Optimal Synchronization and Games on Graphs
Optimal Design for Synchronization & Games on Communication Graphs

Thanks to Jie Huang

Supported by AFOSR, NSF, ARO
Cooperative Control Synchronization: Optimal Design and Games on Communication Graphs

http://ARRI.uta.edu/acs

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

Supported by:
NSF - PAUL WERBOS
ARO, AFOSR
NNSF of China
China Project 111 at NEU
Invited by Derong Liu
Thanks to
Cesare Alippi
Zhang Huaguang
He who exerts his mind to the utmost knows nature’s pattern.
The way of learning is none other than finding the lost mind.

Man’s task is to understand patterns in nature and society.
Kung Tz
500 BC

Archery
Chariot driving

Music
Rites and Rituals

Poetry
Mathematics

Man’s relations to
Family
Friends
Society
Nation
Emperor
Ancestors

Confucius
Control Design Methods for Multi-Agent Systems

Outline

- Optimal Design for Synchronization of Cooperative Systems
- Distributed Observer and Dynamic Regulator
- Discrete-time Optimal Design for Synchronization
- Graphical Games

Acks. to:
- Guanrong Chen – Pinning control
- Lihua Xie – Local nbhd. tracking error
- Zhihua Qu – Lyapunov eq. for di-graphs
Books Coming


Key Point

Lyapunov Functions and Performance Indices
Must depend on graph topology

Hongwei Zhang, F.L. Lewis, and Abhijit Das
“Optimal design for synchronization of cooperative systems: state feedback, observer and output feedback,”
Outline
Cooperative Control
Locally Optimal Design and Synchronization
Globally Optimal Design for Collective Motion
Multi-player Games on Communication Graphs
Reinforcement Learning for Game Solutions
Outline
Cooperative Control
Locally Optimal Design and Synchronization
Globally Optimal Design for Collective Motion
Multi-player Games on Communication Graphs
Reinforcement Learning for Game Solutions
Structure of Natural and Manmade Systems

Local nature of Physical Laws
Peer-to-Peer Relationships in networked systems

Clusters of galaxies

The Internet

J.J. Finnigan, Complex science for a complex world
Synchronized Motion of Biological Groups

- Fish school
- Birds flock
- Locusts swarm
- Fireflies synchronize
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 as 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

Leader or root node
Followers
Communication Graph

\((V,E)\)

N nodes

Adjacency matrix

\[
A = \begin{bmatrix} a_{ij} \end{bmatrix}
\]

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

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
Dynamic Graph- the Distributed Structure of Control

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

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

\[
\begin{align*}
\dot{x}_i &= -x_i \sum_{j \in N_i} a_{ij} + \sum_{j \in N_i} a_{ij} x_j = -d_i x_i + a_{i1} \cdots a_{iN} \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} \\
\end{align*}
\]

\[
\begin{align*}
u &= \begin{bmatrix} u_1 \\ \vdots \\ u_N \end{bmatrix} \\
A &= \begin{bmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{NN} \end{bmatrix} \\
D &= \begin{bmatrix} d_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_N \end{bmatrix} \\
\end{align*}
\]

\[ u = -Dx + Ax = -(D - A)x = -Lx \]

\[ \dot{x} = -Lx \quad \text{Closed-loop dynamics} \]

If \( x \) is an \( n \)-vector then \[ \dot{x} = -(L \otimes I_n)x \]
Communication Graph

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

Synchronization problem

$x_i(t) - x_j(t) \rightarrow 0$
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
Consensus Value and Convergence Rate

Closed-loop system with local voting protocol

\[
\dot{x} = -Lx
\]

Modal decomposition

\[
x(t) = e^{-Lt}x(0) = \sum_{j=1}^{N} v_i e^{-\lambda_i t} w_i^T x(0) = \sum_{j=1}^{N} \left( w_i^T x(0) \right) e^{-\lambda_i t} v_i
\]

Let \( \lambda_1 = 0 \) be simple. Then for large \( t \)

\[
x(t) \to v_2 e^{-\lambda_2 t} w_2^T x(0) + v_1 e^{-\lambda_1 t} w_1^T x(0) = v_2 e^{-\lambda_2 t} w_2^T x(0) + \frac{1}{N} \sum_{j=1}^{N} \gamma_j x_j(0)
\]

\( \lambda_2 \) determines the rate of convergence and is called the \textit{FIEDLER e-value}

There is a big push to find expressions for the left e-vector for \( \lambda_1 = 0 \)

and the Fiedler e-val \( \lambda_2 \)

Let graph have a spanning tree. Then all nodes reach consensus.
Convergence Value and Rate

Closed-loop system with local voting protocol

\[ \dot{x} = -Lx \]

\( L \) has e-val at zero

Modal decomposition

\[ x(t) = e^{-Lt} x(0) = \sum_{j=1}^{N} v_i e^{-\lambda_j t} w_i^T x(0) = \sum_{j=1}^{N} \left( w_i^T x(0) \right) e^{-\lambda_j t} v_i \]

Let \( \lambda_1 = 0 \) be simple. Then for large \( t \)

\[ x(t) \to v_2 e^{-\lambda_2 t} w_2^T x(0) + v_1 e^{-\lambda_1 t} w_1^T x(0) = v_2 e^{-\lambda_2 t} w_2^T x(0) + \frac{1}{N} \sum_{j=1}^{N} \gamma_j x_j(0) \]

\( w_1 = [\gamma_1 \quad \gamma_2 \quad \cdots]^T \) determines the consensus value in terms of the initial conditions

\( \lambda_2 \) determines the rate of convergence - Fiedler e-value

\( \lambda_1 = 0 \) is simple if the graph is strongly connected

Depends on Communication Graph Topology

No freedom to determine the consensus value

We call this the Cooperative Regulator Problem
Graph Eigenvalues for Different Communication Topologies

