



Non-heuristic Identification of the Optimal Number of Oscillation Patterns in Uncertain Power Systems

DOI:
[10.1109/ISGT-Europe47291.2020.9248958](https://doi.org/10.1109/ISGT-Europe47291.2020.9248958)

Document Version
Accepted author manuscript

[Link to publication record in Manchester Research Explorer](#)

Citation for published version (APA):
Morales Alvarado, J., & Milanovic, J. V. (2020). Non-heuristic Identification of the Optimal Number of Oscillation Patterns in Uncertain Power Systems. In *2020 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)* (pp. 424-428). Article 9248958 (IEEE PES Innovative Smart Grid Technologies Conference Europe; Vol. 2020-October). <https://doi.org/10.1109/ISGT-Europe47291.2020.9248958>

Published in:
2020 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)

Citing this paper
Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights
Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy
If you believe that this document breaches copyright please refer to the University of Manchester's Takedown Procedures [<http://man.ac.uk/04Y6Bo>] or contact openresearch@manchester.ac.uk providing relevant details, so we can investigate your claim.



Non-heuristic Identification of the Optimal Number of Oscillation Patterns in Uncertain Power Systems

Juan D. Morales, *Student Member IEEE*
 Department of Electrical and Electronic Engineering
 The University of Manchester
 Manchester, UK
 juan.morales@manchester.ac.uk

Jovica V. Milanović, *Fellow IEEE*
 Department of Electrical and Electronic Engineering
 The University of Manchester
 Manchester, UK
 milanovic@manchester.ac.uk

Abstract—This paper introduces a non-heuristic methodology using Hierarchical Clustering (HC) for identifying the optimal number of oscillation patterns in uncertain power systems due to transient stability phenomena, for subsequent control and stabilization. The HC application parameters are statistically defined after comparing the quality of different clustering partitions using appropriate indicators, so that the most suitable patterns can be automatically obtained. The methodology is illustrated on a modified IEEE 68 bus NETS-NYPS test network with renewable energy sources and realistically modelled uncertainties, resulting in the identification of a reduced (optimal) number of unstable generator patterns which are fully representative of the overall multi-machine transient behavior of the system.

Keywords—*hierarchical clustering, renewables, statistical analysis, transient stability, uncertainties.*

I. INTRODUCTION

Increasing uncertain operating conditions and declining inertia due to a growing integration of Renewable Energy Sources (RES), under already constrained market based operating scenarios, are typical features of modern power systems in most parts around the world, likely driving them to operate closer to their stability limits. Probabilistic approaches seem to be the best option to study these scenarios [1, 2]. These studies generally involve the simulation of large number of operating conditions, facilitating the application of data mining techniques for training or classification purposes so that some specific feature of the system can be assessed or predicted, e.g., clustering algorithms that use rotor angle or speed for coherency identification are reported in [3-5].

In [6-8], Hierarchical Clustering (HC) is used for grouping probabilistically obtained transient rotor angle oscillation patterns (dynamic signature), and it is demonstrated that HC, as a data mining technique, can be effectively used for this purpose. However, the criteria for choosing the algorithm application parameters and how they influence the grouping results were not thoroughly discussed. Results in these studies showed that the number of oscillation patterns obtained for a system is usually large making it difficult to identify critical groups in a straightforward way that could guide actions regarding stabilization control. This is especially important when dealing with multi-machine transient instability, which may involve a large number of unstable generators, i.e., a high blackout risk.

This paper proposes a methodology to find statistically the clustering parameters for the optimal identification of the most suitable multi-machine transient unstable generator

patterns of a power system using HC. The procedure is illustrated on a modified IEEE 68 bus NETS-NYPS test network incorporating RES and realistically modelled uncertainties.

