

Generator Coherency Identification Algorithm Using Modal and Time-Domain Information

Theodoros Kyriakidis ^{#1}, Rachid Cherkaoui ^{*}, Maher Kayal [#]

[#] *Electronics Laboratory, Swiss Federal Institute of Technology
CH-1015 Lausanne, Switzerland*

¹ *theodoros.kyriakidis@epfl.ch*

^{*} *Power Systems Group, Swiss Federal Institute of Technology
CH-1015 Lausanne, Switzerland*

Abstract—Coherency group identification is an integral constituent part of the wider field of reduction techniques in power systems. It consists of separating the machines in the system into groups that feature similar behavior. This paper presents a coherency identification algorithm for dynamic studies. The algorithm combines both modal and time domain techniques in an effort to combine the merits of both approaches. Its outcome is a suggested optimal number of clusters alongside the clustering itself. Tests have been conducted on a sample power system of 39 buses and its validity has been demonstrated.

Index Terms—Generator coherency, k-means clustering, modal analysis, power system dynamics, Principal Component Analysis.

I. INTRODUCTION

Dynamic reduction is the procedure in which the complexity/size of a power system is reduced, trying to retain a dynamic behavior as close as possible to that of the original unreduced system. The restructuring of the power system underway for the last couple of decades, and the soon-to-be advent of the *smart grid* has brought up a renewed interest for reduction techniques [1].

- Power system analysis computations, especially Time-Domain (TD) simulations, are computationally burdensome, and performance scales poorly as the size of the system increases.
- Full-detail models are usually unavailable, as system operators are increasingly unwilling to share full models of their subsystems with neighboring entities for confidentiality reasons, mainly due to market competition.

Reduction can be performed both online and offline and is applied for a plethora of studies: Dynamic Security Assessment (DSA), over-voltage and switching studies [2], AC filter design calculations [2], inter-area oscillation studies [3], Special Protection Schemes (SPS) design (e.g. islanding procedure) [4], post-event vulnerability analysis, etc. Reference [5] provides a comprehensive bibliographical review on various coherency, reduction and aggregation techniques.

This paper is focused on reduction applications for DSA. Transient stability is one of the main concerns of DSA. It has been defined as the ability of the system to retain a state of operating equilibrium after being subjected to a given

set of disturbances [6]. In transient events examined during DSA studies, the behavior of the system is described by complex highly non-linear differential-algebraic equations (DAEs), which are handled by specific solvers known as time-domain (TD) engines. Given that mainly generators contribute to the complexity of the dynamics of the system (electromechanical oscillation modes) generator reduction is the main scope of dynamic reduction for transient stability studies. Generally the following steps apply. 1) The Internal System (IS) and the External System (ES) are defined; 2) the groups of generators that feature similar behavior are determined in the ES; and 3) generators in each coherency group are aggregated.

This paper proposes an algorithm for the point 2 of the above, ie. the *identification of coherent generator groups*. Coherency group identification techniques can be broadly separated in two categories; to the ones that employ information from the solution of the full DAEs that describe the system behavior, and to the ones that extract coherency from a linearized model of the system around a steady-state operating point. We term the former time-domain techniques and the latter modal techniques.

Modal techniques rely on forming the state equations of the linearized system [1], [7]–[14]. This way, the behavior of the system around the operation point is captured well. By expressing its dynamics in a state-space form much can be gained from solid theoretical work on dynamic systems theory. However, linearization renders resulting groupings better suited for small-signal analysis rather than for transient stability analysis.

Time-domain techniques rely on information extracted from trajectories resulting from the time-domain simulation of the system under certain conditions [3], [4], [15]–[19]. They allow for unlimited modelling capabilities and are able of capturing big-signal dynamics. However, usually they employ uninformative (or even uniform) priors regarding the clustering participation. Moreover, they tend to make little (or even no) use of information available by the topology of the system.

The proposed algorithm employs information from both the modal analysis and the time-domain simulation of the power system. The core of the algorithm is the latter, in that, clustering is performed based on the post-fault trajectories

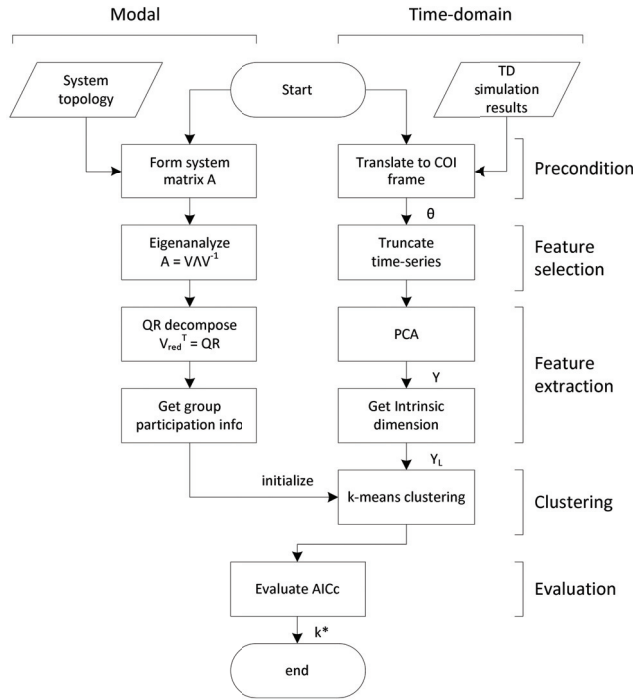


Fig. 1. Algorithm overview

of the angles of the generators. Information from the modal domain is used to facilitate clustering in the time-domain.

