university of groningen

faculty of mathematics and natural sciences

Model Reduction by Clustering of Networked Multi-Agent Systems

Master Project Applied Mathematics

April 2017

Student: J.T. de Jonge

Supervisors: dr. ir. B. Besselink and prof. dr. H.L. Trentelman

Second assessor: dr. K. Efstathiou

Abstract

In this thesis, a clustering-based method of model reduction of networked multi-agent systems will be considered. Starting point is the method described in [3] that uses a one step clustering method for networks with tree graph topology to reduce the dimension of the model by one. It can be applied until an approximated model of desired order is achieved. This thesis generalises the method for a certain class of systems to arbitrary graphs. The mechanism used to select what nodes to cluster, is based on an approximation of the output and control energy. After introducing an edge system, it will be possible to study the level of controllability and observability of edges individually using these energies. Bounds on the output energy and the control energy are determined using a Lyapunov and Riccati inequality respectively. Since for arbitrary networks, these energies are not necessarily finite, we introduce without loss of generality a constraint on the space considered. This guarantees finite energies. Next, the existence of suitable solutions to the matrix inequalities is proven. It is also shown that the clustering method preserves the property of synchronisation.

Contents

1 Introduction

Multi-agent systems, distributed control and complex networks are subjects that have been heavily studied in the last decades. Networked multi-agent systems are networks consisting of a group of systems called agents. The interconnection between these agents is through a communication topology, and can be modeled by a network graph. The agents exchange information only with their neighbours in this topology. Networked dynamical systems are found in a variety of disciplines, ranging from for example electrical power grids to ecological networks [15]. A widely studied problem in networks, is the problem of consensus and synchronisation, where the goal is to have agents in the network agree upon some quantity, while only using information that is locally available. Potentially, these networked systems are large-scale, involving a large set of agents. This complicates analysis, and possibly control, of these networks. Therefore, approximation techniques have been developed to provide systems of lower complexity. For example, we mention the general techniques of balanced truncation [14] and moment matching [2]. These model reduction techniques can be shown to have some satisfactory properties, but they have disadvantages as well. One of these drawbacks is that they do generally not preserve the interconnection topology in the approximation. Finding controllers for the large-scale dynamical systems, based on the approximated system, is therefore complicated. To overcome this problem, alternative reduction techniques have been investigated that preserve the topology in some sense. For example, in [11], edges, representing connections in the topology, that are deemed of lesser importance are removed. Other methods have introduced the concept of clustering-based algorithms, e.g. [10]. A relatively easy physical interpretation is one of the assets of clustering techniques, in which certain vertices are joined to form a single group such that a reduced order network arises. Some research has been done in finding convenient sets of vertices to cluster. In [13] a graph partition called *almost equitable graph partitions* was introduced that allows for an explicit error analysis. However, these partitions require restrictive assumptions on the graph topology.
In this thesis, the approach taken follows that of [3], where a clustering method is introduced that aggregates two vertices that show similar behaviour. The similarity between vertices is determined by an analysis on the edges. To that end, an *edge system* is introduced in which it is possible to analyse properties of individual edges. An edge Laplacian is defined such that the dynamics over edges can be studied. Using such an edge system admits giving a measure on the level of observability and controllability of each edge. These will be called *edge observability* and *edge controllability*. This is usually done by computing the Gramians for the system. The Gramians of a system give a measure on edge observability and edge controllability respectively by using an energy notion, and can be computed by a Lyapunov or Riccati matrix equality. As these Gramians are not easily interpreted, in this thesis an energy bound is found using matrix inequalities. This bound is chosen to have a diagonal structure, which eases interpretation. In the case of tree graphs considered in [3], the edge Laplacian is found to have convenient properties, such as having positive

eigenvalues. In the case of arbitrary connected graphs, these properties generally do not hold, resulting in a possibly infinite energy for some edges. As this has no physical interpretation, these must somehow be excluded. The main contribution of this thesis is the extension of the method in [3] to the case of non-tree graphs for a certain class of interacting systems. This is done by using the fact that for graphs with cycles there exist linearly dependent edges and therefore a vector representing the edges can be restricted to a subspace. Using this linear dependence, Lyapunov and Riccati matrix equalities similar to the case of graphs without cycles can be found that can give the level of observability and controllability for graphs with cycles as well. Matrix inequalities then give energy bounds that will be used to identify the least important edge.
Solutions to the matrix inequalities that give the edge controllability and edge observability are shown to exist for the class of interacting systems considered. More specific, existence is shown for the class of systems defined on undirected weighted graphs, where the dynamics of each system is given as a single integrator with mass. This same class also preserves the network topology, and can be shown to preserve the property of synchronisation as well. After having found an edge that contributes the least to the network, based on its level of controllability and observability, the vertices that are connected by the chosen edge are clustered, thereby reducing the order of the model by one. This leads to a new model that has a new interconnection topology with a suitable physical interpretation. The clustering is done by a projection matrix.
The remainder of this thesis is organised as follows. In the next chapter the problem statement is formulated. In chapter 3 some preliminaries about graph theory and multi-agent systems are introduced. Subsequently, the system that will be used troughout this thesis is introduced in chapter 4, after which the main results are presented in chapter 5. Next we consider how the main results are used in clustering, and an example is given. Finally, conclusions are contained in chapter 8.

2 Problem statement

Consider a network of n identical subsystems Σ_i, where $i \in \{1, 2, \ldots, n\}$. The dynamics of each subsystem is given as

$$\Sigma_i : \begin{cases} m_i \dot{x}_i & = v_i, \\ z_i & = x_i. \end{cases} \tag{1}$$

The state of each subsystem is represented by $x_i(t) \in \mathbb{R}$; $v_i(t) \in \mathbb{R}$ is the input and $z_i(t) \in \mathbb{R}$ the system output. The number m_i can be regarded as the mass and satisfies $m_i > 0$. These subsystems can be interconnected to form a multi-agent system. The interconnection between subsystems is given by

$$v_i = \sum_{j=1, j\neq i}^{n} w_{ij}(z_j - z_i) + \sum_{j=1}^{m} g_{ij} u_j, \tag{2}$$

with g_{ij} describing the strength and location of the external input $u_j(t) \in \mathbb{R}$, and with the strength of the coupling between vertices i and j given by w_{ij}. The interconnection between these subsystems can also be represented by a graph, and by a matrix representation of that graph. Specifically, the corresponding graph Laplacian is defined as

$$L_{ij} = \begin{cases} -w_{ij} & \text{for } j \neq i, \\ \sum_{j=1, j\neq i}^{n} w_{ij} & \text{for } j = i. \end{cases} \tag{3}$$

We proceed by collecting the g_{ij} in matrix form, such that the $\{i, j\}^{th}$ element of G equals $G_{ij} = g_{ij}$. Similarly, we can define the output of the networked system as a function of the output of each subsystem by

$$y_j = \sum_{i=1}^{n} h_{ji} z_i. \tag{4}$$

As before, we collect the output terms h_{ij} in a matrix, such that $H_{ij} = h_{ij}$. Together, the equations $(1) - (4)$ lead to the system of equations

$$\Sigma : \begin{cases} M\dot{x} &= -Lx + Gu \\ y &= Hx, \end{cases} \tag{5}$$

with $x = (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n, y \in \mathbb{R}^p, u \in \mathbb{R}^m$. In the case of large-scale dynamical systems, this system can be too large to compute effectively. In order to tackle this problem, an approximation method is needed that can return a lower dimensional system of arbitrary order, and that preferably preserves some properties of the original system. Particularly, we would like to preserve the property of synchronisation. Since clustering has a very intuitive physical interpretation, we propose to do this model reduction by clustering. We try to find a new system of the following form:

$$\hat{\Sigma} : \begin{cases} \hat{M}\dot{\xi} &= -\hat{L}\xi + \hat{G}u \\ \hat{y} &= \hat{H}\xi, \end{cases} \tag{6}$$

where the state vector $\xi(t)$ of the new system is of lower dimensional order than the state vector $x(t)$ of Σ, and the matrix $\hat{L}$ should be a Laplacian matrix for the reduced order system. The new matrices are $\hat{M}, \hat{L}, \hat{G}$ and $\hat{H}$ are defined via a projection matrix representing the clustering method. This method should be possible to use on any connected graph. This results in the following problem statement

Problem 1: *Given is a system as described in* (5)*. How can we do a clustering-based model reduction on arbitrary undirected weighted graphs that preserves the property of synchronisation and results in a system as described in* (6)*?*

3 Graph Theory

In this section the definitions and notation that will be used troughout the thesis will be introduced. An undirected graph $\mathcal{G}$ can be characterized as $\mathcal{G}(\mathcal{V}, \mathcal{E})$, with $\mathcal{V}$ denoting the set of vertices or nodes, $\mathcal{V} = \{v_1, v_2, \ldots, v_n\}$ and an unordered set $\mathcal{E}$ denoting all edges in the graph, $\mathcal{E} \subseteq \{\{v_i, v_j\} | v_i, v_j \in \mathcal{V}\}$. These edges represent the network topology, for example which nodes can communicate with each other. An undirected weighted graph is defined as $\mathcal{G}(\mathcal{V}, \mathcal{E}, \{w_{ij}\}_{\{v_i,v_j\}\in\mathcal{E}})$, where $\mathcal{G}(\mathcal{V}, \mathcal{E})$ represents an undirected graph, and the set $\{w_{ij}\}_{\{v_i,v_j\}\in\mathcal{E}}$ attributes positive weights w_{ij} to edge $\{v_i, v_j\} \in \mathcal{E}$. In the case of undirected graphs, the weights are symmetric, i.e. $w_{ij} = w_{ji}$. Only undirected weighted graphs are considered in this thesis. For undirected graphs, a path is defined as a series of adjacent nodes connected by edges. In the case of undirected graphs, there is no orientation on the edges. That is to say, the information flow between nodes connected by an edge can go in both directions. A cycle is a path for which the first node and the last node of a path coincide, and only the first and last node coincide. We call a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ connected if for every pair of vertices in $\mathcal{V}$ there exists a path in $\mathcal{G}(\mathcal{V}, \mathcal{E})$ that connects these vertices. A self-loop is an edge that connects a vertex to itself. We consider only graphs without self-loops in this thesis. It is also noted that every graph used in this thesis is assumed to be connected. If not, the theory can be applied to the connected components of the graph. A graph is called a subgraph $\mathcal{G}'(\mathcal{V}', \mathcal{E}')$ of $\mathcal{G}$ if $\mathcal{V}' \subseteq \mathcal{V}$ and $\mathcal{E}' \subseteq \mathcal{E}$. A so-called *spanning tree* $\mathcal{T}$ is a subgraph for which $\mathcal{T} = (\mathcal{V}, \mathcal{E}_\mathcal{T})$ and that furthermore contains no cycles. The elements of the edge set $\mathcal{E}_\tau$ are such that they connect any pair of vertices in $\mathcal{V}$ by exactly one path. The *incidence matrix* is a matrix that conveniently combines the information about the graph topology. To define the incidence matrix, an arbitrary orientation is given to each edge in the graph. This will be denoted as $\mathcal{G}^o$, where the superscript indicates that an orientation is given. Then, the incidence matrix E is defined as

$$E(\mathcal{G}^o) = (e_{ij}) = \begin{cases} +1 & \text{if } v_i \text{ is the head of } e_j \\ -1 & \text{if } v_i \text{ is the tail of } e_j \\ 0 & \text{otherwise,} \end{cases} \tag{7}$$

where e_j is an element of the edge set $\mathcal{E}$. The incidence matrix has dimensions $|\mathcal{V}|\times|\mathcal{E}|$. We denote the number of nodes of a graph as n and the number of edges of a graph as n_e, so that $E(\mathcal{G}^0) \in \mathbb{R}^{n\times n_e}$. For more information on graph theory, see for example [12], [8].

A vector consisting of the same number for all elements is represented by a bold version of that number. For example, a vector of all ones is represented by $\mathbf{1}$. A matrix with every element equal will be represented in the same way, but with a subscript describing the size of the matrix (e.g. $\mathbf{0}_{n\times n}$). A symmetric $n \times n$ matrix is called positive semi-definite if $x^T M x \geq 0$, $\forall x \in \mathbb{R}^n$. If the matrix satisfies the strict inequality for all $x \neq 0$, it is said to be positive definite. The matrix inequality $A \geq B$ is said to hold if $A - B \geq 0$. Also, a matrix A is

called strictly diagonally dominant if $a_{ii} > \sum_{i \neq j} |a_{ij}|, \ \forall i$. The complex right halve plane is denoted by $\mathbb{C}^+ := \{s \in \mathbb{C} \,|\, \mathrm{Re}(s) > 0\}$ and the left halve plane by $\mathbb{C}^- := \{s \in \mathbb{C} \,|\, \mathrm{Re}(s) < 0\}$.

