F.L. Lewis, NAI

Moncrief-O'Donnell Chair, UTA Research Institute (UTARI)
The University of Texas at Arlington, USA

Distributed Synchronization Control of Microgrids:

Optimal Synchronization on Sparse Communication Graphs

Work with Dr. A. Davoudi

Talk available online at
http://www.UTA.edu/UTARI/acs

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

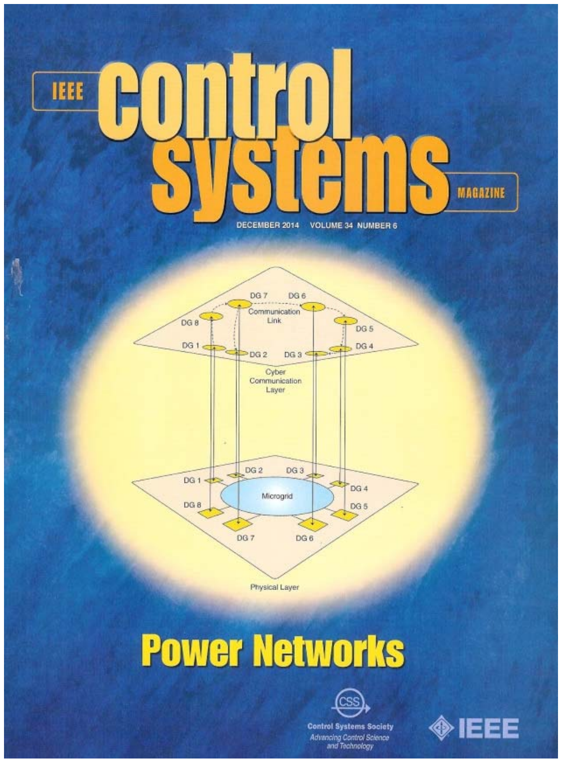
Lyapunov Functions and Performance Indices
Must depend on graph topology

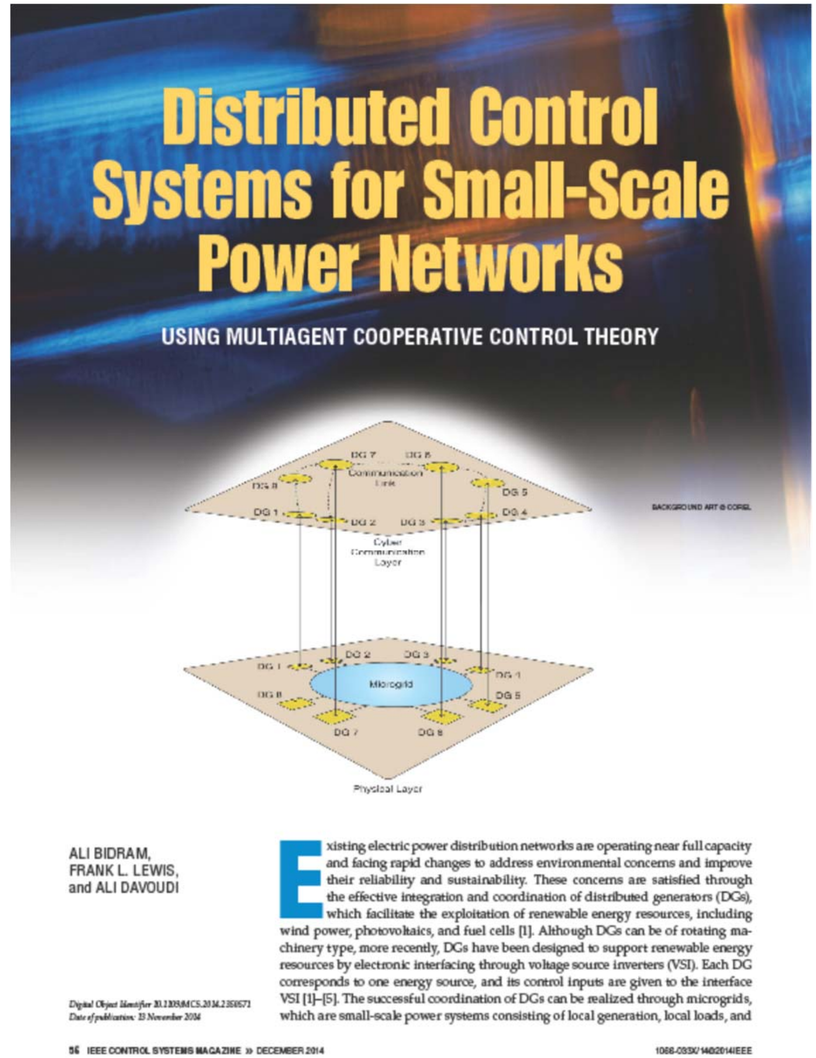
H. Zhang, F.L. Lewis, and Z. Qu, "Lyapunov, Adaptive, and Optimal Design Techniques for Cooperative Systems on Directed Communication Graphs," IEEE Trans. Industrial Electronics, vol. 59, no. 7, pp. 3026-3041, July 2012.

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.

A. Bidram, V. Nasirian, A. Davoudi, and F.L. Lewis, *Cooperative Synchronization in Distributed Microgrid Control*, Springer-Verlag, Berlin, 2017.

A. Bidram, F.L. Lewis, and A. Davoudi, “Distributed Control Systems for Small-scale Power Networks Using Multi-agent Cooperative Control Theory,” IEEE Control Systems Magazine, pp. 56-77, December 2014 ***(Featured cover article).***





Distributed Control Systems for Small-Scale Power Networks

USING MULTIAGENT COOPERATIVE CONTROL THEORY

ALI BIDRAM,
FRANK L. LEWIS,
and ALI DAVOUDI

Existing electric power distribution networks are operating near full capacity and facing rapid changes to address environmental concerns and improve their reliability and sustainability. These concerns are satisfied through the effective integration and coordination of distributed generators (DGs), which facilitate the exploitation of renewable energy resources, including wind power, photovoltaics, and fuel cells [1]. Although DGs can be of rotating machinery type, more recently, DGs have been designed to support renewable energy resources by electronic interfacing through voltage source inverters (VSI). Each DG corresponds to one energy source, and its control inputs are given to the interface VSI [1]–[5]. The successful coordination of DGs can be realized through microgrids, which are small-scale power systems consisting of local generation, local loads, and

Digital Object Identifier 10.1109/MCS.2014.2350571
Date of publication: 13 November 2014

1066-033X/14©2014IEEE


New Research Results
Distributed Cooperative Control on Graphs
Reinforcement Learning for Online Optimal Control
Multi-Player Games on Communication Graphs
Output Synchronization of Heterogeneous MAS
Applications to:
Building HVAC Balancing
AC Microgrid
DC Microgrid

The Power of Synchronization

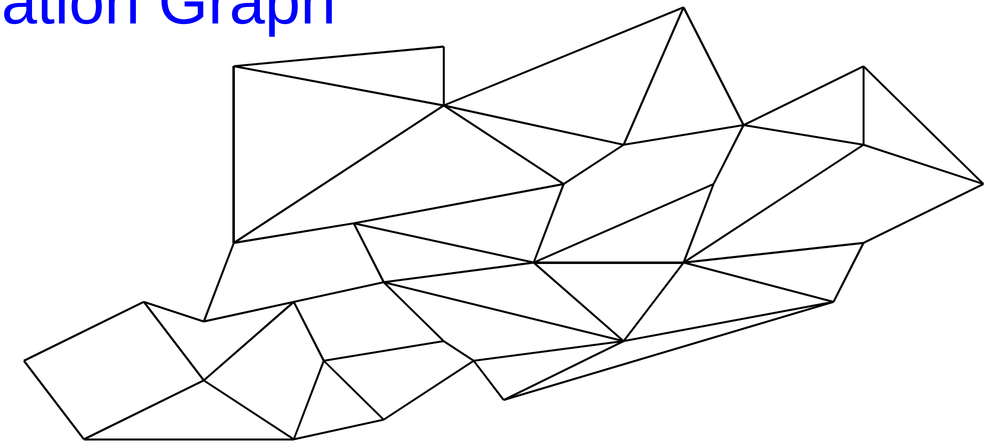
Coupled Oscillators
Diurnal Rhythm

Synchronization on Communication Graph

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

Synchronization problem

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

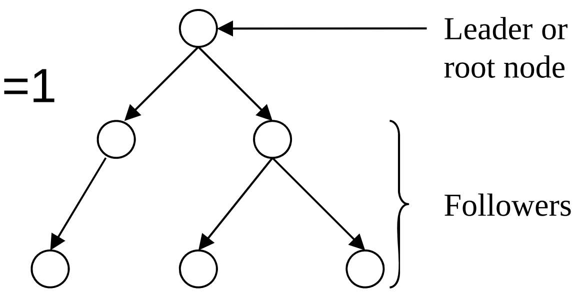


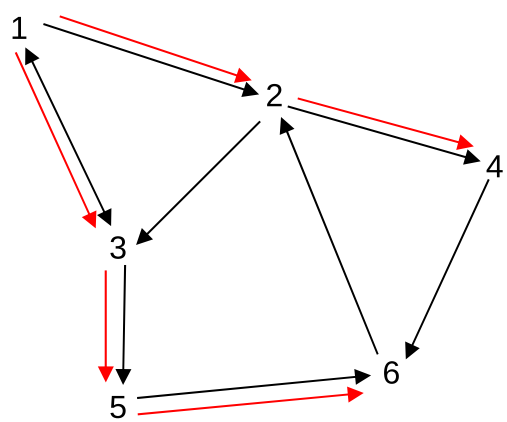
Strongly connected if for all nodes *i* and *j* there is a path from *i* to *j*.

