Model reduction of multi-agent systems

Kanat Camlibel

Johann Bernoulli Institute of Mathematics and Computer Science

University of Groningen

Jan C. Willems Center for Systems and Control

June 4, 2015

joint work with

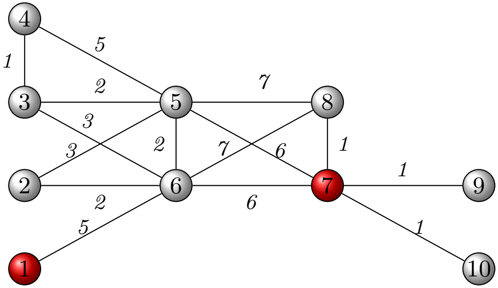
(in alphabetical order)

- Hidde-Jan Jongsma
- Nima Monshizadeh
- Harry Trentelman

outline

1. Introduction
2. Leader-Follower Multi-Agent Systems
3. Network Approximation by Graph Partitions
4. Preservation of Consensus
5. $\mathcal{H}_2$ Analysis of the Model Reduction Error
6. Model Reduction by Edge Removal
7. Directions for Future Research

introduction



- General problem: given a large scale networked multi-agent system with leaders and followers, approximate it by a lower order network, while preserving (some of) the network structure.
- Does the approximated network preserve important properties such as *consensus*?
- How good is the approximation? Can one compute (upper and lower bounds on) the error?

graph theory

- Weighted directed graph $G = (V, E, A)$, $V = \{1, 2, \dots, p\}$, E set of arcs, adjacency matrix $A = [a_{ij}]$.
- Laplacian of G: $L = D - A$, $D = \mathrm{diag}(d_1, d_2, \dots, d_p)$, $d_i = \sum_j a_{ij}$.
- Incidence matrix of G: $R = [r_{ij}]$ with

$$r_{ij} = \begin{cases} 1 & \text{if vertex } i \text{ is the head of arc } j \\ -1 & \text{if vertex } i \text{ is the tail of arc } j \\ 0 & \text{otherwise} \end{cases}$$

  for $i = 1, 2, \dots, p$ and $j = 1, 2, \dots, q$ (q is the total number of arcs).
- Incidence matrix *undirected* graph: assign arbitrary orientation to the edges and take the incidence matrix of the corresponding directed graph.
- Edge weight matrix $W = \mathrm{diag}(w_1, w_2, \dots, w_q)$ with w_j is the weight associated to the edge (arc) j, $j = 1, 2, \dots, q$.
- For undirected graphs Laplacian: $L = RWR^T$.

leader follower multi-agent systems

$G = (V, E, A)$ weighted undirected graph

leaders $V_L = \{v_1, v_2, \ldots, v_\ell\} \subseteq V$ leaders, followers $V_F := V \backslash V_L$

Leader-follower multi-agent system:

$$\dot{x}_i = \begin{cases} \sum_{j=1}^{p} a_{ij}(x_j - x_i) & \text{if } i \in V_F \\ \sum_{j=1}^{p} a_{ij}(x_j - x_i) + u_k & \text{if } i \in V_L \end{cases}$$

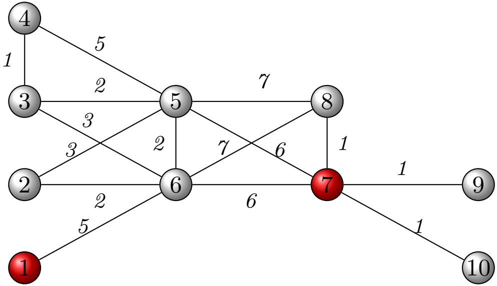
$x_i \in \mathbb{R}$ state of agent i, $u_k \in \mathbb{R}$ external input applied to agent $i = v_k$.
In compact form

$$\dot{x} = -Lx + Mu,$$

with L the Laplacian of G, M is given by

$$M_{ik} = \begin{cases} 1 & \text{if } i = v_k \\ 0 & \text{otherwise.} \end{cases}$$

