Learning Coherent Clusters in Weakly-Connected Network Systems

Hancheng Min HANCHMIN@JHU.EDU
Enrique Mallada MALLADA@JHU.EDU
Department of Electrical and Computer Engineering
Johns Hopkins University, Baltimore, MD, U.S.

Editors: N. Matni, M. Morari, G. J. Pappas

Abstract

We propose a structure-preserving model-reduction methodology for large-scale dynamic networks with tightly-connected components. First, the coherent groups are identified by a spectral clustering algorithm on the graph Laplacian matrix that models the network feedback. Then, a reduced network is built, where each node represents the aggregate dynamics of each coherent group, and the reduced network captures the dynamic coupling between the groups. We provide an upper bound on the approximation error when the network graph is randomly generated from a weight stochastic block model. Finally, numerical experiments align with and validate our theoretical findings.

Keywords: Spectral Clustering; Network Systems; Model Reduction

1. Introduction

In networked dynamical systems, coherence refers to a coordinated behavior from a group of nodes such that all nodes have similar dynamical responses to some external disturbances (Chow, 1982). Coherence analysis is useful in understanding the collective behavior of large networks, including consensus networks (Olfati-Saber and Murray, 2004), transportation networks (Bamieh et al., 2012), and power networks (Ramaswamy et al., 1995). However, little do we know about the underlying mechanism that causes such coherent behavior to emerge in various networks.

Classic slow coherence analyses (Chow, 1982; Ramaswamy et al., 1996; Romeres et al., 2013; Tyuryukanov et al., 2021; Fritzsch and Jacquod, 2022) (with applications mostly to power networks) usually consider the second-order electro-mechanical model without damping: $\ddot{x} = -M^{-1}Lx$, where M is the diagonal matrix of machine inertias, and L is the Laplacian matrix whose elements are synchronizing coefficients between pair of machines. The coherency or synchrony (Ramaswamy et al., 1996) (a generalized notion of coherency) is identified by studying the first few slowest eigenmodes (eigenvectors with small eigenvalues) of $M^{-1}L$. The analysis can be carried over to the case of uniform (Chow, 1982) and non-uniform (Romeres et al., 2013) damping. However, such state-space-based analysis is limited to very specific node dynamics (second order) and does not account for more complex dynamics or controllers that are usually present at a node level; e.g., in the power systems literature (Jiang et al., 2021b,a; Ekomwenrenren et al., 2021). There is, therefore, the need for coherence identification procedures that work for more general network systems.

Recently, it has been theoretically established that coherence naturally emerges when the connectivity of a group of nodes is sufficiently large, regardless of the node dynamics, as long as the interconnection remains stable (Min and Mallada, 2019; Min et al., 2021). The analysis also provides an asymptotically (as the network connectivity increases) exact characterization of the co-

herent response, which amounts to a harmonic sum of individual node transfer functions. Thus, in a sense, coherence identification is closely related to the problem of finding tightly connected components in the network, for which many clustering algorithms based on the spectral embedding of graph adjacency or Laplacian matrices, exist and are theoretically justified (Bach and Jordan, 2004). This leads to the natural question: *Can these graph-based clustering algorithms be adopted for coherence identification in networked dynamical systems?* Intuitively, when we apply those clustering algorithms to identify tightly-connected components in the network, each component should be coherent also in the dynamical sense. Then, applying Min and Mallada (2019); Min et al. (2021) for each cluster should lead to a good model for each coherent group, which, after interconnected with an appropriately chosen reduced graph, should lead to a good network-reduced aggregate model of the dynamic interactions across coherent components. Min and Mallada (2022a) formalizes such an approach exclusively for networks with two coherent components.

In this paper, we extend the result in Min and Mallada (2022a) to networks with an arbitrary number of coherent groups. Specifically, our structure-preserving approximation model for large-scale networks is constructed in two stages: First, the coherent groups are identified by a spectral clustering algorithm solely on the graph Laplacian matrix of the network; Then a reduced network, in which each node represents the aggregate dynamics of one coherent group, approximates the dynamical interactions between the coherent groups in the original network. More importantly, we provide an upper bound on the approximation error when the network graph is randomly generated from a weight stochastic block model, and the numerical results align with our theoretical findings.

Structure-preserving model reduction has been mostly studied for mechanical systems (Li and Bai, 2006; Lall et al., 2003) using Krylov subspace projection, and has only recently been adopted for network systems such as power networks (Safaee and Gugercin, 2021). However, Safaee and Gugercin (2021) assumes second-order nodal dynamics, and the resulting model can not be interpreted as a network. Our approach exploits the natural multi-cluster structure of many network systems, resulting in a reduced network that captures the interaction among the clusters.

The rest of the paper is organized as follows: We formalize the coherence identification problem in Section 2 and also propose our reduction algorithm. Then we show in Section 3 the rationale behind the algorithm and provide theoretical justification in Section 4. Lastly, we validate our model by numerical experiments in Section 5.

Notation: For a real vector x, $\|x\| = \sqrt{x^T x}$ denotes the 2-norm of x, $[x]_i$ denotes its i-th entry, and for a real matrix A, $\|A\|, \|A\|_F$ denotes the spectral norm, and the Frobenius norm, respectively. For a Hermitian matrix A of size n, we let $\lambda_i(A)$ denote its i-th smallest eigenvalue, and $v_i(A)$ the associated unit-norm eigenvector. We let $\mathrm{diag}\{x_i\}_{i=1}^n$ denote a $n \times n$ diagonal matrix with diagonal entries x_i, I_n denote the identity matrix of order n, V^T denote the transpose of matrix V, $\mathbb{1}_n$ denote $[1, \cdots, 1]^T$ with dimension n, and $[n]$ denote the set $\{1, 2, \cdots, n\}$. For non-negative random variables $X(n), Y(n)$, we write $X(n) \sim \mathcal{O}_p(Y(n))$ if $\exists M > 0$, s.t. $\lim_{n\to\infty} \mathbb{P}\left(X(n) \leq MY(n)\right) = 1$. We write $X(n) \sim \Omega_p(Y(n))$ if $\exists M > 0$, s.t. $\lim_{n\to\infty} \mathbb{P}\left(X(n) \geq MY(n)\right) = 1$.