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

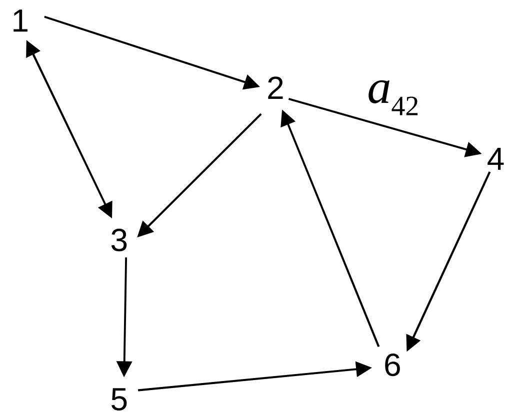
Strongly Connected implies exists Spanning Tree

Communication Graph

Algebraic Graph Theory

(V,E)

N nodes



Adjacency matrix

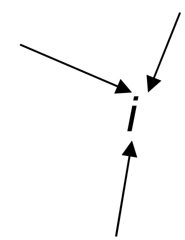
$$A = [a_{ij}]$$

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

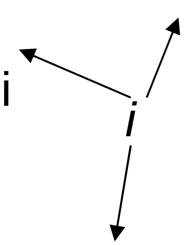
$$a_{ij} > 0 \; if \; (v_j, v_i) \in E$$

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

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

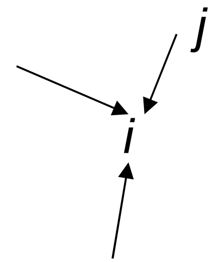


$$d_i^o = \sum_{j=1}^{N} a_{ji}$$ Col sum= out-degree $\quad N_o$ Out-neighbors of node i



Dynamic Graph- the Distributed Structure of Control

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



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

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

x

Global Form

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

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

L=D-A = graph Laplacian matrix

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

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

Distributed Controlled Consensus: Cooperative Tracker

Node state $\dot{x}_i = u_i$

Synchronization problem

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

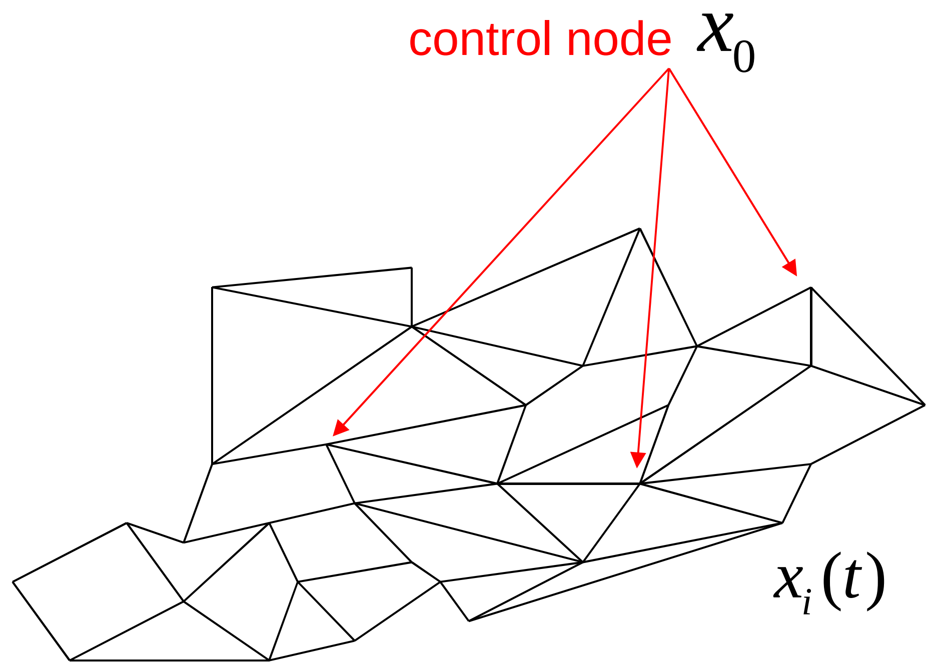
Distributed Local voting protocol

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

N_i is the set of immediate neighbors of agent i

$$u_i = -\left(g_i + \sum_{j \in \bar{N}_i} a_{ij} \right) x_i + \sum_{j \in N_i} a_{ij} x_j \quad + g_i v$$

$$\dot{x} = -(L+G)x + G\underline{1}x_0 \qquad G = diag\{g_i\}$$



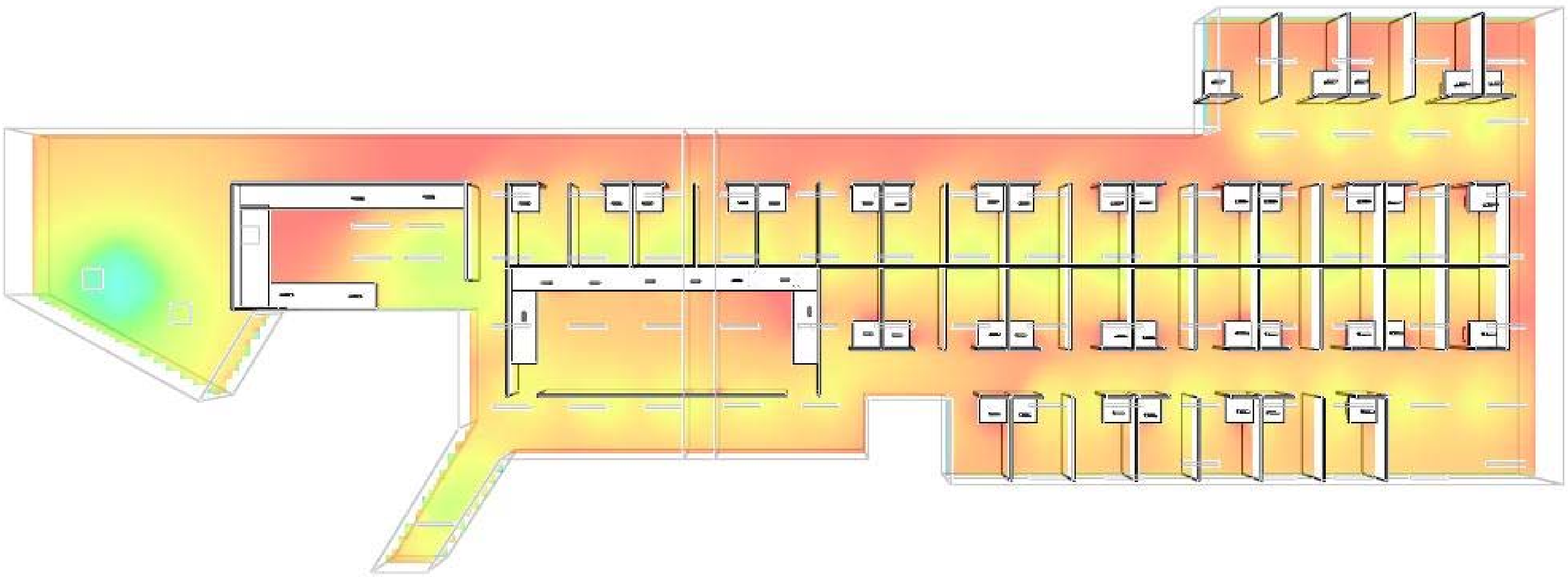
Sparse Communication Graph Topology

Highly efficient fast algorithms
Scalable to any nodes
Low communication overhead

Theorem. Let graph have a spanning tree
and $g_i \neq 0$ for at least one root node.
Then the distributed protocol makes $x_i(t) - x_j(t) \to 0$

Balancing Building HVAC Ventilation Systems

Work with SIMTech – Singapore Inst. Manufacturing Technology



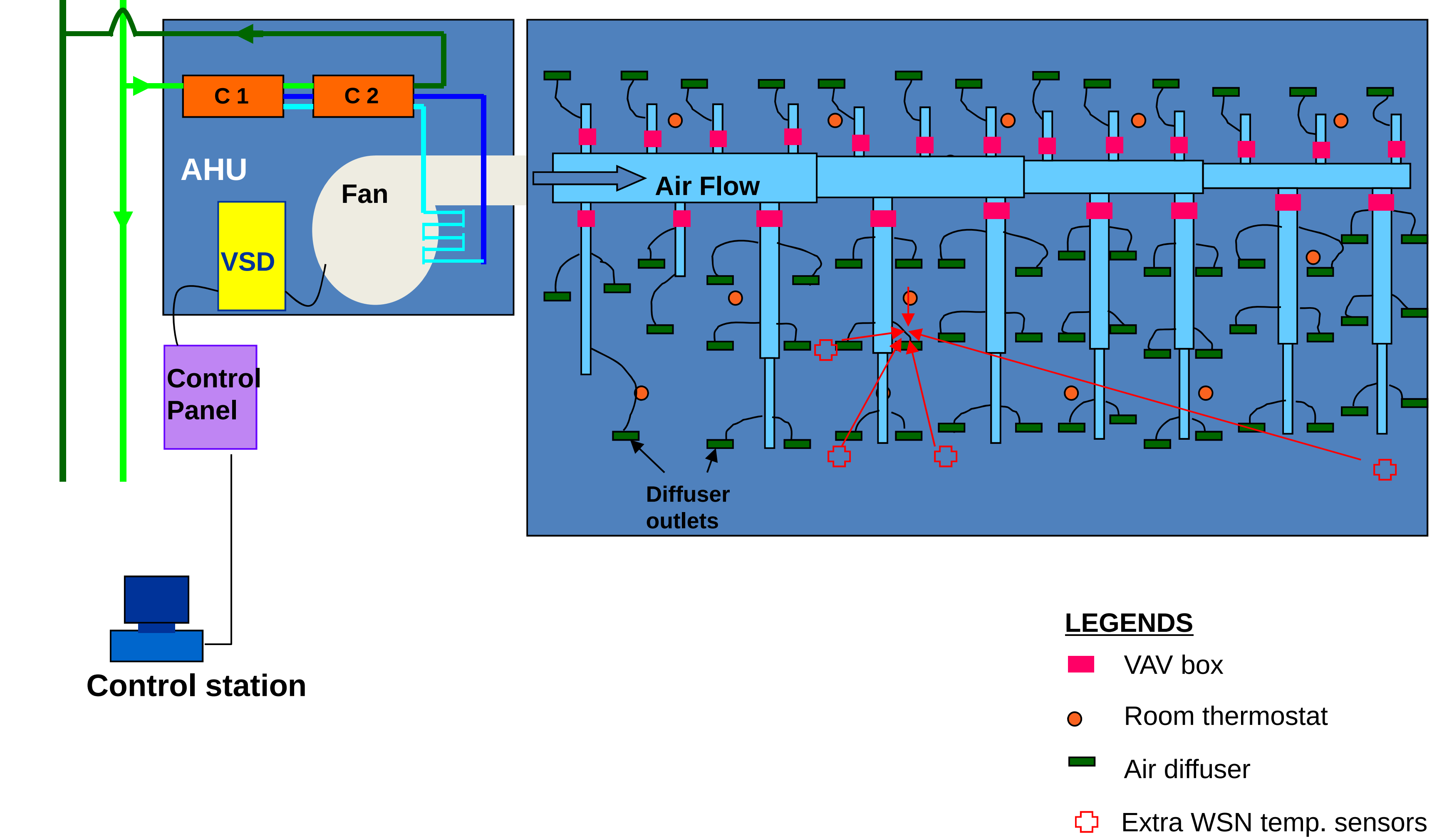
SIMTech 5th floor temperature distribution

