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

Xiaodong Cheng, Lanlin Yu, Dingchao Ren and Jacquielien M.A. Scherpen

Abstract—This paper studies reduced-order modeling of dynamic networks with strongly connected topology. Given a graph clustering of an original complex network, we construct a quotient graph with less number of vertices, where the edge weights are parameters to be determined. The model of the 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 that minimizes 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.

I. INTRODUCTION

With the growing complexity of spatially interconnected dynamic systems, the importance of understanding and managing dynamic networks has been widely recognized. An important class of dynamic systems 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 vehicle formations, electrical networks, synchronization in sensor networks, and opinion dynamics in social networks, see e.g. [1]–[4]. The spatial structures of such systems are usually complex and result in high-dimensional models that cause challenges for analysis, control, and optimization. To effectively capture the collective behaviors of dynamics over networks, 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 paper puts emphasis on the preservation of the network structure, which is necessary for applications e.g., distributed controller design and sensor allocation [5]–[7]. Conventional model reduction methods, e.g., balanced truncation and moment matching, merely focus on approximating the input-output behavior of a given dynamic system [8], while the preservation of the network structure is barely guaranteed. Although a generalized balanced truncation approach in [9] is able to construct an accurate reduced-order model with a network interpretation, the relation between the original and obtained new topologies is not yet clear. Singular perturbation approximation, as alternative network structure-preserving approach, has been applied to the reduction of electric circuits [10] and chemical reaction networks [11]. This class of methods mainly relies on time-scale separation of the states in an autonomous network system, while the external inputs are not considered explicitly. Besides, the resulting reduced topology hardly retains sparsity.

Recently, clustering-based methods have been intensively studied and become the mainstream methodology for reducing network systems, see e.g., [12]–[20]. With graph clustering, the vertices in a large-scale network are partitioned into several disjoint clusters. This class of methods has a clear advantage in retaining the consensus property [12], [19], system positivity [15], and scale-free property [20] in reduced-order models. The model 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 put their main 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 paper, 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 established based on this quotient graph, in which the edge weights are free variables to be optimized. Then, the major problem in this paper 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 (BMI) constraint, which is computationally expensive. Therefore, we devise a novel edge weighting algorithm based on the convex-concave decomposition, which linearizes the nonconvex constraint as a convex one in the form of a linear matrix inequality (LMI). 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

This work of X. Cheng is supported by the European Research Council (ERC), Advanced Research Grant SYSDYNET, under the European Unions Horizon 2020 research and innovation programme (Grant Agreement No. 694504). This work of L. Yu is supported by the National Natural Science Foundation of China Under Project 61761136005. Xiaodong Cheng is with Control Systems Group, Department of Electrical Engineering, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands x.cheng@tue.nl

Lanlin Yu is with the Institute of Advanced Technology, Westlake Institute for Advanced Study, Westlake University, Hangzhou 310024, China and Dingchao Ren is with the Department of Automation, University of Science and Technology of China, Hefei 230026, China. {yulanlin1992, dcrenustc}@gmail.com

Jacquielien M.A. Scherpen is with Jan C. Willems Center for Systems and Control, Engineering and Technology Institute Groningen, Faculty of Science and Engineering, University of Groningen, Nijenborgh 4, 9747 AG Groningen, the Netherlands. j.m.a.scherpen@rug.nl
model is guaranteed a better approximation accuracy than the clustering-based projection methods.

The rest of this paper 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, and Section V finally makes some concluding remarks.

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

For a real matrix $A$, the columns of $A^\perp$ form a basis of the null space of $A$, that is, $AA^\perp = 0$.

II. Preliminaries and Problem Setting

This section provides necessary definitions and concepts in graph theory used in this paper, and we refer to [21] 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 $G := (\mathcal{V}, \mathcal{E})$ consists of a finite and nonempty node set $\mathcal{V} := \{1, 2, \ldots, n\}$ 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 $G$ is called simple, if $G$ 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 introduce several important matrices for characterizing a directed simple graph. Let $m := |\mathcal{E}|$, the incidence matrix $B \in \mathbb{R}^{n \times m}$ is defined by

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

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

$$
L_{ij} = \begin{cases} 
\sum_{j=1, j\neq i}^n A_{ij} & \text{if } i = j, \\
-A_{ij} & \text{otherwise.}
\end{cases}
$$

(1)

Clearly, $L1 = 0$. The diagonal entries of $L$ are strictly positive, and the off-diagonal entries are non-positive. Alternatively, we can characterize the Laplacian matrix using the incidence matrix of $G$ as

$$
L = B^T W B^T, 
$$

(2)

where $B_0$ is a binary matrix obtained by replacing all “$-1$” entries in the incidence matrix $B$ with zeros, and $W := \text{Diag}(w)$, with $w = [w_1 \, w_2 \, \cdots \, w_m]^T$,

and $w_k$ the positive weight associated to the $k$-th edge, for all $k \in \{1, 2, \cdots, |\mathcal{E}|\}$.

