Sparsity-promoting optimal control of systems with invariances and symmetries

Neil K. Dhingra, Xiaofan Wu and Mihailo R. Jovanović

UNIVERSITY OF MINNESOTA

10th IFAC Symposium on Nonlinear Control Systems,

Monterey, CA, 2016

Standard feedback control

$$
\begin{aligned}
\dot{x} &= A\,x + B\,u + d \\
z &= \begin{bmatrix} I \\ 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ I \end{bmatrix} u
\end{aligned}
$$

Design state feedback gain matrix $u = -Fx$

$$
\dot{x} = (A - B\,F)\,x + d
$$

$\mathcal{H}_2$ and $\mathcal{H}_\infty$ optimal controllers known when F unconstrained

Problem formulation

$$\dot{x} = \big(A - K(v)\big)\, x + d$$

- Control parameter $v \in \mathbb{R}^m$
- $K \in \mathbb{R}^{n \times n}$ linear function of v

Problem formulation

$$\dot{x} \;=\; \big(A \;-\; K(v)\big)\, x \;+\; d$$

- Control parameter $v \in \mathbb{R}^m$
- $K \in \mathbb{R}^{n \times n}$ linear function of v

Objectives:

- Optimize closed-loop performance
- Impose structure on v

Related work

Design of system dynamics

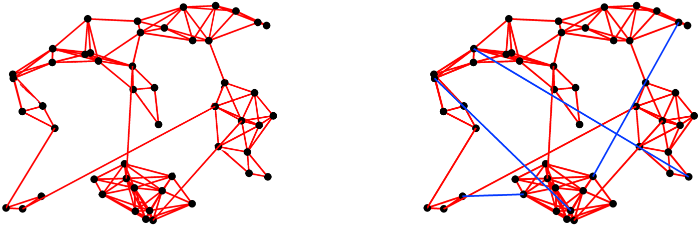
- Convex monotone systems (Rantzer and Bernhardsson '14)
- Combination drug therapy (Jonsson, Rantzer, Matni, and Murray '14)
- Vibrational control of bridges (Nelson, Rajamani, Gastineau, Schultz, and Wojtkiewicz '15)
- $\mathcal{H}_2$ control of symmetric systems (Dhingra and Jovanović '15)
- Decentralized control of positive systems (Dhingra, Colombino, and Jovanović '16)

Sparse feedback synthesis

- Via nonlinear programming (Lin, Fardad, and Jovanovic '13)
- Decentralized control (Schuler, Münz, Allgöwer '14)
- Positive systems (Tanaka and Langbort '11)
- Spatially-invariant state feedback (Zoltowski, Dhingra, and Jovanović '14)
- Quadratically invariant systems (Matni '15)

Examples – Growing consensus networks

$$\dot{x} \;=\; -\big(L \;+\; E\,\mathrm{diag}\{v\}\,E^T\big)\;x \;+\; d$$

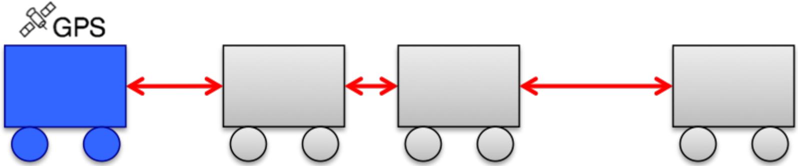


- Nodes x_i determine average via relative information exchange
- L is graph laplacian, E specifies edges, v specifies edge weights

Lin, Fardad, and Jovanovic, Allerton '12
Hassan-Moghaddam and Jovanovic, TCNS '16 (submitted);
also arXiv:1506.03437v2

Examples – Leader selection

$$\dot{x} = -(L + \mathrm{diag}\{v\})\, x + d$$

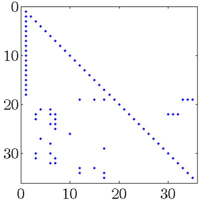


- Some nodes are 'leaders' with access to absolute measurements
- L is graph laplacian, v specifies leaders

