

University of Groningen

Reduced Order Modeling of Diffusively Coupled Network Systems

Cheng, Xiaodong; Yu, Lanlin; Ren, Dingchao; Scherpen, Jacquélien M.A.

Published in:
IEEE Transactions on Automatic Control

DOI:
[10.1109/TAC.2022.3200886](https://doi.org/10.1109/TAC.2022.3200886)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2023

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Cheng, X., Yu, L., Ren, D., & Scherpen, J. M. A. (2023). Reduced Order Modeling of Diffusively Coupled Network Systems: An Optimal Edge Weighting Approach. *IEEE Transactions on Automatic Control*, 68(7), 4233-4240. <https://doi.org/10.1109/TAC.2022.3200886>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).





The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Reduced Order Modeling of Diffusively Coupled Network Systems: An Optimal Edge Weighting Approach

Xiaodong Cheng , Member, IEEE, Lanlin Yu , Member, IEEE, Dingchao Ren ,
and Jacquélien M.A. Scherpen , Fellow, IEEE

Abstract—This article studies reduced-order modeling of dynamic networks with strongly connected topology. Given a graph clustering of an original large-scale network, we construct a quotient graph with less number of vertices, where the edge weights are parameters to be determined. The model of a reduced network is thereby obtained with parameterized system matrices, and then, an edge weighting procedure is devised, aiming to select an optimal set of edge weights to minimize the approximation error between the original and the reduced-order network models in terms of \mathcal{H}_2 -norm. The effectiveness of the proposed method is illustrated by a numerical example.

Index Terms—Interconnected systems, linear systems, network systems, optimization, reduced order systems.

I. INTRODUCTION

With the growing complexity of interconnected dynamic systems, the importance of understanding and managing dynamic networks has been widely recognized. An important class of dynamic networks is given by the so-called diffusively coupled networks, which are commonly used for describing diffusion processes, e.g., information or energy spreading in networks. The examples can be found in, e.g., vehicle formations, electrical networks, and synchronization in sensor networks, see [1], [2], and [3]. The spatial structures of such systems usually result in high-dimensional models that hinder analysis, control, and optimization. Thus, it is desirable to simplify the structure of a complex network without a significant loss of accuracy.

Different from model reduction problems for other types of dynamic systems, the one considered in this article puts emphasis on the preservation of the network structure, which is necessary for applications, e.g., distributed controller design and sensor allocation [4], [5], [6]. Conventional model reduction methods, e.g., balanced truncation and

Manuscript received 27 March 2021; revised 17 January 2022; accepted 7 August 2022. Date of publication 22 August 2022; date of current version 28 June 2023. The work of Lanlin Yu was supported in part by the National Natural Science Foundation of China under Grant 62003276 and in part by the Fundamental Research Funds for the Central Universities of China under Grant JZ2022HG7B0299. Recommended by Associate Editor M.K. Camlibel. (Corresponding author: Xiaodong Cheng.)

Xiaodong Cheng is with the Department of Engineering, University of Cambridge, CB2 1PZ Cambridge, U.K. (e-mail: x.cheng.auto@gmail.com).

Lanlin Yu and Jacquélien M.A. Scherpen are with the School of Electrical Engineering and Automation, Hefei University of Technology, Hefei 230009, China, and also with the Institute of Advanced Technology, Westlake Institute for Advanced Study, Westlake University, Hangzhou 310024, China (e-mail: yulanlin1992@gmail.com).

Dingchao Ren is with the Department of Automation, University of Science and Technology of China, Hefei 230026, China (e-mail: rdc34679@mail.ustc.edu.cn).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TAC.2022.3200886>.

Digital Object Identifier 10.1109/TAC.2022.3200886

moment matching, merely focus on approximating the input–output behavior of a given dynamical system [7], while the preservation of the network structure is barely guaranteed. Although a generalized balanced truncation approach in [8] is able to construct an accurate reduced-order model with a network interpretation, the relation between the original and obtained new typologies is not clear. Singular perturbation approximation has been applied to the reduction of electric circuits [9] and chemical reaction networks [10]. This class of methods mainly relies on time-scale separation of the states in an autonomous system, while the external inputs are not considered explicitly. Besides, the resulting reduced topology hardly retains sparsity.

Recently, clustering-based methods have become the mainstream methodology for reducing network systems, see, e.g., [11], [12], [13], [14], [15], [16], [17], [18], and [19]. This class of methods has a clear advantage in retaining the consensus [12], [19] or positivity [15] in reduced-order models. The reduction procedure can be implemented via the Petrov–Galerkin projection framework, where the projection matrix is formed based on the vertex clusters. However, all the current clustering-based methods only focus on finding suitable clusters. After clusters are found, reduced-order network models are then directly determined by the projection framework, while the freedom to construct a reduced-order network model with higher accuracy is overlooked.

In this article, we will explore the latter freedom and provide a novel method for reduced-order modeling of directed networks. We do not aim to find an optimal clustering. Instead, we assume that the clustering of a network is given, which leads to a *quotient graph*. A parameterized reduced-order model is thereby established with the edge weights to be optimized. Then, the major problem in this article follows: How to tune the edge weights in the parameterized reduced-order model to minimize the approximation error?

This problem can be formulated as an optimization problem with the objective to minimize the \mathcal{H}_2 -norm of the reduction error between original and reduced network systems, in which the edge weights of the reduced network are variables to be optimized. This edge weighting problem is subject to a bilinear matrix inequality constraint, which is computationally expensive. Therefore, we devise a novel edge weighting algorithm based on the *convex–concave decomposition*. Such decomposition presented in [20] is classified as a local optimization method, which aims at finding a local optimum based on solving a sequence of convex semidefinite programming problems. With this decomposition, the nonconvex constraint in our problem can be linearized as a convex one in the form of a linear matrix inequality (LMI), so that an iterative scheme is implemented to search for a set of optimal weights. The convergence of this algorithm is theoretically ensured, and thus at least a local optimum can be reached. Moreover, we initialize the edge weights as the outcome of clustering-based projection, such that the obtained reduced-order network model is guaranteed a better approximation accuracy than the clustering-based projection methods. Compared with the preliminary results presented in [21], this article

provides a generalization to directed networks and improve the optimization scheme, which now guarantees the convergence.

The rest of this article is organized as follows. In Section II, we recap some preliminaries on graph theory and introduce the problem setup. In Section III, the parameterized reduced-order model is formulated, and an edge weighting algorithm is proposed to minimize the approximation error. In Section IV, the proposed method is illustrated by an example. Finally, Section V concludes this article.

Notation: The symbols \mathbb{R} and \mathbb{R}_+ denote the set of real numbers and positive real numbers, respectively. Let S^n be the set of real symmetric matrices of size $n \times n$. I_n is the identity matrix of size n , and $\mathbf{1}_n$ represents the vector in \mathbb{R}^n of all ones. The cardinality of a set S is denoted by $|S|$.

II. PRELIMINARIES AND PROBLEM SETTING

This section provides necessary definitions and concepts in graph theory used in this article, and we refer [22] for more details. The model of a dynamical network is then introduced and the model reduction problem is formulated.