For a vertex in a weighted graph, the indegree and outdegree of the vertex are computed as $\sum_{j \in \mathcal{V}} A_{ij}$ and $\sum_{i \in \mathcal{V}} A_{ij}$, respectively. A strongly connected graph $G$ is called balanced if the indegrees and outdegrees of each vertex in $G$ is equal.

From (1), the following lemma is immediate.

Lemma 1. A weighted strongly connected graph $G$ is balanced if and only if one of the following conditions hold.

1) The edge weights of $G$ satisfies $Bw = 0$.

2) The Laplacian matrix of $G$ satisfies $\ker(L) = \ker(L^\top) = \text{span}(1)$.

The strong connectivity implies that there is only one zero eigenvalue of $L$ [21], and the balance of $G$ 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 $L = BWB^\top$, where $B$ 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]–[17].

Definition 1. Let $G := (\mathcal{V}, \mathcal{E})$ be a directed graph. Then, a graph clustering is a partition of $\mathcal{V}$ into $r$ nonempty disjoint subsets $C_1, C_2, \cdots, C_r$ covering all the elements in $\mathcal{V}$, where $C_i$ is called a cluster of $G$.

Let $\{C_1, C_2, \cdots, C_r\}$ be a clustering of $G$ 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:

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

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

$$
\Pi^T \Pi = I_n \text{ and } I_n^\top \Pi = [|C_1|, |C_2|, \cdots, |C_r|].
$$

(3)

B. Problem Setup

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

$$
\dot{x}_i(t) = -\sum_{j=1}^n A_{ij} [x_i(t) - x_j(t)] + \sum_{j=1}^n f_{ij} u_k(t),
$$

(4)

where $x_i(t) \in \mathbb{R}$ is the state of vertex $i$, and $A_{ij}$ is the $(i, j)$-th entry of the adjacency matrix of $G$, representing the strength of the coupling between vertices $i$ and $j$. $u_k(t) \in \mathbb{R}$ is the external input, and $f_{ij} \in \mathbb{R}$ is the gain of the $j$-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
\[ \dot{\mathbf{x}}(t) = -L \mathbf{x}(t) + \hat{F} u(t), \]
\[ y(t) = H \hat{\mathbf{x}}(t), \]
(5)
with \( x(t) := [x_1, x_2, \ldots, x_n]^T \in \mathbb{R}^n \) and \( u := [u_1, u_2, \ldots, u_p]^T \in \mathbb{R}^p \). The vector \( y \in \mathbb{R}^q \) collects the outputs of the network, and \( H \) is the output matrix.

This paper aims for structure-preserving model reduction of diffusively coupled networks in form of (5), 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 paper is as follows.

**Problem 1.** Given a network system \( \Sigma \) as in (5) and a graph clustering \( \{C_1, C_2, \ldots, C_r\} \), find a reduced-order model
\[ \hat{\Sigma} : \begin{cases} \dot{\hat{\mathbf{x}}}(t) = -\hat{L} \hat{\mathbf{x}}(t) + \hat{F} u(t) \\ \hat{y}(t) = \hat{H} \hat{\mathbf{x}}(t) \end{cases} \]
(6)
with \( \hat{\mathbf{x}} \in \mathbb{R}^r, r \ll n \), such that \( \hat{L} \) is the Laplacian matrix of a reduced directed graph, and the reduction error \( \| \Sigma - \hat{\Sigma} \|_{\mathcal{H}_2} \) is minimized. \( L \in \mathbb{R}^{r \times r}, F \in \mathbb{R}^{r \times p}, H \in \mathbb{R}^{q \times r} \) are matrices depending on the graph clustering.

It is worth emphasizing that Problem 1 does not aim to find an appropriate graph clustering of the network \( \Sigma \). 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], [19], and we do not apply the Petrov-Galerkin projection framework.

### III. MAIN RESULT

In this section, a novel model reduction approach for network systems is presented with two steps. In the first step, a parameterized model of a reduced network is constructed on the basis of graph clustering. Then, the second step computes a set of 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 (5) to its balanced graph representation as follows.

**Lemma 2.** If the underlying graph of \( \Sigma \) in (5) is strongly connected, then there exists a diagonal \( M \in \mathbb{R}^{r \times r} \) with positive diagonal entries such that \( \Sigma \) is equivalent to
\[ \begin{cases} M \dot{x}(t) = -L_0 x(t) + F_b u(t) \\ y(t) = H x(t) \end{cases}, \]
(7)
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 (7) 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 \( \{C_1, C_2, \ldots, C_r\} \), the quotient graph \( \mathcal{G}_b \) is an 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 \( \mathcal{G}_b \) with vertices \( i, j \) in the same cluster, then it will not be presented as an edge in \( \mathcal{G}_0 \). If there exists an edge \((i, j)\) with \( i \in C_k \) and \( j \in C_l \), then there is an edge \((k, l)\) in \( \mathcal{G}_b \).