Fitch and Leonard, CDC '13
Lin, Fardad, and Jovanovic, TAC '14

Examples – Combination drug therapy

$$\dot{x} = \left(A - \sum_{k=1}^{m} v_k D_k \right) x + d$$



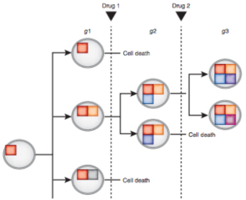


Image credit: Al-Lazikani et al '12

- Mutagen x_i mutates to x_j at rate A_{ji}
- Drug v_k kills x_i at rate $(D_k)_{ii}$

Rantzer and Bernhardsson CDC '14

Jonsson, Rantzer, Matni, and Murray CDC '14

Regularization

minimize $\quad J(v) \quad + \quad \gamma \; g(v)$

$\downarrow$ $\mathcal{H}_2$ **or** $\mathcal{H}_\infty$ **norm**

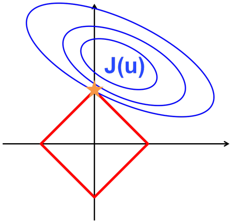
$\downarrow$ **structure-promoting regularizer**

Regularization

$$\text{minimize} \quad J(v) \quad + \quad \gamma \; g(v)$$

$\downarrow$ **$\mathcal{H}_2$ or $\mathcal{H}_\infty$ norm**

$\downarrow$ **structure-promoting regularizer**



- e.g. $g(v) = \sum |v_i|$ for sparsity
- γ specifies importance of sparsity
- $\gamma = 0$ yields dense v

Fardad, Lin, and Jovanovic ACC '11

Lin, Fardad, Jovanovic TAC '13

$\mathcal{H}_2$ norm of system

$$\dot{x} \;=\; \big(A \;-\; K(v)\big)\, x \;+\; d$$

$\mathcal{H}_2$ norm

$$J_2(v) \;:=\; \text{trace}\,(X(v))$$

$X(v)$ is state covariance

$$\big(A \;-\; K(v)\big)\, X \;+\; X\,\big(A \;-\; K(v)\big)^T \;+\; I \;=\; 0$$

In general, $J_2(v)$ nonconvex

$\mathcal{H}_2$ norm of symmetric system

$$\dot{x} \;=\; \big(A_s \;-\; K_s(v)\big)\, x \;+\; d$$

A_s and $K_s(v)$ symmetric $\Longrightarrow$ explicit solution

$$X_s \;=\; -\,\frac{1}{2}\,\big(A_s \;-\; K_s(v)\big)^{-1}$$

$$\boxed{J_{2s}(v) \text{ convex}}$$

$\mathcal{H}_2$ norm of symmetric system

$$\dot{x} = \big(A_s - K_s(v)\big)\, x + d$$

A_s and $K_s(v)$ symmetric $\Longrightarrow$ explicit solution

$$X_s = -\frac{1}{2}\big(A_s - K_s(v)\big)^{-1}$$

$$\boxed{J_{2s}(v) \text{ convex}}$$

SDP characterization

$$\begin{array}{ll} \text{minimize} & \text{trace}\,(X_s) + g(v) \\ \text{subject to} & \begin{bmatrix} X_s & I \\ I & 2(K_s(v) - A_s) \end{bmatrix} \succ 0 \end{array}$$

$\mathcal{H}_\infty$ norm of system

$$\dot{x} \;=\; \big(A \,-\, K(v)\big)\, x \,+\, d$$

$\mathcal{H}_\infty$ norm

$$J_\infty(v) \;:=\; \sup_{\|d\|_{\mathcal{L}_2} \le 1} \frac{\|x\|_{\mathcal{L}_2}}{\|d\|_{\mathcal{L}_2}}$$

Peak of frequency response

$$J_\infty(v) \;=\; \sup_{\omega}\; \sigma_{\max}\big((j\omega I \,-\, (A \,-\, K(v)))^{-1}\big)$$