II. METHODOLOGY

A. Generation of the System Dynamic Response Database

The MC approach developed in [1] is used to generate a representative set of realistic dynamic responses for power systems with the inclusion of RES. The test network (see Fig. 1) used in this paper is a modified IEEE 68 bus - 16 machine network with the inclusion of 10 RES plants. All generators are represented by full 6th order models including automatic voltage regulator and governor controls. Only generator 9 is equipped with a power system stabilizer. Loads are represented using a constant impedance load model. The RES plant mix, per bus, comprises 66.7% of Doubly Fed Induction Generators (DFIGs), representing wind plants, and 33.3% of Full Converter Connected (FCC) generators, of which 30% are wind and 70% are PV generators. DFIGs and FCCs are represented by generic Type 3 and Type 4 models, respectively, suitable for large scale stability studies, and have Fault Ride Through (FRT) capabilities. Complete data of the system components and models used can be found in [1, 2].

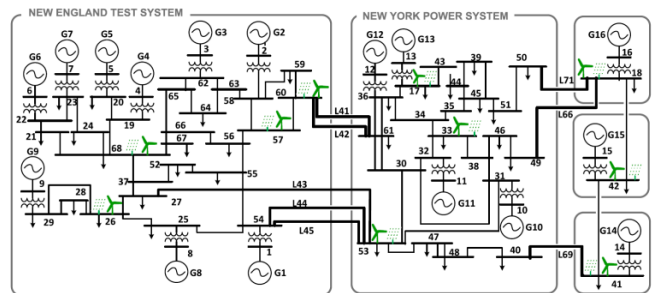


Fig. 1. Modified IEEE 68 bus test network with the inclusion of RES.

The sampling of uncertainties is done independently for each relevant random variable of the system, and according to appropriate probability distributions describing their particular behavior. These include loads, hour of the day, wind/PV production, fault duration and fault location. A typical 24 hour loading curve is used. The hour of the day is sampled randomly using a uniform distribution from the daily loading curve. The uncertainty of loading at each hour is modelled by a normal distribution. Beta distribution is used for modelling the PV generation uncertainty, while a Weibull distribution is used for the wind speed uncertainty. The pre-fault operating conditions are obtained by solving an Optimal Power Flow (OPF), whose dispatch results

This work was supported by the PhD Scholarship Award provided by the Department of Electrical and Electronic Engineering of The University of Manchester.

determine the amount of disconnection of conventional generation and corresponding inertia reduction due to RES penetration, considering here a fixed 15% of spinning reserve for all cases. Regarding system disturbances, only three phase self-clearing faults are considered. The locations of faults along each line and the line itself are sampled with a uniform distribution. A normal distribution with a mean of 14 cycles and a standard deviation of 6.67% is used to model the fault duration. This is longer than typical (approx. 100 ms), in order to generate an appropriate number of stable/unstable cases for testing the methodology [7]. Complete information about the sampling of uncertainties and OPF constraints can be found in [1, 2].

The required number of MC simulations defined in [1] to achieve satisfactory accuracy was 6,000. In this paper, as in [2, 9], the number of simulations is increased to 10,000 to obtain more unstable cases for credible statistical analysis. The case study used for testing has a nominal RES penetration level of 20%. Each MC simulation involves a 5 s time domain simulation of the system with the fault applied at 1 s, since the focus is on first swing instability (transient instability). Simulations are performed using the DigSILENT – Power Factory software. Further, detailed information on data and parameters used for the MC approach can be found in [1, 2]. Rotor angle responses of each generator are used to identify the oscillation patterns of the system using the clustering method proposed in Section II-B.

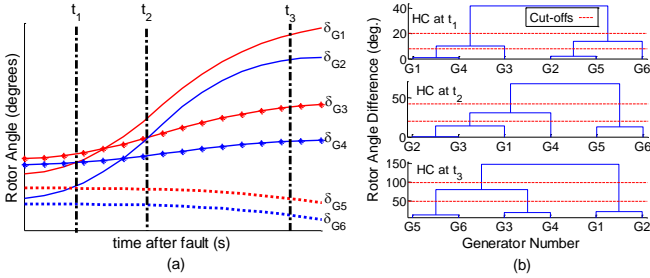


Fig. 2. (a) Post-fault instants for HC. (b) Corresponding dendrograms.

