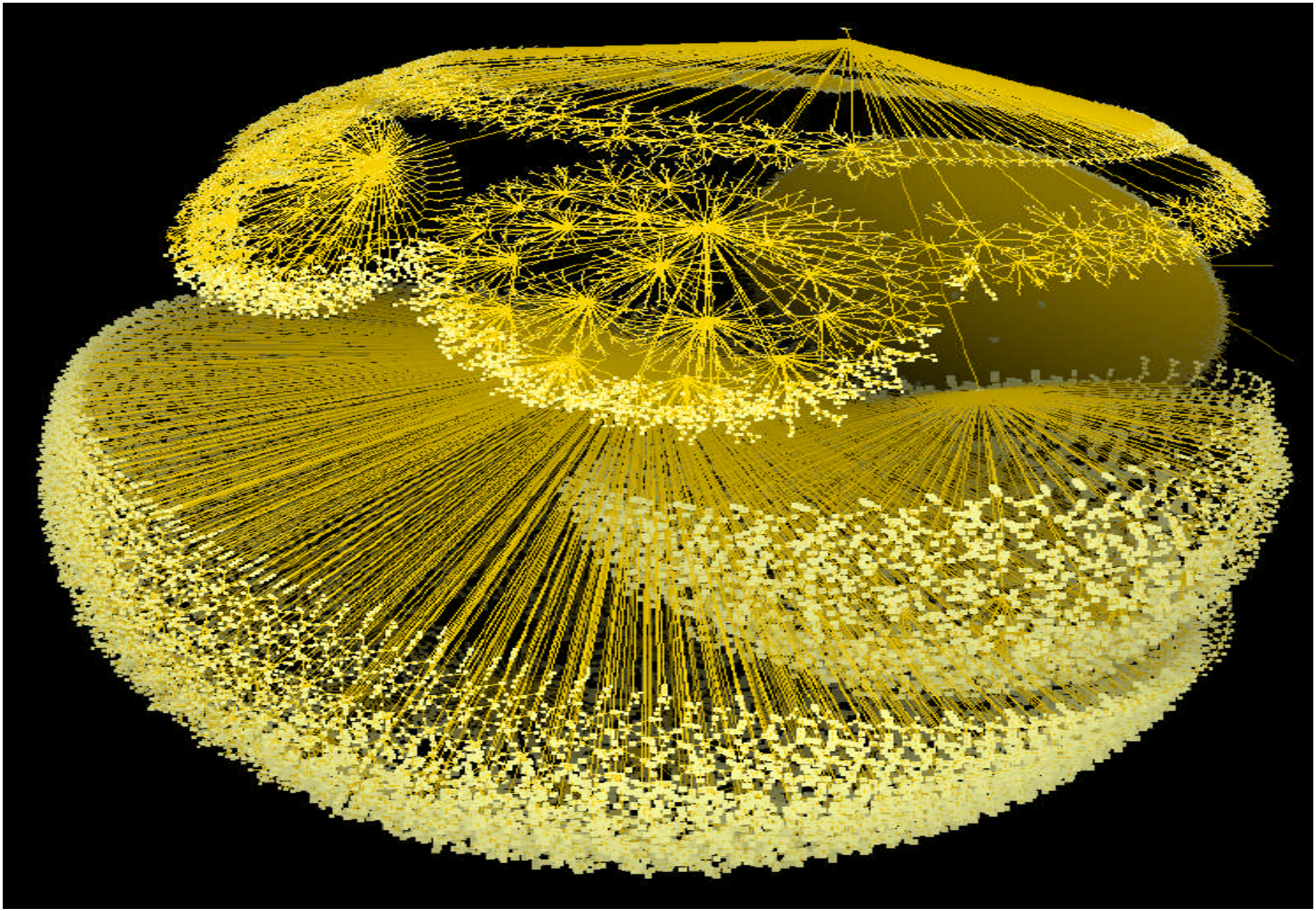
Optimal Synchronization and Games on Graphs







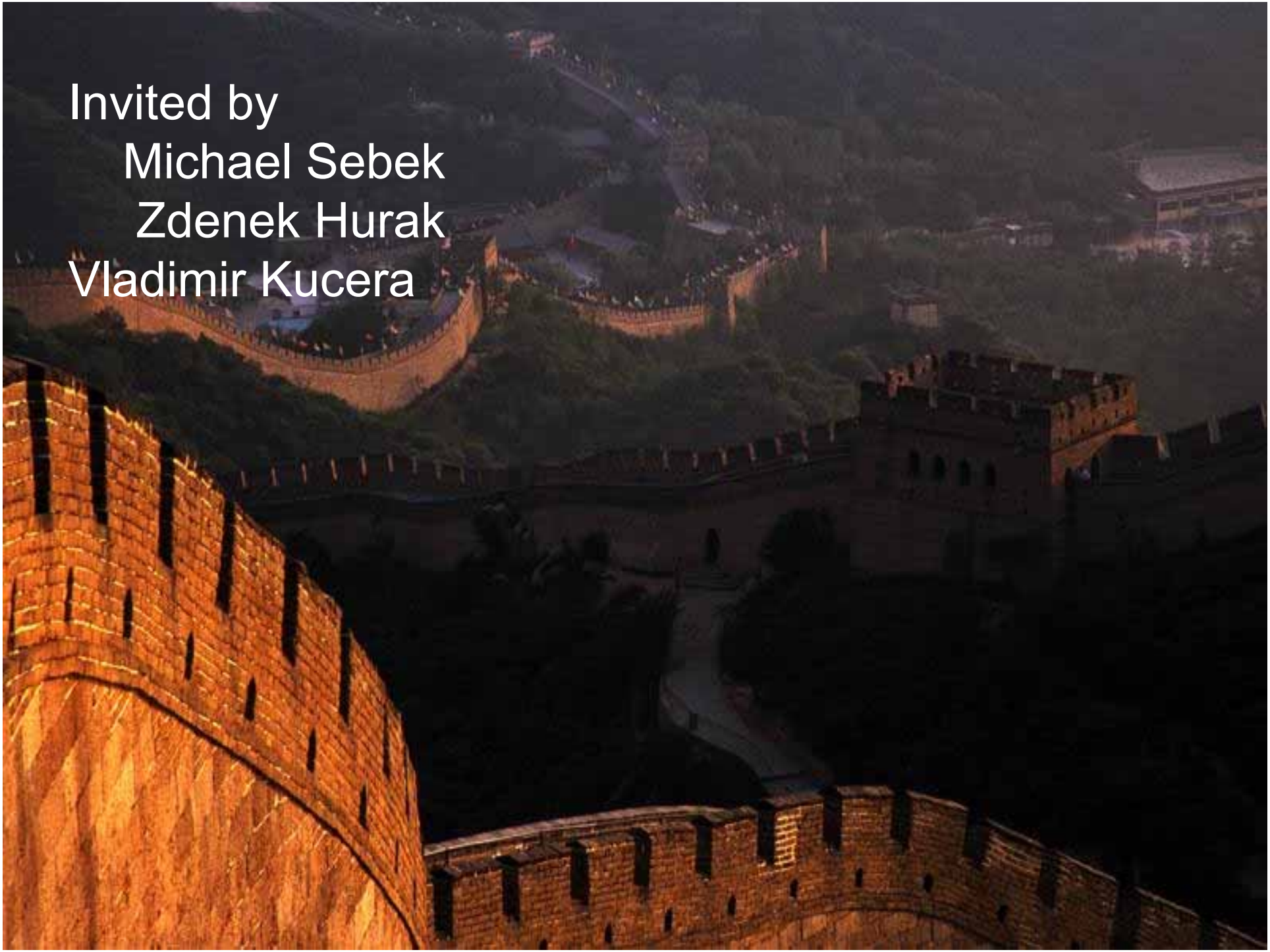
F.L. Lewis
UTA Research Institute (UTARI)
The University of Texas at Arlington

Supported by :
NSF - PAUL WERBOS
ARO, AFOSR, ONR
NNSF of China
China Project 111 at NEU

Cooperative Control: Optimality, Games, and Reinforcement Learning on Communication Graphs



http://www.uta.edu/utari/acs


Invited by
Michael Sebek
Zdenek Hurak
Vladimir Kucera

Kung Tz 500 BC

Confucius

孔子

Tian xia da tong
Harmony under heaven

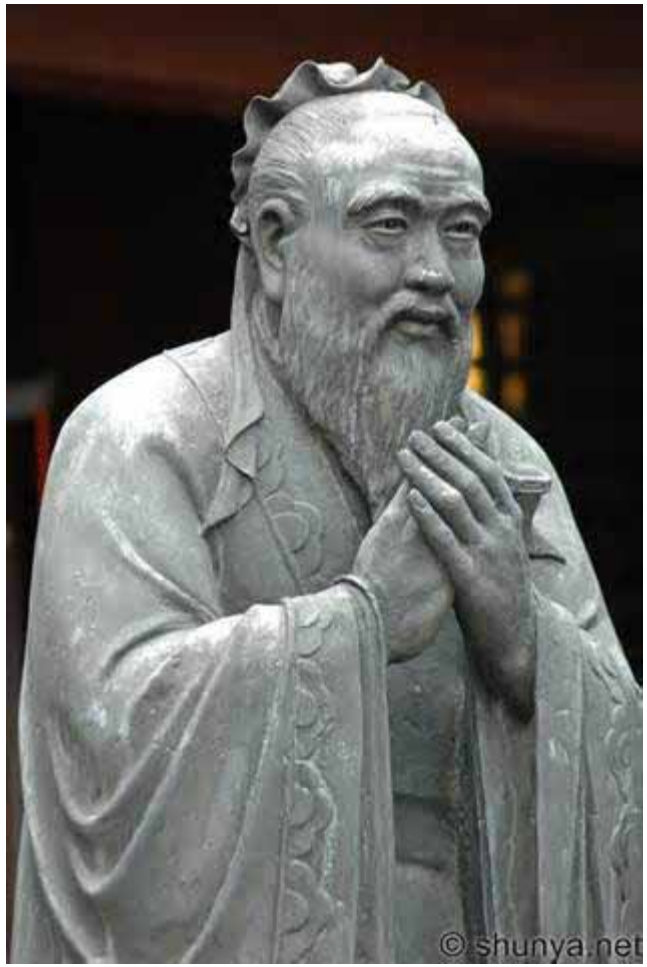
Archery
Chariot driving

Music
Rites and Rituals

Poetry
Mathematics



Man's relations to
- Family
- Friends
- Society
- Nation
- Emperor
- Ancestors

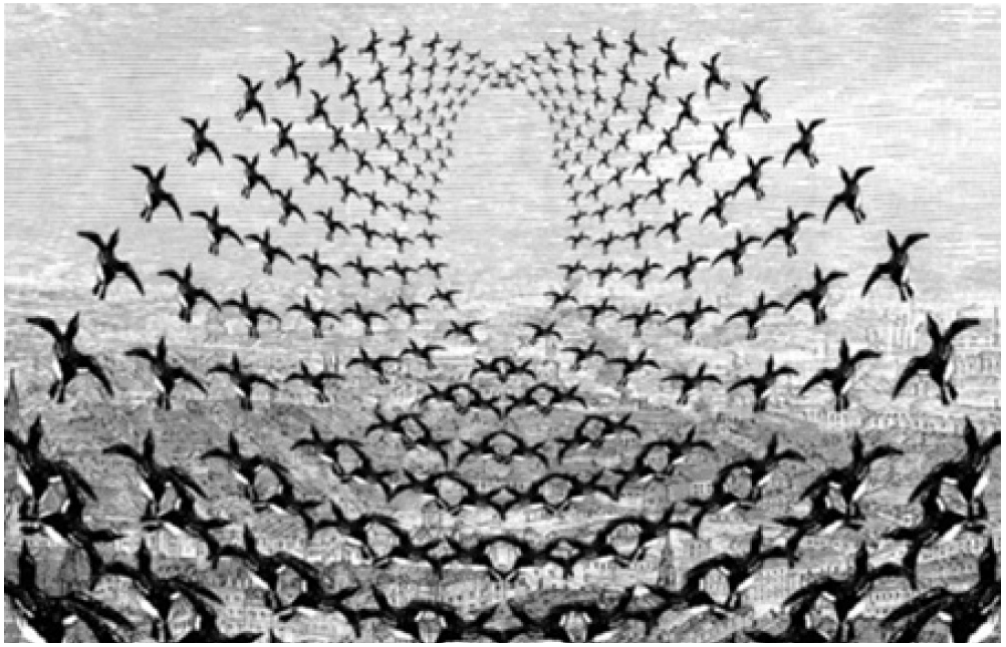


Synchronized Motion of Biological Groups



Fish school

Birds flock





Locusts swarm

Fireflies synchronize



F.L. Lewis, H. Zhang, A. Das, K. Hengster-Movric, *Cooperative Control of Multi-Agent Systems: Optimal Design and Adaptive Control*, Springer-Verlag, 2013



Key Point

Lyapunov Functions and Performance Indices
Must depend on graph topology

H. Zhang, F.L. Lewis, and Z. Qu, "Lyapunov, Adaptive, and Optimal Design Techniques for Cooperative Systems on Directed Communication Graphs," IEEE Trans. Industrial Electronics, vol. 59, no. 7, pp. 3026-3041, July 2012.

Hongwei Zhang, F.L. Lewis, and Abhijit Das
"Optimal design for synchronization of cooperative systems: state feedback, observer and output feedback," IEEE Trans. Automatic Control, vol. 56, no. 8, pp. 1948-1952, August 2011.

Outline

Cooperative Control

Locally Optimal Design and Synchronization

Globally Optimal Design for Collective Motion

Output Feedback H-infinity Cooperative Control

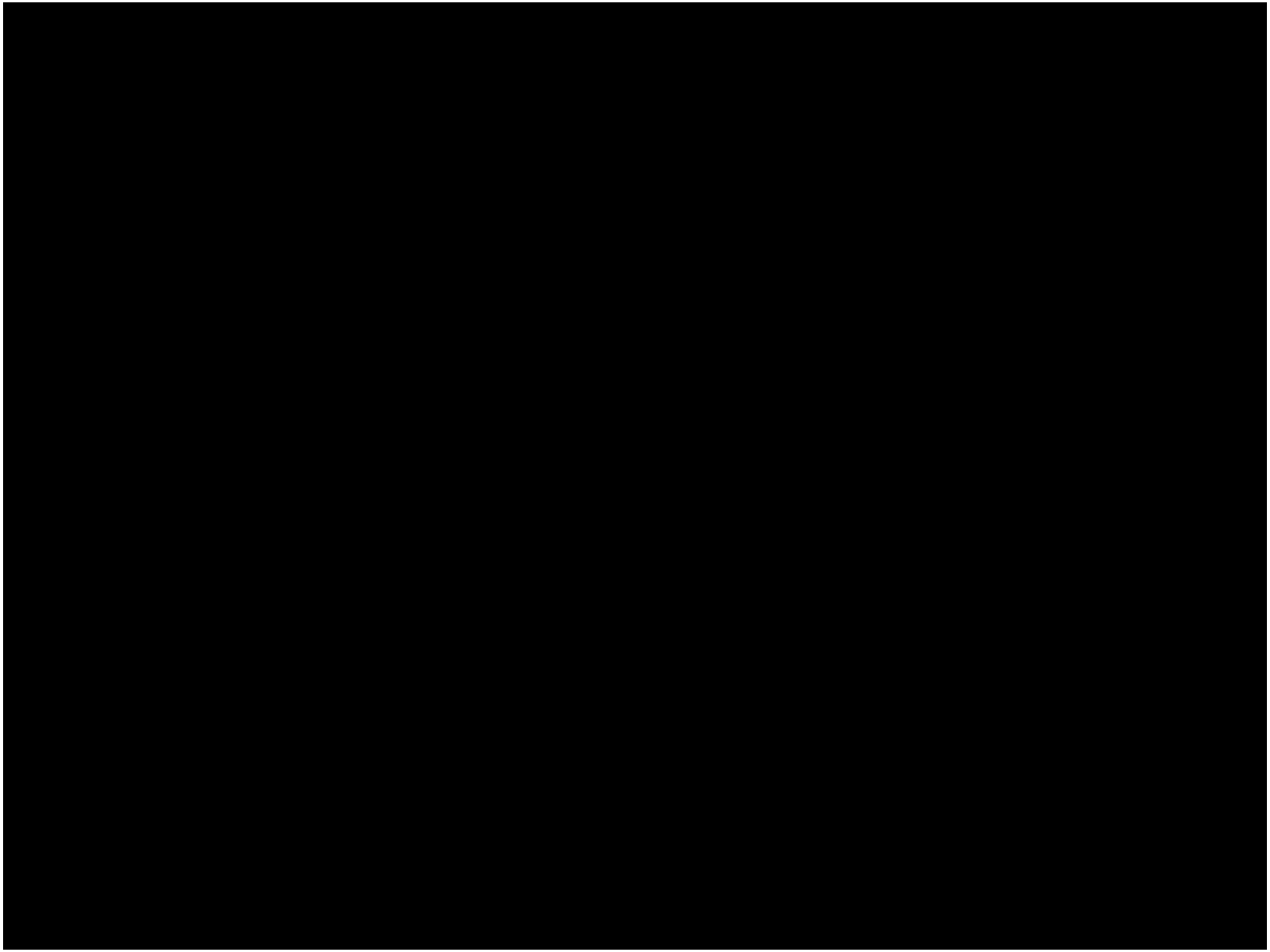
Multi-player Games on Communication Graphs

Reinforcement Learning for Game Solutions



The Power of Synchronization

Coupled Oscillators
Diurnal Rhythm



Stability vs. Optimality of Cooperative Control

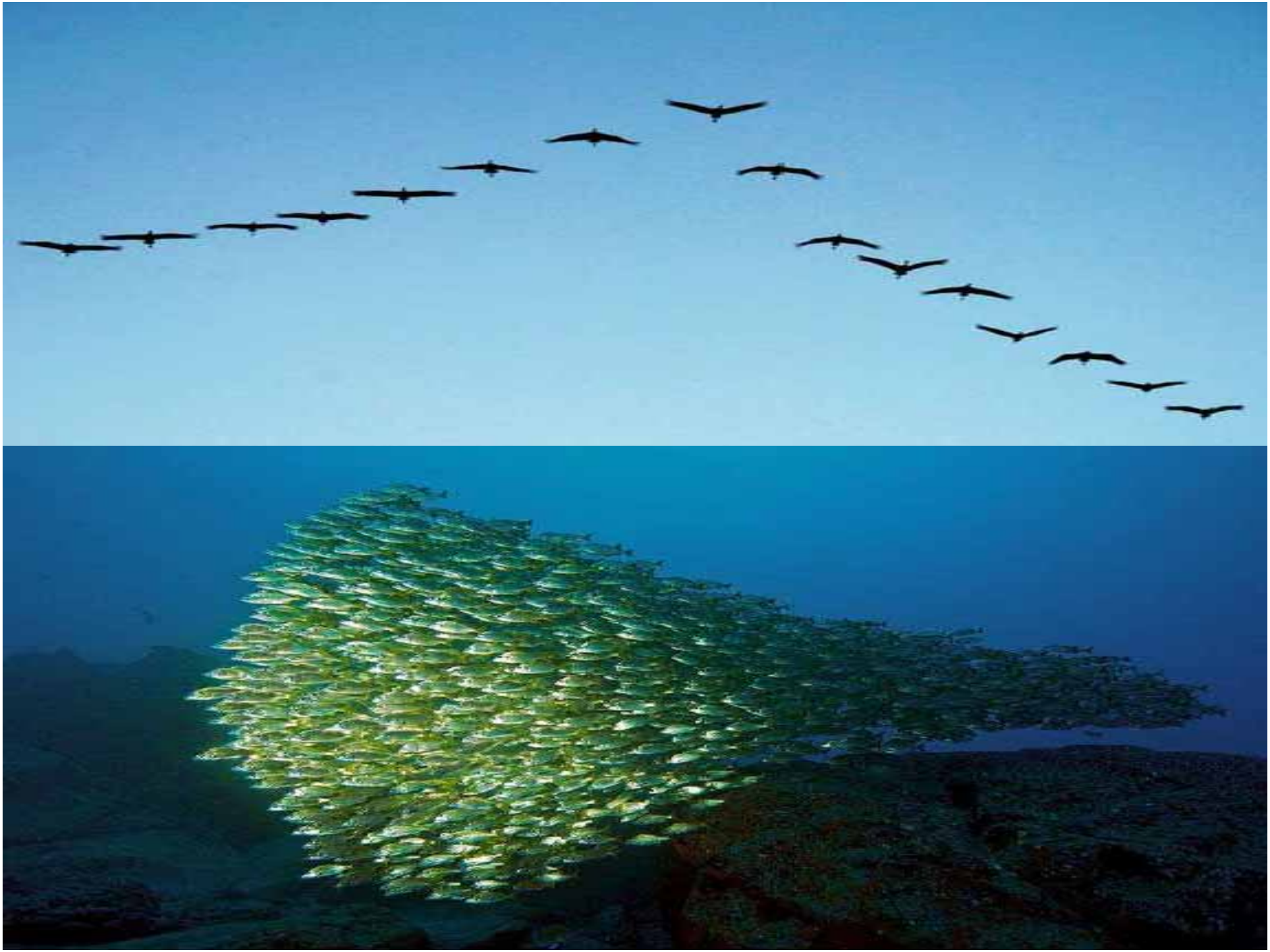
Outline

- A. Stable Design for Synchronization of Cooperative Systems
- B. Global Optimal Design for Collective Group Motion

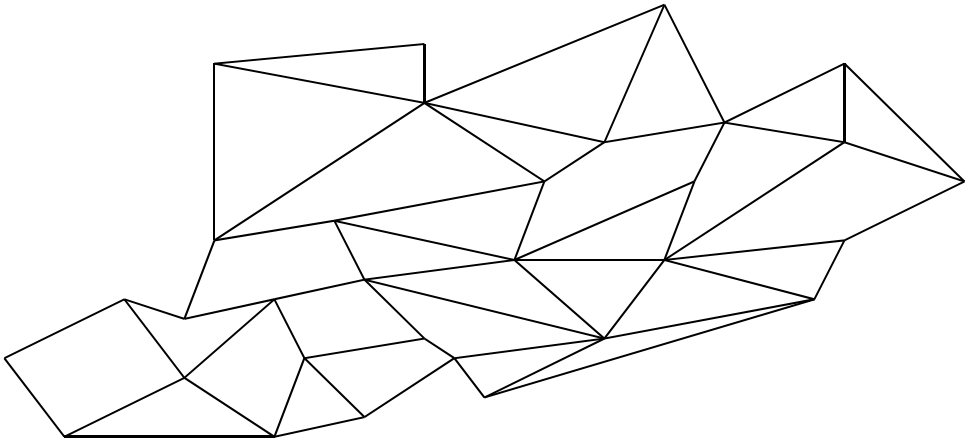
Issues: For cooperative control on graphs -

Local stability of each agent is NOT the same as stable synchronization of the team

Local optimality of each agent is NOT the same a global optimality of the team



Communication Graph

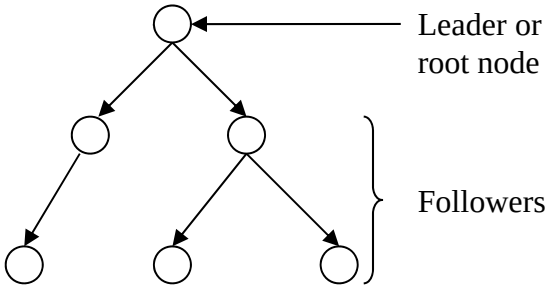


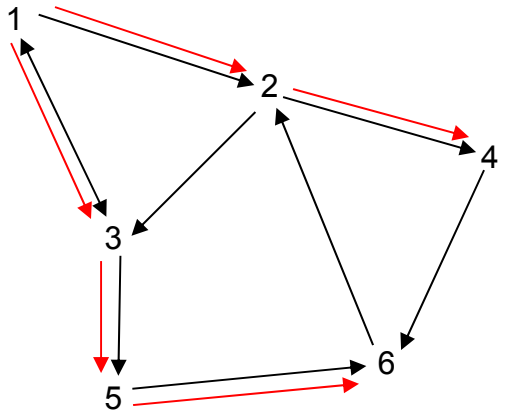
Strongly connected if for all nodes *i* and *j* there is a path from *i* to *j*.