In general, $J_\infty(v)$ nonconvex and nondifferentiable

$\mathcal{H}_\infty$ norm of symmetric system

$$\dot{x} \;=\; \big(A_s \;-\; K_s(v)\big)\,x \;+\; d$$

A_s and $K_s(v)$ symmetric $\implies$ peak at $\omega = 0$

$$J_{\infty s}(v) \;=\; \sigma_{\max}\left(-(A_s \;-\; K_s(v))^{-1}\right)$$

$$\boxed{J_{\infty s}(v) \text{ convex}}$$

$\mathcal{H}_\infty$ norm of symmetric system

$$\dot{x} \;=\; \big(A_s \;-\; K_s(v)\big)\, x \;+\; d$$

A_s and $K_s(v)$ symmetric $\implies$ peak at $\omega = 0$

$$J_{\infty s}(v) \;=\; \sigma_{\max}\left(-(A_s \;-\; K_s(v))^{-1}\right)$$

$$\boxed{J_{\infty s}(v) \text{ convex}}$$

SDP characterization

$$\begin{array}{ll} \text{minimize} & \sigma_{\max}(X_s) \;+\; g(v) \\ \text{subject to} & \begin{bmatrix} X_s & I \\ I & 2(K_s(v) \;-\; A_s) \end{bmatrix} \;\succ\; 0 \end{array}$$

Exploiting symmetry

Central Idea:

Use symmetric components

$$A_s := \frac{1}{2}(A + A^T), \qquad K_s(v) := \frac{1}{2}(K(v) + K^T(v))$$

to inform design for original system

- Convex characterization
- Relationship between A_s and A
  - Stability
  - Performance
  - Fidelity

Stability

$$A_s \text{ Hurwitz} \implies A \text{ Hurwitz}$$

Not a necessary condition

Implication

Designing v for A_s guarantees stability

$$\underbrace{\dot{x} \; = \; \big(A_s \; - \; K_s(v)\big)\, x \; + \; d}_{\text{stable}} \quad \implies \quad \underbrace{\dot{x} \; = \; \big(A \; - \; K(v)\big)\, x \; + \; d}_{\text{stable}}$$

Upper bound on performance

Symmetric system's $\mathcal{H}_2$ and $\mathcal{H}_\infty$ norms upper bound original

$$\underbrace{\dot{x} \;=\; A\,x \;+\; d}_{\begin{array}{c}\|\cdot\|_2\\ \|\cdot\|_\infty\end{array}} \quad \begin{array}{c}\leq\\ \leq\end{array} \quad \underbrace{\dot{x} \;=\; A_s\,x \;+\; d}_{\begin{array}{c}\|\cdot\|_2\\ \|\cdot\|_\infty\end{array}}$$

Convex upper bound on performance

Implication

A controller v designed for A_s will perform better with A

$$\underbrace{\dot{x} \;=\; \big(A \;-\; K(v)\big)\,x \;+\; d}_{\begin{array}{c}J_2(v)\\ J_\infty(v)\end{array}} \quad \begin{array}{c}\leq\\ \leq\end{array} \quad \underbrace{\dot{x} \;=\; \big(A_s \;-\; K_s(v)\big)\,x \;+\; d}_{\begin{array}{c}J_{2s}(v)\\ J_{\infty s}(v)\end{array}}$$

Small perturbations

- Perturb system dynamics of

$$\dot{x} = A_s x + d$$

by antisymmetric matrix A_a

$$A = A_s + \epsilon A_a$$

- First order correction to both $\mathcal{H}_2$ and $\mathcal{H}_\infty$ norms is 0

Implication

A_s is a very good approximation when A is close to symmetric

$$\begin{aligned} J_2(v) &= J_{2s}(v) + O(\epsilon^2) \\ J_\infty(v) &= J_{\infty s}(v) + O(\epsilon^2) \end{aligned}$$

Design procedure

- Form $A_s = \frac{1}{2}(A + A^T)$ and $K_s(v) = \frac{1}{2}(K(v) + K(v)^T)$

- Solve convex symmetric design problem

$$
\begin{aligned}
v_2^* &= \text{argmin} \quad J_{2s}(v) \; + \; \gamma \, g(v) \\
v_\infty^* &= \text{argmin} \quad J_{\infty s}(v) \; + \; \gamma \, g(v)
\end{aligned}
$$