Directed Tree-Chain of command

Directed Ring-Gossip network

OSCILLATIONS
Graph Eigenvalues for Different Communication Topologies

Directed graph-
Better conditioned

Undirected graph-
More ill-conditioned
Synchronization on Good Graphs

Mesh graph
4 neighbors

Chris Elliott fast video
Synchronization on Gossip Rings

Chris Elliott weird video

Ring graph or cycle
10 nodes
These beautiful pictures are from a lecture by **Ron Chen**, City U. Hong Kong
Pinning Control of Graphs
Locally Optimal Design and Synchronization
Controlled Consensus: Cooperative Tracker

Node state \( \dot{x}_i = u_i \)

Distributed Local voting protocol with control node \( v \)

\[
u_i = \sum_{j \in N_i} a_{ij}(x_j - x_i) + b_i(v - x_i)
\]

Local Neighborhood Tracking Error

\( b_i \neq 0 \) If control \( v \) is in the neighborhood of node \( i \)

\[
u_i = -\sum_{j \in \bar{N}_i} a_{ij}x_i + \sum_{j \in N_i} a_{ij}x_j + b_iv
\]

\[
\dot{x} = -(L + B)x + B1v \quad B = \text{diag}\{b_i\}
\]

Theorem. Let graph have a spanning tree and \( b_i \neq 0 \) for at least one root node. Then \( L+B \) is nonsingular with all e-vals positive and \( -(L+B) \) is asymptotically stable

Ron Chen – pinning control
Agent Dynamics and Local Feedback Design

\[ \dot{x}_i = Ax_i + Bu_i \]

\[ u_i = -Kx_i \]

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

\[ K = \begin{bmatrix} 0.5 & -0.5 \end{bmatrix} \]

Couple 6 agents with communication graph

Local neighborhood tracking error

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

\[ u_i = -K \varepsilon_i \]

Nodes synchronize to consensus heading
Agent Dynamics and local Feedback design

\[ \dot{x}_i = Ax_i + Bu_i \]
\[ u_i = -Kx_i \]
\[ A = \begin{bmatrix} -2 & -1 \\ 2 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad K = [0.5 \ -0.5] \]

ADD another comm. Link- more information flow

Local neighborhood tracking error

\[ e_i = \sum_{j \in N_i} e_{ij} (x_j - x_i) + g_i (x_{i_0} - x_i) \]
\[ u_i = K e_i \]

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”
A. State Feedback Design for Cooperative Systems on Graphs

Cooperative Regulator vs. Cooperative Tracker problem

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

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

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

Local neighborhood tracking error

\[ e_i = \sum_{j \in N_i} e_{ij}(x_j - x_i) + g_i(x_0 - x_i) \]

Ron Chen- pinning control \hspace{1cm} Lihua Xie- error

\[ L = D - A \]

Overall error vector \( e = -((L + G) \otimes I_n)(x - x_0) = -((L + G) \otimes I_n)\delta \) = Local quantity

where \( e = [e_1^T, e_2^T, \ldots, e_N^T]^T \in \mathbb{R}^{nN} \), \( x_0 = Ix_0 \in \mathbb{R}^{nN} \), \( L = 1 \otimes I_n \in \mathbb{R}^{nN \times n} \)

Consensus or synchronization error \( \delta = (x - x_0) \in \mathbb{R}^{nN} \) = Global quantity
Local Neighborhood Tracking Error

\[ e = -\left( (L + G) \otimes I_n \right) (x - \bar{x}_0) = -\left( (L + G) \otimes I_n \right) \delta \]

Local quantity \quad \text{Global quantity}

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

\[ \|\delta\| \leq \|e\| / \sigma(L + G) \]

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

\[ x(t) = Ix_0(t) \]
Coop. nbhd SVFB

\[ u_i = cK e_i = cK \left( \sum_{j \in N_i} e_{ij} (x_j - x_i) + g_i (x_0 - x_i) \right) \]

Closed loop system

\[ \dot{x}_i = A x_i + B u_i = A x_i + cBK \left( \sum_{j \in N_i} e_{ij} (x_j - x_i) + g_i (x_0 - x_i) \right) \]

Overall state

\[ x = \begin{bmatrix} x_1^T & x_2^T & \ldots & x_N^T \end{bmatrix}^T \in \mathbb{R}^{nN}, \quad \delta = (x - \bar{x}_0) \]

Distributed form of control

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

Overall c.l. dynamics

\[ \dot{x} = \left( (I_N \otimes A) - c(L + G) \otimes BK \right) x + c \left( (L + G) \otimes BK \right) x_0 \]

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

\[ A - c\lambda_i BK \]

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)
\begin{align}
0 &= A^T P + PA + Q - PBR^{-1}B^T P \\
K &= R^{-1}B^T P
\end{align}
Then the synchronization dynamics (13) are asymptotically stable (AS) for all coupling gains $c > \frac{1}{\lambda(L+G)}$
\begin{align}
\text{with } \lambda(L+G) &= \min_i \text{Re}(\lambda_i(L+G)).
\end{align}

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

**DECOUPLING CONTROL DESIGN FROM COMMUNICATION GRAPH STRUCTURE**

**LOCAL OPTIMAL DESIGN GUARANTEES GLOBAL SYNCHRONIZATION**

minimizes $J_i = \frac{1}{2} \int_0^\infty (x_i^T Q x_i + u_i^T R u_i) \, dt$

Hongwei Zhang, F.L. Lewis, and Abhijit Das, “Optimal design for synchronization of cooperative systems: state feedback, observer and output feedback”

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)