SIMTech

Automated HVAC control system

$$x_i(k+1) = x_i(k) + f_i(x) + u_i(k)$$

$$u_i(k) = \frac{1}{n_i + 1}\gamma_i(k)\sum_{j \in N_i} a_{ij}(x_j(k) - x_i(k))$$



Adjust Dampers for desired Temperature distribution

Temperature dynamics

$$x_i(k+1) = x_i(k) + f_i(x) + u_i(k)$$

Unknown $f_i(x)$

Control damper position based on local voting protocol

$$\gamma_i(k) = 1, \tfrac{1}{2}, \tfrac{1}{4}, \ldots$$

$$u_i(k) = \frac{1}{n_i+1}\gamma_i(k)\sum_{j \in N_i} a_{ij}(x_j(k) - x_i(k))$$

Under mild conditions this converges to steady-state desired temp. distribution

Open Research Topic - HVAC Flow and Pressure control

Cooperative Control for AC Microgrids

What is a micro-grid?

- Micro-grid is a small-scale power system that provides the power for a group of consumers.
- Micro-grid enables local power support for local and critical loads.
- Micro-grid has the ability to work in both grid-connected and islanded modes.
- Micro-grid facilitates the integration of Distributed Energy Resources (DER).

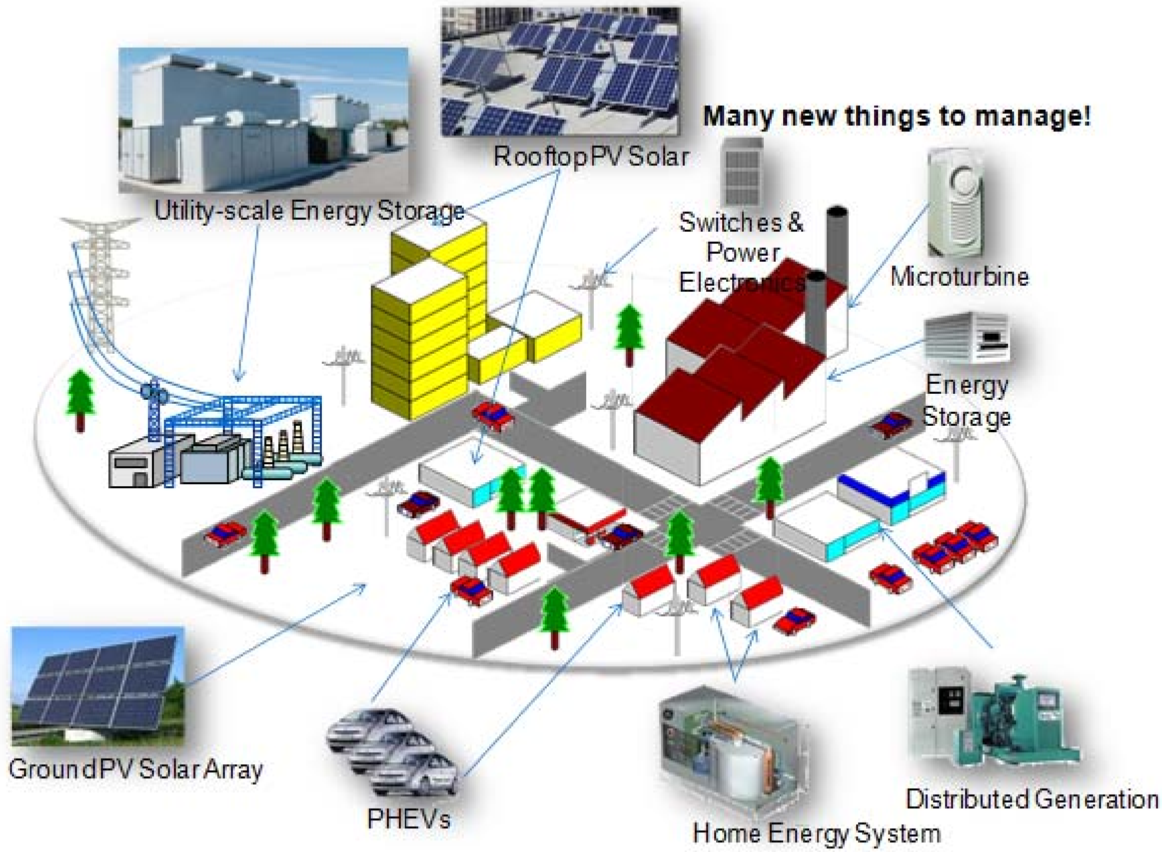
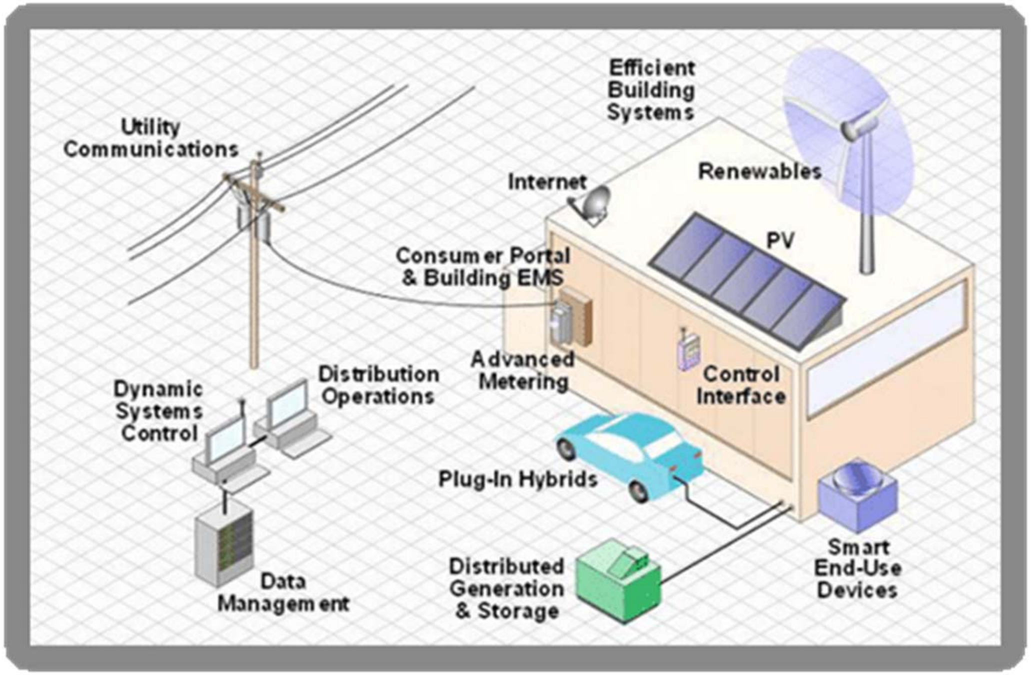


Photo from: http://www.horizonenergygroup.com

An introduction to micro-grids: Micro-grid applications

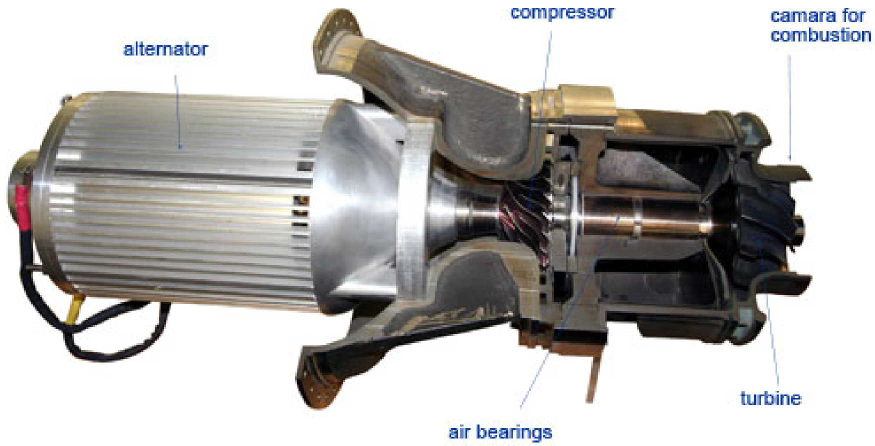
- The main building block of smart-grids
- Rural plants
- Business buildings, hospitals, and factories