2. Preliminaries

2.1. Network Model

We consider a similar network model to the one considered in Min et al. (2021); Min and Mallada (2022a). The network consisting of n nodes ($n \geq 2$), indexed by $i \in [n]$ with the block diagram structure in Fig.1. L is the Laplacian matrix of an undirected, weighted graph that describes the

network interconnection. We further use $f(s)$ to denote the transfer function representing the dynamics of the network coupling, and $G(s) = \mathrm{diag}\{g_i(s)\}_{i=1}^n$ to denote the nodal dynamics, with $g_i(s),\ i \in [n]$, being an SISO transfer function representing the dynamics of node i.

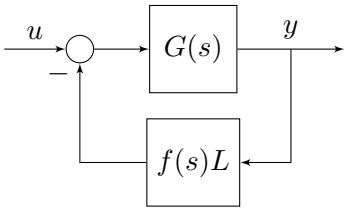

Figure 1: Block Diagram of General Networked Dynamical Systems

The network takes a vector signal $u = [u_1, \cdots, u_n]^T$ as input, whose component u_i is the disturbance or input to node i. The network outputs a vector $y = [y_1, \cdots, y_n]^T$ that contains the individual node outputs $y_i, i = 1, \cdots, n$. We are interested in characterizing and approximating the response of the transfer matrix $T_{yu}(s)$ under certain assumptions on the network topology, i.e., the Laplacian matrix L.

2.2. Network Coherence and Structure-preserving Model Reduction

Recent work (Min and Mallada, 2019; Min et al., 2021) has shown that, under mild assumptions, the following holds[1] for almost any $s_0 \in \mathbb{C}$,

$$\lim_{\lambda_2(L)\to\infty} \|T_{yu}(s_0) - \hat{g}(s_0)\mathbb{1}\mathbb{1}^T\| = 0\,, \tag{1}$$

where

$$\hat{g}(s) = \left(\sum_{i=1}^n g_i^{-1}(s)\right)^{-1}. \tag{2}$$

That is, when the algebraic connectivity $\lambda_2(L)$ of the network is high, one can approximate $T_{yu}(s)$ by a rank-one transfer matrix. Such a rank-one transfer matrix $\hat{g}(s_0)\mathbb{1}\mathbb{1}^T$ precisely describes the coherent behavior of the network: The network takes the aggregated input $\hat{u} = \mathbb{1}^T u = \sum_{i=1}^n u_i$, and responds coherently as $\hat{y}\mathbb{1}$, where $\hat{y} = \hat{g}(s)\hat{u}$. Therefore, it suffices to study $\hat{g}(s)$ to understand the coherent behavior in a tightly-connected network. The aggregate dynamics $\hat{g}(s)$ has been studied for tightly-connected power networks (Min et al., 2021), and Jiang et al. (2021a); Häberle et al. (2022) proposed a control design that leads to desirable response for the entire network by shaping the response of $\hat{g}(s)$.

However, practical networks are not necessarily tightly-connected. Instead, they often contain multiple groups of nodes such that within each group, the nodes are tightly-connected while between groups, the nodes are weakly-connected. Then the network dynamics can be reduced to dynamic interactions among these groups. To approximate such interaction, it is natural first to identify *coherent groups*, or *coherent clusters*, in the network, then apply the aforementioned analysis to obtain the coherent dynamics $\hat{g}(s)$ for each group, and replace the entire coherent group by an

1. In Min et al. (2021), the transfer matrix $\frac{1}{n}\bar{g}(s)\mathbb{1}\mathbb{1}^T$ appeared in the limit, where $\bar{g}(s) = \left(\frac{1}{n}\sum_{i=1}^n g_i^{-1}(s)\right)^{-1}$. It is easy to verify that $\frac{1}{n}\bar{g}(s)\mathbb{1}\mathbb{1}^T = \hat{g}(s_0)\mathbb{1}\mathbb{1}^T$

aggregate node with $\hat{g}(s)$. Lastly, one needs to find a reduced network of the same size as the number of coherent groups, which characterize the interaction among these groups. The aggregate dynamics and the reduced network allow us to build a network model with exactly the same structure as the one in Figure 1 but with a much smaller size, for which we refer to such an approach as *structure-preserving model reduction* and call the resulting reduction model *structure-preserving*. Figure 2 shows our proposed reduced model in the case of three coherent groups, for which the algorithm details are explained later.

In the case of two coherent groups, Min et al. (2021) proposed an algorithm that first uses a simple spectral clustering algorithm to identify the two coherent groups, then shows that the weight of the reduced network (a two-node graph with a single edge) is determined by the algebraic connectivity of the original Laplacian matrix $\lambda_2(L)$ and the size of each group. However, such an approach does not work for networks with more than two coherent groups.

2.3. Our Algorithm

In this paper, we propose a structure-preserving model reduction algorithm for networks with an arbitrary number of groups.