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

$$K = R^{-1}B^TP$$

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

$$\lambda > \frac{1}{\hat{\lambda}(L+G)}$$

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

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 $\lambda_i$ real. For complex $\lambda_i = a + jb$ one has

$$(A - c(a + jb)BR^{-1}B^TP)^T P + P(A - c(a + jb)BR^{-1}B^TP) + Q + (2ca - 1)PBR^{-1}B^TP$$

$$= (A - caBR^{-1}B^TP)^T P + P(A - caBR^{-1}B^TP) + Q + (2ca - 1)PBR^{-1}B^TP$$

$$= A^TP + PA + Q - PBR^{-1}B^TP$$

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 $(A - c\lambda_i BK)$.
Graph Eigenvalues for Different Communication Topologies

Directed Tree-
Chain of command

Directed Ring-
Gossip network
OSCILLATIONS
Example: Unbounded Region of Consensus for Optimal Feedback Gains.

\[
A - c\lambda BK
\]

\[\lambda = \text{E-vals of } (L+G)\]

**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 [Li, Duan, Chen 2009]

**b. Unbounded Consensus Region for Optimal SVFB Gain**

\[Q=I, \ R=1\]

\[
K = \begin{bmatrix} 1.544 & 1.8901 \end{bmatrix}
\]

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

Figure 2. Unbounded consensus region for optimal SVFB gain
Results:

Local Riccati Design yields guaranteed stable synchronization
Decouples Controls Design from Graph Properties
Standard Chartered Bank

Hong Kong$ 20

帳面券

Hong Kong 6.12.53
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 as 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 \( \dot{x}_i = A x_i + B u_i \in \mathbb{R}^n \)

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

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

\( \delta = (x - Ix_0) \)

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

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

So \( Q \) and \( R \) must depend on the graph topology.
Inverse Optimality

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

\[
V(x) = 0 \iff x \in S \\
V(x) \geq \alpha(d(x, S)) \\
L_1(x) \geq \gamma(d(x, S))
\]

\[
L_1(x) + \nabla V(x)^T f(x) - \frac{1}{4} \nabla V(x)^T g(x) R^{-1} g(x)^T \nabla V(x) = 0 \quad (0)
\]

\[
\phi(x) = -\frac{1}{2} R^{-1} g(x)^T \nabla V(x)^T
\]

then \( u = \phi(x) \) is optimal with respect to the performance index with the integrand \( \mathcal{L}(x, u) = L_1(x) + u^T R u \). 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 \( R \)

\[
u = -c \left( (L + G) \otimes K \right) \delta
\]
System \[ \dot{x}_i = Ax_i + Bu_i \in \mathbb{R}^n \]
Leader \[ \dot{x}_0 = Ax_0 \]
\[ \dot{x} = (I \otimes A)x + (I \otimes B)u \]

Global disagreement error \[ \dot{\delta} = (I \otimes A)\delta + (I \otimes B)u \quad \delta = x - \underline{I}x_0 \]

Local nbhd tracking error \[ \varepsilon_i = \sum_{j \in N_i} e_{ij}(x_j - x_i) + g_i(x_0 - x_i) \]
\[ e = -((L + G) \otimes I_n)\delta \]

Distributed Control \[ u_i = cK_2 \varepsilon_i = cK_2 \left( \sum_{j \in N_i} e_{ij}(x_j - x_i) + g_i(x_0 - x_i) \right) \]
\[ u = -c(L + G) \otimes K_2 \delta \]

Closed-loop system \[ \dot{x}_i = Ax_i + Bu_i = Ax_i + cBK_2 \left( \sum_{j \in N_i} e_{ij}(x_j - x_i) + g_i(x_0 - x_i) \right) \]
\[ \dot{x} = [(I_N \otimes A) - c(L + G) \otimes BK_2]x + c[(L + G) \otimes BK_2]x_0 \]

Global synch. error dynamics \[ \dot{\delta} = [(I_N \otimes A) - c(L + G) \otimes BK_2] \delta \]

**Graph structure \[ \otimes \] Control structure**
B.1 Optimal Cooperative Tracker for Single-Integrator Dynamics

System
\[ \dot{x}_i = u_i, \quad x_i \in \mathbb{R} \]

Leader node
\[ \dot{x}_0 = 0 \]

\[ \dot{x} = u \]
\[ x = [x_1 \ldots x_N]^T \quad u = [u_1 \ldots u_N]^T \]

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

\[ e = -(L + G) \delta \]
\[ e = [\varepsilon_1 \ldots \varepsilon_N]^T \quad G = \text{diag} \{g_i\} \]

Global disagreement error
\[ \delta = x - I x_0 \]

control
\[ u_i = \varepsilon_i \]
\[ u = -(L + G) \delta \]

Closed-loop System
\[ \dot{\delta} = u = -(L + G) \delta \]

Graph structure \( \otimes \) Control structure

**Lemma 4.** If the graph is strongly connected, given that there exists at least one non zero pinning gain, then \( L + G > 0 \) i.e. nonsingular, and \( u = -(L + G) \delta \) solves the consensus problem [10].
Theorem 3. Let the error dynamics be given as (0), and the conditions of Lemma 4 be satisfied. Then for some $R = R^T > 0$ the control $u = -(L + G)\delta$ is optimal with respect to the performance index

$$J(\delta_0, u) = \int_0^\infty (\delta^T (L + G)^T R(L + G)\delta + u^T R u) dt,$$ 

$$= \int_0^\infty (e^T R e + u^T R u) dt$$

and is stabilizing to the reference state $x_0$ if there exists a positive definite matrix $P = P^T > 0$ satisfying

$$P = R(L + G).$$

Proof: The Lyapunov function $V(\delta) = \delta^T P \delta > 0$, and $L_1(\delta) = \delta^T Q \delta = \delta^T (L + G)^T R(L + G)\delta > 0$ satisfy the conditions of Lemma 2. The Algebraic Riccati equation

$$(L + G)^T R(L + G) - PR^{-1}P = 0$$

is satisfied by $P$, and $u = -(L + G)\delta = -R^{-1}P\delta$ thus proving the theorem.
Cooperative Tracker for Identical LTI Dynamics