Smart-grid photo from: http://www.sustainable-sphere.com

Distributed Generators (DG)
Distributed Energy Resources (DER)

- Non-renewables
  - Internal combustion engine
  - Micro-turbines
  - Fuel cells
- Renewables
  - Photovoltaic
  - Wind
  - Hydroelectric
  - Biomass



Micro-grid Advantages

- Micro-grid provides high quality and reliable power to the critical consumers

- During main grid disturbances, micro-grid can quickly disconnect form the main grid and provide reliable power for its local loads

- DGs can be simply installed close to the loads which significantly reduces the power transmission line losses

- By using renewable energy resources, a micro-grid reduces CO2 emissions

AC vs. DC Microgrids

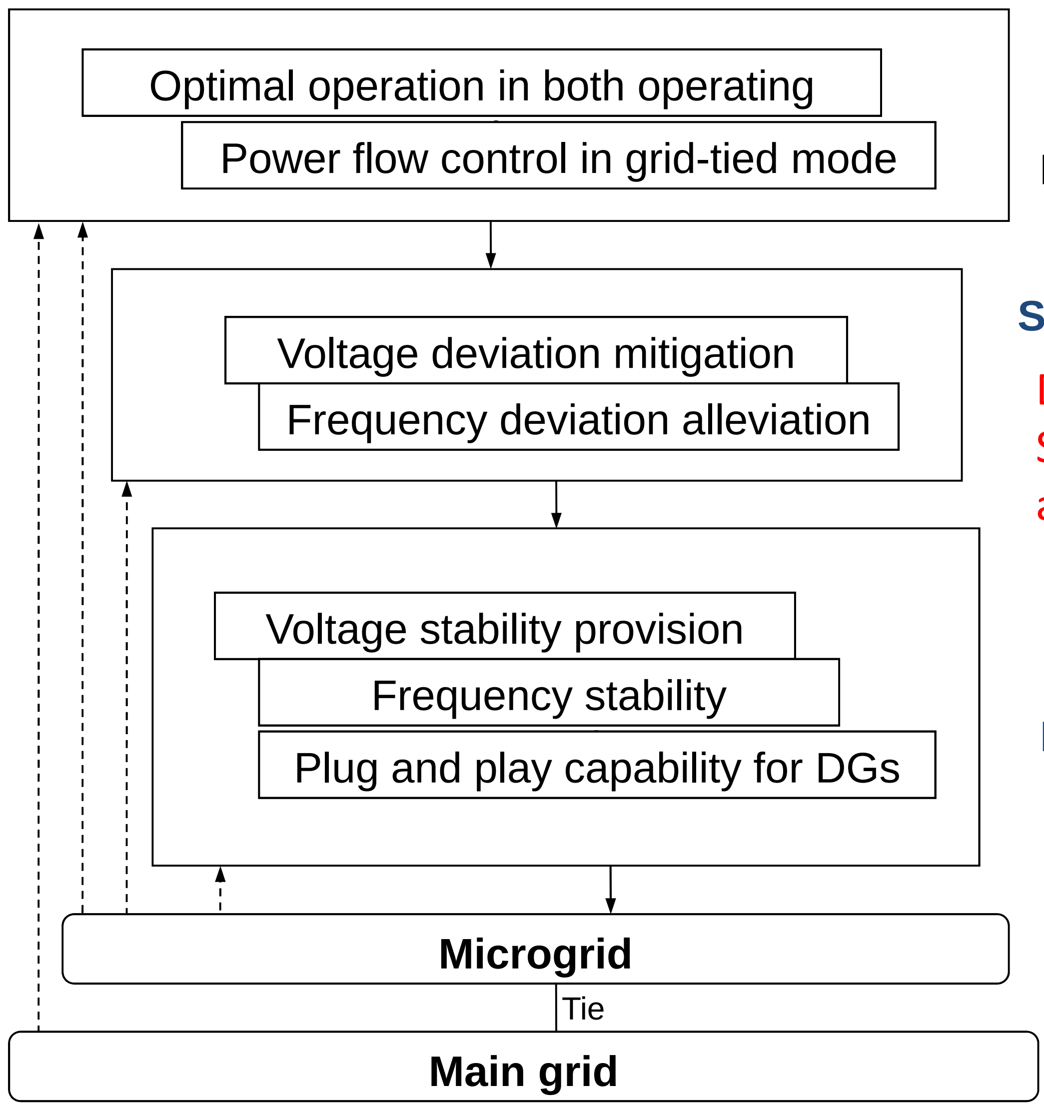
An introduction to AC micro-grids: Micro-grid Objectives

- Voltage and frequency synchronization for both
  grid-connected and islanded operating modes
- Proper load sharing and DG coordination
- Power flow control between the microgrid and the main grid
- Optimizing the microgrid operating cost

Hierarchical control structure

Bidram, A., & Davoudi, A. (2012). Hierarchical structure of microgrids control system. *IEEE Transactions on Smart Grid*, vol. 3, pp. 1963-1976, Dec 2012.

Micro-grid Hierarchical Control Structure



Tertiary Control-
Economic dispatch

Secondary Control

Do coop. ctrl. here to Synchronize frequency and voltage

Primary Control

Maintains Stability of individual DGs

Standard Micro-grid secondary control

- **Secondary control**: The secondary control restores the voltage and frequency of the micro-grid to their nominal value after islanding.

Conventional Secondary control implementation:

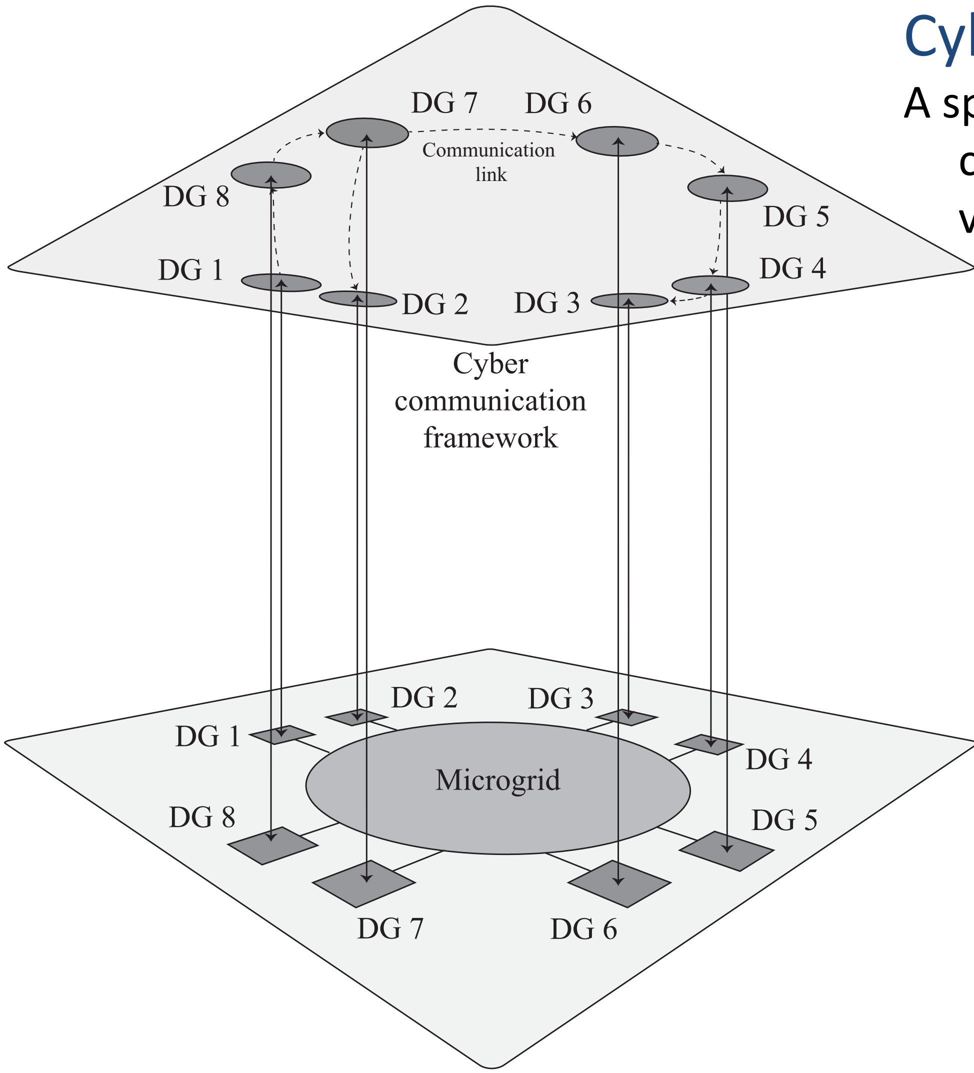
- Centralized structure
  - Low reliability – single point of failure
  - Requires a Central control authority
  - Requires too many communication links
  - Not scalable to many DGs

$$\begin{cases} \delta V_n = K_{PE}(v_{ref} - v_{mag}) + K_{IE}\int (v_{ref} - v_{mag})dt \\ \delta \omega_n = K_{P\omega}(\omega_{ref} - \omega) + K_{I\omega}\int (\omega_{ref} - \omega)dt \end{cases}$$

- We want to develop a new Distributed Control structure
  - Highly reliable
  - Uses sparse communication network

Micro-grid secondary control:
New Distributed CPS structure

Work of Ali Bidram
With Dr. A. Davoudi



Cyber layer

A sparse, efficient communication network to allow cooperative control for synchronization of voltage and frequency