Let \( \hat{B} \) be the incidence matrix of the quotient graph \( \hat{\mathcal{G}}_b \). Algebraically, it can be verified that \( \hat{B} \) is obtained by removing all the zero columns of \( \Pi^T B \), where \( B \) is the incidence matrix of \( \mathcal{G}_b \) (or \( \mathcal{G} \)). Furthermore, we denote
\[ \hat{W} = \text{Diag}(\hat{w}), \quad \hat{w} = [\hat{w}_1, \hat{w}_2, \ldots, \hat{w}_m]^T, \]
(8)
as the edge weight matrix of \( \mathcal{G}_b \), where \( \hat{w}_k \in \mathbb{R}_+ \), and \( m \) is the number of edges in \( \mathcal{G}_b \). In order to maintain \( \mathcal{G}_b \) as a balanced graph, we impose the constraint on its edge weights as
\[ \hat{B} \hat{w} = 0, \]
(9)
according to Lemma 3. Thereby, the dynamics on the balanced quotient graph \( \hat{\mathcal{G}}_b \) is then obtained as
\[ \begin{cases} \hat{M} \dot{\hat{\mathbf{x}}}(t) = -\hat{L}_b(\hat{W}) \hat{\mathbf{x}}(t) + \hat{F}_b u(t) \\ \hat{y}(t) = \hat{H} \hat{\mathbf{x}}(t) \end{cases}, \]
(10)
with the reduced matrices
\[ \hat{M} = \Pi^T M \Pi, \quad \hat{L}_b(\hat{W}) = \hat{B}_b \hat{W} \hat{B}_b^T, \quad \hat{F}_b = \Pi^T F_b, \quad \hat{H} = H \Pi, \]
(11)
where \( \hat{B}_b \) is the binary matrix obtained by replacing all the “-1” entries with zeros in \( \hat{B} \), and \( \hat{M}_0 = \mathbb{R}^{r \times r}, \hat{F}_b = \mathbb{R}^{r \times p}, \) and \( H \in \mathbb{R}^{q \times r} \) are reduced matrices determined by the given clustering of \( \mathcal{G}_b \). Since the graph clustering is given, i.e., \( \Pi \) is known, the only parameters to be decided are the weights in \( \hat{W} \), which satisfy the constraint (9).

From the reduced graph balanced representation (10), we immediately construct a parameterized reduced-order model in the form of (6) with the reduced matrices
\[ \hat{L}(\hat{W}) = \hat{M}^{-1} \hat{B}_b \hat{W} \hat{B}_b^T, \quad \hat{F} = \hat{M}^{-1} \Pi^T F_b, \quad \hat{H} = H \Pi, \]
(12)
where \( \hat{L} \) represents a reduced weighted graph \( \hat{\mathcal{G}} \). In (12), only the weight matrix \( \hat{W} \) is to be determined, which is selected from the following set
\[ \mathcal{M} := \{W = \text{Diag}(\hat{w}) \mid \hat{w} \in \mathbb{R}_+^m, \hat{B} \hat{w} = 0\}. \]
(13)

In the following example, we demonstrate the parameterized modeling of a simplified dynamic network.
Example 1. Consider an network example in vehicle formation [1], where the formation topology $G$ is depicted in Fig. 1a. Clearly, $G$ is balanced, i.e., $G = G_b$, with the incidence matrix

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

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 (7) with

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

and $H = [1 \ 0 \ 0 \ 0 \ 0 \ 0]$. Consider a clustering of $G_b$ as $C_1 = \{1, 2\}$, $C_2 = \{3, 4, 5\}$, $C_3 = \{6\}$, which leads to the characterization matrix as

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

The topology of the quotient graph $\hat{G}_b$ is shown in Fig. 1b 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 $\hat{G}_b$ are positive parameters to be determined, as labeled in Fig. 1b, 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 $\hat{B} \hat{\hat{w}} = 0$, namely, $\hat{w}_3 = \hat{w}_1 - \hat{w}_2$, $\hat{w}_4 = \hat{w}_2$, such that $\hat{G}_b$ is balanced. The other matrices in the reduced-order model (10) are computed as

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

and $\hat{H} = H \Pi = [1 \ 0 \ 0]$. Then, in the parameterized reduced-order model 6, we have

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

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. 1c, which is no longer balanced.

Remark 2. The physical interpretation of the reduced matrices in (10) are explained. $\hat{M}$ is constructed such that the mass of vertex $k$ in $\hat{G}_b$ is equal to the mass sum of all the vertices in $C_k$ in $G_b$. The expression of $\hat{F}_b$ means that if a vertex in a cluster $C_k$ of $G_b$ is controlled by an external input, then vertex $k$ in $\hat{G}_b$ is also controlled. Analogously, $\hat{H}$ indicates that a vertex $k$ in $\hat{G}_b$ is measured if there is a measurement taken from a vertex in $C_k$.

With the reduced matrices in (12) and the constraint in (9), an important property of the reduced-order network model $\hat{\Sigma}$ is that it guarantees the $H_2$ reduction error between the original system $\Sigma$ in (5) and $\hat{\Sigma}$ is always bounded. The computation of the reduction error amounts to find the $H_2$ norm of the following error system:

$$G_e(s) = C_e(sI - A_e)^{-1}B_e, $$

where

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

Note that $A_e$ is not Hurwitz, since $L$ and $\hat{L}$ are both Laplacian matrices containing zero eigenvalues. Thus, $\|G_e(s)\|_{H_2}$ cannot be calculated directly using the state space representation (14). Here, we employ the following matrices

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

which are independent of system dynamics and satisfy

$$S^\top_n I_n = 0, \quad S^\top_r I_r = 0. $$

Let their left pseudo inverses be

$$S_n^+ := (S_n^\top M^{-1} S_n)^{-1} S_n^\top M^{-1} \in \mathbb{R}^{(n-1) \times n}, \quad S_r^+ := (S_r^\top M^{-1} S_r)^{-1} S_r^\top M^{-1} \in \mathbb{R}^{(r-1) \times r}. $$

Then, using the matrices in (15), we show the following result.

