behaviour and the increasing presence of converter-interfaced generation and storage introduce new challenges to the interpretation of the operating conditions. In this respect, also classical model-based approaches need to be revised to incorporate further details in line the evolving context of power systems operations. Some methodological trends include the development of methods based on constrained clustering, correlation-based clustering, and probabilistic clustering.

An insightful direction for future research is to extend data-driven approaches based on real-time measurements, in which the nature of the connections among the data can be constructed by the solver. The efficiency of these approaches will be satisfactory if sufficiently accurate results will be provided with a computational burden consistent with online applications.

REFERENCES

[1] G. Chicco, "Data Consistency for Data-Driven Smart Energy Assessment," *Frontiers in Big Data*, 4:683682, 2021, doi: 10.3389/fdata.2021.683682.

[2] P. Kundur, *Power System Stability and Control*, McGraw-Hill, New York, NY, 1994.

[3] Y. Susuki and I. Mezic, "Nonlinear Koopman Modes and Coherency Identification of Coupled Swing Dynamics," *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 1894–1904, 2011.

[4] F. Raak, Y. Susuki, and T. Hikihara, "Data-driven partitioning of power networks via Koopman mode analysis," *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 2799–2808, Jul. 2016.

[5] L. Ding, Y. Guo, P. Wall, K. Sun, and V. Terzija, "Identifying the timing of controlled islanding using a controlling UEP based method," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 5913–5922, 2018.

[6] M. Naglic, M. Popov, M.A.M.M. van der Meijden, and V. Terzija, "Synchronized Measurement Technology Supported Online Generator Slow Coherency Identification and Adaptive Tracking," *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 3405–3417, July 2020.

[7] J.M. Undrill and A.E. Turner, "Construction of power system electromechanical equivalents by modal analysis," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-90, no. 5, pp. 2049–2059, 1971.

[8] J.M. Undrill, J.A. Cassaza, E.M. Gulachenski, and L.K. Kirchmayer, "Electromechanical equivalents for use in power system stability studies," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-90, no. 5, pp. 2060–2071, 1971.

[9] A. Chang and M.M. Adibi, "Power System Dynamic Equivalents," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-89, no. 8, pp. 1737-1744, 1970.

[10] S.T.Y. Lee and F.C. Schweppe, "Distance measures and coherency recognition for transient stability equivalents," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-82, pp. 1550-1557, 1973.

[11] D.J. Trudnowski and J.W. Pierre, "Overview of algorithms for estimating swing modes from measured responses," Proc. IEEE Power Energy Soc. Gen. Meeting, Jul. 2009.

[12] A. Vahidnia, G. Ledwich, E. Palmer, and A. Ghosh, "Generator coherency and area detection in large power systems," *IET Gener., Transmiss. Distrib.*, vol. 6, no. 9, pp. 874–883, 2012.

[13] G. Giannuzzi, C. Pisani, and W. Sattinger, "Generator Coherency Analysis in ENTSO-E Continental System: Current Status and Ongoing Developments in Italian and Swiss Case," *IFAC-PapersOnLine*, vol. 49, no. 27, pp. 400–406, 2016.

[14] R. Singh, M. Elizondo, and S. Lu, "A review of dynamic generator reduction methods for transient stability studies," Proc. IEEE Power & Energy Society General Meeting, 2011.

[15] Y. Ota, H. Ukai, K. Nakamura, and H. Fujita, "PMU based midterm stability evaluation of wide-area power system," Proc. 2002 IEEE/PES Transmission and Distribution Conference and Exhibition: Asia Pacific, Yokohama, Japan, 6-10 Oct. 2002, vol. 3, pp. 1676–1680.

[16] T.W. Liao, "Clustering of time series data—A survey," *Pattern Recognition*, vol. 38, pp. 1857–1874, 2005.

[17] T. Heiter and H. Manila, "**Computing Discrete Fréchet Distance,**" Technical Report CD-TR 94/64, Christian Doppler Laboratory for Expert Systems, Technical University Vienna, Austria, 1994.