Secondary Control

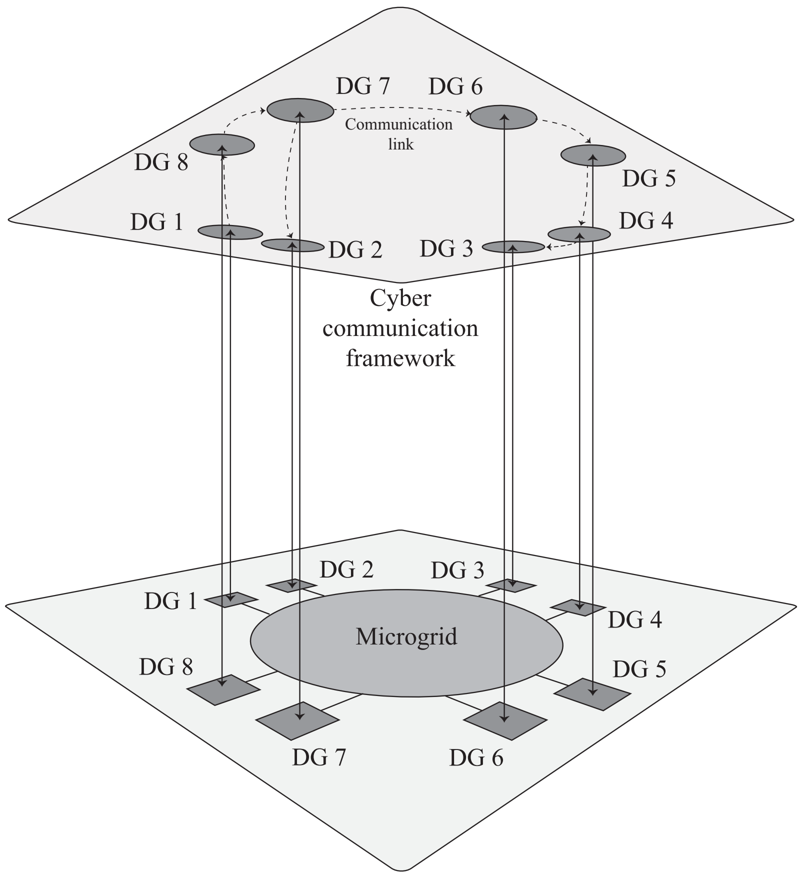
Cyber Physical System (CPS)

Physical Layer

The interconnect structure of the power grid

Primary Control

The Importance of the Communication Network - Interactions Between Communication and Control



Cyber layer

A sparse, efficient communication network to allow cooperative control for synchronization of voltage and frequency

Cyber Physical System (CPS)

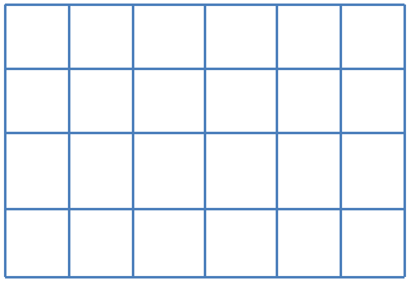
Physical Layer

The interconnect structure of the power grid

Local controller design must take into account the Graph Topology

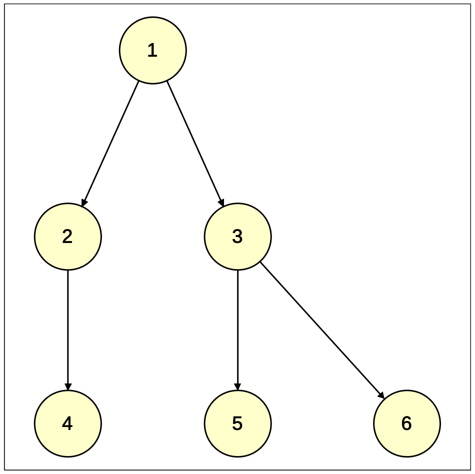
Synchronization on Good Graphs

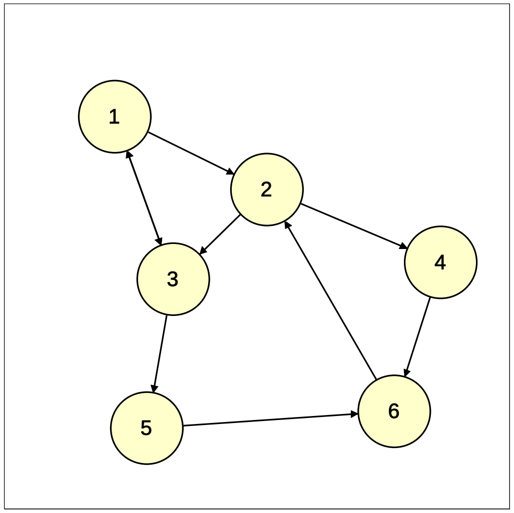
Chris Elliott fast video



Regular mesh

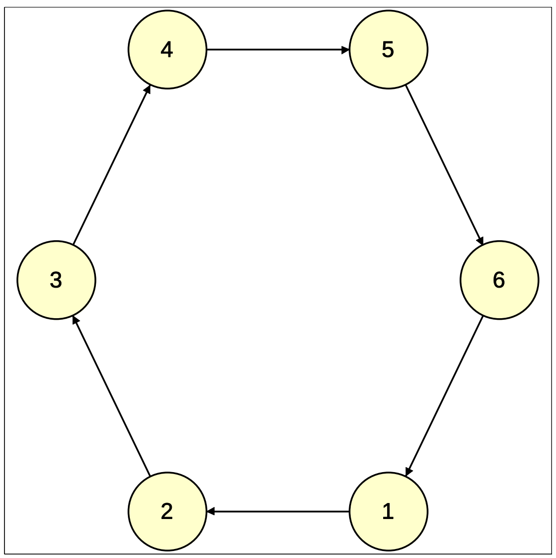
Synchronization Speed depends on communication topology





Synchronization on Gossip Rings

Chris Elliott weird video



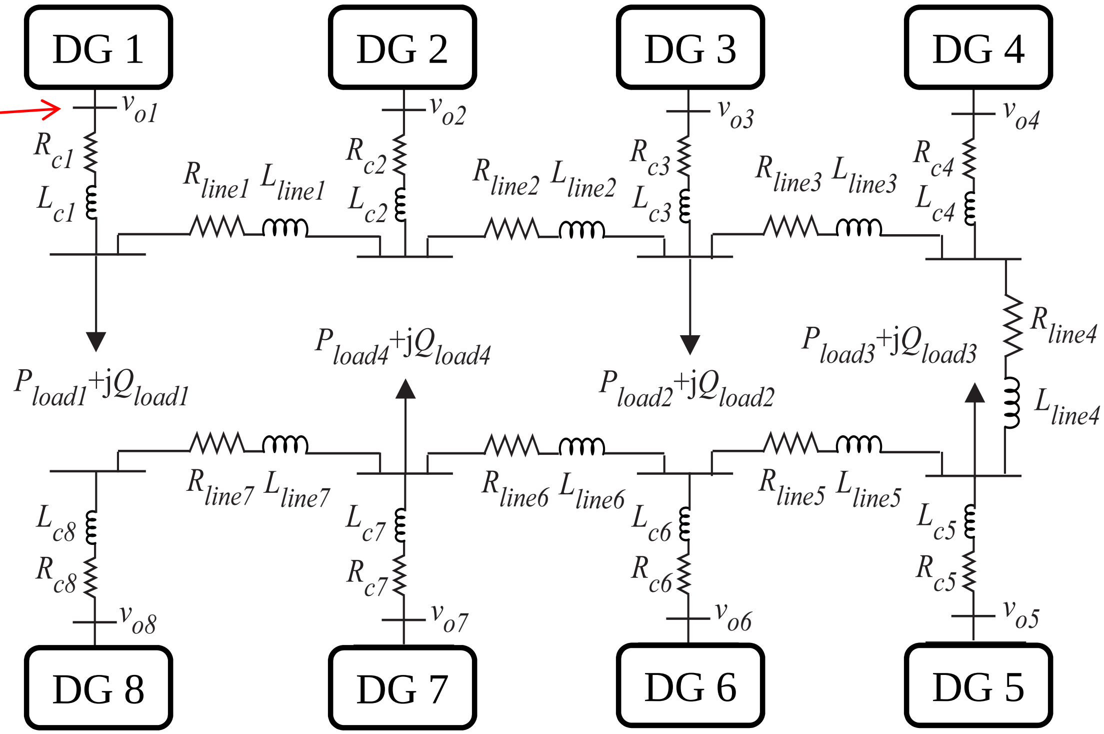
Graph Laplacian L has complex eigenvalues