3.1 Linear Dependence of Cycle Edges

An arbitrary connected graph can be written as the union of two disjoint subgraphs on the same vertex set as $\mathcal{G} = \mathcal{G}_\tau \cup \mathcal{G}_c$. The graph $\mathcal{G}_\tau$ represents an underlying spanning tree, and then $\mathcal{G}_c$ represents the remaining edges, which must necessarily be the edges that close a cycle.
Since the columns of the incidence matrix correspond to the edges, by permutation of the edge numbering, it is always possible to write

$$E(\mathcal{G}) = \begin{pmatrix} E(\mathcal{G}_\tau) & E(\mathcal{G}_c) \end{pmatrix}, \tag{8}$$

which will, for ease of notation, be denoted as $E = \begin{pmatrix} E_\tau & E_c \end{pmatrix}$ with $E_\tau = E(\mathcal{G}_\tau)$ and $E_c = E(\mathcal{G}_c)$. The incidence matrix E has dimensions $E \in \mathbb{R}^{n \times n_e}$. If a graph contains cycles, then $n_e \geq n$ and the columns of E can not be linear independent. In fact, as it is stated in [19], it is possible to write the edges belonging to the cycles as a linear combination of the edges of the chosen spanning tree. Thus, there exists a matrix T such that

$$E_\tau T = E_c, \tag{9}$$

with T given by

$$T = (E_\tau^T E_\tau)^{-1} E_\tau^T E_c.$$

We remark that the choice of spanning tree is not unique and that choosing different spanning trees corresponds to different matrices T. However, results in this thesis are independent of the choice of the spanning tree $\mathcal{G}_\tau$. Using the equation of linear dependence of the cycle edges, another way to represent the incidence matrix of the graph is easily obtained as

$$E = E_\tau \begin{pmatrix} I_{n-1} & T \end{pmatrix} = E_\tau S, \tag{10}$$

where S is defined as $S = \begin{pmatrix} I_{n-1} & T \end{pmatrix}$. This matrix S can be related to a number of structural properties of the graph [19]. For example, the number of spanning trees in a graph can be found using this matrix as

$$\tau(\mathcal{G}) = \det(SS^T). \tag{11}$$

3.2 Graph Laplacian

As explained before, a graph $\mathcal{G}$ can be represented by a matrix called the graph Laplacian. There are several definitions for this matrix representation of the graph. Among other ways, the graph Laplacian of an undirected weighted graph can be written as

$$L = ERE^T = E_\tau S R S^T E_\tau^T, \tag{12}$$

where the last equality is obtained using (10) and where the positive definite diagonal matrix R corresponds to the weight of the corresponding edge in the sense that $R_{ii} = w(e_i)$. The matrix E is the incidence matrix of the graph, given an arbitrary orientation. The fact that the definition of the graph Laplacian in (3) and the above are equivalent can be seen as follows [5]. Recall that the product ERE^T with a diagonal matrix R can be written as $(ERE^T)_{ij} = \sum_k E_{ik}R_{kk}E^T_{kj}$. This can be used to distinguish the two different cases in (3), namely $i = j$ and $i \neq j$. We start with $i \neq j$. In this case, we see that

$$\begin{aligned}(ERE^T)_{ij} &= \sum_k E_{ik}R_{kk}E^T_{kj} \\ &= \sum_k E_{ik}E_{jk}R_{kk} \\ &= (+1)(-1)w_{ij}.\end{aligned}$$

The last step follows from the fact that there is only a non-zero entry in the i^{th} row of incidence matrix, if there is an edge $\{i, j\}$. If $E_{ik} = 1$ it means that the edge k arrives at node i. Then it must mean that the same edge k leaves node j, so that $E_{jk} = -1$. The weigth of this edge is given by w_{ij}, so the last equality holds. For the case where $i = j$, we can proceed similarly. This leads to

$$\begin{aligned}(ERE^T)_{ij} &= \sum_k E_{ik}R_{kk}E^T_{kj} \\ &= \sum_k E^2_{ik}R_{kk} \\ &= \sum_{j\neq i} w_{ij}.\end{aligned}$$

We conclude that the graph Laplacian defined in (3) can indeed be written as $L = ERE^T$. It is well known that the graph Laplacian L is positive semi-definite. Also, the graph is connected if and only if the graph Laplacian has a zero eigenvalue with algebraic multiplicity equal to one [12]. This can easily be seen using (12). We write $x^T Lx = x^T ERE^T x$. This can be written as $x^T Lx = ||R^{1/2}E^T x|| \geq 0$, $\forall x \in \mathbb{R}^n$. From the definition of the incidence matrix E, it is clear that $E^T\mathbf{1} = 0$ for a graph. This turns out to be the only vector in the nullspace of E^T for connected graphs. The graph Laplacian for connected undirected weighted graphs contains a zero eigenvalue with an algebraic multiplicity of one, and all other eigenvalues are positive [12].
We also define an *effective graph Laplacian* as

$$L_{\text{eff}} = M^{-1}ERE^T, \tag{13}$$

where the matrix M is the mass matrix defined earlier. Defining an effective graph Laplacian allows us to conserve the class of undirected weighted networks by the proposed clustering method [17]. The effective graph Laplacian incorporates the weights from both the edges and the vertices. In line with this definition of an effective graph Laplacian, we can also define *effective weights*.

Definition 1. *If w_{ij} is the weight associated to the edge $\{i,j\} \in \mathcal{E}$, and m_i denotes the weight associated to vertex $i \in \mathcal{V}$, then we define the effective weight of vertex i from edge $\{i,j\}$ as $\dfrac{w_{ij}}{m_i}$.*

Then it can be seen that the $\{i,j\}^{th}$ entry of L_{eff} denotes the effective weight of vertex i from edge $\{i,j\}$ [13]. Since the product of two symmetric positive semidefinite matrices has non-negative eigenvalues as well, we see that the eigenvalues of L_{eff} are all non-negative. Also, for connected graphs the zero eigenvalue has an algebraic multiplicity equal to one.
From [19], we introduce the following matrices for a change of basis:

$$S_\tau = \begin{pmatrix} E_\tau^T \\ \alpha\mu^T \end{pmatrix}, \quad S_\tau^{-1} = \begin{pmatrix} M^{-1}E_\tau(E_\tau^T M^{-1} E_\tau)^{-1} & \mathbf{1} \end{pmatrix}. \tag{14}$$

In these matrices, E_τ is the incidence matrix of the graph formed by the spanning tree as before. The vector μ^T is a left eigenvector of L_{eff} corresponding to the eigenvalue 0, i.e. $\mu^T L_{\text{eff}} = 0$, and $\alpha \in \mathbb{R}$ is a scaling factor. Because the algebraic multiplicity of the zero eigenvalue is equal to one, the left eigenvector $\alpha\mu^T$ is unique, up to scaling. It can be seen by inspection that this eigenvector is of the form $\mu = M\mathbf{1}$. It is clear that the elements of the left eigenvector are positive. Therefore, we can safely assume that $\alpha\mu^T\mathbf{1} = 1$ for a suitable choice of α.
Computing $S_\tau L S_\tau^{-1}$ shows that the effective graph Laplacian is similar to

$$L_{\text{eff}} \sim \begin{pmatrix} E_\tau^T M^{-1} E_\tau S R S^T & 0 \\ 0 & 0 \end{pmatrix}. \tag{15}$$

To simplify notation later on, the upper left block matrix is denoted as

$$L_\tau = E_\tau^T M^{-1} E_\tau S R S^T. \tag{16}$$

We call this matrix the *spanning tree edge Laplacian*. From this partitioning of the graph Laplacian, together with the fact that for connected graphs L_{eff} has a zero eigenvalue with multiplicity one, it is clear that L_τ has only positive eigenvalues, and hence $\sigma(L_\tau) \subset \mathbb{C}^+$.

3.3 Edge Laplacian

Another important matrix used in this thesis is the edge Laplacian. For undirected weighted graphs with a corresponding mass matrix M, we define the edge Laplacian as

$$L_e = E^T M^{-1} E R = S^T E_\tau^T M^{-1} E_\tau S R. \tag{17}$$

The edge Laplacian gives information about the adjacency of edges in the sense that edges that are not adjacent, that do not share a vertex, represent a zero value in the edge Laplacian. The non-zero eigenvalues of the edge Laplacian, L_e, are equal to the non-zero eigenvalues of the effective graph Laplacian, L_{eff} [9]. In

turn, these equal the eigenvalues of the graph spanning tree edge Laplacian L_τ as was seen in (15). Next to these eigenvalues, the edge Laplacian contains a zero eigenvalue for every independent cycle that the corresponding graph contains. We note that $L_e \in \mathbb{R}^{n_e \times n_e}$ and that there are $n-1$ non-zero eigenvalues. By comparison, it is seen that the number of zero eigenvalues equals $n_e - n + 1$. This can also be seen by using another similarity transformation [19]. Define

$$S_e = \begin{pmatrix} (SS^T)^{-1}S \\ V_e(\mathcal{G})^T \end{pmatrix}, \quad S_e^{-1} = \begin{pmatrix} S^T & V_e(\mathcal{G}) \end{pmatrix}.$$

The matrix S in these matrices is the same as in (10), and the matrix $V_e(\mathcal{G})$ is a matrix representation of an orthonormal basis for the nullspace of $L_e(\mathcal{G})$. By computing $S_e L_e S_e^{-1}$, we see that

$$L_e \sim \begin{pmatrix} E_\tau^T M^{-1} E_\tau S R S^T & 0 \\ V_e(\mathcal{G})^T L_e S^T & \mathbf{0}_{(n_e-n+1)} \end{pmatrix},$$

and it follows from the block-triangular structure of the right hand side that the non-zero eigenvalues of L_e indeed equal the eigenvalues of L_τ, see (16). A very important observation for this thesis relates the edge Laplacian L_e to the graph spanning tree Laplacian L_τ, which is of importance since it is known that L_τ has strictly positive eigenvalues. This relation reads

$$S^T L_\tau = S^T E_\tau^T M^{-1} E_\tau S R S^T = L_e S^T, \tag{18}$$

and is easily obtained by comparing (16) and (17). It shows that im S^T is L_e − invariant.

4 Edge Dynamics and Synchronisation

We have defined a dynamical system on the graph vertices in (5). However, in a network the information flow between vertices is over the edges present in the graph. The graph topology influences the dynamical behaviour and therefore plays an important role in connected networks. This gives rise to the idea of investigating the behaviour of the system in terms of the dynamics on the edges. Thereto, the *edge state vector* is defined as $x_e = E^T x$, where E is the incidence matrix, and describes the edge dynamics, meaning that x_e represents for each edge the differences in the states between the vertices that the edge connects. Using the partition of $E = \begin{pmatrix} E_\tau & E_c \end{pmatrix}$, we subsequently define $x_\tau = E_\tau^T x$ and $x_c = E_c^T x$. Then a system is needed that describes our original system (5) in terms of the dynamics on the edges. For that we define a system that uses the edge state vector x_e to describe the differences between vertices, and a system that gives the average behaviour of all edges. This can be done using the following coordinate transformation matrices:

$$T = \begin{pmatrix} E^T \\ \alpha \mu^T \end{pmatrix}, \quad T^{-g} = \begin{pmatrix} M^{-1}E_\tau (E_\tau^T M^{-1} E_\tau)^{-1} (SS^T)^{-1} S & \mathbf{1} \end{pmatrix}. \tag{19}$$

It is noted that T^{-g} is a generalized inverse, meaning that, while $TT^{-g} \neq I$, it does hold that $TT^{-g}T = T$. Now define new coordinates as

$$\begin{pmatrix} x_e \\ x_a \end{pmatrix} = \begin{pmatrix} E^T \\ \alpha\mu^T \end{pmatrix} x.$$

The coordinate $x_a(t) \in \mathbb{R}^n$ gives a weighted average of the states of the individual subsystems, and will therefore be called the *average system*. Its dynamics can be written as

$$\Sigma_a : \begin{cases} \dot{x}_a &= \alpha \mathbf{1}^T G u \\ y_a &= H \mathbf{1} x_a, \end{cases} \tag{20}$$

where it is used that $\alpha\mu^T L_{\text{eff}} = 0$. Next to the average system, we have a system representing the dynamics of the variable x_e, which will be called the *edge system*.

$$\Sigma_e : \begin{cases} \dot{x}_e &= -L_e x_e + E^T M^{-1} G u \\ y_e &= H M^{-1} E_\tau (E_\tau^T M^{-1} E_\tau)^{-1} (SS^T)^{-1} S x_e. \end{cases} \tag{21}$$

Here it is used that $E^T L = L_e E^T$. To ease notation we define
$F = M^{-1}E_\tau(E_\tau^T M^{-1} E_\tau)^{-1}(SS^T)^{-1}S$, so that the output of the edge system can be given by $y = HFx_e$.
One of the uses for the edge dynamics is found in studying synchronisation.