Diameter= length of longest path between two nodes

Volume = sum of in-degrees $Vol = \sum_{i=1}^{N} d_i$

Tree- every node has in-degree=1



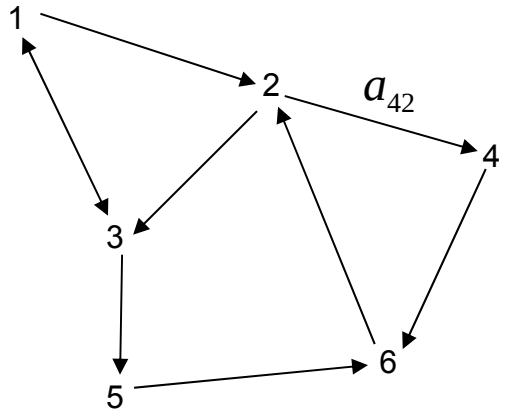


Spanning tree
Root node

Communication Graph

(V,E)

N nodes



Adjacency matrix

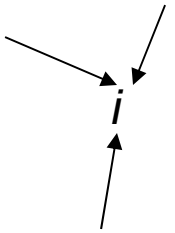
$$A = [a_{ij}]$$

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}$$

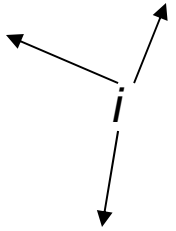
$$a_{ij} > 0 \; if \; (v_j, v_i) \in E$$

$$if \;\; j \in N_i$$

$d_i = \sum_{j=1}^{N} a_{ij}$ Row sum= in-degree N_i In-neighbors of node i



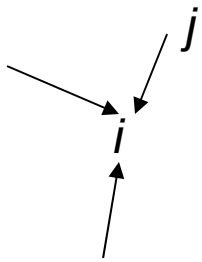
$d_i^o = \sum_{j=1}^{N} a_{ji}$ Col sum= out-degree N_o Out-neighbors of node i



John Baras - social standing

Dynamic Graph- the Distributed Structure of Control

Each node has an associated state $\dot{x}_i = u_i$



Standard local voting protocol $u_i = \sum_{j \in N_i} a_{ij}(x_j - x_i)$

$$u_i = -x_i \sum_{j \in N_i} a_{ij} + \sum_{j \in N_i} a_{ij} x_j = -d_i x_i + \begin{bmatrix} a_{i1} & \cdots & a_{iN} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}$$

x

$$u = \begin{bmatrix} u_1 \\ \vdots \\ u_N \end{bmatrix} \qquad D = \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_N \end{bmatrix} \qquad A = [a_{ij}]$$

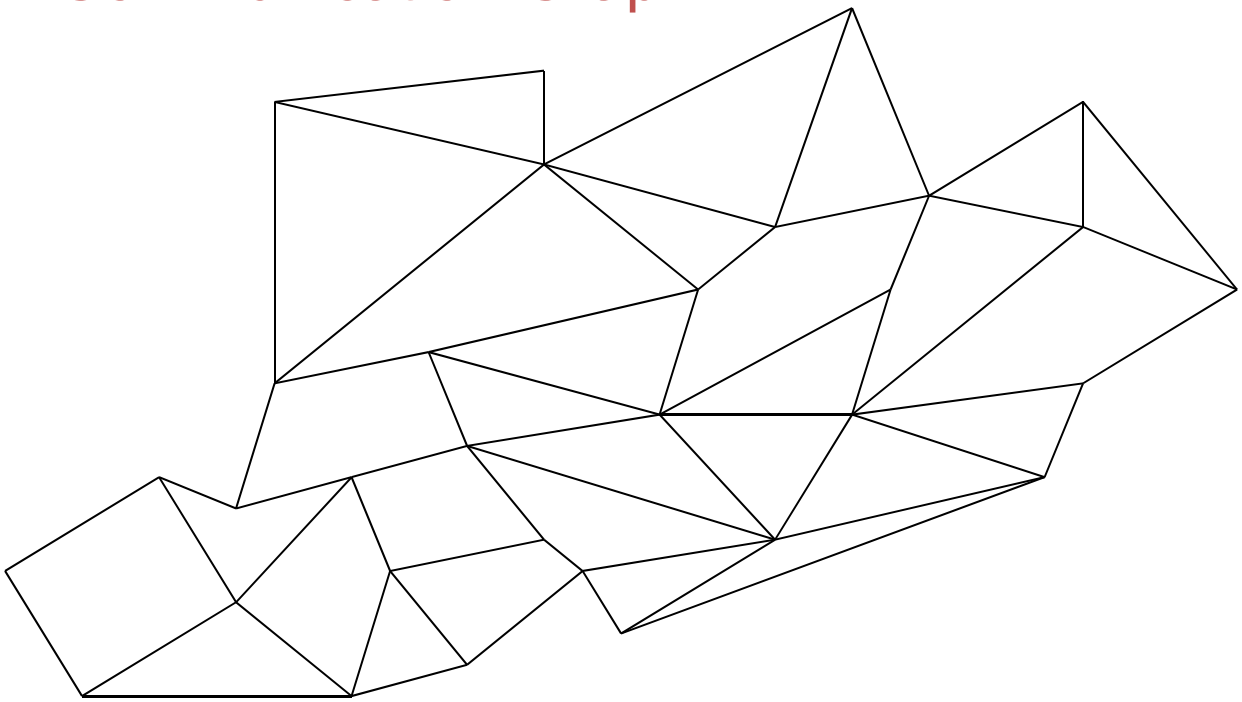
$$u = -Dx + Ax = -(D - A)x = -Lx$$

L=D-A = graph Laplacian matrix

$\dot{x} = -Lx$ Closed-loop dynamics

If x is an n-vector then $\dot{x} = -(L \otimes I_n)x$

Communication Graph



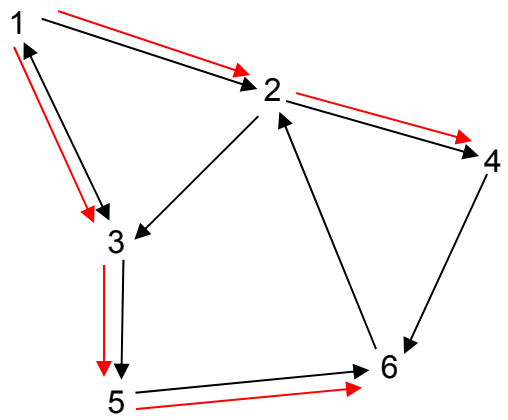
G=(V,E)

N nodes

State at node *i* is $x_i(t)$

Synchronization problem

$$x_i(t) - x_j(t) \to 0$$



Laplacian matrix
L=D-A

Theorem. Graph contains a spanning tree iff e-val of *L* at $\lambda_1 = 0$ is simple.

Then $\lambda_2 > 0$

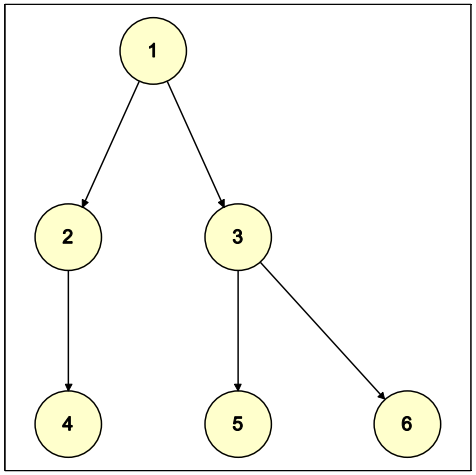
Then *-L* has one e-val at zero and all the rest stable

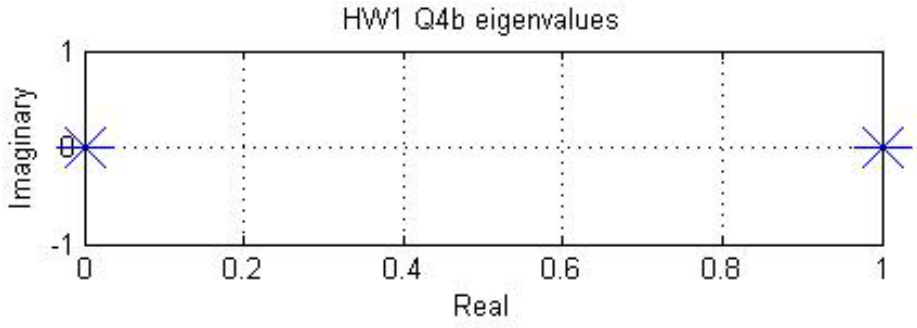
Then, all states synchronize using the local voting protocol

Graph strongly connected implies exists a spanning tree

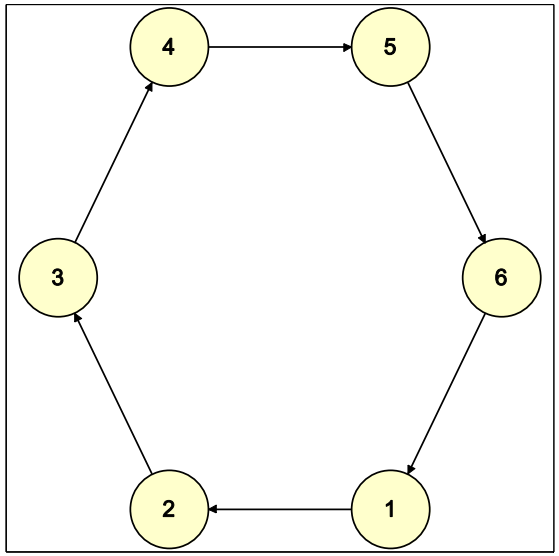
Graph Eigenvalues for Different Communication Topologies

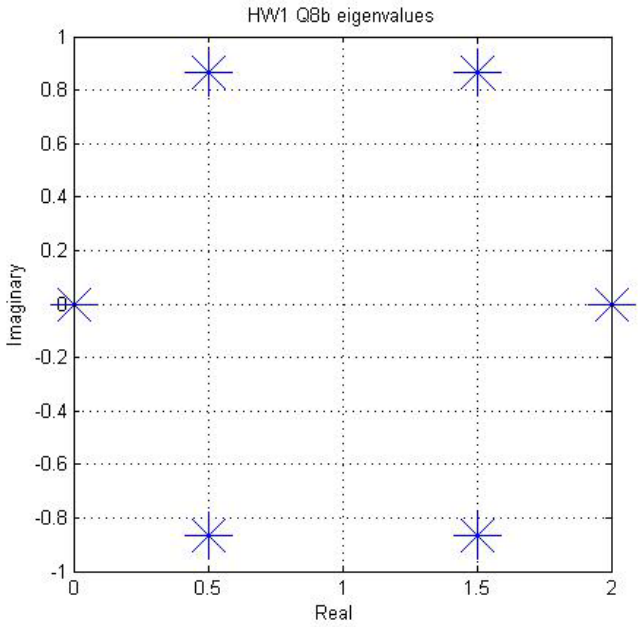
Directed Tree-
Chain of
command





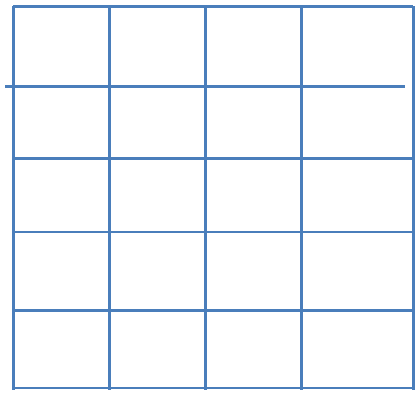
Directed Ring-
Gossip network
OSCILLATIONS



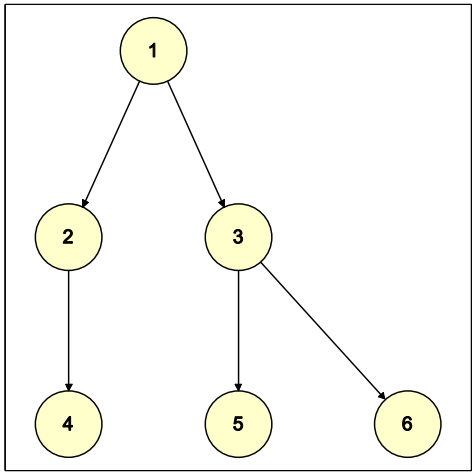


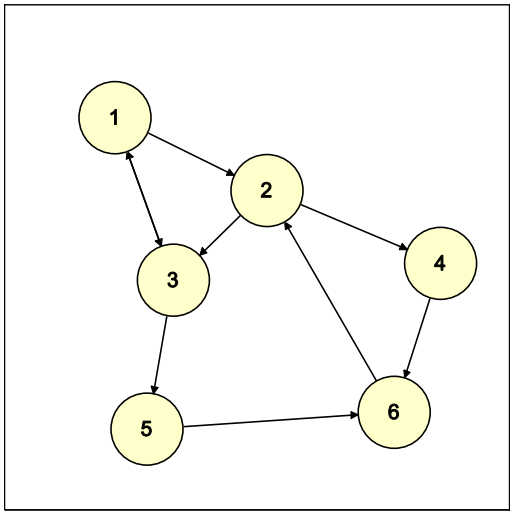
Synchronization on Good Graphs

Chris Elliott fast video



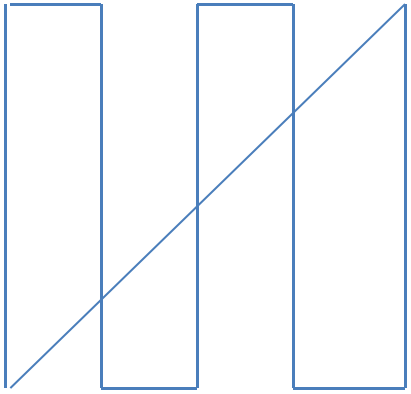
Mesh graph
4 neighbors





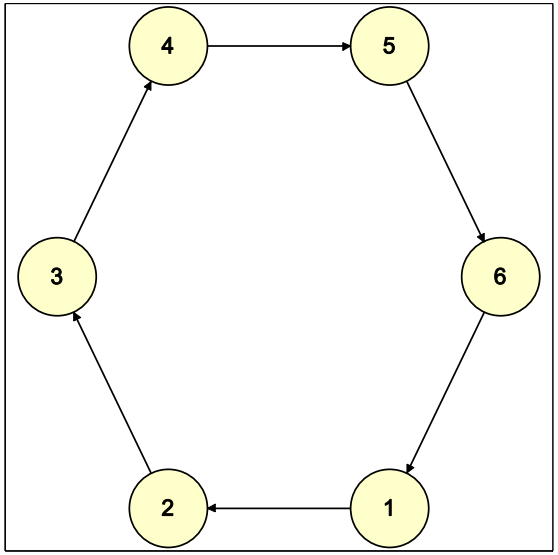
Synchronization on Gossip Rings

Chris Elliott weird video



Ring graph
or cycle

10 nodes



Locally Optimal Design and Synchronization



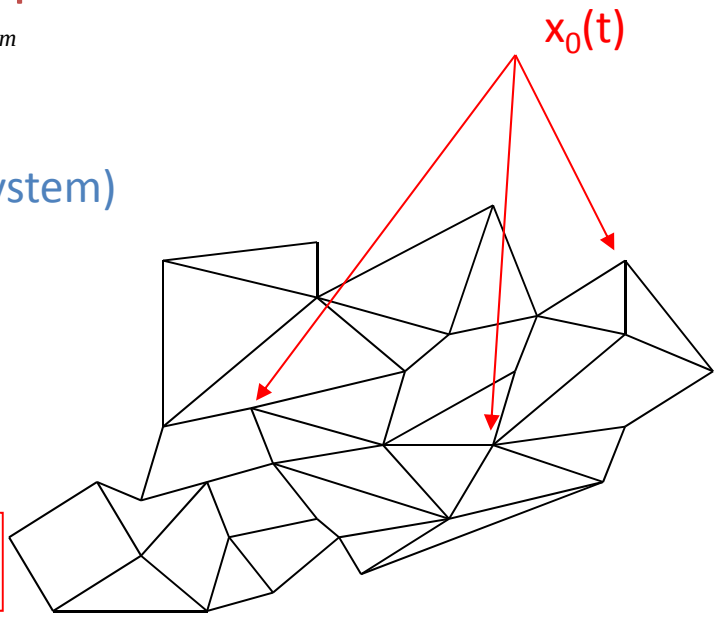
A. State Feedback Design for Cooperative Systems on Graphs

Cooperative Regulator vs. Cooperative Tracker problem

N nodes with dynamics $\dot{x}_i = Ax_i + Bu_i \quad , \quad x_i \in R^n, \quad u_i \in R^m$

Control node or Command generator $\dot{x}_0 = Ax_0$ (Exosystem)

Synchronization Tracker design problem $x_i(t) \to x_0(t), \forall i$

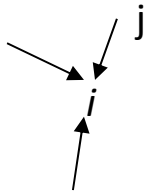


Local neighborhood tracking error

$$\varepsilon_i = \sum_{j \in N_i} e_{ij}(x_j - x_i) \quad + g_i(x_0 - x_i)$$

Ron Chen- pinning control

Lihua Xie & You Keyou - error



$$u_i = cK\varepsilon_i = cK\left(\sum_{j \in N_i} e_{ij}(x_j - x_i) \quad + g_i(x_0 - x_i) \right)$$

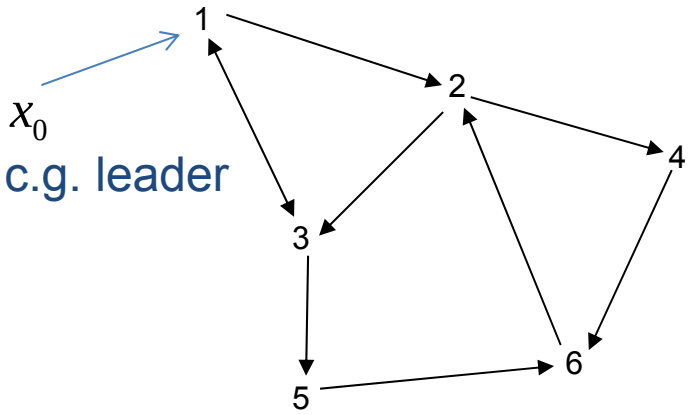
Agent Dynamics and Local Feedback Design

$$\dot{x}_i = Ax_i + Bu_i$$

$$u_i = -Kx_i$$

$$A = \begin{bmatrix} -2 & -1 \\ 2 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad K = \begin{bmatrix} 0.5 & -0.5 \end{bmatrix}$$

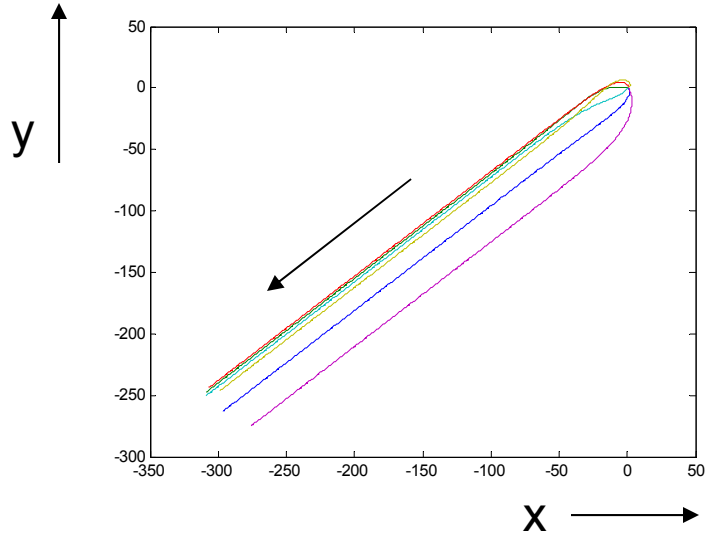
Couple 6 agents with communication graph



Local neighborhood tracking error

$$\varepsilon_i = \sum_{j \in N_i} e_{ij}(x_j - x_i) \quad + g_i(x_0 - x_i)$$

$$u_i = -K\varepsilon_i$$



Nodes synchronize to consensus heading

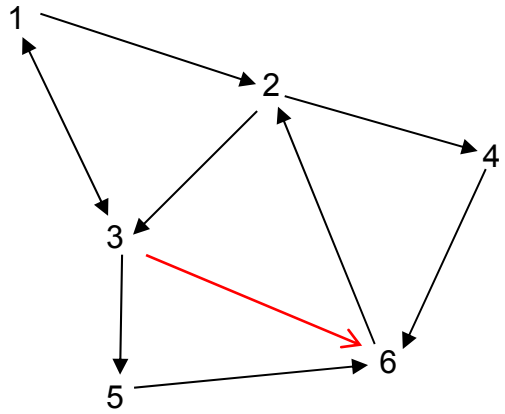
Agent Dynamics and local Feedback design

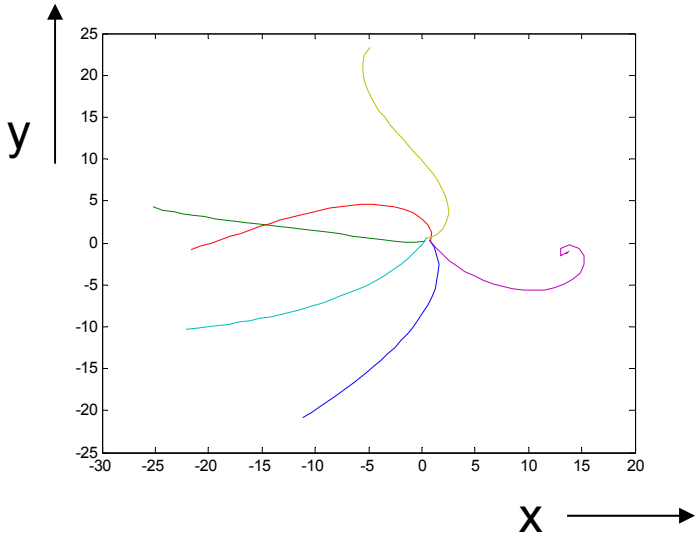
$$\dot{x}_i = Ax_i + Bu_i$$

$$u_i = -Kx_i$$

$$A = \begin{bmatrix} -2 & -1 \\ 2 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad K = \begin{bmatrix} 0.5 & -0.5 \end{bmatrix}$$

ADD another comm. Link- more information flow





Local neighborhood tracking error

$$\varepsilon_i = \sum_{j \in N_i} e_{ij}(x_j - x_i) \quad + g_i(x_0 - x_i)$$

$$u_i = K\varepsilon_i$$

Causes Unstable Formation!

WHY?

A. STABLE DESIGN FOR COOPERATIVE CONTROL ON GRAPHS

We want Design Freedom that overcomes graph topology constraints

Decouple Control Design from Graph Topology constraints

Guaranteed synchronization for general Directed graphs

Guaranteed stability for continuous-time multi-agent systems on graphs -

Hongwei Zhang, F.L. Lewis, and Abhijit Das
"Optimal design for synchronization of cooperative systems: state feedback, observer and output feedback"
IEEE Trans. Automatic Control, vol. 56, no. 8, pp. 1948-1952, August 2011.

We are interested in Fixed graph topologies to explore
relations between classical system theory and communication graph effects

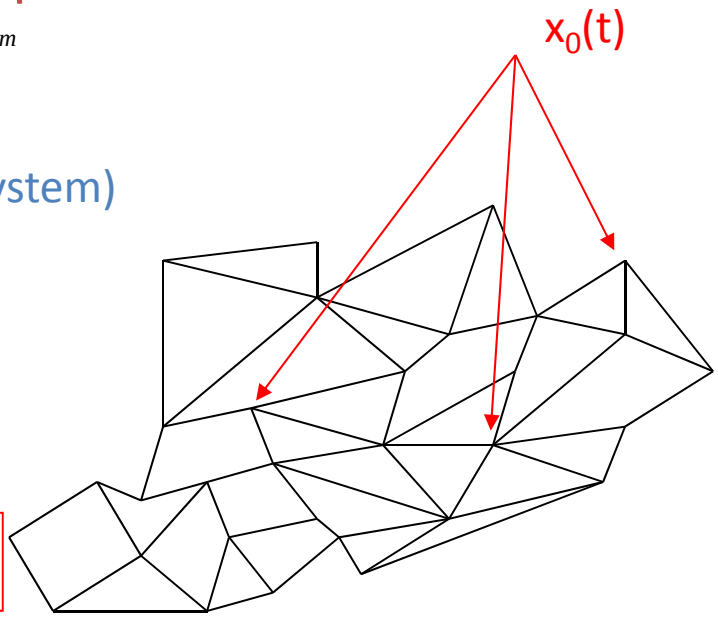
A. State Feedback Design for Cooperative Systems on Graphs

Cooperative Regulator vs. Cooperative Tracker problem

N nodes with dynamics $\dot{x}_i = Ax_i + Bu_i \quad , \quad x_i \in R^n, \quad u_i \in R^m$

Control node or Command generator $\dot{x}_0 = Ax_0$ (Exosystem)

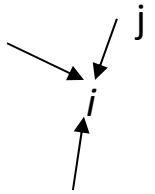
Synchronization Tracker design problem $x_i(t) \to x_0(t), \forall i$



Local neighborhood tracking error

$$\varepsilon_i = \sum_{j \in N_i} e_{ij}(x_j - x_i) \quad + g_i(x_0 - x_i)$$

Ron Chen- pinning control Lihua Xie & You Keyou - error



L=D-A

Overall error vector $e = -\left((L+G)\otimes I_n\right)(x - \underline{x}_0) = -\left((L+G)\otimes I_n\right)\delta$ = Local quantity

where $e = \left[\varepsilon_1^T \quad \varepsilon_2^T \quad \cdots \quad \varepsilon_N^T\right]^T \in R^{nN}, \quad \underline{x}_0 = \underline{I}x_0 \in R^{nN}, \quad \underline{I} = \underline{1} \otimes I_n \in R^{nN \times n}$

Consensus or synchronization error $\delta = \left(x - \underline{x}_0\right) \in R^{nN}$ = Global quantity

Local Neighborhood Tracking Error

$$e = -\left((L+G)\otimes I_n\right)\left(x - \underline{x}_0\right) = -\left((L+G)\otimes I_n\right)\delta$$

Local quantity

Global quantity

Lemma 1. Let the graph be strongly connected and $G \neq 0$. Then

$$\|\delta\| \leq \|e\| / \underline{\sigma}(L+G) \tag{7}$$

with $\underline{\sigma}(L+G)$ the minimum singular value of $(L+G)$, and $e(t) \equiv 0$ if and only if the nodes synchronize, that is

$$x(t) = \underline{I}x_0(t) \tag{8}$$

■

Local control objectives imply global performance

Coop. nbhd SVFB

$$u_i = cK\varepsilon_i = cK\left(\sum_{j\in N_i} e_{ij}(x_j - x_i) \quad + g_i(x_0 - x_i)\right)$$

Closed loop system $\dot{x}_i = Ax_i + Bu_i = Ax_i + cBK\left(\sum_{j\in N_i} e_{ij}(x_j - x_i) \quad + g_i(x_0 - x_i)\right)$

Overall state $x = \begin{bmatrix} x_1^T & x_2^T & \cdots & x_N^T \end{bmatrix}^T \in R^{nN}, \qquad \delta = \left(x - \underline{I}x_0\right)$

Distributed form of control $\boxed{u = -c\left((L+G)\otimes K\right)\delta}$

Overall c.l. dynamics $\dot{x} = \left[(I_N \otimes A) - c(L+G)\otimes BK\right]x + c\left[(L+G)\otimes BK\right]\underline{x}_0$

Global synch. error dynamics $\dot{\delta} = \left[(I_N \otimes A) - c(L+G)\otimes BK\right]\delta$

Graph structure $\otimes$ Control structure

Lemma 2. [6]. Let the graph be strongly connected with at least one pinning gain $g_i > 0$. Let $\lambda_i, i = 1, N$ be the eigenvalues of $(L+G)$. Then the synchronization error dynamics (13) are asymptotically stable (AS) if and only if the matrices

$$A - c\lambda_i BK \qquad (14)$$

are all stable.