1. Distributed secondary frequency control of micro-grids
2. Distributed secondary voltage control of micro-grids

Work of Ali Bidram
With Dr. A. Davoudi

Synchronization in AC Microgrid of Interconnected DG

Voltage synchronization



Frequency synchronization $\quad y_i = \omega_i = \omega_{ni} - m_{Pi} P_i$

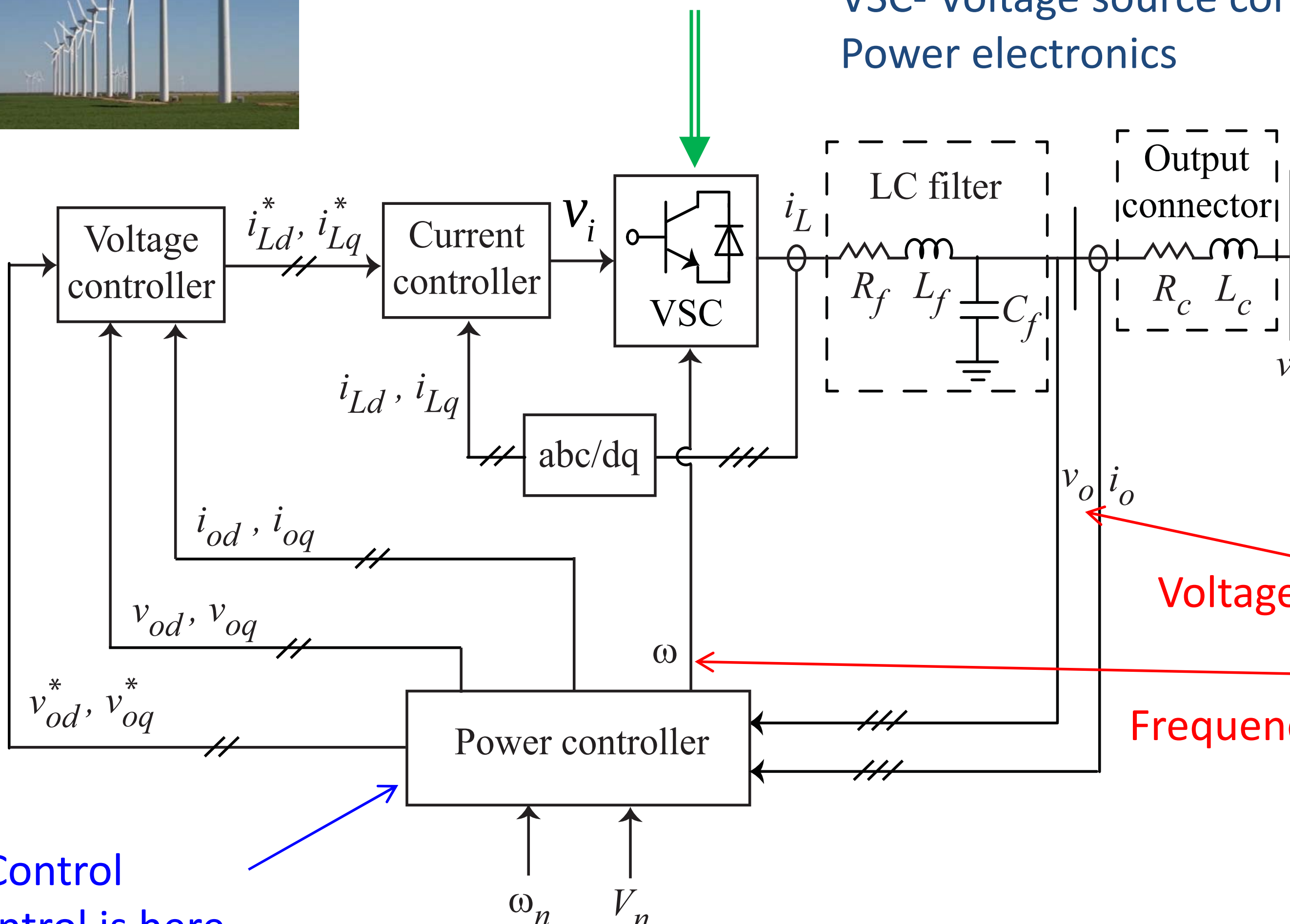
Voltage synchronization (per unit) $\quad E \equiv v_{odi} = V_{ni} - n_{Qi} Q_i$

Dynamical model of a DG

Primary Control Structure

Renewable DER
Provides DC voltage

VSC- Voltage source converter
Power electronics



Microgird
Network
Load
disturbances

Voltage synchronization

Frequency synchronization

Primary Control
Droop control is here

Given load conditions v_0, i_0 - pick $\omega, v^*_{od}, v^*_{oq}$ using Droop to maintain stability

Pogaku, N., Prodanovic, M., & Green, T. C. (2007). Modeling, analysis and testing of autonomous operation of an inverter-based microgrid. *IEEE Transactions on Power Electronics*, *22*(2), 613–625.

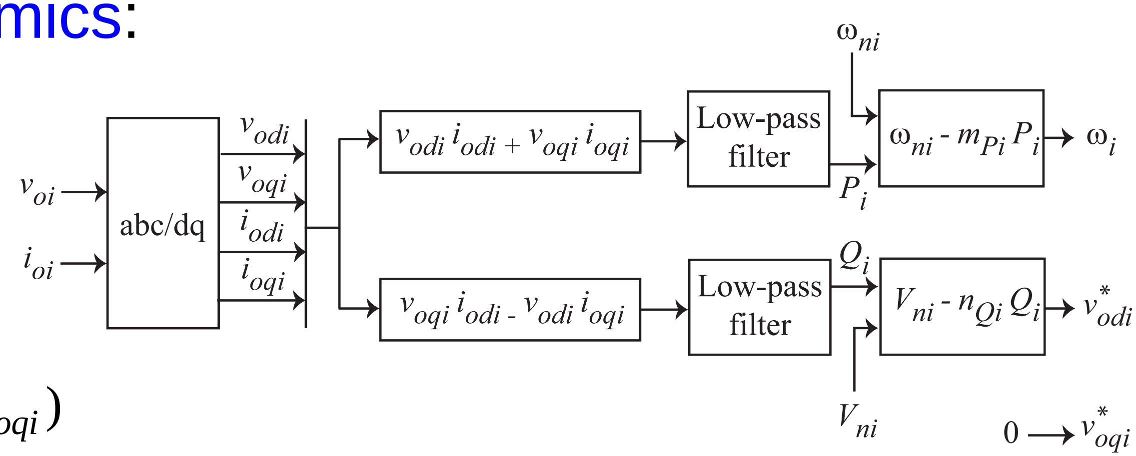
Dynamical model of a DG

$$\begin{cases} \dot{\mathbf{x}}_i = \mathbf{f}_i(\mathbf{x}_i) + \mathbf{k}_i(\mathbf{x}_i)\mathbf{D}_i + \mathbf{g}_i(\mathbf{x}_i)u_i \\ y_i = h_i(\mathbf{x}_i) + d_i u_i \end{cases}$$

Heterogeneous agent dynamics

$$\mathbf{x}_i \in \mathbb{R}^{13} \qquad \mathbf{x}_i = [\delta_i \quad P_i \quad Q_i \quad \phi_{di} \quad \phi_{qi} \quad \gamma_{di} \quad \gamma_{qi} \quad i_{ldi} \quad i_{lqi} \quad v_{odi} \quad v_{oqi} \quad i_{odi} \quad i_{oqi}]^T$$

Power controller dynamics:



$$\dot{\delta}_i = \omega_i - \omega_{com}$$

$$\dot{P}_i = -\omega_{ci}P_i + \omega_{ci}(v_{odi}i_{odi} + v_{oqi}i_{oqi})$$

$$\dot{Q}_i = -\omega_{ci}Q_i + \omega_{ci}(v_{oqi}i_{odi} - v_{odi}i_{oqi})$$

Droop control is here

Dynamical model of a DG

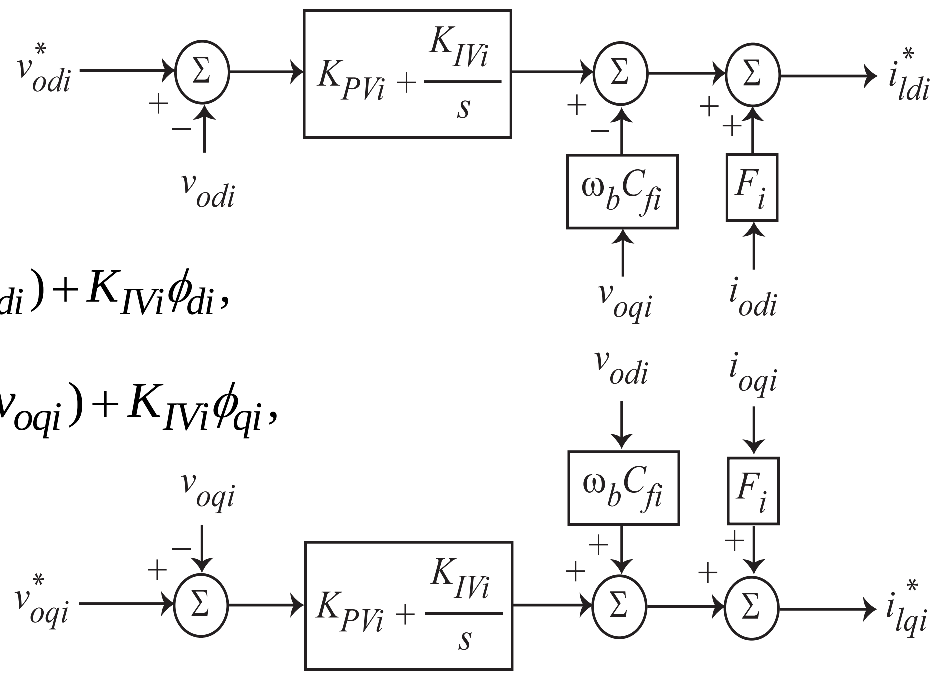
Voltage controller dynamics

$$\dot{\phi}_{di} = v^*_{odi} - v_{odi},$$

$$\dot{\phi}_{qi} = v^*_{oqi} - v_{oqi},$$

$$i^*_{ldi} = F_i i_{odi} - \omega_b C_{fi} v_{oqi} + K_{PVi}(v^*_{odi} - v_{odi}) + K_{IVi}\phi_{di},$$

$$i^*_{lqi} = F_i i_{oqi} + \omega_b C_{fi} v_{odi} + K_{PVi}(v^*_{oqi} - v_{oqi}) + K_{IVi}\phi_{qi},$$