A. Graph Theory

A directed graph $\mathcal{G} := (\mathcal{V}, \mathcal{E})$ consists of a finite and nonempty node set $\mathcal{V} := \{1, 2, \dots, n\}$ and an edge set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$. Each element in \mathcal{E} is an ordered pair of \mathcal{V} , and if $(i, j) \in \mathcal{E}$, we say that the edge is directed from vertex i to vertex j . A directed graph \mathcal{G} is called *simple*, if \mathcal{G} does not contain self-loops (i.e., \mathcal{E} does not contain any edge of the form $(i, i) \forall i \in \mathcal{V}$), and there exists only one edge directed from i to j , if $(i, j) \in \mathcal{E}$. Next, we recap several important matrices for characterizing a directed simple graph from [22]. The *incidence matrix* $\mathcal{B} \in \mathbb{R}^{n \times |\mathcal{E}|}$ is defined by

$$\mathcal{B}_{ij} = \begin{cases} +1 & \text{if edge } j \text{ is directed from vertex } i \\ -1 & \text{if edge } j \text{ is directed to vertex } i \\ 0 & \text{otherwise.} \end{cases}$$

If each edge is assigned a positive value (weight), then the *weighted adjacency matrix* of \mathcal{G} , denoted by $\mathcal{A} \in \mathbb{R}^{n \times n}$, is defined such that $\mathcal{A}_{ij} \in \mathbb{R}_+$ denotes the weight of edge $(i, j) \in \mathcal{E}$, and $\mathcal{A}_{ij} = 0$ if $(i, j) \notin \mathcal{E}$. In the case of a simple graph, \mathcal{A} is a binary matrix with zeros on its diagonal. Then, the *Laplacian matrix* $L \in \mathbb{R}^{n \times n}$ of the graph \mathcal{G} is defined as

$$L_{ij} = \begin{cases} \sum_{k=1, k \neq i}^n \mathcal{A}_{ik} & \text{if } i = j \\ -\mathcal{A}_{ij} & \text{otherwise.} \end{cases} \quad (1)$$

Clearly, $L\mathbf{1} = 0$. When \mathcal{G} is strongly connected, the diagonal entries of L are strictly positive, and the off-diagonal entries are nonpositive. Alternatively, we can characterize the Laplacian matrix using the incidence matrix of \mathcal{G} as $L = \mathcal{B}_0 W \mathcal{B}^\top$, where \mathcal{B}_0 is a binary matrix obtained by replacing all “−1” entries in the incidence matrix \mathcal{B} with zeros, and

$$W := \text{Diag}(w), \text{ with } w = [w_1 \ w_2 \ \dots \ w_{|\mathcal{E}|}]^\top$$

and w_k the positive weight associated to the k th edge, for all $k \in \{1, 2, \dots, |\mathcal{E}|\}$. Note that the adjacency and Laplacian matrices in this article are defined as transposed compared with [23], where the direction of edges corresponds to the flows, rather than the dependency in the network.

For a vertex in a weighted graph, the *indegree* and *outdegree* of vertex i are computed as $\sum_{j \in \mathcal{V}} \mathcal{A}_{ji}$ and $\sum_{j \in \mathcal{V}} \mathcal{A}_{ij}$, respectively. A strongly connected graph \mathcal{G} is called *balanced* if the indegree and outdegree of each vertex in \mathcal{G} is equal. From (1), the following lemma is immediate.

Lemma 1: A weighted strongly connected graph \mathcal{G} is balanced if and only if one of the following conditions hold.

- 1) The edge weights of \mathcal{G} satisfies $\mathcal{B}w = 0$.
- 2) The Laplacian matrix of \mathcal{G} satisfies $\mathbf{1}^\top L = 0$.

The strong connectivity implies that there is only one zero eigenvalue of L , and the balance of \mathcal{G} then indicates that both the row and column sums of L are zero.

Remark 1: Undirected graphs can be viewed as special balanced directed graphs with bidirectional edges. The Laplacian matrix of an undirected graph is $L = \mathcal{B}W\mathcal{B}^\top$, where \mathcal{B} is an incidence matrix obtained by assigning an arbitrary orientation to each edge of the undirected graph, and W is a positive diagonal matrix representing edge weights.

Next, we recapitulate the notion of graph clustering, whose concept can be found in, e.g., [12], [13], [14], [15], [16], and [17].

Definition 1: Let $\mathcal{G} := (\mathcal{V}, \mathcal{E})$ be a directed graph. Then, a graph clustering is a partition of \mathcal{V} into r nonempty disjoint subsets $\mathcal{C}_1, \mathcal{C}_2, \dots, \mathcal{C}_r$ covering all the elements in \mathcal{V} , where \mathcal{C}_i is called a cluster of \mathcal{G} .

Let $\{\mathcal{C}_1, \mathcal{C}_2, \dots, \mathcal{C}_r\}$ be a clustering of \mathcal{G} with n vertices. This graph clustering can be characterized by a binary characteristic matrix $\Pi \in \mathbb{R}^{n \times r}$, whose rows and columns are corresponding to the vertices and clusters, respectively:

$$\Pi_{ij} := \begin{cases} 1 & \text{if vertex } i \in \mathcal{C}_j \\ 0 & \text{otherwise.} \end{cases}$$

Remark 2: Note that all the clusters are nonoverlapping, i.e., each vertex cannot be assigned to distinct clusters. Therefore, each row of the characteristic matrix Π only has one nonzero element. Specifically, we have

$$\Pi \mathbf{1}_r = \mathbf{1}_n \text{ and } \mathbf{1}_n^\top \Pi = [|\mathcal{C}_1|, |\mathcal{C}_2|, \dots, |\mathcal{C}_r|]. \quad (2)$$

Given a graph clustering $\{\mathcal{C}_1, \mathcal{C}_2, \dots, \mathcal{C}_r\}$ of a graph \mathcal{G} , the *quotient graph* $\hat{\mathcal{G}}$ is r -vertex directed graph obtained by aggregating all the vertices in each cluster as a single vertex, while retaining connections between clusters and ignoring the edges within clusters. Specifically, if there is an edge (i, j) in \mathcal{G} with vertices i, j in the same cluster, then it will not be presented as an edge in $\hat{\mathcal{G}}$. If there exists an edge (i, j) with $i \in \mathcal{C}_k$ and $j \in \mathcal{C}_l$, then there is an edge (k, l) in $\hat{\mathcal{G}}$.

B. Problem Setup

In this article, we consider a network system evolving over a directed graph \mathcal{G} , which is simple, weighted, and strongly connected. The dynamics of each vertex is governed by

$$\dot{x}_i(t) = - \sum_{j=1}^n \mathcal{A}_{ij} [x_i(t) - x_j(t)] + \sum_{k=1}^p f_{ik} u_k(t) \quad (3)$$

where $x_i(t) \in \mathbb{R}$ is the state of vertex i , and \mathcal{A}_{ij} is the (i, j) th entry of the adjacency matrix of \mathcal{G} , representing the strength of the coupling between vertices i and j . $u_k(t) \in \mathbb{R}$ is the external input, and $f_{ik} \in \mathbb{R}$ is the gain of the k th input acting on vertex i , which is zero if and only if u_j has no effect on vertex i . Let $F \in \mathbb{R}^{n \times p}$ be the matrix such that $F_{ij} = f_{ij}$. We then present the dynamics of the overall network in a compact form as

$$\Sigma : \begin{cases} \dot{x}(t) = -Lx(t) + Fu(t) \\ y(t) = Hx(t) \end{cases} \quad (4)$$

with $x(t) := [x_1, x_2, \dots, x_n]^\top \in \mathbb{R}^n$ and $u := [u_1, u_2, \dots, u_p]^\top \in \mathbb{R}^p$. The vector $y \in \mathbb{R}^q$ collects the outputs of the network, and H is the output matrix.

This article aims for structure-preserving model reduction of diffusively coupled networks in form of (4), and the reduced-order model not only approximates the input–output mapping of the original network system with a certain accuracy but also inherits an interconnection structure with diffusive couplings. To this end, we adopt graph clustering to build up a reduced-order network model. Specifically, the problem addressed in this article is as follows.

Problem 1: Given a network system Σ as in (4) defined on a directed graph \mathcal{G} and a graph clustering $\{\mathcal{C}_1, \mathcal{C}_2, \dots, \mathcal{C}_r\}$ of \mathcal{G} , find a reduced-order model

$$\hat{\Sigma} : \begin{cases} \dot{\hat{x}}(t) = -\hat{L}\hat{x}(t) + \hat{F}u(t) \\ \hat{y}(t) = \hat{H}\hat{x}(t) \end{cases} \quad (5)$$

with $\hat{x} \in \mathbb{R}^r$, $r \ll n$, such that $\hat{L} \in \mathbb{R}^{r \times r}$ is a Laplacian matrix associated with the quotient graph of \mathcal{G} , $\hat{F} \in \mathbb{R}^{r \times p}$ and $\hat{H} \in \mathbb{R}^{q \times r}$ are reduced input and output matrices, and the reduction error $\|\Sigma - \hat{\Sigma}\|_{\mathcal{H}_2}$ is minimized.