Fax and Murray 2004

MIXES UP CONTROL DESIGN AND GRAPH STRUCTURE

The key to global stability and synchronization of the collective is

Locally optimal design for each agent

Theorem 1. Design of SVFB Gain for Cooperative Tracking Stability

Suppose *(A,B)* is stabilizable and the graph is strongly connected with at least one pinning gain $g_i > 0$. Select design matrices $Q = Q^T > 0, R = R^T > 0$. Compute the SVFB gain K according to the linear quadratic regulator (LQR) control algebraic Riccati equation (CARE)

$$0 = A^T P + PA + Q - PBR^{-1}B^T P \tag{15}$$

$$K = R^{-1}B^T P \tag{16}$$

Then the synchronization dynamics (13) are asymptotically stable (AS) for all coupling gains

$$c > \frac{1}{\underline{\lambda}(L+G)} \tag{17}$$

with $\underline{\lambda}(L+G) = \min_i \mathrm{Re}(\lambda_i(L+G))$.

$$u_i = cK\varepsilon_i = cK\left(\sum_{j\in N_i} e_{ij}(x_j - x_i) \quad + g_i(x_0 - x_i)\right)$$

OPTIMAL Design at Each node

Optimal Control 3rd ed Lewis, Vrabie, Syrmos 2012

DECOUPLES CONTROL DESIGN FROM COMMUNICATION GRAPH STRUCTURE

LOCAL OPTIMAL DESIGN Guarantees Global Synchronization

minimizes $J_i = \frac{1}{2}\int_0^\infty (\mathrm{x}_i^T\, Qx_i + u_i^T Ru_i)\, dt$

Hongwei Zhang, F.L. Lewis, and Abhijit Das, "Optimal design for synchronization of cooperative systems: state feedback, observer and output feedback"
IEEE Trans. Automatic Control, vol. 56, no. 8, pp. 1948-1952, August 2011.

S. Tuna, "LQR-based coupling gain for synchronization of linear systems," *Arxiv preprint arXiv:0801.3390*, 2008.

Theorem 1. Design of SVFB Gain for Cooperative Tracking Stability

Suppose (A,B) is stabilizable and the graph is strongly connected with at least one pinning gain $g_i > 0$. Select design matrices $Q = Q^T > 0, R = R^T > 0$. Compute the SVFB gain K according to the linear quadratic regulator (LQR) control algebraic Riccati equation (CARE)

$$0 = A^T P + PA + Q - PBR^{-1}B^T P \tag{15}$$

$$K = R^{-1}B^T P \tag{16}$$

Then the synchronization dynamics (13) are asymptotically stable (AS) for all coupling gains

$$c > \frac{1}{\underline{\lambda}(L+G)} \tag{17}$$

with $\underline{\lambda}(L+G) = \min_i \mathrm{Re}(\lambda_i(L+G))$.

OPTIMAL Design at Each node

Optimal Control 3rd ed Lewis, Vrabie, Syrmos 2012

Proof:

Follows directly from the infinite gain margin robustness property of the LQR [39]. Specifically, note that LQR design renders $(A-BK)$ AS as well as $(A-kBK)$ AS for all gains $k \geq 1$.

Under the stabilizability assumption a solution $P > 0$ to the CARE exists and the gain (16) renders $(A-BK)$ AS. L irreducible and at least one entry of G is positive means that $(L+G)$ is irreducibly diagonally dominant and hence nonsingular, and its eigenvalues $\lambda_i, i = 1,N$ have positive real parts [36] Hence infinite gain margin of the LQR and condition (17) show stability of (14) for λ_i real. For complex $\lambda_i = a + jb$ one has

$$\begin{aligned}
&(A - c(a+jb)BR^{-1}B^T P)^* P + P(A - c(a+jb)BR^{-1}B^T P) + Q + (2ca-1)PBR^{-1}B^T P \\
&= (A - caBR^{-1}B^T P)^T P + P(A - caBR^{-1}B^T P) + Q + (2ca-1)PBR^{-1}B^T P \\
&= A^T P + PA + Q - PBR^{-1}B^T P
\end{aligned}$$

with * the complex conjugate transpose. According to (15) this is equal to zero and by condition (17) $(2ca-1) > 0$ so that this serves as a Lyapunov equation for $(A - c\lambda_i BK)$.

Li, Duan, Chen- Finsler's Lemma

OPTIMAL Design at each node gives global guaranteed performance on any strongly connected communication graph

Emre Tuna 2008 paper online

Example: Unbounded Region of Consensus for Optimal Feedback Gains.

$A - c\lambda BK$

$\lambda =$ E-vals of (L+G)

a. Bounded Consensus Region for Arbitrarily Chosen Stabilizing SVFB Gain

$$A = \begin{bmatrix} -2 & -1 \\ 2 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$K = \begin{bmatrix} 0.5 & -0.5 \end{bmatrix}$$

Example from [Zhongkui Li, Duan, Ron Chen 2009]

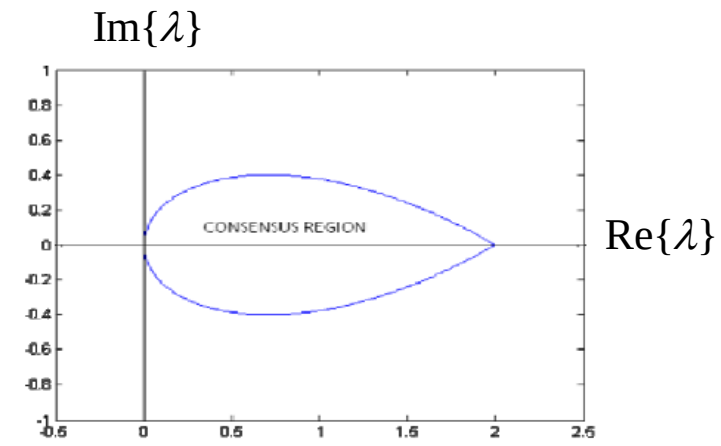


Figure 1. Bounded consensus region for arbitrarily chosen stabilizing SVFB gain

b. Unbounded Consensus Region for Optimal SVFB Gain

Q=I, R=1

$$K = \begin{bmatrix} 1.544 & 1.8901 \end{bmatrix}$$

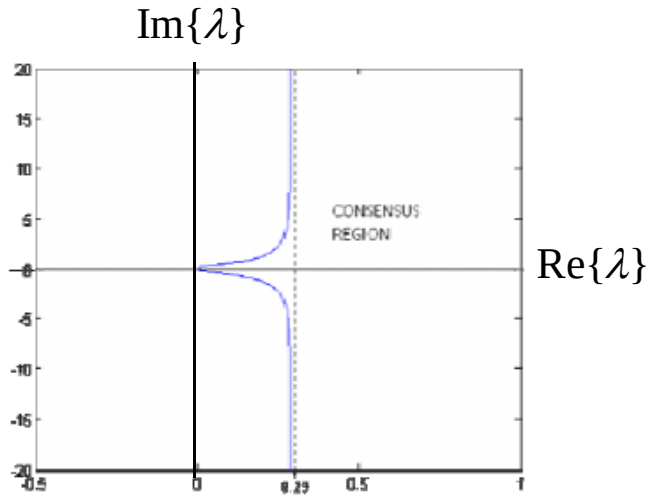
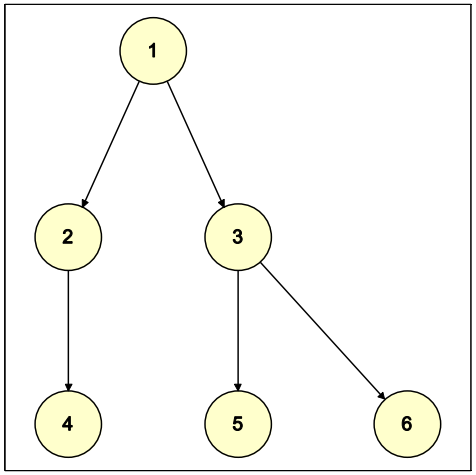
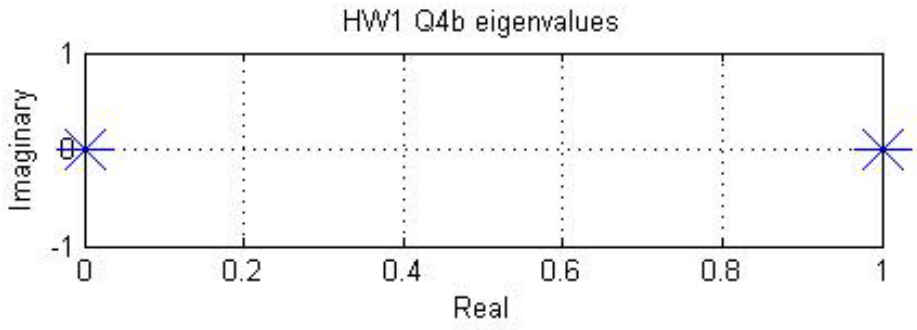


Figure 2. Unbounded consensus region for optimal SVFB gain

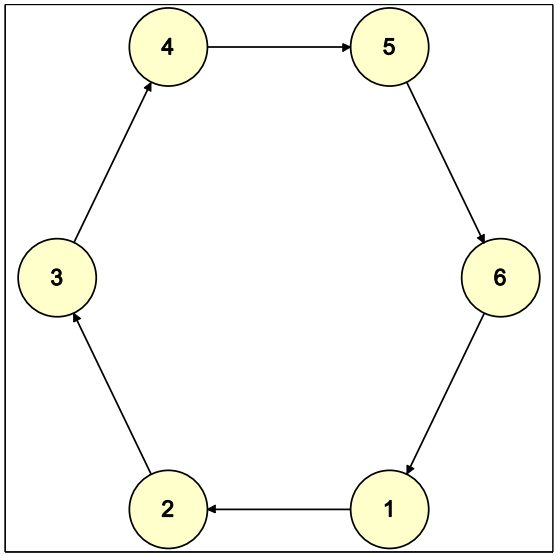
Graph Eigenvalues for Different Communication Topologies

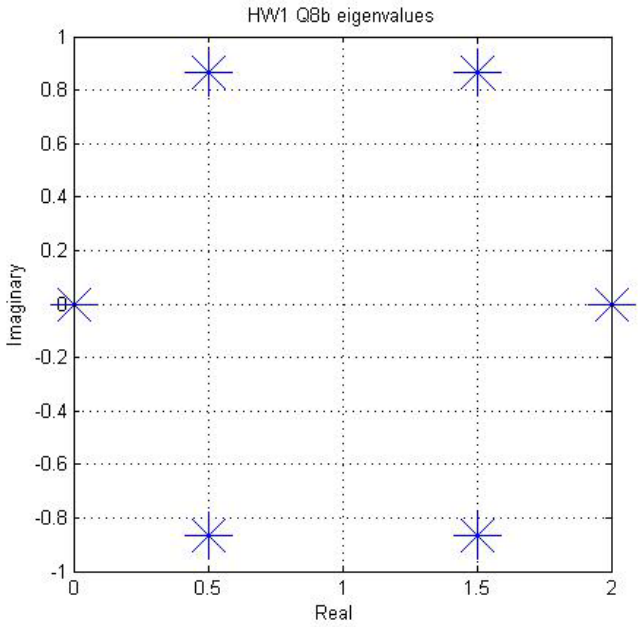
Directed Tree-Chain of command





Directed Ring-Gossip network
OSCILLATIONS

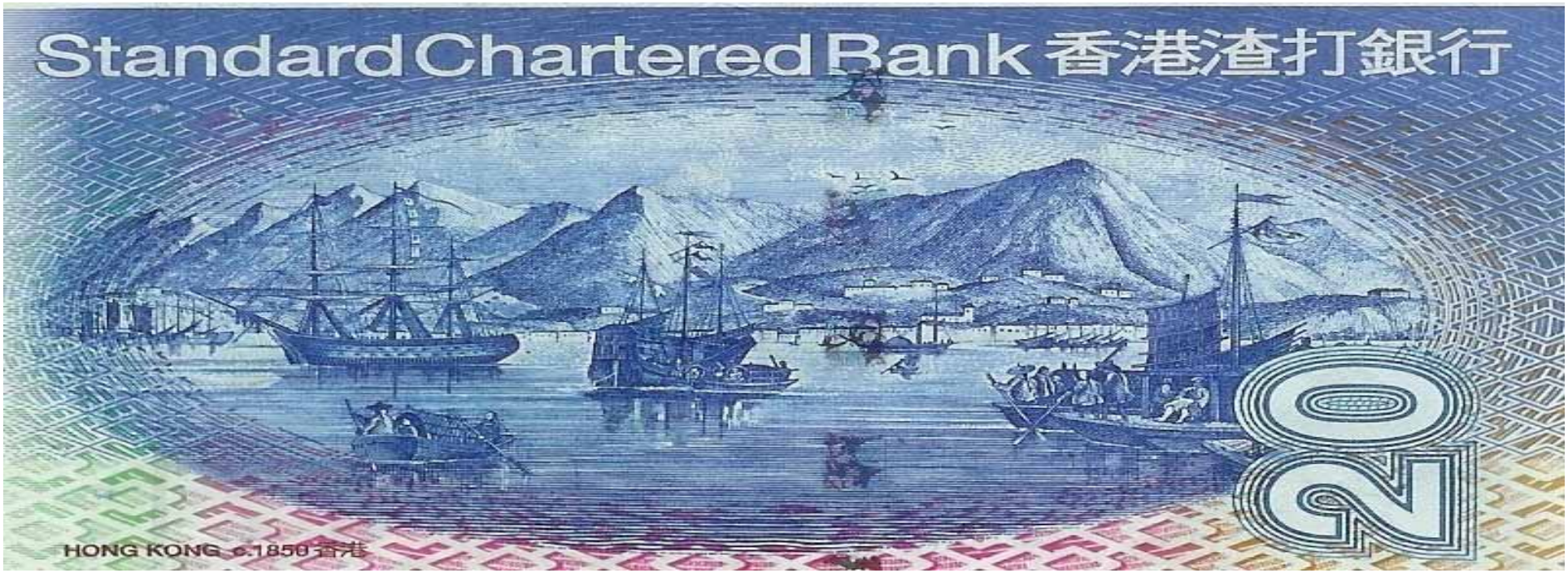


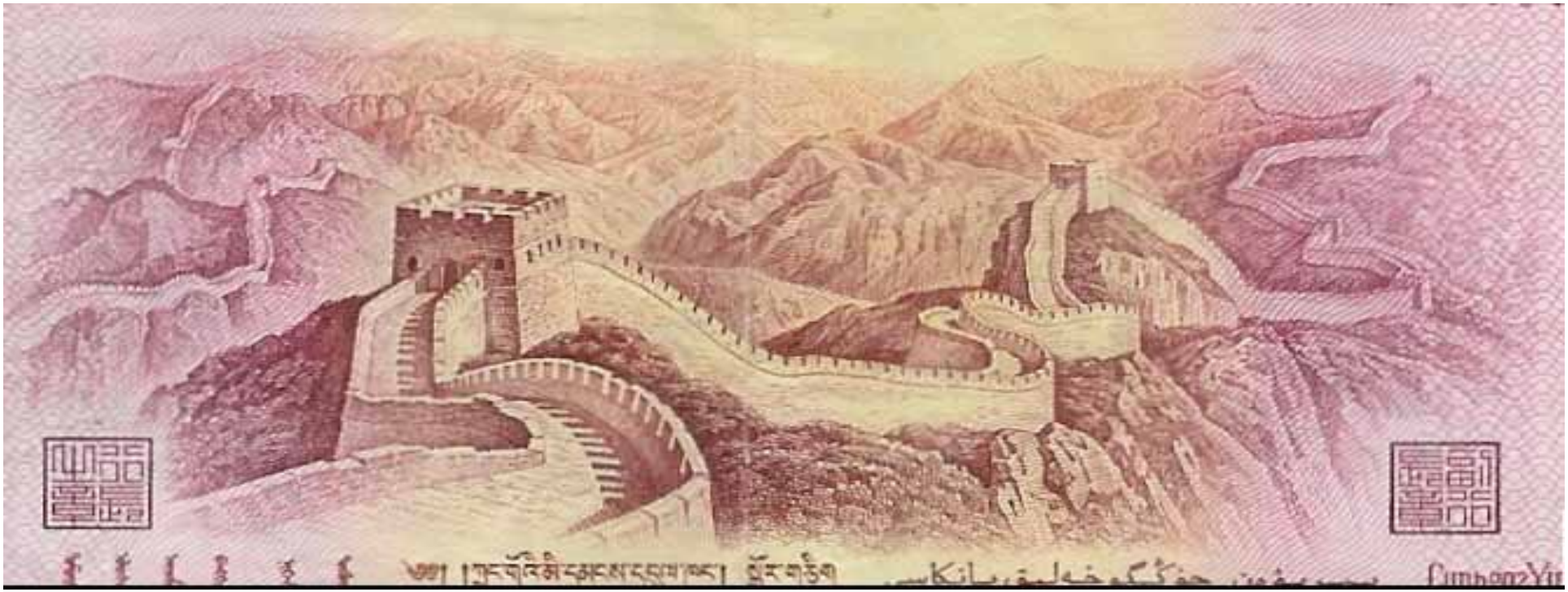


Results:

Local Riccati Design yields guaranteed stable synchronization

Decouples Controls Design from Graph Properties


Standard Chartered Bank 香港渣打銀行
HONG KONG c.1850 香港
20



Cooperative Observers and Dynamic Regulators



1. Cooperative State Feedback, Observers, Duality, and Optimal Design for Synchronization

Results: Distributed dynamic regulator for **synchronization of teams** using only **output measurements**
Duality structure theory extended to networked cooperative feedback systems on graphs
Optimal Design yields synchronization on ANY strongly connected Communication di-graph

CONTROL ⟵ Duality ⟶ DISTRIBUTED ESTIMATION & SENSOR FUSION

Cooperative system node dynamics $\dot{x}_i = Ax_i + Bu_i$

State feedback with local neighborhood tracking error

$$u_i = cK\varepsilon_i = cK\left(\sum_{j\in N_i} e_{ij}(x_j - x_i) \quad + g_i(x_0 - x_i)\right)$$

Pinning control node

Overall Cooperative Team Dynamics

$$\dot{x} = \left[(I_N \otimes A) - c(L+G)\otimes BK\right]x + c\left[(L+G)\otimes BK\right]\underline{x}_0$$

Output measurements at each node $y_i = Cx_i$

Local nbhd. estimation error

$$\varepsilon_i^o = \sum_{j\in N_i} e_{ij}(\tilde{y}_j - \tilde{y}_i) \quad + g_i(\tilde{y}_0 - \tilde{y}_i)_i, \qquad \tilde{y}_i = y_i - \hat{y}_i$$

Use local coop. observer dynamics at each node

$$\dot{\hat{x}}_i = A\hat{x}_i + Bu_i - cF\varepsilon_i^o$$

Overall Team Observer/Sensor Fusion Dynamics

$$\dot{\hat{x}} = \left[(I_N \otimes A) - c(L+G)\otimes FC\right]\hat{x} + c\left[(L+G)\otimes F\right]y + (I_N \otimes B)u$$

Thm 1. Design of SVFB Gain for Coop. Tracking Stability

Use OPTIMAL feedback gain

$$0 = A^T P + PA + Q - PBR^{-1}B^T P, \quad K = R^{-1}B^T P$$

Then synchronization is achieved for
ANY strongly connected digraph

Thm 2. Design of Observer Gain for Coop. Estimation

Use OPTIMAL observer gain

$$0 = AP + PA^T + Q - PC^T R^{-1}CP, \quad F = PC^T R^{-1}$$

Then estimates converge for
ANY strongly connected digraph

Unbounded Region of Consensus for Optimal Gains if coupling gain is $c > \dfrac{1}{\min_i \mathrm{Re}(\lambda_i(L+G))}$

B. Observer Design for Cooperative Systems on Graphs

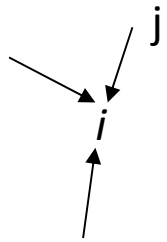
N nodes with dynamics $\dot{x}_i = Ax_i + Bu_i\,, \qquad y_i = Cx_i$

State and output estimates $\hat{x}_i(t) \in R^n, \quad \hat{y}_i(t) = C\hat{x}_i(t) \in R^p$

State and output estimation errors $\tilde{x}_i = x_i - \hat{x}_i, \qquad \tilde{y}_i = y_i - \hat{y}_i$

Control node or Command generator dynamics $\dot{x}_0 = Ax_0, \qquad y_0 = Cx_0$

Cooperative Observer design problem $\tilde{x}_i(t) \to 0, \forall i$ or $\hat{x}_i(t) \to x_i(t), \forall i$



Local neighborhood estimation error $\boxed{\varepsilon_i^o = \sum_{j \in N_i} e_{ij}(\tilde{y}_j - \tilde{y}_i) \quad + g_i(\tilde{y}_0 - \tilde{y}_i)}$

Ren; Beard, Kingston 2008

Overall estimation error $e^o = -\big((L+G)\otimes I_n\big)\big(\tilde{y} - \underline{\tilde{y}}_0\big) = -\big((L+G)\otimes I_n\big)\tilde{y} = -\big((L+G)\otimes C\big)\tilde{x}$

Lemma 3. Let the graph be strongly connected and $G \neq 0$. Then

$$\|\tilde{y}\| \le \|e^0\| / \underline{\sigma}(L+G) \qquad (26)$$

with $\underline{\sigma}(L+G)$ the minimum singular value of $(L+G)$. Then, if (A,C) is observable, $\|\tilde{x}\|$ is bounded in terms of the observability gramian norm and $\|\tilde{y}\|$. Moreover, $e^0(t) \equiv 0$ if and only if $\tilde{y}_i \equiv 0, \forall i$. Then, if (A,C) is observable, $\tilde{x}_i \equiv 0, \forall i$.

■

Local node observers

$$\dot{\hat{x}}_i = A\hat{x}_i + Bu_i - cF\varepsilon_i^o$$

$$\dot{\hat{x}}_i = A\hat{x}_i + Bu_i - cF\left[\sum_{j\in N_i} e_{ij}(\tilde{y}_j - \tilde{y}_i) \quad + g_i(\tilde{y}_0 - \tilde{y}_i)\right]$$

$$\dot{\hat{x}}_i = A\hat{x}_i + Bu_i - cFC\left[\sum_{j\in N_i} e_{ij}(\tilde{x}_j - \tilde{x}_i) \quad + g_i(\tilde{x}_0 - \tilde{x}_i)\right]$$

Overall observer dynamics $\dot{\hat{x}} = \left[(I_N \otimes A) - c(L+G)\otimes FC\right]\hat{x} + c\left[(L+G)\otimes F\right]y + (I_N \otimes B)u$

$$\dot{\hat{x}} = A_o^{Gr}\hat{x} + F_o^{Gr}y + (I_N \otimes B)u$$

where $\hat{x} = \left[\hat{x}_1^T \quad \hat{x}_2^T \quad \cdots \quad \hat{x}_N^T\right]^T \in R^{nN}$

Overall estimation error dynamics $\dot{\tilde{x}} = \left[(I_N \otimes A) - c(L+G)\otimes FC\right]\tilde{x} = A_o^{Gr}\tilde{x}$

Lemma 4. [16]. Let the graph be strongly connected with at least one pinning gain $g_i > 0$. Let $\lambda_i, i = 1, N$ be the eigenvalues of $(L+G)$. Then the observer error dynamics (33) are asymptotically stable (AS) if and only if the matrices

$$A - c\lambda_i FC \tag{34}$$

are all stable.