Definition 2. *A system Σ is said to synchronise if, for $u(t) = 0$,*

$$lim_{t\to\infty} x(t) = c\mathbf{1}, \tag{22}$$

for some $c \in \mathbb{R}$ depending on the initial condition $x(0) = x_0$.

This has a natural edge interpretation. If a system synchronises, in the limit all states are equal. Differences between the states then necessarily converge to zero. It is therefore understood that there is a connection between synchronisation of the system Σ and asymptotic stability of the system Σ_e, which is formalised as follows.

Proposition 1. *A system Σ synchronises if and only if all trajectories $x_e(t)$ of the associated edge system Σ_e with $u(t) = 0$ and $x_e(0) \in im\, E^T$ converge to zero.*

4.1 S^T-invariance and asymptotic stability of Σ_e

For an edge system Σ_e representing the edge dynamics of Σ, the edge state vector x_e can not attain every value in $\mathbb{R}^{n_e}$. There is a linear dependence of x_c on x_τ through (9). As such, it is seen that the edge state vector x_e of edge system Σ_e representing the edge dynamics of system Σ initially satisfies $x_e \in \text{im}\, S^T$. The following theorem deals with the general dynamics.

Theorem 1. *Consider a system Σ as in (5) and let Σ_e in (21) be the corresponding edge system. If the network represented by Σ is connected, then for any trajectory x of Σ, $x_e = E^T x$ is a trajectory of Σ_e. Moreover, im S^T is invariant under the dynamics of (21).*

The following Lemma will be used in the proof of Theorem 1.

Lemma 1. *Consider a connected a graph $\mathcal{G}(V, E)$, its corresponding edge Laplacian and a matrix L_τ defined as in (16). Then, for all $t \in \mathbb{R}$,*

$$e^{-L_e t} S^T = S^T e^{-L_\tau t}.$$

Proof. The matrix exponential $e^{-L_e t}$ by definition equals

$$e^{-L_e t} = \sum_{k=0}^{\infty} \frac{1}{k!} (-L_e)^k t^k. \tag{23}$$

Therefore

$$e^{-L_e t} S^T = \sum_{k=0}^{\infty} \frac{1}{k!} (-L_e)^k t^k S^T.$$

We need to prove that $e^{-L_e t} S^T = S^T e^{-L_\tau t}$. Using the definition of the matrix exponential this holds if and only if

$$\sum_{k=0}^{\infty} \frac{1}{k!} (-L_e)^k t^k S^T = S^T \sum_{k=0}^{\infty} \frac{1}{k!} (-L_\tau)^k t^k. \tag{24}$$

This does hold if and only if, for all $k \geq 0$

$$L_e^k S^T = S^T L_\tau^k. \tag{25}$$

Therefore it is sufficient to prove this. We continue with the principle of mathematical induction. For $k = 0$ (25) reduces to $S^T = S^T$, so the statement is true for $k = 0$. Now assume the equation is true for $k = n$, meaning $L_e^n S^T = S^T L_\tau^n$. For $n+1$ we have $L_e^{n+1} S^T = L_e^n L_e S^T$. Since $L_e S^T = S^T L_\tau$ by (18) it is possible to write $L_e^{n+1} S^T = L_e^n S^T L_\tau$. By the induction hypothesis, this can be rewritten as

$$L_e^{n+1} S^T = S^T L_\tau^{n+1}.$$

This completes the proof. □

We are now in a position to prove Theorem 1.

Proof. From the definition of the systems Σ and Σ_e in (5) and (21) respectively, it follows directly that for any trajectory x in Σ, $x_e = E^T x$ is a trajectory in Σ_e. To prove invariance, we note that a general solution for the edge state vector can be given as

$$x_e(t) = e^{-L_e t} x_e(0) + \int_0^t e^{-L_e (t-\tau)} E^T M^{-1} G u(\tau) d\tau. \tag{26}$$

Now take any $x_e(0) \in \operatorname{im} S^T$ and note that $E^T = S^T E_\tau^T$. It can be seen that

$$x_e(t) = e^{-L_e t} S^T x_\tau(0) + \int_0^t e^{-L_e(t-\tau)} S^T E_\tau^T M^{-1} G u(\tau) d\tau.$$

Using Lemma 1, we see that the general solution can as well be written as

$$x_e(t) = S^T e^{-L_\tau} x_\tau(0) + S^T \int_0^t e^{-L_\tau(t-\tau)} E_\tau^T M^{-1} G u(\tau) d\tau. \tag{27}$$

From this equation, we can conclude that for all initial conditions $x_e(0) \in \operatorname{im} S^T$ and for all inputs $u(t)$, also $x_e(t) \in \operatorname{im} S^T$, $\forall t \geq 0$. So indeed $\operatorname{im} S^T$ is invariant under the dynamics of Σ_e. □

The relevance of Theorem 1 lies in the fact that we can use it to restrict the edge state vector $x_e(t)$ of Σ_e to the part that is repesenting the edge dynamics of the system Σ and study it under the restriction $x_e \in \operatorname{im} S^T$. Using this, we can obtain results for Σ as well. As a matter of fact, it can be shown that $x_e(t) = S^T x_\tau(t)$, $\forall t$. Here $x_\tau(t)$ is the edge state vector of the spanning tree of the graph, i.e. it is defined as $x_\tau = E_\tau^T x$ with E_τ as defined in (8).
Using (27) it is easily seen that, for connected graphs, the system Σ synchronises. This is equivalent to convergence to zero for all trajectories of Σ_e for $u(t) = 0$ and with initial conditions $x_e(0) \in \operatorname{im} S^T$ by Proposition 1. If we take $u(t) = 0$, the general solution (27) becomes

$$x_e(t) = S^T e^{-L_\tau t} x_\tau(0).$$

We have shown that L_τ contains only positive eigenvalues for connected graphs. This was seen in (15). Since all eigenvalues of $-L_\tau$ are strictly negative, the limit for $x_e(t)$ exists for all initial conditions satisfying $x_e(0) \in \operatorname{im} S^T$ and equals $\lim_{t\to\infty} x_e(t) = 0$. This is a well known result, any undirected connected graph synchronises [12].
Most results in this thesis use that $x_e \in \operatorname{im} S^T$ in some way. The following interesting result shows that this is equivalent with the condition $x_e \in \operatorname{im} E^T$.

Lemma 2. *Consider the matrices S and E as defined in* (10)*. Then $\operatorname{im} S^T = \operatorname{im} E^T$.*

Proof. From the definition $E^T = S^T E_\tau^T$ from (10), it is clear that $\operatorname{im} E^T \subset \operatorname{im} S^T$. To show the reverse inclusion $\operatorname{im} S^T \subset \operatorname{im} E^T$ we need to prove that for all $x_e \in \operatorname{im} S^T$, there exists an x such that $x_e = E^T x$. It is known that there exists an x_τ such that $x_e = S^T x_\tau$. Therefore, if we are able to choose $x_\tau = E_\tau^T x$, it necessarily holds that $x_e \in \operatorname{im} E^T$. So the question is whether there exists a solution to $x_\tau = E_\tau x$ for all $x_\tau \in \mathbb{R}^{n_\tau}$.
Consider a matrix with full row rank, such as E_τ^T. A solution x can be found as $x = E_\tau (E_\tau^T E_\tau)^{-1} x_\tau$. Here the inverse exists because E_τ^T has full row rank. Then, $\operatorname{im} S^T \subset \operatorname{im} E^T$ as well, and we conclude that $\operatorname{im} S^T = \operatorname{im} E^T$. □

5 Edge Selection

So far we have identified an edge system, and shown some useful properties of it. However, our goal is to cluster vertices in order to achieve a lower dimensional model. The main purpose of this chapter is to find two vertices that are best to cluster. This is done using the dynamics on the edges, defined in (21). To be specific, we will use the degree of controllability and observability of each edge as a measure of approximately identical behaviour. Using the degree of controllability and observability is a well-tested way to determine which vertices are closely related. If the degree of controllability of an edge between two vertices is small, this implies it is difficult to steer the vertices away from each other. Therefore, when given the same control input, these vertices will show similar behaviour. Then, from a control viewpoint, these vertices are good candidates to be approximated by a single vertex. In other words, clustering these vertices is expected to be a better approximation than the clustering of any other pair of vertices (connected via an edge). From an observability view, a similar reasoning applies. If the degree of observability is low, this implies that it is harder to identify differences between the edges corresponding to the edge, than it is with other pairs of vertices. So in this case it would be beneficial to cluster these two vertices. In this section we will investigate a way to determine what the least important edge is in terms of controllability and observability.

5.1 Observability Gramian

For $u(t) = 0$, the output of the trajectories in Σ_e is given by

$$y_e(t, x_e(0)) = HFe^{-L_e t}x_e(0).$$

We define the *output energy* of the edge system Σ_e as

$$L_o(x_e(0)) = \int_0^\infty \|y_e(t, x_e(0))\|^2 \, dt. \tag{28}$$

This is the energy for zero input that is in the output for a given initial condition $x_e(0)$. The initial condition will also be denoted as $x_{e,0}$. It is noted that we only consider edge states $x_e(0)$ such that $x_e(0) \in \text{im } S^T$. This is done because we want $x_e(t)$ to represent the edge dynamics, the difference between vertices of Σ over the edges. Therefore, $x_e(t)$ cannot be any arbitrary vector in the entire space $\mathbb{R}^{n_e}$, but must be compatible with the state vector $x(t)$ of system Σ. It means we must restrict the edge state to $x_e(0) \in \text{im } S^T$.
We can rewrite the output energy for an arbitrary initial condition as

$$L_0(x_{e,0}) = \int_0^\infty x_{e,0}^T e^{-L_e^T t} F^T H^T H F e^{-L_e t} x_{e,0} \, dt. \tag{29}$$

We would like to define a matrix P_o by

$$P_o = \int_0^\infty e^{-L_e^T t} F^T H^T H F e^{-L_e t} \, dt. \tag{30}$$

However, this integral not necessarily converges. If it does, the output energy for arbitrary initial conditions can be written as

$$L_0(x_{e,0}) = x_{e,0}^T P_o x_{e,0}.$$

This matrix P_o is known as the observability Gramian. The observability Gramian can give information about the observability of each different edge [6]. This is used, for example, in the method of balanced truncation. In this thesis, we use the same idea of Gramians, but the way they are used differs from the method of balanced truncation.
As can be seen from (27), the output of the edge system Σ_e with zero input and $x_e(0) \in \text{im } S^T$ can also be written as

$$y_e(t, x_\tau(0)) = HFS^T e^{-L_\tau t} x_\tau(0),$$

where $x_\tau(0) = E_\tau^T x(0)$. For such initial conditions, the output energy is given by

$$L_0(x_{\tau,0}) = x_{\tau,0}^T \int_0^\infty e^{-L_\tau^T t} SF^T H^T HFS^T e^{-L_\tau t}\, dt\, x_{\tau,0}.$$

We define a matrix W as

$$W = \int_0^\infty e^{-L_\tau^T t} SF^T H^T HFS^T e^{-L_\tau t}\, dt.$$

Because the matrix $-L_\tau$ is Hurwitz, this integral converges. It follows that we can write the output energy as

$$L_0(x_{\tau,0}) = x_{\tau,0}^T W x_{\tau,0}.$$

To indicate the linear dependence between x_τ and x_e, we would like to write the matrix W as $W = SPS^T$ for some matrix P. It can be seen that this is always possible. Recall that $S = \begin{pmatrix} I_{n-1} & T \end{pmatrix}$. For example, if we define $P = \begin{pmatrix} W & \\ & 0_{n_e-n+1} \end{pmatrix}$, then we see that

$$SPS^T = \begin{pmatrix} I_{n-1} & T \end{pmatrix} \begin{pmatrix} W & \\ & 0_{n_e-n+1} \end{pmatrix} \begin{pmatrix} I_{n-1} \\ T^T \end{pmatrix} = W.$$

It is clear that we can define SPS^T as

$$SPS^T = \int_0^\infty e^{-L_\tau^T t} SF^T H^T HFS^T e^{-L_\tau t}\, dt. \tag{31}$$

The matrix SPS^T can be viewed as the observability Gramian P_o, restricted to the relevant subspace where $x_e \in \text{im } S^T$. We will therefore call the matrix SPS^T the *restricted observability Gramian*.
The restricted observability Gramian in (31) can be found using a Lyapunov equation as follows:

Proposition 2. *Consider the edge system Σ_e as in (21) such that the corresponding system Σ represents a connected network. If $-L_\tau$ is Hurwitz, the restricted observability Gramian SPS^T is the unique, symmetric positive semi-definite solution of the Lyapunov equation*

$$L_\tau^T SPS^T + SPS^T L_\tau - SF^T H^T HFS^T = 0. \tag{32}$$

Proof. Substitution of the definition for SPS^T as given in (31) into the Lyapunov equation (32) shows that

$$\begin{aligned} L_\tau^T SPS^T + SPS^T L_\tau = -\int_0^\infty (&-L_\tau^T e^{-L_\tau^T t} SF^T H^T HFS^T e^{-L_\tau t} - \\ &e^{-L_\tau^T t} SF^T H^T HFS^T e^{-L_\tau t} L_\tau)\, dt \end{aligned}$$

This is seen to be equal to

$$\begin{aligned} L_\tau^T SPS^T + SPS^T L_\tau &= -\int_0^\infty \frac{d}{dt}(\, e^{-L_\tau^T t} SF^T H^T HFS^T e^{-L_\tau t})\, dt \\ &= -\, e^{-L_\tau^T t} SF^T H^T HFS^T e^{-L_\tau t}|_0^\infty \\ &= SF^T H^T HFS^T. \end{aligned} \tag{33}$$

The last equality holds due to $-L_\tau$ being Hurwitz. For a proof of uniqueness, we refer to [6]. □

The importance of SPS^T stems from the fact that the matrix P informs about observability of each edge in our network separately, while existence of the matrix SPS^T is guaranteed by the fact that $-L_\tau$ is Hurwitz. The matrix $-L_e$ is not Hurwitz in general, which means that in this framework observability of each edge can not be ensured to be found. Hence, we use the restricted Gramian SPS^T.