Algorithm 1 Structure-Preserving Network Reduction via Spectral Clustering

Data: Network Model $(G(s) = \text{diag}\{g_i(s)\}_{i=1}^n, L, f(s))$; Number of clusters k

Do:

1. $(\{\mathcal{I}_i\}_{i=1}^k, V_k, \Lambda_k) \leftarrow$ SpectralClustering(L); `// Spectral clustering`
   Construct $P_{\{\mathcal{I}_i\}_{i=1}^k}$ as in (5);
2. $\hat{g}_i(s) \leftarrow \left(\sum_{j\in\mathcal{I}_i} g_j^{-1}(s)\right)^{-1}$, $i = 1, \cdots, k$; `// Aggregation`
   $\hat{G}(s) = \text{diag}\{\hat{g}_i(s)\}_{i=1}^k$;
3. $S \leftarrow$(Solution to (8));
   $L_k = (S^{-1})^T \Lambda_k S^{-1}$; `// Construct reduced network`

Result: $\hat{T}_k(s) \leftarrow P_{\{\mathcal{I}_i\}_{i=1}^k}(I_k + \hat{G}_k(s)L_k f(s))^{-1}\hat{G}_k(s)P^T_{\{\mathcal{I}_i\}_{i=1}^k}$

This algorithm, whose rationale will be explained in detail in Section 3, follows the same procedure as we discussed in the previous section: Firstly, we utilize some spectral clustering algorithm to obtain a k-way partition $\{\mathcal{I}_i\}_{i=1}^n$ of $[n]$ that encodes the clustering results. Notice that here any spectral clustering algorithm works. For subsequent steps, we also need to keep the first k smallest eigenvalues of L (in $\Lambda_k = \text{diag}\{\lambda_i(L)\}_{i=1}^k$) and their associated eigenvectors (in $V_k = \begin{bmatrix} v_1(L) & v_2(L) & \cdots v_k(L) \end{bmatrix}$). Then the nodes in the same group $\mathcal{I}_i$ are aggregated into $\hat{g}_i(s)$. Lastly, the Laplacian matrix of the reduced network is constructed after solving an optimization problem (8) that can be viewed as a refinement process on the Laplacian spectral embedding V_k. This algorithm will return a transfer matrix $\hat{T}_k(s)$ as an approximation model of the original transfer matrix $T_{yu}(s)$. The algorithm is illustrated in Figure 2.

In the rest of the paper, we first discuss how our algorithm is constructed based on the aforementioned coherence analysis (Min et al., 2021) in Section 3, then show that our proposed approxima-

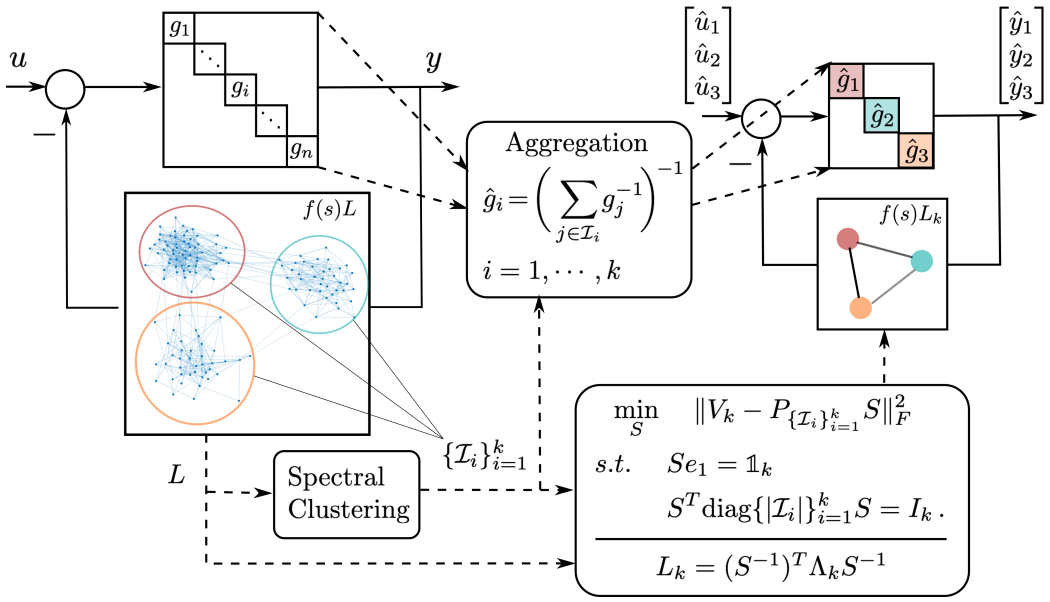

Figure 2: Functional illustration of Algorithm 1.

tion model is asymptotically accurate in a random graph setting where the network graph is sampled from a *weighted stochastic block model* (Ahn et al., 2018) by showing an approximation error bound between the network $T_{yu}(s)$ and the proposed reduced model $\hat{T}_k(s)$ (in Section 4). Lastly, we verify our theoretical findings through a numerical simulation in Section 5.

3. Structure-Preserving Network Reduction via Spectral Clustering

Our algorithm roots in the recent analysis (Min et al., 2021; Min and Mallada, 2022a) showing that the network transfer matrix $T_{yu}(s)$ is approximately low rank for networks with Laplacian matrices satisfying some spectral property. Such a low-rank approximation is generally not structure-preserving, for which we use its closest structure-preserving approximation, obtained by spectral clustering on graph Laplacian L and a refinement process on its eigenvectors V_k, as our final reduction model for the original $T_{yu}(s)$.