Remark 3: It is not necessary to construct a reduced-order network model with graph clustering, while we consider the clustering framework in order to capture the interconnection structure among the clusters in the original graph. This is done through \hat{L} , which represents a quotient graph. Moreover, different from the existing clustering-based model reduction methods that aim to find suitable graph clustering for a small reduction error, Problem 1 is not to seek for clusters of \mathcal{G} . Instead, we focus on how to establish a “good” reduced-order model with given clusters. Thus, it is an essentially different problem from, e.g., [14], [15], [17], and [19], and we do not apply the Petrov–Galerkin projection framework.

III. MAIN RESULT

A novel model reduction approach for network systems is presented in this section. First, a parameterized model of a reduced network is constructed on the basis of graph clustering. Then, we compute the edge weight parameters in an optimal fashion such that the \mathcal{H}_2 -norm of approximation error is minimized.

A. Parameterized Reduced-Order Network Model

Given a graph clustering of the original network, we present a parameterized model for the reduced network, whose interconnection topology is determined by the clustering. An important property of this parameterized model is that it guarantees the boundedness of the reduction error $\|\Sigma - \hat{\Sigma}\|_{\mathcal{H}_2}$ for all positive edge weights.

To derive a parameterized reduced-order network model with such a property, we first convert the system (4) to its *balanced graph representation* as follows.

Lemma 2: If the underlying graph of Σ in (4) is strongly connected, then there exists a diagonal $M \in \mathbb{R}^{n \times n}$ with positive diagonal entries such that Σ is equivalent to

$$\begin{cases} M\dot{x}(t) = -L_b x(t) + F_b u(t) \\ y(t) = Hx(t) \end{cases} \quad (6)$$

where $F_b = MF$, and $L_b = ML$ is the Laplacian of a balanced graph.

The proof follows directly from [19]. Next, we establish a reduced-order model using the representation (6) to guarantee a bounded reduction error $\|\Sigma - \hat{\Sigma}\|_{\mathcal{H}_2}$.

Let \mathcal{G}_b be the balanced graph of \mathcal{G} . Note that \mathcal{G} and \mathcal{G}_b have the same incidence matrix B . Given a graph clustering $\{\mathcal{C}_1, \mathcal{C}_2, \dots, \mathcal{C}_r\}$, we obtain the *quotient graph* $\hat{\mathcal{G}}_b$ with r vertices. Let \hat{B} be the incidence matrix of the quotient graph $\hat{\mathcal{G}}_b$, and it can be verified that \hat{B} is obtained by removing all the zero columns and repeated columns of $\Pi^T B$, where

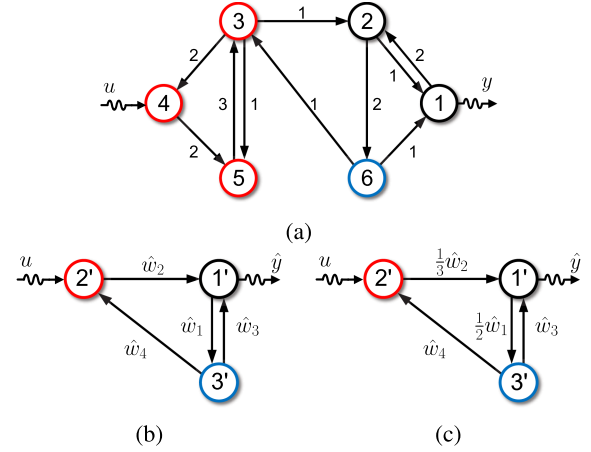


Fig. 1. (a) Directed balanced network \mathcal{G}_b consisting of six vertices, in which vertex 4 is controlled and vertex 1 is measured. The vertices in different clusters are indicated by distinct colors. (b) Quotient graph $\hat{\mathcal{G}}_b$ consisting of three vertices, where the edge weights are parameters to be determined. The quotient graph is balanced, when the constraints $\hat{w}_3 = \hat{w}_1 - \hat{w}_2$ and $\hat{w}_4 = \hat{w}_2$ are imposed. (c) Resulting reduced graph $\hat{\mathcal{G}}$.

B is the incidence matrix of \mathcal{G}_b (or \mathcal{G}). Furthermore, we denote

$$\hat{W} = \text{Diag}(\hat{w}), \text{ with } \hat{w} = [\hat{w}_1 \ \hat{w}_2 \ \dots \ \hat{w}_m]^\top \quad (7)$$

as the edge weight matrix of $\hat{\mathcal{G}}_b$, where $\hat{w}_k \in \mathbb{R}_+$, and m is number of edges in $\hat{\mathcal{G}}_b$. In order to maintain $\hat{\mathcal{G}}_b$ as a balanced graph, we impose the constraint on its edge weights as

$$\hat{B}\hat{w} = 0 \quad (8)$$

according to Lemma 1. Thereby, the dynamics on the balanced quotient graph $\hat{\mathcal{G}}_b$ is then obtained as

$$\begin{cases} \hat{M}\dot{\hat{x}}(t) = -\hat{L}_b(\hat{W})\hat{x}(t) + \hat{F}_b u(t) \\ \hat{y}(t) = \hat{H}\hat{x}(t) \end{cases} \quad (9)$$

with the reduced matrices

$$\begin{aligned} \hat{M} &= \Pi^T M \Pi, \quad \hat{L}_b(\hat{W}) = \hat{B}_0 \hat{W} \hat{B}^\top \\ \hat{F}_b &= \Pi^T F_b, \quad \text{and } \hat{H} = H \Pi \end{aligned} \quad (10)$$

where \hat{B}_0 is the binary matrix obtained by replacing all the “−1” entries with zeros in \hat{B} , and $\hat{M}_b \in \mathbb{R}^{r \times r}$, $\hat{F}_b \in \mathbb{R}^{r \times p}$, and $\hat{H} \in \mathbb{R}^{q \times r}$ are reduced matrices determined by the given clustering of \mathcal{G}_b . Since the graph clustering is given, i.e., Π is known, the only parameters to be decided are the weights in \hat{W} , which satisfy the constraint (8).

From the reduced graph balanced representation (9), we immediately construct a parameterized reduced-order model in the form of (5) with the reduced matrices

$$\hat{L}(\hat{W}) = \hat{M}^{-1} \hat{B}_0 \hat{W} \hat{B}^\top, \quad \hat{F} = \hat{M}^{-1} \Pi^T F_b, \quad \hat{H} = H \Pi \quad (11)$$

where \hat{L} represents a reduced weighted graph $\hat{\mathcal{G}}$. In (11), only the weight matrix \hat{W} is to be determined, which is selected from the following set:

$$\mathcal{M} := \{W = \text{Diag}(\hat{w}) \mid \hat{w} \in \mathbb{R}_+^m, \hat{B}\hat{w} = 0\}. \quad (12)$$

In the following example, we demonstrate the parameterized modeling of a simplified dynamic network.

Example 1: Consider a network example with a directed graph as in [1], where the formation topology \mathcal{G} is depicted in Fig. 1(a). Clearly,

\mathcal{G} is balanced, i.e., $\mathcal{G} = \mathcal{G}_b$, with the incidence matrix

$$B = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ -1 & 1 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & -1 & 0 & -1 \\ 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}.$$

Suppose that each vehicle is modeled as a first-order integrator which has the identical mass, i.e., $M = I_6$. An external control u is applied on vertex 4, and the vertex 1 is measured as the output signal y . Then, the network model is obtained in the form of (6) with

$$L_b = \begin{bmatrix} 2 & -2 & 0 & 0 & 0 & 0 \\ -1 & 3 & 0 & 0 & 0 & -2 \\ 0 & -1 & 4 & -2 & -1 & 0 \\ 0 & 0 & 0 & 2 & -2 & 0 \\ 0 & 0 & -3 & 0 & 3 & 0 \\ -1 & 0 & -1 & 0 & 0 & 2 \end{bmatrix}, F_b = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

$$\text{and } H = [1 \ 0 \ 0 \ 0 \ 0 \ 0].$$

Consider a clustering of \mathcal{G}_b as $\mathcal{C}_1 = \{1, 2\}$, $\mathcal{C}_2 = \{3, 4, 5\}$, $\mathcal{C}_3 = \{6\}$, which gives the characteristic matrix as

$$\Pi = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}^\top$$

The topology of the quotient graph $\hat{\mathcal{G}}_b$ is shown in Fig. 1(b) with the incidence matrix

$$\hat{B} = \begin{bmatrix} 1 & -1 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ -1 & 0 & 1 & 1 \end{bmatrix}.$$