5.2 Generalized Observability Gramian

In the previous section, we found a way to obtain a restricted observability Gramian SPS^T that is expected to be a measure of the level of observability of separate edges. It is however not clear how the level of observability follows from this matrix. Typically, an eigenvalue decomposition of the observability Gramian is used. However, in this thesis a different approach is taken. We try to find a matrix $\tilde{P}$ such that $S\tilde{P}S^T$ serves as an upper bound on the restricted observability Gramian SPS^T, but with a diagonal structure of $\tilde{P}$ that severely eases interpretation of the level of observability of each edge. If a diagonal upper bound can be used, its diagonal entries serve as an upper bound to the level of observability of individual edges. To find such a matrix, we first introduce a *generalized observability Gramian*

Definition 3. *A symmetrix positive semi-definite matrix $\tilde{P}$ is called a generalized observability Gramian if it satisfies the Lyapunov inequality*

$$L_\tau^T S\tilde{P}S^T + S\tilde{P}S^T L_\tau - SF^T H^T HFS^T \geq 0 \tag{34}$$

Solutions to this inequality are not unique. We hope to be able to find a solution that is easy to interpret and can serve as an upper bound to the level of observability of each edge. The level of observability of edges is closely related to the output energy [1]. To be able to use the generalized observability Gramian, we need to show that the matrix $\tilde{P}$ indeed can serve as an upper bound to the ouput energy of each edge. This can be stated in the following theorem.

Theorem 2. *Let a generalized observability Gramian $\tilde{P}$ be defined as a solution to* (34)*. Then $V(x_{e,0}) = x_{e,0}^T \tilde{P} x_{e,0}$ gives an upper bound on the output energy for $x_{e,0} \in im\, S^T$.*

Proof. Assume a solution $\tilde{P}$ to (34) is found. Then an energy function can be defined as $V(x_e(t)) = x_e(t)^T \tilde{P} x_e(t)$. The time derivative of this energy function can be written as

$$\begin{aligned} \dot{V} &= x_\tau^T(-L_\tau^T S\tilde{P}S^T - S\tilde{P}S^T L_\tau)x_\tau \\ &\leq -x_\tau S F^T H^T H F S^T x_\tau = -\|y_e(t)\|^2 \end{aligned} \tag{35}$$

The time dependence is dropped for ease of notation. The inequality holds since $\tilde{P}$ is a solution to (34). If $\dot{V}(x_e) \leq -\|y(t)\|^2$ for all $x_e \in \text{im}\, S^T$, then

$$\begin{aligned} &\int_0^T \dot{V}(x_e(t))\, dt \leq -\int_0^T \|y(t)\|^2\, dt \\ &V(x_e(T)) - V(x_e(0)) \leq -\int_0^T \|y(t)\|^2\, dt \\ &V(x_e(0)) \geq \int_0^T \|y(t)\|^2\, dt + V(x_e(T)), \quad \forall T. \end{aligned}$$

The limit $\lim_{T\to\infty} V(x_e(T)) = 0$ holds for connected graphs by Proposition 1. Then the equation becomes

$$V(x_e(0)) \geq \int_0^\infty \|y(t)\|^2\, dt = L_0(t, x_{e,0}), \quad \forall x_e(0) \in \text{im}\, S^T. \tag{36}$$

Hence the energy function $V(x_e(0)) = x_e(0)^T \tilde{P} x_e(0)$ is indeed an upper bound on the output observability energy. Therefore, the matrix $\tilde{P}$ can be used as an upper bound to the level of observability of edges. □

5.3 Existence of diagonal solutions

Although there might be an infinite number of solutions to the Lyapunov inequality (34), a diagonal solution is generally not guaranteed to exist. However, in the class of undirected weighted graphs that is considered here, existence of a diagonal solution can be proven.

Lemma 3. *Consider the Lyapunov inequality* (34)*. Then, for L_τ as defined in* (16)*, a diagonal solution $\tilde{P}$ does exist.*

Proof. As a candidate solution, the scaled positive diagonal weight matrix $\epsilon^{-1}R$ is proposed. We note that for this choice of solution, the inequality reduces to

$$2\epsilon^{-1}SRS^TE_\tau^TM^{-1}E_\tau SRS^T \geq F^TH^THF.$$

For any positive definite matrix M^{-1} it is true that $(E_\tau SRS^T)^TM^{-1}(E_\tau SRS^T)$ is positive definite, if the matrix $E_\tau SRS^T$ has full column rank [9]. The matrix $E_\tau SRS^T$ is of full rank. This can among other ways be seen from the fact that L_τ is of full rank. Since also the matrix M^{-1} is positive definite, it is seen that

$$2\epsilon^{-1}SRS^TE_\tau^TM^{-1}E_\tau SRS^T > 0.$$

Then, for ϵ sufficiently small, it is true that

$$L_\tau^TS\tilde{P}S^T + S\tilde{P}S^TL_\tau = 2SRS^TE_\tau^TM^{-1}E_\tau SRS^T > \epsilon F^TH^THF.$$

Hence, $\tilde{P} = \epsilon^{-1}R$ is a diagonal solution to (34) for ϵ sufficiently small. □

The diagonal solution found in the lemma is not unique. Finding an, in some sense, optimal diagonal solution is not trivial. Intuitively, it seems clear that a tight upper bound on the observability Gramian will give a better approximation. Without an error analysis, this intuition can not be made mathematically solid. A good heuristic method to find a tight upper bound is to minimize the trace of the solution $\tilde{P}$.

5.4 Controllability Gramian

For the controllability case, a similar approach can be taken. As in the case of observability, there is a direct relation between a certain energy and the level of controllability of each edge. The *control energy* of the edge system Σ_e can be defined as [1]

$$L_c(x_{e,0}) = \min \int_{-\infty}^{0} \|u(t)\|^2 \, dt, \tag{37}$$

where again only $x_{e,0} \in \text{im } S^T$ are considered. This is the minimal energy for the system Σ_e that is needed to steer the system from initial condition $\lim_{t\to-\infty} x_e(t) = 0$, which will be written shortly as $x_e(-\infty) = 0$, to $x_e(0) = x_{e,0}$. In trying to find a minimal control $u(t)$ for the functional $J(x_{e,0}, u) = \int_{-\infty}^{0}\|u(t)\|^2 \, dt$, we recognize the problem setting of linear quadratic optimal control. To be able to use this to its full extent, we first need to rewrite our system Σ_e.

Lemma 4. *Consider the edge system Σ_e, together with the control energy $L_c(x_{e,0}) = min \int_{-\infty}^{0}\|u(t)\|^2 \, dt$. Finding the minimal energy $L_c(x_{e,0})$ and the optimal control $u(t)$ is equivalent to finding an optimal control $\nu(t)$ that minimizes the cost functional*

$$J(\nu) = \int_0^\infty ||\nu(\tau)||^2 \, d\tau \tag{38}$$

for the system

$$\Sigma_e^* : \begin{cases} \dot{\xi}_e(\tau) = L_e\xi_e(\tau) - E^T M^{-1} G\nu(\tau) \\ y_e^* = HF\xi_e(\tau), \end{cases} \tag{39}$$

with initial condition $\xi_e(0) = x_0$ *and a stability requirement* $\xi_e(\infty) = 0$*. The control energy* $L_c(x_{e,0})$ *is equal to the minimal cost function* $J^*(\nu) := min\, J(\nu)$*.*

Proof. Define a new coordinate $\tau = -t$, such that we reverse time. Also define a new input variable $\nu(\tau)$ as $\nu(\tau) = u(-\tau)$, and similarly $\xi_e(\tau) = x_e(-\tau)$. Then the to be minimized energy functional can be written as

$$J(\nu) = -\int_{\infty}^{0} \|\nu(\tau)\|^2 \, d\tau.$$

Reversing the bounds on the integral then returns (38). This shows equivalence between the control energy and the minimal cost. The system Σ_e^* follows from the definition of Σ_e in (21) and the fact that $\dfrac{d}{dt} = -\dfrac{d}{d\tau}$. □

It is noted that $\xi_e(\infty) = 0$ can only be satisfied if the system Σ_e^* is stabilisable. As L_e contains only eigenvalues with non-negative real part, this implies that the system Σ_e^* must be controllable. However, we only require stabilisability for all $\xi_e \in \text{im } S^T$. We note that the linear dependence of the edges is not explicit in Σ_e^*. This can be made explicit by using the equality $\xi_e = S^T\xi_\tau$ for some ξ_τ that represents the dynamics in the chosen spanning tree. Using this, a new system Σ_τ can be defined as follows

$$\Sigma_\tau^* : \begin{cases} \dot{\xi}_\tau = L_\tau\xi_\tau - E_\tau^T M^{-1} Gu, \\ y_\tau^* = HFS^T\xi_\tau. \end{cases} \tag{40}$$

If there exists a control function u for every initial condition $\xi_\tau(0) = \xi_{\tau,0}$ such that $\lim_{\tau\to\infty}\xi_\tau(t) = 0$, then also $\lim_{\tau\to\infty}\xi_e(\tau) = S^T\xi_\tau(\tau) = 0$ using the same control. Therefore we make the following assumption.

Assumption 1. *The system described by* $\left(L_\tau, E_\tau^T M^{-1} G\right)$ *is stabilisable.*

Next we use a lemma from [16].

Lemma 5. *Consider the system* $\dot{x}(t) = Ax(t) + Bu(t)$ *together with the cost function*

$$J(x_0, u) := \int_0^\infty x(t)^T W x(t) + u(t)^T u(t)\, dt,$$

with $W \geq 0$*. Factorize* $W = C^T C$*. Then the following statements are equivalent:*

1. *For every* $x_0 \in \mathcal{X}$ *there exists* $u \in \mathcal{U}$ *such that* $J(x_0, u) < \infty$*.*

2. *The algebraic Riccati equation*

$$A^T P + PA - PBB^T P + W = 0$$

has a real symmetric positive semidefinite solution P.

3. *The system*

$$\Sigma = \left(A, \quad B, \quad \begin{pmatrix} C \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ I \end{pmatrix} \right)$$

is output stabilisable.

4. $\langle ker\, C \mid A \rangle + \mathcal{X}_{stab} = \mathcal{X}$.