example



$$\dot{x} = -Lx + Mu$$

$$L = \begin{bmatrix} 5 & 0 & 0 & 0 & 0 & -5 & 0 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 & -3 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 6 & -1 & -2 & -3 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 6 & -5 & 0 & 0 & 0 & 0 & 0 \\ 0 & -3 & -2 & -5 & 25 & -2 & -6 & -7 & 0 & 0 \\ -5 & -2 & -3 & 0 & -2 & 25 & -6 & -7 & 0 & 0 \\ 0 & 0 & 0 & 0 & -6 & -6 & 15 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & -7 & -7 & -1 & 15 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 \end{bmatrix}, \qquad M = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

problem statement

Problem

Approximate the leader-follower multi-agent system

$$\dot{x} = -Lx + Mu$$

by a reduced order multi-agent system

$$\dot{\hat{x}} = -\widehat{L}\hat{x} + \widehat{M}u$$

such that the spatial structure of the network is preserved.

Idea

We want to do this by reducing the size of the network graph through clustering of the agents/vertices

Petrov-Galerkin projections

Consider an arbitrary system

$$\dot{x} = Ax + Bu \qquad y = Cx$$

with state space $\mathbb{R}^n$. Let $W, V \in \mathbb{R}^{n \times r}$ (with $r < n$) such that $W^T V = I_r$. Then VW^T is a projection.

A reduced order model (projected model) is then given by

$$\dot{\widehat{x}} = W^T AV\widehat{x} + W^T Bu \qquad y = CV\widehat{x}$$

with reduced order state space $\mathbb{R}^r$.

General reduction framework: Krylov based, truncation methods, moment matching use this projection with appropriate choice of matrices V and W.

Idea

Put idea of clustering of network graph into the Petrov-Galerkin framework

graph partitions

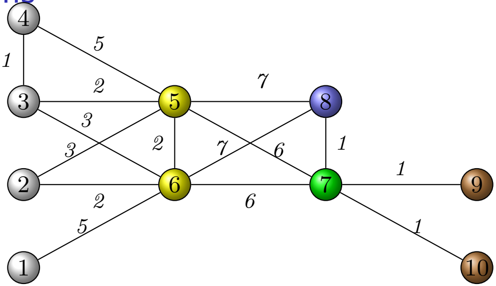
- $V = \{1, 2, \ldots, p\}$ vertex set of a graph G.
- Any nonempty subset of V is called a cell of V.
- Collection of cells, given by $\pi = \{C_1, C_2, \ldots, C_r\}$, called a partition of V if $\cup_i C_i = V$ and $C_i \bigcap C_j = \varnothing$ whenever $i \neq j$.
- For a cell $C \subset V$, the characteristic vector of C is the p-dimensional column vector $p(C)$ with

$$p_i(C) = \begin{cases} 1 & \text{if } i \in C, \\ 0 & \text{otherwise.} \end{cases}$$

- For a partition $\pi = \{C_1, C_2, \ldots, C_r\}$, the characteristic matrix of π is

$$P(\pi) = \begin{bmatrix} p(C_1) & p(C_2) & \cdots & p(C_r) \end{bmatrix}.$$

graph partitions



- Vertex set: $V = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$.
- Cell: any nonempty subset of V
- Partition: $\pi = \{\{1, 2, 3, 4\}, \{5, 6\}, \{7\}, \{8\}, \{9, 10\}\}$
- Characteristic matrix:

$$P(\pi) = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}^T .$$

projection by graph partitions

- A direct application of Petrov-Galerkin projections will destroy the spatial structure of the network (e.g. balanced truncation, moment matching)
- Instead: choose a graph partition, characteristic matrix $P(\pi)$.

$$W := P(\pi)(P^T(\pi)P(\pi))^{-1}$$

$$V := P(\pi).$$

- Corresponding reduced order system:

$$\dot{\hat{x}} = -\hat{L}\hat{x} + \widehat{M}u,$$

  - $\hat{L} = W^T L V = (P^T P)^{-1} P^T L P$
  - $\widehat{M} = W^T M = (P^T P)^{-1} P^T M$