Dynamical model of a DG

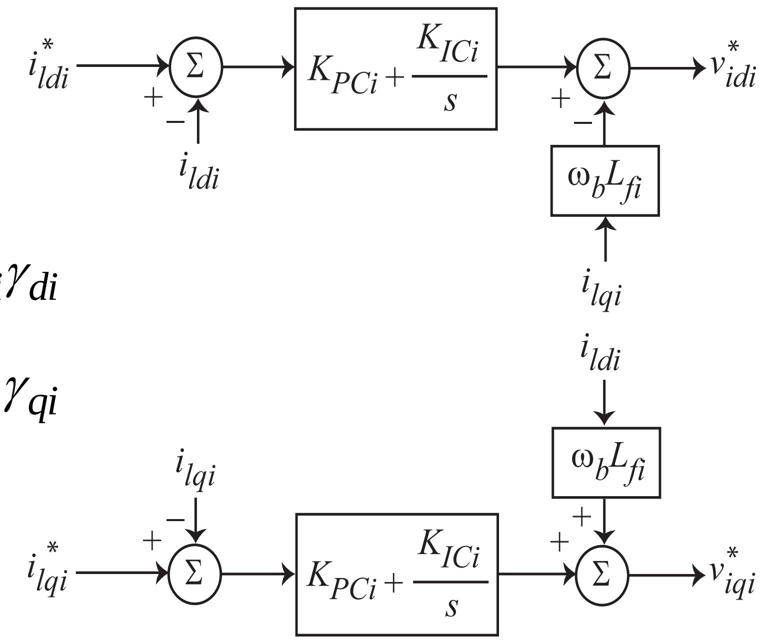
Current controller dynamics

$$\dot{\gamma}_{di} = i_{ldi}^* - i_{ldi}$$

$$\dot{\gamma}_{qi} = i_{lqi}^* - i_{lqi}$$

$$v_{idi}^* = -\omega_b L_{fi} i_{lqi} + K_{PCi}(i_{ldi}^* - i_{ldi}) + K_{ICi}\gamma_{di}$$

$$v_{iqi}^* = \omega_b L_{fi} i_{ldi} + K_{PCi}(i_{lqi}^* - i_{lqi}) + K_{ICi}\gamma_{qi}$$



Dynamical model of a DG

Output filter dynamics

$$\dot{i}_{ldi} = -\frac{R_{fi}}{L_{fi}} i_{ldi} + \omega_i i_{lqi} + \frac{1}{L_{fi}} v_{idi} - \frac{1}{L_{fi}} v_{odi}$$

$$\dot{i}_{lqi} = -\frac{R_{fi}}{L_{fi}} i_{lqi} - \omega_i i_{ldi} + \frac{1}{L_{fi}} v_{iqi} - \frac{1}{L_{fi}} v_{oqi}$$

$$\dot{v}_{odi} = \omega_i v_{oqi} + \frac{1}{C_{fi}} i_{ldi} - \frac{1}{C_{fi}} i_{odi}$$

$$\dot{v}_{oqi} = -\omega_i v_{odi} + \frac{1}{C_{fi}} i_{lqi} - \frac{1}{C_{fi}} i_{oqi}$$

Output connector dynamics

$$\dot{i}_{odi} = -\frac{R_{ci}}{L_{ci}} i_{odi} + \omega_i i_{oqi} + \frac{1}{L_{ci}} v_{odi} - \frac{1}{L_{ci}} v_{bdi}$$

$$\dot{i}_{oqi} = -\frac{R_{ci}}{L_{ci}} i_{oqi} - \omega_i i_{odi} + \frac{1}{L_{ci}} v_{oqi} - \frac{1}{L_{ci}} v_{bqi}$$

Depends on microgrid conditions and loads

Voltage disturbances

The nonlinear dynamics of the i^{th} DG, while neglecting the fast dynamics of voltage and current controllers

$$
\begin{cases}
\dot{x}_{i,1} = \omega_{ni} - m_{Pi} x_{i,2} - \omega_{com} \\
\dot{x}_{i,2} = \omega_{ci} (x_{i,6} x_{i,8} + x_{i,7} x_{i,9} - x_{i,2}) \\
\dot{x}_{i,3} = \omega_{ci} (x_{i,6} x_{i,9} - x_{i,7} x_{i,8} - x_{i,3}) \\
\dot{x}_{i,4} = \dfrac{-r_{fi}}{L_{fi}} x_{i,4} + \omega_{com} x_{i,5} + \dfrac{V_{ni} - n_{Qi} x_{i,3} - x_{i,6}}{L_{fi}} \\
\dot{x}_{i,5} = \dfrac{-r_{fi}}{L_{fi}} x_{i,5} - \omega_{com} x_{i,4} - \dfrac{x_{i,7}}{L_{fi}} \\
\dot{x}_{i,6} = \omega_{com} x_{i,7} + \dfrac{x_{i,4} - x_{i,8}}{C_{fi}} \\
\dot{x}_{i,7} = -\omega_{com} x_{i,6} + \dfrac{x_{i,5} - x_{i,9}}{C_{fi}} \\
\dot{x}_{i,8} = \dfrac{-r_{ci}}{L_{ci}} x_{i,8} + \omega_{com} x_{i,9} + \dfrac{x_{i,6} - v_{bdi}}{L_{ci}} \\
\dot{x}_{i,9} = \dfrac{-r_{ci}}{L_{ci}} x_{i,9} - \omega_{com} x_{i,8} + \dfrac{x_{i,7} - v_{bqi}}{L_{ci}}
\end{cases}
$$

$$x_i = [\alpha_i \quad P_i \quad Q_i \quad i_{Ldi} \quad i_{Lqi} \quad v_{odi} \quad v_{oqi} \quad i_{odi} \quad i_{oqi}]^T .$$

DG Microgrid Model and Synchronization Control Objectives

Heterogeneous Agent Dynamics – different dynamics

$$\begin{cases} \dot{\mathbf{x}}_i = \mathbf{f}_i(\mathbf{x}_i) + \mathbf{k}_i(\mathbf{x}_i)\mathbf{D}_i + \mathbf{g}_i(\mathbf{x}_i)u_i \\ y_i = h_i(\mathbf{x}_i) + d_i u_i \end{cases} \qquad \mathbf{x}_i \in \mathbb{R}^{13}$$

1. For secondary frequency control: synchronize ω_i

$$y_i = \omega_i = \omega_{ni} - m_{Pi}P_i$$

$$u_i = \omega_{ni} \quad \Rightarrow \quad d_i \neq 0$$

2. For secondary voltage control: synchronize v_{odi}

$$E \equiv v_{odi} = V_{ni} - n_{Qi}Q_i \quad \Rightarrow \quad y_i = v_{oi}, \quad u_i = V_{ni} \quad \Rightarrow \quad d_i = 0$$

Droop Control in Primary Loop

Micro-grid Primary Control

How to Synchronize? Look at Power controller dynamics

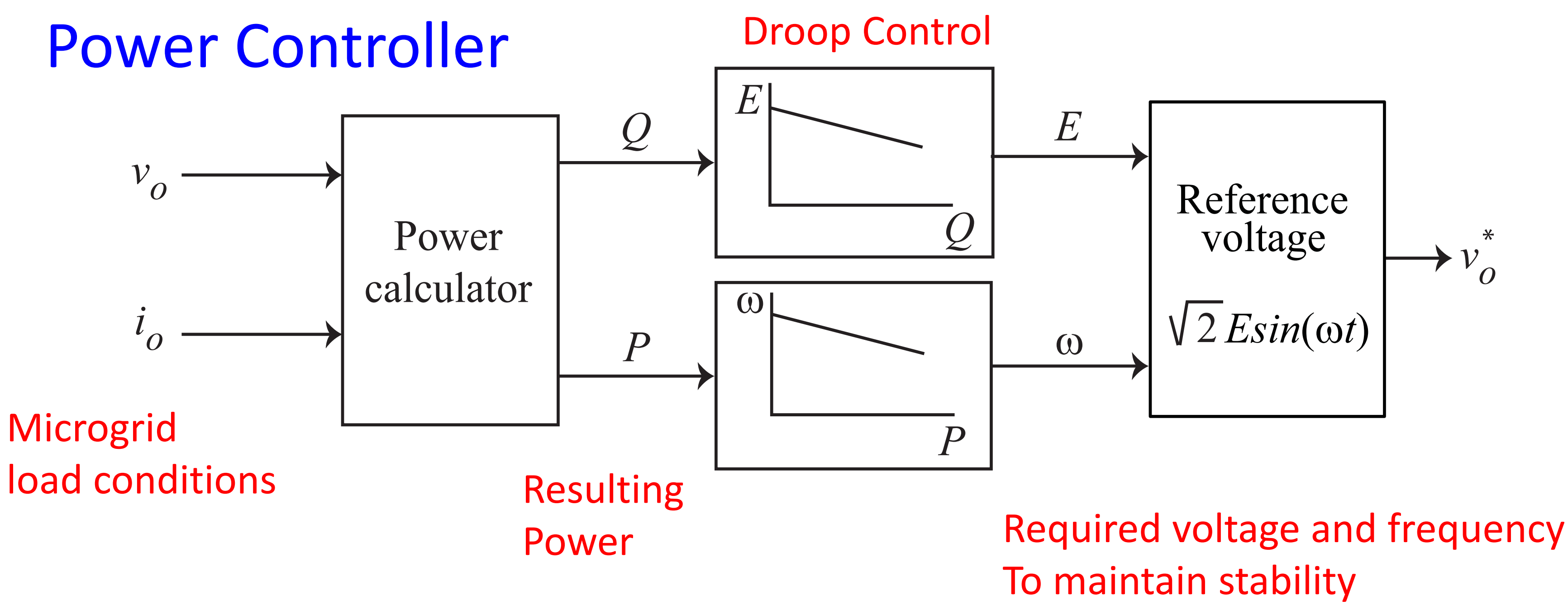
- ❑ **Primary control**: The primary control maintains voltage and frequency stability
- ❑ **Conventional primary control**: Droop techniques

$$\begin{cases} \omega = \omega_n - m_P P \\ E \equiv v_{od} = V_n - n_Q Q \end{cases}$$