All the edge weights of $\hat{\mathcal{G}}_b$ are positive parameters to be determined, as labeled in Fig. 1(b), which leads to the parameterized Laplacian matrix as

$$\hat{L}_b(\hat{W}) = \begin{bmatrix} \hat{w}_1 & 0 & -\hat{w}_1 \\ -\hat{w}_2 & \hat{w}_2 & 0 \\ -\hat{w}_3 & -\hat{w}_4 & \hat{w}_3 + \hat{w}_4 \end{bmatrix}.$$

The weights satisfy the constraint $\hat{B}\hat{w} = 0$, namely, $\hat{w}_3 = \hat{w}_1 - \hat{w}_2$, $\hat{w}_4 = \hat{w}_2$, such that $\hat{\mathcal{G}}_b$ is balanced. The other matrices in the reduced-order model (9) are computed as

$$\hat{M} = \Pi^\top M \Pi = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \hat{F}_b = \Pi^\top F_b = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$\text{and } \hat{H} = H \Pi = [1 \ 0 \ 0].$$

Then, in the parameterized reduced-order model (5), we have

$$\hat{L}(\hat{W}) = \hat{M}^{-1} \hat{L}_b(\hat{W}) = \begin{bmatrix} \frac{1}{2} \hat{w}_1 & 0 & -\frac{1}{2} \hat{w}_1 \\ -\frac{1}{3} \hat{w}_2 & \frac{1}{3} \hat{w}_2 & 0 \\ -\hat{w}_3 & -\hat{w}_4 & \hat{w}_3 + \hat{w}_4 \end{bmatrix}$$

with $\hat{w}_3 = \hat{w}_1 - \hat{w}_2$, and $\hat{w}_4 = \hat{w}_2$. The corresponding reduced graph is depicted in Fig. 1(c), which is no longer balanced.

Remark 4: The physical interpretation of the reduced matrices in (10) are explained. \hat{M} is constructed such that the mass of vertex k in $\hat{\mathcal{G}}_b$ is equal to the mass sum of all the vertices in \mathcal{C}_k in \mathcal{G}_b . The expression of \hat{F}_b means that if a vertex in a cluster \mathcal{C}_k of \mathcal{G}_b is controlled by an external input, then vertex k in $\hat{\mathcal{G}}_b$ is also controlled. Analogously, \hat{H} indicates that a vertex k in $\hat{\mathcal{G}}_b$ is measured if there is a measurement taken from a vertex in \mathcal{C}_k .

With the reduced matrices in (11) and the constraint in (8), an important property of the reduced-order network model $\hat{\Sigma}$ is that it guarantees the \mathcal{H}_2 reduction error between the original system Σ in (4) and $\hat{\Sigma}$ is always bounded.

The computation of the reduction error amounts to find the \mathcal{H}_2 norm of the following error system:

$$G_e(s) = C_e(sI - A_e)^{-1} B_e \quad (13)$$

where

$$A_e = - \begin{bmatrix} L & 0 \\ 0 & \hat{L} \end{bmatrix}, B_e = \begin{bmatrix} F \\ \hat{F} \end{bmatrix}, C_e = [H \ -\hat{H}].$$

Note that A_e is not Hurwitz, since L and \hat{L} are both Laplacian matrices containing zero eigenvalues. Thus, $\|G_e(s)\|_{\mathcal{H}_2}$ cannot be calculated directly using the state space representation (13). Here, we employ the following matrices:

$$S_n = \begin{bmatrix} -I_{n-1} \\ \mathbf{1}_{n-1}^\top \end{bmatrix} \in \mathbb{R}^{n \times (n-1)}, S_r = \begin{bmatrix} -I_{r-1} \\ \mathbf{1}_{r-1}^\top \end{bmatrix} \in \mathbb{R}^{r \times (r-1)} \quad (14)$$

which are independent of system dynamics and satisfy $S_n^\top \mathbf{1}_n = 0$, and $S_r^\top \mathbf{1}_r = 0$. Consider their left pseudoinverses as $S_n^+ = (S_n^\top M^{-1} S_n)^{-1} S_n^\top M^{-1}$, $S_r^+ = (S_r^\top \hat{M}^{-1} S_r)^{-1} S_r^\top \hat{M}^{-1}$, respectively. Then, using the matrices in (14), we show the following result.

Lemma 3: Consider the network system Σ in (4) and the reduced-order network model $\hat{\Sigma}$ in (5) with matrices in (11). Then, $\Sigma - \hat{\Sigma} \in \mathcal{H}_2$ holds for all $\hat{W} \in \mathcal{M}$.

Proof: With S_n and S_r in (14), we construct a nonsingular $(n+r) \times (n+r)$ matrix as

$$U_e = \begin{bmatrix} \sigma_M^{-1} \mathbf{1}_n & 0 & M^{-1} S_n & 0 \\ 0 & \sigma_M^{-1} \mathbf{1}_r & 0 & \hat{M}^{-1} S_r \end{bmatrix} \quad (15)$$

where $\sigma_M := \mathbf{1}_n^\top M \mathbf{1}_n = \mathbf{1}_r^\top \Pi^\top M \Pi \mathbf{1}_r = \mathbf{1}_r^\top \hat{M} \mathbf{1}_r$. The inverse of U_e is given as

$$U_e^{-1} = \begin{bmatrix} \mathbf{1}_n^\top M & 0 \\ 0 & \mathbf{1}_r^\top \hat{M} \\ S_n^+ M & 0 \\ 0 & S_r^+ \hat{M} \end{bmatrix}.$$

Note that $L = M^{-1} L_b$ and $\hat{L} = \hat{M}^{-1} \hat{L}_b$, where both L_b and \hat{L}_b are the Laplacian matrices of balanced graphs, satisfying $\mathbf{1}_n^\top L_b = 0$, $L_b \mathbf{1}_n = 0$, and $\mathbf{1}_r^\top \hat{L}_b = 0$, $\hat{L}_b \mathbf{1}_r = 0$. Using these properties, we obtain

$$\begin{aligned} G_e(s) &= C_e U_e (sI - U_e^{-1} A_e U_e)^{-1} U_e^{-1} B_e \\ &= [\bar{C}_e \ C_e] \begin{bmatrix} I/s & 0 \\ 0 & (sI - A_e)^{-1} \end{bmatrix} \begin{bmatrix} \bar{B}_e \\ B_e \end{bmatrix} \\ &= 1/s \bar{C}_e \bar{B}_e + C_e (sI - A_e)^{-1} B_e \end{aligned} \quad (16)$$

where

$$A_e = - \begin{bmatrix} S_n^+ L_b M^{-1} S_n & 0 \\ 0 & S_r^+ \hat{L}_b \hat{M}^{-1} S_r \end{bmatrix} \quad (17)$$

$$B_e = \begin{bmatrix} S_n^+ F_b \\ S_r^+ \hat{F}_b \end{bmatrix}, C_e = [H M^{-1} S_n \ -\hat{H} \hat{M}^{-1} S_r]$$

$$\bar{B}_e := \begin{bmatrix} \mathbf{1}_n^\top F_b \\ \mathbf{1}_r^\top \hat{F}_b \end{bmatrix}, \text{ and } \bar{C}_e := \sigma_M^{-1} [H \mathbf{1}_n \ -\hat{H} \mathbf{1}_r]. \quad (18)$$

It follows from (2) that $\mathbf{1}_r^\top \hat{F}_b = \mathbf{1}_n^\top F_b$, and $\hat{H} \mathbf{1}_r = H \mathbf{1}_n$, which yield $\bar{C}_e \bar{B}_e = \sigma_M^{-1} (H \mathbf{1}_n \mathbf{1}_n^\top F_b - \hat{H} \mathbf{1}_r \mathbf{1}_r^\top \hat{F}_b) = 0$. Thus, (16) becomes

$$G_e(s) = C_e (sI - A_e)^{-1} B_e. \quad (19)$$

It is not hard to verify that both the matrices $-S_n^+ L_b M^{-1} S_n$ and $-S_r^+ \hat{L}_b \hat{M}^{-1} S_r$ are Hurwitz. Consequently, $G_e(s)$ in (19) is asymptotically stable, i.e., $\Sigma - \hat{\Sigma} \in \mathcal{H}_2$, for all $\hat{W} \in \mathcal{M}$.