3.1. Low-rank Approximation of Network Transfer Matrix

Given the network Laplacian L and its first k smallest eigenvalues (in a diagonal matrix) $\Lambda_k = \mathrm{diag}\{\lambda_i(L)\}_{i=1}^k$ and the associated eigenvectors $V_k = \begin{bmatrix} v_1(L) & v_2(L) & \cdots v_k(L) \end{bmatrix}$ (we also refer it as *Laplacian spectral embedding*), we define the following rank-k transfer matrix

$$T_k(s) = V_k(V_k^T G^{-1}(s) V_k + f(s)\Lambda_k)^{-1} V_k^T\,, \tag{3}$$

and we have the following result:

Theorem 1 *For $s_0 \in \mathbb{C}$ that is not a pole of $f(s)$ and has these two quantities*

$$\|T_k(s_0)\| := M_1, \textit{and } \max_{1\le i\le n} |g_i^{-1}(s_0)| := M_2\,,$$

finite. Then whenever $|f(s_0)|\lambda_{k+1}(L) > M_2 + M_1 M_2^2$*, the following inequality holds:*

$$\|T_{yu}(s_0) - T_k(s_0)\| \leq \frac{(M_1 M_2 + 1)^2}{|f(s_0)|\lambda_{k+1}(L) - M_2 - M_1 M_2^2}\,. \tag{4}$$

We refer the readers to Min and Mallada (2022b) for the proof. Previous work presented similar approximation results for the case $k = 1$ (Min et al., 2021) and $k = 2$ (Min and Mallada, 2022a). Theorem 1 shows that in the large $\lambda_{k+1}(L)$ regime, one can somewhat approximate the original transfer matrix $T_{yu}(s)$ by a low-rank one $T_k(s)$, but the approximation result in (4) is weaker than that the two transfer matrices $T_{yu}(s)$ and $T_k(s)$ are close in the $\mathcal{H}_\infty$ sense. It heavily depends on the choice of s_0, the frequency of interest, as we should not expect $T_{yu}(s)$ and $T_k(s)$ to behave similarly under input of any frequency. For the case of $k = 1$, Min et al. (2021) have shown that if $\sup_{s\in(-j\eta,+j\eta)} \|T_{yu}(s) - T_k(s)\|$ is small for some $\eta > 0$, then one can show, provided that $T_{yu}(s)$ and $T_k(s)$ are stable, the time domain responses of the two transfer matrices under low-frequency inputs (characterized by η) are close to each other.

Following such observation, we consider any $\hat{T}_k(s)$ with $\sup_{s\in(-j\eta,+j\eta)} \|T_{yu}(s) - \hat{T}_k(s)\|$ being small for some $\eta > 0$ as a good approximation for the original network. Applying (4) uniformly over $\{s : s \in (-j\eta, +j\eta)\}$, one can show that $T_k(s)$ is such a good approximation when $\lambda_{k+1}(L)$ is large. However, $T_k(s)$ is, in general, not structure-preserving, and thus may not be interpreted as a reduced network of aggregate nodes. Therefore, we need to find a structure-preserving $\hat{T}_k(s)$ that is close to $T_k(s)$.

3.2. Structured Low-rank Approximation via Spectral Embedding Refinement

We first discuss the case when $T_k(s)$ is structure-preserving. We show that a special property on the Laplacian spectral embedding V_k suffices. For some $\mathcal{I} \subseteq [n]$, we let $\mathbb{1}_\mathcal{I}$ be an $n \times 1$ vector such that $[\mathbb{1}_\mathcal{I}]_i = \begin{cases} 1, & i \in \mathcal{I} \\ 0, & i \notin \mathcal{I} \end{cases}$.

Definition 2 *A Laplacian matrix L is said to be* ***k-block-ideal*** *with respect to a k-way partition $\{\mathcal{I}_1, \cdots, \mathcal{I}_k\}$ of $[n]$, if there exists some invertible matrix $S \in \mathbb{R}^{k\times k}$ such that*

$$V_k := \begin{bmatrix} v_1(L) & v_2(L) & \cdots & v_k(L) \end{bmatrix} = \begin{bmatrix} \mathbb{1}_{\mathcal{I}_1} & \mathbb{1}_{\mathcal{I}_2} & \cdots & \mathbb{1}_{\mathcal{I}_k} \end{bmatrix} S\,.$$

We also say V_k is k-block-ideal in this case.

A k-block-ideal spectral embedding V_k, together with Λ_k containing the bottom k eigenvalues of L, would immediately lead to a reduced network: the k coherent groups are determined by the k-way partition $\{\mathcal{I}_i\}_{i=1}^k$, and the invertible matrix S, combined with Λ_k, characterize the interconnection in the reduced network, as show in the following theorem:

Theorem 3 *Given a k-block-ideal Laplacian L associated with a partition $\{\mathcal{I}_1, \cdots, \mathcal{I}_k\}$ and an invertible matrix S, and we define*

$$P_{\{\mathcal{I}_i\}_{i=1}^k} := \begin{bmatrix} \mathbb{1}_{\mathcal{I}_1} & \mathbb{1}_{\mathcal{I}_2} & \cdots & \mathbb{1}_{\mathcal{I}_k} \end{bmatrix}, \tag{5}$$

then

$$T_k(s) = P_{\{\mathcal{I}_i\}_{i=1}^k}(I_k + \hat{G}_k(s)L_k f(s))^{-1}\hat{G}_k(s)P^T_{\{\mathcal{I}_i\}_{i=1}^k}, \tag{6}$$

where $\hat{G}(s) = \mathrm{diag}\{\hat{g}_i(s)\}_{i=1}^k$, $\hat{g}_i(s) = \left(\sum_{j\in\mathcal{I}_i} g_j^{-1}(s)\right)^{-1}$ *and* $L_k = (S^{-1})^T \Lambda_k S^{-1}$.