Assume that one of the above conditions holds. Then there exists a smallest real symmetric positive semidefinite solution of the algebraic Riccati equation, i.e. there exists a real symmetric solution $P^- \geq 0$ such that for every real symmetric solution $P \geq 0$ we have $P^- \leq P$. For every x_0 we have

$$J^*(x_0) := inf\{J(x_0, u) \mid u \in \mathcal{U}\} = x_0^T P^- x_0.$$

Furthermore, for every x_0, there exists exactly one optimal input function, i.e. a function $u^* \in \mathcal{U}$ such that $J(x_0, u^*) = J^*(x_0)$. This optimal input is generated by the time-invariant feedback law

$$u(t) = -B^T P^- x(t).$$

This lemma we apply to our special situation where $W = 0$. This lemma can be applied because it can be seen that Assumption 1 implies that the conditions in the lemma are satisfied.
By applying Lemma 5 to Σ_τ, it follows that the minimal cost is given by

$$J^*(\xi_{\tau,0}) = \xi_{\tau,0}^T Q_\tau \xi_{\tau,0},$$

where the matrix Q_τ satisfies the equation

$$L_e^T Q_\tau + Q_\tau L_e - Q_\tau E^T M^{-1} G G^T M^{-1} E Q_\tau = 0, \tag{41}$$

for all $\xi_\tau \in \mathbb{R}^{n-1}$. This solution Q_τ can be written as $Q_\tau = SQS^T$, such that it compares with the solution to the Lyapunov inequality (32). This can be seen as follows. The solution Q_τ is symmetric positive semi-definite and can therefore be written as $Q_\tau = \Delta\Delta^T$, for some matrix Δ. We try to find a solution Q that is also symmetric positive semi-definite, such that $Q = DD^T$ for some matrix D. So the question is whether we can find a matrix D in such a way that $\Delta\Delta^T = SDD^T S^T$. Or equivalently, can we find a matrix D such that $\Delta = SD$. If we take a column vector δ_i of Δ, then $\delta_i \in \mathbb{R}^{n-1}$. Since S is defined as $S = \begin{pmatrix} I & T \end{pmatrix}$ in (10), it can be seen that $\mathbb{R}^{n-1} \subset \text{im}\, S$. Therefore, for every column vector δ_i of Δ, there exists a vector d_i such that $\delta_i = Sd_i$. Apply this for every column vector in Δ and the result follows that $Q_\tau = SQS^T$, for a matrix Q.

Using this, the Riccati inequality (41) can be rewritten as

$$L_\tau^T SQS^T + SQS^T L_\tau + SQS^T E_\tau^T M^{-1} G G^T M^{-1} E_\tau SQS^T = 0, \qquad (42)$$

for all $\xi_\tau \in \mathbb{R}^{n-1}$. To this Riccati equation, the minimal cost $J^*(\xi_{\tau,0})$ is given as

$$J^*(\xi_{\tau,0}) = \xi_{\tau,0}^T SQS^T \xi_{\tau,0},$$

Now, if a symmetric positive semi-definite solution SQS^T can be found that also stabilises Σ_e^*, then by Lemma 4 the control energy is given as $L_c(x_\tau) = x_{\tau,0}^T SQS^T x_{\tau,0}$. We claim that the inverse of the controllability Gramian is the unique stabilising solution to the algebraic Riccati equation, under the assumption of stabilisability [16]. The controllability Gramian in this case is defined as

$$P_c = \int_0^\infty e^{-L_\tau t} S^T E_\tau^T M^{-1} G G^T M^{-1} E_\tau S e^{-L_\tau^T t}\, dt,$$

and forms a dual of the observability Gramian in (30). This is the reason the controllabilty Gramian is commonly used to give measures on the controllability of each node. In the case forehand, it would mean identifying the controllability Gramian with the inverse of SQS^T, i.e. $P_c = (SQS^T)^{-1}$. With the use of the inverse controllability Gramian, the Riccati equation (42) could be written as the Lyapunov equation $P_c L_\tau^T + L_\tau P_c + E_\tau^T M^{-1} G G^T M^{-1} E_\tau = 0$. In the case of tree graphs with $M = I$, from here results in [3] follow. However, because the inverse P_c is not as easily partitioned as SQS^T itself, we continue using the Riccati equation instead of the Lyapunov equation.

5.5 Lower bound on the controllability energy

We have shown that the minimal energy needed to steer our original edge system from $x_e(-\infty) = 0$ to $x_e(0) = x_{e,0}$ equals

$$L_c(x_{\tau,0}) = x_{\tau,0}^T SQS^T x_{\tau,0}\,, \qquad (43)$$

where SQS^T is a solution to (42) and $x_{\tau,0}$ is such that $x_e(0) = S^T x_{\tau,0}$. For the clustering method used in this thesis, we propose as a candidate the edge that has the highest minimal energy to steer the nodes it connects apart. A solution Q to (42) is in general not easily interpreted, just as in the observability case. Therefore, a lower bound on the energy is needed that is not as difficult to interpret. This can be achieved with a Riccati inequality.

Proposition 3. *Let a generalized controllability matrix $\tilde{Q}$ be defined as a solution to the following Riccati inequality:*

$$L_\tau^T S\tilde{Q}S^T + S\tilde{Q}S^T L_\tau - S\tilde{Q}S^T E_\tau^T M^{-1} G G^T M^{-1} E_\tau S\tilde{Q}S^T \geq 0. \qquad (44)$$

Then $V(x_{\tau,0}) = x_{\tau,0}^T S\tilde{Q}S^T x_{\tau,0}$ gives a lower bound on the control energy $L_c(x_{\tau,0})$ in (43).

Proof. This proof is based upon a result in [18]. Suppose inequality (44) holds, then for some symmetric positive semi-definite matrix K it is true that

$$L_\tau^T S\tilde{Q}S^T + S\tilde{Q}S^T L_\tau - S\tilde{Q}S^T E_\tau^T M^{-1} G G^T M^{-1} E_\tau S\tilde{Q}S^T = K.$$

Substracting this equation from (42), results in a new equation. To ease notation, we define the matrices $G_e = E_\tau^T M^{-1} G$ and $\Delta Q_S = S(Q - \tilde{Q})S^T$. Then this equation can be written as

$$(L_\tau - G_e G_e^T SQS^T)^T \Delta Q_S + \Delta Q_S (L_\tau - G_e G_e^T SQS^T) = \\ - \Delta Q_S G_e G_e^T \Delta Q_S - K. \quad (45)$$

As we assume the solution SQS^T is the solution used in the unique stabilising control for system Σ_τ, we see that $L_\tau - G_e G_e^T SQS^T$ is a stability matrix. Then a solution to (45) is given by

$$S(Q - \tilde{Q})S^T = \int_0^\infty e^{(L_\tau - G_e G_e^T SQS^T)^T t} \times \\ (S(Q - \tilde{Q})S^T G_e G_e^T S(Q - \tilde{Q})S^T + K) e^{(L_\tau - G_e G_e^T SQS^T)t}\, dt, \quad (46)$$

where it used that $\Delta Q_S = S(Q - \tilde{Q})S^T$.
Since both $K \geq 0$ and $S(Q - \tilde{Q})S^T G_e G_e^T S(Q - \tilde{Q})S^T \geq 0$ it follows that the solution $S(Q - \tilde{Q})S^T \geq 0$. Hence, $S\tilde{Q}S^T \leq SQS^T, \forall x_\tau \in \mathbb{R}^{n-1}$, including the initial condition $x_\tau(0) = x_{\tau,0}$.
This proves that a solution to inequality (44) provides a lower bound on the control energy, $V(x_{\tau,0}) = x_{\tau,0}^T S\tilde{Q}S^T x_{\tau,0} \leq x_{\tau,0}^T SQS^T x_{\tau,0} = L_c(x_{\tau,0})$. □

A diagonal matrix $\tilde{Q}$ that solves (44) can be used as an approximation for the control energy of each edge. We will call such a diagonal solution to (44) a *control energy bound*. In the same manner, a generalized observability Gramian will also be called an *output energy bound*. As in the observability case, a diagonal solution to the inequality can be shown to exist.

Proposition 4. *A diagonal solution $\tilde{Q}$ to inequality* (44) *does exist.*

Proof. Consider $\tilde{Q} = \alpha R$. Then, given the definition of L_τ, the inequality can be rewritten as

$$SRS^T E_\tau^T (2\alpha M^{-1} - \alpha^2 M^{-1} G G^T M^{-1}) E_\tau SRS^T \geq 0.$$

This is true if $2\alpha M^{-1} - \alpha^2 M^{-1} G G^T M^{-1} > 0$. The diagonal entries of this matrix have a leading term that decreases linearly with decreasing α, whereas the off-diagonal entries decrease quadratically. For α sufficiently small the matrix becomes strictly diagonally dominant. This implies that the matrix has positive eigenvalues and therefore is positive definite. Then $\tilde{Q} = \alpha R$ is indeed a diagonal, positive semi-definite solution to the Riccati inequality, for α sufficiently small. □

6 Model reduction

So far we have developed a method that provides a way of determining the level of controllability and observability of each edge. The least important edge in the network ideally has a low output energy and a high control energy. In the absence of an error analysis, it is not clear how to combine the information on the level of controllability and observability in order to identify the least important edge. Therefore, the idea of [3] is taken. Suppose that we have found the matrices $\tilde{P}$ and $\tilde{Q}$ that give bounds on the energies. Then we compare by taking the product $\tilde{P}\tilde{Q}^{-1}$. The edge corresponding to the lowest value of the product $\tilde{P}\tilde{Q}^{-1}$ is expected to be the least contributing edge. This is a result of the intuition that an edge that has a low output energy and a high control energy give the best pair of vertices to cluster. Assume now that we have found a least contributing edge in the system. By a suitable permutation of the vertex numbering, we can make sure that the least important edges are given by i and j, with $i = n-1$ and $j = n$. Consider the projection $\Gamma = VV^T$, where $V \in \mathbb{R}^{n\times(n-1)}$ is defined as

$$V = \begin{pmatrix} I & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}. \tag{47}$$

This leads to a projected system of the form

$$\hat{\Sigma}_{n-1} : \begin{cases} V^T M V \dot{\xi} & = -V^T L V \xi + V^T G u \\ \hat{y} & = HV\xi. \end{cases} \tag{48}$$

The state vector x is approximated by $x \approx V\xi$, where $\xi \in \mathbb{R}^{n-1}$ is of lower order. We remark that this ξ has no relation with the ξ_e that was defined before. To be able to do this one-step clustering method repeatedly, it is necessary to check whether the reduced order system also represents a multi-agent network on a undirected weighted graph. So it is needed that this reduced order model is brought to the form of the original system Σ in (5). To that end, we first of all note that V^TMV has the same diagonal structure as M. The main difference is the addition in V^TMV of the diagonal elements of M corresponding to the clustered vertices. Therefore, we define a new inertia matrix of the reduced system,

$$\hat{M} = V^T M V.$$

The new graph Laplacian can be written as

$$\hat{L} = V^T E R E^T V, \tag{49}$$

which also represents a weighted undirected graph if V^TE is indeed a valid incidence matrix. To study the properties of V^TE, we assume that the edge connecting the clustered vertices i and j, is given by the $l-$th column of the incidence matrix E, where $l = n_e$ represents the last column. This can always be achieved by renumbering the edges. In this case, we see that the original

incidence matrix can be partitioned as

$$E = \begin{pmatrix} E_{00} & 0 \\ E_{i0} & E_{il} \\ E_{j0} & E_{jl} \end{pmatrix}. \tag{50}$$

The zero entry in the incidence matrix comes from the fact that the $l-$th column represents the edge connecting the to be clustered vertices i and j. Also note that $E_{il}+E_{jl} = 0$, because of the strucure of the incidence matrix. Computation of the product $V^T E$ then results in

$$V^T E = \begin{pmatrix} E_{00} & 0 \\ E_{i0} + E_{j0} & 0 \end{pmatrix}.$$

Due to the zero column in this matrix $V^T E$ is not a valid incidence matrix. However, if we now define the incidence matrix of the reduced order model as

$$\hat{E} = \begin{pmatrix} E_{00} \\ E_{i0} + E_{j0} \end{pmatrix}, \tag{51}$$

then the reduced order graph Laplacian (49) is also given by

$$\hat{L} = \hat{E}\hat{R}\hat{E}^T,$$

where $\hat{R}$ is defined such that

$$R = \begin{pmatrix} \hat{R} & 0 \\ 0 & r_{ll} \end{pmatrix}.$$

Here r_{ll} is the weigth of the edge that is deleted in the clustering process. It remains to be checked whether $\hat{E}$ indeed is an incidence matrix. It is true that $\hat{E}$ has the right dimensions, $\hat{E} \in \mathbb{R}^{(n-1)\times(n_e-1)}$. Also, $\hat{E}_{ij} \in \{-1, 0, 1\}$, $\forall i$, and every column also contains only one pair of ± 1. This can be seen from the partition in (50) and the definition of the incidence matrix (7). Because $\hat{E}$ has the same structure as E and has the right dimensions, we conclude that it is an incidence matrix for the reduced order model. Note that also for the reduced order incidence matrix it is possible that $\hat{n}_e \geq \hat{n}$, meaning that there are typically more edges than vertices in the reduced order model. Therefore, also the incidence matrix $\hat{E}$ has linearly dependent columns. This implies that

$$\hat{E} = \hat{E}_\tau \hat{S}, \tag{52}$$

as in the non-reduced case. Here $\hat{S}$ is defined similarly to (9).
We remark that it is possible that the resulting incidence matrix is representing more than one edge between two vertices. An example is given by the following graph when vertices 3 and 4 are clustered.

This can be overcome by representing the original two edges as one edge with a weigth equal to the sum of the weights of the edges it is substituted for. This needs to be checked at every reduction step.