Next, we discuss the consensus property of the reduced-order network (5) with the matrices in (11). Consensus is a typical property of diffusively coupled networks, and it implies the nodal states converge to a common value in the absence of the external input. More precisely, the network system in (4) reaches consensus if $\lim_{t \rightarrow \infty} [x_i(t) - x_j(t)] = 0$ holds for all $i, j \in \mathcal{V}$ and all initial conditions.

Proposition 1: Consider the network system Σ in (4) which reaches consensus. Then, the reduced-order model $\hat{\Sigma}$ in (5) also reaches consensus, for any clustering Π and $\hat{W} \in \mathcal{M}$.

Proof: It can be verified that the parameterized Laplacian matrix \hat{L} defined in (11) characterizes a strongly connected graph. Thus, \hat{L} has only one zero eigenvalue. Then, the proof immediately follows from [19], [22].

The parameterized modeling of the reduced dynamic network using the graph balanced representation in (6) guarantees the stability of the error system (13), whose \mathcal{H}_2 norm can be evaluated via the transfer function (19) with the Hurwitz matrix A_e . Note that in (19), the matrices S_n and S_r in (14) is only dependent on the sizes of the networks, and Π is known for a given graph clustering, then the weights in \hat{W} become the only unknown parameters to be determined in the follow-up procedure. In the following section, we aim for an optimal selection of the edges weights in the reduced network.

B. Optimal Edge Weighting

In this section, we aim for an optimization scheme for determining $\hat{W} \in \mathcal{M}$ that minimizes the approximation error $\|G_e(s)\|_{\mathcal{H}_2}$. Thereby, the following problem is addressed.

Problem 2: Consider the original network system Σ in (4). Given a graph clustering Π , find a $\hat{W} \in \mathcal{M}$ such that $\|G_e(s)\|_{\mathcal{H}_2}$ is minimized, where $\hat{\Sigma}$ is the reduced network model defined in (5) with the matrices (11).

To solve this problem, we apply an optimization technique based on the *convex-concave decomposition*, which can be implemented to search for a set of optimal weights iteratively. A fundamental step toward the implementation is to develop a necessary and sufficient condition for characterizing $\|G_e(s)\|_{\mathcal{H}_2}$, which leads to suitable constraints for the optimization problem.

Theorem 1: Given the network system Σ in (4). There exists a reduced-order network model $\hat{\Sigma}$ in (5) such that $\|G_e(s)\|_{\mathcal{H}_2} < \hat{\gamma}$, if and only if there exist matrices $\hat{Q} = \hat{Q}^\top > 0$ with $\hat{Q} \in \mathbb{R}^{(n+r-2) \times (n+r-2)}$, $\hat{R} = \hat{R}^\top > 0$ with $\hat{R} \in \mathbb{R}^{q \times q}$, $\hat{W} \in \mathcal{M}$, and $\hat{\delta} \in \mathbb{R}_+$, such that

$$\begin{bmatrix} \hat{Q}\bar{A} + \bar{A}^\top \hat{Q} & \hat{Q}\hat{B}_e & \hat{Q}E \\ B_e^\top \hat{Q} & -\hat{\delta}I & 0 \\ E^\top \hat{Q} & 0 & 0 \end{bmatrix} + \begin{bmatrix} -\bar{A}_r^\top \bar{A}_r & 0 & \bar{A}_r^\top \\ 0 & 0 & 0 \\ \bar{A}_r & 0 & -I \end{bmatrix} < 0 \quad (20)$$

$$\begin{bmatrix} \hat{Q} & \hat{\delta}C_e^\top \\ \hat{\delta}C_e & \hat{R} \end{bmatrix} > 0 \quad (21)$$

$$\text{tr}(\hat{R}) < \hat{\gamma} \quad (22)$$

where B_e, C_e are defined in (17), and

$$\begin{aligned} \bar{A} &= \begin{bmatrix} -S_n^+ L_b M^{-1} S_n & 0 \\ 0 & 0 \end{bmatrix}, \quad E = \begin{bmatrix} 0 & 0 \\ I & 0 \end{bmatrix} \\ \bar{A}_r &= \begin{bmatrix} 0 & -S_r^+ \hat{B}_0 \hat{W} \hat{B}^\top M^{-1} S_r \\ 0 & 0 \end{bmatrix}. \end{aligned} \quad (23)$$

Proof: Consider the error system $G_e(s)$ in (19), which is asymptotically stable. Following, e.g., [24], we have $\|G_e(s)\|_{\mathcal{H}_2}^2 < \gamma$, with $\gamma \in \mathbb{R}_+$, if and only if there exist matrices $Q = Q^\top > 0$ and $R = R^\top > 0$ such that

$$\begin{bmatrix} QA_e + A_e^\top Q & QB_e \\ B_e^\top Q & -I_p \end{bmatrix} < 0 \quad (24)$$

$$\begin{bmatrix} Q & C_e^\top \\ C_e & R \end{bmatrix} > 0 \quad (25)$$

$$\text{tr}(R) < \gamma \quad (26)$$

where A_e, B_e , and C_e are defined in (17).

In the following, we prove that the three inequalities are equivalent to (20), (21), and (22), respectively. First, it is not hard to verify that (24) is equivalent to

$$\begin{bmatrix} QA_e + A_e^\top Q & QB_e & QE \\ B_e^\top Q & -I & 0 \\ E^\top Q & 0 & -\delta I \end{bmatrix} < 0 \quad (27)$$

for a sufficiently large scalar $\delta \in \mathbb{R}_+$, where E is defined in (23). Consider a nonsingular matrix

$$T = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ -\bar{A}_r & 0 & I \end{bmatrix}.$$

Pre and postmultiplying by T^\top and T , respectively, (27) then becomes (20), where the equation $A_e = \bar{A} + E\bar{A}_r$, and the substitutions $\hat{\delta} = 1/\delta > 0$ and $\hat{Q} = (1/\delta)Q > 0$ are used.

Next, we observe that the following implications hold:

$$\begin{bmatrix} Q & C_e^\top \\ C_e & R \end{bmatrix} > 0 \Leftrightarrow \begin{bmatrix} \frac{1}{\hat{\delta}}\hat{Q} & C_e^\top \\ C_e & R \end{bmatrix} > 0 \Leftrightarrow \begin{bmatrix} \hat{Q} & \hat{\delta}C_e^\top \\ \hat{\delta}C_e & \hat{R} \end{bmatrix} > 0$$

$$\text{tr}(R) < \gamma \Leftrightarrow \frac{1}{\hat{\delta}}\text{tr}(\hat{R}) < \gamma \Leftrightarrow \text{tr}(\hat{R}) < \hat{\gamma}$$

with $\hat{R} = \hat{\delta}R$ and $\hat{\gamma} = \hat{\delta}\gamma$. As a result, (21) and (22) are equivalent to (25) and (26), respectively.

Based on Theorem 1, we reformulate Problem 2 more explicitly as the following minimization problem

$$\begin{aligned} \min_{\hat{Q} > 0, \hat{W} \in \mathcal{M}} \quad & \text{tr}(R) \\ \text{s.t.} \quad & (20) \text{ and } (21) \text{ hold} \end{aligned} \quad (28)$$

where $\hat{R} = \hat{\delta}R$ with a given $\hat{\delta} \in \mathbb{R}_+$. Note that the constraint (21) can be solved efficiently using standard LMI solvers, while (20), due to the nonlinearity term $\bar{A}_r^\top \bar{A}_r$, is a bilinear matrix inequality, which causes the major challenge in solving the problem (28).

To handle the bilinear constraint (20), we adopt the technique called *psd-convex-concave decomposition* [20].