B. Hierarchical Clustering for Multi-Machine Instability

The HC method as proposed in [6] is used as the clustering algorithm in this study. It clusters the angle data by the agglomerative (bottom-up) strategy, first finding the similarity between a pair of objects (rotor angles) and then linking sets of objects for merging in upper levels into a hierarchical tree (or dendrogram), by means of a predefined linkage criterion. Euclidean distance is used as the similarity measure and complete linkage is chosen as the linkage criterion (distance between clusters is equal to the distance between the objects in each cluster that are the farthest away from each other). Different instants for applying HC within a system post-fault oscillation, e.g. t_1 , t_2 , and t_3 in Fig. 2(a), will yield different dendrograms, see Fig. 2(b). Moreover, for each dendrogram obtained at a particular clustering instant, also different cluster arrangements or Grouping Patterns (GPs) can be formed depending on the value of the threshold for cutting-off the hierarchical tree, represented by the red dashed lines in Fig. 2(b). In this paper, the estimation of the optimal instant of clustering and cut-off threshold values are done using statistics which measure the appropriateness and quality of cluster partitions, over the set of representative rotor angle responses obtained as described in Section II-A. This ensures that the critical oscillation patterns are identified.

The clustering of generators should be done at an instant close to the loss of synchronism. Hence, it is better to define this instant in terms of an angle threshold for transient instability identification rather than an actual point in time after a disturbance. In [9], a thorough analysis on the subject was done using thresholds based on the Maximum Angle Difference Between Two Generators (MADBTG) and on the angular distance from the Center Of Inertia (COI) reference frames. Both will be used in this study to measure system separation for evaluating the clustering results. For the MADBTG, instability is reached if condition (1) is fulfilled.

$$\delta_T < \delta_{MAX} = \max(|\delta_i - \delta_j|) \quad (1)$$

where δ_T is the instability threshold and δ_{MAX} is the maximum angle difference (degrees) between any two rotor angles δ_i and δ_j at the same point in time after the occurrence of a large disturbance. Similarly, for the COI reference frame based thresholds, instability is reached if condition (2) is fulfilled.

$$\delta_T < |\delta_i - \delta_{COI}| \quad (2)$$

where δ_{COI} is the angular position of the COI, defined in (3).

$$\delta_{COI} = \frac{\sum_{i=1}^{N_G} H_i \times \delta_i}{\sum_{i=1}^n H_i} \quad (3)$$

where H_i is the inertia constant of generator i , and N_G is the total number of generators. For the MADBTG and COI frames based thresholds, the value of δ_T in (1) and (2) for the clustering assessment exploration is defined for the ranges between 100° - 360° and 100° - 180° , respectively [9]. With respect to the cut-off thresholds (δ_C), values up to 100° are explored, so an overlap with the δ_T values (which start at 100°) is avoided. Both parameters are evaluated in steps of 10° .

C. Statistical Assessment of Clustering Partitions

The choosing of the cut-off threshold in HC algorithms can be considered a special case of the general problem of finding the most appropriate number of clusters for any clustering method. A thorough comparison of 30 different procedures for choosing cluster partitions and/or the optimal number of clusters is performed in [10], with some of the top performing ones further studied in [11, 12]. However, there is a lack of consensus on which method to use, and it is recommended not to depend entirely on a single way for finding the optimal number of clusters [11]. It can be safely concluded that subjective criteria based on expert judgement still have a substantial influence on a dataset clustering results. The common denominator for all methods is that they always involve some degree of heuristics in their definition. Nevertheless, there is a consensus in literature, e.g. [11-14], that most methods rely on the behavior of the change of the defined within-cluster dissimilarity W_K as a function of the K number of clusters, since the value of W_K will decrease as the number of clusters K increases because having more clusters allows to group elements that are more similar to each other. However, above the optimal number of clusters K^* , any new partition will split a “true” cluster containing elements already similar enough, and the effect on the decrease of W_K will become significantly less than before reaching K^* , and more marginal as K is further increased. The criterion for finding K^* is then to identify this turning point or “elbow” in the plot of W_K as a function of K . This is still an approach involving heuristics since deciding how substantial the change defining the elbow should be is

unclear in some cases and subjected to user's judgement and prior expectation. Two proposed statistics to find the optimal number of clusters are the Average Silhouette Width (ASW) [15] and the Gap Statistic (GS) [12]. Both are chosen to assess statistically the rotor angles clustering arrangements as proposed in this paper.

