CrossMark

Clustering approach to model order reduction of power networks with distributed controllers

Xiaodong Cheng[1] (ID) · **Jacquelien M. A. Scherpen[1]**

Abstract This paper considers the network structure preserving model reduction of power networks with distributed controllers. The studied system and controller are modeled as second-order and first-order ordinary differential equations, which are coupled to a closed-loop model for analyzing the dissimilarities of the power units. By transfer functions, we characterize the behavior of each node (generator or load) in the power network and define a novel notion of dissimilarity between two nodes by the $\mathcal{H}_2$-norm of the transfer function deviation. Then, the reduction methodology is developed based on separately clustering the generators and loads according to their behavior dissimilarities. The characteristic matrix of the resulting clustering is adopted for the Galerkin projection to derive explicit reduced-order power models and controllers. Finally, we illustrate the proposed method by the IEEE 30-bus system example.

Communicated by: Peter Benner

✉ Xiaodong Cheng
x.cheng@rug.nl

Jacquelien M. A. Scherpen
j.m.a.scherpen@rug.nl

[1] Jan C. Willems Center for Systems and Control, Faculty of Science and Engineering, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands

1 Introduction

Power networks, nowadays, are experiencing the penetration and integration of a wide array of new electronic devices and renewable energy sources (see [6, 16] for an overview). In the foreseeable future, power networks will become more and more complex with more variation in the generators, uncertain loads, and denser transmission lines. Centralized power generation is being replaced by a more distributed generation.

The immense size of power grids yields mathematical models with high dimensions complicating further analysis. In most cases, a complete model of the power network is neither practical nor necessary for, e.g., transient analysis, failure detection, distributed controller design, or system simulation. Therefore, we need to construct a reduced-order model that can approximate the behavior of the original complex power system with an acceptable accuracy. More importantly, we desire to preserve the network structure in the reduced-order model such that it can be interpreted as a reduced network with a less complex topology. Specifically, the reduced power system is evolving over a simpler network that consists of fewer generator and load buses and sparser transmission lines. The network structure is necessary for the application of sensor allocation and the management of distributed power generation. The main interest of this paper is to investigate the problem of model order reduction for power networks with distributed controllers with the preservation of network structures.

Literature review Conventional model reduction techniques, including balanced truncation and Krylov subspace methods, have been extended to the dynamic reduction of power systems (see, e.g., [10, 23, 24, 32, 42, 45, 49]). In these papers, the power network systems are modeled in first-order state space representations, and the reduced-order models are constructed within the framework of Petrov-Galerkin projection. Even though these methods provide systematic procedures to produce lower-dimensional models, which benefits the computation and simulation of power systems, the network structure is not retained through the reductions, and the projected states in the reduced-order model often lack a physical interpretation. A structure-preserving approach is proposed in [15], where the network model is firstly reduced by the balanced truncation, which still preserves the semistability such that the reduced-order model can be reformed as a network system. In spite of fewer nodes, the obtained reduced network has a complete interconnection. This restricts the use of the reduced-order models in further applications, especially the implementation of distributed control laws. In a distributed control network, control laws are implemented locally by each generator, and these controllers send and receive data through a communication network, which is only connecting generators and in general, that has a different topology from the power network (see [1, 21, 31, 34, 36, 47] for more details). In this case, each generator can only receive information from its adjacent generators through the communication network, which means that the control signals have to be generated by using only part of the system states. However, in the reduced-order model resulting from the conventional approaches, each state is composed of a linear combination of all the original states. As a result, the

controllers designed based on such models are not to be realized in a distributed fashion.

Another network simplification method is the Kron reduction, which has been widely used in e.g. classic circuit theory, smart grid monitoring, and transient stability assessment. The extension of this approach for model order reduction of electric networks can be found in [11, 19, 36] and in the references therein. The network structure of a power system is preserved as the Schur complement of the Laplacian matrix of the original network is again a Laplacian matrix, which represents a smaller network. Despite the simplicity of implementation, the Kron reduction modeling only eliminates the load buses in the power network [19]. To reduce the complexity of a network of synchronous generators, which is modeled by a second-order swing equation, [20, 41] propose a method based on time-scale separation and singular perturbation analysis. This approach identifies the sparsely and densely connected areas of power grids and then aggregates the state variables of the coherent areas. Similar to Kron reduction, the algebraic structure of the Laplacian matrix is maintained through the singular perturbation approximation. However, the method in [41] ignores the fluctuation of power demands from the loads.

Clustering analysis of power networks, derived from the concept of *diakoptics*, has been intensively explored in the study of coherency recognition, area aggregation, and behavior approximation in the power grid. A large and complex power network could be more easily analyzed and managed when it is decomposed into several smaller components. The coherency-based approach is most popular for the clustering analysis of electric networks. Many related results have been reported in the literature (see [3, 16, 17, 26, 33, 39, 48] and the references therein). The generator coherency is introduced as the tendency of generators when the disturbances are affecting the system. The reduced model of the power network then results from aggregating those coherent generators. However, the coherency identification is a data-driven process, which heavily relies on the accuracy of the sampling data generated from the time-domain voltage angle responses of the generator buses.

Differently, [43] proposes a Petrov-Galerkin approximation framework based on graph clustering, which is applied to physical network systems. However, the methods for cluster selection are still not well-explored. As an extension, [37] suggests to use the almost equitable partition (AEP) as a clustering of the underlying graph, but finding AEPs is generally a very difficult problem and computationally expensive, which limits this method in practical applications. The results in [9, 35] can be regarded as the combination of graph clustering and conventional model reduction techniques, balanced truncation, and Krylov subspace methods, respectively. However, it is still a challenge to apply these ideas to second-order systems, which are common model settings in the context of power networks. The approach in [28] offers another feasible solution for the clustering-based model reduction. The notion of cluster reducibility is generalized as the uncontrollability of the system states. Then, the reducible states are merged to reduce the dimension of the system as well as to preserve the network structure. To cope with second-order networks, the results are then extended in [27]. Nevertheless, the considered model is a second-order model, which is assumed to be asymptotically stable. This restricts this approach in some applications, such as the coupled swing dynamics in power networks [20].

Contribution The presented work provides a novel model order reduction scheme for power networks with the associated distributed controllers. The reduction procedure is developed based on our previous results in [12, 14], where the reduction of first-order network systems is studied. An extension to second-order models of networks is proposed in [13]. This paper explores the application of power networks that contain many generator and load buses, and the proposed method aims to *simultaneously* simplify the structures of the transmission network and its distributed controller.

By linearizing the nonlinear *structure preserving model* of the power network [7], we obtain the network model in a second-order representation. Together with the distributed controller, the closed-loop system is derived, which is semistable and contains the Laplacian matrices of the power transmission network and the communication graph of the controllers. The clustering analysis of the networks is based on the closed-loop system. Similar to the idea in [13], that a reduced-order model is generated by clustering the nodes behaving similarly. More specifically, the behaviors of network nodes are characterized by transfer functions, and an $\mathcal{H}_2$-norm characterization of the dissimilarity of nodal behaviors is adopted such that the hierarchical clustering algorithm in [13] can find suitable clusterings for generator and load nodes, respectively. This notion of dissimilarity is data-independent, which means that we do not need the time-domain sample data to evaluate the differences among the nodes. Compared with [13], the current paper mainly focuses on the application of power networks whose models couple the generator and load dynamics with the distributed controllers. The proposed method need to reduce the power network and communication network simultaneously, instead of only simplifying the network system as in [13]. The definition of nodal behaviors cannot be directly applied, since the influence of communication networks of the distributed controllers has to be concerned, and the behaviors of generator and load nodes are supposed to be treated differently.

As the dissimilarity is defined in terms of the $\mathcal{H}_2$-norm, we use the controllability Gramian of the closed-loop system to evaluate the dissimilarities between all pairs of generator and load buses. A numerical difficulty comes from the computation of the controllability Gramian. Due to the semistability, the corresponding algebraic Lyapunov equation may not have a unique solution [5]. This paper extends the idea of [14] to obtain a controllability Gramian by solving two algebraic equations rather than a single Lyapunov equation.

Besides, the hierarchical clustering algorithm is adapted to generate appropriate network clusterings for both generators and loads, so that the nodes with similar behaviors are grouped into the same clusters. Then, the Galerkin method is applied to yield a reduced-order power network with a simplified distributed controller. The characteristic matrices of the resulting clusterings are used as the projection matrices. It is shown that the algebraic structures of Laplacian matrices are retained in the reduced-order model. Consequently, the interconnection topologies of the power transmission network and communication network of the controller are simplified.

2 Models of power network and distributed controller

Consider a connected power network consisting of n synchronous generators and m loads. Denote $\mathcal{V}_g = \{1, \cdots, n\}$ and $\mathcal{V}_l = \{n + 1, \cdots, n + m\}$ as the index sets of generator and load buses, respectively. Then, the interconnection structure of the power grid can be represented by a connected undirected graph

$$\mathcal{G} = (\mathcal{V}_g \cup \mathcal{V}_l, \mathcal{E}), \text{ with } \mathcal{E} \subset (\mathcal{V}_g \cup \mathcal{V}_l) \times (\mathcal{V}_g \cup \mathcal{V}_l).$$

$\mathcal{E}$ is the set of unordered pairs (i, j) representing transmission lines between nodes i and j, which are assumed to be inductive. Notice that $|\mathcal{V}_g| = n$ and $|\mathcal{V}_l| = m$; therefore, $\mathcal{G}$ contains $n + m$ nodes.

The dynamics of generators and loads in a power network are characterized by nonlinear structure-preserving models [7, 18] as follows.

