Model Reduction of Complex Dynamical Networks based on State Aggregation

Takayuki Ishizaki (Tokyo Institute of Technology)
Contents

- Overview of model reduction problem (preliminary)
  - what is model reduction?
  - error evaluation in terms of $\mathcal{H}_\infty/\mathcal{H}_2$-norms

- Network structure preserving model reduction (main)
  - system description of dynamical networks
  - key idea of state aggregation with an $\mathcal{H}_\infty$-error evaluation
  - numerical example

- State aggregation with an $\mathcal{H}_2$-error evaluation (additional)
  - controllability gramian
Overview of Model Reduction

Given stable system

\[ \Sigma : \begin{cases} \dot{x} = Ax + Bu \\ y = Cx \end{cases} \]

Find

\[ P x = \hat{x}, \quad P \in \mathbb{R}^{\Delta \times n} \]

\[ \hat{A} = P A P^\dagger, \quad \hat{B} = P B, \quad \hat{C} = C P^\dagger \]

\[ \text{Find } P \text{ such that } \|\Sigma - \hat{\Sigma}\| \text{ is small enough} \]

Find stable reduced model

\[ \hat{\Sigma} : \begin{cases} \dot{\hat{x}} = \hat{A}\hat{x} + \hat{B}u \\ \hat{y} = \hat{C}\hat{x} \end{cases} \]

Dimension of state: \( \Delta < n \)
Typical System Norms

stable system \( \Sigma : \begin{cases} \dot{x} = Ax + Bu \\ y = Cx \end{cases} \)

solution \( y(t) = \int_0^\infty h(t - \tau)u(\tau)d\tau \)

impulse response \( h(t) := Ce^{At}B \)

transfer function \( H(s) := \mathcal{L}[h] = C(sI_n - A)^{-1}B \)

\( H_\infty\)-norm

\[
\|\Sigma\|_{H_\infty} = \sup_{u \neq 0} \frac{\|y(t)\|_{L^2}}{\|u(t)\|_{L^2}} = \sup_{\omega \in \mathbb{R}} \|H(j\omega)\|
\]

\( \mathcal{H}_2\)-norm \( \|\Sigma\|_{\mathcal{H}_2} = \|h\|_{L^2} = \left( \frac{1}{2\pi} \int_{-\infty}^{\infty} \|H(j\omega)\|_F^2 \, d\omega \right)^{\frac{1}{2}} \)

\( \mathcal{L}_2\)-norm of \( f : \mathbb{R}_+ \rightarrow \mathbb{R}^{p \times m} \)

\[
\|f\|_{L^2} := \left( \int_0^\infty \|f(t)\|_F^2 \, dt \right)^{\frac{1}{2}}
\]

\( \mathcal{L}[\cdot] \) : Laplace transform of \( \cdot \)

\( \|\cdot\|_F \) : Frobenius norm of \( \cdot \)
Contents

- Overview of model reduction problem (preliminary)
  - what is model reduction?
  - error evaluation in terms of $H_\infty/H_2$-norms

- Network structure preserving model reduction (main)
  - system description of dynamical networks
  - key idea of state aggregation with an $H_\infty$-error evaluation
  - numerical example

- State aggregation with an $H_2$-error evaluation (additional)
  - controllability gramian

5/22
**System Description**

[Definition] (single-input) **Dynamical Network** \((A, b)\)

A linear system \(\dot{x} = Ax + bu\) with \(A = \{a_{i,j}\} \in \mathbb{R}^{n \times n}\) and \(b = \{b_i\} \in \mathbb{R}^n\) is said to be a **dynamical network** \((A, b)\) if \(A\) is **stable** and **symmetric**.

A generalization of **reaction-diffusion** systems:

\[
\dot{x}_i = -r_i x_i + \sum_{j=1, j \neq i}^{n} a_{i,j} (x_j - x_i) + b_i u
\]

where \[
\begin{align*}
    r_i &\geq 0 : \text{intensity of reaction (chemical dissolution)} \\
    a_{i,j} = a_{j,i} &\geq 0 : \text{intensity of diffusion between } x_i \text{ and } x_j
\end{align*}
\] are stable, symmetric.
Drawback of Traditional Model Reduction Methods

- Traditional model reduction methods
  - Balanced truncation, Krylov projection, Hankel norm approximation
  - No specific structure in transformation matrix \( P \)

**Drawback:** Network structure (spatial information) is destroyed

\[
\begin{align*}
\text{Given } & (A, b) \\
\text{Reduced model } & (PAP^\dagger, Pb)
\end{align*}
\]

\[
\begin{align*}
P & \in \mathbb{R}^{\Delta \times n} : \text{full matrix} \\
&PAP^\dagger, Pb : \text{dense} \\
&\text{no physical interpretation...}
\end{align*}
\]
State Aggregation Approach