The rest of the paper is organized as follows. A theoretical background on coherency group identification alongside the proposed algorithm is presented in section II. In section III the application of the algorithm on a sample test system is demonstrated and a characterization of its validity and consistency is performed. Finally, section IV gives conclusions and suggests issues to be tackled in future work.

II. PROPOSED ALGORITHM

Dividing the machines of the system into coherency groups is actually forming a set of similarity classes. In order to form these classes a criterion needs to be established. This work is primarily concerned with ϵ -coherency. That is, between any two generators i and j of a coherency group, (1) needs to be satisfied.

$$\delta_i(t) - \delta_j(t) = 0(\epsilon), \forall t \quad (1)$$

A synoptical overview of the algorithm is given in fig. 1. Each block of the diagram is explained in detail in sections that follow.

There are two flows in the algorithm, the modal and the time-domain one. The latter is the core of it. The system is subjected to a perturbation, so as to excite its dynamics. The perturbation can be chosen according to engineering judgement/experience of the system operator. Scrutiny of this selection is out of scope of this work; however, in [18] it

is suggested that perturbations that result in a stable post-fault response are used. The response of the system to the perturbation is simulated with a conventional TD engine, and then clustering takes place based on the behavior of the generators. Modal information is used for the initialization of the clustering algorithm. Finally, given that the number of clusters to be determined is not known a priori, the algorithm examines different alternatives and proposes an optimal value.

A. Time domain flow

The input of the time-domain flow of the algorithm is the time response of the system to the perturbation. Notice that there is no constraint on the modelling of the power system.

Trajectories of n_{ts} time samples can be viewed as points on a n_{ts} dimensional space. As n_{ts} is normally expected to be high, the problem suffers from the *curse of dimensionality*. In an effort to alleviate the problem, *dimension reduction* is performed. The goal of dimension reduction is to reduce the dimension of the data set of the problem down as close as possible to the *intrinsic dimension* of the data. The intrinsic dimension of a signal is the minimum number of independent variables necessary to describe the signal. Dimension reduction is performed in two stages, *feature selection* and *feature extraction*. Feature selection refers to what part/form of the original input to retain, and feature extraction refers to how this data can be manipulated to approximate as closely as possible the intrinsic dimension.

In this problem, the time series are first preconditioned to remove the trend component. Feature selection refers to retaining only a part of the time series that capture the dynamics of the system. Finally, feature extraction is done by manipulating the reduced time series using Principal Component Analysis (PCA).

1) *Time series preconditioning*: The angle time series of each generator i is decomposed into three constituent parts.

$$\delta_i(t) = \theta_i^r(t) + \underbrace{\theta_i^{pf} + \delta_{COI}(t)}_{T_i^t(t)} \quad (2)$$

In the above, $\delta_{COI}(t)$ is the trajectory of the angle of the Center Of Inertia (COI) system reference frame.

$$\delta_{COI}(t) = M_{COI}^{-1} \sum_{j=1}^{n_g} M_j \delta_j(t) \quad (3)$$

$$M_{COI} = \sum_{j=1}^{n_g} M_j$$

θ_i^{pf} is the post-fault steady-state internal angle of generator i expressed in the COI frame. The sum $T_i^t(t) = \theta_i^{pf} + \delta_{COI}(t)$ can be viewed as the trend component of $\delta_i(t)$. When $T_i^t(t)$ is subtracted, the remaining $\theta_i^r(t)$ captures all the relevant dynamics of generator i .

2) *Feature selection*: An index called aggregate Angle Radius (AR) is employed to retain a number of oscillations of the machine angles.

Firstly, the machines of the system G are separated into two groups, the non-critical machines that retain their stability N ,

and the critical ones that lose it C . A machine i is non-critical if and only if its angular separation from the least advanced machine in angle terms remains bounded by δ_{th} , for each time instant t_k . In this work a threshold value of $\delta_{th} = 180^\circ$ is used.

$$N = \{G_i : |\delta_i(t_k) - \min_{j \in G}(\delta_j(t_k))| \leq \delta_{th}, \forall t_k\} \quad (4)$$

$$C = G - N$$

Thereafter, AR is calculated for N -machines as follows.

$$AR(t) = \sum_{i \in N} M_i(\theta_i^r(t))^2 \quad (5)$$

AR can be viewed as a measure of the weighted aggregate angular distance of the system from its Stable Equilibrium Point (SEP). Critical machines are not taken into account in this calculation as normally they are not expected to feature periodicity after the loss of their synchronism with the rest of the grid (aperiodic oscillations).

Peak detection is run on $AR(t)$ to determine the third peak, at $t = t_{3p}$. Then, all time series θ^r are truncated to n samples, retaining values for $t \in [0, t_{3p}]$. Testing has shown that the three first oscillations of the AR index are enough to capture the dynamics of the system. This is in accord with the fact that any coherency/interdependency between the trajectories of machine angles is manifested more strongly during the first moments after the perturbation.

3) *Feature extraction*: Feature extraction is done using Principal Component Analysis (PCA) [16]. In this case the application of PCA is much like Multi-Dimensional Scaling (MDS) in the sense that it projects θ^r points from the n -dimensional space to a space of a much lower dimension n_{dim} that is an approximation \hat{n}_{intr} of the intrinsic dimension n_{intr} of the data set.

$$n_{dim} = \hat{n}_{intr} \approx n_{intr} \quad (6)$$

PCA uses an orthogonal transformation to transform a set of correlated variables to a set of uncorrelated ones called *principal components*. The latter are a linear combination of the original ones and are derived in decreasing order of importance, in that the component determined at each step accounts for as much as possible of the variation of the original data.