Definition 2: A matrix-valued mapping $\Phi: \mathbb{R}^n \rightarrow \mathcal{S}^\ell$ is called positive semidefinite convex-concave (psd-convex-concave) if Φ can be expressed as $\Phi = \Phi_1 - \Phi_2$, where Φ_k , with $k = 1, 2$, are positive semidefinite convex (psd-convex), i.e.,

$$\Phi_k(\lambda w_1 + (1-\lambda)w_2) \leq \lambda \Phi_k(w_1) + (1-\lambda)\Phi_k(w_2) \quad (29)$$

holds for all $\lambda \in [0, 1]$ and $w_1, w_2 \in \mathbb{R}^n$. The pair (Φ_1, Φ_2) is called a psd-convex-concave decomposition of Φ .

Consider the bilinear inequality (20), and define the following matrix-valued mapping:

$$\Phi(\hat{Q}, \hat{\delta}, \hat{W}) = \psi(\hat{Q}, \hat{\delta}) + \varphi(\hat{W}) \quad (30)$$

where

$$\psi(\hat{Q}, \hat{\delta}) = \begin{bmatrix} \hat{Q}\bar{A} + \bar{A}^\top\hat{Q} & \hat{Q}B_e & \hat{Q}E \\ B_e^\top\hat{Q} & -\hat{\delta}I & 0 \\ E^\top\hat{Q} & 0 & 0 \end{bmatrix} \quad (31)$$

$$\varphi(\hat{W}) = \begin{bmatrix} -\bar{A}_r^\top\bar{A}_r & 0 & \bar{A}_r^\top \\ 0 & 0 & 0 \\ \bar{A}_r & 0 & -I \end{bmatrix}. \quad (32)$$

Then, the following lemma shows that the pair $(\psi, -\varphi)$ is a psd-convex-concave decomposition of Φ .

Lemma 4: The matrix-valued mapping $\Phi(\hat{Q}, \hat{\delta}, \hat{W})$ in (30) is psd-convex-concave.

Proof: Note that the matrix $\psi(\hat{Q}, \hat{\delta})$ in (30) is linear with respect to \hat{Q} and $\hat{\delta}$. Thus, it is immediate that $\psi(\hat{Q}, \hat{\delta})$ is psd-convex. Then, the claim holds if $-\varphi(\hat{W})$ in (30) is psd-convex.

With the structure of \bar{A}_r in (23), the only nonlinear submatrix in $-\varphi(\hat{W})$ can be expressed as

$$\bar{A}_r^\top\bar{A}_r = \begin{bmatrix} 0 & 0 \\ 0 & \varphi_a(\hat{W}) \end{bmatrix} \quad (33)$$

with $\varphi_a(\hat{W}) := S_r^\top \hat{M}^{-1} \hat{B} \hat{W} \hat{B}_0^\top (S_r^+)^T S_r^+ \hat{B}_0 \hat{W} \hat{B}^\top \hat{M}^{-1} S_r$. Then, showing the psd-convexity of $-\varphi(\hat{W})$ in (32) is equivalent to prove that $\varphi_a(\hat{W})$ is psd-convex.

Let $W_1, W_2 \in \mathcal{M}$, and denote $W_\lambda = \lambda W_1 + (1 - \lambda) W_2$. For any $\lambda \in [0, 1]$, we have

$$\begin{aligned} & \varphi_a(W_\lambda) - \lambda \varphi_a(W_1) - (1 - \lambda) \varphi_a(W_2) \\ &= S_r^\top \hat{M}^{-1} \hat{B} W_\lambda \hat{B}_0^\top (S_r^+)^T S_r^+ \hat{B}_0 W_\lambda \hat{B}^\top \hat{M}^{-1} S_r \\ & \quad - \lambda S_r^\top \hat{M}^{-1} \hat{B} W_1 \hat{B}_0^\top (S_r^+)^T S_r^+ \hat{B}_0 W_1 \hat{B}^\top \hat{M}^{-1} S_r \\ & \quad - (1 - \lambda) S_r^\top \hat{M}^{-1} \hat{B} W_2 \hat{B}_0^\top (S_r^+)^T S_r^+ \hat{B}_0 W_2 \hat{B}^\top \hat{M}^{-1} S_r \\ &= -\lambda(1 - \lambda) (V_1(\hat{W}_1 - \hat{W}_2) V_2(\hat{W}_1 - \hat{W}_2) V_1^\top) \leq 0 \end{aligned} \quad (34)$$

where $V_1 = S_r^\top \hat{M}^{-1} \hat{B}$ and $V_2 = \hat{B}_0^\top (S_r^+)^T S_r^+ \hat{B}_0$. Since $-\lambda(1 - \lambda) \leq 0$ and $V_1(\hat{W}_1 - \hat{W}_2) V_2(\hat{W}_1 - \hat{W}_2) V_1^\top \geq 0$, it holds that

$$\varphi_a(\lambda W_1 + (1 - \lambda) W_2) \leq \lambda \varphi_a(W_1) + (1 - \lambda) \varphi_a(W_2)$$

which implies that the mapping $\varphi_a(\hat{W})$ is psd-convex from (29), i.e., $-\varphi(\hat{W})$ is psd-convex. As a result, it follows from Definition 2 that the matrix-valued mapping $\Phi(\hat{Q}, \hat{\delta}, \hat{W})$ in (30) is psd-convex-concave.

The psd-convex-concave decomposition in (30) allows us to linearize the optimization problem (28) at a stationary point $\hat{W} \in \mathcal{M}$. To simplify the optimization procedure, we introduce a new optimization variable μ to eliminate the equality constraint $\hat{B}\hat{w} = 0$ in (12), where $\hat{W} = \text{Diag}(\hat{w})$, and m is the number of edges in the reduced network.

Let $\bar{B} \in \mathbb{R}^{\bar{r} \times m}$ be a full row rank matrix obtained by removing linearly dependent rows of the $\hat{B} \in \mathbb{R}^{r \times m}$, and it still holds that $\bar{B}\hat{w} = 0$. Then, there exists a column permutation matrix $P \in \mathbb{R}^{m \times m}$ such that

$$\bar{B}\hat{w} = [\bar{B}_a \ \bar{B}_b] P\hat{w} = \bar{B}_a \mu_a + \bar{B}_b \mu = 0, \text{ with } P\hat{w} = \begin{bmatrix} \mu_a \\ \mu \end{bmatrix}$$

where $\bar{B}_a \in \mathbb{R}^{\bar{r} \times \bar{r}}$ is full rank, $\mu \in \mathbb{R}_+^{\bar{m}}$, and $\bar{m} = m - \bar{r}$, is defined as the new optimization variable. Note that

$$\hat{w} = P^\top \begin{bmatrix} -\bar{B}_a^{-1} \bar{B}_b \\ I_{\bar{m}} \end{bmatrix} \mu \quad (35)$$

which projects the weights \hat{w} into $\ker(\hat{B})$. Thereby, we rewrite the constraint $\hat{W} \in \mathcal{M}$ as $\mu \in \mathbb{R}_+^{\bar{m}}$ (35). Now, we redefine the matrix-valued

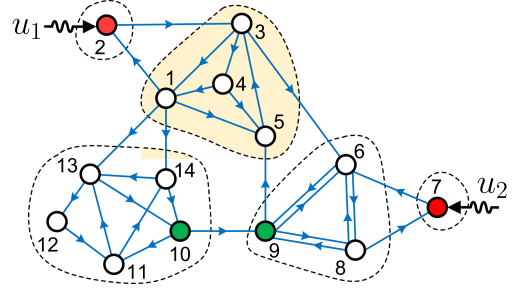


Fig. 2. Connected directed sensor network containing 14 vertices, in which the red vertices are controlled, and the shadowed ones are measured.

mapping $\varphi(\hat{W})$ in (30) as

$$\phi(\mu) = \varphi(\hat{W}) \quad (36)$$

which remains psd-convex due to the linear relation in (35). The derivative of the matrix-valued mapping $\phi(\mu)$ at μ is a linear mapping $D\phi: \mathbb{R}_+^{\bar{m}} \rightarrow \mathcal{S}^\ell$, with $\ell = n + 2r + p + pq - 3$, which is defined as

$$D\phi(\mu)[h] = \sum_{i=1}^{\bar{m}} h_i \frac{\partial \phi}{\partial \mu_i}(\mu) \quad \forall h \in \mathbb{R}_+^{\bar{m}}. \quad (37)$$