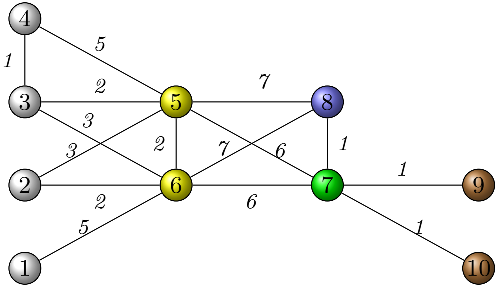
properties of partition-based projection

- $\widehat{L}$ is the Laplacian of a weighted *directed* graph, $\widehat{G} = (\widehat{V}, \widehat{E}, \widehat{A})$.
- Each cell of π in G becomes a vertex in $\widehat{G}$.
- There is an arc from vertex α to vertex β in $\widehat{G}$ if and only if there exist $i \in C_\alpha$ and $j \in C_\beta$ with $\alpha \neq \beta$ such that $\{i, j\} \in E$.
- Therefore, $\widehat{G}$ is a *symmetric directed* graph, i.e. $(\alpha, \beta) \in \widehat{E} \Leftrightarrow (\beta, \alpha) \in \widehat{E}$.
- Relationship between A and $\widehat{A} = [\widehat{a}_{\alpha\beta}]$ given by

$$\widehat{a}_{\alpha\beta} = \frac{1}{|C_\alpha|} \sum_{i \in C_\alpha, j \in C_\beta} a_{ij},$$

  for $\alpha \neq \beta$.
- The input weights in $\widehat{M}$ depend on the cardinality of the leader cells.

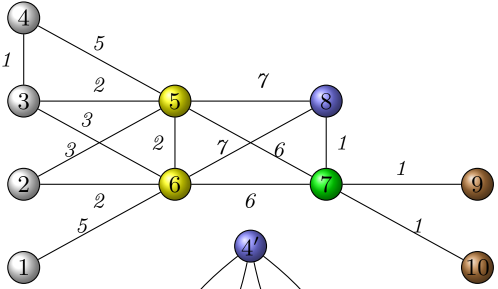
relationship between the original and the reduced model



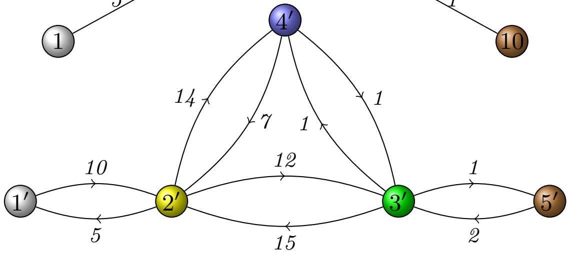
- The reduced order model: $\dot{\hat{x}} = -\widehat{L}\hat{x} + \widehat{M}u$
  - $$\widehat{L} = \begin{bmatrix} 5 & -5 & 0 & 0 & 0 \\ -10 & 23 & -6 & -7 & 0 \\ 0 & -12 & 15 & -1 & -2 \\ 0 & -14 & -1 & 15 & 0 \\ 0 & 0 & -1 & 0 & 1 \end{bmatrix}, \qquad \widehat{M} = \begin{bmatrix} 0.25 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

relationship between the original and the reduced model

- Original graph:



- Reduced graph:



preservation of consensus

- If the weighted undirected graph G is connected, then the multi-agent system $\dot{x} = -Lx$ reaches *consensus*, i.e. for all $i, j \in V$ we have $x_i(t) - x_j(t) \to 0$ as $t \to \infty$.
- Since $\widehat{L} = (P^T P)^{-1} P^T L P$, we see that

$$(P^T P)^{\frac{1}{2}} \widehat{L} (P^T P)^{-\frac{1}{2}} = (P^T P)^{-\frac{1}{2}} P^T L P (P^T P)^{-\frac{1}{2}}.$$