[18] Y. Chen, Z. Zhang, X.Song, J. Liu, M. Hou, G. Li, W. Xu, and J. Ma, "Coherent Clustering Method Based on Weighted Clustering of Multi-Indicator Panel Data," *IEEE Access*, vol. 7, pp. 43462-43472, 2019.

[19] M.H. Rezaeian, S. Esmaeili, and R. Fadaeinedjad, "Generator coherency and network partitioning for dynamic equivalencing using subtractive clustering algorithm," *IEEE Systems Journal*, vol. 12, no. 4, pp. 3085–3095, Dec. 2018.

[20] J.H. Chow, "Slow Coherency and Aggregation," in J. Chow (ed.) *Power System Coherency and Model Reduction*, Power Electronics and Power Systems, vol. 94, Springer-Verlag, New York, NY, 2013.

[21] Z. Lin, F. Wen, Y. Ding, Y. Xue, S. Liu, Y. Zhao, and S. Yi, "WAMS-based coherency detection for situational awareness in power systems with renewables," *IEEE*

Transactions on Power Systems, vol. 33, no. 5, pp. 5410–5426, Sep. 2018.

[22] *R. Yadav, A. K. Pradhan, and I. Kamwa, "A Spectrum Similarity Approach for Identifying Coherency Change Patterns in Power System Due to Variability in Renewable Generation," IEEE Transactions on Power Systems, vol. 34, no. 5, pp. 3769–3779, 2019.*

[23] M. Klein, G. J. Rogers, and P. Kundur, "A fundamental study of interarea oscillations in power systems," *IEEE Transactions on Power Systems*, vol. 6, no. 3, pp. 914–921, Aug. 1991.

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

[25] R. Podmore, "Identification of coherent generators for dynamic equivalents," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-97, no. 4, pp. 1344–1354, Jul. 1978.

[26] A.J. Germond and R. Podmore, "Dynamic Aggregation of Generating Unit Models," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-97, no. 4, pp. 1060-1069, July 1978.

[27] J.H. Chow (ed.), *Power System Coherency and Model Reduction*, Power Electronics and Power Systems, vol. 94, Springer-Verlag, New York, NY, 2013.

[28] J.B. Ward, "Equivalent circuits for power flow studies," *AIEE Trans. Power Appl. Syst.*, vol. 68, no. 1, pp. 373–382, 1949.

[29] R. Agrawal and D. Thukaram, "Support vector clustering-based direct coherency identification of generators in a multi-machine power system," *IET Gener., Transm., Distrib.*, vol. 7, no. 12, p. 1357, 1366, Dec. 2013.

[30] M. Babaei, S.M. Muyeen, and S. Islam, "Identification of Coherent Generators by Support Vector Clustering with an Embedding Strategy," *IEEE Access*, vol. 7, pp. 105420-105431, 2019.

[31] V. Venkatasubramanian and R.G. Kavasseri, "Direct computation of generator internal dynamic states from terminal measurements," Proc. 37th Annu. Hawaii Int. Conf. System Sciences, January 2004.

[32] Z. Lin, F. Wen, Y. Ding, and Y. Xue, "Data-driven coherency identification for generators based on spectral clustering", *IEEE Transactions on Industrial Informatics*, vol. 14, no. 3, pp. 1275–1285, Mar. 2018.

[33] S. Sastry and P. Varaiya, "Coherency for interconnected power systems," *IEEE Transactions on Automatic Control*, vol. AC-26, no. 1, pp. 218–226, 1981.

[34] I. Kamwa, A. K. Pradhan, and G. Joos, "Automatic segmentation of large power systems into fuzzy coherent areas for dynamic vulnerability Assessment," *IEEE Transactions on Power Systems*, vol. 22, no. 4, pp. 1974–1985, Nov. 2007.

[35] I. Kamwa and R. Grondin, "PMU configuration for system dynamic performance measurement in large, multiarea power systems," *IEEE Transactions on Power Systems*, vol. 17, no. 2, pp. 385–394, May 2002.