The ASW is the average of the silhouette coefficient $s(\delta_i)$ over all elements in a dataset and is defined in (4).

$$s(\delta_i) = \frac{b(\delta_i) - a(\delta_i)}{\max\{a(\delta_i), b(\delta_i)\}} \quad (4)$$

with values varying from -1 to 1. For each element δ_i , the value $a(\delta_i)$ is the average distance to all other elements in its cluster, which reflects the cluster compactness, while $b(\delta_i)$ is its minimum average distance to all clusters to which it does not belong, which reflects the degree of cluster separation. Thus, large $s(\delta_i)$ values indicate a well clustered element. The silhouette coefficient is a metric for assessing the quality of clusters. Nevertheless, choosing a partition that maximizes the ASW is a formal criterion for selecting the number of clusters [11]. In [15], it is stated that a reasonable classification will yield an ASW above 0.5, while values below 0.2 should be interpreted as a lack of substantial cluster structure. For some clustering algorithms, having fewer elements per cluster could achieve better results [13], considering that the granularity of any dataset would be maximized having one element per cluster. This, however, does not provide any useful partition of the data and might unrealistically increase the value of the ASW, hence it is undesirable. The ASW metric properties will be useful when comparing the patterns obtained from the different clustering parameters as proposed in this study.

The GS formalizes the heuristic elbow criterion for estimating the optimal number of clusters to statistically locate when the elbow occurs. The gap curve $G(K)$ is defined in (5).

$$G(K) = E\{\log(W_K^*)\} - \log(W_K) \quad (5)$$

where W_K^* is the within cluster dissimilarity of clustering data from an appropriate null distribution and E the mathematical expectation. For the general case of a multivariate dataset of elements with more than one feature (or dimension), a suitable null reference distribution is generated using MC sampling from data uniformly distributed over a box containing the original dataset elements over each feature. It estimates the optimal number of clusters K^* to be where the gap curve is the largest. The formal definition for estimating K^* is given in (6).

$$K^* = \underset{K}{\operatorname{argmin}}\{K | G(K) \geq G(K+1) - s'_{K+1}\} \quad (6)$$

where s'_K is an error term due to the MC sampling [14]. Since based on the elbow criterion, some parameters for testing also rely on user's judgement, i.e., the number of MC samples, the value of s'_K , the choice of the reference distribution for the multivariate case, and the choice for W_K (squared error, squared distances around cluster means, etc.). On the other hand, for the present application, where each element (generator) only has its rotor angle as a feature for clustering (one dimension), *Theorem 1* proposed in [12] for the univariate case can be applied, which states that for data in one dimension, the uniform distribution is the most likely to produce spurious clusters by means of the GS test. Hence, a single standard uniform distribution over the data can be used directly for generating the null reference distribution

without any MC sampling, also removing the error term in (6) and thus further reducing heuristics. Another benefit of the GS, is that if the produced gap curve exhibits non-monotone behavior, a smaller well-separated cluster arrangement within a larger well-separated one, i.e., the optimal obtained from (6), may exist, implying that more than one natural partition trend is present in the dataset. In such cases, examination of the entire gap curve is suggested, rather than only finding the position of its maximum. It has been observed that the GS has less heuristics involved in estimating the number of clusters, making it a good choice for this purpose. In this study, the GS test (considering within cluster sum of squares as W_K and $K^* > 1$) is used as a filter to allow comparison using the ASWs values only between valid results according to the test.

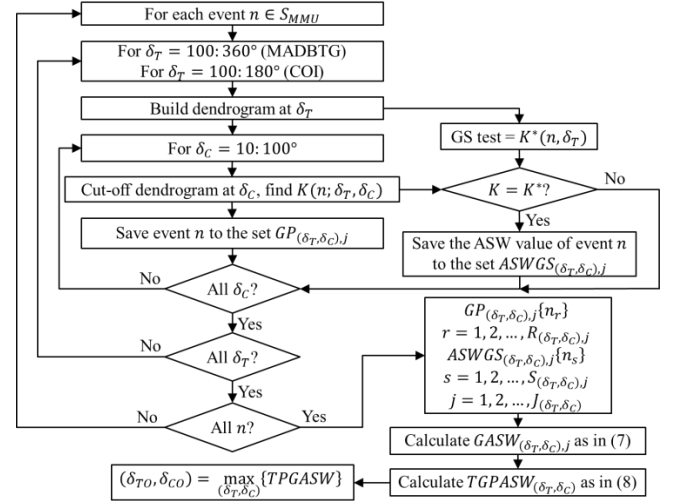


Fig. 3. Flowchart of the proposed statistical clustering assessment.