Cooperative Regulator vs. Cooperative Tracker problem

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

Control node or Command generator \[ \dot{x}_0 = Ax_0 \quad \text{ (Exosystem)} \]

Synchronization Tracker design problem \[ x_i(t) \rightarrow x_0(t), \forall i \]

Local neighborhood tracking error

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

Ron Chen - pinning control \quad \text{Lihua Xie - error}

\[ L = D - A \]

Overall error vector \[ e = -((L + G) \otimes I_n)(x - x_0) = -((L + G) \otimes I_n)\delta \quad \text{= Local quantity} \]

where \[ e = \begin{bmatrix} \varepsilon_1^T & \varepsilon_2^T & \cdots & \varepsilon_N^T \end{bmatrix}^T \in \mathbb{R}^{nN}, \quad x_0 = I x_0 \in \mathbb{R}^{nN}, \quad L = I \otimes I_n \in \mathbb{R}^{nN \times n} \]

Consensus or synchronization error \[ \delta = (x - x_0) \in \mathbb{R}^{nN} \quad \text{= Global quantity} \]
Cooperative Tracker for Identical LTI Dynamics

System
\[ \dot{x}_i = A x_i + B u_i \in \mathbb{R}^n \]
\[ \dot{x} = (I \otimes A)x + (I \otimes B)u \]
\[ \dot{\delta} = (I \otimes A)\delta + (I \otimes B)u \]
\[ \delta = x - I x_0 \]

Local nbhd tracking error
\[ e_i = \sum_{j \in N_i} e_{ij} (x_j - x_i) + g_i (x_0 - x_i) \]
\[ e = -((L + G) \otimes I_n)\delta \]
\[ e = [e_1 \ldots e_N]^T \]

Control
\[ u_i = cK_2 e_i = cK_2 \sum_{j \in N_i} e_{ij} (x_j - x_i) + g_i (x_0 - x_i) \]
\[ u = -c(L + G) \otimes K_2 \delta \]

Closed-loop system
\[ \dot{x}_i = A x_i + B u_i = A x_i + cBK_2 \sum_{j \in N_i} e_{ij} (x_j - x_i) + g_i (x_0 - x_i) \]
\[ \dot{x} = [(I_N \otimes A) - c(L + G) \otimes BK_2] x + c[(L + G) \otimes BK_2] x_0 \]
\[ \dot{\delta} = [(I_N \otimes A) - c(L + G) \otimes BK_2] \delta \]

Graph structure \( \otimes \) Control structure
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

\[
P_1 = c R_1 (L + G),
\]

\[
A^T P_2 + P_2 A + Q_2 - P_2 B R^{-1} B^T P_2 = 0,
\]

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

\[
K_2 = R_2^{-1} B^T P_2.
\]

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

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

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

\[
c > \frac{\sigma_{\text{max}} (R_1 (L + G) \otimes (Q_2 - K_2^T R_2 K_2))}{\sigma_{\text{min}} ((L + G)^T R_1 (L + G) \otimes K_2^T R_2 K_2)}
\]
**Proof:**

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

**ARE**

\[
\bar{A}^T P + P\bar{A} + Q - PBR^{-1}\bar{B}^T P = 0
\]

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

Select \( P = P_1 \otimes P_2 \quad R = R_1 \otimes R_2 \)

**ARE**

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

\[
P_1 \otimes (A^T P_2 + P_2 A) + Q - PR^{-1}_1 P_1 \otimes (P_2 BR^{-1}_2 B^T P_2) = 0
\]

Choose \( Q \)

\[
Q = c^2 ((L + G) \otimes K_2)^T (R_1 \otimes R_2)((L + G) \otimes K_2) - cR_1(L + G) \otimes (A^T P_2 + P_2 A)
\]

\[
= c^2 (L + G)^T R_1 (L + G) \otimes K_2 R_2 K_2 + cR_1(L + G) \otimes (Q_2 - P_2 BR^{-1}_2 B^T P_2)
\]

\[
= P_1 R_1^{-1} P_1 \otimes P_2 BR^{-1}_2 B^T P_2 + P_1 \otimes (Q_2 - P_2 BR^{-1}_2 B^T P_2)
\]

\[
Q_1 = c^2 (L + G)^T R_1 (L + G)
\]

**ARE**

\[
P_1 \otimes (A^T P_2 + P_2 A^T + Q_2 - P_2 BR^{-1}_2 B^T P_2)
\]

\[
+ (Q_1 - P_1 R_1^{-1} P_1) \otimes (P_2 BR^{-1}_2 B^T P_2) = 0
\]

*2 conditions:*

\[
A^T P_2 + P_2 A + Q_2 - P_2 BR^{-1}_2 B^T P_2 = 0
\]

\[
P_1 = cR_1(L + G)
\]

Control

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

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

Distributed !!
Two Conditions for global optimal design on the graph

1. Condition on graph topology

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

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

\[ A^T P_2 + P_2 A + Q_2 - P_2 B R_2^{-1} B^T P_2 = 0 \]

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

Always holds if \((A,B)\) reachable

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

\[ P_1 = c R_1 (L + G) \quad \text{Equivalent to} \quad R_1 (L + G) = (L + G)^T R_1 \]

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. \( R_1 (L + G) = (L + G) R_1 \quad \text{iff} \quad R_1, (L + G) \quad \text{have the same eigenvectors} \)