Let the input matrix be:

$${}_{(n_g \times n)}X = \begin{bmatrix} \theta_1^r(t) \\ \vdots \\ \theta_{n_g}^r(t) \end{bmatrix} \quad (7)$$

Let the Singular Value Decomposition (SVD) of X^T be:

$$X^T = W^T \Sigma V^T \quad (8)$$

Then, the PCA preserving dimensionality is given by:

$$Y = XW^T \quad (9)$$

Matrix W^T is by definition of the SVD an orthogonal matrix $W^T W = W W^T = I$. It is also the matrix of eigenvectors of the covariance matrix $C = X X^T$. Rearranging (9):

$$X = YW = [y_{*1} | \dots | y_{*n}] \begin{bmatrix} w_{1*}(t) \\ \vdots \\ w_{n*}(t) \end{bmatrix} \quad (10)$$

In (10), Y is the *scores* matrix and W is the *principal components* matrix. Each original time series i is decomposed to a weighted sum of the principal components.

$$\theta_i^r(t) = y_{i1}w_{1*}(t) + \dots + y_{in}w_{n*}(t) \quad (11)$$

As said, principal components w_{j*} are calculated in a decreasing order of importance. So, a reduced-dimensionality representation can be achieved by projecting X down to the reduced space defined by only the first n_{dim} principal components.

$$\hat{X} = {}_{(n_g \times n_{dim})}Y_L \times {}_{(n_{dim} \times n)}W_L \quad (12)$$

The singular values in Σ are the eigenvalues of the covariance matrix C . Each value in Σ is proportional to the portion of the variance that is correlated with each eigenvector (row of W). So, based on Σ values, n_{dim} is selected to retain at least 99% of the original variance. This results in an acceptable reconstruction of X by \hat{X} and (6) holds.

Rows of Y_L correspond to generators, and columns to scores of PCA the principal components. Y_L is the time-domain input to the core of the clustering procedure that is presented in detail in section II-C.

B. Modal flow

Modal information is used to initialize the time-domain-based clustering algorithm. For this a linearization of the system equations is performed around the steady state operating point.

1) *Derivation of state-space equations*: Loads are expressed as constant impedances at their respective buses.

$$\bar{Y}_{di} = \frac{P_{di} - jQ_{di}}{V_{0i}^2} \quad (13)$$

Using (13) the admittance matrix of the system is augmented to account for the loads (constant impedances to ground) $Y_0 \xrightarrow{(13)} Y_0^+$.

For the generators, the classical generator model is used. Generators are modeled with their Norton equivalent of a transient reactance x'_i in parallel to a current source $\bar{I}'_i = \frac{E'_i}{x'_i} e^{j(\delta_i - \frac{\pi}{2})}$. The swing equation applies for the dynamic electromechanical behavior. Notice that nominal synchronous speed is assumed and damping is neglected.

$$\begin{aligned} \dot{\delta}_i &= \Omega_r \omega_i = \omega_i \\ \dot{\omega}_i &= \frac{P_{mi} - P_{ei}}{M_i} - D_i \omega_i = \frac{P_{mi} - P_{ei}}{M_i} \end{aligned} \quad (14)$$

The load-augmented admittance matrix is further augmented with the (Norton) transient reactances of the generators $Y_0^+ \xrightarrow{x'_i}$

Y_0^{++} . If a different model is used for the generators, this step needs to be modified accordingly.

The nodal equations are rearranged to group injection nodes as follows.

$$\begin{bmatrix} n_g \times 1 I_g \\ (n-n_g) \times 1 0 \end{bmatrix} = Y_p^{++} \begin{bmatrix} V_g \\ V_{rest} \end{bmatrix} \quad (15)$$

Where:

$$Y_p^{++} = \begin{bmatrix} Y_{gg} & Y_{gr} \\ Y_{rg} & Y_{rr} \end{bmatrix} \quad (16)$$

Passive nodes are eliminated from (16) using Kron Reduction [20].

$$Y = Y_{gg} - Y_{gr} Y_{rr}^{-1} Y_{rg} \quad (17)$$

The electrical power is $S = VI^* = V(YV)^* = VV^*Y^*$ so for each generator.

$$P_{ei} = \Re\left\{ \sum_{j=1}^{n_g} \bar{V}_i \bar{V}_j^* Y_{ij}^* \right\} = \sum_{j=1}^{n_g} V_i V_j Y_{ij} \cos(\delta_i - \delta_j - \phi_{ij}) \quad (18)$$

By combining (14) and (18) the system is written in state space form as follows.

$$\begin{bmatrix} \dot{\delta} \\ \dot{\omega} \\ y \end{bmatrix} = \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} f(x, u) \\ g(x, u) \end{bmatrix} \rightarrow \begin{bmatrix} \Delta \dot{x} \\ \Delta \dot{y} \end{bmatrix} = \begin{bmatrix} A \Delta x + B \Delta u \\ C \Delta x + D \Delta u \end{bmatrix} \quad (19)$$

The system matrix A is the Jacobian of the system calculated for steady state conditions.

$$A = J(x)|_{x=x_0} = \begin{bmatrix} \frac{\partial \delta}{\partial \delta} & \frac{\partial \delta}{\partial \omega} \\ \frac{\partial \omega}{\partial \delta} & \frac{\partial \omega}{\partial \omega} \end{bmatrix}_{x=x_0} = \begin{bmatrix} J_{\delta\delta} & J_{\delta\omega} \\ J_{\omega\delta} & J_{\omega\omega} \end{bmatrix} \quad (20)$$