(1) Generator bus $i \in \mathcal{V}_g$:

$$\begin{aligned} \dot{\theta}_i &= \omega_i, \\ M_i \dot{\omega}_i &= -D_i \omega_i - \sum_{j=1}^{n+m} V_i V_j X_{ij}^{-1} \sin(\theta_i - \theta_j) + P_i^m; \end{aligned} \tag{1}$$

(2) Load bus $i \in \mathcal{V}_l$:

$$\begin{aligned} \dot{\theta}_i &= \omega_i \\ 0 &= -D_i \omega_i - \sum_{j=1}^{n+m} V_i V_j X_{ij}^{-1} \sin(\theta_i - \theta_j) - P_i^l. \end{aligned} \tag{2}$$

All the symbols in the models are described as follows.

- θ_i Voltage phase angle
- ω_i Voltage frequency w.r.t the nominal reference ω^* (typically 50 or 60 Hz)
- $M_i > 0$ Angular momentum of generator i
- $D_i > 0$ Damping coefficient
- $V_i > 0$ The voltage magnitude at node i
- $X_{ij} \geq 0$ The inductance of the transmission line connecting nodes i and j
- P_i^m Controllable power generation
- P_i^l Unknown power demand

The above structure-preserving power network models are commonly used in stability analysis and controller design of power networks, including the analysis of network synchronization and frequency regulation (see [7, 18, 47]). For example, [22] studies the full-order description of synchronous generators in a grid setting. However, controlling such model turns out to be rather complicated. Furthermore, it should be remarked that the models in (1) and (2) are simplifications for the real power systems based on the assumption that the influence of the windings of the generators is negligible and the power lines are assumed to be lossless. Therefore, only very slow electromechanical transients will be affected by the type of feedback that is proposed for use in this paper. For example, transients as described in [30] would not be applicable.

To obtain a linear model based on (1) and (2), we follow [36, 44] and assume that the differences of the phase angles are relatively small, i.e., $\theta_i - \theta_j \approx 0$ for any $i, j \in \mathcal{V}_g \cup \mathcal{V}_l$, which is satisfied in a vicinity of the nominal condition. This assumption then leads to a linearization of the power network model in the following compact second-order form

$$\mathbf{\Sigma}_p : \begin{bmatrix} M_g & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \ddot{\theta}_g \\ \ddot{\theta}_l \end{bmatrix} + \begin{bmatrix} D_g & 0 \\ 0 & D_l \end{bmatrix} \begin{bmatrix} \dot{\theta}_g \\ \dot{\theta}_l \end{bmatrix} + \begin{bmatrix} L_1 & L_{12} \\ L_{12}^T & L_2 \end{bmatrix} \begin{bmatrix} \theta_g \\ \theta_l \end{bmatrix} = \begin{bmatrix} P_m \\ -P_l \end{bmatrix}, \quad (3)$$

where $\theta_g \in \mathbb{R}^n$, $\theta_l \in \mathbb{R}^m$ are the state vectors collecting θ_i with $i \in \mathcal{V}_g$ and $i \in \mathcal{V}_l$, respectively. Matrices M_g, D_g, and D_l are diagonal and positive definite, which are defined as $M_g := \text{diag}(M_1, M_2, \cdots, M_n) \in \mathbb{R}^{n \times n}$, $D_g := \text{diag}(D_1, D_2, \cdots, D_n) \in \mathbb{R}^{n \times n}$, and $D_l := \text{diag}(D_{n+1}, D_{n+1}, \cdots, D_{n+m}) \in \mathbb{R}^{m \times m}$. Moreover, P_m and P_l are collections of P_i^m with $i \in \mathcal{V}_g$ and P_i^l with $i \in \mathcal{V}_l$, respectively. The weighted Laplacian matrix $\mathcal{L} \in \mathbb{R}^{(n+m) \times (n+m)}$ of the network topology $\mathcal{G}$ is partitioned as

$$\mathcal{L} = \begin{bmatrix} L_1 & L_{12} \\ L_{12}^T & L_2 \end{bmatrix}. \quad (4)$$

The (i, j)-th entry of $\mathcal{L}$ is given by

$$W_{ij} = W_{ji} := V_i V_j X_{ij}^{-1}, \quad (5)$$

which is interpreted as the maximum real power transfer between any two nodes i and j with constant voltage levels. The value of W_{ij} is positive if there is a physical cable directly connecting nodes i and j and 0 otherwise. $\mathcal{L}$ has a very special structure as stated in the following lemma [12, 40].

Lemma 1 *If $\mathcal{L}$ is a Laplacian matrix associated to a connected weighted undirected graph $\mathcal{G}$, then ①. The diagonal entries of $\mathcal{L}$ are strictly positive while the off-diagonal components are negative or 0. ② $\mathcal{L}^T = \mathcal{L} \geq 0$ and $\mathcal{L}$ has only one 0 eigenvalue. ③ $\mathbf{1}^T \mathcal{L} = 0$ and $\mathcal{L} \mathbf{1} = 0$.*

It is a crucial task to maintain the frequencies of generators and loads in the power network close to the nominal value. However, with the growth of the network size, centralized controllers become increasingly expensive due to the need for the information of all generators. Recently, distributed averaging PI controllers have been proposed for the frequency control of power networks (see [1, 21, 36] for the details). The distributed controllers exchange information over a communication network

$$\mathcal{G}_c := (\mathcal{V}_g, \mathcal{E}_c), \text{ with } \mathcal{E}_c \subset \mathcal{V}_g \times \mathcal{V}_g,$$

which is assumed to be undirected and connected. It should be emphasized that the topology of the communication network $\mathcal{V}_g$ is generally different from the topology of the power network $\mathcal{G}$. For each generator $i \in \mathcal{V}_g$, the controller takes the form

$$\begin{aligned} Q_i \dot{\xi}_i &= -\omega_i - \sum_{j \in \mathcal{N}_i} c_{ij}(z_i - z_j), \\ P_i^m &= \xi_i. \end{aligned} \quad (6)$$

Here, $\mathcal{N}_i$ represents the set of neighboring generators of node i in $\mathcal{G}_c$, i.e., the collection of generators that generator i communicates with. The coefficients $c_{ij} = c_{ji} > 0$

reflect the strengths of the connection between generators i and j. Q_i is a positive gain. The controller (6) is designed to regulate the frequency deviation to 0 and enforce the controller states ξ_i to reach a consensus, i.e., the generated power deviation of all generators becomes equal at steady state.

Notice that the controller (6) only need the information from the neighbors of generator i; therefore, it can be implemented in a distributed fashion. We now write the controller in the vector form

$$\Sigma_c : \begin{cases} Q\dot{\xi} = -\mathcal{L}_c\xi - \omega_g, \\ P_m = \xi, \end{cases} \quad (7)$$

where $\mathcal{L}_c$ is the Laplacian matrix of the communication network whose (i, j) entry is given by c_{ij}. The vector $\xi \in \mathbb{R}^n$ is the collection of ξ_i, $i \in \mathcal{V}_g$, and $Q := \mathrm{diag}(Q_1, Q_2, \cdots, Q_n)$. $\omega_g := \dot{\theta}_g \in \mathbb{R}^n$ represent the frequencies of all the generators.

By substituting the distributed controller (7) for P_m in (3), we obtain the closed-loop system

$$\Sigma_{cl} : E\dot{x} = Ax + Bd \quad (8)$$

where the state variable $x^T := \begin{bmatrix} \theta_g^T & \theta_l^T & \omega_g^T & \xi^T \end{bmatrix} \in \mathbb{R}^{3n+m}$ and

$$E = \begin{bmatrix} D_g & 0 & M_g & 0 \\ 0 & D_l & 0 & 0 \\ M_g & 0 & 0 & 0 \\ 0 & 0 & 0 & Q \end{bmatrix}, \quad A = \begin{bmatrix} -L_1 & -L_{12} & 0 & I_n \\ -L_{12}^T & -L_2 & 0 & 0 \\ 0 & 0 & M_g & 0 \\ 0 & 0 & -I_n & -\mathcal{L}_c \end{bmatrix}, \quad \text{and } B = \begin{bmatrix} 0 \\ -I_m \\ 0 \\ 0 \end{bmatrix}. \quad (9)$$

Here, $d := P_l$ is the power consumption of the m loads, which is uncontrollable and regarded as the stochastic disturbance of the closed-loop system (8). We analyze the stability of Σ_{cl} in the following theorem.

Theorem 1 *The closed-loop power system Σ_{cl} is semistable, and the trajectories of its impulse response converge to 0.*

Proof Observe that $\det(E) = \det(D_g)\det(D_l)\det(D_g^{-1}M_g) \neq 0$. Therefore, E is nonsingular, and its inverse reads as

$$E^{-1} = \begin{bmatrix} 0 & 0 & M_g^{-1} & 0 \\ 0 & D_l^{-1} & 0 & 0 \\ M_g^{-1} & 0 & -M_g^{-1}D_gM_g^{-1} & 0 \\ 0 & 0 & 0 & Q^{-1} \end{bmatrix}. \quad (10)$$

which gives

$$E^{-1}A = \begin{bmatrix} 0 & 0 & I_n & 0 \\ -D_l^{-1}L_{12}^T & -D_l^{-1}L_2 & 0 & 0 \\ -M_g^{-1}L_1 & -M_g^{-1}L_{12} & -M_g^{-1}D_g & M_g^{-1} \\ 0 & 0 & -Q^{-1} & -Q^{-1}\mathcal{L}_c \end{bmatrix}. \quad (11)$$

By [36], interconnecting the controller Σ_c and the power network system Σ_p results in a zero frequency deviation and consensus of the controller states. It then implies

that, with impulse signals as disturbances of the closed-loop system Σ_{cl}, all the state trajectories converge to constant values, i.e., $\lim_{t\to\infty} e^{E^{-1}At}$ exists. Thus, Σ_{cl} is semistable by the definition in [8].