- Model reduction by **aggregating disjoint sets of states** \( \{x[l]\}_{l \in \{1, \ldots, L\}} \)
- **Block-diagonally structured** aggregation matrix \( P = \text{Diag}(p[1], \ldots, p[L]) \)
- Network topology among clusters is to be preserved 😊

\[
\begin{align*}
    x[1] & \in \mathbb{R} \\
    x & = Px \\
    \begin{bmatrix}
        p[1] \\
        p[2]
    \end{bmatrix} & \in \mathbb{R}^{1 \times 2} \\
    \begin{bmatrix}
        p[2]
    \end{bmatrix} & \in \mathbb{R}^{1 \times 4}
\end{align*}
\]

**How to find “reducible” clusters?**

\[
x = P x \iff \begin{bmatrix} x[1] \\ x[2] \end{bmatrix} = \begin{bmatrix} p[1] \\ p[2] \end{bmatrix} \begin{bmatrix} x[1] \\ x[2] \end{bmatrix}
\]

※ For simplicity of notation, suppose the indices are permuted accordingly to clusters
Key Observation to Construct Reducible Clusters

Given \((A, b)\)

Clustering accordingly to behavior

50th order

50th order

50th order

50th order

\[ x \in \mathbb{R}^{50} \]

about 7 trajectories

\[ x \text{ can be aggregated into 7 dimensional variables??} \]

[State behavior under a input signal]

Let \(I_{[u]}\) be the index set of the \(l\)-th cluster.

The \(l\)-th cluster is said to be reducible if

\[ \exists \rho_{i,j} \in \mathbb{R} \text{ s.t. } g_i(t) \equiv \rho_{i,j} g_j(t), \forall i, j \in I_{[u]} \text{ where any input signal } u(t). \]

[Definition] Reducible Cluster

\( l \)-th cluster

\( x_i \) \( x_j \) \( I_{[u]} = \{i, j, \ldots\} \)
Basis Expansion via Positive Tri-diagonalization

Dynamical network \((A, b)\)

\[
x = Ax + bu,
\]

\[
\begin{aligned}
A &= A^T \in \mathbb{R}^{n \times n} \\
b &\in \mathbb{R}^n
\end{aligned}
\]

basis expansion \(g_i(s) = \sum_{j=1}^{n} h_{i,j} g_j(s)\)

Positive tri-diagonal pair \((\mathcal{A}, b)\)

\[
H = \{h_{i,j}\} \in \mathbb{R}^{n \times n}
\]

\[
\mathcal{A} = \begin{bmatrix}
\alpha_1 & \beta_1 & \beta_2 \\
\beta_1 & \alpha_2 & \beta_2 \\
\vdots & \ddots & \ddots \\
\beta_{n-1} & \cdots & \beta_1 & \alpha_n \\
\end{bmatrix}
\]

\[
b = \begin{bmatrix}
\beta_0 \\
0 \\
0 \\
\vdots \\
0
\end{bmatrix}
\]

DC-gain is maximal

\[
\|g_i(s)\|_{\mathcal{H}_\infty} = g_i(0)
\]

\[
\|g_i(s) - \rho_{i,j} g_j(s)\|_{\mathcal{H}_\infty} = 0 \iff \text{row}_i [H \text{diag} (g)] = \rho_{i,j} \text{row}_j [H \text{diag} (g)]
\]

where \(g := -\mathcal{A}^{-1}b = [g_1(0), \ldots, g_n(0)]^T\)
Aggregation of Reducible Clusters

Cluster reducibility: \( \exists \rho_{i,j} \in \mathbb{R} \text{ s.t. } g_i(s) = \rho_{i,j} g_j(s), \forall i, j \in \mathcal{I}_l \)

**[Theorem]** Let \( \mathcal{I}_l \) be the index set of the \( l \)-th cluster.

The \( l \)-th cluster is reducible if and only if

\[ \exists \rho_{i,j} \in \mathbb{R} \text{ s.t. } \text{row}_i \left[ H \text{diag} \left( g \right) \right] = \rho_{i,j} \text{row}_j \left[ H \text{diag} \left( g \right) \right], \forall i, j \in \mathcal{I}_l. \]

Furthermore, \( g(s) = g(s) \) holds for \( p[l] = \frac{p[u]}{\|p[u]\|}, p[l] = \{\rho_{i,j}\}_{j \in \mathcal{I}_l} \in \mathbb{R}^{1 \times |\mathcal{I}_l|}. \)

\[ P: \text{unitary, i.e., } P^T = P^\dagger \]

Elimination of uncontrollable subspace with block-diagonally structured \( P \)

Aggregated model \((PA^TP, Pb)\)
\[ g(s) = P^T \left( sI_\Delta - PA^TP \right)^{-1} Pb \]