Let $L = [l_{ij}] = J_{\omega\delta}$ with:

$$\begin{aligned} l_{ij} &= -M_i^{-1} V_i V_j Y_{ij} \sin(\delta_i - \delta_j - \phi_{ij}) \\ l_{ii} &= M_i^{-1} \sum_{j=1}^{n_g} V_i V_j Y_{ij} \sin(\delta_i - \delta_j - \phi_{ij}) \end{aligned} \quad (21)$$

Since $J_{\delta\delta} = 0$, $J_{\omega\omega} = 0$ and $J_{\delta\omega} = I$, then A can be rewritten as follows.

$$A = \begin{bmatrix} 0 & I \\ L & 0 \end{bmatrix} \quad (22)$$

Notice that unless damping D had been neglected, it would be $J_{\omega\omega} = -diag(D)$.

2) *Eigenanalysis*: Let the eigendecomposition of L be:

$$L = V \Lambda V^{-1} \quad (23)$$

Where the eigenvalues are $\Lambda = diag(\lambda_0, \lambda_1, \dots, \lambda_{n-1})$ and the corresponding eigenvectors $V = [V_0 | V_1 | \dots | V_{n-1}]$.

Matrix L can be used to analyse A , since L captures all electromechanical dynamics of A (the original system). It holds that the eigenvalues of A are complex pairs of the eigenvalues of L : $\Lambda_A = \pm \sqrt{\Lambda}$.

Eigenvalues of L are non-positive. It is supposed without loss of generality that they are sorted $0 = \lambda_0 > \lambda_1 > \dots > \lambda_{n-1}$. Zero-frequency eigenvalue $\lambda_0 = 0$ corresponds to the aggregate motion of the machine speeds and angles.

Suppose that we want to extract cluster initialization information for k clusters. Then a reduced version of the eigenvector set is retained.

$$V_{red} = [V_1 | \dots | V_k] \quad (24)$$

In this matrix, rows correspond to generators and columns to system modes. V_{red} is then normalized row-wise, as it has been noted that this improves separability of the participation of generators between different modes. This normalization implies that coherency is decided upon angular proximity of the generator vectors in the system-modes space. Reference [9] suggests a stricter criterion requiring match for the magnitude as well.

3) *Clustering initialization*: Next, the n_g row elements of V_{red} need to be assigned to k clusters in the k -dimensional space. Reference [8] proposes LU-decomposition with complete pivoting to determine a set of k reference generators, according to the pivot elements selected at each step of the decomposition. Then, V rows are translated to the coordinate system defined by the reference generators. Cluster participation is determined based on the translated coordinates of the rows of V .

This work employs QR decomposition instead. The orthonormality of the base resulting from the QR algorithm is preferable as it results in better separability between the machines. This is not necessarily true with LU decomposition as in that case the base of the new coordinate system is not necessarily orthogonal.

QR decomposition of a matrix A is its decomposition to the product of an orthogonal matrix Q and an upper triangular matrix R . If column pivoting is used and A is of size $(m_1 \times m_2)$ then the following holds.

$$A = QRP^T = Q [R_1 \ R_2] P^T \quad (25)$$

Where P is the permutation matrix of size $(m_2 \times m_2)$, R_1 is upper triangular of size $(m_1 \times m_2)$ and R_2 a matrix of size $(m_2 - m_1 \times m_1)$. Supposing that $rank(A) = m_1$, i.e. A has m_1 linearly independent columns, then columns of Q form an orthonormal basis of the column space of A . For $c \leq m_1$, the c^{th} column of $A \times P$ depends of the first c columns of Q , and thus R_1 is upper triangular.

The analysis described hereinabove is performed for $A = V_{red}^T$. Generators relate to matrix columns as V_{red} is transposed. Q columns are selected as the base for the new coordinate system and R columns contain the translated coordinates of $V_{red}^T P$ columns.

Finally, cluster membership info is extracted from R^T . Let sgn_j be the sign of the j^{th} element in the diagonal of R_1 . Let elements of column j of R^T have the maximum absolute value signed with sgn_j , in row i . This signifies that the generator corresponding to this row belongs to the cluster whose behavior is dominated by the j^{th} column of the base matrix Q . In other words, generators corresponding to all rows that have a maximum absolute element at column j belong to the same cluster. The permutation matrix P links R_T rows with real system generators.

The result of this procedure is an array of sets $C_i^0, \forall i \in [1, k]$ that contain the indexes of generators participating in each cluster. For each cluster, the corresponding centroid c_0 in the Y_L space is computed.

$$c_i^0 = \underset{j \in C_i^0}{ave}(Y_L[j, :]), \forall i \in [1, k] \quad (26)$$

C. Clustering

Coherency group identification can be classified as unsupervised learning, in that no knowledge on generator cluster participation is available a priori. In this work the *k-means clustering* algorithm is used. It is a method which aims to partition n_s observations into k clusters trying to minimize the within-cluster sum of squares (see also (28)). In our case Y_L is the input of the algorithm, and n_s is the number of generators n_g .

A major factor affecting the performance of k-means is the initial position of the centroids. Unless additional information is provided, initialization is done in a random manner, fact that has detrimental effects on the quality of the clustering. In our case, centroid initialization information is extracted from the modal analysis, as detailed in section II-B

Another issue, distinct from the clustering proper but closely related to it, is the *determination of the number of clusters k*. For this, a series of different cluster numbers are investigated in a loop. In [21], $k_0^* = \sqrt{n/2}$ is proposed as a rule-of-thumb optimal number of clusters. In this work candidate cluster numbers between 2 and $k_{max} = 3k_0^*$ are investigated as per the algorithm below.