Furthermore, using the third property of Laplacian matrices in Lemma 1, it is easy to check that

$$\ker(E^{-1}A) = \ker(A) = \mathrm{span}\left(\begin{bmatrix} \mathbf{1}_n^T & \mathbf{1}_m^T & 0_n^T & 0_n^T \end{bmatrix}^T \right). \tag{12}$$

Together with semistability, we conclude that $E^{-1}A$ only has one 0 eigenvalue at the origin, and all the other eigenvalues have strict negative real parts. Notice that the vectors

$$v_R \in \mathrm{span}\left(\begin{bmatrix} \mathbf{1}_n \\ \mathbf{1}_m \\ 0_n \\ 0_n \end{bmatrix} \right) \quad \text{and} \quad v_L \in \mathrm{span}\left(\begin{bmatrix} \mathbf{1}_n \\ 0_m \\ 0_n \\ Q\mathbf{1}_n \end{bmatrix} \right) \tag{13}$$

are the right and left eigenvectors of $E^{-1}A$ corresponding to the 0 eigenvalue, respectively, i.e., $E^{-1}Av_R = 0$ and $v_L^T E^{-1}A = 0$. Based on this, we can obtain the following decomposition

$$E^{-1}A = \mathcal{U}D\mathcal{U}^{-1} = [v_R, V_R] \begin{bmatrix} 0 & \\ & \bar{D} \end{bmatrix} \begin{bmatrix} v_L^T \\ V_L^T \end{bmatrix}, \tag{14}$$

where $\mathcal{U}$ is unitary, and $\bar{D} \in \mathbb{R}^{3n+m-1}$ is Hurwitz. Hence, $\|v_R\|_2 = 1$ and $v_L^T v_R = 1$, which yields

$$v_R = \frac{1}{\sqrt{n+m}} \begin{bmatrix} \mathbf{1}_n \\ \mathbf{1}_m \\ 0_m \\ 0_n \end{bmatrix} \quad \text{and} \quad v_L = \frac{\sqrt{n+m}}{n} \begin{bmatrix} \mathbf{1}_n \\ 0_m \\ 0_n \\ Q\mathbf{1}_n \end{bmatrix}, \tag{15}$$

Then, the convergence value of the impulse response is computed as

$$\begin{aligned}
\lim_{t\to\infty} \left(e^{E^{-1}At} \right) E^{-1}B &= v_R v_L^T E^{-1}B \\
&= \frac{1}{n} \begin{bmatrix} \mathbf{1}_n\mathbf{1}_n^T & 0_{n\times m} & 0_{n\times n} & \mathbf{1}_n\mathbf{1}_n^T Q \\ \mathbf{1}_m\mathbf{1}_n^T & 0_{m\times m} & 0_{m\times n} & \mathbf{1}_m\mathbf{1}_n^T Q \\ 0_{n\times n} & 0_{n\times m} & 0_{n\times n} & 0_{n\times n} \\ 0_{n\times n} & 0_{n\times m} & 0_{n\times n} & 0_{n\times n} \end{bmatrix} \begin{bmatrix} 0_n \\ D_l^{-1} \\ 0_n \\ 0_n \end{bmatrix} = 0.
\end{aligned} \tag{16}$$

That completes the proof.   □

Remark 1 Due to the singularity of matrix $E^{-1}A$, the closed-loop system Σ_{cl} is not asymptotically stable. However, Theorem 1 indicates that, for any initial condition, the unforced system responses can reach the steady states at 0. Theorem 1 also offers a physical interpretation for the power network. The distributed controller in (7) eliminates the effects of disturbances, namely a sudden impulse change in the power demand, and steers the frequencies of all the nodes (i.e., the generator and load buses) to the nominal value ω^*.

3 Clustering of power networks

To uncover the community structure of generators and loads, a novel clustering method is proposed and applied in this section, where the clusters are constructed based on the differences of node behaviors.

3.1 Concepts of clustering

As the physical components of a power network can be interpreted by an undirected weighted graph [38], the techniques of graph analysis help us achieve a better understanding of the essential properties of the power system. Especially, the graph clustering can be applied to identify the communities of generators and loads, which group the units with similar behaviors. In this subsection, we first recapitulate the notion of network clustering (or graph partition) from, e.g., [12, 13].

Definition 1 Consider a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{1, 2, \cdots, k\}$ is the index set of nodes. A nonempty index subset of $\mathcal{V}$, denoted by $\mathcal{C}$, is called a cluster of graph $\mathcal{G}$. Then, network clustering is to partition $\mathcal{V}$ into r $(r \leq k)$ disjoint clusters which cover all the elements in $\mathcal{V}$.

A network clustering is denoted by $\{\mathcal{C}_1, \mathcal{C}_2, \cdots, \mathcal{C}_r\}$, where $\mathcal{C}_i \subseteq \mathcal{V}$, and $\mathcal{C}_i \cap \mathcal{C}_j = \emptyset, \forall i, j \in \{1, 2, \cdots, r\}$. Mathematically, the clustering of a graph $\mathcal{G}$ can be represented by a binary characteristic matrix, which is defined as follows.

Definition 2 Let $|\mathcal{V}| = k$ and $\{\mathcal{C}_1, \mathcal{C}_2, \cdots, \mathcal{C}_r\}$ be the clustering of $\mathcal{G}$ with node set $\mathcal{V}$. The characteristic vector of the cluster $\mathcal{C}_i$ is defined by binary vector $\pi(\mathcal{C}_i) \in \mathbb{R}^k$ where $\mathbf{1}_k^T \pi(\mathcal{C}_i) = |\mathcal{C}_i|$, and the jth element of $\pi(\mathcal{C}_i)$ is 1 when $j \in \mathcal{C}_i$ and 0 otherwise. Then, the characteristic matrix of the clustering is a binary matrix denoted by

$$\Pi := [\pi(\mathcal{C}_1), \pi(\mathcal{C}_2), \cdots, \pi(\mathcal{C}_r)] \in \mathbb{R}^{k \times r}. \tag{17}$$

In the following subsections, we propose a new method for cluster selection in power networks. The choices of clustering will determine the accuracy of network approximation. i.e., how close the behaviors of the reduced and original networks are. Hence, the cluster selection algorithm is the most crucial part of the clustering-based model reduction. This paper provides a new way of cluster selection, which involves a particular notion of dissimilarity and an adaption of the hierarchical clustering algorithm. The details of our method are described hereafter.

3.2 Characterization and computation of dissimilarity

In the context of clustering, node dissimilarities intuitively describe how different the nodes are from each other. In contrast with the existing results in the literature, such as [16, 29, 39], the clusters of a power network in this paper are identified without sampling data from a real system. We calculate the *dissimilarities* of network

nodes by the $\mathcal{H}_2$-norms of transfer function discrepancies, rather than the Euclidean distances of the relative positions. Then, potentially, we can place the nodes with similar behaviors into the same clusters.

More precisely, by considering the state variables of the closed-loop power system $\mathbf{\Sigma}_{cl}$ in (8), the behavior of generator i ($i \in \mathcal{V}_g$) is characterized by its responses of the voltage angles θ_i, the frequencies ω_i, and the controller states ξ_i. We include ξ_i into behavior characterization of generators since we also approximate the controller $\mathbf{\Sigma}_c$. When the generators are clustered, the communication network of the distributed controllers is automatically simplified. Hence, the effects of the controller states ξ_i are also needed to be included. As for the behavior of load i ($i \in \mathcal{V}_l$), it can be simply represented by its voltage angle responses θ_i with respect to the stochastic disturbances d.

The behaviors of generators and loads can be expressed in the complex frequency domain by the transfer functions from d to the indicative state variables. Denote two new binary vectors

$$\boldsymbol{\alpha}_i = \mathbf{e}_i + \mathbf{e}_{i+n+m} + \mathbf{e}_{i+2n+m}, \ \boldsymbol{\beta}_i = \mathbf{e}_{i+n}, \tag{18}$$

where $\mathbf{e}_i$ is the ith unit base vector. Then, the behavior of generator i is given by

$$\Psi_i^g(s) = \frac{\left[\theta_i(s) \ \omega_i(s) \ \xi_i(s) \right]^T}{d(s)} = \boldsymbol{\alpha}_i^T (sI - \mathcal{A})\mathcal{B}, \ i = \{1, 2, \cdots, n\}, \tag{19}$$

where $\theta_i(s)$, $\omega_i(s)$, $\xi_i(s)$, and $d(s)$ are the Laplace transforms of the states θ_i, ω_i, and ξ_i, and the input d, respectively, in the closed-loop system. Furthermore, the behavior of load $i \in \mathcal{V}_l$ is represented by

$$\Psi_i^l(s) = \frac{\theta_{i+n}(s)}{d(s)} = \boldsymbol{\beta}_i^T (sI - \mathcal{A})\mathcal{B}, \ i = \{1, 2, \cdots, m\}. \tag{20}$$

Thereupon, a pairwise dissimilarity of nodes i and j is defined for generators and loads, respectively.

$$\mathcal{D}_{ij}^g = \|\Psi_i^g(s) - \Psi_j^g(s)\|_{\mathcal{H}_2}, \ \mathcal{D}_{ij}^l = \|\Psi_i^l(s) - \Psi_j^l(s)\|_{\mathcal{H}_2}. \tag{21}$$

Moreover, the *dissimilarity matrices* for the generators and loads are denoted by $\mathcal{D}^g \in \mathbb{R}^{n \times n}$ and $\mathcal{D}^l \in \mathbb{R}^{m \times m}$, which are constructed by collecting $\mathcal{D}_{ij}^g$ and $\mathcal{D}_{ij}^l$. Clearly, $\mathcal{D}^g$ and $\mathcal{D}^l$ are symmetric matrices with nonnegative entries and zero diagonal elements. Besides, the boundedness of $\mathcal{D}^g$ and $\mathcal{D}^l$ is also guaranteed by Theorem 1. The reasons of the boundedness are explained as follows.