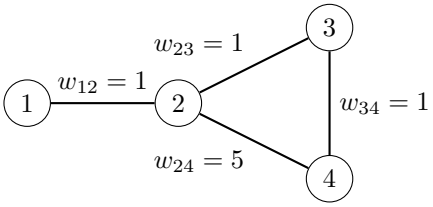

Figure 1: Graph that loses a cycle by clustering

Since determining what vertices are best to cluster is done using an edge system, a reduced order edge system must also be defined. Therefore, a transformation matrix $\hat{T}$ is defined such that

$$\hat{T} = \begin{pmatrix} \hat{E} \\ \hat{\mu} \end{pmatrix}.$$

This transformation matrix is of the same form as (19). For this matrix, the vector $\hat{\mu}^T$ is the normalized left eigenvector of the reduced order effective graph Laplacian, i.e. $\hat{\mu}^T \hat{M}^{-1} \hat{L} = 0$ and $\hat{\mu}^T \mathbf{1} = 1$. By inspection it is seen that $\hat{\mu} = \gamma \hat{M} \mathbf{1}$ for some $\gamma \in \mathbb{R}$. The matrix $\hat{E}$ is the incidence matrix as defined in (51). Then the reduced order average system is given by

$$\hat{\Sigma}_a : \begin{cases} \dot{\xi}_a &= \mathbf{1}^T V^T G u \\ \hat{y}_a &= HV\mathbf{1}\xi_a. \end{cases} \tag{53}$$

The reduced order edge system is defined as

$$\hat{\Sigma}_e : \begin{cases} \dot{\xi}_e &= -\hat{L}_e \xi_e + \hat{E}^T \hat{M}^{-1} V^T G u \\ \hat{y}_e &= HV\hat{M}^{-1}\hat{E}_\tau (\hat{E}_\tau^T \hat{M}^{-1} \hat{E}_\tau)^{-1} (\hat{S}\hat{S}^T)^{-1} \hat{S} \xi_e. \end{cases} \tag{54}$$

Here it is used that $\hat{L}_e$ is defined as $\hat{L}_e = \hat{E}^T \hat{M}^{-1} \hat{E} \hat{R}$.

6.1 Preservation of synchronisation

Next to knowing how to do a clustering based model reduction, we would also like it to have several properties. In the problem statement it was explicitly formulated that the clustered model should preserve synchronisation. It turns out that this is the case. It is known that a connected graph synchronises. Intuitively, clustering of vertices does not change connectedness of a network. Therefore, the reduced order system should synchronise as well. This intuition is made precise in the following theorem.

Theorem 3. *Consider the system Σ as in (5). Let $\hat{\Sigma}_{n-1}$ be the one step reduced order system. If the network the system Σ represents is a connected network, then the clustered system $\hat{\Sigma}_{n-1}$ achieves synchronisation as well.*

Proof. The system $\hat{\Sigma}_{n-1}$ synchronises if $\mathcal{N}(\hat{E}^T) = \text{span}\{\mathbf{1}\}$ [12]. It is clear from the definition of $\hat{E}$ in (51) that $\text{span}\{\mathbf{1}\} \subset \mathcal{N}(\hat{E}^T)$. If this is the only

vector in the kernel of $\hat{E}^T$, then we have proven synchronisation of the reduced order model. It is assumed that the system Σ synchronises, implying that $\mathcal{N}(E^T) = \text{span}\{\mathbf{1}\}$.
The proof follows from a contradiction argument. Consider a vector $\begin{pmatrix}\alpha^T & \beta\end{pmatrix}^T \neq c\mathbf{1}^T$ in the kernel of $\hat{E}^T$, i.e. a vector such that

$$\begin{pmatrix}E_{00}^T & E_{i0}^T + E_{j0}^T\end{pmatrix}\begin{pmatrix}\alpha \\ \beta\end{pmatrix} = 0.$$

Or equivalently

$$E_{00}^T\alpha + E_{i0}^T\beta + E_{j0}^T\beta = 0.$$

Then, using the partition in (50), it is seen that the vector $\begin{pmatrix}\alpha^T & \beta & \beta\end{pmatrix}^T$ is in the kernel of the incidence matrix E.

$$\begin{pmatrix}E_{00}^T & E_{i0}^T & E_{j0}^T \\ 0 & E_{il}^T & E_{jl}^T\end{pmatrix}\begin{pmatrix}\alpha \\ \beta \\ \beta\end{pmatrix} = 0,$$

However, we assumed that the system Σ synchronises, which implies the only vector in the kernel $\mathcal{N}(E^T) = \text{span}\{\mathbf{1}\}$. Therefore, it must be true that $\begin{pmatrix}\alpha^T & \beta\end{pmatrix}^T = c\mathbf{1}^T$, which is a contradiction. Therefore, we see that the only vector in the kernel of $\hat{E}^T$ is indeed the scaled vector of all ones, $\mathcal{N}(\hat{E}^T) = \text{span}\{\mathbf{1}\}$. This proves synchronisation of the system $\hat{\Sigma}$. □

Using these results, an algorithm for model reduction can be given. After identifying the least important edge using the energy bounds provided by the Riccati and Lyapunov inequalities, the one-step clustering is done by using a projection $\Gamma = VV^T$. Theorem 3 allows us to conclude that the clustering method preserves the class of systems considered and achieves synchronisation as well. The one step clustering method can be repeated indefinitely, until the desired order is achieved. However, at every step it must be checked if there is no more than one edge connecting two vertices. If there are two edges connecting the same vertices, these must be replaced with one edge where the weight equals the sum of the edges it replaces.

7 Illustrative example

As a first example of the clustering algorithm, we consider a network containing a square, where the vertices and edges all have the same weight, equal to one. So, $M = I_5$ and $R = I_5$ as well. The edge and vertex numbering is given in Figure 2.

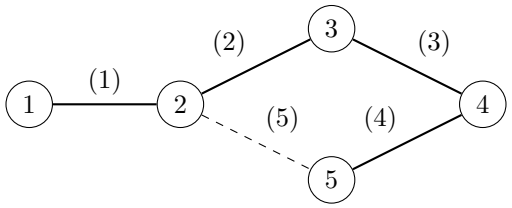

Figure 2: Graph containing a cycle

Here the dashed edge denotes the edge that is not in the spanning tree. It can readily be checked that the incidence matrix can be given by

$$E = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -1 & -1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{pmatrix}.$$

The spanning tree is given as $\mathcal{G}(\mathcal{V}_\tau, \mathcal{E}_\tau)$, where $\mathcal{V}_\tau = \{1, 2, 3, 4, 5\}$ and $\mathcal{E}_\tau = \{\{1, 2\}, \{2, 3\}, \{3, 4\}, \{4, 5\}\}$. The output and input is put on vertex 1, such that the output and control matrices G, H are given by

$$H = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad G = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \end{pmatrix}^T.$$

Using these matrices, the Lyapunov (34) and Riccati (44) inequalities can be solved. Minimising the trace of the possible solutions gives the following solutions that are expected to be the best approximations:

$$P = \operatorname{diag}\begin{pmatrix} 0.67 & 0.40 & 0.17 & 0.17 & 0.40 \end{pmatrix}, \tag{55}$$

$$Q = \operatorname{diag}\begin{pmatrix} 1.32 & 3.81 & 39.91 & 39.91 & 3.81 \end{pmatrix}. \tag{56}$$

Calculating the product PQ^{-1} gives a measure on the least contributing edge. It follows that the edge $\{4, 5\}$ and $\{3, 4\}$ are the least contributing. Intuitively, it makes sense that the least contributing edges are the furthest away from the vertex that is controlled and observed. Also, due to symmetry we would expect the same value for $\{3, 4\}$ and $\{4, 5\}$. Minimalisation of the trace of the solutions does not uniquely describe the solution. Symmetry arguments must be added to obtain the result presented here. This is a drawback of the method, as it is

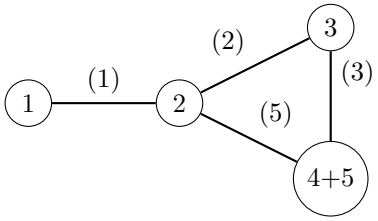

Figure 3: Graph after a one-step reduction

not always easy to identify these symmetry constraints. Choosing $\{4, 5\}$ for a one-step reduction results in the network in Figure 3.
To identify the reduced order effective Laplacian $\hat{L}_{\text{eff}}$, we first need to redefine our edge and vertex numbering such that the edge that is clustered will be represented in the incidence matrix by the entries E_{45} and E_{55}. Then the reduced effective graph Laplacian is given by

$$\hat{L}_{\text{eff}} = (VMV^T)^{-1}(V^T E E^T V), \tag{57}$$

where it is used that the weights of all edges are equal to 1. This results in

$$\hat{L}_{\text{eff}} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 3 & -1 & -1 \\ 0 & -1 & 2 & -1 \\ 0 & -0.5 & -0.5 & 1 \end{pmatrix}.$$

Using the reduced order effective graph Laplacian $\hat{L}_{\text{eff}}$, it can be seen that the reduced model also shows synchronisation. This is depicted in Figure 4.

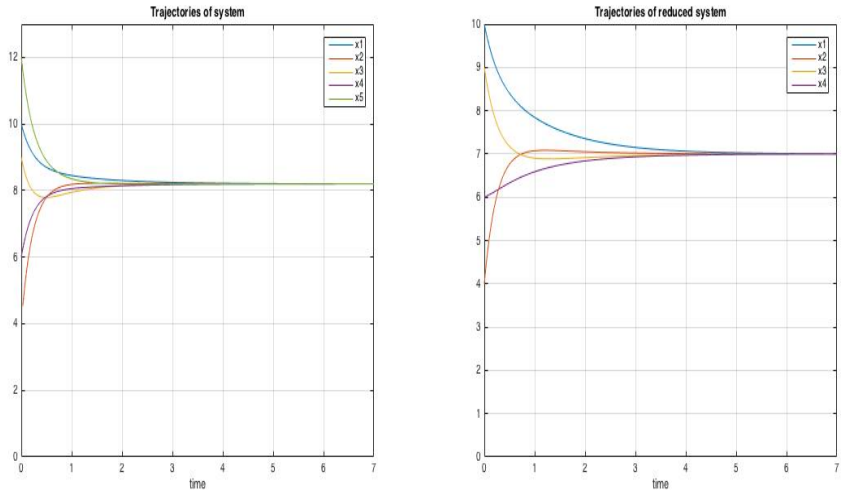

Figure 4: Trajectories of a system and its one step reduced system

In these figures, the initial condition is chosen to be $x_0 = \begin{pmatrix}10 & 4 & 9 & 6 & 12\end{pmatrix}$. This is the one-step clustering method. After completing these calculations, new control and output energy bounds must be calculated to do another one-step clustering. Computing these shows that the next vertices that should be clustered are those connected by edge (3). This results in the graph given in Figure 5.

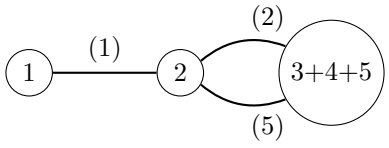

Figure 5: Graph of the two times reduced model

At this point, we need to add the edges (2) and (5) connecting the most right nodes before continuing. We also observe that we have lost the cycle and are now considering a network with a tree structure. After this point, the richer method of [3] can be used.
A more involved example shows some of the downfalls of the theory. Consider a graph as in Figure 6. Again the edges not in the spanning tree are represented by dashed lines.

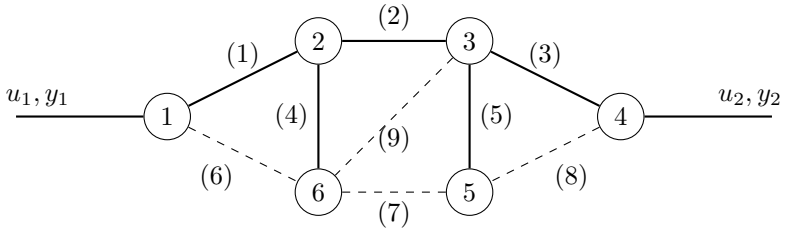

Figure 6: A more complicated graph

In this case, we do consider vertices and edges with different weights, such that

$$M = \operatorname{diag}\begin{pmatrix}1 & 2 & 8 & 1 & 9 & 2\end{pmatrix}, \quad R = \operatorname{diag}\begin{pmatrix}1 & 3 & 2 & 1 & 3 & 1 & 1 & 2 & 1\end{pmatrix}.$$

We take inputs and outputs on vertices 1 and 4, implying

$$G = \begin{pmatrix}1 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 1 & 0 & 0\end{pmatrix}^T, \quad H = \begin{pmatrix}1 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 1 & 0 & 0\end{pmatrix}.$$

As before, it is possible to solve the corresponding inequalities to obtain a measure on the output en control energies. These are given by

$$P = \operatorname{diag}\begin{pmatrix}0.44 & 0.62 & 0.26 & 0 & 0 & 0.27 & 0.44 & 0.01 & 0\end{pmatrix}, \tag{58}$$

$$Q = \operatorname{diag}\begin{pmatrix} 0 & 26.8 & 7.6 & 6.4 & 1267.9 & 1.9 & 0 & 0 & 6.5 \end{pmatrix}. \tag{59}$$

The zero eigenvalues of the matrix Q do not allow us to take the inverse of Q. However, the diagonal structure of Q makes it possible to compare edges with non-zero eigenvalues for Q. If an edge corresponds to a zero eigenvalue, the vertices on that edge take no energy to steer apart. Therefore, we consider these vertices as an improper choice to cluster. Then, if a control energy bound is found that is singular, we neglect these values and compare only the edges for which the product $(PQ^{-1})_i$ can be defined. In the example, this corresponds to the following.

$$PQ^{-1} = \operatorname{diag}\begin{pmatrix} \star & 23 & 33 & 0 & 0 & 135 & \star & \star & 0 \end{pmatrix} \times 10^{-3}. \tag{60}$$

This would suggest clustering of vertices 2 and 6, 3 and 6, or 3 and 5, should be done first. We choose to cluster edge (5) by the heuristic argument that it also has the highest value for Q. In doing so, we create a new graph with multiple edges between nodes 3 and 6 and nodes 3 and 4.