Algorithm 1 Investigation of different numbers of clusters k

```

k ← 2
while k ≤ kmax do
  MODALANALYSIS(k)
  ↓ init
  K-MEANS(k)
  k ← k + 1
end while
EVALUATE(k = 2 . . . kmax)

```

D. Clustering Evaluation

Determining the optimal k is done according to the corrected Akaike Information Criterion (AICc). The original AIC measures the goodness of fit of a model, with respect to alternative ones [22]. In this application, “model” refers to the number of clusters k . Given a set of candidate models, the preferred is the one with the minimum AIC value. When the number of samples is small, as in our case with $n_s = n_g$, the use of the corrected AIC (AICc) is recommended, as it accounts for the finite sample size [23].

$$AICc(k) = 2p - 2\ln(L) + \frac{2p(p+1)}{n_s - p - 1} \quad (27)$$

In (27):

- $n_s = n_g$ is the number of samples in the data set. It coincides with the number of generators.
- $p = k \cdot n_{dim}$ is the number of the parameters in the model. In our case, there are as many parameters as there are centroid coordinates in the n_{dim} -dimensional space, whose values are fixed.
- $\ln(L)$ is the maximum log-likelihood of the data for k clusters and it is essentially a measure of distortion. If the model underlying the k-means is assumed to be a Gaussian mixture with hard assignment, uniform cluster priors and identical spherical covariance matrices, then the following equality holds [24].

$$-2\ln(L) = RSS = \sum_{i=1}^k RSS_i = \sum_{i=1}^k \sum_{j \in C_i} (Y_L[j, :] - c_i)^2 \quad (28)$$

Where RSS is the overall residual sum of squares, which by the way is the objective function of the k-means minimization problem. RSS_i is the residual sum of squares within the cluster i .

After the calculation of $AICc(k)$ for different k values the relative likelihood is calculated for each model.

$$L_{AICc}(k) = e^{\frac{\min(AICc) - AICc(k)}{2}} \quad (29)$$

The model k^* that is the best fit to our data is the one with $L_{AICc}(k^*) = 1$.

III. NUMERICAL RESULTS

The algorithm described in section II has been implemented in MATLAB. It has been applied on the 39-bus 10-generator New England test system found in [25] and shown in fig. 2. The generators of the system are modeled with the classical second-order synchronous machine model. It has been noted that coherency would manifest in a similar manner regardless of the modelling detail in the system [26]. All simulations were performed using the MatDyn time-domain simulator [27] for MATLAB, in a typical modern desktop PC (Intel Core i7 4 × 2.80GHz, 8GB RAM) running Microsoft Windows 7 x64.

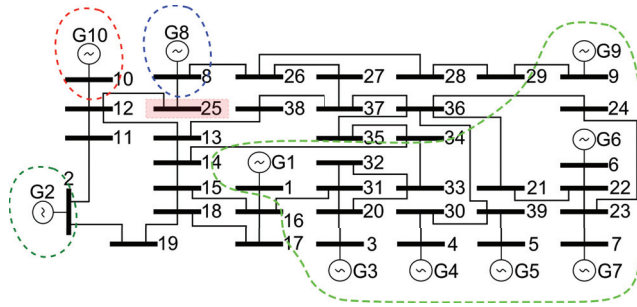


Fig. 2. New England 39-bus 10-generator system and clustering results for a fault at bus #25

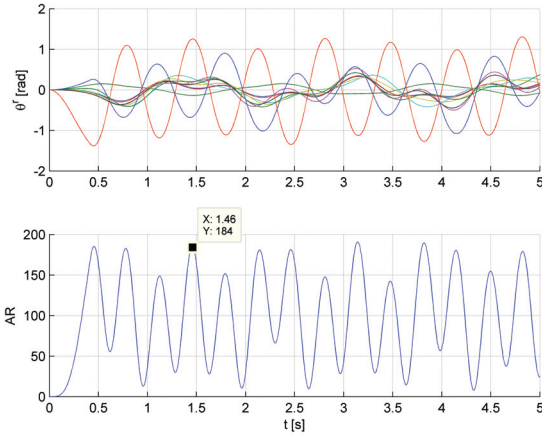


Fig. 3. θ^r and Angle Radius values

A. Detailed application on a specific perturbation

The perturbation is chosen to be a perfect 3ϕ fault to ground on bus 25, applied at $t_f = 0s$ and cleared at time $t_{cl} = 0.420s$. The faulted bus is highlighted in light red in fig. 2. The simulation is performed for $t_{sim} = 10s$. For a timestep of $t_{step} = 0.005s$ this results in $n_{ts} = 2003$ samples for the simulated trajectories. Fig. 3 shows θ^r and AR values as per (2) and (5) respectively, for the first 5s.

The fourth peak of AR is detected on sample $n = 294$ at time $t_{3p} = 1.455s$. θ^r trajectories are truncated and PCA is performed according to section II-A3. Scores for the three first principal components are visualized in the scatter plot of fig. 4.

The approximation of the intrinsic dimension of the data set, i.e. the determination of how many principal components to retain is done according to table I and section II-A3. $n_{dim} = 3$ is chosen, as this results in retaining 99.6% of the variance of the original data.