For simplicity, let

$$\mathcal{A} = E^{-1}A, \ \mathcal{B} = E^{-1}B, \tag{22}$$

where E, A, and B are system coefficients of the closed-loop system $\mathbf{\Sigma}_{cl}$ in (8). Note that impulse response of node i in $\mathbf{\Sigma}_{cl}$ is expressed by $e^{\mathcal{A}t}\mathcal{B}$, which is a bounded smooth function of t and exponentially converges to 0. Thus, the boundedness of $\mathcal{D}_{ij}^g$ and $\mathcal{D}_{ij}^l$ can be seen from the the definition of $\mathcal{H}_2$-norm [2].

$$\begin{aligned}
\|\Psi_i^g(s) - \Psi_j^g(s)\|_{\mathcal{H}_2}^2 &= \int_0^\infty (\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_j)^T e^{\mathcal{A}t} \mathcal{B}\mathcal{B}^T e^{\mathcal{A}^T t} (\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_j) dt \\
&= (\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_j)^T \mathcal{P}(\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_j),
\end{aligned} \tag{23}$$

and similarly,

$$\|\Psi_i^l(s) - \Psi_j^l(s)\|_{\mathcal{H}_2}^2 = (\boldsymbol{\beta}_i - \boldsymbol{\beta}_j)^T \mathcal{P}(\boldsymbol{\beta}_i - \boldsymbol{\beta}_j). \tag{24}$$

In both equations, $\mathcal{P}$ is the controllability Gramian of the closed-loop system $\boldsymbol{\Sigma}_{cl}$. For large-scale networks, (23) and (24) provide an efficient method to compute matrix $\mathcal{D}$, since we first calculate $\mathcal{P}$ and then just apply vector-matrix multiplication to obtain all the entries of $\mathcal{D}$. However, the semistability of $\boldsymbol{\Sigma}_{cl}$ poses a challenge to implement this idea. For asymptotically stable systems, their controllability Gramians can be uniquely determined by solving the associated continuous-time algebraic Lyapunov equation

$$\mathcal{AP} + \mathcal{P}\mathcal{A}^T + \mathcal{B}\mathcal{B}^T = 0. \tag{25}$$

In the semistable case, $\mathcal{A}$ is not Hurwitz, which results in multiple solutions of the above Lyapunov equation [5]. To characterize the controllability Gramian from the solutions of (25), the following lemma is proven.

Lemma 2 *The controllability Gramian $\mathcal{P}$ of the semistable power system $\boldsymbol{\Sigma}_{cl}$ is positive semidefinite, and it is uniquely determined by the combination of (25) and*

$$v_L^T \mathcal{P} v_L = 0, \tag{26}$$

where v_L is the left eigenvector of $\mathcal{A}$ given in (15).

Proof Since $v_L^T \mathcal{A} = 0$, for any t, we have

$$
\begin{aligned}
v_L^T e^{\mathcal{A}t} \mathcal{B} &= v_L^T \left(I + \mathcal{A}t + \frac{\mathcal{A}^2}{2}t^2 + \cdots \right) \mathcal{B} \\
&= v_L^T \mathcal{B} = \frac{\sqrt{n+m}}{n} \begin{bmatrix} \mathbf{1}_n^T & \mathbf{0}_m^T & \mathbf{0}_n^T & \mathbf{1}_n^T Q \end{bmatrix} \begin{bmatrix} 0_n \\ D_l^{-1} \\ 0_n \\ 0_n \end{bmatrix} = 0.
\end{aligned}
\tag{27}
$$

Thus, (26) holds. It then yields $v_L^T \mathcal{P} v_L = 0$, which means $\mathcal{P}$ is positive semidefinite.

Next, we show that solving the algebraic Lyapunov equation (25) together with (26) will yield a unique solution. A contradiction proof is provided. Assume that two distinct symmetric matrices $\mathcal{P}_1$ and $\mathcal{P}_2$ are both the solutions of (25) and (26). From (25), we have

$$\mathcal{A}(\mathcal{P}_1 - \mathcal{P}_2) + (\mathcal{P}_1 - \mathcal{P}_2)\mathcal{A}^T = 0, \tag{28}$$

which leads to

$$e^{\mathcal{A}t} \left[\mathcal{A}(\mathcal{P}_1 - \mathcal{P}_2) + (\mathcal{P}_1 - \mathcal{P}_2)\mathcal{A}^T \right] e^{\mathcal{A}^T t} = \frac{d}{dt} \left[e^{\mathcal{A}t}(\mathcal{P}_1 - \mathcal{P}_2)e^{\mathcal{A}^T t} \right] = 0. \tag{29}$$

Therefore,

$$\int_0^\infty \frac{d}{dt} \left[e^{\mathcal{A}t}(\mathcal{P}_1 - \mathcal{P}_2)e^{\mathcal{A}^T t} \right] dt = 0. \tag{30}$$

Note that $\lim_{t \to \infty} e^{\mathcal{A}t} = v_R v_L^T$. Then, we obtain from (30) that

$$\mathcal{P}_1 - \mathcal{P}_2 = v_R v_L^T (\mathcal{P}_1 - \mathcal{P}_2) v_L v_R^T, \tag{31}$$

which is equal to 0 due to (26). Thus, it contradicts the fact that $\mathcal{P}_1 \neq \mathcal{P}_2$. As a result, the common solution of (25) and (26) is unique. □

In the following theorem, we then provide a method to determine the controllability Gramian of the semistable system $\mathbf{\Sigma}_{cl}$ by solving linear matrix equations.

Theorem 2 *Let $\mathcal{P}_a$ be an arbitrary solution that fulfills (25) and (26) simultaneously. Then, the controllability Gramian $\mathcal{P}$ of the closed-loop power system $\mathbf{\Sigma}_{cl}$ in (8) is computed as*

$$\mathcal{P} = \mathcal{P}_a - \mathcal{J}\mathcal{P}_a\mathcal{J}^T, \tag{32}$$

where $\mathcal{J}$ is a constant matrix defined by

$$\mathcal{J} := \lim_{t \to \infty} e^{\mathcal{A}t} = \frac{1}{n} \begin{bmatrix} \mathbf{1}_n\mathbf{1}_n^T & 0_{n \times m} & 0_{n \times n} & \mathbf{1}_n\mathbf{1}_n^T Q \\ \mathbf{1}_m\mathbf{1}_n^T & 0_{m \times m} & 0_{m \times n} & \mathbf{1}_m\mathbf{1}_n^T Q \\ 0_{n \times n} & 0_{n \times m} & 0_{n \times n} & 0_{n \times n} \\ 0_{n \times n} & 0_{n \times m} & 0_{n \times n} & 0_{n \times n} \end{bmatrix} \in \mathbb{R}^{(3n+m) \times (3n+m)}. \tag{33}$$

Proof Equation (27) implies that $\mathcal{J}\mathcal{P}\mathcal{J}^T = 0$. From (31) in the proof of Lemma 2, we have

$$\mathcal{P}_a - \mathcal{P} = \mathcal{J}(\mathcal{P}_a - \mathcal{P})\mathcal{J}^T = \mathcal{J}\mathcal{P}_a\mathcal{J}^T, \tag{34}$$

which leads to (32). □

It should be remarked that $\mathcal{P}$ in (32) does not depend on the choice of $\mathcal{P}_a$, which can be any symmetric real solution of (25). Besides, we can modify the Hessenberg-Schur method (SB04MD (SLICOT)) in [25] to solve the Lyapunov equation with a nonsingular $\mathcal{A}$ matrix. The output of the adapted Hessenberg-Schur algorithm will give one solution of (25), and we then use the relation in (32) to obtain the real value of $\mathcal{P}$. Hereafter, we apply (23) and (24) to determine all the dissimilarities of different pairs of generators or loads in the power grid.

3.3 Hierarchical clustering of generator and load buses

The idea of hierarchical clustering has been extensively explored in many fields, including pattern recognition, data compression, and network science (see [4, 29]). Hierarchical clustering, in principle, is a greedy algorithm, whose running time grows polynomially with the size of the studied networks. Compared to the K-means algorithm, another popular clustering method, the result of hierarchical clustering is not affected by the initializations of the graph partitions. Due to the above qualities, we choose and adapt the hierarchical algorithm to cluster the generator and load buses for the purpose of model order reduction.

In this paper, the generators and loads are grouped separately on account of different evaluation criteria of dissimilarities (see Section 3.2), yet the same hierarchical clustering procedures can be applied for both. It is because the hierarchical clustering only needs the dissimilarity matrices $\mathcal{D}^g$ and $\mathcal{D}^l$, whose entries already contain the dissimilarities of all pairs of generator and load nodes. Hereby, we only analyze the

clustering of generators and the clustering of loads in parallel. Both procedures go in the same way; hence, we only present it for any $\mathcal{D} \in \{\mathcal{D}^g, \mathcal{D}^l\}$.