Dynamical network \((A, b)\)
\[ g(s) = (sI_n - A)^{-1} b \]
Simple Example

Given $g(s) = (sI_n - A)^{-1}b$

Reducibility:

$\exists \rho_{i,j}$ s.t. $\text{row}_i[H\text{diag}(g)] = \rho_{i,j} \text{row}_j[H\text{diag}(g)]$

where $g = -A^{-1}b$

positive tri-diagonal pair $(\mathcal{A}, b)$

$H\text{diag}(g) =$

\[
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & 1.2 & -0.2 & 0 & 0 \\
0 & 0.6 & 0.4 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\]

aggregation

$x = P^T(sI_\Delta - PAP^T)^{-1}Pb$

$g(s) = P^T(sI_\Delta - PAP^T)^{-1}Pb$

$g(s) = g(s)$ holds

$\iff P^T x \rightarrow x, \forall u, x(0), x(0)$
Relaxation of Reducibility Condition

Dynamical network \((A, b)\)

\(H_{\text{diag}}(g) =
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & 1.2 & -0.2 & 0 & 0 \\
0 & 1.2 & -0.2 & 0 & 0 \\
0 & 0.6 & 0.4 & 0 & 0 \\
0 & 0.6 & 0.4 & 0 & 0
\end{bmatrix}\)

row\(_i\) \([[H_{\text{diag}}(g)]] = \rho_{i,j} \text{ row}\(_j\) [[H_{\text{diag}}(g)]]

similar behavior

\(H_{\text{diag}}(g) =
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & 1.23 & -0.25 & 0.03 & -0.00 \\
0 & 1.17 & -0.14 & -0.03 & 0.01 \\
0 & 0.58 & 0.39 & 0.02 & 0.01 \\
0 & 0.64 & 0.38 & -0.01 & -0.01
\end{bmatrix}\)

row\(_i\) \([[H_{\text{diag}}(g)]] \sim \rho_{i,j} \text{ row}\(_j\) [[H_{\text{diag}}(g)]]\)
Aggregation of Weakly Reducible Clusters

Denote \( \text{row}_i [H \text{diag} (g)] = h_i^g \) and define \( p = \{p_i\} = (-A^{-1} b)^T \in \mathbb{R}^{1 \times n} \).

**[Definition]** \( \theta \)-weakly Reducible Cluster

The cluster \( \mathcal{I}_{[\ell]} \) is \( \theta \)-weakly reducible if

\[
\exists i \in \mathcal{I}_{[\ell]} \text{ s.t. } \left\| h_i^g - \rho_{i,j} h_j^g \right\| \leq \theta, \quad \forall j \in \mathcal{I}_{[\ell]}.
\]

\[
\text{Denote } \quad \rho_{i,j} := p_i / p_j
\]

**[Theorem]** Let \( p_{[\ell]} = \{\rho_{i,j}\}_{j \in \mathcal{I}_{[\ell]}} \in \mathbb{R}^{1 \times |\mathcal{I}_{[\ell]}|} \).

Suppose all clusters are \( \theta \)-weakly reducible, and take \( p_{[\ell]} = \frac{p_{[\ell]}}{\|p_{[\ell]}\|} \in \mathbb{R}^{1 \times |\mathcal{I}_{[\ell]}|} \).

Then, \( g(0) = g(0), \quad \|g(s) - g(s)\|_{H_\infty} \leq \alpha \theta \) hold for an \( \alpha \) determined by \( A \).
Algorithm to Construct Weakly Reducible Cluster Set

Given network

- Give $\theta \in \mathbb{R}_+$, Initialize $\{\mathcal{I}_l\}_{l \in L} = \emptyset$, $L = \emptyset$, $l = 0$

While $\{\mathcal{I}_l\}_{l \in L} \neq \{1, \ldots, n\}$
- $l++$, $L \leftarrow \{L, l\}$
- Choose $i \in \{1, \ldots, n\} \setminus \{\mathcal{I}_l\}_{l \in L}$, Set $\mathcal{I}_l = \{i\}$
- For all $j \in \{1, \ldots, n\} \setminus \{\mathcal{I}_l\}_{l \in L}$, if $(i, j)$ satisfies $(*)$, then $\mathcal{I}_l \leftarrow \{\mathcal{I}_l, j\}$

$\theta$-weak reducibility condition:

$$\left\| h_i^g - \rho_{i,j} h_j^g \right\| \leq \theta \quad \cdots \cdots \quad (*)$$

where $h_i^g := \text{row}_i[H \text{diag}(g)]$ and $\rho_{i,j} = p_i/p_j$
Numerical Example

- Holme-Kim model of 1000 nodes
  - a famous complex network model
  - non-zero edge weight is randomly chosen from \((0, 1]\)

\[
\frac{\|g - \hat{g}\|_{\mathcal{H}_\infty}}{\|g\|_{\mathcal{H}_\infty}} = \begin{cases} 
5.93 \times 10^{-2} \% & (\theta = 1.5) \\
9.09 \times 10^{-2} \% & (\theta = 3.0)
\end{cases}
\]
Remarks on $\mathcal{H}_\infty$-aggregation