The procedure for the assessment is shown in Fig. 3. A search over all the predefined values for the instants for clustering δ_T and cut-off thresholds δ_C is performed for all the events from the set of multi-machine unstable events S_{MMU} . Each pair (δ_T, δ_C) will produce a number of $J(\delta_T, \delta_C)$ Grouping Patterns (GPs), each of which consists of $R(\delta_T, \delta_C, j)$ events, which are grouped into the sets $GP(\delta_T, \delta_C, j)$, with $j = 1, 2, \dots, J(\delta_T, \delta_C)$. In parallel, for each event n and instant of clustering δ_T , a GS test is performed, yielding an estimation of the optimal number of clusters $K^*(n; \delta_T)$. This result is compared to the number of clusters $K(n; \delta_T, \delta_C)$ obtained directly for each of the predefined cut-off threshold values δ_C , then only if $K = K^*$, the ASW of the event's GP is calculated and stored in the $ASWGS_{(\delta_T, \delta_C), j}$ set, each of which finally ends having a total of $S_{(\delta_T, \delta_C), j}$ elements. With this information, a quality measure can be assigned to all GPs according to (7).

$$GPASW_{(\delta_T, \delta_C), j} = \frac{\sum_{s=1}^{S_{(\delta_T, \delta_C), j}} ASWGS_{(\delta_T, \delta_C), j}\{n_s\}}{R_{(\delta_T, \delta_C), j}} \quad (7)$$

where n_s is the ASW value of each element (or event) of the set $ASWGS_{(\delta_T, \delta_C), j}$, with $j = 1, 2, \dots, J(\delta_T, \delta_C)$. The calculated index $GPASW_{(\delta_T, \delta_C), j}$ in (7) is an average over all $R_{(\delta_T, \delta_C), j}$ elements (events) of the $GP_{(\delta_T, \delta_C), j}$ sets instead of over the total $S_{(\delta_T, \delta_C), j}$ elements (events) of the $ASWGS_{(\delta_T, \delta_C), j}$ sets. It considers only clustered events whose GPs were correctly estimated with the GS test, and quantifies the degree to which that identified GP represents a true oscillation pattern

for the system. In order to compare the HC results from all pairs of clustering parameters (δ_T, δ_C) , a total quality measure is defined in (8).

$$TGPASW_{(\delta_T, \delta_C)} = \frac{\sum_{j=1}^{J_{\delta_T, \delta_C}} GPASW_{(\delta_T, \delta_C), j}}{J_{\delta_T, \delta_C}} \quad (8)$$

where the index $TGPASW_{(\delta_T, \delta_C)}$ is a simple average of the $GPASW_{(\delta_T, \delta_C), j}$ values, since each of the $J_{(\delta_T, \delta_C)}$ obtained

GPs would represent a true oscillation pattern according to the clustering parameters (δ_T, δ_C) , and hence must be assigned the same level of importance. The highest quality solution is then simply found by identifying the clustering parameters yielding the maximum value of the $TGPASW_{(\delta_T, \delta_C)}$ index over all pairs (δ_T, δ_C) . The obtained GPs will represent the optimal multi-machine unstable oscillation patterns for the system. The optimal clustering parameters are denoted as $(\delta_{TO}, \delta_{CO})$.