Let \( L = T \Lambda T^T \quad \text{Jordan form} \)

Select \( R = T \Theta T^T > 0 \quad \text{For any} \quad \Theta > 0 \quad \text{diagonal} \)

R depends on graph topology- ALL e-vectors
2. Detail Balanced Graphs

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

\[ P_1 = P_1^T > 0, \quad R_1 = R_1^T > 0 \]

\[ \lambda_i e_{ij} = \lambda_j e_{ji} \quad \text{for} \quad \lambda_1...\lambda_N > 0 \]

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

\[ L = D\bar{P}, \quad D = diag\{1/\lambda_i\} > 0 \quad \text{with} \quad \bar{P} \quad \text{a symmetric graph Laplacian matrix} \]

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

\[ P = D^{-1}(L + G) = R(L + G) \]

\( R \) 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

\[ P_1 = c R_1 (L + G) \quad \text{Equivalent to} \quad R_1 (L + G) = (L + G)^T R_1 \]

\[ P_1 = P_1^T > 0, \quad R_1 = R_1^T > 0 \]

3. Directed Graphs with Simple Graph Laplacian \( L + G \)

\[ T (L + G) T^{-1} = \Lambda \quad \text{Diagonal Jordan form} \]

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

\[ T^T T (L + G) = (L + G)^T T^T T \]

Select \( R = T^T T = R^T > 0 \)

\[ R \text{ depends on graph topology - ALL e-vectors} \]

**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
A.2 Discrete-Time Optimal Design for Synchronization


Distributed systems
\[ x_i(k+1) = Ax_i(k) + Bu_i(k) \]

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
\[ \Lambda_k, \quad k = 1, N \]
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 \( \Lambda_k \), \( k = 1, N \)

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

**MIXES UP CONTROL DESIGN AND GRAPH STRUCTURE**
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

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

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

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

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 the

$$K = (B^T P B)^{-1} B^T P A$$

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 $\Gamma$ matrix of the graph $G(V,E)$ have all eigenvalues real. Select $Q$ as in Corollary 4. Then the synchronization condition becomes

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

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.

If graph eigenvalues are real

For SI systems, for proper choice of $Q$

Mahler measure

$$\log_2 \prod_u \lambda^u_i(A)$$

intrinsic entropy rate = minimum data rate in a networked control system that enables stabilization of an unstable system – Baillieul and others.

$$\frac{r_0}{c_0} = \frac{\Lambda_{\text{max}} - \Lambda_{\text{min}}}{\Lambda_{\text{max}} + \Lambda_{\text{min}}}.$$

$$r = \frac{1}{\prod_u |\lambda^u(A)|}$$

Eigen-ratio = ‘condition number’ of the communication graph

Work on log quantization- Elia & Mitter, Lihua Xie
Single-Input case with Real Graph Eigenvalues

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

$$\prod_i |\lambda_i(A)| < \frac{\Lambda_{\max} + \Lambda_{\min}}{\Lambda_{\max} - \Lambda_{\min}}.$$ 

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.

<table>
<thead>
<tr>
<th>Mahler Measure</th>
<th>Graph Condition Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M(A) = \prod_{\text{unstable } \lambda_i}</td>
<td>\lambda_i(A)</td>
</tr>
<tr>
<td>Synchronization guaranteed if $h(A) = \log(M(A))$</td>
<td>Like to have $\kappa(G) \approx 1$ Varshney</td>
</tr>
<tr>
<td>$\Lambda_{\min}$ large means fast convergence</td>
<td>$M(A) &lt; \frac{1 + \kappa^{-1}}{1 - \kappa^{-1}}$</td>
</tr>
<tr>
<td>$\Lambda_{\max}$</td>
<td></td>
</tr>
</tbody>
</table>

New definition- Graph Channel Capacity

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

Graph Eigenvalues for Different Communication Topologies

Directed Tree-Chain of command

Directed Ring-Gossip network OSCILLATIONS

HW1 Q4b eigenvalues

HW1 Q8b eigenvalues
Graph Eigenvalues for Different Communication Topologies

Directed graph-
Better conditioned

Undirected graph-
More ill-conditioned
Single-Input case with Real Graph Eigenvalues

\[ x_i(k+1) = Ax_i(k) + Bu_i(k) \quad u_i = cK \left[ \sum_{j \in N_i} e_{ij}(x_j - x_i) + g_i(x_0 - x_i) \right] \]

\[ \mu(A) \equiv \prod_{u} |\lambda^u(A)| < \frac{\Lambda_{\text{max}}}{} + \frac{\Lambda_{\text{min}}}{} \]

Is equivalent to

\[ \frac{\Lambda_{\text{max}}}{} \frac{\mu(A)+1}{} < \frac{\mu(A)-1}{} \]

Add stable filter

\[ u_i = cF(z)K \left[ \sum_{j \in N_i} e_{ij}(x_j - x_i) + g_i(x_0 - x_i) \right] \]

Filtered protocol gives synch. if

\[ \frac{\Lambda_{\text{max}}}{} \frac{\mu(A)+1}{} \frac{\gamma^2}{} < \left( \frac{\mu(A)-1}{} \right)^2 \]

select \( \gamma > \mu(A) \)

\[ P \geq 0 \quad \text{the stabilizing solution to} \]

\[ P = A^T P(I + (1-\gamma^{-2})\lambda_{\text{min}}^2 BB^T P)^{-1} A, \quad B^T PB < \frac{\gamma^2}{\lambda_{\text{min}}^2} \]

\[ K = \lambda_{\text{min}}(I + (1-\gamma^{-2})\lambda_{\text{min}}^2 B^T PB)^{-1} B^T PA \]

\[ T(z) = \lambda_{\text{min}}K(zI - A + \lambda_{\text{min}}BK)^{-1} B \]

\[ F(z) = \frac{(1-\gamma^{-1})^2}{1-\gamma^{-2}T(z)} \]