  Thus $\widehat{L}$ is similar to a symmetric matrix, and hence has *only real eigenvalues*.
- Moreover: the eigenvalues $\widehat{\lambda}_i$ of $\widehat{L}$ *interlace* the eigenvalues λ_i of L:

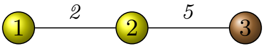
$$\lambda_i \leqslant \widehat{\lambda}_i \leqslant \lambda_{p-r+i}$$

  for $i = 1, 2 \ldots, r$.
- In particular $\lambda_2 \leqslant \widehat{\lambda}_2$ so the reduced order system $\dot{\widehat{x}} = -\widehat{L}\widehat{x}$ reaches consensus with rate of convergence at least as fast as the original network.

input-output approximation of networks

- $G = (V, E, A)$ undirected weighted graph with incidence matrix R, and edge weight matrix W.
- Original model: $\dot{x} = -Lx + Mu$
- We add output variables as: $y = W^{\frac{1}{2}} R^T x$. Note that $\|y\|^2 = x^T L x$.
- Example:

$$y = \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix}^{\frac{1}{2}} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$



- Output variables after projection: $\widehat{y} = W^{\frac{1}{2}} R^T P \widehat{x}$
- In this example: $$\widehat{y} = \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix}^{\frac{1}{2}} \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \widehat{x}_1 \\ \widehat{x}_2 \end{bmatrix}$$

model reduction error

Original model:

$$\dot{x} = -Lx + Mu$$
$$y = W^{\frac{1}{2}} R^T x$$

Reduced model:

$$\dot{\widehat{x}} = -\widehat{L}\widehat{x} + \widehat{M}u$$
$$\widehat{y} = W^{\frac{1}{2}} \widehat{R}^T \widehat{x}$$

- Transfer matrices

$$S(s) = W^{\frac{1}{2}} R^T (sI + L)^{-1} M \qquad \text{and} \qquad \widehat{S}(s) = W^{\frac{1}{2}} \widehat{R}^T (sI + \widehat{L})^{-1} \widehat{M}$$

- Note S and $\widehat{S}$ are in $\mathcal{H}_2$. Approximation error (squared $\mathcal{H}_2$-norm):

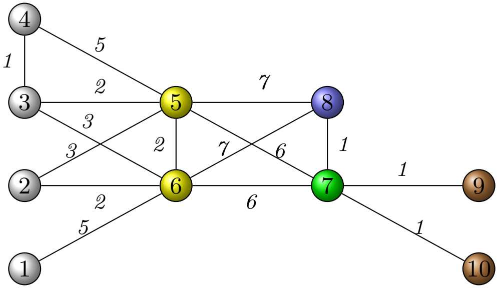
$$\|S - \widehat{S}\|_2^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} \|S(i\omega) - \widehat{S}(i\omega)\|^2 d\omega$$

- Problem: for a given partition, find an expression for this error. Find a priori upper or lower bounds.

almost equitable partitions

Def. For an unweighted undirected graph, a partition $\pi = \{C_1, C_2, \ldots, C_r\}$ is said to be *almost equitable* if for any pair (i, j) with $1 \leqslant i \neq j \leqslant r$, there exists an integer b_{ij} such that any vertex $v \in C_i$ has b_{ij} neighbors in C_j.

Def. Let $G = (V, E, A)$ be a weighted undirected graph. A partition $\pi = \{C_1, C_2, \ldots, C_r\}$ is called *almost equitable* if for any pair (i, j) with $1 \leqslant i \neq j \leqslant r$, there exists a real number b_{ij} such that $\sum_{\substack{(v,w)\in E \\ w\in C_j}} a_{ij} = b_{ij}$ for all $v \in C_i$.



model reduction error

Thm. Let

$\pi = \{C_1, C_2, \ldots, C_r\}$ be an AEP of G,

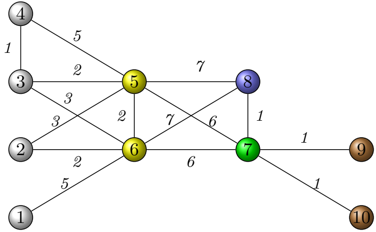
$V_L = \{v_1, v_2, \ldots, v_\ell\}$ leader vertices,