The numbers of clusters to investigate is according to II-C.

$$2 \leq k \leq k_{max} = 3k_0^* = 3\sqrt{\frac{n_{gen}}{2}} = 9 \quad (30)$$

Algorithm 1 is applied for ks in question. Cluster initial-

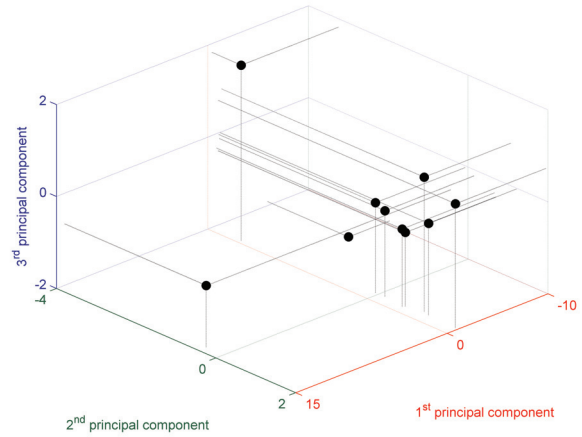


Fig. 4. Scores of the 3 first principal components

TABLE I
RETAINED VARIANCE AS A PERCENTAGE AGAINST
THE NUMBER OF RETAINED PRINCIPAL COMPONENTS

Retained PCA components	1	2	3	4	5	6
Retained variance	0.887	0.9665	0.9964	0.9997	~1	~1

ization information by the modal analysis, as well as final clustering in the time domain by the k-means is given in table II.

Determination of the final number of clusters is done according to the relative likelihood of the different models, as per section II-D. Fig. 5 shows AIC_c and L_{AIC_c} values for different cluster number options. From the figure it is obvious that the optimal number of clusters suggested by the algorithm is $k^* = 4$.

This clustering is projected back to the time (θ^r) domain in fig. 6. Three clusters are single machine ones: $G2$ in dark green, $G8$ in blue and $G10$ in red. The cluster in light green aggregates the response of generators $G1$, $G3$, $G4$, $G5$, $G6$, $G7$ and $G9$.

TABLE II
CLUSTERING SUGGESTED BY MODAL INITIALIZATION
AND FINALLY COMPUTED BY K-MEANS

k	Modal initialization	k-means clustering
2	{1,2,3,4,5,6,7} {8,9,10}	{1,2,3,4,5,6,7,8,9} {10}
3	{1,2,3} {4,5,6,7} {8,9,10}	{1,2,3,4,5,6,7,9} {8} {10}
4	{1,2,3} {4,5} {6,7} {8,9,10}	{1,3,4,5,6,7,9} {2} {8} {10}
5	{1,3} {2,9} {4,5} {6,7}	{1,3,5,7} {2} {4,6,9} {8} {10}
6	{1,2} {3} {4,5} {6,7} {8,10}	{1,3,5,7} {2} {4} {6,9} {8} {10}
7	{1,2} {3} {4,5} {6,7} {8}	{1,3} {2} {4} {5,6,7} {8} {9} {10}
8	{1,2} {3} {4,5} {6} {7} {8}	{1,3} {2} {4} {5,7} {6} {8} {9} {10}
9	{1} {2,4} {3} {5} {6} {7}	{1,3} {2} {4} {5} {6} {7} {8} {9} {10}

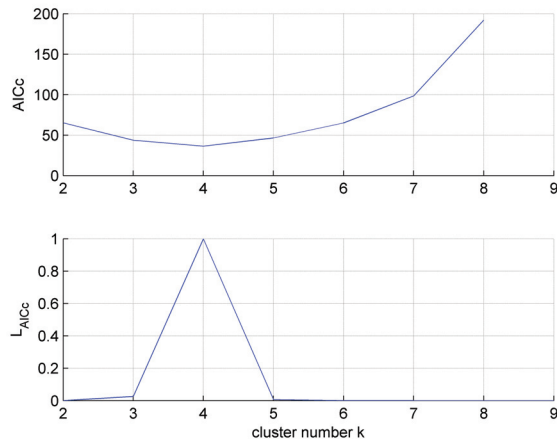
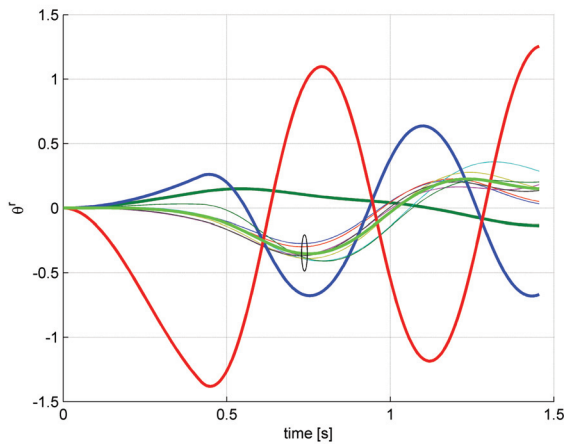
Fig. 5. $AICc$ and L_{AICc} values for different cluster numbers k 

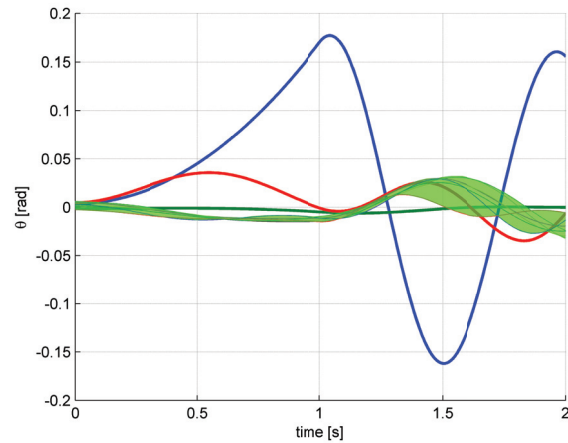
Fig. 6. Clustering projected to the time domain

Qualitatively we can say that both the clustering number choice and the clustering itself are quite well-placed. From the response of the machines it can be clearly seen that there are indeed four groups of machines with different response patterns to the perturbation.