Proof Outline: Exactly the dual to the proof of Lemma 2. See [16]. ■

Li, Duan, Ron Chen 2009

OPTIMAL Design at Each node

Theorem 2. Design of Observer Gain for Cooperative Estimation

Suppose (A,C) is detectable and the graph is strongly connected with at least one pinning gain $g_i > 0$. Select design matrices $Q = Q^T > 0, R = R^T > 0$. Compute the observer gain F according to the observer algebraic Riccati equation (OARE)

$$0 = AP + PA^T + Q - PC^T R^{-1} CP \tag{35}$$

$$F = PC^T R^{-1} \tag{36}$$

Then the estimation error dynamics (33) are asymptotically stable for all coupling gains

$$c > \frac{1}{\underline{\lambda}(L+G)} \tag{37}$$

with $\underline{\lambda}(L+G) = \min_i \mathrm{Re}(\lambda_i(L+G))$.

Proof: Based on the robust infinite gain margin property of the Kalman filter [41], exactly dual to proof of Theorem 1.

■

OPTIMAL Design at Each node gives guaranteed performance
On any strongly connected communication di-graph topology

c.f. Finsler Lemma design in Li, Duan, Ron Chen 2009

Emre Tuna 2008 paper online
(without observer design)

C. Control/Observer Duality on Graphs

SVFB

$$u_i = cK\varepsilon_i = cK\left(\sum_{j\in N_i} e_{ij}(x_j - x_i) \quad + g_i(x_0 - x_i)\right),$$

Observer

$$\dot{\hat{x}}_i = A\hat{x}_i + Bu_i - cFC\left[\sum_{j\in N_i} e_{ij}(\tilde{x}_j - \tilde{x}_i) \quad + g_i(\tilde{x}_0 - \tilde{x}_i)\right]$$

Must use local nbhd. Tracking error and local nbhd. Estimation error or duality does not work

Converse, reverse, or transpose graph = reverse edge arrows

Theorem 3. SVFB Control/Observer Duality on Graphs (balanced)

Given dynamical nodes with input-coupling structure (A,B) and communication coupling graph Gr, suppose SVFB K stabilizes the synchronization dynamics (13). Then observer gain K^T stabilizes the estimation error dynamics (33) for systems with output-coupling structure (A^T, B^T) and communication coupling graph Gr'.

Proof:

The synchronization error dynamics is

$$\dot{\delta} = \left[(I_N \otimes A) - c(L+G)\otimes BK\right]\delta \equiv A_c^G \delta$$

and the estimation error dynamics is

$$\dot{\tilde{x}} = \left[(I_N \otimes A) - c(L+G)\otimes FC\right]\tilde{x} \equiv A_o^{Gr}\tilde{x}$$

Standard operations on the Kronecker product [38] reveal that

$$\left[(I_N \otimes A) - c(L+G)\otimes BK\right]^T = \left[(I_N \otimes A^T) - c(L+G)^T \otimes K^T B^T\right]$$

Now the result follows from the standard dynamical system duality relation based on $(A-BK)^T = (A^T - K^T B^T) \approx (A - FC)$ [43].

■

D. Dynamic Tracker for Synchronization of Cooperative Systems Using Output Feedback

N nodes with dynamics $\dot{x}_i = Ax_i + Bu_i, \quad y_i = Cx_i$

Command generator dynamics (exosystem) $\dot{x}_0 = Ax_0, \quad y_0 = Cx_0$

Observers at each node

$$\dot{\hat{x}}_i = A\hat{x}_i + Bu_i - cF\varepsilon_i^o$$

$$\dot{\hat{x}}_i = A\hat{x}_i + Bu_i - cF\left[\sum_{j\in N_i} e_{ij}(\tilde{y}_j - \tilde{y}_i) \quad + g_i(\tilde{y}_0 - \tilde{y}_i)\right]$$

Estimated SVFB $u_i = cK\hat{\varepsilon}_i = cK\left[\sum_{j\in N_i} e_{ij}(\hat{x}_j - \hat{x}_i) \quad + g_i(x_0 - \hat{x}_i)\right]$

Closed-loop systems $\dot{x}_i = Ax_i + cBK\left[\sum_{j\in N_i} e_{ij}(\hat{x}_j - \hat{x}_i) \quad + g_i(x_0 - \hat{x}_i)\right]$

Overall system/observer dynamics $\dot{x} = (I_N \otimes A)x - c(L+G)\otimes BK\hat{x} + c\left[(L+G)\otimes BK\right]\underline{x}_0$

$$\dot{\hat{x}} = \left[(I_N \otimes A) - c(L+G)\otimes FC\right]\hat{x} + c\left[(L+G)\otimes F\right]y$$
$$- c(L+G)\otimes BK\hat{x} + c\left[(L+G)\otimes BK\right]\underline{x}_0$$

Must use local nbhd. Tracking error and local nbhd. Estimation error or it is not nice.

Theorem 4. OPFB Regulator Synchronization Performance

Let *(A,B,C)* be stabilizable and detectable, the communication graph be strongly connected, and at least one pinning gain nonzero. Select the SVFB gain K and pinning gain c according to Theorem 1 and the observer gain F according to Theorem 2. Then the cooperative dynamic OPFB trackers given by (41), (42) yield synchronization of all nodes to the control state $x_0(t)$.

Proof:

Write (44) as

$$\dot{x} = \left[(I_N \otimes A) - c(L+G) \otimes BK\right] x + c\left[(L+G) \otimes BK\right] \tilde{x} + c\left[(L+G) \otimes BK\right] \underline{x}_0 \qquad (46)$$

or in terms of the synchronization error (6)

$$\dot{\delta} = \left[(I_N \otimes A) - c(L+G) \otimes BK\right] \delta + c\left[(L+G) \otimes BK\right] \tilde{x} = A_c^{Gr} \delta + B_c^{Gr} \tilde{x} \qquad (47)$$

where the observer error dynamics are

$$\dot{\tilde{x}} = \left[(I_N \otimes A) - c(L+G) \otimes FC\right] \tilde{x} = A_o^{Gr} \tilde{x} \qquad (48)$$

Then one has

$$\begin{bmatrix} \dot{\delta} \\ \dot{\tilde{x}} \end{bmatrix} = \begin{bmatrix} A_c^{Gr} & B_c^{Gr} \\ 0 & A_o^{Gr} \end{bmatrix} \begin{bmatrix} \delta \\ \tilde{x} \end{bmatrix}$$

Theorem 1 shows the stability of A_c^{Gr} while Theorem 2 shows the stability of A_o^{Gr}.

■

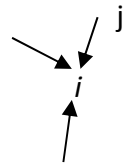
OPTIMAL Design at Each node gives guaranteed performance
On any strongly connected communication di-graph topology

Three Regulator Designs

1. Neighborhood Controller and Neighborhood Observer

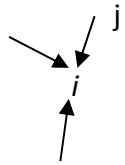
Nbhd Observers

$$\dot{\hat{x}}_i = A\hat{x}_i + Bu_i - cF\left[\sum_{j\in N_i} e_{ij}(\tilde{y}_j - \tilde{y}_i) \quad + g_i(\tilde{y}_0 - \tilde{y}_i)\right]$$



Nbhd Controls

$$u_i = cK\hat{\varepsilon}_i = cK\left[\sum_{j\in N_i} e_{ij}(\hat{x}_j - \hat{x}_i) \quad + g_i(x_0 - \hat{x}_i)\right]$$



2. Neighborhood Controller and Local Observer

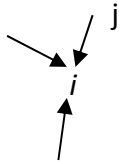
Local Observers

$$\dot{\hat{x}}_i = A\hat{x}_i + Bu_i - cF\,\tilde{y}_i.$$



Nbhd Controls

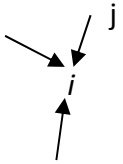
$$u_i = cK\hat{\varepsilon}_i = cK\left[\sum_{j\in N_i} e_{ij}(\hat{x}_j - \hat{x}_i) \quad + g_i(x_0 - \hat{x}_i)\right]$$



3. Local Controller and Neighborhood Observer

Nbhd Observers

$$\dot{\hat{x}}_i = A\hat{x}_i + Bu_i - cF\left[\sum_{j\in N_i} e_{ij}(\tilde{y}_j - \tilde{y}_i) \quad + g_i(\tilde{y}_0 - \tilde{y}_i)\right]$$



Local Controls

$$u_i = K\hat{x}_i$$



Discrete-Time Optimal Design for Synchronization



Discrete-Time Optimal Design for Synchronization

K. Hengster-Movric, Keyou You, F.L. Lewis, and Lihua Xie, "Synchronization of Discrete-Time Multi-agent Systems on Graphs Using Riccati Design," Automatica, vol. 49, no. 2, pp. 414-423, Feb. 2013.

Distributed systems

Decouple controls design from graph topology

$$x_i(k+1) = Ax_i(k) + Bu_i(k)$$

Command generator

$$x_0(k+1) = Ax_0(k)$$

Local Nbhd Tracking Error

$$\varepsilon_i = \sum_{j \in N_i} e_{ij}(x_j - x_i) + g_i(x_0 - x_i)$$

Local cooperative SVFB - weighted

$$u_i = c(1 + d_i + g_i)^{-1} K\varepsilon_i$$

Local closed-loop dynamics

$$x_i(k+1) = Ax_i(k) + c(1 + d_i + g_i)^{-1} BK\varepsilon_i(k)$$

Global disagreement error dynamics $\delta(k) = x(k) - \bar{x}_0(k)$

$$\delta(k+1) = A_c\delta(k) = \left[I_N \otimes A - c(I + D + G)^{-1}(L + G) \otimes BK \right] \delta(k)$$

Weighted Graph Matrix

$$\Gamma = (I + D + G)^{-1}(L + G)$$

Weighted graph eigenvalues $\Lambda_k, \quad k = 1, N$

Bounded Stability region of Sampled-data Systems

The Unit Circle in the z-plane

Synchronization error dynamics

$$\delta(k+1) = A_c\delta(k) = \left[I_N \otimes A - c(I+D+G)^{-1}(L+G)\otimes BK \right]\delta(k)$$

Weighted Graph Matrix

$$\Gamma = (I + D + G)^{-1}(L + G)$$

Weighted graph eigenvalues $\Lambda_k, \quad k = 1, N$

Lemma 1. The multi-agent systems (5) synchronize if and only if $\rho(A - c\Lambda_k BK) < 1$ for all eigenvalues Λ_k, $k = 1...N$, of graph matrix (10).

MIXES UP CONTROL DESIGN AND GRAPH STRUCTURE

Decouple Controls Design From Graph Topology

Kristian Movric

Theorem 2. H_2 **Riccati Design for Synchronization.** Assume the interaction graph contains a spanning tree with at least one pinning gain nonzero that connects into the root node. Let $P>0$ be a solution of the discrete-time Riccati-like equation

$$A^T PA - P + Q - A^T PB(B^T PB)^{-1} B^T PA = 0$$

for some prescribed $Q=Q^T>0$. Define

$$r := \left[\sigma_{\max}(Q^{-1/2} A^T PB(B^T PB)^{-1} B^T PAQ^{-1/2})\right]^{-1/2}.$$

Then protocol guarantees synchronization of the multi-agent systems for some K if there exists a covering circle $C(c_0, r_0)$ of the graph matrix eigenvalues $\Lambda_k, k=1\ldots N$ such that

$$\frac{r_0}{c_0} < r.$$

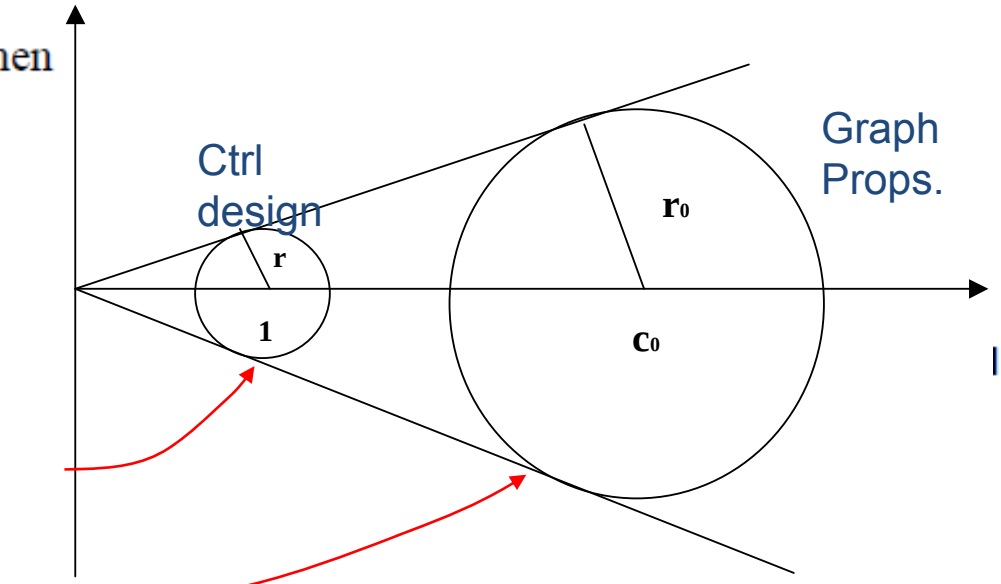
Moreover, if this condition is satisfied then

$$K = (B^T PB)^{-1} B^T PA$$

and coupling gain

$$c = \frac{1}{c_0}$$

guarantee synchronization.



Synchronization region contains this circle

Covering circle of graph eigenvalues

Single-Input case with Real Graph Eigenvalues

Corollary 5. Let the distributed systems be single-input and let the Γ matrix of the graph $G(V,E)$ have all eigenvalues real. Select Q as in Corollary 4. Then the synchronization condition becomes

$$\prod_u |\lambda^u(A)| < \frac{\Lambda_{\max}+\Lambda_{\min}}{\Lambda_{\max}-\Lambda_{\min}}$$

Moreover this condition is necessary and sufficient for synchronization for any choice of the feedback matrix K if all the eigenvalues of A lie on or outside the unit circle.

Mahler Measure

$$M(A) = \prod_{unstable\ \lambda_i} |\lambda_i(A)|$$

Intrinsic Entropy Rate (Baillieul &c)

minimum data rate in networked systems

$$\log_2 \prod_i \lambda_i^u(A)$$

Synchronization guaranteed if $M(A) < \frac{1+\kappa^{-1}}{1-\kappa^{-1}}$

Topological Entropy

$$h(A) = \log(M(A))$$

Graph Condition Number

$$\kappa(L+G) = \frac{\Lambda_{\max}}{\Lambda_{\min}}$$

Like to have $\kappa(G) \approx 1$ Varshney

$\Lambda_{\min}$ large means fast convergence

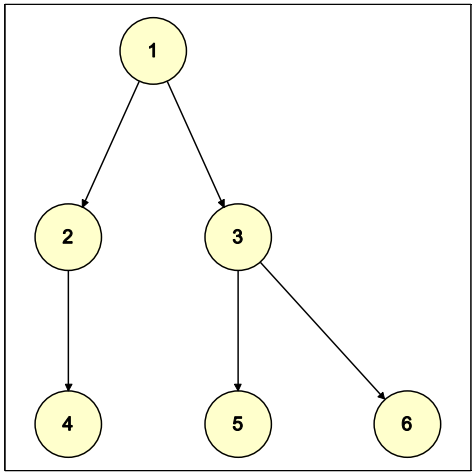
New definition- Graph Channel Capacity

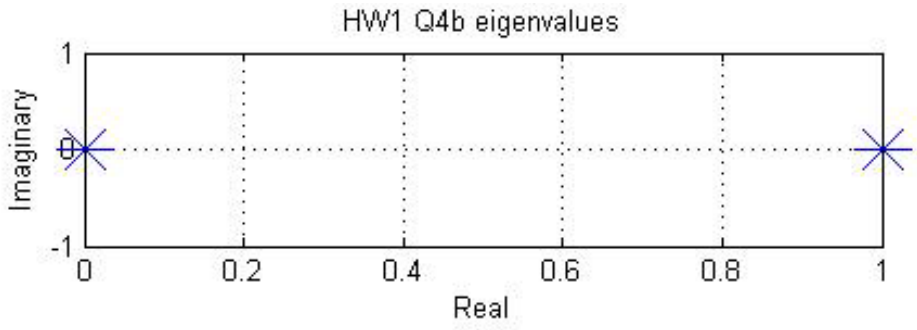
$$C(L+G) = \log \frac{1+\kappa^{-1}}{1-\kappa^{-1}}$$

K. Hengster-Movric, Keyou You, F.L. Lewis, and Lihua Xie, "Synchronization of Discrete-Time Multi-agent Systems on Graphs Using Riccati Design," Automatica, vol. 49, no. 2, pp. 414-423, Feb. 2013.

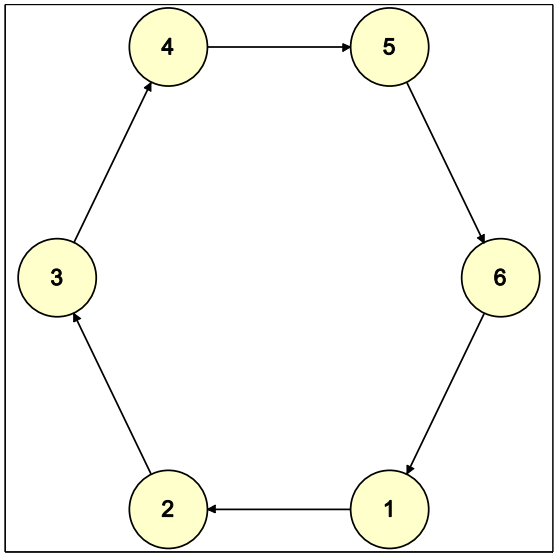
Graph Eigenvalues for Different Communication Topologies

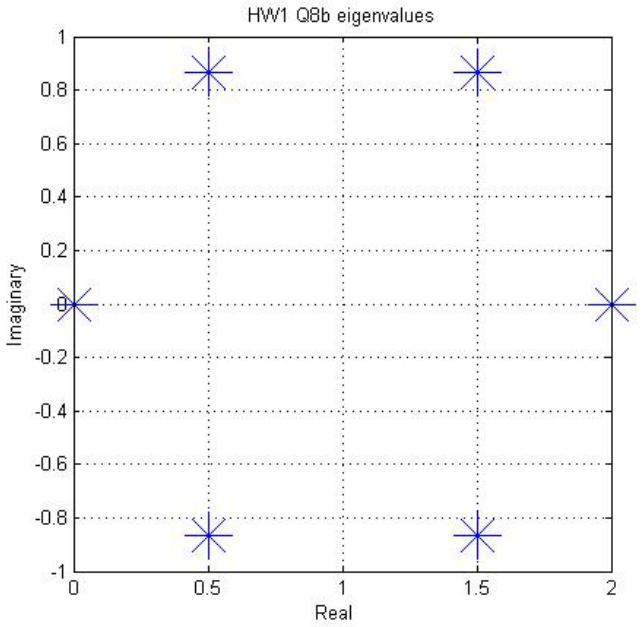
Directed Tree-Chain of command





Directed Ring-Gossip network
OSCILLATIONS





Kung Tz 500 BC

Confucius

孔子

Tian xia da tong
Harmony under heaven

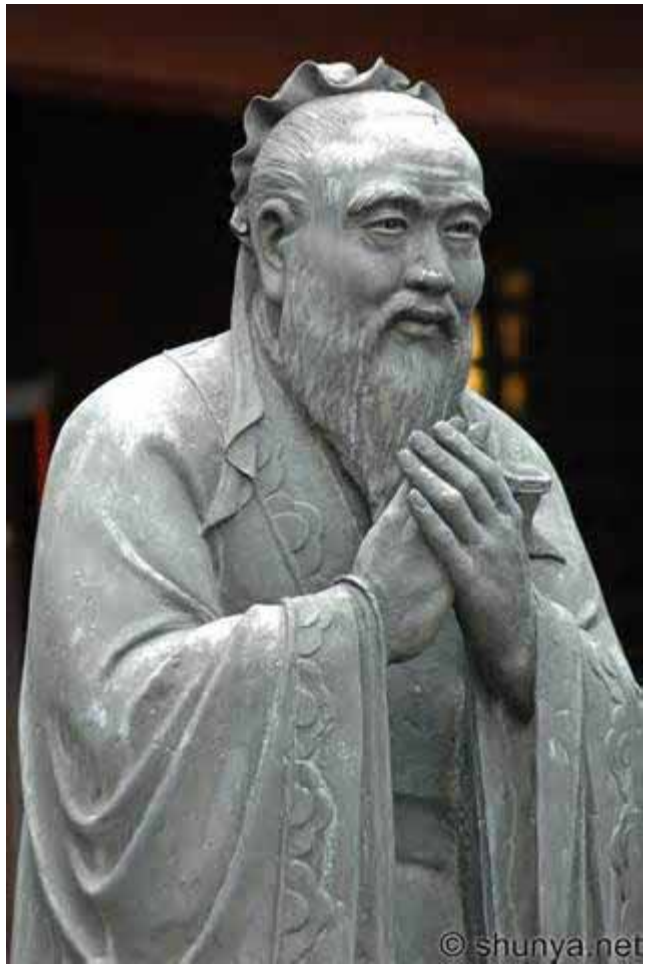
Archery
Chariot driving

Music
Rites and Rituals

Poetry
Mathematics

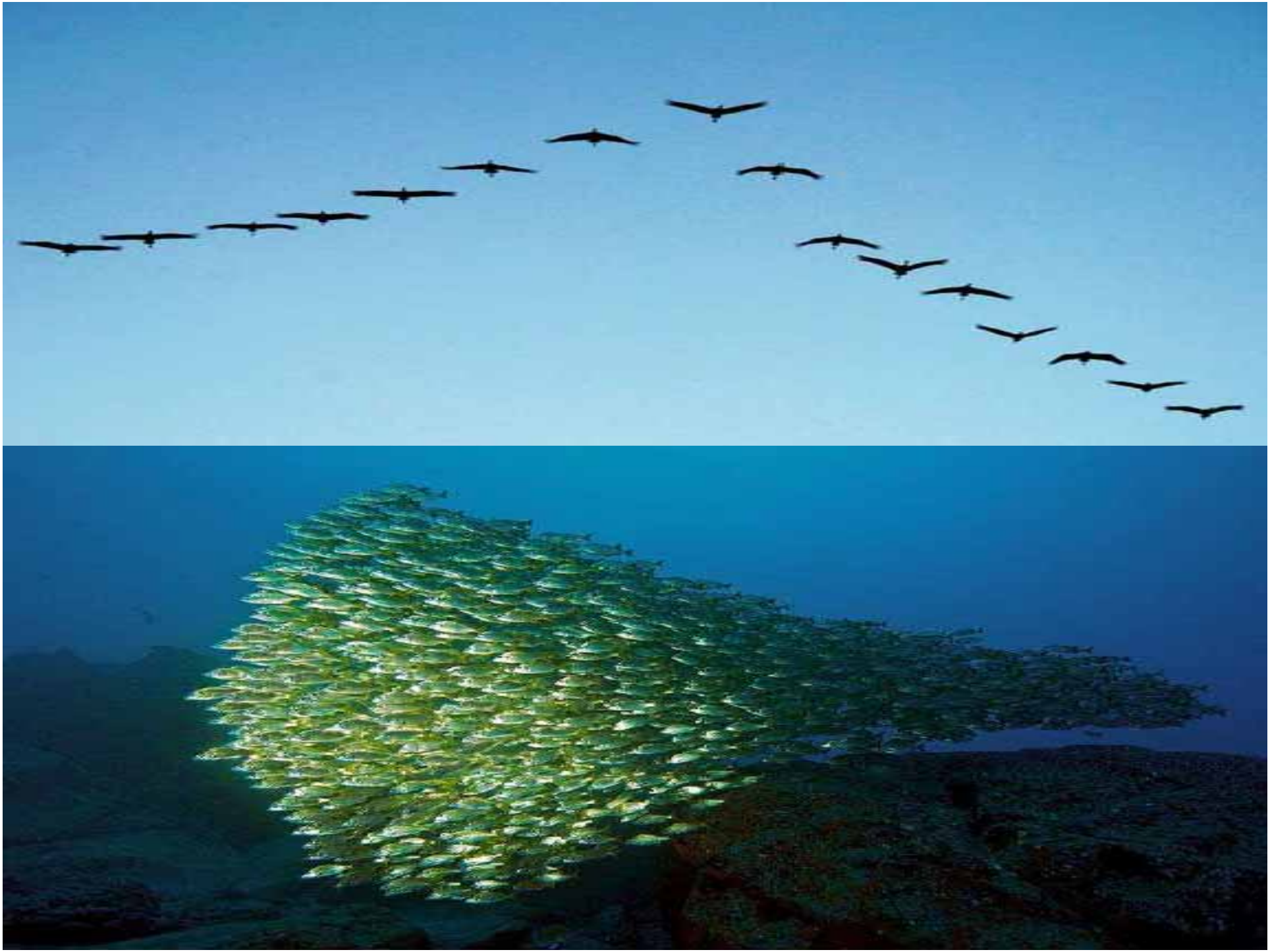


Man's relations to
- Family
- Friends
- Society
- Nation
- Emperor
- Ancestors



Globally Optimal Design for Collective Group Motions





Stability vs. Optimality of Cooperative Control

Outline

- A. Stable Design for Synchronization of Cooperative Systems
- B. Optimal Design for Collective Group Motion

Issues: For cooperative control on graphs -

Local stability of each agent is NOT the same as stable synchronization of the team



Local optimality of each agent is NOT the same a global optimality of the team

Have seen that LOCAL OPTIMAL DESIGN Guarantees Global Synchronization

B. GLOBAL OPTIMAL DESIGN FOR COLLECTIVE MOTION ON GRAPHS

The method just shown guarantees synchronization on arbitrary graphs
It is a LOCAL OPTIMAL DESIGN at each agent

What about **Global Optimality** of cooperative control on graphs?