Improvement

Complementary sensitivity
Graph Condition Number

\[ \kappa(G) = \frac{\Lambda_{\text{max}}}{\Lambda_{\text{min}}} \]

\[ \text{eigenratio} = \frac{\Lambda_{\text{min}}}{\Lambda_{\text{max}}} \]

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

\( \Lambda_{\text{min}} \) large means fast convergence

L.R. Varshney, “Distributed inference with costly wires”
Games on Communication Graphs
Kyriakos Vamvoudakis, Mohammed Abouheaf
Graphical Coalitional Games

Sun Tzu

500 BC

Sun Tzu's

The Art of War

孫子兵法

孫子兵法

Sun Tz bin fa
Games on Communication Graphs

http://ARRI.uta.edu/acs

F.L. Lewis, K. Vamvoudakis, M. Abouheaf
UTA Research Institute (UTARI)
The University of Texas at Arlington

Supported by:
NSF - PAUL WERBOS
ARO, AFOSR
Manufacturing as the Interactions of Multiple Agents

Each machine has its 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)
\[ \delta_i = \sum_{j \in N_i} e_{ij}(x_i - x_j) + g_i(x_i - x_0), \quad \text{Pinning gains } g_i \geq 0 \quad \text{(Ron Chen)} \]

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
\( x_i(t) \rightarrow x_0(t), \forall i \)

Local neighborhood tracking error (Lihua Xie)
\[ \delta_i = \sum_{j \in N_i} e_{ij} (x_i - x_j) + g_i (x_i - x_0), \quad \text{Pinning gains } g_i \geq 0 \quad \text{(Ron Chen)} \]

Standard way =

Global neighborhood tracking error
\[ \delta = \begin{bmatrix} \delta_1^T & \delta_2^T & \cdots & \delta_N^T \end{bmatrix}^T \in \mathbb{R}^{nN}, \quad x = \begin{bmatrix} x_1^T & x_2^T & \cdots & x_N^T \end{bmatrix}^T \in \mathbb{R}^{nN}, \quad x_0 = I x_0 \in \mathbb{R}^{nN} \]
\[ \delta = \left( (L + G) \otimes I_n \right) (x - x_0) = \left( (L + G) \otimes I_n \right) \zeta, \quad \zeta = (x - x_0) \in \mathbb{R}^{nN} \]

Lemma. Let graph be strongly connected and at least one pinning gain nonzero. Then
\[ \| \zeta \| \leq \| \delta \| / \sigma (L + G) \]
and agents synchronize iff \( \delta(t) \rightarrow 0 \)
Graphical Game: Games on Graphs

Local nbhd. tracking error dynamics

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

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

Define Local nbhd. performance index

$$J_i(\delta_i(0), u_i, u_{-i}) = \frac{1}{2} \int_0^\infty (\delta_i^T Q_i \delta_i + u_i^T R_i u_i + \sum_{j \in N_i} u_j^T R_j u_j) \, dt = \frac{1}{2} \int_0^\infty L_i(\delta_i(t), u_i(t), u_{-i}(t)) \, dt$$

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

Local value functions for fixed policies $u_i$

$$V_i(\delta_i(t)) = \frac{1}{2} \int_t^\infty (\delta_i^T Q_i \delta_i + u_i^T R_i u_i + \sum_{j \in N_i} u_j^T R_j u_j) \, dt$$

Static Graphical Game

$$(G, U, \nu)$$

$$G = (V, E), \quad \nu = [v_1 \ldots v_N]^T$$

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

Standard N-player differential game

$$\dot{z} = Az + \sum_{i=1}^N B_i u_i$$

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

Kyriakos Vamvoudakis

Local agent dynamics driven by neighbors’ controls

Values driven by neighbors’ controls

Value depends only on neighbors

Dynamics depend on all other agents

Values depend on all other agents
**Graphical Game: Games on Graphs**

**Local nbhd. tracking error dynamics**

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

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

Define Local nbhd. performance index

\[
J_i(\delta_i(0), u_i, u_{-i}) = \frac{1}{2} \int_0^\infty (\delta_i^T Q_i \delta_i + u_i^T R_i u_i + \sum_{j \in N_i} u_j^T R_{ij} u_j) \, dt
\]

\[
u_{-i}(t) = \{ u_j : j \in N_i \}\]

Local value functions for fixed policies \( u_i \)

\[
V_i(\delta_i(t)) = \frac{1}{2} \int_t^\infty (\delta_i^T Q_i \delta_i + u_i^T R_i u_i + \sum_{j \in N_i} u_j^T R_{ij} u_j) \, dt
\]

**Static Graphical Game**

\[(G, U, v)\]

\[G = (V, E), \quad v = [v_1 \cdots v_N]^T\]

\[v_i(U_i, \{U_j : j \in N_i\}) \in R\]
New Differential Graphical Game

Control action of player $i$  

State dynamics of agent $i$  
\[ \dot{\delta}_i = A\delta_i + (d_i + g_i)B_iu_i - \sum_{j \in N_i} e_{ij}B_ju_j \]

Value function of player $i$  
\[ J_i(\delta_i(0), u_i, u_{-i}) = \frac{1}{2} \int_0^\infty (\delta_i^T Q_{ii} \delta_i + u_i^T R_{ii}u_i + \sum_{j \in N_i} u_j^T R_{ij}u_j) \, dt \]
Standard Multi-Agent Differential Game