**Lemma 3.** Consider the network system $\Sigma$ in (5) and the reduced-order network model $\hat{\Sigma}$ in (6) with matrices in (12). Then, $\Sigma - \hat{\Sigma} \in \mathcal{H}_2$ holds for all $\hat{W} \in \mathcal{M}$. Fig. 1: (a) A directed balanced network $G_b$ consisting of 6 vertices, in which vertex 3 is controlled and vertex 4 is measured. The vertices in different clusters are indicated by distinct colors. (b) The quotient graph $\hat{G}_b$ consisting of 3 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) The resulting reduced graph $\hat{G}_b$. 
Proof. With $S_n$ and $S_r$ in (15), we construct a nonsingular $(n + r) \times (n + r)$ matrix as

$$U_e = \begin{bmatrix} \sigma_M 1_n & 0 & M^{-1}S_n & 0 \\ 0 & \sigma_M^{-1} 1_r & 0 & M^{-1}S_r \end{bmatrix},$$

(16)

where $\sigma_M := 1_n^T M 1_n = 1_r^T \Pi^T \Pi 1_r = 1_r^T M 1_r$. The inverse of $U_e$ is given as

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

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

$$G_e(s) = C_c U_e (s I - U_e^{-1} A_c U_e^{-1})^{-1} U_e^{-1} B_e,$$

$$= [\bar{C}_e \ C_e] \begin{bmatrix} 0 & 0 \\ 0 & (s I - A_e)^{-1} \end{bmatrix} \begin{bmatrix} \bar{B}_e \\ B_e \end{bmatrix},$$

$$= \bar{C}_e \bar{B}_e + C_e (s I - A_e)^{-1} B_e,$$

(17)

where

$$A_e = \begin{bmatrix} S_n^+ L_b M^{-1} S_n & 0 \\ 0 & S_r^+ L_b M^{-1} S_r \end{bmatrix},$$

$$B_e = \begin{bmatrix} S_n^+ F_b \\ S_r^+ F_b \end{bmatrix},$$

$$C_e = \begin{bmatrix} H M^{-1} S_n - \hat{H} M^{-1} S_r \end{bmatrix},$$

(18)

and

$$\bar{B}_e := \begin{bmatrix} 1_n^T F_b \\ 1_r^T F_b \end{bmatrix},$$

and $\bar{C}_e := \sigma_M^{-1} [H 1_n - \hat{H} 1_r].$

(19)

It follows from (3) that $1_n^T \hat{F}_b = 1_n^T F_b$, and $\hat{H} 1_r = H 1_n$, which yield $\bar{C}_e \bar{B}_e = \sigma_M^{-1} (H 1_n 1_n^T F_b - \hat{H} 1_r 1_r^T \hat{F}_b) = 0$. Thus, (17) becomes

$$G_e(s) = C_c (s I - A_e)^{-1} B_e.$$  

(20)

It is not hard to verify that both the matrices $-S_n^+ L_b M^{-1} S_n$ and $-S_r^+ L_b M^{-1} S_r$ are Hurwitz. Consequently, $G_e(s)$ in (20) is asymptotically stable, i.e., $\Sigma - \Sigma \in \mathcal{H}_2$, for all $\hat{W} \in \mathcal{M}$.

Next, we discuss the consensus property of the reduced-order network (6) with the matrices in (12). 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 (5) reaches consensus if

$$\lim_{t \to \infty} |x_i(t) - x_j(t)| = 0$$

holds for all $i, j \in \mathcal{V}$ and all initial conditions.

**Proposition 1.** Consider the network system $\Sigma$ in (5) which reaches consensus. Then, the reduced-order model $\Sigma$ in (6) also reaches consensus, for any clustering $\Pi$ and $\hat{W} \in \mathcal{M}$.

**Proof.** It can verified that the parameterized Laplacian matrix $\hat{L}$ defined in (12) characterizes a strongly connected graph. Thus, $\hat{L}$ has only one zero eigenvalue. Then, the proof immediately follows from [19], [21].

The parameterized modeling of the reduced dynamic network using the graph balanced representation in (7) guarantees the stability of the error system (14), whose $\mathcal{H}_2$ norm can be evaluated via the transfer function (20) with the Hurwitz matrix $A_e$. Note that in (20), the matrices $S_n$ and $S_r$ in (15) is only dependent on the sizes of the networks, and $\Pi$ 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 $\Sigma$ in (5). Given a graph clustering $\Pi$, find a $\hat{W} \in \mathcal{M}$ such that $\|G_e(s)\|_{\mathcal{H}_2}$ is minimized, where $\Sigma$ is the reduced network model defined in (6) with the matrices (12).