[36] A.K. Singh and M. Fozdar, "DBSCAN-based coherency identification of generators using complex wavelet transform," *The Journal of Engineering*, vol. 2019, no. 18, pp. 4780–4784, 2019.

[37] A.M. Khalil and R. Iravani, "A dynamic coherency identification method based on frequency deviation signals," *IEEE Transactions on Power Systems*, vol. 31, no. 3, pp. 1779–1787, May 2016.

[38] H.H. Wilfert, K. Voigtländer, I. Erlich, "Dynamic coherency identification of generators using self-organising feature maps," *Control Engineering Practice*, vol. 9, pp. 769–775, 2001.

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

[40] T. Jiang, H. Jia, H. Yuan, N. Zhou, and F. Li, "Projection pursuit: A general methodology of wide-area coherency detection in bulk power grid," *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 2776–2786, Jul. 2016.

[41] S.S. Lamba and R. Nath, "Coherency identification by the method of weak coupling," *International Journal of Electrical Power & Energy Systems*, vol. 7, no. 4, pp. 233-242, 1985.

[42] M.A.M. Ariff and B.C. Pal, "Coherency identification in interconnected power system—An independent component analysis approach," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1747–1755, May 2013.

[43] J.A. de la O Serna, J.M. Ramirez, A. Zamora Mendez, and M.R.A. Paternina, "Identification of Electromechanical Modes Based on the Digital Taylor-Fourier Transform," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 206-215, 2016.

[44] 44 M.R. Arrieta Paternina, A. Zamora-Mendez, J. Ortiz-Bejar, J.H. Chow, and J.M. Ramirez, "Identification of coherent trajectories by modal characteristics and hierarchical agglomerative clustering," *Electric Power Systems Research*, vol. 158, pp. 170–183, 2018.

[45] E. Barocio, B.C. Pal, N.F. Thornhill, and A.R. Messina, "A Dynamic Mode Decomposition Framework for Global Power System Oscillation Analysis," *IEEE Transactions on Power Systems*, vol. 30, no. 6, pp. 2902–2912, 2015.

[46] H.A. Alsafih and R. Dunn, "Determination of coherent clusters in a multi-machine power system based on wide-area signal measurements," IEEE Power and Energy Society General Meeting, Minneapolis, MN, USA 2010.

[47] D.E.T. Hiyama, "Identification of coherent generators using frequency response," *IEE Proceedings C Generation, Transmission and Distribution*, vol. 128, no. 5, pp. 262 – 268, 1981.

[48] S.C. Wang and P.H. Huang, "Direct coherency identification of synchronous generators in Taiwan

power system based on fuzzy c-means clustering," *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.*, vol. E90A, no. 10, pp. 2223–2231, 2007.

[49] F. Wu and N. Narasimhamurthi, "Coherency identification for power system dynamic equivalents," *IEEE Transactions on Circuits and Systems*, vol. 30, no. 3, pp. 140–147, March 1983.

[50] E.J.S. Pires de Souza, "Identification of coherent generators considering the electrical proximity for drastic dynamic equivalents," *Electric Power Systems Research*, vol. 78, no. 7, pp. 1169-1174, 2008.

[51] M.R. Aghamohammadi and S.M. Tabandeh, "A new approach for online coherency identification in power systems based on correlation characteristics of generators rotor oscillations," *International Journal of Electrical Power & Energy Systems*, vol. 83, pp. 470-484, 2016.

[52] S.K. Joo, C.C. Liu, and J.W. Choe, "Enhancement of coherency identification techniques for power system dynamic equivalents," IEEE Power Eng. Soc. Summer Meet., Vancouver, BC, Canada, 2001, 3, pp. 1811–1816.

[53] K. Mei, S. M. Rovnyak, and C.M. Ong, «Clustering-Based Dynamic Event Location Using Wide-Area Phasor Measurements," *IEEE Transactions on Power Systems*, vol. 23, no. 2, pp. 673–679, May 2008.