Problem- the global optimal control is not distributed

The global optimal control is generally distributed only on a complete graph – Wei Ren

Agent dynamics $\dot{x}_i = Ax_i + Bu_i \in \mathbb{R}^n$

Global dynamics $\dot{x} = (I \otimes A)x + (I \otimes B)u = \bar{A}x + \bar{B}u$

$\dot{\delta} = (I \otimes A)\delta + (I \otimes B)u \equiv \bar{A}\delta + \bar{B}u$ $\qquad \delta = \left(x - \underline{I}x_0\right)$

LQR $J = \frac{1}{2}\int_0^\infty (\delta^T Q\delta + u^T Ru)\, dt$

ARE $\bar{A}^T P + P\bar{A} + Q - P\bar{B}R^{-1}\bar{B}^T P$

Control $u = R^{-1}B^T P\delta$ is distributed only on a complete graph- Wei Ren

BUT- a distributed control must have the form $u = -c\left((L+G)\otimes K\right)\delta$

So *Q* and *R* must depend on the graph topology

Inverse Optimality for Coop control

Inspired by Wassim Haddad book

Kristian Hengster-Movric

Stability wrt Synchronization Manifold
noncompact set

Lemma 2a. (*Inverse optimality*) Consider the control affine system (1). Let $u=\phi(x)$ be a stabilizing control, with respect to a manifold S. If there exist scalar functions $V(x)$ and $L_1(x)$ satisfying the following conditions

$$V(x)=0 \Leftrightarrow x \in S$$
$$V(x) \geq \alpha(d(x,S))$$
$$L_1(x) \geq \gamma(d(x,S))$$
$$L_1(x)+\nabla V(x)^T f(x)-\frac{1}{4}\nabla V(x)^T g(x)R^{-1}g(x)^T \nabla V(x)=0 \quad (0)$$
$$\phi(x)=-\frac{1}{2}R^{-1}g(x)^T \nabla V(x)^T$$

then $u=\phi(x)$ is optimal with respect to the performance index with the integrand $\mathcal{L}(x,u)=L_1(x)+u^T Ru$. Moreover the optimal value of the performance criterion equals $J(x_0,\phi(x))=V(x_0)$.

LQR case- ARE $A^T P+PA+Q-PBR^{-1}B^T P$

Given *A, B*, and the distributed control form, find *Q* and *R*

$$u=-c\left((L+G)\otimes K\right)\delta$$

K. Hengster-Movric and F.L. Lewis, "Cooperative Optimal Control for Multi-agent Systems on Directed Graph Topologies", IEEE Trans. Automatic Control, to appear, 2013.

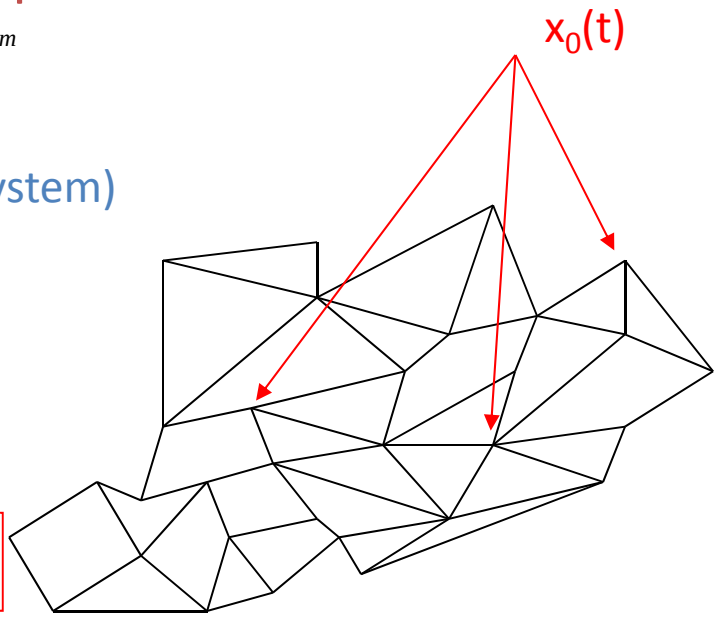
Cooperative Tracker for Identical LTI Dynamics

Cooperative Regulator vs. Cooperative Tracker problem

N nodes with dynamics $\dot{x}_i = Ax_i + Bu_i \quad , \quad x_i \in R^n, \quad u_i \in R^m$

Control node or Command generator $\dot{x}_0 = Ax_0$ (Exosystem)

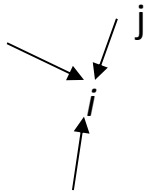
Synchronization Tracker design problem $x_i(t) \to x_0(t), \forall i$



Local neighborhood tracking error

$$\varepsilon_i = \sum_{j \in N_i} e_{ij}(x_j - x_i) \quad + g_i(x_0 - x_i)$$

Ron Chen- pinning control Lihua Xie- error



L=D-A

Overall error vector $e = -\left((L+G)\otimes I_n\right)(x - \underline{x}_0) = -\left((L+G)\otimes I_n\right)\delta$ = Local quantity

where $e = \left[\varepsilon_1^T \quad \varepsilon_2^T \quad \cdots \quad \varepsilon_N^T\right]^T \in R^{nN}, \quad \underline{x}_0 = \underline{I}x_0 \in R^{nN}, \quad \underline{I} = \underline{1} \otimes I_n \in R^{nN \times n}$

Consensus or synchronization error $\delta = \left(x - \underline{x}_0\right) \in R^{nN}$ = Global quantity

Cooperative Tracker for Identical LTI Dynamics

System $\dot{x}_i = Ax_i + Bu_i \in \mathbb{R}^n$

leader $\dot{x}_0 = Ax_0$

$$\dot{x} = (I \otimes A)x + (I \otimes B)u$$

$$\dot{\delta} = (I \otimes A)\delta + (I \otimes B)u \qquad \delta = x - \underline{I}x_0$$

Local nbhd tracking error $\varepsilon_i = \sum_{j \in N_i} e_{ij}(x_j - x_i) \quad + g_i(x_0 - x_i)$

$$e = -\big((L+G) \otimes I_n\big)\delta \qquad e = \begin{bmatrix} \varepsilon_1 & \dots & \varepsilon_N \end{bmatrix}^T$$

Control $u_i = cK_2\varepsilon_i = cK_2\left(\sum_{j \in N_i} e_{ij}(x_j - x_i) \quad + g_i(x_0 - x_i)\right)$

$$u = -c(L+G) \otimes K_2\delta$$

Closed-loop system $\dot{x}_i = Ax_i + Bu_i = Ax_i + cBK_2\left(\sum_{j \in N_i} e_{ij}(x_j - x_i) \quad + g_i(x_0 - x_i)\right)$

$$\dot{x} = \left[(I_N \otimes A) - c(L+G) \otimes BK_2\right]x + c\left[(L+G) \otimes BK_2\right]\underline{x}_0$$

$$\dot{\delta} = \left[(I_N \otimes A) - c(L+G) \otimes BK_2\right]\delta$$

Graph structure $\otimes$ Control structure

Kristian Hengster-Movric

Theorem 5. Let the error dynamics be given as (0), and conditions of Lemma 4 be satisfied. Suppose there exist a positive definite matrix $P_1 = P_1^T > 0$, and a positive definite matrix $P_2 = P_2^T > 0$ satisfying

$$P_1 = cR_1(L+G) \qquad (0)$$

$$A^T P_2 + P_2 A + Q_2 - P_2 B R_2^{-1} B^T P_2 = 0, \qquad (0)$$

for some $Q_2 = Q_2^T > 0, R_1 = R_1^T > 0, R_2 = R_2^T > 0$ and a coupling gain $c > 0$. Define the feedback gain matrix K_2 as

$$K_2 = R_2^{-1} B^T P_2. \qquad (0)$$

Then the control $u = -c(L+G) \otimes K_2 \delta$ is optimal with respect to the performance index

$$J(\delta_0, u) = \int_0^\infty \delta^T \left[c^2 ((L+G) \otimes K_2)^T (R_1 \otimes R_2)((L+G) \otimes K_2) - cR_1(L+G) \otimes (A^T P_2 + P_2 A) \right] \delta + u^T (R_1 \otimes R_2) u dt$$

,(0)

and is stabilizing to the origin for sufficiently high coupling gain c satisfying (0).

A new condition on the graph

TWO CONDITIONS

The local optimal design from before

$$u = -c(L+G) \otimes K_2 \delta$$

Q depends on Graph topology

$$\delta^T Q \delta$$

$$c > \frac{\sigma_{\max}(R_1(L+G) \otimes (Q_2 - K_2^T R_2 K_2))}{\sigma_{\min}((L+G)^T R_1 (L+G) \otimes K_2^T R_2 K_2)}$$

K. Hengster-Movric and F.L. Lewis, "Cooperative Optimal Control for Multi-agent Systems on Directed Graph Topologies", IEEE Trans. Automatic Control, to appear, 2013.

Kristian Hengster-Movric

Proof: System $\dot{\delta} = (I \otimes A)\delta + (I \otimes B)u \equiv \bar{A}\delta + \bar{B}u$

ARE $\bar{A}^T P + P\bar{A} + Q - P\bar{B}R^{-1}\bar{B}^T P$

$$(I \otimes A)^T P + P(I \otimes A) + Q - P(I \otimes B)R^{-1}(I \otimes B)^T P = 0$$

Select $P = P_1 \otimes P_2 \qquad R = R_1 \otimes R_2$

ARE $P_1 \otimes A^T P_2 + P_1 \otimes P_2 A + Q - (P_1 \otimes P_2 B)(R_1^{-1} \otimes R_2^{-1})(P_1 \otimes B^T P_2) = 0$

$$P_1 \otimes (A^T P_2 + P_2 A) + Q - P_1 R_1^{-1} P_1 \otimes (P_2 B R_2^{-1} B^T P_2) = 0$$

Choose Q

$$\begin{aligned} Q &= c^2((L+G) \otimes K_2)^T (R_1 \otimes R_2)((L+G) \otimes K_2) - cR_1(L+G) \otimes (A^T P_2 + P_2 A) \\ &= c^2 (L+G)^T R_1 (L+G) \otimes K_2^T R_2 K_2 + cR_1(L+G) \otimes (Q_2 - P_2 B R_2^{-1} B^T P_2) \\ &= P_1 R_1^{-1} P_1 \otimes P_2 B R_2^{-1} B^T P_2 + P_1 \otimes (Q_2 - P_2 B R_2^{-1} B^T P_2) \\ &= Q_1 \otimes P_2 B R_2^{-1} B^T P_2 + P_1 \otimes (Q_2 - P_2 B R_2^{-1} B^T P_2) \end{aligned}$$

$$Q_1 = c^2 (L+G)^T R_1 (L+G)$$

ARE $P_1 \otimes (A^T P_2 + P_2 A^T + Q_2 - P_2 B R_2^{-1} B^T P_2)$

$$+(Q_1 - P_1 R_1^{-1} P_1) \otimes (P_2 B R_2^{-1} B^T P_2) = 0$$

2 conditions:

$$A^T P_2 + P_2 A + Q_2 - P_2 B R_2^{-1} B^T P_2 = 0$$

$$P_1 = cR_1(L+G)$$

Control $u = R^{-1}\bar{B}^T P = -R^{-1}(I \otimes B)^T P\delta = -(R_1^{-1} \otimes R_2^{-1})(I \otimes B^T)(P_1 \otimes P_2)\delta$

$$= -R_1^{-1} P_1 \otimes R_2^{-1} B^T P_2 \delta = -c(L+G) \otimes K_2 \delta$$

Distributed !!

Two Conditions for global optimal design on the graph

1. Condition on graph topology

$$P_1 = cR_1(L+G) \qquad \text{For some} \qquad P_1 = P_1^T > 0, \qquad R_1 = R_1^T > 0$$

2. Local agent control design condition – Same as before-local optimal control

$$A^T P_2 + P_2 A + Q_2 - P_2 B R_2^{-1} B^T P_2 = 0$$

$$\text{For some} \quad P_2 = P_2^T > 0, \qquad R_2 = R_2^T > 0, \qquad Q_2 = Q_2^T > 0$$

Always holds if *(A,B)* reachable

Locally optimal design is also globally optimal on the graph if condition 1 holds

Kristian Hengster-Movric

Condition on Graph Topology

$P_1 = cR_1(L+G)$ Equivalent to $R_1(L+G) = (L+G)^T R_1$

$P_1 = P_1^T > 0, \qquad R_1 = R_1^T > 0$

1. Undirected Graphs

$L+G = (L+G)^T$

The condition becomes a Commutativity Requirement $R_1(L+G) = (L+G)R_1$

Case 1. $R_1 = I$

For single-integrator dynamics $J(\delta_0, u) = \int_0^\infty (\delta^T (L+G)^T (L+G)\delta + u^T u)dt = \int_0^\infty (e^T e + u^T u)dt$

Case 2. $R_1(L+G) = (L+G)R_1$ Iff $R_1, (L+G)$ have the same eigenvectors

Let $L = T\Lambda T^T$ Jordan form

Select $R = T\Theta T^T > 0$ For any $\Theta > 0$ diagonal

R depends on graph topology- ALL e-vectors

$$P_1 = cR_1(L+G) \quad \text{Equivalent to} \quad R_1(L+G) = (L+G)^T R_1$$

$$P_1 = P_1^T > 0, \qquad R_1 = R_1^T > 0$$

2. Detail Balanced Graphs

$$\lambda_i e_{ij} = \lambda_j e_{ji} \quad \text{for} \quad \lambda_1 \ldots \lambda_N > 0$$

Then $\begin{bmatrix} \lambda_1 & \cdots & \lambda_N \end{bmatrix}^T$ is a left eigenvector for *L for e-val= 0*

$L = D\bar{P}, \qquad D = diag\{1/\lambda_i\} > 0$ with $\bar{P}$ a symmetric graph Laplacian matrix

$$L + G = D\bar{P} + G = D(\bar{P} + D^{-1}G) \equiv DP$$

$$P = D^{-1}(L+G) = R(L+G)$$

R depends on graph topology – principal left e-vector

Detail balanced implies reversibility of an associated Markov Process

Detail balanced implies balanced

A new class of digraphs

$$P_1 = cR_1(L+G) \quad \text{Equivalent to} \quad R_1(L+G) = (L+G)^T R_1$$

$$P_1 = P_1^T > 0, \qquad R_1 = R_1^T > 0$$

3. Directed Graphs with Simple Graph Laplacian *L+G*

$$T(L+G)T^{-1} = \Lambda \qquad \text{Diagonal Jordan form}$$

$$T(L+G)T^{-1} = \Lambda = \Lambda^T = T^{-T}(L+G)^T T^T$$

$$T^T T(L+G) = (L+G)^T T^T T$$

Select $R = T^T T = R^T > 0$

Dennis Bernstein
Matrix book

$$(L+G) = R^{-1}(L+G)^T R, \quad R = R^T > 0$$

R depends on graph topology- ALL e-vectors

Theorem 6. Let L be a positive semi-definite matrix (generally not symmetric). Then there exists a positive definite symmetric matrix $R = R^T > 0$ such that $RL = P$ is a symmetric positive semi-definite matrix if and only if L is simple, *i.e.* there exists a basis of eigenvectors of L.

Distributed Systems



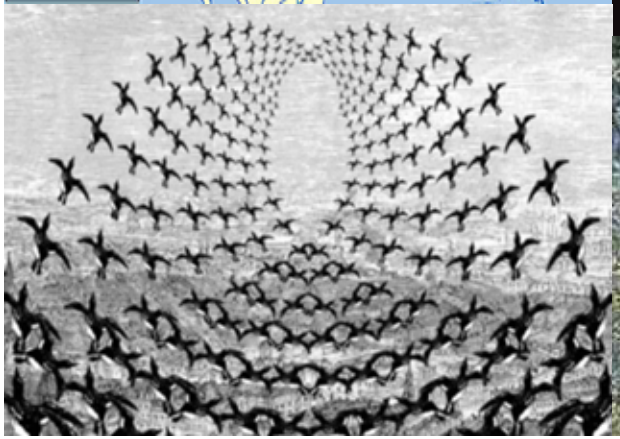




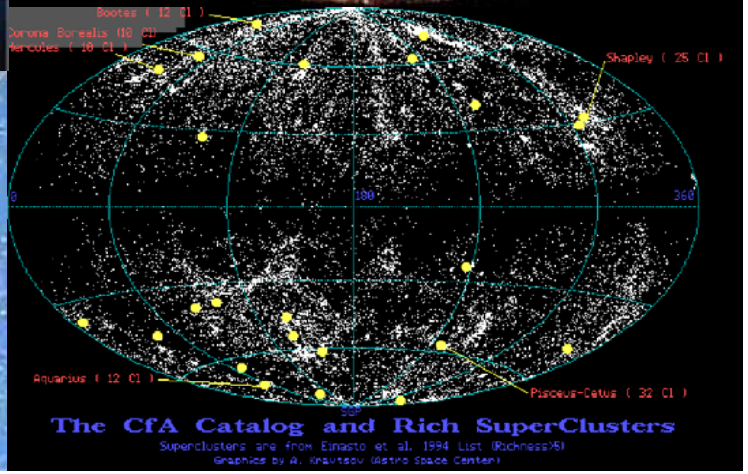






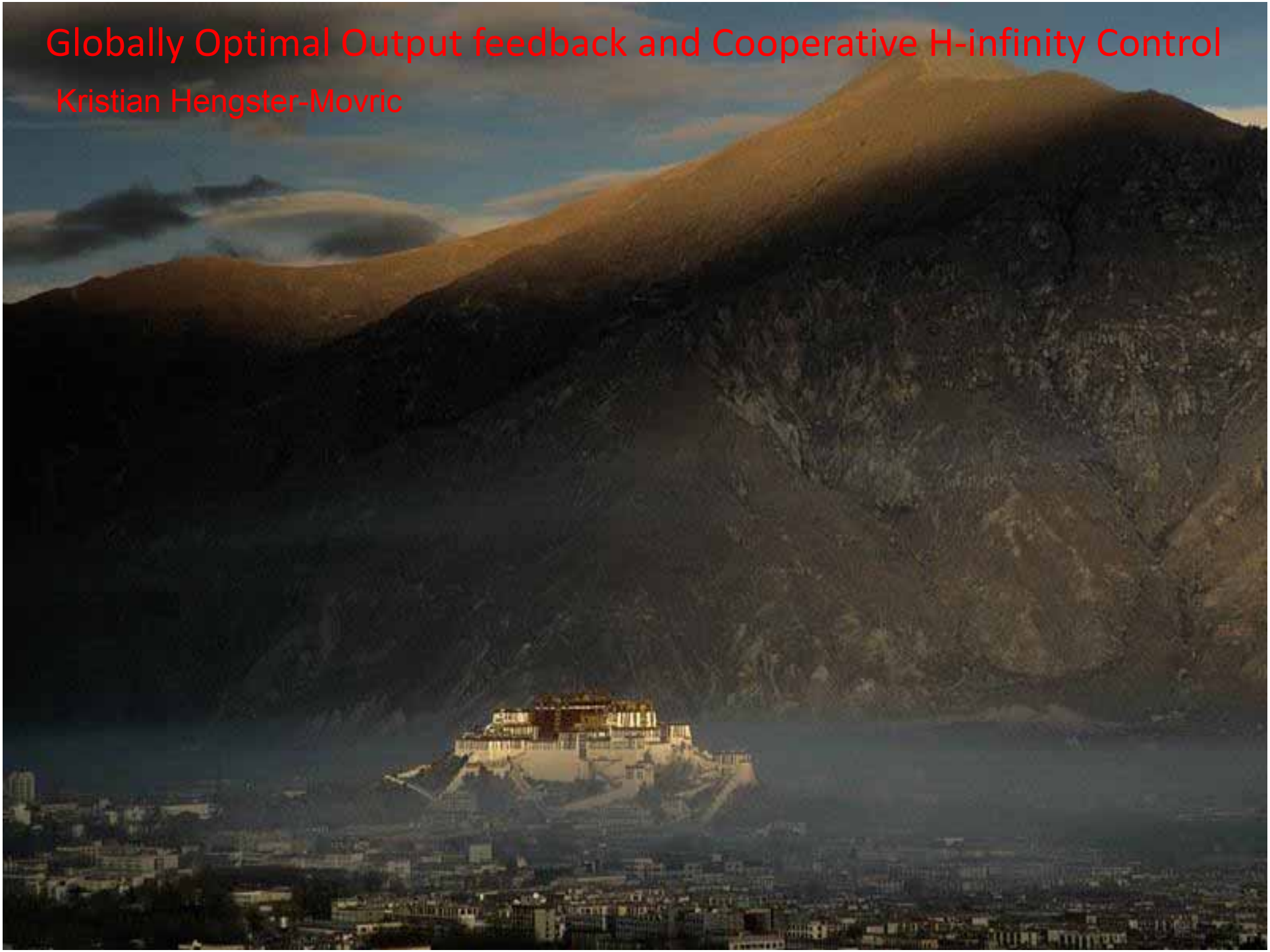






Globally Optimal Output feedback and Cooperative H-infinity Control

Kristian Hengster-Movric



Cooperative Static Output Feedback Tracker

N nodes with dynamics $\dot{x}_i = Ax_i + Bu_i$, $y_i = Cx_i$, $x_i \in R^n$, $u_i \in R^m$, $y_i \in R^p$

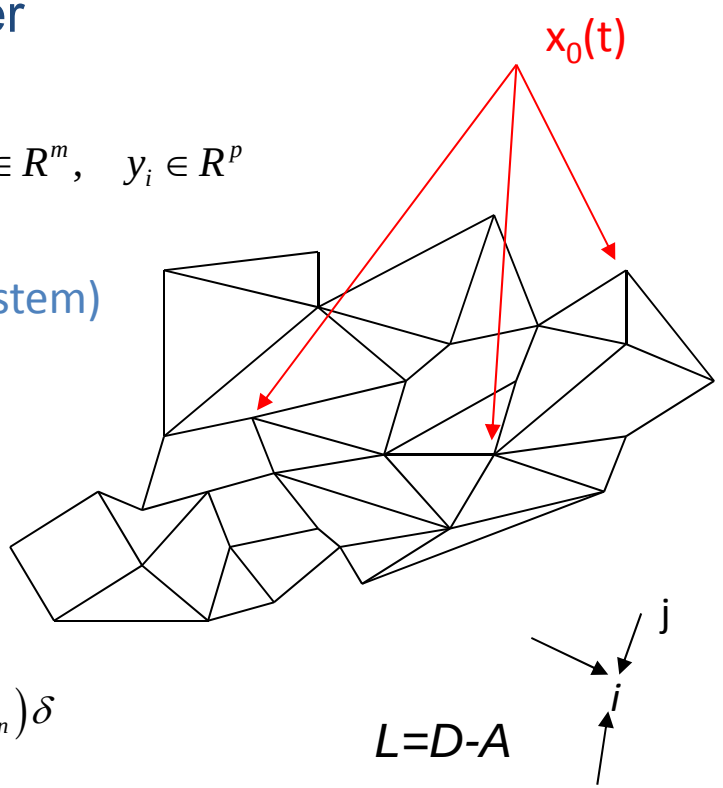
Control node or Command generator $\dot{x}_0 = Ax_0$, $y_0 = Cx_0$ (Exosystem)

Synchronization Tracker design problem $x_i(t) \to x_0(t), \forall i$

Local neighborhood tracking error

$$\varepsilon_i = \sum_{j \in N_i} e_{ij}(x_j - x_i) + g_i(x_0 - x_i)$$

Overall error vector $e = -\big((L+G)\otimes I_n\big)\big(x - \underline{x}_0\big) = -\big((L+G)\otimes I_n\big)\delta$



Local neighborhood output error

$$e_{yi} = C\varepsilon_i = \sum_{j \in N_i} e_{ij}(y_j - y_i) + g_i(y_0 - y_i)$$

Overall output error vector $\boxed{e_y = -\big((L+G)\otimes C\big)\delta}$ = Local quantity

Coarse global OPFB structure

Fine Local OPFB structure

The graph topology and the OPFB have exactly the same effects in limiting the information

We seek a unified design theory that treats both effects similarly

Coop. nbhd OPFB

$$u_i = cKC\varepsilon_i = cKe_{yi} = cK\left(\sum_{j\in N_i} e_{ij}(y_j - y_i) \quad + g_i(y_0 - y_i) \right)$$

Closed loop system $\dot{x}_i = Ax_i + Bu_i = Ax_i + cBKC\left(\sum_{j\in N_i} e_{ij}(x_j - x_i) \quad + g_i(x_0 - x_i) \right)$

Overall state $x = \begin{bmatrix} x_1^T & x_2^T & \cdots & x_N^T \end{bmatrix}^T \in R^{nN}, \qquad \delta = \left(x - \underline{I} x_0 \right)$

Distributed form of control $\boxed{u = -c\left((L+G) \otimes KC \right) \delta}$

Overall c.l. dynamics $\dot{x} = \left[(I_N \otimes A) - c(L+G) \otimes BKC \right] x + c\left[(L+G) \otimes BKC \right] \underline{x}_0$

Global synch. error dynamics $\dot{\delta} = \left[(I_N \otimes A) - c(L+G) \otimes BKC \right] \delta$

Graph structure $\otimes$ Control structure

Lemma 1. The matrix $I_N \otimes A - c(L+G) \otimes BKC$ is Hurwitz if and only if all the matrices $A - c\lambda_j BKC$ are stable, where λ_j are the eigenvalues of the graph matrix $(L+G)$.

c.f. Fax and Murray 2004

MIXES UP CONTROL DESIGN AND GRAPH STRUCTURE

Static OPFB for (single-agent) systems

Proposition 1. Let the linear time-invariant system be given as

$$\begin{aligned} \dot{x} &= Ax + Bu, \\ y &= Cx. \end{aligned} \tag{15}$$

Let there exist a positive definite solution, $P = P^T$, of the output algebraic Riccati-type equation

$$A^T P + PA + Q - PBR^{-1}B^T P + M^T R^{-1} M = 0, \tag{16}$$

where $Q, R > 0$. Then, the output feedback,

$$u = -Ky = -KCx, \tag{17}$$

with the output gain, K, satisfying

$$KC = R^{-1}(B^T P + M), \tag{18}$$

for some matrix M, stabilizes the linear time invariant system (15), *i.e.*, the closed loop system matrix $A - BKC$ is Hurwitz.