As nodes are merged into small clusters, we must measure how similar two clusters are. In this paper, we consider the *average cluster dissimilarity*. It characterizes the dissimilarity of two clusters as the average of $\mathcal{D}_{ij}$ over all node pairs i and j that belong to distinct clusters [4]. More precisely, the dissimilarity of clusters $\mathcal{C}_\mu$ and $\mathcal{C}_\nu$ is defined as follows.

$$\Gamma(\mathcal{C}_\mu, \mathcal{C}_\nu) = \frac{1}{|\mathcal{C}_\mu| \cdot |\mathcal{C}_\nu|} \sum_{i \in \mathcal{C}_\mu} \sum_{j \in \mathcal{C}_\nu} \mathcal{D}_{ij}. \tag{35}$$

The value of $\Gamma(\mathcal{C}_\mu, \mathcal{C}_\nu)$ is computed only using the entries of $\mathcal{D}$ that correspond to the nodes in $\mathcal{C}_\mu$ and $\mathcal{C}_\nu$. Hereupon, the clustering approach links the pairs of nodes that are in close proximity and place them into binary clusters. Then, the newly formed clusters are merged into larger clusters according to the cluster dissimilarity. In Algorithm 1, the pseudocode of hierarchical clustering algorithm is presented. The inputs are the dissimilarity matrix $\mathcal{D}$ and the desired number of clusters r. The characteristic matrix Π of the clustering is considered as the output of the algorithm.

Algorithm 1 Hierarchical clustering

1: Compute the dissimilarity matrix $\mathcal{D}$ by (23) or (24)
2: Place each node into its own singleton cluster, that is $\mathcal{C}_i \leftarrow \{i\}$
3: $k \leftarrow n$
4: **while** $k > r$ **do**
5:     Set Γ_m to be an arbitrary large number
6:     **for** $i = 1 : k - 1$ and $j = 2 : i - 1$ **do**
7:         Compute $\Gamma(\mathcal{C}_i, \mathcal{C}_j)$ by (35)
8:         **if** $\Gamma_m > \Gamma(\mathcal{C}_i, \mathcal{C}_j)$ **then**
9:             $\mu \leftarrow i, \nu \leftarrow j, \Gamma_m \leftarrow \Gamma(\mathcal{C}_i, \mathcal{C}_j)$
10:         **end if**
11:     **end for**
12:     Merge cluster μ and ν into a single cluster
13:     $k \leftarrow k - 1$
14: **end while**
15: Compute $\Pi \in \mathbb{R}^{n \times r}$

In principle, hierarchical clustering does not require preliminary knowledge about the number and the size of clusters. In practice, it generates a *dendrogram* that organizes a hierarchy of clusters in a tree structure. Any cut of the hierarchical tree offers a potentially valid clustering of the studied network.

It is also worth mentioning that he formation of clusters in Algorithm 1 does not intend to only cluster the adjacent nodes. Even if two nodes are not neighbors, their dissimilarity still can be computed and thus their difference of behaviors is measured, and two nonadjacent vertices can be clustered when they have relatively similar behaviors, as the obtained aggregated network will achieve a small approximation

error (see some numerical examples in e.g. [12, 13]). However, our method does not exclude the possibility of modifying the algorithm to only aggregate adjacent nodes. To this end, we need to update the computation of dissimilarity matrix, i.e., we add a penalty factor if two nodes are not neighbors.

3.4 Reduced model of power network

In Section 3.2, the dissimilarities are calculated in the closed-loop systems, and the generator and load buses are grouped separately by the hierarchical clustering in Section 3.3. Then, the controllers are automatically clustered as each of them is associating with a generator.

Suppose that we acquire r clusters of generator buses and q clusters of load buses. Correspondingly, the characteristic matrices are denoted by $\Pi_g \in \mathbb{R}^{n \times r}$ and $\Pi_l \in \mathbb{R}^{m \times q}$, respectively. Let

$$\Pi_p = \begin{bmatrix} \Pi_g & 0 \\ 0 & \Pi_l \end{bmatrix}. \tag{36}$$

Then, a simplified weighted power network $\hat{\mathcal{G}}$ can be obtained by aggregating all the nodes with the same clusters in the original complex network. Mathematically, the reduced Laplacian matrix of $\hat{\mathcal{G}}$ is generated by the following projection.

$$\hat{\mathcal{L}} = \Pi_p^T \mathcal{L} \Pi_p = \begin{bmatrix} \Pi_g^T L_1 \Pi_g & \Pi_g^T L_{12} \Pi_g \\ \Pi_l^T L_{12}^T \Pi_l & \Pi_l^T L_2 \Pi_l \end{bmatrix} := \begin{bmatrix} \hat{L}_{11} & \hat{L}_{12} \\ \hat{L}_{12}^T & \hat{L}_{22} \end{bmatrix} \in \mathbb{R}^{(r+q) \times (r+q)}, \tag{37}$$

where $\mathcal{L}$ is the weighted Laplacian matrix representing the original network. Similarly, the inertia and damping coefficients in the reduced network are given by

$$\hat{M}_g = \Pi_g^T M_g \Pi_g, \ \hat{D}_g = \Pi_g^T D_g \Pi_g \in \mathbb{R}^{r \times r}, \text{ and } \hat{D}_l = \Pi_l^T D_l \Pi_l \in \mathbb{R}^{q \times q}. \tag{38}$$

This clustering-based projection also has a physical interpretation, as shown in Fig. 1, which shows the aggregation of generator and load buses, respectively. Notice that the nodes in the same cluster can be aggregated even if they do not have a direct connection with each other.

The transmission lines contained in a cluster are neglected, while those connecting two clusters are aggregated. Moreover, the maximum real power transfer on the new branch between the clusters $\mathcal{C}_\mu$ and $\mathcal{C}_\nu$ is given by

$$\hat{W}_{\mu,\nu} = \sum_{i \in \mathcal{C}_\mu} \sum_{j \in \mathcal{C}_\nu} W_{ij}, \tag{39}$$

where W_{ij} is the maximum real power transfer between nodes i and j. In the reduced network, a node represents a cluster of nodes in the original network, and its angular momentum and damping coefficient are the sum of these parameters of generators and loads in the original power network. Specifically, the inertia and damping coefficients of the cluster $\mathcal{C}_i$ is computed as

$$\hat{M}_\mu = \sum_{i \in \mathcal{C}_\mu} M_i \text{ and } \hat{D}_\mu = \sum_{i \in \mathcal{C}_\mu} D_i. \tag{40}$$
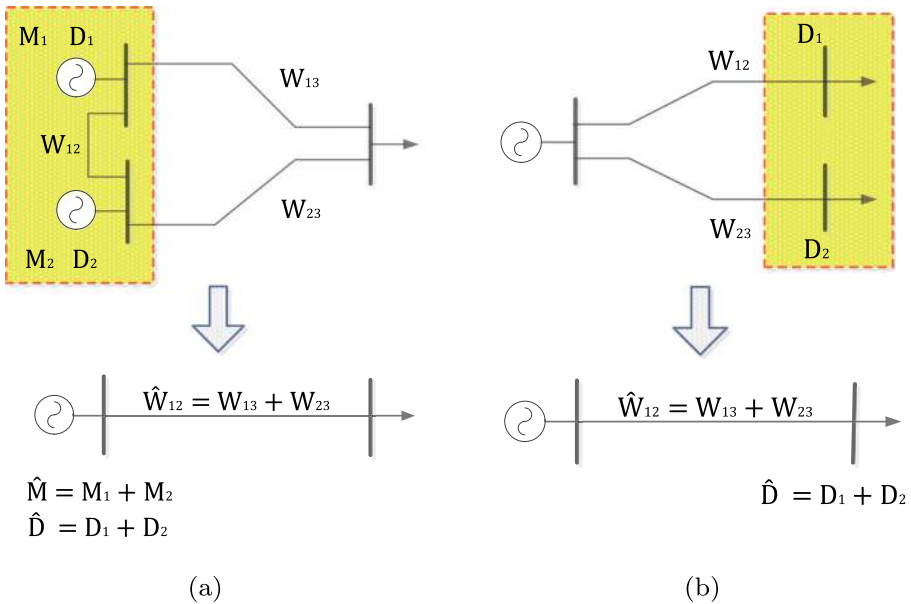
Fig. 1 Illustration of aggregation of generator buses (**a**) and load buses (**b**). The nodes in the same cluster are included in the dashed boxes

Now, we are ready to present the reduced-order power system of Σ_p in (3), which is obtained by Galerkin projection.

$$
\hat{\Sigma}_p : \begin{bmatrix} \hat{M}_g & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \ddot{\hat{\theta}}_g \\ \ddot{\hat{\theta}}_l \end{bmatrix} + \begin{bmatrix} \hat{D}_g & 0 \\ 0 & \hat{D}_l \end{bmatrix} \begin{bmatrix} \dot{\hat{\theta}}_g \\ \dot{\hat{\theta}}_l \end{bmatrix} + \begin{bmatrix} \hat{L}_1 & \hat{L}_{12} \\ \hat{L}_{12}^T & \hat{L}_2 \end{bmatrix} \begin{bmatrix} \hat{\theta}_g \\ \hat{\theta}_l \end{bmatrix} = \begin{bmatrix} \Pi_g^T P_m \\ -\Pi_l^T P_l \end{bmatrix},
$$
(41)

where the Laplacian matrix and $\hat{M}_g$, $\hat{D}_g$, and $\hat{D}_l$ are defined in (37) and (38), respectively. $\hat{\theta}_g \in \mathbb{R}^r$ and $\hat{\theta}_l \in \mathbb{R}^q$ are the voltage phase angles of the aggregated generators and loads in the reduced-order power network, which are used to approximate the states in the original system:

$$
\theta_g \approx \Pi_g \hat{\theta}_g, \quad \omega_g \approx \Pi_g \hat{\omega}_g = \Pi_g \dot{\hat{\theta}}_g, \quad \text{and } \theta_l \approx \Pi_l \hat{\theta}_l.
$$
(42)

Based on (41), the reduced-order nonlinear power system can be also constructed, which follows the same form as (1) or (2). Here, the reduced-order nonlinear model is omitted.