We refer the readers to Min and Mallada (2022b) for the proof. Theorem 3 shows that under k-block-ideal V_k, the dynamical behavior of T_k is structure-preserving since it is fully characterized by a reduced network with k nodes, with nodal dynamics $\hat{G}(s)$ and network coupling L_k. Each node $\hat{g}_i(s)$ represents the aggregate dynamics for nodes in $\mathcal{I}_i$. Any input u to $T_k(s)$ is aggregated into $\begin{bmatrix}\hat{u}_1 & \cdots & \hat{u}_k\end{bmatrix}^T = P^T_{\{\mathcal{I}_i\}_{i=1}^k} u$ as the input to the reduced network. Then the output $\begin{bmatrix}\hat{y}_1 & \cdots & \hat{y}_k\end{bmatrix}^T$ is "broadcast" to the original nodes via $P_{\{\mathcal{I}_i\}_{i=1}^k}$ such that every node in the same $\mathcal{I}_i$ has the same response.

Notice that such structure-preserving property only depends on the Laplacian spectral embedding V_k. For V_k that is not k-block-ideal, we should be able to find a $\hat{V}_k$ close to V_k and is k-block-ideal. This gives rise to the following optimization problem:

$$\min_{S,\{\mathcal{I}_i\}_{i=1}^k} \|V_k - P_{\{\mathcal{I}_i\}_{i=1}^k} S\|_F^2, \quad s.t. \quad Se_1 = \mathbb{1}_k/\sqrt{n},\ S^T \mathrm{diag}\{|\mathcal{I}_i|\}_{i=1}^k S = I_k\,. \tag{7}$$

The resulting $\hat{V}_k = P_{\{\mathcal{I}_i\}_{i=1}^k} S$ is a refinement of V_k that is k-ideal, and the constraints in (7) ensures that the first column of $\hat{V}_k$ is $\mathbb{1}_n/\sqrt{n}$ and that $\hat{V}_k^T \hat{V}_k = I_k$. Now

$$\hat{T}_k(s) = \hat{V}_k(\hat{V}_k^T G^{-1}(s) \hat{V}_k + f(s)\Lambda_k)^{-1} \hat{V}_k^T$$

is structure-preserving by Theorem 3. In the optimization problem (7), the need for identifying coherent groups is implicitly suggested by the fact that we are optimizing over all possible k-way partitions of n, and the reduced network interconnection is constructed by jointly optimizing over invertible S.

Generally, (7) is hard to solve. Notice, however, that given a fixed partition $\{\mathcal{I}_i\}_{i=1}^k$, one can find a closed-form solution (We refer the readers to Min and Mallada (2022b) for the details) to the following optimization problem

$$\min_{S} \|V_k - P_{\{\mathcal{I}_i\}_{i=1}^k} S\|_F^2, \quad s.t. \quad Se_1 = \mathbb{1}_k/\sqrt{n},\ S^T \mathrm{diag}\{|\mathcal{I}_i|\}_{i=1}^k S = I_k\,. \tag{8}$$

This suggests that a computationally efficient way to find a sub-optimal solution to (7): First, we use any spectral clustering algorithm to find a good partition/clustering $\{\mathcal{I}_i\}_{i=1}^k$, then refine the spectral embedding V_k by optimizing (8) with the obtained partition, resulting in our Algorithm 1.

4. Performance Analysis

In this section, we provide an error bound on $\sup_{s\in(-j\eta,j\eta)} \|T_{yu}(s) - \hat{T}_k(s)\|$ for our proposed approximation model $\hat{T}_k(s)$ from Algorithm 1. As we discussed in Section 3.1, such error measure is related to how close the time-domain response of $\hat{T}_k(s)$ is to the one of $T_{yu}(s)$ when subjected to low-frequency inputs. We consider a Laplacian sampled from a stochastic weighted block model.

4.1. Weighted Stochastic Block Model

We first discuss how we sample our Laplacian matrix from a weighted stochastic block model $(\{\mathcal{I}_i\}_{i=1}^k, Q, W)$. Here, $\{\mathcal{I}_i\}_{i=1}^k$ is a k-way partition of $[n]$, $Q \in [0,1]^{k\times k}$, and $W \in \mathbb{R}_{\geq 0}^{k\times k}$, where $Q_{ij} = Q_{ji}, W_{ij} = W_{ji}$. We let (j) denote the *block membership* of node j: when $j \in \mathcal{I}_i$, then $(j) = i$. The adjacency matrix A is sampled as follows:

$$A_{ij} = \begin{cases} W_{(i),(j)}, & \text{with probability } Q_{(i),(j)} \\ 0, & \text{with probability } 1 - Q_{(i),(j)} \end{cases}, \quad i \geq j, \qquad A_{ij} = A_{ji}, \quad i < j\,. \tag{9}$$

That is, each (undirected) edge i, j appears independently with probability $Q_{(i),(j)}$ that is determined by the block membership of node i, j, and has weight $W_{(i),(j)}$ if it appears. Then we have the Laplacian matrix L:

$$L = D_A - A, \quad D_A = \mathrm{diag}\{A\mathbb{1}\}\,. \tag{10}$$

4.2. Approximation Error Bound

Given the network model $(G(s), L, f(s))$ with L sampled from a weighted stochastic block model $(\{\mathcal{I}_i\}_{i=1}^k, Q, W)$, we show that under certain assumptions, the error $\sup_{s\in(-j\eta,j\eta)} \|T_{yu}(s) - \hat{T}_k(s)\|$ is small with high probability when the network size is sufficiently large. We start by stating our assumptions.