[54] I. Tyuryukanov, M. Popov, M.A.M.M. Van Der Meijden, and V. Terzija, "Slow Coherency Identification and Power System Dynamic Model Reduction by using Orthogonal Structure of Electromechanical Eigenvectors," *IEEE Transactions on Power Systems*, vol. 36, no. 2, pp. 1482–1492, 2021.

[55] E. Barocio, P. Korba, W. Sattinger, and F.R. Segundo Sevilla, "Online coherency identification and stability condition for large interconnected power systems using an unsupervised data mining technique," *IET Generation, Transmission & Distribution*, vol. 13, no. 15, pp. 3323-3333, 2019.

[56] J.T. Tou and R.C. Gonzalez, *Pattern Recognition Principles*, Addison-Wesley, 1974.

[57] K. Tang and G. K. Venayagamoorthy, "Online coherency analysis of synchronous generators in a power system," Proc. 2014 IEEE PES Innovative Smart Grid Technologies Conf. (ISGT), Feb. 19–22, 2014.

[58] Y. Wei, P. Arunagirinathan, A. Arzani, and G.K. Venayagamoorthy, "Situational awareness of coherency behavior of synchronous generators in a power system with utility-scale photovoltaics," *Electric Power Systems Research*, vol. 172, pp. 38–49, 2019.

[59] J.C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York, 1981.

[60] S.C. Wang and P.H. Huang, "Fuzzy c-means clustering for power system coherency," Proc. Int. Conf. Syst., Man, Cybern., Waikoloa, HI, Oct. 10–12, 2005, vol. 3, pp. 2850–2855.

[61] S.L. Chiu, "Fuzzy model identification based on cluster estimation," *J. Intell. Fuzzy Syst.*, vol. 2, pp. 267–278, 1994.

[62] R. Krishnapuram, A. Joshi, and O. Nasraoui, "Low-complexity fuzzy relational clustering algorithms for web mining," *IEEE Transactions on Fuzzy Systems*, vol. 9, no. 4, pp. 595–607, Aug. 2001.

[63] G. Pyo, J. Park, and S. Moon, "A new method for dynamic reduction of power system using PAM algorithm," Proc. IEEE PES General Meeting, 2010, pp. 1–7.

[64] M. Ester, H.P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," Proc. of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96). AAAI Press, pp. 226–231, 1996.

[65] S. Avdaković, E. Becirović, A. Nuhanović, and M. Kusljugić, "Generator coherency using the wavelet phase difference approach," *IEEE Transactions on Power Systems*, vol. 29, no. 1, pp. 271–278, Jan. 2014.

[66] T. Kohonen, *Self-organisation and associative memory*, Springer Verlag, Heidelberg, Germany, 1989.

[67] B.P. Padhy, S.C. Srivastava, and N.K. Verma, "A coherency-based approach for signal selection for wide area stabilizing control in power systems," *IEEE Systems Journal*, vol. 7, no. 4, pp. 807–816, 2013.

[68] B.J. Frey and D. Dueck, "Clustering by passing messages between data points," *Science*, vol. 315, no. 5814, pp. 972–976, 2007.

[69] M.R. Anderberg, *Cluster Analysis for Applications*, Academic Press, New York, 1973.

[70] *C. Juarez, A.R. Messina, R. Castellanos, and A.R. Espinosa-Pérez, "Characterization of multi-machine system behavior using a hierarchical trajectory cluster analysis," IEEE Transactions on Power Systems, vol. 26, no. 3, pp. 972–981, 2011.*

[71] T. Guo and J.V. Milanović, "Online identification of power system dynamic signature using PMU measurements and data mining," *IEEE Transactions on Power Systems*, vol. 31, no. 3, pp. 1760–1768, 2016.

[72] J.H. Ward, "Hierarchical grouping to optimize an objective function," *Journal of the American Statistical Association*, vol. 58, no. 301, pp. 235–244, 1963.

[73] Z. Lin, F. Wen, Y. Ding, and Y. Xue, "Wide-area coherency identification of generators in interconnected power systems with renewables," *IET Gener., Transmiss. Distrib.*, vol. 11, no. 18, pp. 4444–4455, 2017.