Given a point $\mu^{(k)} \in \mathbb{R}_+^{\bar{m}}$, the linearized formulation of the problem (28) at $\mu^{(k)}$ is formulated as

$$\begin{aligned} & \min_{\hat{Q} > 0, \mu \in \mathbb{R}_+^{\bar{m}}} f(\mu) = \text{tr}(R) \\ & \text{s.t.} \quad \begin{bmatrix} \hat{Q} & \hat{\delta} C_e^\top \\ \hat{\delta} C_e & \hat{R} \end{bmatrix} > 0, \hat{\delta} \in \mathbb{R}_+, \hat{R} = \hat{\delta} R > 0 \\ & \psi(\hat{Q}, \hat{\delta}) + \varphi(\hat{W}^{(k)}) + D\phi(\mu^{(k)})[\mu - \mu^{(k)}] < 0 \end{aligned} \quad (38)$$

where the derivative of $\phi(\mu^{(k)})$ is given as

$$D\phi(\mu^{(k)})[\mu - \mu^{(k)}] := \sum_{i=1}^{\bar{m}} (\mu_i - \mu_i^{(k)}) \frac{\partial \phi}{\partial \mu_i^{(k)}}(\mu^{(k)})$$

with $j = 1, \dots, \bar{m}$. Note that the optimization problem (38) is convex, of which the global optimum can be solved efficiently using standard SDP solvers, e.g., SeDuMi [25]. Moreover, the computational complexity of Algorithm 1 mainly depends on solving the problem (38), more precisely, on adopting which solver to solve the LMIs in (38). Thus, different solvers will lead to different complexity of Algorithm 1.

Based on Lemma 4 and (38), we are now ready to present an algorithmic approach for solving the minimization problem in (28) in an iterative fashion, see Algorithm 1, where the main procedure follows [20] and $\varepsilon \in \mathbb{R}_+$ is a prefixed error tolerance determining whether to terminate the iteration loop.

The initial condition $\mu^{(0)}$ can be chosen as an arbitrary vector with all strictly positive entries. With (35), it will guarantee $\hat{W}^0 \in \mathcal{M}$, i.e., the reduced graph is balanced. Furthermore, we can also initialize μ using the outcome of graph clustering projection in [18], and [19]. Specifically, from a given clustering Π , we construct an initial reduced Laplacian matrix in (10) as $\hat{L}_b^{(0)} := \Pi^\top L_b \Pi$, with L_b the Laplacian matrix of the balanced graph \mathcal{G}_b . Then, the initial weight of the edge (i, j) in the quotient graph $\hat{\mathcal{G}}_b$ is the (i, j) th entry of $\hat{L}_b^{(0)}$. By doing so, $\mu^{(0)}$ can be formed such that $\hat{W}^0 \in \mathcal{M}$.

The convergence analysis of Algorithm 1 follows naturally from [20], and it means that a local optimum can be obtained. More

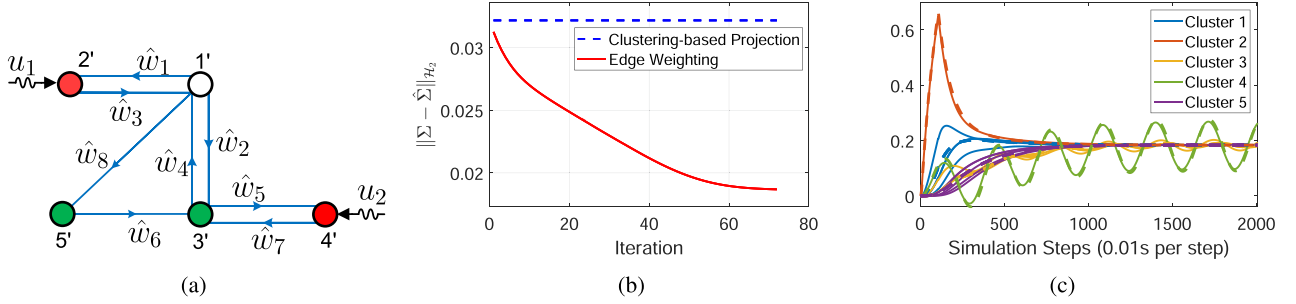


Fig. 3. (a) Quotient graph obtained by the clustering, where the controlled vertices are labeled with red color, and measured vertices are filled by green color. The weights of the edges are parameters to be determined. (b) Approximation errors of clustering-based projection and the proposed edge weighting method. (c) State trajectories of the original and reduced networks, indicated by solid and dash lines, respectively.

Algorithm 1: Iterative Edge Weighting.

Input: L, F, H, Π , and a small scalar $\hat{\delta} \in \mathbb{R}_+$

Output: \hat{W}^* .

- 1: Compute the incidence matrix \hat{B} of the quotient graph \hat{G}_b .
 - 2: Choose an initial vector $\mu^{(0)} \in \mathbb{R}_+^{\hat{m}}$.
 - 3: Set iteration step: $k \leftarrow 0$.
 - 4: **repeat**
 - 5: Solve (38) to obtain the optimal solution μ^* .
 - 6: $k \leftarrow k + 1$, and $\mu^{(k)} \leftarrow \mu^*$.
 - 7: **until** $|f(\mu^{(k+1)}) - f(\mu^{(k)})| \leq \varepsilon$.
 - 8: Compute \hat{w}^* using (35), and return $\hat{W}^* \leftarrow \text{Diag}(\hat{w}^*)$.
-

importantly, if we select the initial condition from the clustering-based projection, it is guaranteed that, through iteration, the approximation accuracy of reduced-order network model with the weights obtained by Algorithm 1 will be improved. In this sense, the approximations obtained by the proposed method is at least better than the ones obtained by clustering-based projection methods in, e.g., [18] and [19]. We will show this merit from a numerical example in the next section.

IV. ILLUSTRATIVE EXAMPLE

To illustrate the effectiveness of the proposed edge weighting approach, we implement it to a sensor network example from [3] and [18]. The topology of this network is shown in Fig. 2, which consists of 14 strongly connected vertices, and all the edge weights are 1. In this example, two external input signals are injected into the network via vertices 2 and 7, respectively, and the states of vertices 9 and 10 are measured.

Suppose that five clusters are given for this directed network as $C_1 = \{1, 3, 4, 5\}$, $C_2 = \{2\}$, $C_3 = \{6, 8, 9\}$, $C_4 = \{7\}$, and $C_5 = \{10, 11, 12, 13, 14\}$, which leads to the quotient network in Fig. 3(a), with incidence matrix

$$\hat{B} := \begin{bmatrix} 1 & 1 & -1 & -1 & 0 & 0 & 0 & -1 \\ -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 \end{bmatrix}$$

There are eight edges in the quotient graph, and each edge is assigned with a symbolic weight as labeled in Fig. 3(a). These variables,

determining the reduction error, are to be determined by our optimization approach.

First, the parameterized reduced model in (9) of the quotient graph is generated with matrices

$$\hat{L}_b = \begin{bmatrix} \hat{w}_1 + \hat{w}_2 & -\hat{w}_1 & & -\hat{w}_2 & & 0 & 0 & \\ & -\hat{w}_3 & \hat{w}_3 & & 0 & & 0 & 0 \\ & -\hat{w}_4 & 0 & \hat{w}_4 + \hat{w}_5 + \hat{w}_6 & -\hat{w}_5 & -\hat{w}_6 & & \\ & 0 & 0 & & -\hat{w}_7 & \hat{w}_7 & 0 & \\ & -\hat{w}_8 & 0 & & 0 & 0 & 0 & \hat{w}_8 \end{bmatrix}$$

$$\hat{M} = [2.1921 \ 0.6803 \ 0.3779 \ 0.0756 \ 0.4157]$$

$$\hat{F}_b = \begin{bmatrix} 0 & 0.6803 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.0756 & 0 \end{bmatrix}^\top, \hat{H} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

and the weight vector \hat{w} in (7) satisfy the following constraints for a balanced graph: $\hat{w}_1 = \hat{w}_3$, $\hat{w}_2 = \hat{w}_4 + \hat{w}_8$, $\hat{w}_6 = \hat{w}_8$, and $\hat{w}_5 = \hat{w}_7$.