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 $\Sigma$ in (5). There exists a reduced-order network model $\Sigma$ in (6) such that $\|G_e(s)\|_{\mathcal{H}_2} < \gamma$ if and only if there exist matrices $\hat{Q} = \hat{Q}^\top > 0$ with dimension $\hat{Q} \in \mathbb{R}^{(n+r-2) \times (n+r-2)}$, $\hat{R} = \hat{R}^\top > 0$ with dimension $\hat{R} \in \mathbb{R}^{q \times q}$, $\hat{W} \in \mathcal{M}$, and $\hat{\delta} \in \mathbb{R}_+$, such that the following inequalities are satisfied,

$$\begin{bmatrix} \hat{Q} \bar{A} + \bar{A}^\top \hat{Q} & \hat{Q} B_e \\ \bar{B}_e^\top \hat{Q} & -\delta I \end{bmatrix} > 0,$$

$$E^\top \hat{Q} > 0,$$

$$\begin{bmatrix} -\bar{A}_r^\top \hat{A}_r & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -I \end{bmatrix} < 0,$$

$$\begin{bmatrix} \hat{Q} & \hat{C}_e \hat{C}_e^\top \\ \hat{C}_e & \hat{R} \end{bmatrix} > 0,$$

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

(21)

(22)

(23)

where $B_e$, $C_e$ are defined in (18), and

$$\bar{A} = \begin{bmatrix} -S_n^+ L_b M^{-1} S_n & 0 \\ 0 & 0 \end{bmatrix}, E = \begin{bmatrix} 0 \\ I \end{bmatrix},$$

$$\bar{A}_r = \begin{bmatrix} 0 & -S_r^+ \hat{B}_0 \hat{W} \hat{B}_0^\top M^{-1} S_r \\ 0 & 0 \end{bmatrix}.$$

(24)

**Proof.** Consider the error system $G_e(s)$ in (20), which is asymptotically stable. Following e.g., [22], we have $\|G_e(s)\|_{\mathcal{H}_2} < \gamma$, with $\gamma \in \mathbb{R}_+$, if and only if there exist matrices $\hat{Q} = \hat{Q}^\top > 0$ and $\hat{R} = \hat{R}^\top > 0$ such that

$$\begin{bmatrix} Q A_e + A_e^\top Q & Q B_e \\ B_e^\top Q & -I_p \end{bmatrix} < 0,$$

$$\begin{bmatrix} Q & C_e^\top \\ C_e & R \end{bmatrix} > 0,$$

$$\text{tr}(R) < \gamma,$$

(25)

(26)

(27)
where $A_e$, $B_e$, $C_e$ are defined in (18).

In the following, we prove that the three inequalities are equivalent to (21), (22), and (23), respectively. First, it is not hard to verify that (25) is equivalent to

\[
\begin{bmatrix}
QA_e + A_e^T Q & QB_e & QE
\
B_e^T Q & -I & 0
\
E^T Q & 0 & -\delta I
\end{bmatrix} < 0
\]  

for a sufficiently large scalar $\delta \in \mathbb{R}_+$, where $E$ is defined in (24). Consider a nonsingular matrix

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

Pre- and post-multiplying by $T^T$ and $T$, respectively, (28) then becomes (21), where the equation $A_e = A + E\bar{A}_r$, and the substitutions $\delta = \frac{1}{\delta} > 0$, $\bar{Q} = \frac{1}{\delta}Q > 0$ are used.

Next, we observe that the following implications hold.

\[
\begin{bmatrix}
Q & C_e^T \\
C_e & R
\end{bmatrix} > 0 \iff \begin{bmatrix}
\hat{Q} & \delta C_e^T
\end{bmatrix} > 0
\]

\[
\text{tr}(R) < \gamma \iff \frac{1}{\delta} \text{tr}(\hat{R}) < \gamma \iff \text{tr}(\hat{R}) < \gamma,
\]

with $\hat{R} = \delta R$ and $\gamma = \delta \gamma$. As a result, (22) and (23) are equivalent to (26) and (27), respectively.

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

\[
\min_{Q > 0, \hat{W} \in \mathcal{M}} \text{tr}(R)
\]

s.t. (21) and (22) hold,

where $\hat{R} = \delta R$ with a given $\delta \in \mathbb{R}_+$. Note that the constraint (22) can be solved efficiently using standard LMI solvers, while (21), due to the nonlinearity term $A_e^T \bar{A}_r$, is a bilinear matrix inequality, which causes the major challenge in solving the problem (29).

To handle the bilinear constraint (21), we adopt the technique called psd-convex-concave decomposition [23].

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

\[
\Phi_k(\lambda w_1 + (1 - \lambda)w_2) \leq \lambda \Phi_k(w_1) + (1 - \lambda)\Phi_k(w_2),
\]

holds for all $\lambda \in [0, 1]$ and $w_1, w_2 \in \mathbb{R}^n$. The pair $(\Phi_1, \Phi_2)$ is called a psd-convex-concave decomposition of $\Phi$.