J. Gadewadikar, Frank L. Lewis, L. Xie, V. Kucera and M. Abu-Khalaf, "Parameterization of all stabilizing H_∞ static state-feedback gains: Application to output-feedback design," *Automatica*, vol. 43, no. 9, pp. 1597-1604, September 2007.

J. Gadewadikar, M. Abu-Khalaf, and F.L. Lewis, "Necessary and sufficient conditions for H-infinity static output-feedback control," J. Guidance, Control, and Dynamics, vol. 29, no. 4, pp. 915-920, July-Aug. 2006.

Synchronizing Region for Static OPFB

Definition 2. Given matrices (A,B,C), and the output feedback gain K, the synchronizing region for the matrix pencil $A-\sigma BKC$ is a subset of the complex plane, $S_y=\{\sigma\in\mathbb{C}: A-\sigma BKC \text{ is stable}\}$. This we call the *synchronizing region for output feedback.*

Hence, the matrix (11) is stable if and only if all the scaled graph matrix eigenvalues, $c\lambda_j$, are in the synchronizing region for output feedback of the matrix pencil

$$A-\sigma BKC. \quad (12)$$

Recall the UNBOUNDED synchronizing region for local optimal SVFB design

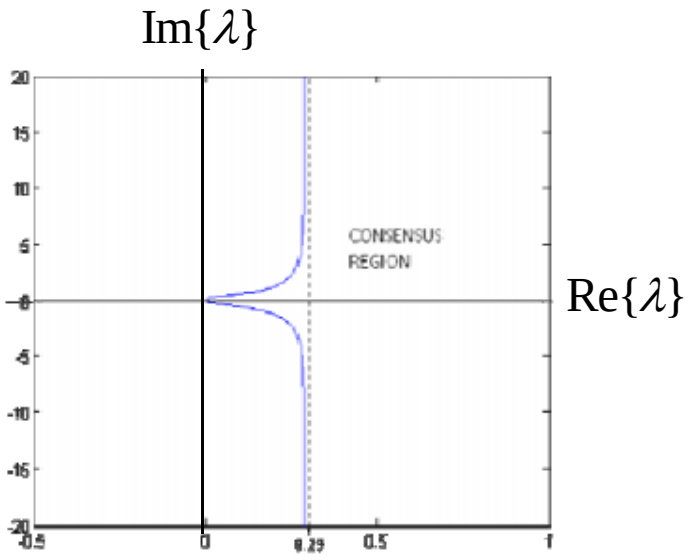


Figure 2. Unbounded consensus region for optimal SVFB gain

Synch. Region for Locally Optimal OPFB Design

$$A^T P + PA + Q - PBR^{-1}B^T P + M^T R^{-1} M = 0$$

$$KC = R^{-1}(B^T P + M)$$

$$u = -Ky = -KCx$$

$$A - \sigma BKC$$

Theorem 1. Let the multi-agent system be given by (1), (2). Let the graph have a spanning tree with at least one non-zero pinning gain connecting to a root note. Choose the distributed static output feedback control (4), where the output feedback gain, K, satisfies (18). Let the following abbreviations be introduced

$$a := \lambda_{\min}(Q^{-T/2} C^T K^T R K C Q^{-1/2}) > 0, \tag{19}$$

where $\lambda_{\min}$ denotes the smallest eigenvalue on the complement of the kernel of $KCQ^{-1/2}$,

$$2b := \left\| Q^{-T/2} C^T K^T M Q^{-1/2} \right\| + \left\| Q^{-T/2} M^T K C Q^{-1/2} \right\| \geq 0, \tag{20}$$

and

$$A_1 := 4(a^2 - b^2),\ A_2 := 4(a - a^2 + 2b^2),\ A_3 := 4b^2,\ A_4 := (1-a)^2 - (2b)^2. \tag{21}$$

The form of the guaranteed synchronizing region for output feedback depends on the matrix M, in that

i. The guaranteed synchronizing region for output feedback is an interior of an elliptic region in $\mathbb{C}$ if $A_1 < 0$ and $A_4 - \frac{A_2^2}{4A_1} > 0$.

ii. The guaranteed synchronizing region for output feedback reduces to the empty set if $A_1 < 0$ and $A_4 - \frac{A_2^2}{4A_1} < 0$.

iii. The guaranteed synchronizing region for output feedback is a hyperbolic region in $\mathbb{C}$, having two connected components, if $A_1 > 0$ and $A_4 - \frac{A_2^2}{4A_1} < 0$.

iv. The guaranteed synchronizing region for output feedback is a hyperbolic region in $\mathbb{C}$, having a single connected component if $A_1 > 0$ and $A_4 - \frac{A_2^2}{4A_1} > 0$.

a depends on the optimal OPFB design

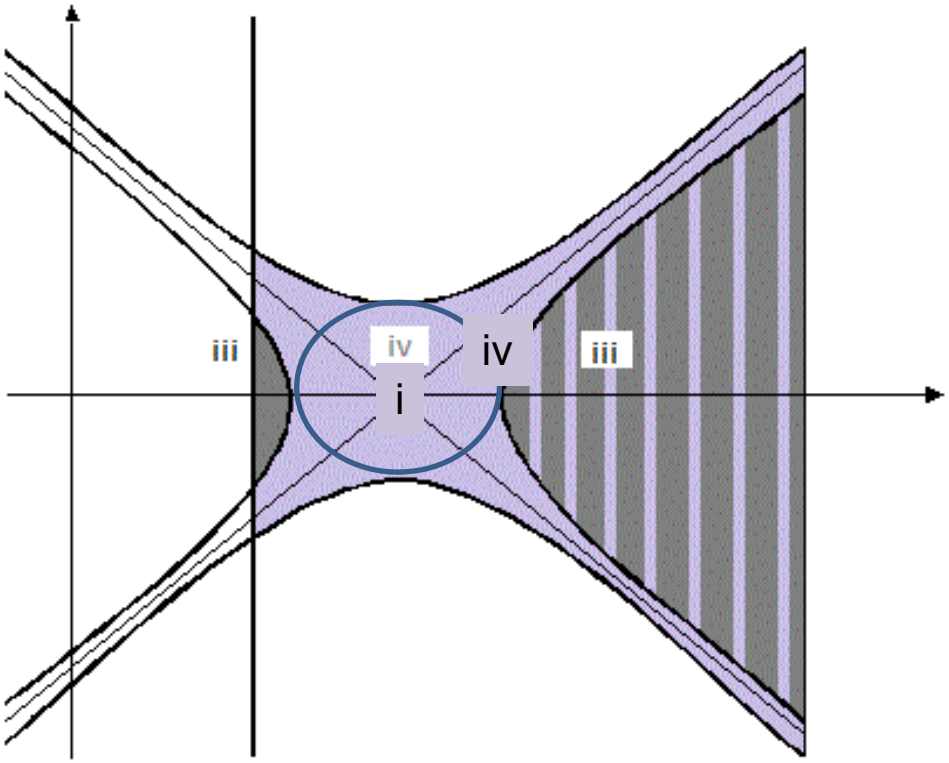
b is small if *M* is small

Kristian Hengster-Movric

Synch. Region for Locally Optimal OPFB Design

Bounded by hyperbolae

$$A^T P + PA + Q - PBR^{-1}B^T P + M^T R^{-1} M = 0$$



iii. The guaranteed synchronizing region for output-feedback is a hyperbolic region in $\mathbb{C}$, having two connected components, if $A_1 > 0$ and

$$A_4 - \frac{A_2^2}{4A_1} < 0 .$$

iv. The guaranteed synchronizing region for output-feedback is a hyperbolic region in $\mathbb{C}$, having a single connected component if $A_1 > 0$ and

$$A_4 - \frac{A_2^2}{4A_1} > 0 .$$

Synch. region is
unbounded in case iii or iv.
M is not too big

Remark 1. Notice that in case $M = 0$ one recovers the unbounded synchronizing region characteristic of the full state feedback, [12],[13],

$$-Q + (1 - 2\,\mathrm{Re}\,\sigma) C^T K^T R K C < 0 .$$

Cooperative H_∞ Synchronization

System

$$\dot{x}_i = Ax_i + Bu_i + Dd_i$$
$$y_i = Cx_i$$

leader

$$\dot{x}_0 = Ax_0$$
$$y_0 = Cx_0$$

$$\dot{\delta} = (I \otimes A)\delta + (I \otimes B)u + (I \otimes D)\mathrm{d}$$
$$\mathrm{y} = (\mathrm{I} \otimes C)\delta$$

$$\delta = x - \underline{I}x_0$$

Performance output $\quad \|z(t)\|^2 = \delta^T Q\delta + u^T Ru$

Bounded L_2 synchronization error

$$\frac{\int_0^\infty \|z(t)\|^2 \, dt}{\int_0^\infty \|d(t)\|_T^2 \, dt} = \frac{\int_0^\infty (\delta^T Q\delta + u^T Ru)dt}{\int_0^\infty d(t)^T Td(t)dt} \le \gamma^2$$

H-inf performance index $\quad J(\delta, u, d) = \int_0^\infty (\delta^T Q\delta + u^T Ru - \gamma^2 d^T Td)dt$

Note we introduced the matrix *T*. It is required for coop *H-inf* synchronization.

Coop. nbhd OPFB

$$u_i = cKC\varepsilon_i = cKe_{yi} = cK\left(\sum_{j\in N_i} e_{ij}(y_j - y_i) \quad + g_i(y_0 - y_i)\right)$$

Closed loop system $\dot{x}_i = Ax_i + Bu_i + Dd_i = Ax_i + cBKC\left(\sum_{j\in N_i} e_{ij}(x_j - x_i) \quad + g_i(x_0 - x_i)\right) + Dd_i$

Overall state $x = \begin{bmatrix} x_1^T & x_2^T & \cdots & x_N^T \end{bmatrix}^T \in R^{nN}, \qquad \delta = \left(x - \underline{I}x_0\right)$

Distributed form of control $\boxed{u = -c\left((L+G)\otimes KC\right)\delta}$

Overall c.l. dynamics

$$\dot{x} = \left[(I_N \otimes A) - c(L+G)\otimes BKC\right]x + c\left[(L+G)\otimes BKC\right]\underline{x}_0 + (I_N \otimes D)d$$

Global synch. error dynamics $\dot{\delta} = \left[(I_N \otimes A) - c(L+G)\otimes BKC\right]\delta + (I_N \otimes D)d$

Graph structure $\otimes$ Control structure

WANT TO DECOUPLE
CONTROL DESIGN AND GRAPH STRUCTURE

Cooperative H_∞ Synchronization

Kristian Hengster-Movric

Theorem 2. Let the synchronization error dynamics with disturbances acting upon the agents be given in global form as (29). Let the graph have a spanning tree with at least one non-zero pinning gain connecting to a root node. Suppose there exist matrices P_1, P_2, symmetric and positive definite, satisfying

$$P_1 = cR_1 (L+G) \tag{36}$$

$$A^T P_2 + P_2 A + Q_2 - P_2 B R_2^{-1} B^T P_2 + \gamma^{-2} P_2 D D^T P_2 + M_2^T R_2^{-1} M_2 = 0 \tag{37}$$

for some $Q_2 = Q_2^T > 0$, $R_1 = R_1^T > 0$, $R_2 = R_2^T > 0$, and the coupling gain, $c > 0$. Define the output feedback gain matrix, $K_2 \equiv K_2^*$ as the one satisfying

$$K_2 C = R_2^{-1}(B^T P_2 + M_2), \tag{38}$$

then the control

$$u = -c(L+G) \otimes K_2 C\delta \tag{39},$$

is the solution of the H_∞ static output distributed feedback problem, guaranteeing boundedness of the L_2 gain (30) by γ, defined by the matrices Q, R, T given as

$$R = R_1 \otimes R_2,$$

and either

i. $T = P_1 \otimes I = cR_1(L+G) \otimes I$ with

$$Q = c^2 (L+G)^T R_1 (L+G) \otimes (Q_2 + A^T P_2 + P_2 A + \gamma^2 P_2 D D^T P_2) - cR_1 (L+G) \otimes (A^T P_2 + P_2 A + \gamma^2 P_2 D D^T P_2) > 0 \tag{40}$$

or

ii. $T = R_1 \otimes I$ with

$$Q = c^2 (L+G)^T R_1 (L+G) \otimes (Q_2 + A^T P_2 + P_2 A) - cR_1 (L+G) \otimes (A^T P_2 + P_2 A) > 0 \tag{41}$$

Condition on graph topology

$$R_1 (L+G) = \mathrm{P}_1$$

Condition for existence of local OPFB

$$R_2 K_2 C = B^T P_2 + M_2$$

Symmetry between OPFB and Graph Topology Information Restrictions

$$e_y = -\left(\left(L+G\right)\otimes C\right)\delta$$

Coarse global OPFB structure

Fine Local OPFB structure

Condition on graph topology $R_1\,(L+G) = \mathrm{P}_1$

Condition for existence of local OPFB $R_2 K_2 C = B^T P_2 + M_2$

Same Same

2-player Zero-sum Cooperative Nash Equilibrium

Command generator

$$\dot{x}_0 = Ax_0,$$
$$y_0 = Cx_0,$$

Agent Dynamics

$$\dot{x}_i = Ax_i + Bu_i + Dd_i,$$
$$y_i = Cx_i,$$

Global Dynamics

$$\dot{\delta} = (I_N \otimes A)\delta + (I_N \otimes B)u + (I_N \otimes D)d,$$
$$y = (I_N \otimes C)\delta.$$

H-inf performance index

$$J(\delta, u, d) = \int_0^\infty (\delta^T Q\delta + u^T Ru - \gamma^2 d^T Td)dt$$

worse. The two-player zero-sum game has a unique solution if a game theoretic saddle point, (u^*, d^*), exists, that is if the value of the performance criterion at the Nash equilibrium, $V(x_0) = J(x_0, u^*, d^*)$, satisfies

$$V^*(x_0) = \min_u \max_d J(x(0), u, d) = \max_d \min_u J(x(0), u, d)$$

Nash Equilibrium

Definition 3. Given a performance criterion, $J(x_0, u, d)$, the policies u^*, d^* are in the Nash equilibrium if

$$J(x(0), u^*, d) \le J(x(0), u^*, d^*) \le J(x(0), u, d^*).$$

$$J(\delta,u,d)=\int_0^\infty(\delta^T Q\delta+u^T Ru-\gamma^2 d^T Td)dt$$

$$P_1=cR_1\ (L+G)$$

$$A^T P_2+P_2A+Q_2-P_2BR_2^{-1}B^T P_2+\gamma^{-2}P_2DD^T P_2+M_2^T R_2^{-1}M_2=0$$

$$K_2^*C=R_2^{-1}(B^T P_2+M_2)$$

Theorem 3. Under the hypotheses of Theorem 2, suppose that in addition matrix M_2 satisfies

$$C^T(K_2-K_2^*)^T R_2(K_2-K_2^*)C+M_2^T(K_2-K_2^*)C+C^T(K_2-K_2^*)^T M_2\geq 0$$

for any $K_2\neq K_2^*$

Then a Nash equilibrium is provided by

$$u^*=-c(L+G)\otimes K_2^*C\delta$$

Optimal *H-inf* control is distributed

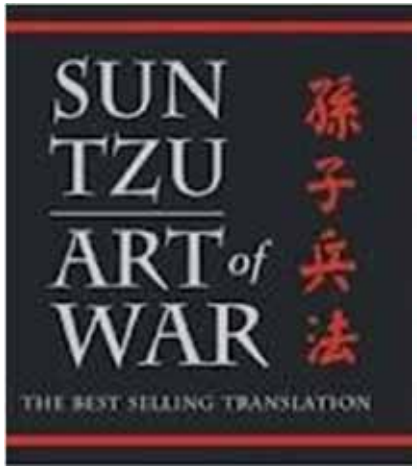
$$d^*=\gamma^{-2}T^{-1}D^T P\delta$$

Worst case disturbance is NOT distributed

with *Q,R,T* given in Theorem 2 and $P=P_1\otimes P_2$


Games on Communication Graphs
Kyriakos Vamvoudakis, Mohammed Abouheaf

Graphical Coalitional Games









500 BC

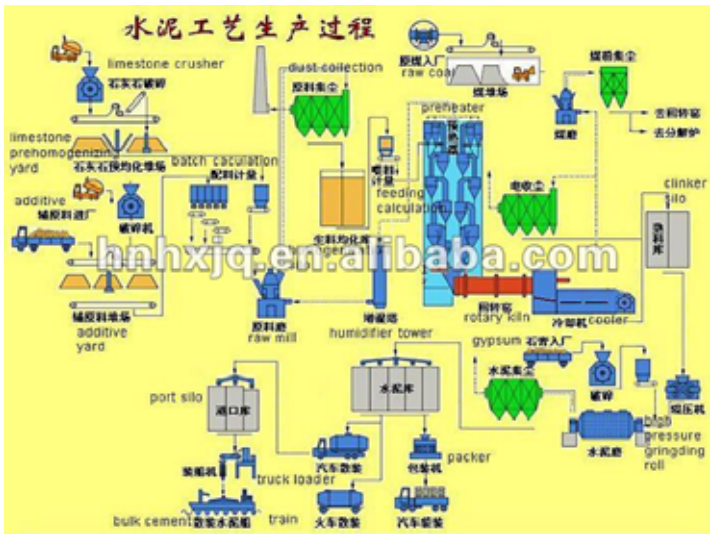
孙**子兵法**

Sun Tz bin fa

Manufacturing as the Interactions of Multiple Agents

Each machine has it own dynamics and cost function
Neighboring machines influence each other most strongly
There are local optimization requirements as well as global necessities

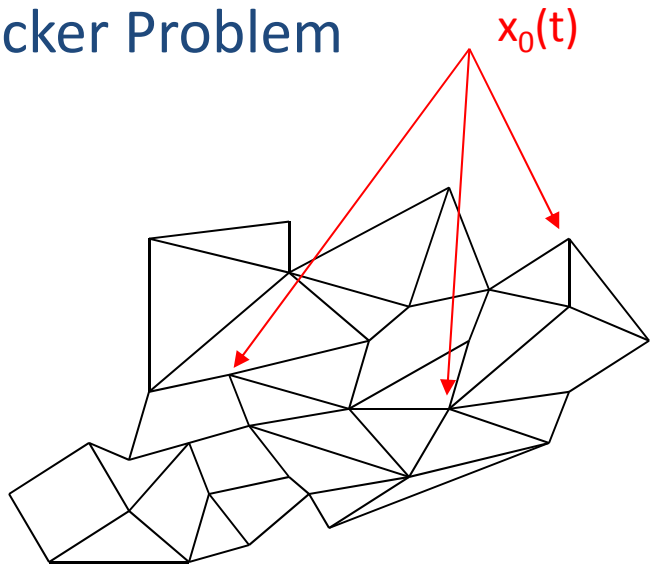








Graphical Games
Synchronization- Cooperative Tracker Problem



Node dynamics

$$\dot{x}_i = Ax_i + B_i u_i, \quad x_i(t) \in \mathbb{R}^n, \quad u_i(t) \in \mathbb{R}^{m_i}$$

Target generator dynamics

$$\dot{x}_0 = Ax_0$$

Synchronization problem

$$x_i(t) \to x_0(t), \forall i$$

Local neighborhood tracking error (Lihua Xie)

$$\delta_i = \sum_{j \in N_i} e_{ij}(x_i - x_j) + g_i(x_i - x_0),$$ Pinning gains $g_i \geq 0$ (Ron Chen)

K.G. Vamvoudakis, F.L. Lewis, and G.R. Hudas, "Multi-Agent Differential Graphical Games: online adaptive learning solution for synchronization with optimality," Automatica, vol. 48, no. 8, pp. 1598-1611, Aug. 2012.