The respective topographical clustering is superimposed in color on the network topology in fig. 2. Again, we can see that the results are consistent with the scenario and the topology in question. All three single-generator clusters are electrically near the faulted bus. Whereas, all generators that are relatively far from it exhibit similar response to the disturbance, and are thus grouped in the same cluster (in light green).

The clustering results retrieved in the previous section can be validated by subjecting the system to a different perturbation that is expected to excite it in a similar way. The event $E = \{500ms\ 3\phi\ \text{fault on bus}\#11\}$ is chosen and the resulting generator trajectories are shown in fig. 7.

Color coding remains consistent with the previous example.

Fig. 7. Generator θ responses of a 500ms 3ϕ fault on bus #11TABLE III
CLUSTERING RESULTS FOR DIFFERENT EVENTS ON THE SYSTEM

Event	Number of clusters	Clustering
E_1	1s 3ϕ fault at B12	5 {1,3,4,5,6,7} {2} {8} {9} {10}
E_2	1s 3ϕ fault at B19	1 {1,2,3,4,5,6,7,8,9,10}
E_3	1s 3ϕ fault at B25	4 {1,3,4,5,6,7,9} {2} {8} {10}
E_4	1s 3ϕ fault at B27	2 {1,3,4,5,6,7,8,9,10} {2}
E_5	1s 3ϕ fault at B33	2 {1,2,3} {4,5,6,7,8,9,10}
E_6	1s 3ϕ fault at B36	2 {1,2,3,6,7,8,9,10} {4,5}

G_2 in dark green, G_8 in blue and G_{10} in red are the single-generator clusters, which indeed have a distinct dynamic behavior. Generators G_1 , G_3 , G_4 , G_5 , G_6 , G_7 and G_9 are clustered together, and their responses are shaded in light green. It is apparent that the proposed clustering successfully captures the similar behavior of the machines in this group.

B. Application on different perturbations

The algorithm has been applied for a number of different events and respective results can be found in table III.

E_2 takes place near the slack of the system G_2 and the algorithm proposes a single cluster, result that is deemed reasonable. The slack generator buffers the shock of the perturbation, as its huge inertia (mechanical starting time $M_2 = 500s$ for modelling purposes) absorbs most of the kinetic energy injected in the system during the fault-on period. This results in dynamics that are satisfactorily enough approximated even with a single cluster. This can be useful e.g. when the system under study is a subsystem of a greater entity.

In events E_5 and E_6 two clusters are enough to closely describe the behavior of the system. In the case of E_5 the system is split in two groups: G_1 , G_2 and G_3 who are electrically closer to the faulted bus, and thus more perturbed. The rest of the generators swing together, following the same weak oscillation mode. For E_6 the same applies; the system is

split between two clusters: one electrically closer to the faulted bus consisting of G_4 and G_5 , and another encompassing the remaining machines.

Faults at the electrical region around buses 12, 25, etc. seem to be able to excite more complex dynamics on the system. In both cases generators G_2 , G_8 and G_{10} that are electrically closer to the disturbance form single-machine clusters.

IV. CONCLUSIONS

In this paper, a generator coherency group identification algorithm has been presented. The algorithm derives information from both the modal analysis of the system and the time-domain simulation, in an effort to combine merits of the two approaches.

In the modal domain, system equations are linearized around the steady state operating point and clustering initialization is derived from the eigen-analysis of the dynamics of generators. In the time domain, a perturbation is selected to excite the system. The resulting trajectories are truncated and dimensionally reduced with PCA, and the resulting features are fed to a k-means clustering algorithm that performs the clustering proper. Clusterings for different cluster numbers are evaluated using the corrected Akaike Information Criterion. The final outcome of the algorithm is a suggested optimal number of clusters alongside the clustering proper.

The algorithm has been tested on a sample test system of 39 buses and has been proved to derive consistent results both in the proposed number of clusters and in the clustering itself.

This work can be extended in two ways. Firstly, the exciting perturbation selection needs to be systematized as noted in section II. Then, it needs to be put in the frame of a complete reduction methodology (see also section I). This would involve the selection of the internal and external area of studies, as well as the aggregation of the generators in the resulting coherency groups.

ACKNOWLEDGMENTS

This work is part of the Innovation program on Energy of EPFL Middle East. The authors would also like to thank S. Perdakis, currently with CNBI at the EPFL, for the fruitful discussions on clustering algorithms.