Consider the bilinear inequality (21), and define the following matrix-valued mapping:

\[
\Phi(\hat{Q}, \hat{\delta}, \hat{W}) = \psi(\hat{Q}, \hat{\delta}) + \varphi(\hat{W}),
\]

where

\[
\psi(\hat{Q}, \hat{\delta}) = \begin{bmatrix}
\hat{Q}A + A^T \hat{Q} & \hat{Q}B_e & \hat{Q}E
\
B_e^T \hat{Q} & -I & 0
\
E^T \hat{Q} & 0 & -\delta I
\end{bmatrix},
\]

\[
\varphi(\hat{W}) = \begin{bmatrix}
-A_e^T \bar{A}_r & 0 & \hat{A}_r^T
\end{bmatrix},
\]

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

**Lemma 4.** The matrix-valued mapping $\Phi(\hat{Q}, \hat{\delta}, \hat{W})$ in (31) is psd-convex-concave.

**Proof.** Note that the matrix $\psi(\hat{Q}, \hat{\delta})$ in (31) 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 (31) is psd-convex.

With the structure of $\bar{A}_r$ in (24), the only nonlinear submatrix in $-\varphi(\hat{W})$ can be expressed as

\[
\bar{A}_r^T \bar{A}_r = \begin{bmatrix}
0 & 0 \\
0 & \varphi_a(\hat{W})
\end{bmatrix},
\]

with $\varphi_a(\hat{W}) := S_r^T M^{-1} \hat{B}_e \hat{W} \hat{B}_o^T (S_r^T)^T S_r^T \hat{B}_o \hat{W} \hat{B}_e^T M^{-1} S_r$.

Then, showing the psd-convexity of $-\varphi(\hat{W})$ in (33) 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

\[
\varphi_a(W_\lambda) - \lambda \varphi_a(W_1) - (1 - \lambda) \varphi_a(W_2)
\]

\[
= S_r^T M^{-1} \hat{B}_e \hat{W}_\lambda \hat{B}_o^T (S_r^T)^T S_r^T \hat{B}_o \hat{W}_\lambda \hat{B}_e^T M^{-1} S_r
\]

\[=- t S_r^T M^{-1} \hat{B}_e \hat{W}_1 \hat{B}_o^T (S_r^T)^T S_r^T \hat{B}_o \hat{W}_1 \hat{B}_e^T M^{-1} S_r
\]

\[=- (1 - \lambda) S_r^T M^{-1} \hat{B}_e \hat{W}_2 \hat{B}_o^T (S_r^T)^T S_r^T \hat{B}_o \hat{W}_2 \hat{B}_e^T M^{-1} S_r
\]

\[=(1 - \lambda)(1 - \lambda) S_r^T M^{-1} \hat{B}_e \hat{W}_1 \hat{B}_o^T (S_r^T)^T S_r^T \hat{B}_o \hat{W}_1 \hat{B}_e^T M^{-1} S_r
\]

\[= -\lambda(1 - \lambda) V_1 (W_1 - W_2) V_2 (W_1 - W_2) V_2^T \leq 0,
\]

where $V_1 = S_r^T M^{-1} \hat{B}_e$, and $V_2 = \hat{B}_o^T (S_r^T)^T S_r^T \hat{B}_o$. Since $-t(1 - t) \leq 0$, and $V_1 (W_1 - W_2) V_2 (W_1 - W_2) V_1^T \geq 0$, 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 (30), 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 (31) is psd-convex-concave.

The psd-convex-concave decomposition in (31) allows us to linearize the optimization problem (29) at a stationary point $\hat{W} \in \mathcal{M}$. To simplify the optimization procedure, we introduce a new optimization variable $\mu$ to eliminate the equality constraint $\hat{B}_o \hat{W} = 0$ in (13), where $\hat{W} = \text{Diag}(\hat{w})$, and $m$ is the number of edges in the reduced network.

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

\[
\hat{B} \hat{w} = (\hat{B}_a \hat{B}_b) P \hat{w} = (\hat{B}_a \mu_a + \hat{B}_b \mu_b) = 0,
\]

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

\[
\hat{w} = P^T \begin{bmatrix}
-I_{\bar{m}} \\
\hat{B}_a^{-1} \hat{B}_b
\end{bmatrix} \mu,
\]

which projects the weights $\hat{w}$ into $\ker(\hat{B})$. Thereby, we rewrite the constraint $\hat{W} \in \mathcal{M}$ as $\mu \in \mathbb{R}^{\mu_a}$ (36). Now, we redefine the matrix-valued mapping $\varphi(\hat{W})$ in (31) as

\[
\phi(\mu) = \varphi(\hat{W}),
\]
where the derivative of $\phi$ which is defined as

$$ D\phi(\mu)[h] = \sum_{i=1}^{m} h_{i} \frac{\partial \phi}{\partial \mu_{i}}(\mu), \quad \forall h \in \mathbb{R}^{m}. \tag{38} $$

Given a point $\mu^{(k)} \in \mathbb{R}^{m}_{+}$, the linearized formulation of the problem (29) at $\mu^{(k)}$ is formulated as

$$ \min_{Q>0, \mu \in \mathbb{R}^{m}_{+}} f(\mu) = \text{tr}(R) $$

s.t. $$ \frac{\hat{Q}}{\hat{e}} > 0, \quad \hat{R} = \hat{\delta}R > 0 $$

$$ \psi(\hat{Q}, \hat{\delta}) + \varphi(\hat{W}^{(k)} + D\phi(\mu^{(k)})[\mu - \mu^{(k)}]) < 0, $$

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

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

with $j = 1, \ldots, m$. Notice that the optimization problem (39) is convex, of which the global optimum can be solved efficiently using standard SDP solvers e.g., SeDuMi [24].