As the generators are clustered, the size of the communication network $\mathcal{G}_c$ for the distributed controller in (7) is also simplified. As a result, the dimension of the controller is reduced simultaneously. Analogously, the representation of the lower-dimensional controller is written as

$$
\hat{\Sigma}_c : \begin{cases} \hat{Q}\dot{\hat{\xi}} = -\hat{\mathcal{L}}_c \hat{\xi} - \Pi_g^T \Pi_g \hat{\omega}_g, \\ P_m = \Pi_g \hat{\xi}, \end{cases}
$$
(43)

where $\hat{Q} := \Pi_g^T Q \Pi_g$ and $\hat{\xi} \in \mathbb{R}^r$ are states of the reduced-order controller. $\hat{\mathcal{L}}_c :=$ $\Pi_g^T \mathcal{L}_c \Pi_g$ represents the simplified communication network.

Combining the reduced versions of the power network $\hat{\boldsymbol{\Sigma}}_p$ and the distributed controller $\hat{\boldsymbol{\Sigma}}_c$, we derive the reduced-order closed-loop system as follows.

$$\boldsymbol{\Sigma}_{cl} : \ \hat{E}\dot{\hat{x}} = \hat{A}\hat{x} + \hat{B}d, \tag{44}$$

where $\hat{x}^T := \begin{bmatrix} \hat{\theta}_g^T & \hat{\theta}_l^T & \hat{\omega}_g^T & \hat{\xi} \end{bmatrix} \in \mathbb{R}^{3r+q}$ and

$$\hat{E} = \begin{bmatrix} \hat{D}_g & 0 & \hat{M}_g & 0 \\ 0 & \hat{D}_l & 0 & 0 \\ \hat{M}_g & 0 & 0 & 0 \\ 0 & 0 & 0 & \hat{Q} \end{bmatrix}, \ \hat{A} = \begin{bmatrix} -\hat{L}_1 & -\hat{L}_{12} & 0 & \Pi_g^T \Pi_g \\ -\hat{L}_{12}^T & -\hat{L}_2 & 0 & 0 \\ 0 & 0 & \hat{M}_g & 0 \\ 0 & 0 & -\Pi_g^T \Pi_g & -\hat{\mathcal{L}}_c \end{bmatrix}, \text{ and } \hat{B} = \begin{bmatrix} 0 \\ -\Pi_l^T \\ 0 \\ 0 \end{bmatrix}. \tag{45}$$

Here, $d := P_l$ is the uncontrollable power demand in (8). The Galerkin projection matrix for the closed-loop system $\boldsymbol{\Sigma}_{cl}$ is

$$\Pi = \begin{bmatrix} \Pi_g & 0 & 0 & 0 \\ 0 & \Pi_l & 0 & 0 \\ 0 & 0 & \Pi_g & 0 \\ 0 & 0 & 0 & \Pi_g \end{bmatrix}, \tag{46}$$

which satisfies

$$\hat{E} = \Pi^T E \Pi, \ \hat{A} = \Pi^T A \Pi, \text{ and } \hat{B} = \Pi^T B. \tag{47}$$

Next, the properties of the reduced-order models are discussed in the following theorem.

Theorem 3 *The reduced power network $\boldsymbol{\Sigma}_p$ in (41) and the distributed controller $\boldsymbol{\Sigma}_c$ in (43) preserve the network structures. The closed-loop system $\boldsymbol{\Sigma}_{cl}$ in (44) is also semistable, and the trajectories of its impulse response converge to 0.*

Proof In [13], we have proven that the Galerkin projection based on the characteristic matrix of network clustering which can preserve the algebraic structure of a Laplacian matrix, which means $\hat{\mathcal{L}}$ in (37) and $\hat{\mathcal{L}}_c$ in (43) are the reduced Laplacian matrices representing the simplified power network and communication links. Besides, (40) shows that $\hat{M}_g$, $\hat{D}_g$, and $\hat{D}_l$ are diagonal positive definite. Therefore, the reduced models in (41) and (43) have the same structures as (8) and (7), respectively.

Following the same reasoning line of Theorem 1, the closed-loop system $\boldsymbol{\Sigma}_{cl}$ is semistable. It has only one pole at the origin, and all the other poles are located in the open-left half plane. The vectors

$$\hat{v}_L = \frac{\sqrt{r+q}}{n} \begin{bmatrix} \Pi_g^T \Pi_g \mathbf{1}_r \\ 0_q \\ 0_r \\ \hat{Q}\mathbf{1}_r \end{bmatrix} \text{ and } \hat{v}_R = \frac{1}{\sqrt{r+q}} \begin{bmatrix} \mathbf{1}_r \\ \mathbf{1}_q \\ 0_r \\ 0_r \end{bmatrix} \tag{48}$$

are the left and right eigenvectors of $\hat{E}^{-1}\hat{A}$ corresponding to the only zero eigenvalue, which can be verified by $\hat{v}_L^T \hat{E}^{-1}\hat{A} = 0$, and $\hat{E}^{-1}\hat{A}\hat{v}_R = 0$. Note that the characteristic matrix of the graph clustering has the property $\Pi_g \mathbf{1}_r = \mathbf{1}_n$. Hence, $\|\hat{v}_R\|_2 = 1$, and $\hat{v}_L^T \hat{v}_R = 1$. We further obtain

$$
\lim_{t \to \infty} \left(e^{\hat{E}^{-1}\hat{A}t} \right) \hat{E}^{-1}\hat{B} = \hat{v}_R \hat{v}_L^T \hat{E}^{-1}\hat{B}
$$
$$
= \frac{1}{n} \begin{bmatrix} \mathbf{1}_r \mathbf{1}_r^T \Pi_g^T \Pi_g & 0_{r \times q} & 0_{r \times r} & \mathbf{1}_r \mathbf{1}_r^T \hat{Q} \\ \mathbf{1}_q \mathbf{1}_r^T \Pi_g^T \Pi_g & 0_{q \times q} & 0_{q \times r} & \mathbf{1}_q \mathbf{1}_r^T \hat{Q} \\ 0_{r \times r} & 0_{r \times q} & 0_{r \times r} & 0_{r \times r} \\ 0_{r \times r} & 0_{r \times q} & 0_{r \times r} & 0_{r \times r} \end{bmatrix} \begin{bmatrix} 0_r \\ \hat{D}_l^{-1}\Pi_l^T \\ 0_r \\ 0_r \end{bmatrix} = 0.
$$

(49)

That completes the proof. □

4 Case study

We illustrate the proposed method on the IEEE 30-bus test system [46], which contains 6 generator buses and 41 transmission lines. The graph representation of this power system is depicted in Fig. 2a. In order to fit the network data into the index setting of the states in our model (3), we use the bus numberings that are different from the original system date. Now, the generator buses are with the indices from 1 to 6. Assume that the distributed controllers of the generators are identical and connected based on the communication network in Fig. 2b. The control parameters in (6) are given by $Q_i = 0.1$ for all generators and $c_{ij} = 1$. Thereby, the closed-loop power system in (44) is established, which include 42 states, namely the voltage phase angles of generators and loads, $\theta_g \in \mathbb{R}^6$ and $\theta_l \in \mathbb{R}^{24}$, the generator voltage
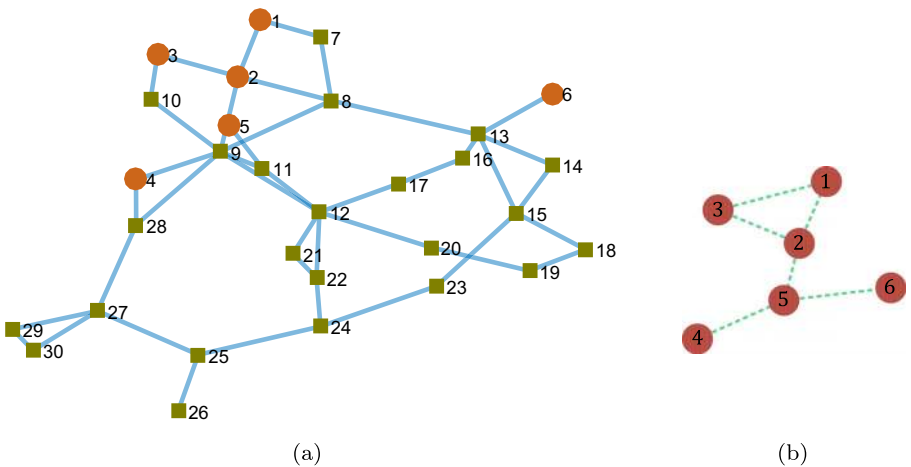

(a) (b)

Fig. 2 **a** The topology of IEEE 30-bus test system. The generators and load buses are represented by circles and squares, respectively. **b** The communication network that links the distributed controllers on 6 generators

frequencies $\omega_g \in \mathbb{R}^6$, and the controller states $\xi \in \mathbb{R}^6$. The positive semidefinite controllability Gramian is then obtained by Theorem 2. Consequently, the dissimilarity matrices of generator and load behaviors are computed using (23) and (24). Due to space reasons, we only give the results for generators as follows.

$$\mathcal{D}^g = \begin{bmatrix} 0 & 0.0737 & 0.0824 & 0.3147 & 0.1401 & 0.2319 \\ 0.0737 & 0 & 0.0854 & 0.2899 & 0.1071 & 0.2121 \\ 0.0824 & 0.0854 & 0 & 0.3124 & 0.1445 & 0.2418 \\ 0.3147 & 0.2899 & 0.3124 & 0 & 0.2472 & 0.3374 \\ 0.1401 & 0.1071 & 0.1445 & 0.2472 & 0 & 0.2041 \\ 0.2319 & 0.2121 & 0.2418 & 0.3374 & 0.2041 & 0 \end{bmatrix}. \tag{50}$$

We then group the generator and load nodes by Algorithm 1, which divides the node sets in the original network into several subsets that contain nodes with small dissimilarities to each other. The clustering results of the generator and load nodes are straightforwardly interpreted by the dendrograms in Fig. 3. The leaves are the bottom vertical lines representing the buses, and the clustering of two nodes is indicated by merging two leaves into a single branch. The dissimilarity can be read from the horizontal position of each fusion. In this example, we cluster node sets of the generators and loads as in Table 1. The clustering results are reasonable owing to the hierarchical structures of the dendrograms.