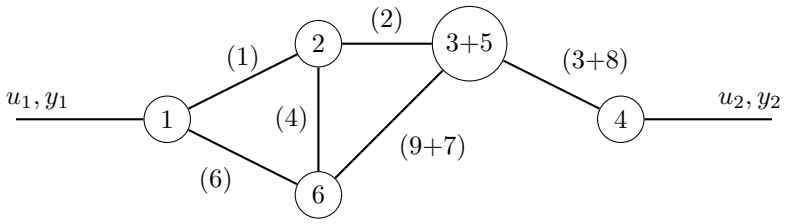

Figure 7: Graph after a one-step reduction

For this reduced model, we can again calculate upper and lower bounds using the Riccati and Lyapunov inequalities (44) and (34). These can be seen to be equal to

$$P_{new} = \operatorname{diag}\begin{pmatrix} 0.61 & 0.75 & 0.22 & 0 & 0.16 & 0 \end{pmatrix},$$
$$Q_{new} = \operatorname{diag}\begin{pmatrix} 0 & 34.8 & 4.7 & 1010.3 & 0 & 2.2 \end{pmatrix}.$$

Using the same arguments as before, we choose to cluster vertices 2 and 6 corresponding to edge 4, after which a simple line graph, without cycles, is again recovered.

Inspection of the output and control energy bounds in (55) shows an interesting property. In the first example, the energy bounds suggest clustering of edges (3) and (4). This is the same clustering we did by calculating new bounds for the reduced model. Taking into account that we need to add edge weights when they connect the same vertices, the second example suggests clustering of edges (5) and (4), just as was done by doing the calculations once more. This is not a coincidence. In [3] it was shown that for tree graphs the control and output bounds remain bounds for the reduced model, after deleting the value corresponding to the clustered edge. For graphs with cycles this does not hold

in general. Some thoughts on when this nonetheless can be applied are given in the appendix.
We remark that the error minimalisation procedure followed in this method stems from heuristic arguments and similarities with the method in [3]. An error analysis for model reduction by clustering for graphs with cycles may show a different procedure should be taken. Therefore, an error analysis would be highly beneficial to the clustering method considered.

8 Conclusion

In this thesis we have looked at a clustering-based model reduction method to approximate networked multi-agent systems. A technique introduced in [3] is extended for a certain class of networked systems to include graphs with cycles. Comparing the levels of controllability and observability in an edge system determines the least contributing edge, which has its connecting vertices clustered. The main contribution of this thesis is the idea that the dynamics of the edge system, and therefore the energy levels as well, can be restricted to a subspace, ensuring asymptotic stability and convergence. An asset of this method is that it aggregates vertices into clusters, such that the reduced order model has a neat physical interpretation. It is also shown that the property of synchronisation is preserved in the approximation. Not every question is answered in this thesis. A problem that is left open is in which cases the bounds determining the level of observability and controllability of a system can be used in the reduced order model as well. This can be investigated in future research. Also, the procedure would highly benefit from an error analysis. Other topics that can be considered relate to extending the method to the class of directed graphs. Although the idea of restricting the edge state vector to a subspace is not limited to the class of undirected graphs, some difficulties are expected to surface in a more general case. Particularly, it is not directly clear how a non-singular edge Laplacian can be obtained that, for example, will guarantee convergence. Extending this clustering method to a system with arbitrary linear dynamics can also be considered.

9 Appendix

Computing the upper and lower bounds on the level of observability and controllability respectively, is the main time-consuming task in the clustering-based reduction procedure discussed in this thesis. Being able to know what vertices to cluster for every step by only calculating the diagonal bounds once would make the method more useful. Therefore, we would like to extend a result in [3] to the case of arbitrary graphs. We would like to show that if a control energy bound is given by $\tilde{Q} = \text{diag}(\tilde{q}_1,\ \tilde{q}_2,\ \ldots, \tilde{q}_{n_e})$, then a control energy bound for the reduced model is given by $\hat{Q} = \text{diag}(\tilde{q}_1,\ \tilde{q}_2,\ \ldots, \tilde{q}_{n_e-1})$.
However, this does not hold in general for arbitrary graphs. A counter example would be the graph given in Figure 8.

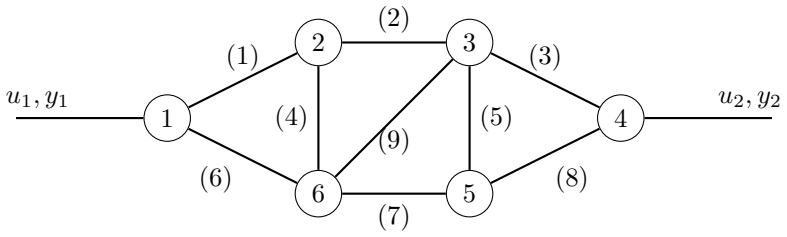

Figure 8: Graph of a counter example

Suppose that we cluster vertices 1 and 2. Then, after comparing with (59), we see that the supposed new bound would be

$$Q = \text{diag}\begin{pmatrix}26.8 & 7.6 & 8.4 & 1267.9 & 0 & 0 & 6.5\end{pmatrix}.$$

Assume the above result can be extended to this graph. Then this bound $\hat{Q}$ must be a solution to the algebraic Riccati inequality for the reduced model,

$$\hat{L}_\tau^T \hat{S}\hat{Q}\hat{S}^T + \hat{S}\hat{Q}\hat{S}^T\hat{L}_\tau - \hat{S}\hat{Q}\hat{S}^T\hat{E}_\tau^T\hat{M}^{-1}\hat{G}\hat{G}^T\hat{M}^{-1}\hat{E}_\tau\hat{S}\hat{Q}\hat{S}^T \geq 0, \qquad (61)$$

where $\hat{E}_\tau$, $\hat{M}$, $\hat{L}_\tau$ and $\hat{S}$ are defined with respect to the reduced graph.

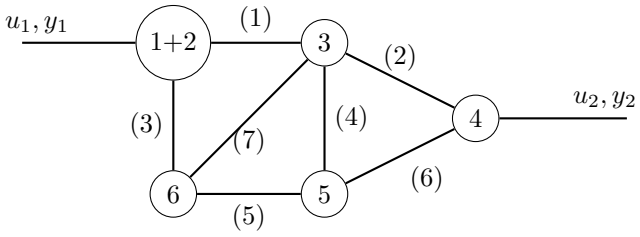

Figure 9: Reduced order graph

Computing the eigenvalues of the matrix on the left-hand side of inequality (61), shows that its eigenvalues have different signs. From this we can conclude that

the inequality is not fulfilled with this choice of $\hat{Q}$. So we see that this result does not extend to graphs with cycles. However, as was shown in the examples, there are cases where it is not necessary to calculate the energy bounds at every reduction. The remainder of this appendix is aimed at providing some insights on this observation.
Consider a system Σ as in (5) and a reduced system $\hat{\Sigma}_{n-1}$ as in (48). If the generalized controllability bound for system Σ_e, $\tilde{Q}$ can be written as

$$\tilde{Q} = \text{diag}(\tilde{q}_1,\ \tilde{q}_2,\ \dots, \tilde{q}_{n_e}),$$

then what we want is a generalized controllability bound, $\hat{Q}$ for the reduced order system $\hat{\Sigma}_{n-1}$ to be given by

$$\hat{Q} = \text{diag}(\tilde{q}_1,\ \tilde{q}_2,\ \dots, \tilde{q}_{n_e-1}).$$

We know that this does not hold in general, so there are other conditions that need to be satisfied. It can be shown that the above is true if the deleted $\tilde{q}_{n_e}$, the bound on the edge that was determined to be the least contributing, is sufficiently large.
To see this, we first note that the Riccati inequality (44) solves

$$x_e^T(L_e^TQ + QL_e - QE^TM^{-1}GG^TM^{-1}EQ)x_e \geq 0, \tag{62}$$

for all $x_e \in \text{im}\, S^T$. Here it is used that $x_e = S^Tx_\tau$ and that $L_eS^T = S^TL_\tau$. Using Lemma 2, it follows that (62) holds for all $x_e \in \text{im}\, E^T$, which is more useful to work with.
Giving a control energy bound for the reduced system, based on the original system means a relation between these systems must be found. To that end, define an input matrix G_e as $G_e = E^TM^{-1}G$. Then partition it as

$$G_e = \begin{pmatrix} G_{e,1} \\ G_{e,2} \end{pmatrix},$$

where $G_{e,2} \in \mathbb{R}$. Also partition the edge Laplacian accordingly as

$$L_e = \begin{pmatrix} L_{e,11} & L_{e,12} \\ L_{e,21} & L_{e,22} \end{pmatrix}.$$

It holds that

$$\hat{G}_e = G_{e,1} - L_{e,12}L_{e,22}^{-1}G_{e,2}. \tag{63}$$

This can be seen as follows. First, we note that $\hat{G}_e$ is defined as $\hat{G}_e = \hat{E}^T\hat{M}^{-1}V^TG$ and the original input matrix G_e likewise as $G_e = E^TM^{-1}G$. As we have partitioned the incidence matrix in (50), we partition the matrices M, R correspondingly as

$$M = \begin{pmatrix} M_1 & & \\ & m_{l-1} & \\ & & m_l \end{pmatrix}, \quad R = \begin{pmatrix} R_1 & \\ & r_{ll} \end{pmatrix}.$$

It is noted that r_{ll}, m_{l-1} and m_l are real positive numbers. Using these partitions and (51) the matrix $\hat{G}_e$ can now be calculated to be

$$\begin{aligned}\hat{G}_e &= \begin{pmatrix} E_{00}^T & E_{i0}^T + E_{j0}^T \end{pmatrix} \begin{pmatrix} M_1^{-1} & \\ & (m_{l-1}+m_l)^{-1} \end{pmatrix} \begin{pmatrix} I & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} G_1 \\ G_{l-1} \\ G_l \end{pmatrix} \\ &= E_{00}^T M_1^{-1} G_1 + \frac{1}{m_{l-1}+m_l}(E_{i0}^T + E_{j0}^T)(G_{l-1}+G_l).\end{aligned} \tag{64}$$

Also G_e can be written as a product of partitioned matrices

$$\begin{aligned}G_e &= \begin{pmatrix} E_{00}^T & E_{i0}^T & E_{j0}^T \\ 0 & E_{il}^T & E_{jl}^T \end{pmatrix} \begin{pmatrix} M_1^{-1} & & \\ & m_{l-1}^{-1} & \\ & & m_l^{-1} \end{pmatrix} \begin{pmatrix} G_1 \\ G_{l-1} \\ G_l \end{pmatrix} \\ &= \begin{pmatrix} E_{00}^T M_1^{-1} G_1 + \frac{1}{m_{l-1}} E_{i0}^T G_{l-1} + \frac{1}{m_l} E_{j0}^T G_l \\ \frac{1}{m_{l-1}} E_{il}^T G_{l-1} - \frac{1}{m_l} E_{il}^T G_l \end{pmatrix}.\end{aligned} \tag{65}$$

Here it is used that $E_{jl} = -E_{il}$. So now we know $G_{e,1}$ is the top entry of (65) and and $G_{e,2}$ the lower entry. In the same manner, the edge Laplacian can be computed using $L_e = E^T M^{-1} E R$, and the partitioning of the included matrices.