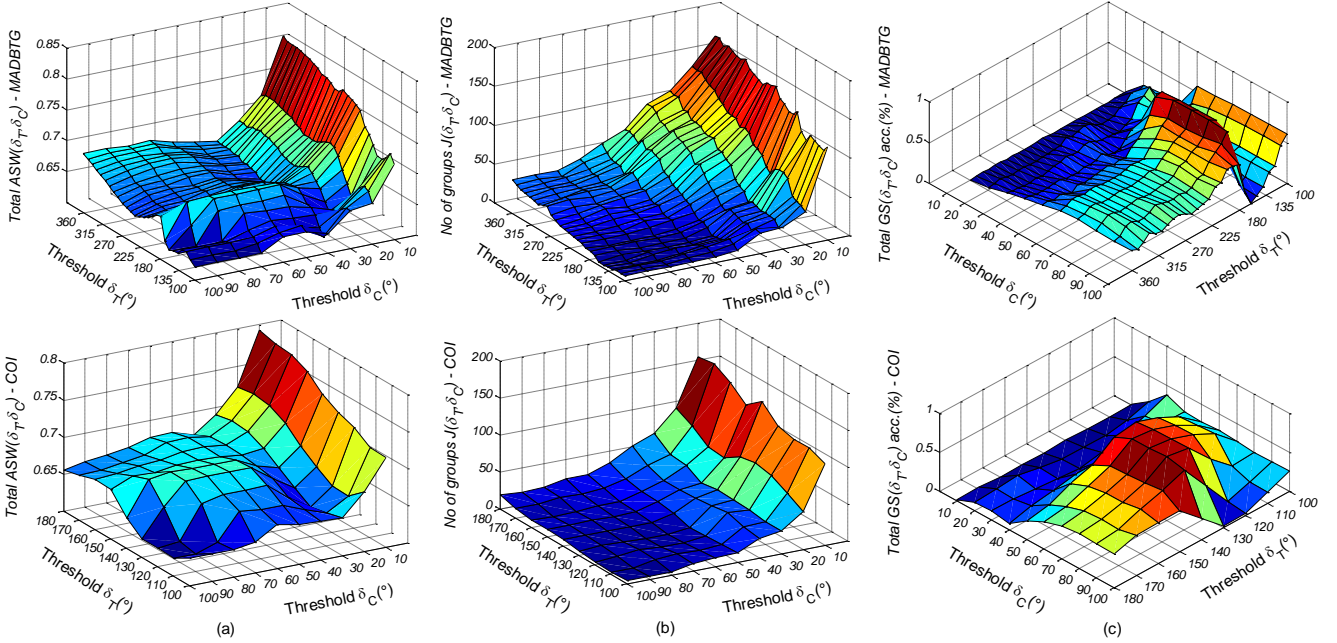


Fig. 4. Main metrics for the MADBTG (above) and COI (below). (a) ASW. (b) Number of Grouping Patterns - GP. (c) Percentage of cases that passed the GS test (accuracy).

III. APPLICATION RESULTS

The MC assessment of the testing scenario, described in Section II-A, results in 925 unstable cases out of 10,000 simulated faults. From the 925 unstable cases, 701 (76%) are single-machine and 224 (24%) are multi-machine unstable cases. Only the latter are further analyzed as proposed in Section II-C. The surface plots of Fig. 4 show the ASW values, total number $J_{(\delta_T, \delta_C)}$ of GPs, and the percentage of events that passed the GS test, respectively, as a function of (δ_T, δ_C) , for both the MADBTG and COI frames. It can be seen that all values of the ASW are greater than 0.62, indicating that the data does present classifiable trends over the entire clustering region defined. For low values of δ_C , the ASW is the highest, however, in this region, the number $J_{(\delta_T, \delta_C)}$ of GPs is also the highest and very large (around 100-160). The ASW values are therefore artificially increased there, as discussed in Section II-C. For higher values of δ_C , the number of GPs reduces considerably, which is not the case for the ASW, whose values also decrease but show an increasing trend over the region $\delta_C = 50^\circ-80^\circ$ and $\delta_T = 160^\circ-170^\circ$ for the MADBTG and $\delta_C = 50^\circ-90^\circ$ and $\delta_T = 120^\circ-150^\circ$ for the COI. In these areas, the $J_{(\delta_T, \delta_C)}$ values are less than 20, which suggests that for such regions better clusters might be obtained. More insights are provided in Fig. 4(c). For the MADBTG, relatively high values of clustering accuracy according to the GS test are observed for δ_T values of 100° and δ_C values of $50^\circ-100^\circ$, though they decrease considerably as δ_T increases. However, the highest

values of accuracy (around 70-80%) are obtained for the region $\delta_C = 50^\circ-80^\circ$ and $\delta_T = 160^\circ-170^\circ$ for the MADBTG, and $\delta_C = 50^\circ-80^\circ$ and $\delta_T = 130^\circ-150^\circ$ for the COI, all of which agrees with the trend observed in the ASW surfaces in Fig. 4(a).