Using the characteristic matrices of the resulting clusterings for Galerkin projection, we obtain the reduced-order models of the power networks and distributed controller in forms of (41) and (43). The simplified networks are depicted in Fig. 4.
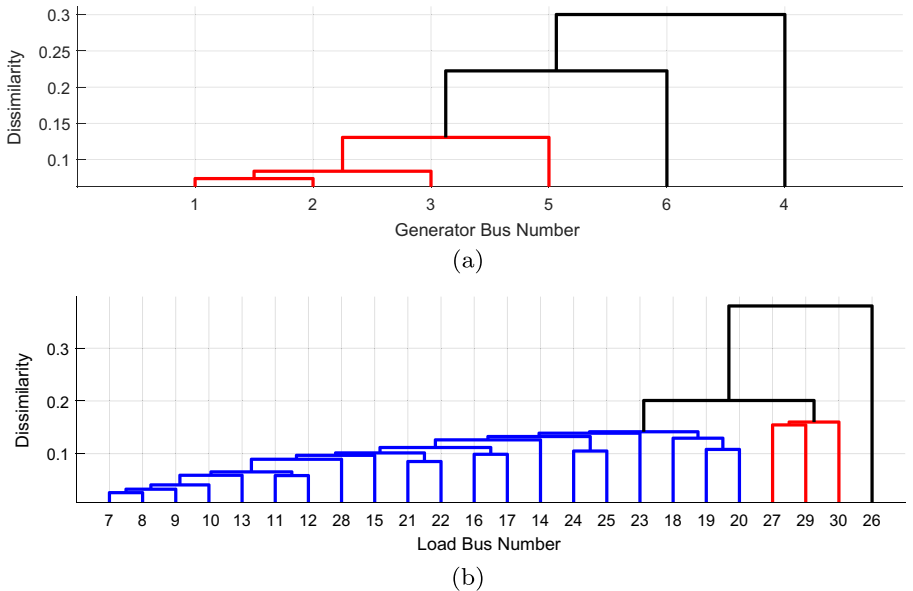

Fig. 3 Dendrograms showing the clusterings of 6 generator buses (**a**) and 24 load buses. The horizontal axis are labeled by bus numberings, and the dissimilarity data are read from vertical axis

Table 1 Clustering results of generator and load buses

Clusters	$\mathcal{C}_1$	$\mathcal{C}_2$	$\mathcal{C}_3$
Generator buses	1, 2, 3, 5	4	6

Clusters	$\mathcal{C}_4$	$\mathcal{C}_5$	$\mathcal{C}_6$
Load buses	7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 28	27, 29, 30	26

Therefore, the network structures are preserved through the reduction process, which means that we can use the power network with much smaller size to approximate the behavior of the original one.

Next, the quality of the approximation is evaluated. We consider the outputs of the closed-loop system Σ_{cl} as θ_g, θ_l, ω_g, and ξ, respectively. Correspondingly, they are approximated by $\Pi_g\hat{\theta}_g$, $\Pi_l\hat{\theta}_l$, $\Pi_g\hat{\omega}_g$, and $\Pi_g\hat{\xi}$, which are assumed to be the outputs of reduced-order model and controller. Based on this, we compute the approximation errors of those variables in terms of $\mathcal{H}_2$-norms (see Table 2). The errors are not significant if we think of the values of the node dissimilarities in (50) and the dimension of the reduced closed-loop model which is considerably lower compared to the original system.

Next, the performance of the reduced-order model is further demonstrated in the time domain. The states of both the original and simplified networks are initialized at zero. From 0 to 20 s, the power demands of the 24 loads, P_l, are assumed by a random vector with all the entries in the range [0, 3]. After 20 s, the demands are set to be a new random constant vector, which is generated from [0, 1]. The unit of P_l is 100 MVA. In Fig. 5, the state trajectories of both systems are compared. Evolving over time, each state trajectory of the reduced network shows a similar directional tendency to the trajectories representing a cluster of nodes in the original network. Furthermore, the most of the response curves of the full-order system are approximated by those of the reduced model with acceptable errors. Besides, from

Fig. 4 The reduced topology of the power transmission lines and communication links, which are represented by solid and dashed edges, respectively
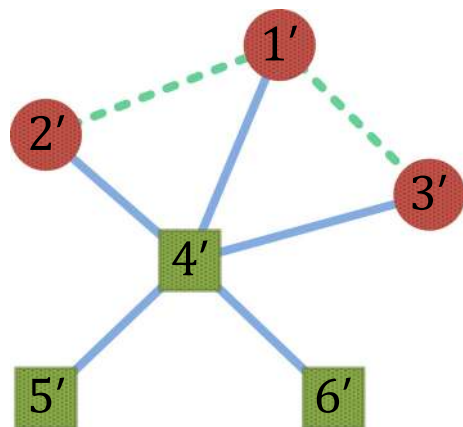
Table 2 The approximation errors evaluated by $\mathcal{H}_2$-norms

Variables	θ_g	θ_l	ω_g	ξ
Errors	0.0791	0.2088	0.2379	0.1187

Fig. 5c, we can see how the full-order and reduced-order distributed controllers regulate the generator frequency deviations back to 0 in both systems. Moreover, Fig. 5d illustrates that the states of both controllers are synchronized to the same trajectories.

We implement this numerical test by MatLab 2016a in the environment of a 64-bit operating system with Intel Core i5-3470 CPU at 3.20 GHz, RAM 8.00 GB. We observe that the evaluation of dissimilarities takes 0.0564 s, while the clustering algorithm only uses 0.0033 s to find three clusters for both generator and load buses. Therefore, the total computation time is mainly consumed by the calculation of the controllability Gramian for dissimilarity evaluation.
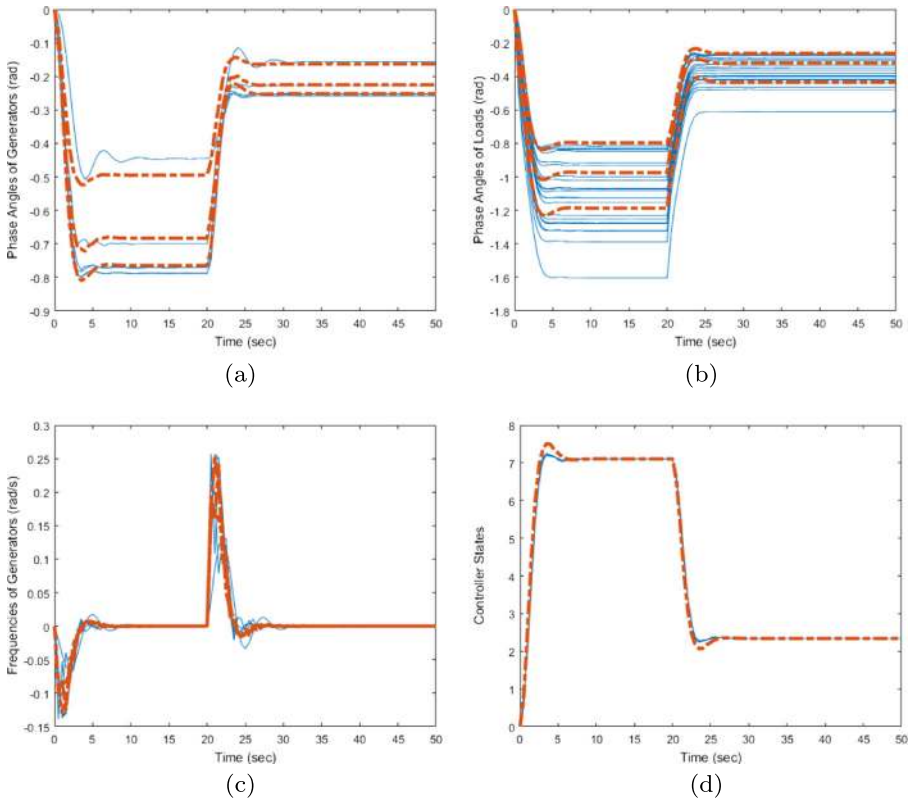

Fig. 5 The comparisons of the state responses of the full-order and reduced-order power systems, where the solid and dashed lines are representing the trajectories of the original and simplified networks

5 Conclusion

We have considered the reduced-order modeling of power networks and distributed controllers, which are expressed as semistable second-order and first-order differential algebraic equations, respectively. By exploiting the controllability Gramian of the semistable closed-loop system, we have proposed a novel notion of node dissimilarity and apply a hierarchical clustering approach to divide the generator and load buses into several subsets. Towards the preservation of network structures in the reduced-order models, the characteristic matrices of the resulting clusters are adopted for the Galerkin projections of both power system and its controller. The explicit reduced-order models are established in the same forms of the original models, which inherit a network interpretation for the interconnections of the power units. A numerical example, at last, has shown the performance of the reduced-order model and controller. This paper provides an idea to explore clusterings of controlled power networks based on dynamical models. As a further extension, other representations of interconnections, generators, loads, and controllers may be investigated, e.g., including resistors on the transmission lines, or more complicated nonlinearities in the generators, or different types of controllers.