Assumption 1 *For our network model $(G(s), L, f(s))$ with L sampled from a weighted stochastic block model $(\{\mathcal{I}_i\}_{i=1}^k, Q, W)$, we assume the following:*

1. *All $g_i(s), f(s)$ are rational. Moreover, node dynamics are **output strictly passive**: There exists $\gamma > 0$, such that for $i = 1, \cdots, n$, $Re(g_i(s)) \geq \frac{1}{\gamma}|g_i(s)|^2, \forall Re(s) > 0$, and network coupling $f(s)$ is **positive real**: $Re(f(s)) > 0, \forall Re(s) > 0$, and $Im(f(s)) = 0, \forall Re(s) = 0$*

2. *The node dynamics satisfies that for any $\eta > 0$, there exists $M(\eta)$ such that for $i = 1, \cdots, n$*

$$\sup_{s\in(-j\eta,+j\eta)} |g_i^{-1}(s)| \leq M(\eta)\,. \tag{11}$$

   The network coupling $f(s)$ satisfies that $F_l(\eta) := \inf_{s\in(-j\eta,+j\eta)} |f(s)|$ is positive for all $\eta > 0$.

3. *The blocks are approximately balanced:*

$$\frac{n_{\max}}{n_{\min}} \leq \rho\,, \tag{12}$$

   for some $\rho \geq 1$, where $n_{\max} := \max_{1\leq i\leq k} |\mathcal{I}_i|$ and $n_{\min} := \min_{1\leq i\leq k} |\mathcal{I}_i|$,

4. *The network has a stronger intra-block connection than the inter-block one:*

$$\min_i B_{ii} - 2\rho \max_i \sum_{j\neq i} B_{ij} \geq \Delta\,, \tag{13}$$

   for some $\Delta > 0$, where $B = Q \odot W$. ($\odot$ is the Hadamard product)

The first assumption ensures the network $T_{yu}(s)$ and our approximation model $\hat{T}_k(s)$ are stable. The second assumption ensures that our low-rank approximation $T_k(s)$ in Theorem 1 is valid on the interval of our interest $(-j\eta, +j\eta)$. The third assumption ensures our problem is non-degenerate: if the size of one block is too small, the network effectively has $k-1$ clusters. Such an assumption is standard in analyzing the consistency of spectral clustering algorithms on stochastic block models (Lyzinski et al., 2014; Ahn et al., 2018). Lastly, since we are interested in networks containing multiple groups of nodes such that within each group, the nodes are tightly-connected while between groups, the nodes are weakly-connected, the fourth assumption formally characterizes such a property.

In our algorithm, a spectral clustering algorithm is used to find a partition $\{\mathcal{I}_i\}_{i=1}^k$ that is used for aggregating node dynamics and constructing the reduced network. Ideally, we want some consistency property on the obtained partition.

Assumption 2 *Given L sampled from a weighted stochastic block model $(\{\mathcal{I}_i\}_{i=1}^k, Q, W)$ satisfying Assumption 1, we have an asymptotically consistent spectral clustering algorithm in Algorithm 1: For any $\delta > 0$, there exists $\tilde{N}(\delta)$ such that for network with size $n > \tilde{N}(\delta)$, the spectral clustering algorithm on L returns the true $\{\mathcal{I}_i\}_{i=1}^k$ partition with probability at least $1 - \delta$.*

Formally justifying this assumption for some spectral clustering algorithms is an interesting future research topic. Nonetheless, such a consistency result has been studied for spectral clustering algorithms on the adjacency matrix from the stochastic block model (Lyzinski et al., 2014) and the weighted stochastic block model (Ahn et al., 2018).

With these assumptions, we have the following theorem regarding the error bound.

Theorem 4 *Consider the network model $(G(s), L, f(s))$ with L sampled from a weighted stochastic block model $(\{\mathcal{I}_i\}_{i=1}^k, Q, W)$. If Assumption 1 and Assumption 2 hold, then Algorithm 1 returns a $\hat{T}_k(s)$ such that*

1. *$\|T_{yu}(s)\|_{\mathcal{H}_\infty} \leq \gamma$, $\|\hat{T}_k(s)\|_{\mathcal{H}_\infty} \leq \gamma$;*
2. *For any $\eta > 0$, $0 < \delta < 1$ and $\epsilon > 0$, there exists an $N(\delta, \epsilon, \tilde{N}(\delta/2), \gamma, M(\eta), F_l(\eta), \rho, Q, W)$ such that for network with size $n \geq N$, with probability at least $1 - \delta$, we have*

$$\sup_{s \in (-j\eta, +j\eta)} \|T_{yu}(s) - \hat{T}_k(s)\| \leq \epsilon \,. \tag{14}$$

Proof [Proof Sketch] For the stability of $T_{yu}(s)$ and $\hat{T}_k(s)$, the proof is similar to the one in Min et al. (2021) and uses the assumption $g_i(s)$ are output strictly passive and $f(s)$ is positive real. The error bound relies on that the sampled Laplacian matrix L is close to one that is easy to analyze: Let A_{blk} be the expected value of the adjacency matrix A from the block model, and we can construct a Laplacian matrix $L_{\text{blk}} = D_{A_{\text{blk}}} - A_{\text{blk}}$, $D_{A_{\text{blk}}} = \text{diag}\{A_{\text{blk}}\mathbb{1}\}$. L_{blk} has (by (12)(13) in Assumption 1) all the desired properties: 1) $\lambda_{k+1}(L_{\text{blk}})$ grows linearly in network size n; 2) L_{blk} is k-block-ideal. Min and Mallada (2022a) has shown that under the weighted stochastic block model, $\|L - L_{\text{blk}}\| \sim \mathcal{O}_p(\sqrt{n \log n})$, which is sufficient to show that 1) $\lambda_{k+1}(L) \sim \Omega_p(n)$ by Weyl's inequality (Horn and Johnson, 2012); 2) L is approximately k-block-ideal by Davis-Khan theorem (Yu et al., 2014). The former shows that the error between $T_{yu}(s)$ and $T_k(s)$ is small w.h.p. by Theorem 1 and the latter ensures the error between $T_k(s)$ and $\hat{T}_k(s)$ is small w.h.p. We refer the readers to Min and Mallada (2022b) for the full proof. ■

Theorem 4 shows that our algorithms perform well for large networks with multiple coherent clusters, it also implies that the collective dynamic behavior of such networks can be modeled as a structured reduced network. This suggests a new avenue for data-driven system identification for such networks where only the reduced network model is learned from the data collected from the network.