Based on Lemma 4 and (39), we are now ready to present an algorithmic approach for solving the minimization problem in (29) in an iterative fashion, see Algorithm 1, in which $\varepsilon \in \mathbb{R}_{+}$ 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 (36), it will guarantee $\hat{W}^{0} \in \mathcal{M}$, i.e., the reduced graph is balanced. Furthermore, we can also initialize $\mu$ using the outcome of graph clustering projection in [18], [19]. Specifically, from a given clustering II, we construct an initial reduced Laplacian matrix in (11) as $\hat{L}_{b}^{(0)} := \Pi^{T} L_{b} \Pi$, with $L_{b}$ the Laplacian matrix of the balanced graph $\hat{G}_{b}$. Then, the initial weight of the edge $(i,j)$ in the quotient graph $\hat{G}_{b}$ 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 [23], and it means that a local optimum can be obtained. More 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], [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], [18]. The topology of the 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 5 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, with incidence matrix

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

There are 8 edges in the quotient graph, and each edge is assigned with a symbolic weight as labeled in Fig. 3. These variables, determining the reduction error, are to be determined by our optimization approach.

First, the parameterized reduced model in (10) of the quotient graph is generated with matrices

$$ \hat{L}_{b} = \begin{bmatrix} w_{1} + \bar{w}_{1} & w_{2} & w_{1} & w_{2} & 0 & 0 & -w_{3} & w_{3} & 0 & 0 \\ -w_{3} & w_{3} & 0 & 0 & 0 & 0 & w_{4} & w_{5} & w_{6} & -w_{5} & -w_{6} \\ -w_{5} & w_{5} & 0 & 0 & 0 & 0 & \bar{w}_{7} & w_{7} & w_{8} & -w_{8} & w_{7} \\ -w_{6} & w_{6} & 0 & 0 & 0 & 0 & w_{8} & w_{7} & w_{5} & \bar{w}_{7} & w_{8} \end{bmatrix}, $$

$$ \hat{M} = \begin{bmatrix} 2.1921 & 0.6803 & 0.3779 & 0.0756 & 0.4157 \\ 0.6803 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.0756 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \hat{H} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, $$

and the weight vector $\hat{w}$ in (8) satisfies the following constraints for a balanced graph: $\hat{w}_{1} = \hat{w}_{3}$, $\hat{w}_{2} = \hat{w}_{4} + \hat{w}_{8}$, $\hat{w}_{6} = \hat{w}_{8}$, $\hat{w}_{5} = \hat{w}_{7}$.

Next, we implement Algorithm 1 to solve the optimization problem (29) with $\mu = [\hat{w}_{1}, \hat{w}_{2}, \hat{w}_{5}, \hat{w}_{6}]^{T} \in \mathbb{R}_{+}^{4}$ as the optimization variable. Particularly, the SeDuMi solver [24] is adopted to solve the convex problem (39).

We choose the initial edge weights obtained by the clustering-based projection [18], [19], which gives $w^{(0)} = [0.6803, 0.2268, 0.6803, 0.0756, 0.0756, 0.1512, 0.0756, 0.1512]$ and the approximation error $\|G_{e}(s)\|_{H_{2}} = 0.0322$. With $\hat{\delta} = \varepsilon = 10^{-5}$, Algorithm 1 stops after 72 iterations. The convergence trajectory of the resulting $H_{2}$ reduction error is shown in Fig. 4. 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 $\|G_{e}(s)\|_{H_{2}} = 0.0187$. Through iteration, the edge weighting method further reduces the error by 41.93%, compared to the clustering-based projection. Therefore, our method can provide a reduced network systems with a better $H_{2}$ approximation error.
networked high-order linear subsystems are also of interest. The advantage of this proposed model reduction method in [12], the feasibility of this method is illustrated by an example. The advantage of this proposed model reduction method in [12], the feasibility of this method is illustrated by an example. 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 $H_2$ norm of the approximation error is small. Finally, compared with the projection-based method in [11], the $H_\infty$ error bounds, “Mathematics of Control, Signals, and Systems,” vol. 30, no. 1, p. 6, 2018.

X. Cheng and J. M. A. Scherpen, “Clustering-based model reduction of laplacian dynamics with weakly connected topology,” IEEE Transactions on Automatic Control, 2019.

N. Martin, P. Frasca, and C. Canudas-de Wit, “Large-scale network reduction towards scale-free structure,” IEEE Transactions on Network Science and Engineering, vol. 6, no. 4, pp. 711–723, 2018.