References

1. Andreasson, M., Tegling, E., Sandberg, H., Johansson, K.H.: Coherence in synchronizing power networks with distributed integral control. arXiv:1703.10425 (2017)
2. Antoulas, A.C.: Approximation of Large-Scale Dynamical Systems. SIAM (2005)
3. Avramovic, B., Kokotovic, P.V., Winkelman, J.R., Chow, J.H.: Area decomposition for electromechanical models of power systems. Automatica **16**(6), 637–648 (1980)
4. Barabási, A.L.: Network Science. Cambridge University Press (2016)
5. Bartels, R.H., Stewart, G.W.: Solution of the matrix equation $AX + XB = C$. Commun. ACM **15**(9), 820–826 (1972)
6. Beaulieu, A., de Wilde, J., Scherpen, J.M.A.: Smart Grids from a Global Perspective. Springer (2016)
7. Bergen, A.R., Hill, D.J.: A structure preserving model for power system stability analysis. IEEE Trans. Power Apparatus Syst. **PAS-100**(1), 25–35 (1981)
8. Bernstein, D.S., Bhat, S.P.: Lyapunov stability, semistability, and asymptotic stability of matrix second-order systems. J. Mech. Des. **117**(B), 145–153 (1995)
9. Besselink, B., Sandberg, H., Johansson, K.H.: Clustering-based model reduction of networked passive systems. IEEE Trans. Autom. Control **61**(10), 2958–2973 (2016)
10. Bettayeb, M., Al-Saggaf, U.M.: Practical model reduction techniques for power systems. Electr. Power Syst. Res. **25**(3), 169–176 (1992)

11. Caliskan, S.Y., Tabuada, P.: Towards Kron reduction of generalized electrical networks. Automatica **50**(10), 2586–2590 (2014)

12. Cheng, X., Kawano, Y., Scherpen, J.M.A.: Graph structure-preserving model reduction of linear network systems. In: Proceedings of 2016 European Control Conference, pp. 1970–1975 (2016)

13. Cheng, X., Kawano, Y., Scherpen, J.M.A.: Reduction of second-order network systems with structure preservation. IEEE Trans. Autom. Control **62**, 5026–5038 (2017)

14. Cheng, X., Scherpen, J.M.A.: Introducing network Gramians to undirected network systems for structure-preserving model reduction. In: Proceedings of 55th IEEE Conference on Decision and Control, pp. 5756–5761 (2016)

15. Cheng, X., Scherpen, J.M.A.: Balanced truncation approach to linear network system model order reduction. In: Proceedings of the 20th World Congress of the International Federation of Automatic Control (IFAC), pp. 2506–2511. Toulouse (2017)

16. Chow, J.H.: Power System Coherency and Model Reduction. Springer (2013)

17. Chow, J.H., Galarza, R., Accari, P., Price, W.W.: Inertial and slow coherency aggregation algorithms for power system dynamic model reduction. IEEE Trans. Power Syst. **10**(2), 680–685 (1995)

18. Dörfler, F., Bullo, F.: Topological equivalence of a structure-preserving power network model and a non-uniform kuramoto model of coupled oscillators. In: Proceedings of IEEE 50th Conference on Decision and Control and European Control Conference (CDC-ECC), pp. 7099–7104. IEEE, Orlando (2011)

19. Dörfler, F., Bullo, F.: Kron reduction of graphs with applications to electrical networks. IEEE Trans. Circ. Syst. I: Regular Papers **60**(1), 150–163 (2013)

20. Dörfler, F., Jovanovic, M.R., Chertkov, M., Bullo, F.: Sparsity-promoting optimal wide-area control of power networks. IEEE Trans. Power Syst. **29**(5), 2281–2291 (2014)

21. Dörfler, F., Simpson-Porco, J.W., Bullo, F.: Breaking the hierarchy: distributed control and economic optimality in microgrids. IEEE Trans. Control Netw. Syst. **3**(3), 241–253 (2016)

22. Fiaz, S., Zonetti, D., Ortega, R., Scherpen, J.M.A., van Der Schaft, A.J.: A port-Hamiltonian approach to power network modeling and analysis. Eur. J. Control. **19**(6), 477–485 (2013)

23. Freitas, F.D., Rommes, J., Martins, N.: Gramian-based reduction method applied to large sparse power system descriptor models. IEEE Trans. Power Syst. **23**(3), 1258–1270 (2008)

24. Ghosh, S., Senroy, N.: Balanced truncation approach to power system model order reduction. Electric Power Comp. Syst. **41**(8), 747–764 (2013)

25. Golub, G., Nash, S., Van Loan, C.: A Hessenberg-Schur method for the problem $AX + XB = C$. IEEE Trans. Autom. Control **24**(6), 909–913 (1979)

26. Hogan, E., Cotilla-Sanchez, E., Halappanavar, M., Wang, S., Mackey, P., Hines, P., Huang, Z.: Towards effective clustering techniques for the analysis of electric power grids. In: Proceedings of the 3rd International Workshop on High Performance Computing, Networking and Analytics for the Power Grid, p. 1. ACM (2013)

27. Ishizaki, T., Imura, J.I.: Clustered model reduction of interconnected second-order systems. IEICE Nonlin. Theory Appl. **6**(1), 26–37 (2015)

28. Ishizaki, T., Kashima, K., Girard, A., Imura, J.i., Chen, L., Aihara, K.: Clustered model reduction of positive directed networks. Automatica **59**, 238–247 (2015)

29. Jain, A.K., Murty, M.N., Flynn, P.J.: Data clustering: a review. ACM Comput. Surv. (CSUR) **31**(3), 264–323 (1999)

30. Kundur, P., Balu, N.J., Lauby, M.G.: Power System Stability and Control, vol. 7. McGraw-Hill, New York (1994)

31. Larsen, G.K., van Foreest, N.D., Scherpen, J.M.A.: Power supply–demand balance in a smart grid: an information sharing model for a market mechanism. Appl. Math. Model. **38**(13), 3350–3360 (2014)

32. Li, Y., Geng, G., Jiang, Q.: An efficient parallel Krylov-Schur method for Eigen-analysis of large-scale power systems. IEEE Trans. Power Syst. **31**(2), 920–930 (2016)

33. Mehrjerdi, H., Lefebvre, S., Asber, D., Saad, M.: Graph partitioning of power network for emergency voltage control. In: 2013 9th Asian Control Conference (ASCC), pp. 1–6. IEEE (2013)

34. Michele, C., Trip, S., Persis, C.D., Cheng, X., Ferrara, A., Schaft, A.v.d.: A robust consensus algorithm for current sharing and voltage regulation in dc microgrids. To appear in IEEE Transactions on Control Systems Technology (2018)

35. Mlinarić, P., Grundel, S., Benner, P.: Efficient model order reduction for multi-agent systems using QR decomposition-based clustering. In: Proceedings of 54th IEEE Conference on Decision and Control (CDC), pp. 4794–4799 (2015)

36. Monshizadeh, N., De Persis, C., van der Schaft, A.J., Scherpen, J.M.A.: A novel reduced model for electrical networks with constant power loads. IEEE Transactions on Automatic Control (2017)
37. Monshizadeh, N., Trentelman, H.L., Camlibel, M.K.: Projection-based model reduction of multi-agent systems using graph partitions. IEEE Trans. Control Netw. Syst. **1**(2), 145–154 (2014)
38. Pagani, G.A., Aiello, M.: The power grid as a complex network: a survey. Physica Statist. Mech. Appl. **392**(11), 2688–2700 (2013)
39. Pyo, G.C., Park, J.W., Moon, S.I.: A new method for dynamic reduction of power system using pam algorithm. In: IEEE PES General Meeting, pp. 1–7. https://doi.org/10.1109/PES.2010.5589996 (2010)
40. Brualdi, R.A., Ryser, H.J.: Combinatorial Matrix Theory. Cambridge University Press (1991)
41. Romeres, D., Dörfler, F., Bullo, F.: Novel results on slow coherency in consensus and power networks. In: Proceedings of the 2013 European Control Conference, pp. 742–747. IEEE, Zürich (2013)
42. Scarciotti, G.: Low computational complexity model reduction of power systems with preservation of physical characteristics. IEEE Trans. Power Syst. **32**(1), 743–752 (2017)
43. van der Schaft, A.J.: On model reduction of physical network systems. In: Proceedings of 21st International Symposium on Mathematical Theory of Networks and Systems (MTNS), pp. 1419–1425. Groningen (2014)
44. Schiffer, J., Goldin, D., Raisch, J., Sezi, T.: Synchronization of droop-controlled microgrids with distributed rotational and electronic generation. In: Proceedings of 2013 IEEE 52nd Annual Conference on Decision and Control (CDC), pp. 2334–2339. IEEE (2013)
45. Sturk, C., Vanfretti, L., Chompoobutrgool, Y., Sandberg, H.: Coherency-independent structured model reduction of power systems. IEEE Trans. Power Syst. **29**(5), 2418–2426 (2014)
46. Totonchi, I., Al Akash, H., Al Akash, A., Faza, A.: Sensitivity analysis for the IEEE 30 bus system using load-flow studies. In: 3rd International Conference on Electric Power and Energy Conversion Systems (EPECS), pp. 1–6. IEEE (2013)
47. Trip, S., Bürger, M., De Persis, C.: An internal model approach to (optimal) frequency regulation in power grids with time-varying voltages. Automatica **64**, 240–253 (2016)
48. Xu, G., Vittal, V.: Slow coherency based cutset determination algorithm for large power systems. IEEE Trans. Power Syst. **25**(2), 877–884 (2010)
49. Zhang, Z., Hu, X., Cheng, C.K., Wong, N.: A block-diagonal structured model reduction scheme for power grid networks. In: Design, Automation & Test in Europe Conference & Exhibition (DATE), 2011, pp. 1–6. IEEE (2011)