- Implement v^* on original system A

Design procedure

- Form $A_s = \frac{1}{2}(A + A^T)$ and $K_s(v) = \frac{1}{2}(K(v) + K(v)^T)$

- Solve convex symmetric design problem

$$
\begin{aligned}
v_2^* &= \operatorname{argmin} \quad J_{2s}(v) \; + \; \gamma \, g(v) \\
v_\infty^* &= \operatorname{argmin} \quad J_{\infty s}(v) \; + \; \gamma \, g(v)
\end{aligned}
$$

- Implement v^* on original system A

- Note: can jointly minimize/constrain $\mathcal{H}_2$ and $\mathcal{H}_\infty$ performance

Undirected edge addition in directed consensus networks

$$\dot{x} \;=\; -\left(L \;+\; E\,\mathrm{diag}\{v\}\,E^T\right)\,x \;+\; d$$

Regularized problem

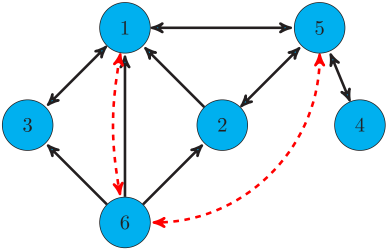
$$\text{minimize} \quad J_s(v) \;+\; v^T v \;+\; \gamma \sum |v_i|$$

Undirected edge addition in directed consensus networks

$$\dot{x} \;=\; -\left(L \;+\; E\,\mathrm{diag}\{v\}\,E^T\right)\,x \;+\; d$$

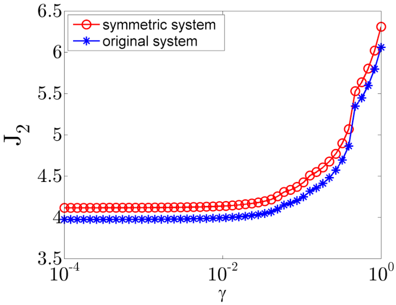
Regularized problem

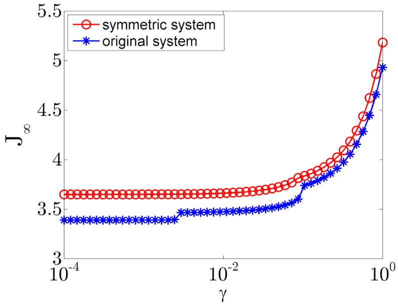
$$\text{minimize} \quad J_s(v) \;+\; v^T v \;+\; \gamma \sum |v_i|$$



Edges added for $\gamma = 1$

Undirected edge addition in directed consensus networks





Closed-loop performance with respect to γ

Computational advantage from invariances

SDPs scale poorly with constraint size

$$\begin{bmatrix} X & I \\ I & -A \end{bmatrix} \succeq 0$$

Computational advantage from invariances

SDPs scale poorly with constraint size

$$\begin{bmatrix} X & I \\ I & -A \end{bmatrix} \succeq 0$$

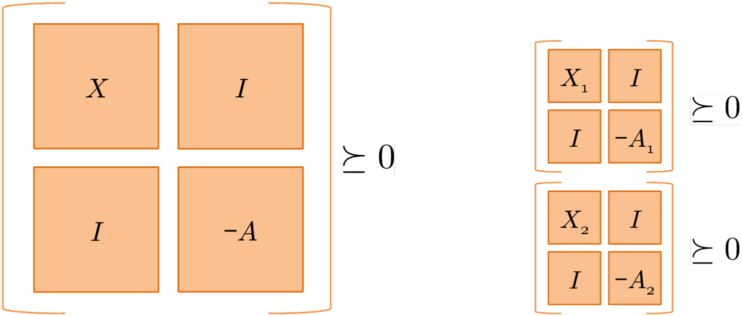
Jointly block-diagonalize A and K

$$PAP^T = \text{blkdiag}\{\hat{A}_i\}, \qquad PK(v)P^T = \text{blkdiag}\{\hat{K}_i(v)\}$$

E.g. spatially-invariant control of spatially-invariant PDEs

Computational advantage from invariances

SDPs scale poorly with constraint size



Jointly block-diagonalize A and K $\implies$ many small SDP constraints

$$PAP^T = \text{blkdiag}\{\hat{A}_i\}, \qquad PK(v)P^T = \text{blkdiag}\{\hat{K}_i(v)\}$$