[74] A. Ben-Hur, D. Horn, H.T. Siegelmann, and V. Vapnik, "Support vector clustering," *J. Mach. Learning Res.*, vol. 2, pp. 125–137, 2001.

[75] U. von Luxburg, "A tutorial on spectral clustering," *Statist. Comput.*, vol. 17, no. 4, pp. 395–416, 2007.

[76] R.J. Sánchez-García, M. Fennelly, S. Norris, N. Wright, G. Niblo, J. Brodzki, and J.W. Bialek, *"*Hierarchical Spectral Clustering of Power Grids," *IEEE Transactions on Power Systems*, vol. 29, no. 5, pp. 2229–2237, 2014.

[77] A. Thakallapelli, S. J. Hossain, and S. Kamalasadan, "Coherency and online signal selection based wide area control of wind integrated power grid," *IEEE*

Transactions on Industry Applications, vol. 54, no. 4, pp. 3712–3722, 2018.

[78] A.M. Khalil and R. Iravani, "Power system coherency identification under high depth of penetration of wind power," *IEEE Transactions on Power Systems*, vol. 33, no. 5, pp. 5401–5409, Sep. 2018.

[79] M. Ali, I. Ilie, J.V. Milanović, and G. Chicco, "Wind Farm Model Aggregation Using Probabilistic Clustering," *IEEE Transactions on Power Systems*, vol. 28, no. 1, pp. 309-316, 2013.

[80] P. Wang, Z. Zhang, Q. Huang and W. Lee, "Wind Farm Dynamic Equivalent Modeling Method for Power System Probabilistic Stability Assessment," *IEEE Transactions on Industry Applications*, vol. 56, no. 3, pp. 2273–2280, 2020.

[81] P.N. Papadopoulos, T. Guo and J.V. Milanović, "Probabilistic framework for online identification of dynamic behavior of power systems with renewable generation," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 45–54, Jan. 2018.

[82] J.V. Milanović, "Probabilistic stability analysis: the way forward for stability analysis of sustainable power systems," *Phil. Trans. R. Soc. A*, vl. 375, art. 20160296, 2017.

[83] A.K. Singh and B.C. Pal, Report on the 68-bus, 16-machine, 5-area system. Technical Report. IEEE Power and Energy Society Task Force on Benchmark Systems for Stability Controls, 2013. http://eioc.pnnl.gov/benchmark/ieeess/NETS68/New_England_New_York_68_Bus_System_study_report.pdf

[84] I. Kamwa, A. K. Pradhan, G. Joos, and S.R. Samantaray, "Fuzzy partitioning of a real power system for dynamic vulnerability assessment," *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1356–1365, Aug. 2009.

[85] D.L. Davies and D.W. Bouldin, "A cluster separation measure," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAM-1, no. 2, pp. 224–227, 1979.

[86] P.J. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," *J. Comput. Appl. Math.*, vol. 20, no. 1, pp. 53–65, 1987.

[87] X.L. Xie and G. Beni, "A validity measure for fuzzy clustering," *IEEE Trans. Pattern Anal. Mach. Intell.*, 1991, 13, (8), pp. 841–847.

[88] J.H. Chow, *Time-Scale Modeling of Dynamic Networks with Applications to Power Systems*, Lecture Notes in Control and Information Sciences, vol. 46. Berlin, Germany: Springer-Verlag, 1982.

[89] J.H. Chow, R. Galarza, P. Accari, and W. W. Price, "Inertial and slow coherency aggregation algorithms for power system dynamic model reduction," *IEEE Transactions on Power Systems*, vol. 10, no. 2, pp. 680–685, May 1995.

Gianfranco Chicco received the Ph.D. degree in Electrotechnics Engineering from Politecnico di Torino (POLITO), Turin, Italy. He is a Full Professor of Electrical Energy Systems at POLITO. His research interests include power system and distribution system analysis and optimization, multi-energy system optimization and planning, electrical load management, data analytics, and artificial intelligence applications to energy systems.