Fig. 5 shows the $TGPASW$ index values allowing a better and definitive assessment of the actual quality of different GPs or clustering partitions using the properties of both the ASW and GS. It is shown that the optimal clustering parameters displaying the maximum $TGPASW$ value are $\delta_{TO} = 160^\circ$ and $\delta_{CO} = 60^\circ-80^\circ$ for the MADBTG and $\delta_{TO} = 130^\circ$ and $\delta_{CO} = 70^\circ$ for the COI. The GPs obtained with these parameters are shown in Table I and Table II, also indicating the different cluster's composition per group. By comparing these results obtained from the $TGPASW$ values shown in Fig. 5, with others obtained with (δ_T, δ_C) pairs in the region also showing high clustering accuracy according to Fig. 4(c), it can be seen that the GPs obtained with $(\delta_{TO}, \delta_{CO})$ present the minimum number of patterns and the minimum number of clusters per pattern (three for each GP). This makes them the best ones fully discriminating the unstable oscillation patterns having fewer groups. All these features are an indication of the optimality of the obtained results. In addition, it can also be seen that both the MADBTG and COI frames yield practically the same GPs, only differing slightly on $GP_6=6$, the least frequent one with less than 1% participation in the identified unstable patterns.

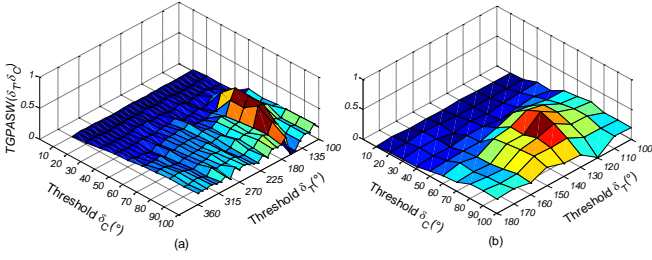


Fig. 5. Total Grouping Pattern Average Silhouette Width (TGPASW) index values. (a) MADBTG. (b) COI.

TABLE I. MULTI-MACHINE UNSTABLE GROUPING PATTERNS – MADBTG

GP_j	R_j	Freq.	Multi-machine Unstable Generator Clusters
1	111	49.6%	(G4-G5)(G2-G3/G6-G11)(G1/G12-G16)
2	74	33.0%	(G4-G7)(G2-G3/G8-G11)(G1/G12-G16)
3	19	8.5%	(G6-G7)(G2-G5/G8-G11)(G1/G12-G16)
4	17	7.6%	(G4-G7/G9)(G2-G3/G8/G10-G11)(G1/G12-G16)
5	2	0.9%	(G4/G6-G7)(G2-G3/G5/G8-G11)(G1/G12-G16)
6	1	0.4%	(G4-G5)(G2-G3/G6-G11/G16)(G1/G12-G15)

TABLE II. MULTI-MACHINE UNSTABLE GROUPING PATTERNS – COI

GP_j	R_j	Freq.	Multi-machine Unstable Generator Clusters
1	113	50.4%	(G4-G5)(G2-G3/G6-G11)(G1/G12-G16)
2	74	33.0%	(G4-G7)(G2-G3/G8-G11)(G1/G12-G16)
3	17	7.6%	(G6-G7)(G2-G5/G8-G11)(G1/G12-G16)
4	16	7.1%	(G4-G7/G9)(G2-G3/G8/G10-G11)(G1/G12-G16)
5	3	1.3%	(G4/G6-G7)(G2-G3/G5/G8-G11)(G1/G12-G16)
6	1	0.4%	(G6-G7)(G2-G5/G8-G11/G16)(G1/G12-G15)