E.g. spatially-invariant control of spatially-invariant PDEs

Swift-Hohenberg equation

$$\partial_t \psi(t,x) \;=\; \beta\, \psi(t,x) \,-\, (1+\partial_{xx})^2\, \psi(t,x) \,-\, v(x)\psi(t,x)$$

Swift-Hohenberg equation

$$\partial_t \psi(t,x) \;=\; \beta\,\psi(t,x) \,-\, (1+\partial_{xx})^2\,\psi(t,x) \,-\, v(x)\psi(t,x)$$

Discretized at N points $\implies \psi \in \mathbb{R}^N$

$$\dot{\psi} \;=\; (\beta I \,-\, (I+D^2)^2 \,-\, V)\psi$$

Swift-Hohenberg equation

$$\partial_t \psi(t,x) \;=\; \beta\,\psi(t,x) \,-\, (1+\partial_{xx})^2\,\psi(t,x) \,-\, v(x)\psi(t,x)$$

Discretized at N points $\implies \psi \in \mathbb{R}^N$

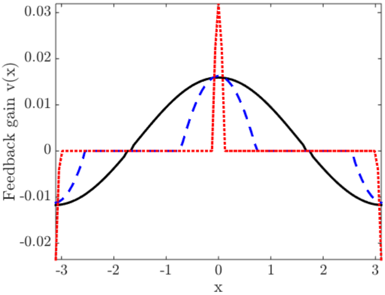
$$\dot{\psi} \;=\; (\beta I \,-\, (I+D^2)^2 \,-\, V)\psi$$

Spatial DFT $\implies N$ decoupled $\psi_\kappa \in \mathbb{R}$

$$\dot{\hat{\psi}}_\kappa \;=\; (\beta-(1-\kappa^2)^2 \,-\, \hat{v}_\kappa)\hat{\psi}_\kappa$$

Swift-Hohenberg equation

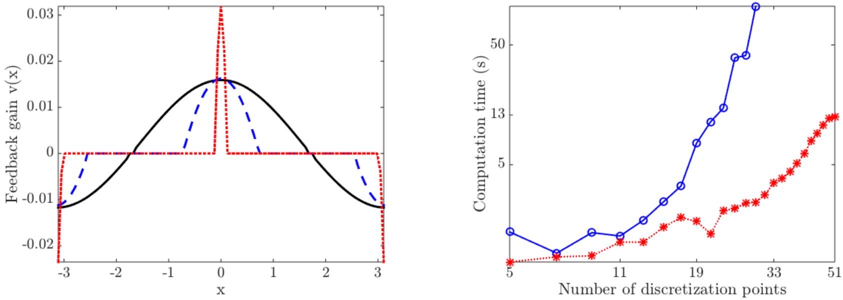
$$\text{minimize} \quad J_2(V) \; + \; \|V\|_F^2 \; + \; \gamma \sum |V_{ij}|$$



Left: $v(x)$ with $N = 101$ for $\gamma = \mathbf{0}$, $\gamma = \mathbf{0.1}$ and $\gamma = \mathbf{10}$

Swift-Hohenberg equation

$$\text{minimize} \quad J_2(V) \; + \; \|V\|_F^2 \; + \; \gamma \sum |V_{ij}|$$



Left: $v(x)$ with $N = 101$ for $\gamma = \mathbf{0}$, $\gamma = \mathbf{0.1}$ and $\gamma = \mathbf{10}$

Right: Computation time in V and $\hat{v}$ coordinates

Conclusions

Regularized design

- Exploit convex characterization using symmetric component
- Implement on original system
  - Stability guarantee
  - Upper bound on performance
  - Good approximation when 'mostly' symmetric

Exploiting invariance

- A and K jointly block-diagonalizable
- Linear scaling with number of SDP blocks

Acknowledgements

Support:

- NSF ECCS-1407958
- University of Minnesota Doctoral Dissertation Fellowship
- NASA JPFP Fellowship
- MnDRIVE Graduate Scholars Program