5. Numerical Experiments

The frequency response of synchronous generator (including grid-forming inverters) networks, linearized at its equilibrium point (Zhao et al., 2013), can be modeled exactly as the network model in Fig 1 with $f(s) = \frac{1}{s}$ and second order node dynamics $g_i(s)$. We validate our algorithm with a

synthetic test case, where the coefficients of generator dynamics are randomly sampled. The network adjacency matrix A is sampled from our weighted stochastic block model $(\{\mathcal{I}_i\}_{i=1}^k, Q, W)$ with $k = 3$, and

$$\begin{bmatrix} |\mathcal{I}_1| & 0 & 0 \\ 0 & |\mathcal{I}_2| & 0 \\ 0 & 0 & |\mathcal{I}_3| \end{bmatrix} = \begin{bmatrix} 20 & 0 & 0 \\ 0 & 40 & 0 \\ 0 & 0 & 20 \end{bmatrix}, Q = \begin{bmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.8 & 0.1 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}, W = \begin{bmatrix} 20 & 0.4 & 0.8 \\ 0.4 & 20 & 0.7 \\ 0.8 & 0.7 & 20 \end{bmatrix}. \quad (15)$$

We use the spectral clustering algorithm proposed in Bach and Jordan (2004). Since the network size is not sufficiently large for the algorithm to return a true partition with high probability, when we run the experiments with multiple random seeds, we see a small fraction of the runs in which the algorithm fails to cover the true partition. For the case when the spectral clustering algorithm succeeds, we inject a step disturbance $u_2(t) = \chi(t)$ at the second node of the network and plot the step response of $T_{yu}(s)$ in Fig 3, along with the response $\hat{y}$ of our approximate model $\hat{T}_3(s)$ from Algorithm 1. There is a clear difference between the dynamical response of generators from different groups, and the aggregate responses $\hat{y}$ capture such difference while providing a good approximation to the actual node responses. Due to space constraints, we only present the result of running Algorithm 1 on one instance of the randomly generated networks, but the results are consistent across multiple runs as long as the spectral clustering succeeds.

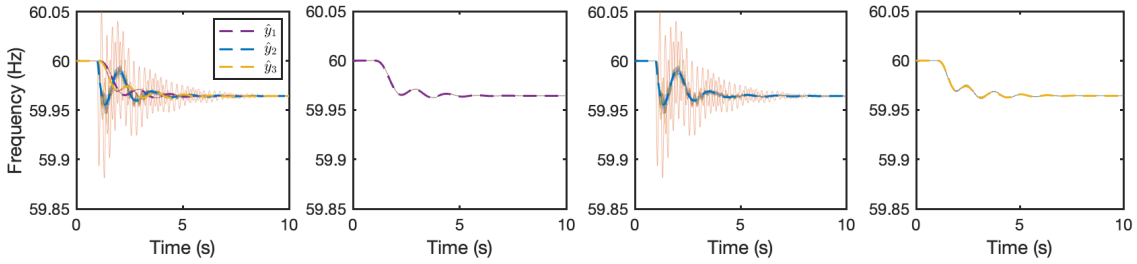

Figure 3: Left most plot shows the step response of $T_{yu}(s)$ (solid lines) and $\hat{T}_k(s)$ (dashed lines) from algorithm 1. The three plots on the right show the response for each identified group $\mathcal{I}_i$. The node injected with step disturbance is in the group 2.

6. Conclusion

In this paper, we propose a structure-preserving model-reduction methodology for large-scale dynamic networks based on a recent frequency-domain characterization of coherent dynamics in networked systems. Our analysis shows that networks with multiple coherent groups can be well approximated by a reduced network of the same size as the number of coherent groups, and we provide an upper bound on the approximation error when the network graph is randomly generated from a weight stochastic block model. We believe our proposed model can be applied to power networks for studying the inter-area oscillation in the frequency response and allows new control designs based on the reduced network.

Acknowledgments

This work was supported by the NSF HDR TRIPODS Institute for the Foundations of Graph and Deep Learning (NSF grant 1934979), the ONR MURI Program (ONR grant 503405-78051), the NSF CPS Program (NSF grant 2136324), and the NSF CAREER Program (NSF grant 1752362). The authors thank Professor Steven Low for the insightful discussion that inspired this work. The authors also thank Bala Kameshwar Poolla and Dhananjay Anand for carefully reading our manuscript and for providing many insightful comments and suggestions.

References

Kwangjun Ahn, Kangwook Lee, and Changho Suh. Hypergraph spectral clustering in the weighted stochastic block model. *IEEE Journal of Selected Topics in Signal Processing*, 12(5):959–974, 2018.

Francis R Bach and Michael I Jordan. Learning spectral clustering. In *Advances in Neural Information Processing Systems*, pages 305–312, 2004.