$$L_e = \begin{pmatrix} E_{00}^T & E_{i0}^T & E_{j0}^T \\ 0 & E_{il}^T & E_{jl}^T \end{pmatrix} \begin{pmatrix} M_1^{-1} & & \\ & m_{l-1}^{-1} & \\ & & m_l^{-1} \end{pmatrix} \begin{pmatrix} E_{00} & 0 \\ E_{i0} & E_{il} \\ E_{j0} & E_{jl} \end{pmatrix} \begin{pmatrix} R_1 & \\ & r_{ll} \end{pmatrix}$$

A careful computation gives that

$$L_{e,12} = \frac{1}{m_{l-1}} E_{i0}^T E_{il} r_{ll} - \frac{1}{m_l} E_{j0}^T E_{il} r_{ll},$$

and

$$\begin{aligned}L_{e,22} &= \frac{1}{m_{l-1}} E_{il}^T E_{il} r_{ll} + \frac{1}{m_l} E_{jl}^T E_{jl} r_{ll} \\ &= \frac{r_{ll}(m_{l-1}+m_l)}{m_{l-1} m_l}.\end{aligned}$$

Now all the required terms are defined, and we compute that

$$\begin{aligned}G_{e,1} - L_{e,12} L_{e,22}^{-1} G_{e,2} = E_{00}^T M_1^{-1} G_1 + (\frac{1}{m_{l-1}} - \frac{m_l}{m_{l-1}(m_{l-1}+m_l)}) E_{i0}^T G_{l-1} + \\ (\frac{1}{m_l} - \frac{m_{l-1}}{m_l(m_{l-1}+m_l)})) E_{j0}^T G_l + \frac{1}{m_{l+1}+m_l}(E_{i0}^T G_l + E_{j0}^T G_{l-1}).\end{aligned} \tag{66}$$

It is easily verified that

$$\frac{1}{m_l} - \frac{m_{l-1}}{m_l(m_{l-1}+m_l)} = \frac{1}{m_{l-1}+m_l},$$

and similarly for the other expression between brackets. It can now be seen that (66) equals (64). So we conclude that indeed it is possible to relate the new input matrix to the original one via the equation

$$\hat{G}_e = G_{e,1} - L_{e,12}L_{e,22}^{-1}G_{e,2}.$$

The same partitioning also allows identifying the new edge Laplacian $\hat{L}_e$ as

$$\hat{L}_e = L_{e,11} - L_{e,12}L_{e,22}^{-1}L_{e,21}. \tag{67}$$

This can be shown similarly to the equation for $\hat{G}_e$ (see [4] for a proof, where it is noted that in our case $w_{ij} = m_i r_j$). Since we compare the reduced model with the non-reduced model it is natural to make the distinction

$$Q = \begin{pmatrix} \hat{Q} & 0 \\ 0 & \tilde{q}_{n_e} \end{pmatrix}, \tag{68}$$

where $\tilde{q}_{n_e}$ is the bound on the edge that is lost by clustering. Also, by a suitable numbering of edges, write $x_e = \begin{pmatrix} \hat{x}_e^T & \bar{x}_e^T \end{pmatrix}^T$. Here $\bar{x}_e$ is the edge that connects the vertices that will be clustered. Now, we define a transformation matrix to connect the two systems by

$$T_c = \begin{pmatrix} I & -L_{e,12}L_{e,22}^{-1} \\ 0 & 1 \end{pmatrix}.$$

Then, by defining a new coordinate $\eta_e = Q^{-1}T_c^{-T}Qx_e$, it can be seen that (62) can be written as

$$\begin{pmatrix} \hat{\eta}_e^T & \bar{\eta}_e^T \end{pmatrix} \begin{pmatrix} \hat{Q} & 0 \\ 0 & \tilde{q}_{n_e} \end{pmatrix} T_c(Q^{-1}L_e^T + L_eQ^{-1} - G_e^TG_e)T_c^T \begin{pmatrix} \hat{Q} & 0 \\ 0 & \tilde{q}_{n_e} \end{pmatrix} \begin{pmatrix} \hat{\eta}_e \\ \bar{\eta}_e \end{pmatrix} \geq 0, \tag{69}$$

for all η_e such that $Q^{-1}T_c^TQ\eta_e \in \text{im } E^T$. It is remarked that we assume that the matrix Q is invertible. This is not necessarily satisfied, as was seen in the example. However, at the expense of the accuracy of the approximation, the energy bound Q can be made non-singular. An example is the matrix constructed to prove the existence of a diagonal positive semi-definite matrix in Lemma 4. Using a result in [7], the matrix T_c finally relates both reduced and original systems with each other. In particular,

$$\begin{aligned} &T_c(Q^{-1}L_e^T + L_eQ^{-1} - G_e^TG_e)T_c^T = \\ &\quad \begin{pmatrix} \hat{Q}^{-1}\hat{L}_e^T + \hat{L}_e\hat{Q}^{-1} - \hat{G}_e^T\hat{G}_e & \hat{Q}^{-1}L_{e,21}^T - L_{e,12}\tilde{q}_{n_e}^{-1} - \hat{G}_eG_{e,12}^T \\ L_{e,21}\hat{Q}^{-1} - \tilde{q}_{n_e}^{-1}L_{e,12}^T - G_{e,12}\hat{G}_e^T & \tilde{q}_{n_e}^{-1}L_{e,22}^T + \hat{L}_{e,22}\tilde{q}_{n_e}^{-1} - G_{e,2}^TG_{e,2} \end{pmatrix}. \end{aligned}$$

This we substitute in inequality (69). First, we note that for the inequality (69) we only consider η_e such that $Q^{-1}T_c^T Q\eta_e \in \operatorname{im} E^T$, so that we would like to know how these are characterised. It is known that

$$\begin{pmatrix}\hat{x}_e\\ \bar{x}_e\end{pmatrix} = \begin{pmatrix}\hat{Q}^{-1} & 0\\ 0 & \tilde{q}_{n_e}^{-1}\end{pmatrix}\begin{pmatrix}I & 0\\ -L_{e,22}^{-1,T}L_{e,12}^T & 1\end{pmatrix}\begin{pmatrix}\hat{Q} & 0\\ 0 & \tilde{q}_{n_e}\end{pmatrix}\begin{pmatrix}\hat{\eta}_e\\ \bar{\eta}_e\end{pmatrix}.$$

Therefore,

$$\begin{pmatrix}\hat{x}_e\\ \bar{x}_e\end{pmatrix} = \begin{pmatrix}I & 0\\ -\tilde{q}_{n_e}^{-1}L_{e,22}^{-1}L_{e,12}^T\hat{Q} & 1\end{pmatrix}\begin{pmatrix}\hat{\eta}_e\\ \bar{\eta}_e\end{pmatrix} = \begin{pmatrix}E_{00}^T & E_{i0}^T & E_{j0}^T\\ 0 & E_{il}^T & E_{jl}^T\end{pmatrix}\begin{pmatrix}z_1\\ z_2\\ z_3\end{pmatrix}, \tag{70}$$

for some vector $z = \begin{pmatrix}z_1^T & z_2 & z_3\end{pmatrix}^T$. An edge controllability bound for the reduced model is given by

$$\hat{x}_e^T(\hat{L}_e^T\hat{Q} + \hat{Q}\hat{L}_e - \hat{Q}\hat{G}_e^T\hat{G}_e\hat{Q})\hat{x}_e \geq 0, \tag{71}$$

for all $\hat{x}_e \in \operatorname{im}\hat{E}^T$. This is equivalent to the upper left entry of (69). To let (71) be satisfied for all $\hat{x}_e \in \operatorname{im}\hat{E}^T$, it can be seen by comparing (70) and (51) that necessarily $z_2 = z_3$. For this choice of z_2, and z_3 it follows that $\bar{x}_e = 0$ and therefore this implies that

$$\bar{\eta}_e = \tilde{q}_{n_e}^{-1}L_{e,22}^{-1}L_{e,12}^T\hat{Q}\hat{\eta}_e. \tag{72}$$

For ease of notation, we define

$$\chi = \tilde{q}_{n_e}^{-1}L_{e,22}^{-1}L_{e,12}^T\hat{Q}.$$

Now we are in a position to write the inequality (69) purely in terms of $\hat{x}_e$ as follows:

$$\begin{aligned}\hat{x}_e^T(\hat{L}_e^T\hat{Q} + \hat{Q}\hat{L}_e - \hat{Q}\hat{G}_e^T\hat{G}_e\hat{Q})\hat{x}_e&\\ + \hat{x}_e^T\chi^T(\tilde{q}_{n_e}^{-1}L_{e,22}^T + \hat{L}_{e,22}\tilde{q}_{n_e}^{-1} - G_{e,2}^TG_{e,2})\chi\hat{x}_e&\\ + 2\hat{x}_e^T(\hat{Q}^{-1}L_{e,21}^T - L_{e,12}\tilde{q}_{n_e}^{-1} - \hat{G}_eG_{e,12}^T)\chi\hat{x}_e \geq 0.&\end{aligned} \tag{73}$$

If the value $\tilde{q}_{n_e}^{-1}$ is sufficiently small, then this inequality gives a contradiction if (71) is not satisfied.
This reasoning lacks mathematical rigour, because conditions on $\tilde{q}_{n_e}^{-1}$ for being sufficiently small are not provided. However, it gives a first indication that in some cases, the controllability bound does not need to be computed at every step.
For the observability case, a complementing result is expected to exist. However, due to the definition of the output as $y_e = HE_\tau(E_\tau^TE_\tau)^{-1}(SS^T)^{-1}Sx_e$, a relation similar to (63) is not as easily attained. If it is true that $\hat{H}_e = H_{e,1} - L_{e,12}L_{e,22}^{-1}H_{e,2}$, then using the same method the result can be extended to the observability case.

References

[1] A. Antoulas. *Approximation of Large-Scale Dynamical Systems*. Society for Industrial and Applied Mathematics, 2005.

[2] Alessandro Astolfi. Model reduction by moment matching for linear and nonlinear systems. *IEEE Transactions on Automatic Control*, 55(10):2321–2336, 2010.

[3] B. Besselink, H. Sandberg, and K. H. Johansson. Clustering-based model reduction of networked passive systems. *IEEE Transactions on Automatic Control*, 61(10):2958–2973, Oct 2016.

[4] Bart Besselink, Henrik Sandberg, and Karl Henrik Johansson. Model reduction of networked passive systems through clustering. In *Control Conference (ECC), 2014 European*, pages 1069–1074. IEEE, 2014.

[5] F. Bullo. *Lectures on Network Systems*. Version 0.95, 2017. With contributions by J. Cortes, F. Dorfler, and S. Martinez.

[6] Geir E Dullerud and Fernando Paganini. *A course in robust control theory: a convex approach*, volume 36. Springer Science & Business Media, 2013.

[7] K Fernando and H Nicholson. Singular perturbational model reduction of balanced systems. *IEEE Transactions on Automatic Control*, 27(2):466–468, 1982.

[8] Chris Godsil and Gordon F Royle. *Algebraic graph theory*, volume 207. Springer Science & Business Media, 2013.

[9] Roger A Horn and Charles R Johnson. *Matrix analysis*. Cambridge university press, 2012.

[10] Takayuki Ishizaki, Kenji Kashima, Jun-ichi Imura, and Kazuyuki Aihara. Model reduction and clusterization of large-scale bidirectional networks. *IEEE Transactions on Automatic Control*, 59(1):48–63, 2014.

[11] Hidde-Jan Jongsma, Harry L Trentelman, and M Kanat Camlibel. Model reduction of consensus networks by graph simplification. In *Decision and Control (CDC), 2015 IEEE 54th Annual Conference on*, pages 5340–5345. IEEE, 2015.

[12] Mehran Mesbahi and Magnus Egerstedt. *Graph theoretic methods in multiagent networks*. Princeton University Press, 2010.

[13] Nima Monshizadeh, Harry L Trentelman, and M Kanat Camlibel. Projection-based model reduction of multi-agent systems using graph partitions. *IEEE Transactions on Control of Network Systems*, 1(2):145–154, 2014.

[14] Bruce Moore. Principal component analysis in linear systems: Controllability, observability, and model reduction. *IEEE transactions on automatic control*, 26(1):17–32, 1981.

[15] Steven H Strogatz. Exploring complex networks. *Nature*, 410(6825):268–276, 2001.

[16] Harry L Trentelman, Anton A Stoorvogel, and Malo Hautus. Control theory for linear systems. *Communications and Control Engineering Series, Springer*, 2002.

[17] AJ van der Schaft. On model reduction of physical network systems. In *Proceedings of the 21st International Symposium on Mathematical Theory of Networks and Systems*, pages 1419–1425, 2014.

[18] Jan Willems. Least squares stationary optimal control and the algebraic riccati equation. *IEEE Transactions on Automatic Control*, 16(6):621–634, 1971.

[19] D. Zelazo and M Mesbahi. Edge agreement: Graph-theoretical performance bounds and passivity analysis. *IEEE Transactions on Automatic Control*, 56(3):544–555, March 2011.