Graphical Game: Games on Graphs

Kyriakos Vamvoudakis

Local nbhd. tracking error dynamics

$$\dot{\delta}_i = \sum_{j\in N_i} e_{ij}(\dot{x}_i - \dot{x}_j) \quad + g_i(\dot{x}_i - \dot{x}_0)$$

$$\dot{\delta}_i = A\delta_i + (d_i + g_i)B_i u_i - \sum_{j\in N_i} e_{ij}B_j u_j$$

Local agent dynamics driven by neighbors' controls

Define Local nbhd. performance index

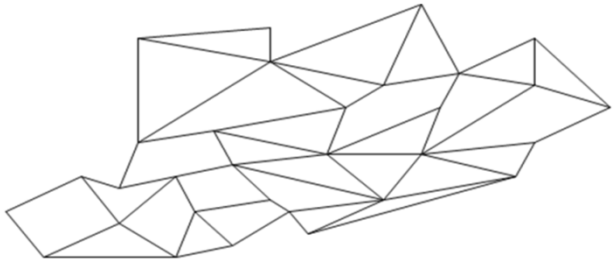
$$J_i(\delta_i(0), u_i, u_{-i}) = \tfrac{1}{2}\int_0^\infty (\delta_i^T Q_{ii}\delta_i + u_i^T R_{ii}u_i + \sum_{j\in N_i} u_j^T R_{ij}u_j)\,dt \quad \equiv \tfrac{1}{2}\int_0^\infty L_i(\delta_i(t), u_i(t), u_{-i}(t))\,dt$$

$$u_{-i}(t) \equiv \{u_j : j \in N_i\}$$

Values driven by neighbors' controls

Local value functions for fixed policies u_i

$$V_i(\delta_i(t)) = \tfrac{1}{2}\int_t^\infty (\delta_i^T Q_{ii}\delta_i + u_i^T R_{ii}u_i + \sum_{j\in N_i} u_j^T R_{ij}u_j)\,dt$$



Static Graphical Game

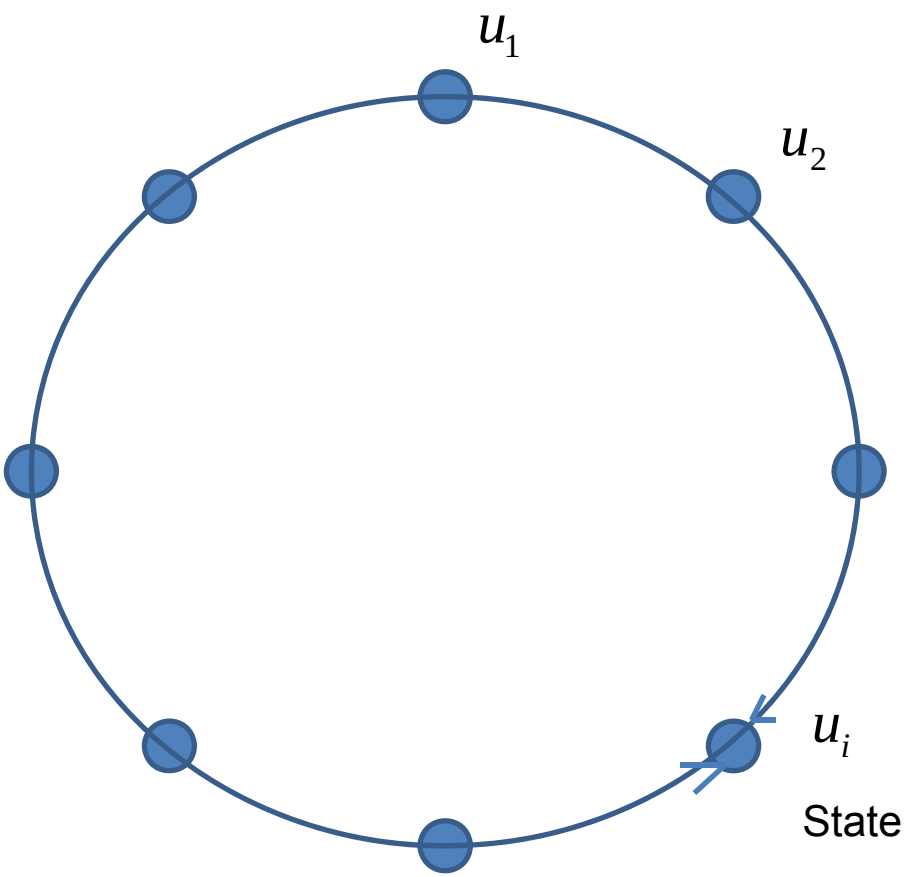
$$(G, U, v)$$

$$G = (V, E), \qquad v = \begin{bmatrix} v_1 & \cdots & v_N \end{bmatrix}^T$$

$$v_i(U_i, \{U_j : j \in N_i\}) \in R$$

Value depends only on neighbors

New Differential Graphical Game



Local Dynamics
Local Value Function
Only depends on
graph neighbors

u_i Control action of player *i*

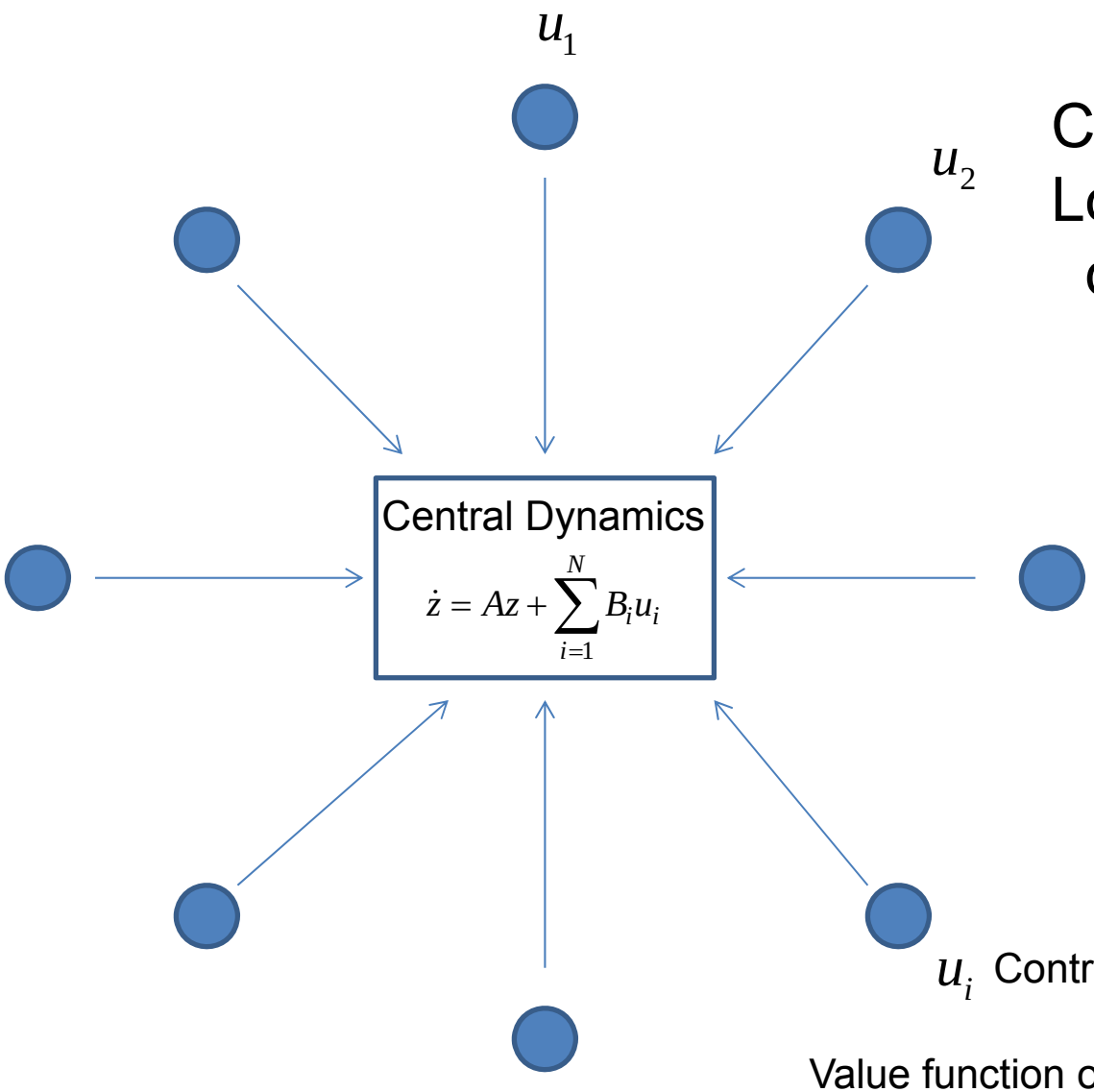
State dynamics of agent *i*

$$\dot{\delta}_i = A\delta_i + (d_i + g_i)B_i u_i - \sum_{j\in N_i} e_{ij} B_j u_j$$

Value function of player *i*

$$J_i(\delta_i(0), u_i, u_{-i}) = \tfrac{1}{2}\int_0^\infty (\delta_i^T Q_{ii}\delta_i + u_i^T R_{ii} u_i + \sum_{j\in N_i} u_j^T R_{ij} u_j)\, dt$$

Standard Multi-Agent Differential Game



Central Dynamics
Local Value Function
depends on ALL
other control actions

u_i Control action of player *i*

Value function of player *i*

$$J_i(z(0), u_i, u_{-i}) = \frac{1}{2}\int_0^{\infty} (z^T Q z + \sum_{j=1}^{N} u_j^T R_{ij} u_j)\, dt$$

Problems with Nash Equilibrium Definition on Graphical Games

Game objective

$$V_i^*(\delta_i(t)) = \min_{u_i} \int_t^\infty \tfrac{1}{2}(\delta_i^T Q_{ii}\delta_i + u_i^T R_{ii} u_i + \sum_{j\in N_i} u_j^T R_{ij} u_j)\, dt$$

Define $u_{-i}(t), \quad \{u_j : j \in N_i\}$ Neighbors of node *i*

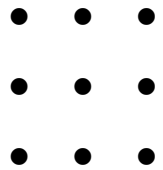
$u_{G-i} = \{u_j : j \in N, j \neq i\}$ All other nodes in graph

Def: Nash equilibrium

$\{u_1^*, u_2^*, \ldots, u_N^*\}$ are in Nash equilibrium if

$$J_i^* \triangleq J_i(u_i^*, u_{G-i}^*) \leq J_i(u_i, u_{G-i}^*), \ \forall i \in N$$

Counterexample. Disconnected graph



Then, each agent's cost does not depend on any other agent

$$J_i(u_i) = J_i(u_i, u_{G-i}) = J_i(u_i, u'_{G-i}), \ \forall i$$

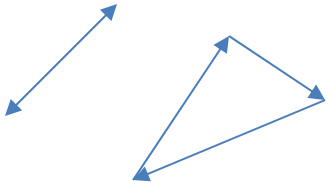
Let each node play his optimal control

$$J_i^* = J_i(u_i^*)$$

Then all agents are in Nash equilibrium

Note- this Nash is also coalition-proof

Another example



New Definition of Nash Equilibrium for Graphical Games

Def. Local Best response.

u_i^* is said to be agent *i's* local best response to fixed policies u_{-i} of its neighbors if

$$J_i(u_i^*, u_{-i}) \leq J_i(u_i, u_{-i}), \quad \forall u_i$$

Def: Interactive Nash equilibrium

$\{u_1^*, u_2^*, \ldots, u_N^*\}$ are in Interactive Nash equilibrium if

1. $J_i^* \triangleq J_i(u_i^*, u_{G-i}^*) \leq J_i(u_i, u_{G-i}^*), \ \forall i \in N$ — They are in Nash equilibrium

2. There exists a policy u_j such that — Interaction Condition

$$J_i(u_j, u_{G-j}^*) \neq J_i(u_j^*, u_{G-j}^*), \quad \forall i, j \in N$$

That is, every player can find a policy that changes the value of every other player.

A restriction on what sorts of performance indices can be selected in multi-player graph games.

A condition on the reaction curves (Basar and Olsder) of the agents

This rules out the disconnected counterexample.

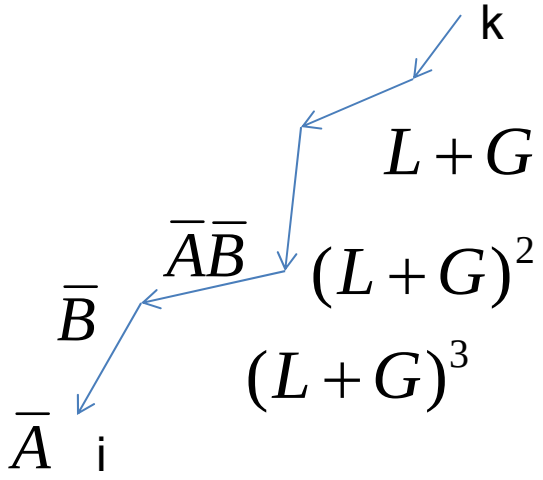
Theorem 3. Let (A,B_i) be reachable for all i.
Let agent i be in local best response

$$J_i\,(u_i^*, u_{-i}) \le J_i\,(u_i, u_{-i}), \quad \forall i$$

Then $\{u_1^*, u_2^*, \ldots, u_N^*\}$ are in global Interactive Nash iff the graph is strongly connected.

$$u_i = u_i(V_i) \equiv -(d_i + g_i) R_{ii}^{-1} B_i^T \frac{\partial V_i}{\partial \delta_i} \equiv -K_i\, p_i$$

$$u_k = -K_k\, p_k - v_k$$



Hamiltonian System

$$\begin{bmatrix} \dot{\delta} \\ \dot{p} \end{bmatrix} = \begin{bmatrix} (I_N \otimes A) & ((L+G)\otimes I_n)\,diag(B_i K_i) \\ -diag(Q_{ii}) & -(I_N \otimes A^T) \end{bmatrix} \begin{bmatrix} \delta \\ p \end{bmatrix} + \begin{bmatrix} ((L+G)\otimes I_n)\underline{B}_k \\ 0 \end{bmatrix} \bar{v}_k \equiv \bar{A} \begin{bmatrix} \delta \\ p \end{bmatrix} + \bar{B}\bar{v}_k$$

$$\begin{bmatrix} \bar{B} & \bar{A}\bar{B} & \bar{A}^2\bar{B} & \cdots \end{bmatrix}$$

Picks out the shortest path from node k to node i

Graphical Game Solution Equations

Value function

$$V_i(\delta_i(t)) = \tfrac{1}{2}\int_t^\infty (\delta_i^T Q_{ii}\delta_i + u_i^T R_{ii} u_i + \sum_{j\in N_i} u_j^T R_{ij} u_j)\, dt$$

Differential equivalent (Leibniz formula) is Bellman's Equation

$$H_i(\delta_i, \frac{\partial V_i}{\partial \delta_i}, u_i, u_{-i}) \equiv \frac{\partial V_i}{\partial \delta_i}^T \left(A\delta_i + (d_i + g_i) B_i u_i - \sum_{j\in N_i} e_{ij} B_j u_j \right) + \tfrac{1}{2}\delta_i^T Q_{ii}\delta_i + \tfrac{1}{2} u_i^T R_{ii} u_i + \tfrac{1}{2}\sum_{j\in N_i} u_j^T R_{ij} u_j = 0$$

Stationarity Condition

$$0 = \frac{\partial H_i}{\partial u_i} \quad \Rightarrow \quad u_i = -(d_i + g_i) R_{ii}^{-1} B_i^T \frac{\partial V_i}{\partial \delta_i}$$

1. Coupled HJ equations

$$\frac{\partial V_i}{\partial \delta_i}^T A_i^c + \tfrac{1}{2}\delta_i^T Q_{ii}\delta_i + \tfrac{1}{2}(d_i + g_i)^2 \frac{\partial V_i}{\partial \delta_i}^T B_i R_{ii}^{-1} B_i^T \frac{\partial V_i}{\partial \delta_i} + \tfrac{1}{2}\sum_{j\in N_i} (d_j + g_j)^2 \frac{\partial V_j}{\partial \delta_j}^T B_j R_{jj}^{-1} R_{ij} R_{jj}^{-1} B_j^T \frac{\partial V_j}{\partial \delta_j} = 0, i \in N$$

$$H_i(\delta_i, \frac{\partial V_i}{\partial \delta_i}, u_i^*, u_{-i}^*) = 0$$

where $A_i^c = A\delta_i - (d_i + g_i)^2 B_i R_{ii}^{-1} B_i^T \frac{\partial V_i}{\partial \delta_i} + \sum_{j\in N_i} e_{ij}(d_j + g_j) B_j R_{jj}^{-1} B_j^T \frac{\partial V_j}{\partial \delta_j}, i \in N$

---

2. Best Response HJ Equations – other players have fixed policies u_j

$$0 = H_i(\delta_i, \frac{\partial V_i}{\partial \delta_i}, u_i^*, u_{-i}) \equiv \frac{\partial V_i}{\partial \delta_i}^T A_i^c + \tfrac{1}{2}\delta_i^T Q_{ii}\delta_i + \tfrac{1}{2}(d_i + g_i)^2 \frac{\partial V_i}{\partial \delta_i}^T B_i R_{ii}^{-1} B_i^T \frac{\partial V_i}{\partial \delta_i} + \tfrac{1}{2}\sum_{j\in N_i} u_j^T R_{ij} u_j$$

where $A_i^c = A\delta_i - (d_i + g_i)^2 B_i R_{ii}^{-1} B_i^T \frac{\partial V_i}{\partial \delta_i} - \sum_{j\in N_i} e_{ij} B_j u_j$

Theorem 1. Stability and Solution for Cooperative Nash Equilibrium.

Let $V_i > 0 \in C^1$, $i \in N$ be smooth solutions to HJ equations (23) and control policies u_i^*, $i \in N$ be given by (22) in terms of these solutions V_i. Then

a. Systems (8) are asymptotically stable.

b. u_i^*, u_{-i}^* are in cooperative Nash equilibrium and the corresponding game values are

$$J_i^*(\delta_i(0)) = V_i \ , i \in N \qquad (34)$$

Kyriakos Vamvoudakis

Theorem 2. Solution for Best Response Policy

Given fixed neighbor policies $u_{-i} = \{u_j : j \in N_i\}$, assume there is an admissible policy u_i. Let $V_i > 0 \in C^1$ be a smooth solution to the best response HJ equation (36) and let control policy u_i^* be given by (22) in terms of this solution V_i. Then

a. System (8) is asymptotically stable.

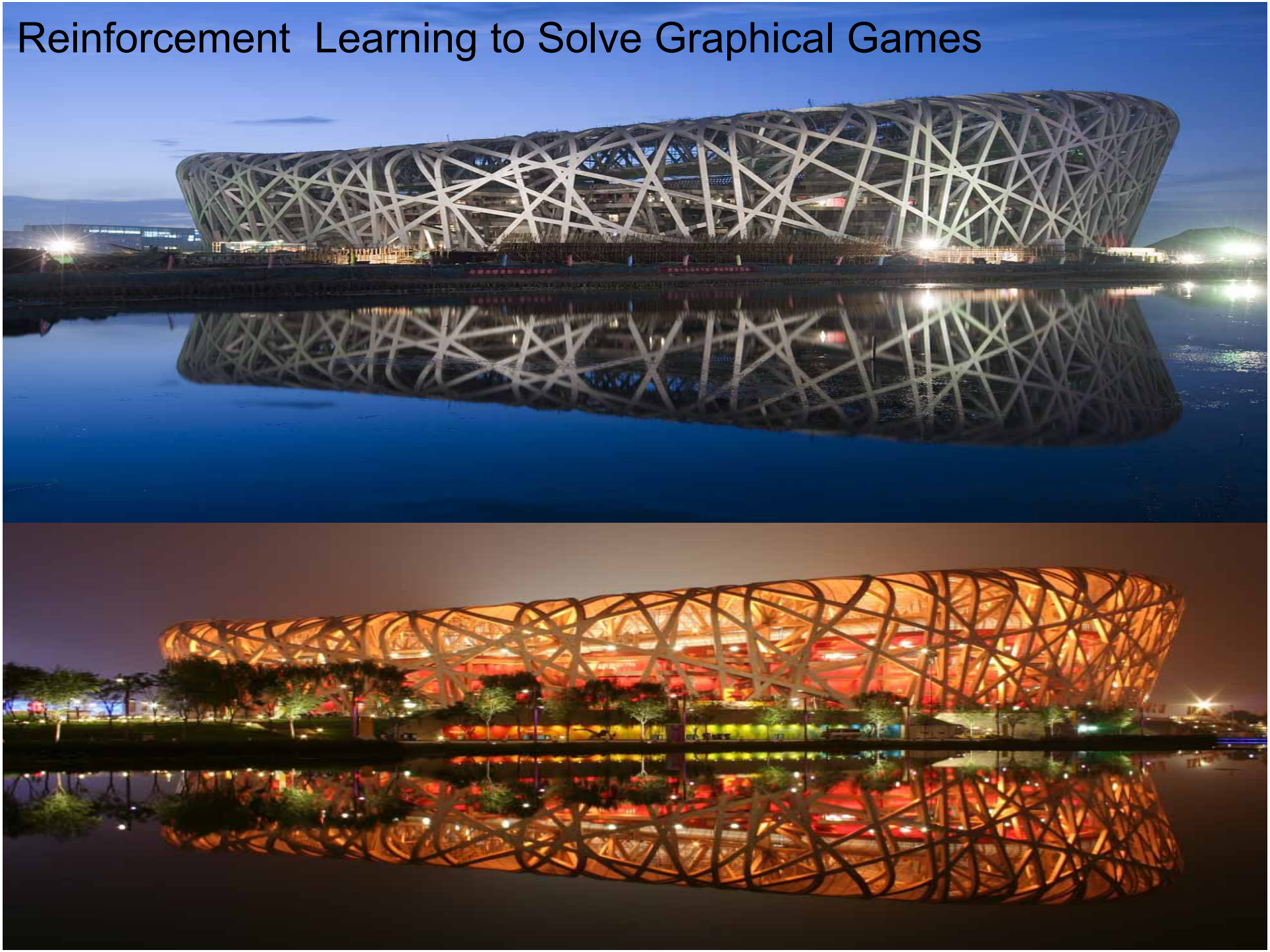
b. u_i^* is the best response to the fixed policies u_{-i} of its neighbors.

How to solve Graphical Games

Online in real time
Using data measured from the system
Without knowing full system dynamics

??

Reinforcement Learning to Solve Graphical Games





Adaptive Control is online and works for unknown systems.
Generally not Optimal

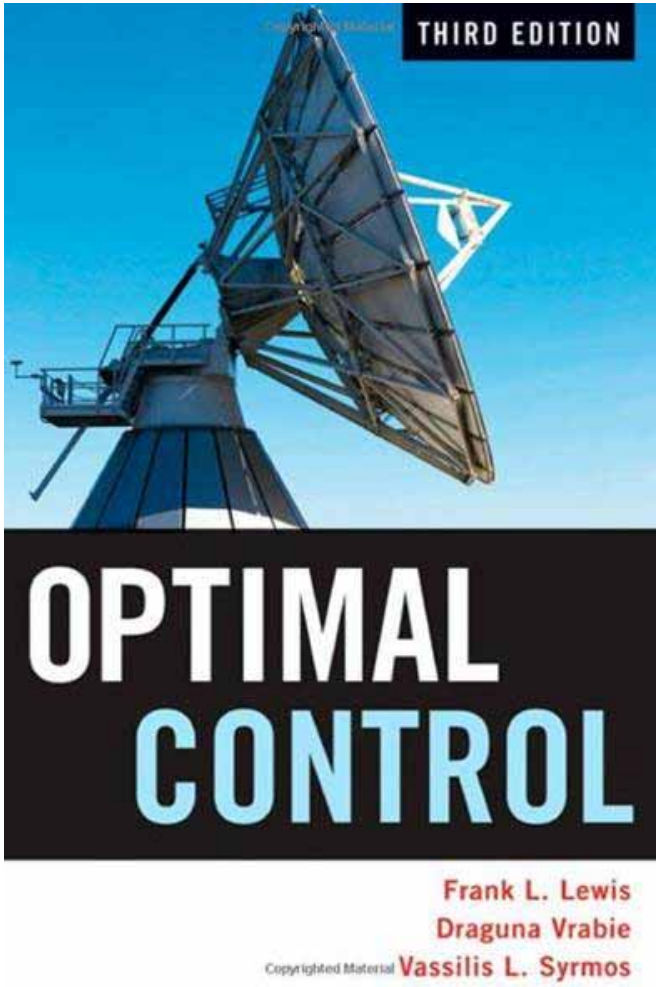
Optimal Control is off-line,
and needs to know the system dynamics to solve design eqs.

We want ONLINE DIRECT ADAPTIVE OPTIMAL Control
For any performance cost of our own choosing

Reinforcement Learning turns out to be the key to this!

F. Lewis, Draguna Vrabie, and Vassilis Syrmos, "Optimal Control," 3rd edition, 2012. Chapter 11

F.L. Lewis and D. Vrabie, "Reinforcement learning and adaptive dynamic programming for feedback control," IEEE Circuits & Systems Magazine, Invited Feature Article, pp. 32-50, Third Quarter 2009.



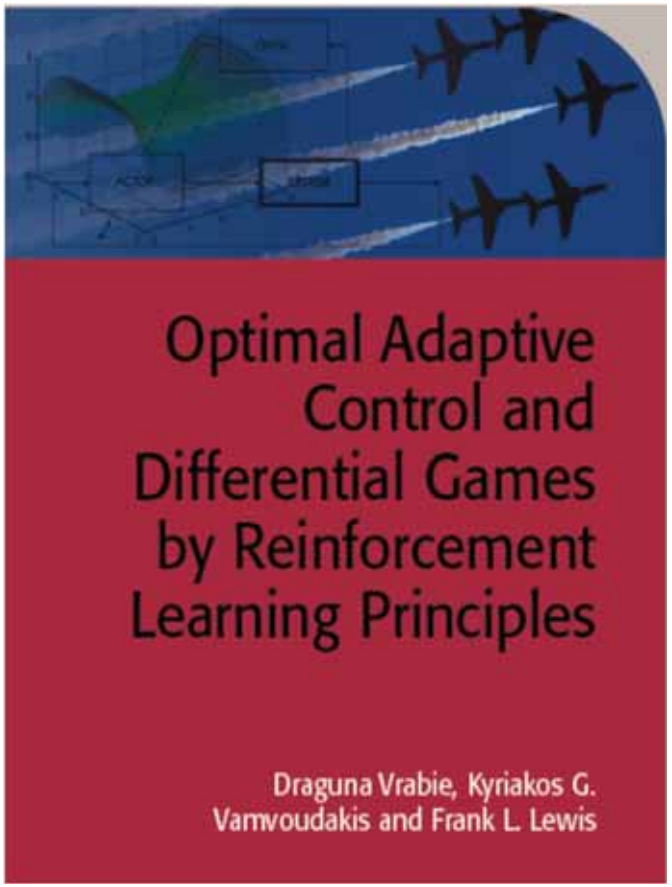
Books

F.L. Lewis, D. Vrabie, and V. Syrmos, *Optimal Control*, third edition, John Wiley and Sons, New York, 2012.