Central Dynamics

Value function of player $i$

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

Control action of player $i$

Central Dynamics depends on ALL other control actions

$$\dot{z} = A z + \sum_{i=1}^{N} B_i u_i$$
Team Interest vs. Self Interest

Cooperation vs. Collaboration

The objective functions of each player can be written as a team average term plus a conflict of interest term:

\[ J_1 = \frac{1}{3} (J_1 + J_2 + J_3) + \frac{1}{3} (J_1 - J_2) + \frac{1}{3} (J_1 - J_3) \equiv J_{team} + J_{1^{coi}} \]

\[ J_2 = \frac{1}{3} (J_1 + J_2 + J_3) + \frac{1}{3} (J_2 - J_1) + \frac{1}{3} (J_2 - J_3) \equiv J_{team} + J_{2^{coi}} \]

\[ J_3 = \frac{1}{3} (J_1 + J_2 + J_3) + \frac{1}{3} (J_3 - J_1) + \frac{1}{3} (J_3 - J_2) \equiv J_{team} + J_{3^{coi}} \]

For N-players

\[ J_i = \frac{1}{N} \sum_{j=1}^{N} J_j + \frac{1}{N} \sum_{j=1}^{N} (J_i - J_j) \equiv J_{team} + J_{i^{coi}}, \quad i = 1, N \]

For N-player zero-sum games, the first term is zero, i.e. the players have no goals in common.
Problems with Nash Equilibrium Definition on Graphical Games

Game objective

\[
V_i^* (\delta_i(t)) = \min_{u_i} \int_0^\infty \left( \frac{1}{2} (\delta_i^T \tilde{Q}_i \delta_i + u_i^T \tilde{R}_i u_i + \sum_{j \in N_i} u_j^T R_{ij} u_j) \right) dt
\]

Define 

\[ u_{-i}(t), \quad \{u_j : j \in N_i\} \quad \text{Neighbors of node } i \]

\[ u_{G-i} = \{u_j : j \in N, j \neq i\} \quad \text{All other nodes in graph} \]

Def: Nash equilibrium

\[ \{u_1^*, u_2^*, \ldots, u_N^*\} \quad \text{are in Nash equilibrium if} \]

\[ J_i^* = J_i (u_i^*, u_{G-i}^*) \leq J_i (u_i, u_{G-i}^*), \quad \forall i \in N \]

Counterexample. Disconnected graph

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

\[ J_i (u_i) = J_i (u_i, u_{G-i}) = J_i (u_i, u_{G-i}^*), \quad \forall i \]

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
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

\( \{u^*_1, u^*_2, ..., u^*_N\} \) are in Interactive Nash equilibrium if

1. \( J^*_i \triangleq J_i (u^*_i, u_{G-i}^*) \leq J_i (u_i, u_{G-i}^*), \quad \forall i \in N \)

2. There exists a policy \( u_j \) such that

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

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 multiplayer graph games.

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

This rules out the disconnected counterexample.
Theorem 3. Let $(A,B)$ be reachable for all $i$.
Let agent $i$ be in local best response

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

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

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

$$u_k = -K_kp_k - v_k$$

Hamiltonian System

$$\begin{bmatrix}
\dot{\delta} \\
\dot{p}
\end{bmatrix} = \begin{bmatrix}
(I_n \otimes A) & (L+G) \otimes I_n \text{diag}(B_iK_i) \\
-diag(Q_{ii}) & -(I_n \otimes A^T)
\end{bmatrix} \begin{bmatrix}
\delta \\
p
\end{bmatrix} + \begin{bmatrix}
(L+G) \otimes I_n \text{diag}(B_k) \\
0
\end{bmatrix} \begin{bmatrix}
v_k
\end{bmatrix} \equiv \bar{A} \begin{bmatrix}
\delta \\
p
\end{bmatrix} + \bar{B}v_k$$

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

Picks out the shortest path from node $k$ to node $i$
Graphical Game Solution Equations

Value function

\[ V_i(\delta_i(t)) = \frac{1}{2} \int_{\tau}^{\infty} \left( \delta_i^T Q_i \delta_i + u_i^T R_i u_i + \sum_{j \in N_i} u_j^T R_{ij} u_j \right) dt \]

Differential equivalent (Leibniz formula) is Bellman’s Equation

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

Stationarity Condition

\[ 0 = \frac{\partial H_i}{\partial u_i} \quad \Rightarrow \quad u_i = - (d_i + g_i) R_i^{-1} B_i^T \frac{\partial V_i}{\partial \delta_i} \]

1. Coupled HJ equations

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

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

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

\[ 0 = H_i(\delta_i, \frac{\partial V_i}{\partial \delta_i}, u_i^*, u_{-i}) = \frac{\partial V_i^T}{\partial \delta_i} A_i^c + \frac{1}{2} \delta_i^T Q_i \delta_i + \frac{1}{2} (d_i + g_i)^2 \frac{\partial V_i^T}{\partial \delta_i} B_i R_i^{-1} B_i^T \frac{\partial V_i}{\partial \delta_i} + \frac{1}{2} \sum_{j \in N_i} u_j^T R_{ij} u_j \]

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

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

(34)

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.
Reinforcement Learning to Solve Graphical Games
Books


New Chapters on:
- Reinforcement Learning
- Differential Games


Different methods of learning

Reinforcement learning
Ivan Pavlov 1890s

We want OPTIMAL performance
- ADP- Approximate Dynamic Programming

Actor-Critic Learning
Paul Werbos

Desired performance

Sutton & Barto book
Figure 1. Learning-oriented specialization of the cerebellum, the basal ganglia, and the cerebral cortex [1, 2]. The cerebellum is specialized for supervised learning based on the error signal encoded in the climbing fibers from the inferior olive. The basal ganglia are specialized for reinforcement learning based on the reward signal encoded in the dopaminergic fibers from the substantia nigra. The cerebral cortex is specialized for unsupervised learning based on the statistical properties of the input signal.