- Similar aggregation approach admits to multi-input systems
  - positive tri-diagonalization with respect to each input signal

- Aggregation of positive networks preserves the positivity
  - positivity: non-negativity of off-diagonal entries of $A$ and entries of $b$

- Generalization to positive directed networks is also possible
  - asymmetric $A$ with non-negative off-diagonal entries
  - tri-diagonalization is replaced with Hessenberg transformation
  - however, higher computational costs are to be required
Contents

- Overview of model reduction problem (preliminary)
  - what is model reduction?
  - error evaluation in terms of $\mathcal{H}_\infty/\mathcal{H}_2$-norms

- Network structure preserving model reduction (main)
  - system description of dynamical networks
  - key idea of state aggregation with an $\mathcal{H}_\infty$-error evaluation
  - numerical example

- State aggregation with an $\mathcal{H}_2$-error evaluation (additional)
  - controllability gramian
$\mathcal{H}_2$-aggregation based on Controllability Gramian

$$
\Sigma : \begin{cases} 
\dot{x} = Ax + Bu \\
y = Cx 
\end{cases} \quad \mathcal{H}_2\text{-norm: } \|\Sigma\|_{\mathcal{H}_2} := \left( \int_0^\infty \|Ce^{At}B\|_F^2 \, dt \right)^{\frac{1}{2}}
$$

Positive semi-definite solution $\Phi$ of Lyapunov equation $A\Phi + \Phi A^T + BB^T = 0$

$\Phi :$ controllability gramian

$$
\begin{align*}
\|\Sigma\|_{\mathcal{H}_2} &= \left( \text{trace}(C\Phi C^T) \right)^{\frac{1}{2}} \\
\Phi \text{ is non-singular } &\iff (A, B) \text{ is controllable}
\end{align*}
$$

$\Phi$ relates to $\mathcal{H}_2$-norm & controllability

[Theorem]

The cluster $\mathcal{I}_{\{i\}}$ is reducible if and only if

$$
\exists \rho_{i,j} \in \mathbb{R} \text{ s.t. } \text{row}_i[\Phi^{\frac{1}{2}}] = \rho_{i,j} \text{ row}_j[\Phi^{\frac{1}{2}}], \ \forall i, j \in \mathcal{I}_{\{i\}}.
$$
Aggregation of Weakly Reducible Clusters

Denote $\text{row}_i[\Phi^{1/2}] = \phi_i$ and define $p = \{p_i\} = (-A^{-1}b)^\top \in \mathbb{R}^{1 \times n}$.

[Definition] $\theta$-weakly Reducible Cluster

The cluster $\mathcal{I}[u]$ is $\theta$-weakly reducible if $\exists i \in \mathcal{I}[u]$ s.t. $\|\phi_i - \rho_{i,j} \phi_j\| \leq \theta$, $\forall j \in \mathcal{I}[u]$.

[Theorem] Let $p[u] = \{\rho_{i,j}\}_{j \in \mathcal{I}[u]} \in \mathbb{R}^{1 \times |\mathcal{I}[u]|}$, $\rho_{i,j} := p_i/p_j$.

Suppose all clusters are $\theta$-weakly reducible, and take $p[u] = \frac{p[u]}{\|p[u]\|} \in \mathbb{R}^{1 \times |\mathcal{I}[u]|}$.

Then, $g(0) = g(0)$, $\|g(s) - g(s)\|_{\mathcal{H}_2} \leq \alpha \theta$ hold for a positive constant $\alpha$. 
Remarks on $\mathcal{H}_2$-aggregation

- Lyapnov equation $A\Phi + \Phi A^T + BB^T = 0$ is identical regardless of
  - single- or multi-input systems
  - moreover, regardless of symmetry of $A$

- Generalization to positive directed networks is also possible
  - preserving network topology, stability, and positivity
  - that to systems with arbitrary stable $A$ is difficult
    - not only error evaluation, but also guaranteeing stability of reduced model
Summary

- Model reduction based on state aggregation is proposed
  - preserving network structure
  - aggregation of reducible clusters
    - characterization via positive tri-diagonalization leads to $\mathcal{H}_\infty$-aggregation
    - characterization via controllability gramian leads to $\mathcal{H}_2$-aggregation
  - Aggregation techniques have ability to preserve other properties
    - system positivity, second order (oscillator) structure, etc...

Aggregated model $(PAP^T, Pb)$
$$g(s) = P^T (sI_\Delta - PAP^T)^{-1} Pb$$

Dynamical network $(A, b)$
$$g(s) = (sI_n - A)^{-1} b$$