M. Meshahi and M. Egerstedt, Graph Theoretic Methods in Multiagent Networks. Princeton University Press, 2010.

G. Pipeleers, B. Demuynck, J. Swevers, and L. Vandenberghe, “Extended LMI characterizations for stability and performance of linear systems,” Systems & Control Letters, vol. 58, no. 7, pp. 510–518, 2009.

Q. T. Dinh, S. Gunussoy, W. Michiels, and M. Diehl, “Combining convex-concave decompositions and linearization approaches for solving BMIs, with application to static output feedback,” IEEE Transactions on Automatic Control, vol. 57, no. 5, pp. 1377–1390, 2011.

J. F. Sturm, “Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones,” Optimization Methods and Software, vol. 11, no. 1-4, pp. 625–653, 1999.

V. CONCLUSIONS

In this paper, the $H_2$ 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 $H_2$ 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 works, we will improve the effectiveness of the iterative algorithm such that the obtained solution is not restricted to a local optimum. Moreover, an extension to networked high-order linear subsystems are also of interest.

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

Fig. 3: The quotient graph obtained by clustering, where the controlled vertices are labeled with red color, and measured vertices are indicated by shadow. The weights of the edges are parameters to be determined.

Fig. 4: Approximation errors of clustering-based projection and the proposed edge weighting method.

REFERENCES

[1] J. A. Fax and R. M. Murray, “Graph Laplacians and stabilization of vehicle formations,” 2001.

[2] F. Dörfler, J. W. Simpson-Porco, and F. Bullo, “Electrical networks and algebraic graph theory: Models, properties, and applications,” Proceedings of the IEEE, vol. 106, no. 5, pp. 977–1005, 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 Transactions on Signal Processing, vol. 56, no. 4, pp. 1667–1684, 2008.

[4] A. V. Proskurnikov, A. S. Matveev, and M. Cao, “Opinion dynamics in social networks with hostile camps: Consensus vs. polarization,” IEEE Transactions on Automatic Control, vol. 61, no. 6, pp. 1524–1536, 2015.

[5] T. H. Summers and J. Lygeros, “Optimal sensor and actuator placement in complex dynamical networks,” IFAC Proceedings Volumes, vol. 47, no. 3, pp. 3784–3789, 2014.

[6] A. J. Gates and L. M. Rocha, “Control of complex networks requires both structure and dynamics,” Scientific Reports, vol. 6, no. 1, pp. 1–11, 2016.

[7] T. Ishizaki, A. Chakrabortty, and J.-I. Imura, “Graph-theoretic analysis of power systems,” Proceedings of the IEEE, vol. 106, no. 5, pp. 951–982, 2018.

[8] A. C. Antoulas, Approximation of Large-Scale Dynamical Systems. Philadelphia, USA: SIAM, 2005.

[9] X. Cheng, J. M. A. Scherpen, and B. Besselink, “Balanced truncation of networked linear passive systems,” Automatica, vol. 104, pp. 17–25, 2019.

[10] E. Biyik and M. Arcak, “Area aggregation and time-scale modeling for sparse nonlinear systems,” Systems & Control Letters, vol. 57, no. 2, pp. 142–149, 2008.

[11] 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,” Journal of Mathematical Chemistry, vol. 51, no. 9, pp. 2401–2422, 2013.

[12] N. Monshizadeh, H. L. Trentelman, and M. K. Camlibel, “Projection-based model reduction of multi-agent systems using graph partitions,” IEEE Transactions on Control of Network Systems, vol. 1, pp. 145–154, Jun. 2014.

[13] B. Besselink, H. Sandberg, and K. H. Johansson, “Clustering-based model reduction of networked passive systems,” IEEE Transactions on Automatic Control, vol. 61, no. 10, pp. 2958–2973, Oct 2016.

[14] T. Ishizaki, K. Kashima, J. I. Imura, and K. Aihara, “Model reduction and clustering of large-scale bidirectional networks,” IEEE Transactions on Automatic Control, vol. 59, pp. 48–63, 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. Mlinarič, S. Grundel, P. Benner, and H. L. Trentelman, “Model reduction of linear multi-agent systems by clustering with $H_2$ and $H_\infty$ error bounds,” Mathematics of Control, Signals, and Systems, vol. 30, no. 1, p. 6, 2018.

[17] X. Cheng, Y. Kawano, and J. M. A. Scherpen, “Model reduction of multi-agent systems using dissimilarity-based clustering,” IEEE Transactions on Automatic Control, vol. 64, no. 4, pp. 1663–1670, April 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 Proceedings of IEEE 56th Conference on Decision and Control, Melbourne, Australia, Dec 2017, pp. 3823–3828.

[19] X. Cheng and J. M. A. Scherpen, “Clustering-based model reduction of laplacian dynamics with weakly connected topology,” IEEE Transactions on Automatic Control, 2019.

[20] N. Martin, P. Frasca, and C. Canudas-de Wit, “Large-scale network reduction towards scale-free structure,” IEEE Transactions on Network Science and Engineering, vol. 6, no. 4, pp. 711–723, 2018.

[21] M. Mesbah and M. Egerstedt, Graph Theoretic Methods in Multiagent Networks. Princeton University Press, 2010.