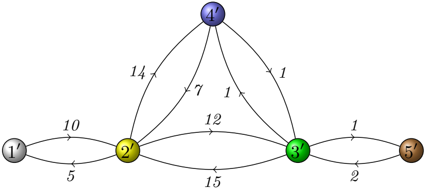
k_i be s.t. $v_i \in C_{k_i} \quad (i = 1, 2 \ldots, \ell)$.

The (normalized) model reduction error is given by

$$E(\pi) = \frac{\|S - \widehat{S}\|_2^2}{\|S\|_2^2} = \frac{\sum_{i=1}^{\ell}(1 - \frac{1}{|C_{k_i}|})}{\ell(1 - \frac{1}{p})}.$$

example





$$V_L = \{1, 7\}$$

$$E(\pi) = \frac{\sum_{i=1}^{\ell}(1 - \frac{1}{|C_{k_i}|})}{\ell(1 - \frac{1}{p})} = \frac{(1 - \frac{1}{4}) + (1 - \frac{1}{1})}{2(1 - \frac{1}{10})} = \frac{5}{12}$$

.

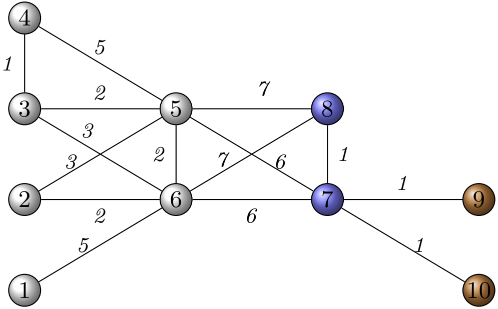
remarks

- $0 \leqslant E(\pi) \leqslant 1$
- $E(\pi) = 0$ for $\pi = \{\{1\}, \{2\}, \dots, \{p\}\}$
- $E(\pi) = 1$ for $\pi = \{V\}$
- $E(\pi)$ determined by the cardinalities of those cells in π that contain the leaders.

Cor. If π is an almost equitable partition with the property that $\{v_i\} \in \pi$ for all $i = 1, 2 \dots, \ell$ (i.e. each leader is the unique element in its cell) then $E(\pi) = 0$.

arbitrary partitions

- What if we start with an *arbitrary* (not necessarily AEP) partition? In that case the closed form expression fails. But: can we make an estimate of $E(\pi)$?



- $\pi = \{\{1,2,3,4,5,6\},\{7,8\},\{9,10\}\}$ is not an an almost equitable partition.

partial ordering of partitions

- Let $G = (V, E, A)$ be a weighted undirected graph. Partial ordering on all partitions of G. Let π_1 and π_2 be two partitions. We say $\pi_1 \leqslant \pi_2$ if every cell in π_1 is a subset of a cell in π_2: π_1 is *finer* then π_2, π_2 is *coarser* than π_1.
- In terms of the characteristic matrices: $\pi_1 \leqslant \pi_2$ if and only if $\operatorname{im} P(\pi_2) \subseteq \operatorname{im} P(\pi_1)$.
- Using the fact that for any AEP π_0, $\operatorname{im} P(\pi_0)$ is L-invariant we obtain:

Thm. Let π_0 be an AEP of G. Then for every partition π that is coarser than π_0 we have $E(\pi_0) \leqslant E(\pi)$.

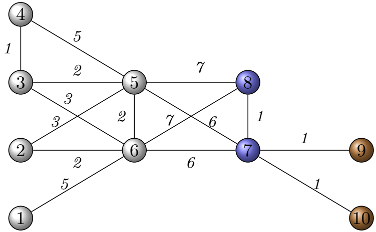
$\pi^*_{\mathrm{AEP}}(\pi)$: *the maximal almost equitable partition finer than* π.