Next, we implement Algorithm 1 to solve the optimization problem (28) with $\mu = [\hat{w}_1, \hat{w}_2, \hat{w}_5, \hat{w}_6]^\top \in \mathbb{R}_+^4$ as the optimization variable. Particularly, the SeDuMi solver [25] is adopted to solve the convex problem (38). We choose the initial edge weights obtained by the clustering-based projection [18], [19], which gives $\hat{w}^{(0)} = [0.6803, 0.2268, 0.6803, 0.0756, 0.0756, 0.1512, 0.0756, 0.1512]$ and the approximation error $\|G_e(s)\|_{\mathcal{H}_2} = 0.0322$. With $\hat{\delta} = \varepsilon = 10^{-5}$, Algorithm 1 stops after 72 iterations. The convergence trajectory of the resulting \mathcal{H}_2 reduction error is shown in Fig. 3(b). The final solution of the edge weights are given as $\hat{w}^* = [0.6826, 0.2394, 0.6826, 0.0948, 0.0537, 0.1446, 0.0537, 0.1446]$, which provides the approximation error $\|G_e(s)\|_{\mathcal{H}_2} = 0.0187$. Through iteration, the edge weighting method further reduces the error by 41.93%, compared with the clustering-based projection. Therefore, our method can provide a reduced network systems with a better \mathcal{H}_2 approximation error. Then, we compare the state trajectories of reduced and original networks in Fig. 3(c). The data are obtained with u_1 as a step signal when $u_1(t) = 1$, if $t \in [0, 1]$, and $u_1 = 0$, otherwise, and $u_2 = 0.2 \sin(2t)$. We observe that each trajectory of the reduced network provides an approximation of the average states of the corresponding cluster in the original network.

V. CONCLUSION

In this article, the \mathcal{H}_2 model reduction problem for dynamical networks consisting of diffusively coupled agents has been formulated as a minimization problem, in which the edge weights in the reduced network are parameters to be chosen. Necessary and sufficient conditions have been proposed for constructing a set of optimal edge

weights. An iterative algorithm has been provided to search for the desired edge weights such that the \mathcal{H}_2 norm of the approximation error is small. Finally, compared with the projection-based method in [12], the feasibility of this method is illustrated by an example. The advantage of this proposed model reduction method is that not only the structure of the original network has been preserved but also the approximation error has been optimized. For future work, the performance of the iterative algorithm can be studied further to benefit from applying more efficient optimization methods. Moreover, an extension to networked high-order linear subsystems is also of interest.

REFERENCES

- [1] J. A. Fax and R. M. Murray, "Graph Laplacians and stabilization of vehicle formations," *IFAC Proc. Volumes*, vol. 35, no. 1, pp. 55–60, 2002.
- [2] F. Dörfler, J. W. Simpson-Porco, and F. Bullo, "Electrical networks and algebraic graph theory: Models, properties, and applications," *Proc. IEEE*, vol. 106, no. 5, pp. 977–1005, May 2018.
- [3] G. Scutari, S. Barbarossa, and L. Pescosolido, "Distributed decision through self-synchronizing sensor networks in the presence of propagation delays and asymmetric channels," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1667–1684, Apr. 2008.
- [4] T. H. Summers and J. Lygeros, "Optimal sensor and actuator placement in complex dynamical networks," *IFAC Proc. Volumes*, vol. 47, no. 3, pp. 3784–3789, 2014.
- [5] A. J. Gates and L. M. Rocha, "Control of complex networks requires both structure and dynamics," *Sci. Rep.*, vol. 6, no. 1, pp. 1–11, 2016.
- [6] T. Ishizaki, A. Chakraborty, and J.-I. Imura, "Graph-theoretic analysis of power systems," *Proc. IEEE*, vol. 106, no. 5, pp. 931–952, May 2018.
- [7] A. C. Antoulas, *Approximation of Large-Scale Dynamical Systems*. Philadelphia, PA, USA: SIAM, 2005.
- [8] X. Cheng, J. M. A. Scherpen, and B. Besselink, "Balanced truncation of networked linear passive systems," *Automatica*, vol. 104, pp. 17–25, 2019.
- [9] E. Biyik and M. Arcak, "Area aggregation and time-scale modeling for sparse nonlinear networks," *Syst. Control Lett.*, vol. 57, no. 2, pp. 142–149, 2008.
- [10] S. Rao, A. J. van der Schaft, and B. Jayawardhana, "A graph-theoretical approach for the analysis and model reduction of complex-balanced chemical reaction networks," *J. Math. Chem.*, vol. 51, no. 9, pp. 2401–2422, 2013.
- [11] X. Cheng and J. M. A. Scherpen, "Model reduction methods for complex network systems," *Annu. Rev. Control, Robot. Auton. Syst.*, vol. 4, no. 1, pp. 425–453, 2021.
- [12] N. Monshizadeh, H. L. Trentelman, and M. K. Camlibel, "Projection-based model reduction of multi-agent systems using graph partitions," *IEEE Trans. Control Netw. Syst.*, vol. 1, no. 2, pp. 145–154, Jun. 2014.
- [13] B. Besselink, H. Sandberg, and K. H. Johansson, "Clustering-based model reduction of networked passive systems," *IEEE Trans. Autom. Control*, vol. 61, no. 10, pp. 2958–2973, Oct. 2016.
- [14] T. Ishizaki, K. Kashima, J. I. Imura, and K. Aihara, "Model reduction and clusterization of large-scale bidirectional networks," *IEEE Trans. Autom. Control*, vol. 59, no. 1, pp. 48–63, Jan. 2014.
- [15] T. Ishizaki, K. Kashima, A. Girard, J.-I. Imura, L. Chen, and K. Aihara, "Clustered model reduction of positive directed networks," *Automatica*, vol. 59, pp. 238–247, 2015.
- [16] H.-J. Jongsma, P. Mlinarić, S. Grundel, P. Benner, and H. L. Trentelman, "Model reduction of linear multi-agent systems by clustering with h_2 and H_∞ error bounds," *Math. Control, Signals, Syst.*, vol. 30, no. 1, 2018, Art. no. 6.
- [17] X. Cheng, Y. Kawano, and J. M. A. Scherpen, "Model reduction of multi-agent systems using dissimilarity-based clustering," *IEEE Trans. Autom. Control*, vol. 64, no. 4, pp. 1663–1670, Apr. 2019.
- [18] X. Cheng and J. M. A. Scherpen, "A new controllability Gramian for semistable systems and its application to approximation of directed networks," in *Proc. IEEE 56th Conf. Decis. Control*, Melbourne, Australia, 2017, pp. 3823–3828.
- [19] X. Cheng and J. M. A. Scherpen, "Clustering-based model reduction of Laplacian dynamics with weakly connected topology," *IEEE Trans. Autom. Control*, vol. 65, no. 10, pp. 4393–4399, Oct. 2020.
- [20] Q. T. Dinh, S. Gumussoy, W. Michiels, and M. Diehl, "Combining convex-concave decompositions and linearization approaches for solving BMIs, with application to static output feedback," *IEEE Trans. Autom. Control*, vol. 57, no. 6, pp. 1377–1390, Jun. 2012.
- [21] X. Cheng, L. Yu, and J. M. A. Scherpen, "Reduced order modeling of linear consensus networks using weight assignments," in *Proc. 18th Eur. Control Conf.*, 2019, pp. 2005–2010.
- [22] F. Bullo, *Lectures on Network Systems*. Seattle, WA, USA: Kindle Direct Publishing, 2022. [Online]. Available: <http://motion.me.ucsb.edu/books>
- [23] M. Mesbahi and M. Egerstedt, *Graph Theoretic Methods in Multiagent Networks*. Princeton, NJ, USA: Princeton Univ. Press, 2010.
- [24] G. Pipeleers, B. Demeulenaere, J. Swevers, and L. Vandenberghe, "Extended LMI characterizations for stability and performance of linear systems," *Syst. Control Lett.*, vol. 58, no. 7, pp. 510–518, 2009.
- [25] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optim. Methods Softw.*, vol. 11, no. 1/4, pp. 625–653, 1999.