Doya, Kimura, Kawato 2001
Summary of Motor Control in the Human Nervous System

Hierarchy of multiple parallel loops
Online Solution of Graphical Games

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)

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

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

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

which explicitly is

\[
u_i^{k+1} = -(d_i + g_i)R_i^{-1}B_i^T \frac{\partial V_i^k}{\partial \delta_i}, \forall i = 1, \ldots, N.
\]  
(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 \( \rho_{ij} = \overline{\sigma}(R_{ij}^{-1}R_{ij}) \), where \( \overline{\sigma}(R_{ij}^{-1}R_{ij}) \) is the maximum singular value of \( R_{ij}^{-1}R_{ij} \).

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 \( \rho_{ij} \), \( \mu_i \) converges to the global Nash equilibrium and for all \( i \), and the values converge to the optimal game values \( V_i^k \rightarrow \bar{V}_i^* \).
Online Solution of Graphical Games

Policy Iteration gives structure needed for online graph games

Solve simultaneously online:

Bellman equation

\[
H_i(\delta_i, \frac{\partial V_i}{\partial \delta_i}, u_i, u_{-i}) \equiv \frac{\partial V_i}{\partial \delta_i} \left( A\delta_i + (d_i + g_i)B_iu_i - \sum_{j \in N_i} e_{ij}B_ju_j \right) + \frac{1}{2} \delta_i^TQ_{ii}\delta_i + \frac{1}{2}u_i^TR_iu_i + \frac{1}{2} \sum_{j \in N_i} u_j^TR_ju_j = 0
\]

\[
u_i = -(d_i + g_i)R_i^{-1}B_i^T\frac{\partial V_i}{\partial \delta_i}
\]

Weierstrass Approximator structures- 2 at each node

\[
\hat{V}_i = \hat{W}_i^T \phi_i \quad \text{critic}
\]

\[
\hat{u}_i = -\frac{1}{2} (d_i + g_i)R_i^{-1}B_i^T \frac{\partial \phi_i}{\partial \delta_i} \hat{W}_{i+N}
\quad \text{actor}
\]

Bellman equation becomes an algebraic equation in the parameters

\[
\sigma_i^T \hat{W}_i + \delta_i^TQ_{ii}\delta_i + \frac{1}{4} \hat{W}_{i+N}^T D_i \hat{W}_{i+N} + \frac{1}{4} \sum_{j \in N_i} (d_j + g_j)^2 \hat{W}_{j+N}^T \frac{\partial \phi_j}{\partial \delta_j} B_j R_j^{-T} R_j R_j^{-1} B_j \frac{\partial \phi_j}{\partial \delta_j} \hat{W}_{j+N}
\]

\[
\sigma_i = \frac{\partial \phi_i}{\partial \delta_i} (A\delta_i + (d_i + g_i)B_i\hat{u}_i - \sum_{j \in N_i} e_{ij}B_j\hat{u}_j)
\]
Online Solution of Graphical Games Using Value Function Approximation

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 $i$

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

Tuning law for Critic parameters

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

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

Tuning law for Actor parameters

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

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

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

\[
V(x(t)) = \int_{t}^{t+T} r(x,u) \, d\tau + V(x(t+T)), \quad V(0) = 0
\]

Another form for the CT Bellman eq.

Solves Lyapunov equation without knowing \( f(x,u) \)
Can avoid knowledge of drift term $f(x)$ by using \textit{Integral Reinforcement Learning (IRL)}

Draguna Vrabie

Then HJ equations are solved online without knowing $f(x)$

Coupled AREs are solved online without knowing $A$
Lemma 1 – Draguna Vrabie

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

Is equivalent to

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

Another form for the CT Bellman eq.

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

**Proof:**

\[
\frac{d(V(x))}{dt} = \left( \frac{\partial V}{\partial x} \right)^T f(x,u) = -r(x,u)
\]

\[
\int_t^{t+T} r(x,u) \, d\tau = - \int_t^{t+T} d(V(x)) = V(x(t)) - V(x(t+T))
\]

Allows definition of temporal difference error for CT systems

\[
e(t) = -V(x(t)) + \int_t^{t+T} r(x,u) \, d\tau + V(x(t+T))
\]
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 \[ \dot{x}_i = f(x_i) + g(x_i)u_i \]


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
Motions of Biological Groups

Local / Peer-to-Peer Relationships in socio-biological systems

- Fish school
- Birds flock
- Locusts swarm
- Fireflies synchronize
Our revels now are ended. These our actors,  
As I foretold you, were all spirits, and  
Are melted into air, into thin air.

The cloud-capped towers, the gorgeous palaces,  
The solemn temples, the great globe itself,  
Yea, all which it inherit, shall dissolve,  
And, like this insubstantial pageant faded,  
Leave not a rack behind.

We are such stuff as dreams are made on,  
and our little life is rounded with a sleep.

Prospero, in The Tempest,  
act 4, sc. 1, l. 152-6, Shakespeare
Policy Iteration is Reinforcement Learning

Adaptive Critics

The Adaptive Critic Architecture

Critic and Actor tuned simultaneously
Leads to ONLINE FORWARD-IN-TIME implementation of optimal control

Optimal Adaptive Control

Adaptive Critics

The Adaptive Critic Architecture

Control policy update

\[ \hat{u}_i = -\frac{1}{2} (d_i + g_i) R_{ii}^{-1} B_i^T \frac{\partial \phi_i^T}{\partial \hat{\phi}_i} \hat{W}_{i+N} \]

Cost

Policy Evaluation

Value update

Critic and Actor tuned simultaneously

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

A new adaptive control architecture