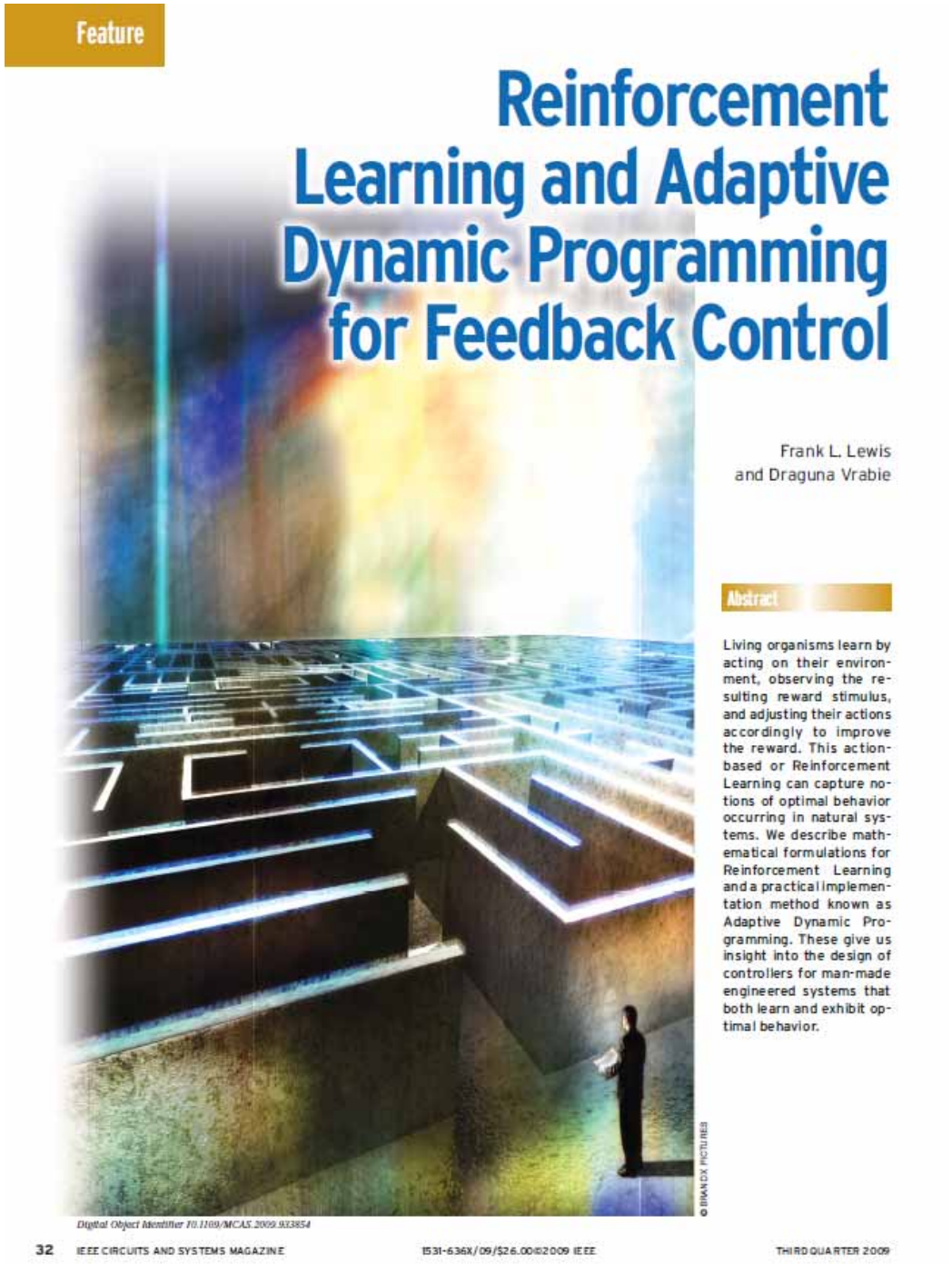
New Chapters on:

Reinforcement Learning

Differential Games



D. Vrabie, K. Vamvoudakis, and F.L. Lewis, *Optimal Adaptive Control and Differential Games by Reinforcement Learning Principles*, IET Press, 2012.



Feature

Reinforcement Learning and Adaptive Dynamic Programming for Feedback Control

Frank L. Lewis
and Draguna Vrabie

Abstract

Living organisms learn by acting on their environment, observing the resulting reward stimulus, and adjusting their actions accordingly to improve the reward. This action-based or Reinforcement Learning can capture notions of optimal behavior occurring in natural systems. We describe mathematical formulations for Reinforcement Learning and a practical implementation method known as Adaptive Dynamic Programming. These give us insight into the design of controllers for man-made engineered systems that both learn and exhibit optimal behavior.

© BRAND X PICTURES

Digital Object Identifier 10.1109/MCAS.2009.933854

 1531-636X/09/$26.00©2009 IEEE

F.L. Lewis and D. Vrabie, "Reinforcement learning and adaptive dynamic programming for feedback control," IEEE Circuits & Systems Magazine, Invited Feature Article, pp. 32-50, Third Quarter 2009.

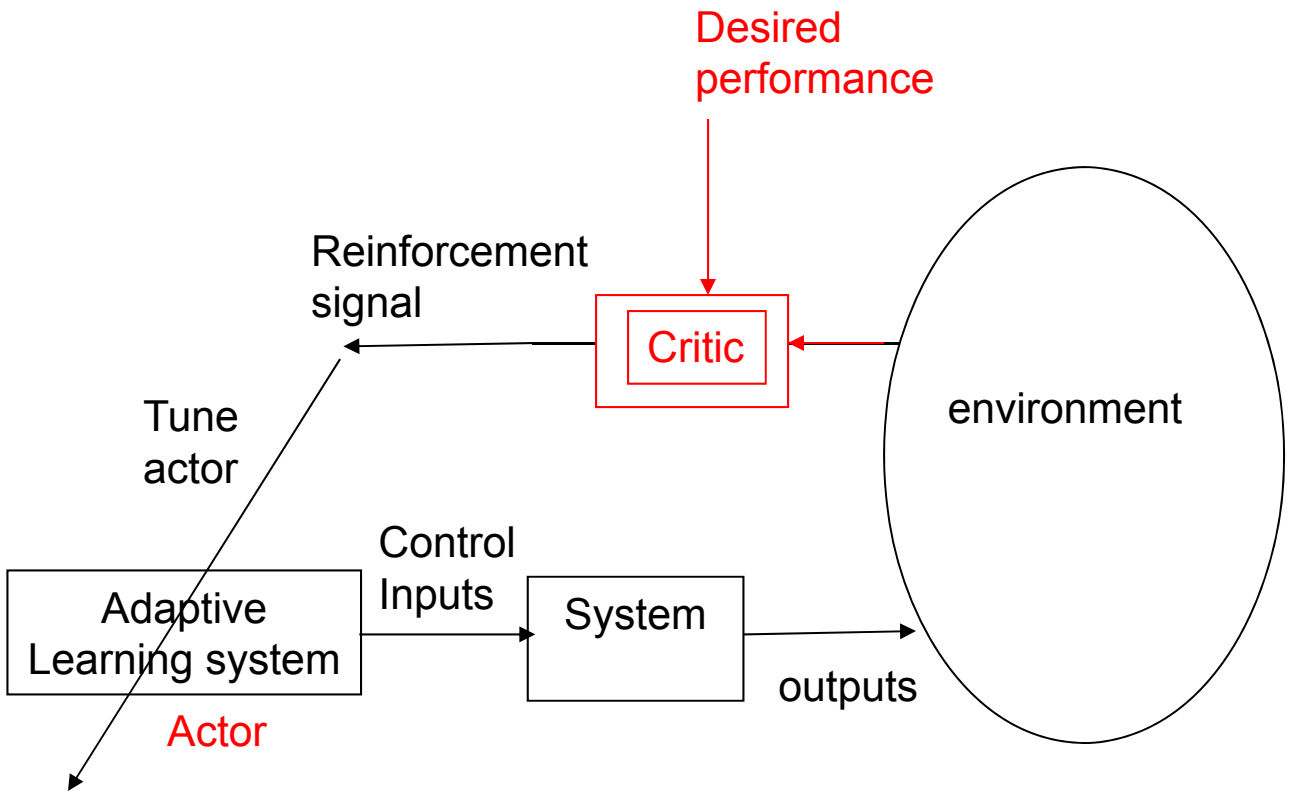
IEEE Control Systems Magazine, F. Lewis, D. Vrabie, and K. Vamvoudakis, "Reinforcement learning and feedback Control," Dec. 2012

Different methods of learning

Reinforcement learning
Ivan Pavlov 1890s

We want OPTIMAL performance
- ADP- Approximate Dynamic Programming

Actor-Critic Learning
Paul Werbos



Sutton & Barto book

Graphical Game Solution Equations

Value function

$$V_i(\delta_i(t)) = \tfrac{1}{2}\int_t^{\infty} (\delta_i^T Q_{ii}\delta_i + u_i^T R_{ii} u_i + \sum_{j\in N_i} u_j^T R_{ij} u_j)\, dt$$

Differential equivalent (Leibniz formula) is Bellman's Equation

$$H_i(\delta_i, \frac{\partial V_i}{\partial \delta_i}, u_i, u_{-i}) \equiv \frac{\partial V_i}{\partial \delta_i}^T \left(A\delta_i + (d_i + g_i) B_i u_i - \sum_{j\in N_i} e_{ij} B_j u_j \right) + \tfrac{1}{2}\delta_i^T Q_{ii}\delta_i + \tfrac{1}{2} u_i^T R_{ii} u_i + \tfrac{1}{2}\sum_{j\in N_i} u_j^T R_{ij} u_j = 0$$

Stationarity Condition

$$0 = \frac{\partial H_i}{\partial u_i} \quad \Rightarrow \quad u_i = -(d_i + g_i) R_{ii}^{-1} B_i^T \frac{\partial V_i}{\partial \delta_i}$$

1. Coupled HJ equations

$$\frac{\partial V_i}{\partial \delta_i}^T A_i^c + \tfrac{1}{2}\delta_i^T Q_{ii}\delta_i + \tfrac{1}{2}(d_i + g_i)^2 \frac{\partial V_i}{\partial \delta_i}^T B_i R_{ii}^{-1} B_i^T \frac{\partial V_i}{\partial \delta_i} + \tfrac{1}{2}\sum_{j\in N_i} (d_j + g_j)^2 \frac{\partial V_j}{\partial \delta_j}^T B_j R_{jj}^{-1} R_{ij} R_{jj}^{-1} B_j^T \frac{\partial V_j}{\partial \delta_j} = 0, i \in N$$

$$H_i(\delta_i, \frac{\partial V_i}{\partial \delta_i}, u_i^*, u_{-i}^*) = 0$$

where $A_i^c = A\delta_i - (d_i + g_i)^2 B_i R_{ii}^{-1} B_i^T \frac{\partial V_i}{\partial \delta_i} + \sum_{j\in N_i} e_{ij}(d_j + g_j) B_j R_{jj}^{-1} B_j^T \frac{\partial V_j}{\partial \delta_j}, i \in N$

---

2. Best Response HJ Equations – other players have fixed policies u_j

$$0 = H_i(\delta_i, \frac{\partial V_i}{\partial \delta_i}, u_i^*, u_{-i}) \equiv \frac{\partial V_i}{\partial \delta_i}^T A_i^c + \tfrac{1}{2}\delta_i^T Q_{ii}\delta_i + \tfrac{1}{2}(d_i + g_i)^2 \frac{\partial V_i}{\partial \delta_i}^T B_i R_{ii}^{-1} B_i^T \frac{\partial V_i}{\partial \delta_i} + \tfrac{1}{2}\sum_{j\in N_i} u_j^T R_{ij} u_j$$

where $A_i^c = A\delta_i - (d_i + g_i)^2 B_i R_{ii}^{-1} B_i^T \frac{\partial V_i}{\partial \delta_i} - \sum_{j\in N_i} e_{ij} B_j u_j$

Online Solution of Graphical Games

Kyriakos Vamvoudakis

Use Reinforcement Learning

POLICY ITERATION

Algorithm 1. Policy Iteration (PI) Solution for N-player distributed games.

Step 0: Start with admissible initial policies $u_i^0, \forall i$.

Step 1: (Policy Evaluation) Solve for V_i^k using (14)

$$H_i(\delta_i, \frac{\partial V_i}{\partial \delta_i}^k, u_i^k, u_{-i}^k) = 0, \forall i = 1, \ldots, N \qquad (38)$$

Step 2: (Policy Improvement) Update the N-tuple of control policies using

$$u_i^{k+1} = \arg\min_{u_i} H_i(\delta_i, \frac{\partial V_i}{\partial \delta_i}^k, u_i, u_{-i}^k), \forall i = 1, \ldots, N$$

which explicitly is

$$u_i^{k+1} = -(d_i + g_i) R_{ii}^{-1} B_i^T \frac{\partial V_i}{\partial \delta_i}^k, \forall i = 1, \ldots, N. \qquad (39)$$

Go to step 1.

On convergence End ■

Convergence Results

Theorem 3. Convergence of Policy Iteration algorithm when only i^{th} agent updates its policy and all players u_{-i} in the neighborhood do not change. Given fixed neighbors policies u_{-i}, assume there exists an admissible policy u_i. Assume that agent i performs Algorithm 1 and the its neighbors do not update their control policies. Then the algorithm converges to the best response u_i to policies u_{-i} of the neighbors and to the solution V_i to the best response HJ equation (36).

The next result concerns the case where all nodes update their policies at each step of the algorithm. Define the relative control weighting as $\rho_{ij} = \bar{\sigma}(R_{jj}^{-1} R_{ij})$, where $\bar{\sigma}(R_{jj}^{-1} R_{ij})$ is the maximum singular value of $R_{jj}^{-1} R_{ij}$.

Theorem 4. Convergence of Policy Iteration algorithm when all agents update their policies. Assume all nodes i update their policies at each iteration of PI. Then for small enough edge weights e_{ij} and ρ_{ij}, μ_i converges to the global Nash equilibrium and for all i, and the values converge to the optimal game values $V_i^k \to V_i^*$.

Online Solution of Graphical Games

Policy Iteration gives structure needed for online graph games

Solve simultaneously online:

Bellman equation

$$H_i(\delta_i, \frac{\partial V_i}{\partial \delta_i}, u_i, u_{-i}) \equiv \frac{\partial V_i}{\partial \delta_i}^T \left(A\delta_i + (d_i + g_i) B_i u_i - \sum_{j \in N_i} e_{ij} B_j u_j \right) + \tfrac{1}{2} \delta_i^T Q_{ii} \delta_i + \tfrac{1}{2} u_i^T R_{ii} u_i + \tfrac{1}{2} \sum_{j \in N_i} u_j^T R_{ij} u_j = 0$$

$$u_i = -(d_i + g_i) R_{ii}^{-1} B_i^T \frac{\partial V_i}{\partial \delta_i}$$

Weierstrass Approximator structures- 2 at each node

$$\hat{V}_i = \hat{W}_i^T \phi_i \qquad \text{critic}$$

$$\hat{u}_i = -\tfrac{1}{2} (d_i + g_i) R_{ii}^{-1} B_i^T \frac{\partial \phi_i}{\partial \delta_i}^T \hat{W}_{i+N} \qquad \text{actor}$$

Bellman equation becomes an algebraic equation in the parameters

$$\sigma_i^{\mathrm{T}} \hat{W}_i + \delta_i^T Q_{ii} \delta_i + \tfrac{1}{4} \hat{W}_{i+N}^T \bar{D}_i \hat{W}_{i+N} + \tfrac{1}{4} \sum_{j \in N_i} (d_j + g_j)^2 \hat{W}_{j+N}^T \frac{\partial \phi_j}{\partial \delta_j} B_j R_{jj}^{-T} R_{ij} R_{jj}^{-1} B_j^T \frac{\partial \phi_j}{\partial \delta_j}^T \hat{W}_{j+N}$$

$$\sigma_i = \frac{\partial \phi_i}{\partial \delta_i} (A\delta_i + (d_i + g_i) B_i \hat{u}_i - \sum_{j \in N_i} e_{ij} B_j \hat{u}_j)$$

Online Solution of Graphical Games Using Value Function Aproximation

Approximate values by a Critic Network at each node *i*

$$\hat{V}_i = \hat{W}_i^T \phi_i$$

Approximate control policies by an Actor Network at each node *i*

$$\hat{u}_{i+N} = -\tfrac{1}{2}(d_i + g_i) R_{ii}^{-1} B_i^T \frac{\partial \phi_i}{\partial \delta_i}^T \hat{W}_{i+N}$$

Tuning law for Critic parameters

$$\dot{\hat{W}}_i = -a_i \frac{\partial E_1}{\partial \hat{W}_i}$$

$$= -a_i \frac{\sigma_i}{(1+\sigma_i^{\mathrm{T}}\sigma_i)^2}[\sigma_i^{\mathrm{T}}\hat{W}_i + \delta_i^T Q_{ii}\delta_i + \tfrac{1}{4}\hat{W}_{i+N}{}^T \bar{D}_i \hat{W}_{i+N} + \tfrac{1}{4}\sum_{j\in N_i}(d_j + g_j)^2 \hat{W}_{j+N}{}^T \frac{\partial \phi_j}{\partial \delta_j} B_j R_{jj}^{-T} R_{ij} R_{jj}^{-1} B_j^T \frac{\partial \phi_j}{\partial \delta_j}^T \hat{W}_{j+N}]$$

Tuning law for Actor parameters

$$\dot{\hat{W}}_{i+N} = -a_{i+N}\{(F_{i+1}\hat{W}_{i+N} - F_i \bar{\sigma}_{i+N}^T \hat{W}_i) - \frac{1}{4}\bar{D}_i \hat{W}_{i+N} \frac{\bar{\sigma}_{i+N}}{m_{si}}^T \hat{W}_i - \frac{1}{4}\hat{W}_{i+N}{}^T \sum_{\substack{j\in N_i \\ j\neq i}} (d_j + g_j)^2 \hat{W}_j \frac{\bar{\sigma}_{i+N}}{m_{s_i}}^T \frac{\partial \phi_j}{\partial \delta_j} B_j R_{jj}^{-T} R_{ij} R_{jj}^{-1} B_j^T \frac{\partial \phi_j}{\partial \delta_j}^T\}$$

Converges to solution of coupled HJ equations, and Nash equilibrium and keeps states stable while learning

Need PE of $\sigma_i = \frac{\partial \phi_i}{\partial \delta_i}(A\delta_i + (d_i + g_i)B_i\hat{u}_i - \sum_{j\in N_i} e_{ij} B_j \hat{u}_j)$

Integral Reinforcement Form of Bellman Equation

Draguna Vrabie

Can avoid knowledge of drift term *f(x) by* using
Integral Reinforcement Learning (IRL)

Then HJ equations are solved online without knowing *f(x)*
Coupled AREs are solved online without knowing *A*

Lemma 1 – Draguna Vrabie

Bellman Equation

$$0 = \left(\frac{\partial V}{\partial x}\right)^T f(x,u) + r(x,u) \equiv H(x, \frac{\partial V}{\partial x}, u), \quad V(0) = 0$$

Is equivalent to

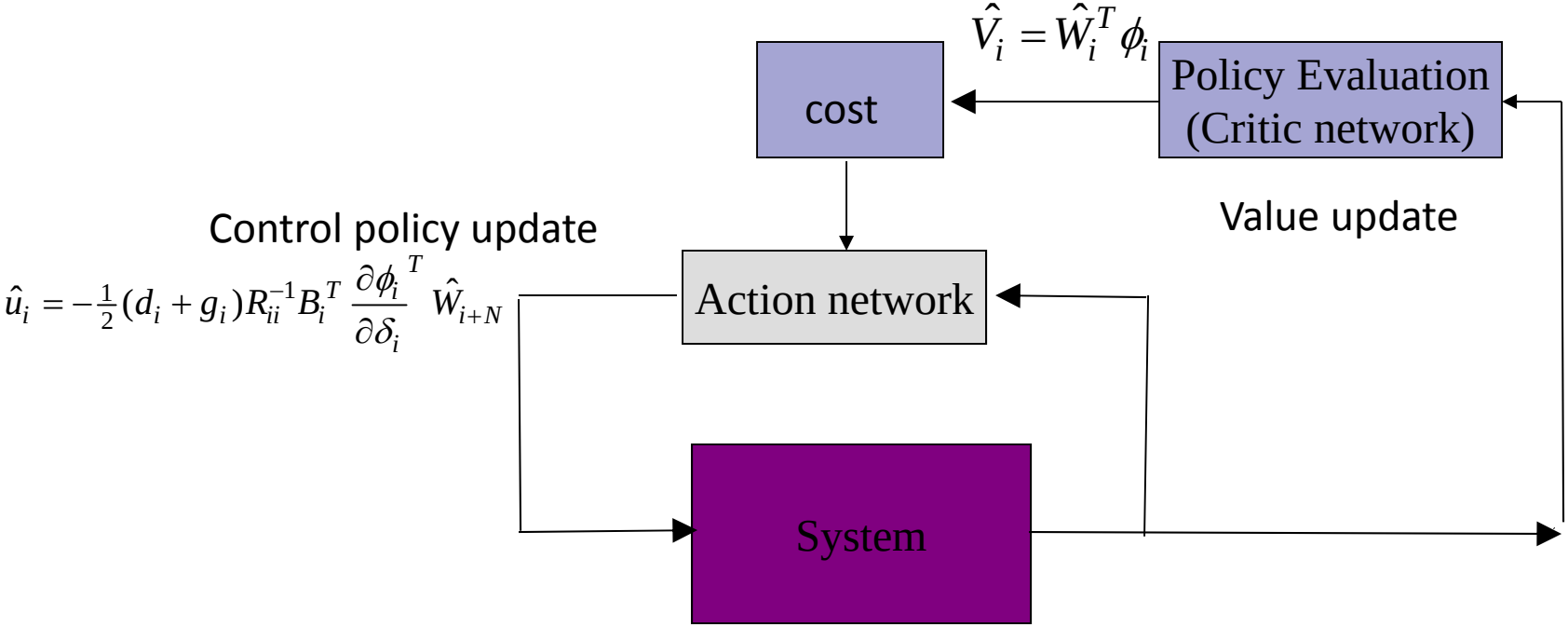
$$V(x(t)) = \int_t^{t+T} r(x,u)\, d\tau \quad + \quad V(x(t+T)), \quad V(0) = 0$$

Another form for the CT Bellman eq.

Solves Lyapunov equation without knowing *f(x,u)*

Online Solution of Graphical Games

New Structure of Adaptive Controller
Reinforcement Learning Adaptive Critic



Critic and Actor tuned simultaneously

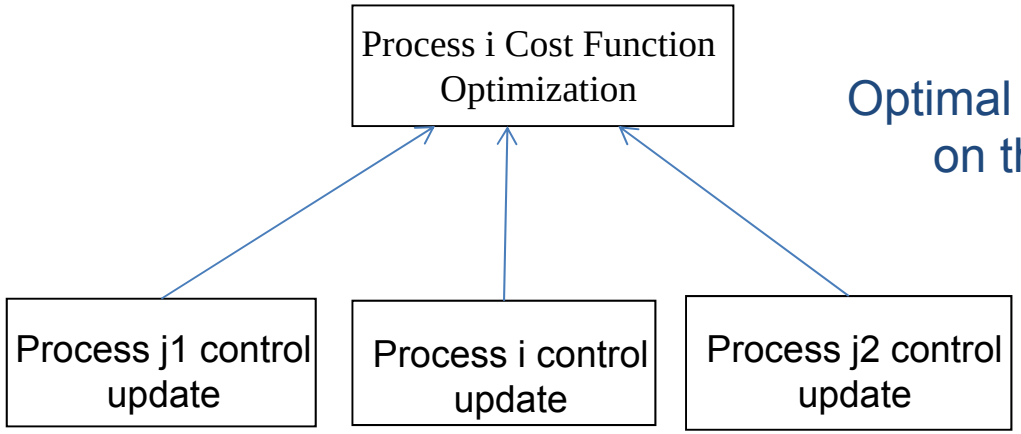
Leads to ONLINE FORWARD-IN-TIME implementation of optimal control

Do not need to know system drift dynamics $\dot{x}_i = f(x_i) + g(x_i)u_i$

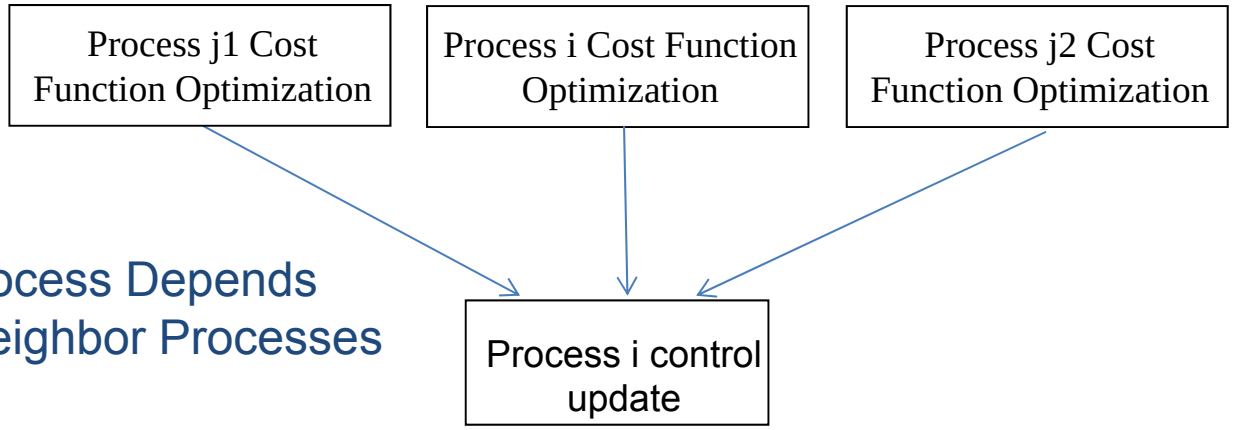
Best Paper Award, Int. Joint Conf. Neural Networks, Barcelona, 2010. **Draguna Vrabie** and F.L. Lewis, "Adaptive Dynamic Programming Algorithm for Finding Online the Equilibrium Solution of the Two-Player Zero-Sum Differential Game."

F.L. Lewis and D. Vrabie, "Reinforcement learning and adaptive dynamic programming for feedback control," IEEE Circuits & Systems Magazine, Invited Feature Article, pp. 32-50, Third Quarter 2009.

Graphical Games for Multi-Process Optimal Control



Optimal Performance of Each Process Depends on the Control of its Neighbor Processes



Control Policy of Each Process Depends on the Performance of its Neighbor Processes


09 . 06 . 2005