REFERENCES

- [1] B. Marinescu, B. Mallem, and L. Rouco, "Large-scale power system dynamic equivalents based on standard and border synchrony," *Power Systems, IEEE Transactions on*, vol. 25, no. 4, pp. 1873–1882, 2010.
- [2] X. Lei, D. Povh, and O. Ruhle, "Industrial approaches for dynamic equivalents of large power systems," in *Power Engineering Society Winter Meeting, 2002. IEEE*, vol. 2, 2002, pp. 1036–1042 vol.2.
- [3] M. Jonsson, M. Begovic, and J. Daalder, "A new method suitable for real-time generator coherency determination," *Power Systems, IEEE Transactions on*, vol. 19, no. 3, pp. 1473–1482, 2004.
- [4] N. Senroy, "Generator coherency using the hilbert-huang transform," *Power Systems, IEEE Transactions on*, vol. 23, no. 4, pp. 1701–1708, 2008.
- [5] R. Singh, M. Elizondo, and S. Lu, "A review of dynamic generator reduction methods for transient stability studies," in *Power and Energy Society General Meeting, 2011 IEEE*, 2011, pp. 1–8.
- [6] P. Kundur, J. Paserba, V. Ajjarapu, G. Andersson, A. Bose, C. Canizares, N. Hatziaargyriou, D. Hill, A. Stankovic, C. Taylor, T. Van Cutsem, and V. Vittal, "Definition and classification of power system stability ieee/cigre joint task force on stability terms and definitions," *Power Systems, IEEE Transactions on*, vol. 19, no. 3, pp. 1387–1401, Aug. 2004.
- [7] D. Chaniotis and M. A. Pai, "Model reduction in power systems using krylov subspace methods," *Power Systems, IEEE Transactions on*, vol. 20, no. 2, pp. 888–894, 2005.
- [8] J. H. Chow, J. Cullum, and R. A. Willoughby, "A sparsity-based technique for identifying slow-coherent areas in large power systems," *Power Apparatus and Systems, IEEE Transactions on*, vol. PAS-103, no. 3, pp. 463–473, 1984.
- [9] I. Erlich, Z. Kasztel, and P. Schegner, "Enhanced modal based technique for construction of power system dynamic equivalents," 24–28 June 2002 2002.
- [10] H. Kim, G. Jang, and K. Song, "Dynamic reduction of the large-scale power systems using relation factor," *Power Systems, IEEE Transactions on*, vol. 19, no. 3, pp. 1696–1699, 2004.
- [11] R. Nath, S. S. Lamba, and K. S. Prakasa Rao, "Coherency based system decomposition into study and external areas using weak coupling," *Power Apparatus and Systems, IEEE Transactions on*, vol. PAS-104, no. 6, pp. 1443–1449, 1985.
- [12] G. N. Ramaswamy, G. C. Verghese, L. Rouco, C. Vialas, and C. L. DeMarco, "Synchrony, aggregation, and multi-area eigenanalysis," *Power Systems, IEEE Transactions on*, vol. 10, no. 4, pp. 1986–1993, 1995.
- [13] S. Yi, F. Wen, and S. Yi, "A new approach for identifying coherent generator groups in large scale power systems," in *Electric Utility Deregulation and Restructuring and Power Technologies (DRPT), 2011 4th International Conference on*, 1993, pp. 332–336.
- [14] S. B. Yusuf, G. J. Rogers, and R. T. H. Alden, "Slow coherency based network partitioning including load buses," *Power Systems, IEEE Transactions on*, vol. 8, no. 3, pp. 1375–1382, 1993.
- [15] H. A. Alsafih and R. Dunn, "Determination of coherent clusters in a multi-machine power system based on wide-area signal measurements," in *Power and Energy Society General Meeting, 2010 IEEE*, 1974, pp. 1–8.
- [16] K. K. Anaparthi, B. Chaudhuri, N. F. Thornhill, and B. C. Pal, "Coherency identification in power systems through principal component analysis," *Power Systems, IEEE Transactions on*, vol. 20, no. 3, pp. 1658–1660, 2005.
- [17] M. Davodi, H. R. Modares, E. Reihani, and A. Sarikhani, "Coherency approach by hybrid pso, k-means clustering method in power system," in *Power and Energy Conference, 2008. PECon 2008. IEEE 2nd International*, 2008, pp. 1203–1207.
- [18] K. Mei, S. M. Rovnyak, and C.-M. Ong, "Clustering-based dynamic event location using wide-area phasor measurements," *Power Systems, IEEE Transactions on*, vol. 23, no. 2, pp. 673–679, 2008.
- [19] Y. Suzuki and I. Mezic, "Nonlinear koopman modes and coherency identification of coupled swing dynamics," *Power Systems, IEEE Transactions on*, vol. 26, no. 4, pp. 1894–1904, 2011.
- [20] F. Dorfler and F. Bullo, "Kron reduction of graphs with applications to electrical networks," *Circuits and Systems I: Regular Papers, IEEE Transactions on*, vol. 60, no. 1, pp. 150–163, Jan.
- [21] K. V. Mardia, J. T. Kent, and J. M. Bibby, *Multivariate analysis*, ser. Probability and mathematical statistics. London ; New York: Academic Press, 1979.
- [22] H. Akaike, "A new look at the statistical model identification," *Automatic Control, IEEE Transactions on*, vol. 19, no. 6, pp. 716–723, 1974.
- [23] K. P. Burnham and D. R. Anderson, "Multimodel inference: Understanding aic and bic in model selection," *Sociological Methods and Research*, vol. 33, pp. 261–304, 2004.
- [24] C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to information retrieval*. New York: Cambridge University Press, 2008.
- [25] K. R. Padiyar, *Power system dynamics : stability and control*. Singapore New York :: John Wiley, 1996.
- [26] R. Podmore, "Identification of coherent generators for dynamic equivalents," *Power Apparatus and Systems, IEEE Transactions on*, vol. PAS-97, no. 4, pp. 1344–1354, 1978.
- [27] S. Cole and R. Belmans, "Matdyn, a new matlab-based toolbox for power system dynamic simulation," *Power Systems, IEEE Transactions on*, vol. 26, no. 3, pp. 1129–1136, 2011.