Cor. For any partition π, we have $E(\pi) \geqslant E\big(\pi^*_{\mathrm{AEP}}(\pi)\big)$.

example

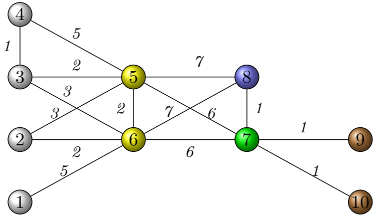
$V_L\{1,7\}$

π



$E(\pi) \geqslant \frac{5}{12}$

$\pi^*_{\mathrm{AEP}}(\pi)$



$E(\pi^*_{\mathrm{AEP}}(\pi)) = \frac{5}{12}$

model reduction by edge removal

so far clustering of the nodes

$$G = (V, E, A) \xrightarrow{P(\pi)} \widehat{G} = (\widehat{V}, \widehat{E}, \widehat{A})$$

undirected weighted → directed symmetric weighted

$$\dot{x} = -Lx + Mu \qquad \dot{\widehat{x}} = -\widehat{L}\widehat{x} + \widehat{M}u$$

$$y = W^{\frac{1}{2}} R^T x \qquad \widehat{y} = W^{\frac{1}{2}} \widehat{R}^T \widehat{x}$$

$$\mathrm{card}(V) > \mathrm{card}(\widehat{V})$$

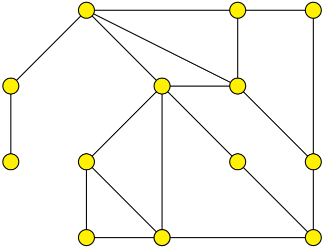
from now on removing edges

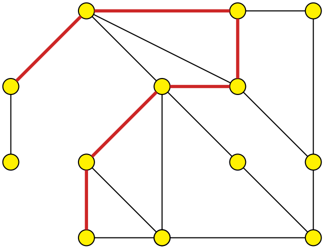
$$G = (V, E) \xrightarrow{\text{edge removal}} \widehat{G} = (V, \widehat{E})$$

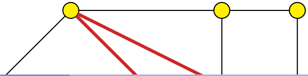
undirected weighted → undirected weighted

$$\mathrm{card}(E) > \mathrm{card}(\widehat{E})$$

paths, cycles, trees







$G = (V, E)$ undirected unweighted.

A path is a sequence of nodes such that sequential nodes are connected.

A cycle is a path where the first and last node in the path are the same.

error bound for reduction by edge removal

Thm. Let $G = (V, E)$ be an undirect unweighted graph with the Laplacian L and incidence matrix R. Consider the system Σ

$$\dot{x} = -Lx + d \qquad y = R^T x.$$

Let T be a spanning tree of G with missing independent cycles $c_1, c_2, \ldots, c_k$. Consider the system $\widehat{\Sigma}$

$$\dot{\widehat{x}} = -\widehat{L}\widehat{x} + d \qquad \widehat{y} = \widehat{R}^T \widehat{x}$$

where $\widehat{L}$ is the Laplacian and $\widehat{R}$ is the incidence matrix of T. Then, the approximation error is given by

$$\|\Sigma - \widehat{\Sigma}\|_{\mathcal{H}_2}^2 \leqslant \frac{1}{2} \sum_{i=1}^{k} \left(\frac{\ell(c_i) - 1}{1 + \frac{\lambda_2}{\lambda_2 + \lambda_p}\big(\ell(c_i) - 1\big)} + \frac{1}{\ell(c_i)} - 1 \right)$$

where λ_2 and λ_p are the smallest nonzero and the largest eigenvalues of $\widehat{L}$.

directions for future research

- Extension to directed weighted graphs.
- Instead of a lower bound on the error associated with an arbitrary partition we would like to find *upper bounds*.
- Extension to the $\mathcal{H}_\infty$ norm of the error transfer matrix.
- Extension to multi-agent systems with arbitrary homogeneous (heterogeneous) linear dynamics at the vertices.
- Combination of clustering method with edge removal.