Bassam Bamieh, Mihailo R Jovanovic, Partha Mitra, and Stacy Patterson. Coherence in large-scale networks: Dimension-dependent limitations of local feedback. *IEEE Trans. Automat. Contr.*, 57 (9):2235–2249, 2012.

Joe H Chow. *Time-scale modeling of dynamic networks with applications to power systems*. Springer, 1982.

Etinosa Ekomwenrenren, Zhiyuan Tang, John W. Simpson-Porco, Evangelos Farantatos, Mahendra Patel, and Hossein Hooshyar. Hierarchical coordinated fast frequency control using inverter-based resources. *IEEE Transactions on Power Systems*, 36(6):4992–5005, 2021. doi: 10.1109/TPWRS.2021.3075641.

Julian Fritzsch and Philippe Jacquod. Long wavelength coherency in well connected electric power networks. *IEEE Access*, 10:19986–19996, 2022. doi: 10.1109/ACCESS.2022.3151687.

Verena Häberle, Eduardo Prieto-Araujo, Ali Tayyebi, and Florian Dörfler. Grid-forming control design of dynamic virtual power plants. *arXiv preprint arXiv:2202.02057*, 2022.

Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, New York, NY, USA, 2nd edition, 2012. ISBN 0521548233, 9780521548236.

Yan Jiang, Andrey Bernstein, Petr Vorobev, and Enrique Mallada. Grid-forming frequency shaping control in low inertia power systems. *IEEE Control Systems Letters (L-CSS)*, 5(6):1988–1993, 12 2021a. doi: 10.1109/LCSYS.2020.3044551.

Yan Jiang, Richard Pates, and Enrique Mallada. Dynamic droop control in low inertia power systems. *IEEE Transactions on Automatic Control*, 66(8):3518–3533, 8 2021b. doi: 10.1109/TAC.2020.3034198.

Sanjay Lall, Petr Krysl, and Jerrold E Marsden. Structure-preserving model reduction for mechanical systems. *Physica D: Nonlinear Phenomena*, 184(1):304–318, 2003. ISSN 0167-2789. doi:

https://doi.org/10.1016/S0167-2789(03)00227-6. Complexity and Nonlinearity in Physical Systems – A Special Issue to Honor Alan Newell.

Ren-Cang Li and Zhaojun Bai. Structure-preserving model reduction. In Jack Dongarra, Kaj Madsen, and Jerzy Waśniewski, editors, *Applied Parallel Computing. State of the Art in Scientific Computing*, pages 323–332, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg. ISBN 978-3-540-33498-9.

Vince Lyzinski, Daniel L Sussman, Minh Tang, Avanti Athreya, and Carey E Priebe. Perfect clustering for stochastic blockmodel graphs via adjacency spectral embedding. *Electronic journal of statistics*, 8(2):2905–2922, 2014.

H. Min and E. Mallada. Dynamics concentration of large-scale tightly-connected networks. In *IEEE 58th Conf. on Decision and Control*, pages 758–763, 2019.

H. Min, F. Paganini, and E. Mallada. Accurate reduced-order models for heterogeneous coherent generators. *IEEE Contr. Syst. Lett.*, 5(5):1741–1746, 2021. doi: 10.1109/LCSYS.2020.3043733.

Hancheng Min and Enrique Mallada. Spectral clustering and model reduction for weakly-connected coherent network systems. *arXiv preprint arXiv:2209.13701*, 2022a.

Hancheng Min and Enrique Mallada. Learning coherent clusters in weakly-connected network systems. *arXiv preprint arXiv:2211.15301*, 2022b.

Hancheng Min, Richard Pates, and Enrique Mallada. Coherence and concentration in tightly-connected networks. *arXiv preprint arXiv:2101.00981*, 2021.

R. Olfati-Saber and R.M. Murray. Consensus problems in networks of agents with switching topology and time-delays. *IEEE Trans. Automat. Contr.*, 49(9):1520–1533, 2004. doi: 10.1109/TAC.2004.834113.

G.N. Ramaswamy, G.C. Verghese, L. Rouco, C. Vialas, and C.L. DeMarco. Synchrony, aggregation, and multi-area eigenanalysis. *IEEE Transactions on Power Systems*, 10(4):1986–1993, 1995. doi: 10.1109/59.476067.

G.N. Ramaswamy, L. Rouco, O. Fillatre, G.C. Verghese, P. Panciatici, B.C. Lesieutre, and D. Peltier. Synchronic modal equivalencing (sme) for structure-preserving dynamic equivalents. *IEEE Transactions on Power Systems*, 11(1):19–29, 1996. doi: 10.1109/59.485977.

Diego Romeres, Florian Dörfler, and Francesco Bullo. Novel results on slow coherency in consensus and power networks. In *2013 European Control Conference (ECC)*, pages 742–747, 2013. doi: 10.23919/ECC.2013.6669400.

Bita Safaee and Serkan Gugercin. Structure-preserving model reduction of parametric power networks. In *2021 American Control Conference (ACC)*, pages 1824–1829, 2021. doi: 10.23919/ACC50511.2021.9483168.

Ilya Tyuryukanov, Marjan Popov, Mart A. M. M. van der Meijden, and Vladimir Terzija. Slow coherency identification and power system dynamic model reduction by using orthogonal structure of electromechanical eigenvectors. *IEEE Transactions on Power Systems*, 36(2):1482–1492, 2021. doi: 10.1109/TPWRS.2020.3009628.

Yi Yu, Tengyao Wang, and Richard J. Samworth. A useful variant of the davis–kahan theorem for statisticians, 2014.

Changhong Zhao, Ufuk Topcu, Na Li, and S Low. Power system dynamics as primal-dual algorithm for optimal load control. *arXiv preprint arXiv:1305.0585*, 2013.