IV. CONCLUSIONS

A methodology for identifying the optimal oscillation patterns for multi-machine transient instability in power systems is proposed in this paper. Hierarchical Clustering (HC) is used as the clustering algorithm to illustrate the approach for grouping rotor angle responses obtained probabilistically (though any other clustering algorithm could have been used equally well). Two parameters require an advanced setting, the instant of clustering formation, defined in terms of angle thresholds for measuring system separation, and the threshold for cutting-off the dendrograms produced by the algorithm. Non-specific or not clearly defined criteria for choosing both parameters have been used in past work. A statistical assessment of the clustering results obtained with an appropriately defined range of clustering parameters was performed, based on a comparison using a proposed metric which combines the properties of the well-known Average Silhouette Width (ASW) and Gap Statistic (GS) indicators, both used for measuring clustering quality and optimal selection of the number of clusters in a dataset. The efficacy of the method was proven by finding a reduced number of groups properly differentiating the multi-machine unstable oscillation patterns in a large uncertain test network, while at the same time minimizing the heuristics involved to achieve this.

The same oscillation patterns are obtained using two reference frames (Maximum Angle Difference Between Two Generators and Center Of Inertia), therefore, they have no influence on the results of the proposed clustering application. For both frames, cut-offs of around 70° were

found to yield the optimal number of groups, implying that such interval is also a good reference threshold value for general coherency identification. The methodology presented here will contribute, among others, to improved data training, classification, and forecast accuracy of on-line dynamic signature identification techniques, coherency identification, etc. Most importantly, it will facilitate the design of proper network stabilization and control strategies, focusing on the effective deployment of corrective control measures. This task being particularly facilitated by the method capability of identifying the minimum (optimal) number of oscillation patterns that are the most representative of the overall dynamic behavior of a power system.

REFERENCES

- [1] P. N. Papadopoulos and J. V. Milanovic, "Probabilistic Framework for Transient Stability Assessment of Power Systems With High Penetration of Renewable Generation," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 3078-3088, 2017.
- [2] J. D. Morales, P. N. Papadopoulos, and J. V. Milanović, "Statistical Assessment of the Impact of Renewable Energy Sources on Transient Stability," in *2018 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, 2018, pp. 1-6.
- [3] 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, 2007.
- [4] R. Agrawal and D. Thukaram, "Identification of coherent synchronous generators in a Multi-Machine Power System using Support Vector Clustering," in *2011 International Conference on Power and Energy Systems*, 2011, pp. 1-6.
- [5] 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, 2013.
- [6] T. Guo and J. V. Milanović, "Identification of power system dynamic signature using hierarchical clustering," in *2014 IEEE PES General Meeting | Conference & Exposition*, 2014, pp. 1-5.
- [7] T. Guo and J. V. Milanovic, "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.
- [8] 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, 2018.
- [9] J. D. Morales, J. V. Milanović, and P. N. Papadopoulos, "Analysis of Angular Threshold Criteria for Transient Instability Identification in Uncertain Power Systems," in *2019 IEEE Milan PowerTech*, 2019, pp. 1-6.
- [10] G. W. Milligan and M. C. Cooper, "An examination of procedures for determining the number of clusters in a data set," *Psychometrika*, Article vol. 50, no. 2, pp. 159-179, 1985.
- [11] B. S. Everitt, S. Landau, M. Leese, and D. Stahl, *Cluster Analysis*, 5th ed. Chichester, West Sussex, U.K.: John Wiley & Sons Ltd, 2011.
- [12] R. Tibshirani, G. Walther, and T. Hastie, "Estimating the number of clusters in a data set via the gap statistic," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Article vol. 63, no. 2, p. 411, 2001.
- [13] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed. Waltham, MA, USA: Morgan Kaufmann, 2012.
- [14] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. New York, USA: Springer, 2009.
- [15] L. Kaufman and P. J. Rousseeuw, *Finding Groups in Data: An Introduction to Cluster Analysis*. Hoboken, NJ, USA: John Wiley & Sons, Inc., 1990.