$$m_{P1} P_{\max 1} = \cdots = m_{PN} P_{\max N}$$

$$n_{Q1} Q_{\max 1} = \cdots = n_{QN} Q_{\max N}$$

Power Controller

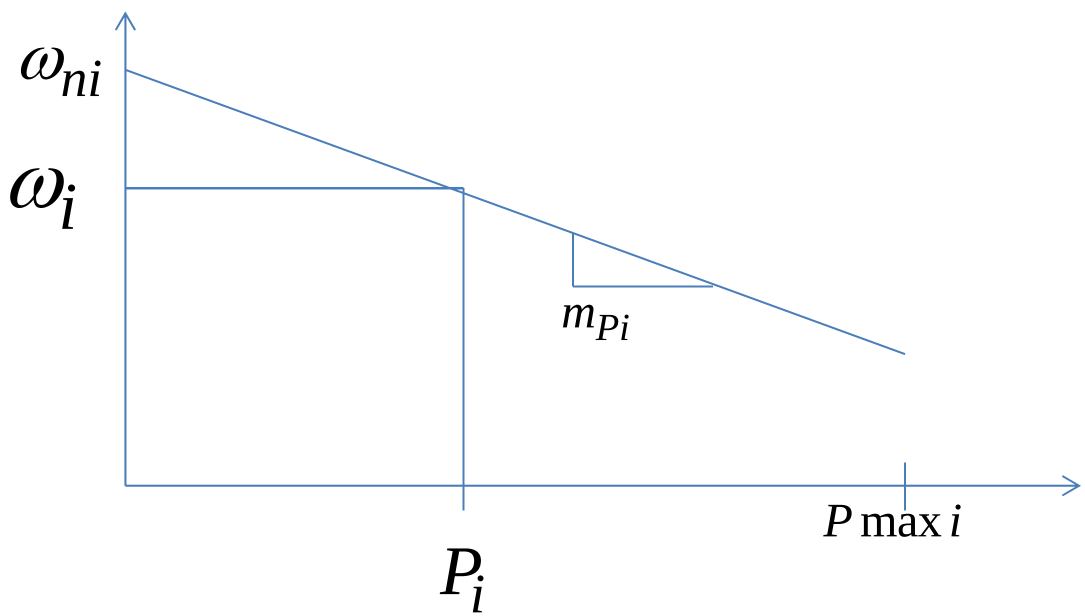


1. Secondary Frequency Control

New Secondary Control Input for Frequency Synchronization

$$\omega_i = \omega_{ni} - m_{Pi} P_i$$

Secondary
Control input



Prescribed frequency
e.g. 60 Hz

Change ω_{ni}

To synchronize ω_i

Existing power conditions in the microgrid

Primary Droop Control

1. Secondary Frequency Control

Droop control relationship

$$\omega_i = \omega_{ni} - m_{Pi} P_i$$

Using input-output feedback linearization

$$\dot{\omega}_i = \dot{\omega}_{ni} - m_{Pi} \dot{P}_i \equiv u_{\omega i}$$

$$u_{\omega i} = -c_{\omega i} e_{\omega i}$$

$$e_{\omega i} = \sum_{j \in N_i} a_{ij} (\omega_i - \omega_j) + g_i (\omega_i - \omega_{ref})$$

Then $$\dot{\omega}_{ni} = m_{Pi} \dot{P}_i + u_{\omega i} = m_{Pi} \dot{P}_i - c_{\omega i} \sum_{j \in N_i} a_{ij} (\omega_i - \omega_j) + g_i (\omega_i - \omega_{ref})$$

Theorem . Let the digraph of the communication network have a spanning tree and the pinning gain be nonzero for at least one DG placed on a root node.

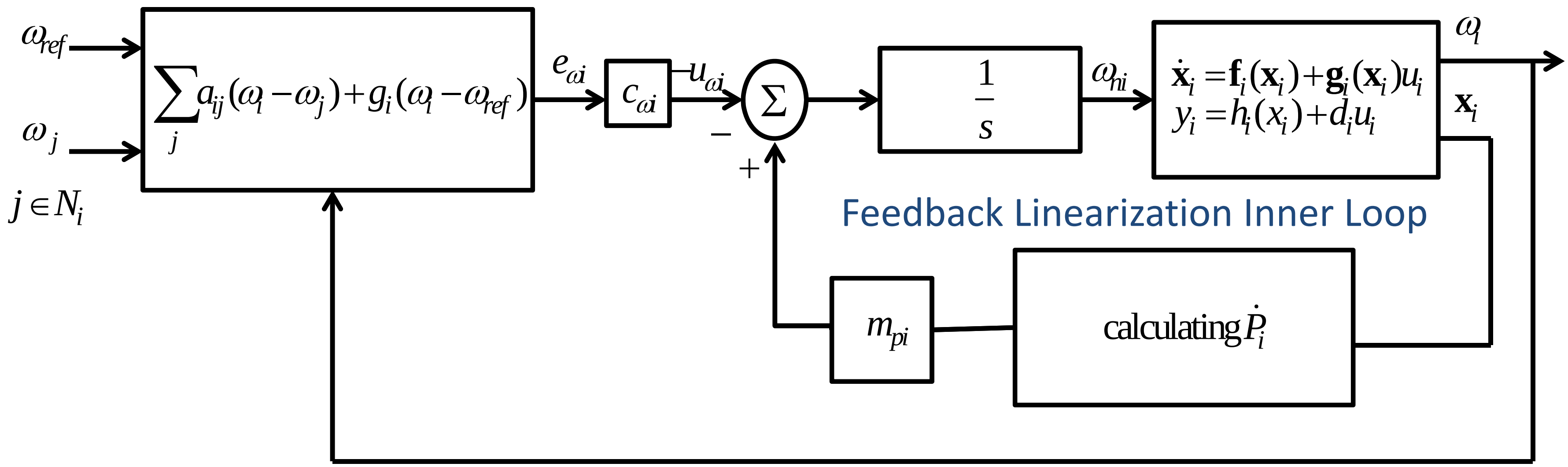
Let the auxiliary control $u_{\omega i}$ be chosen as above.

Then, the global neighborhood error is asymptotically stable.
Moreover, the DG frequencies synchronize to ω_{ref}

1. Secondary Frequency Control

Restores Frequency Synchronization after islanding

$$\dot{\omega}_i = \dot{\omega}_{ni} - m_{Pi}\dot{P}_i \equiv u_{\omega i}$$

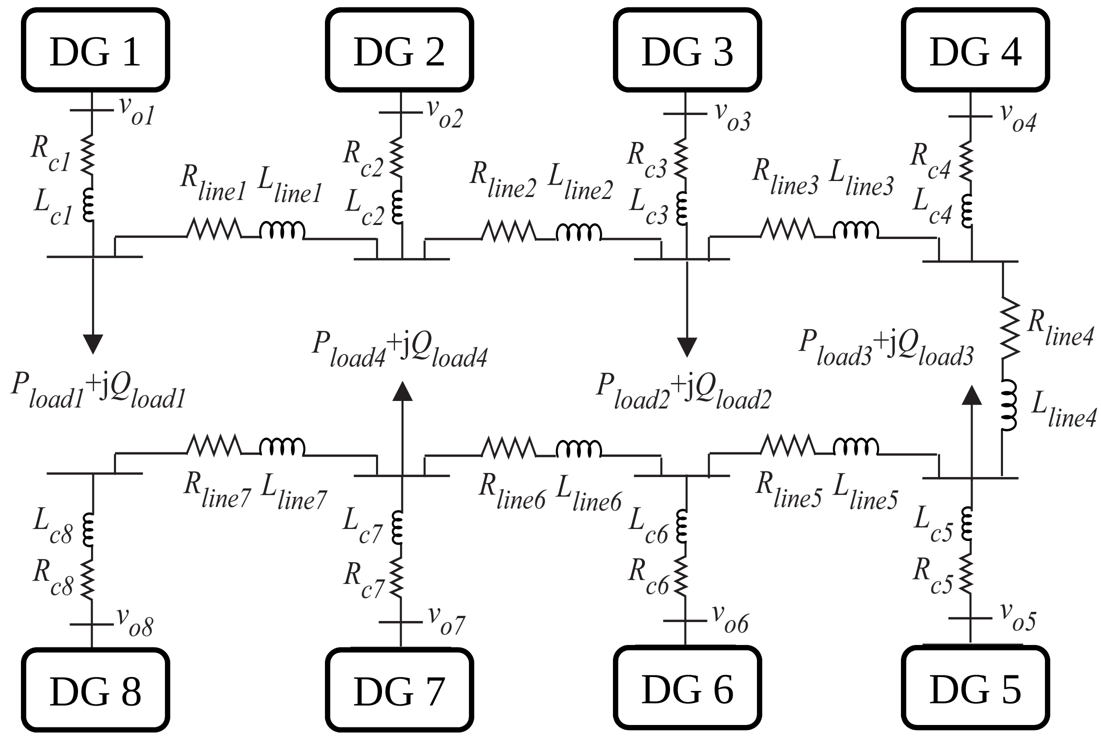


Distributed Cooperative Tracker

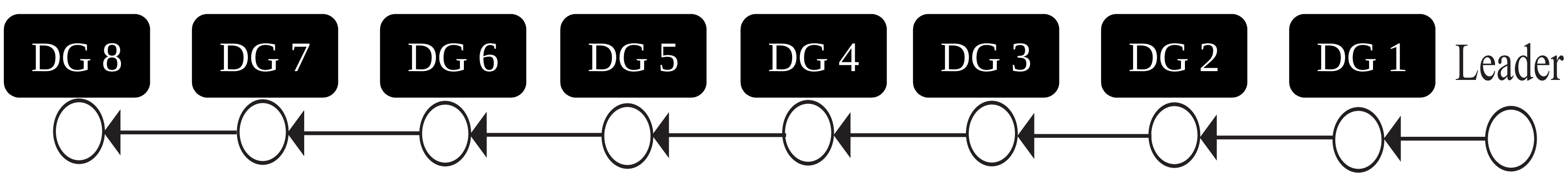
1. Secondary frequency control

Simulation Example

Physical Microgrid Network

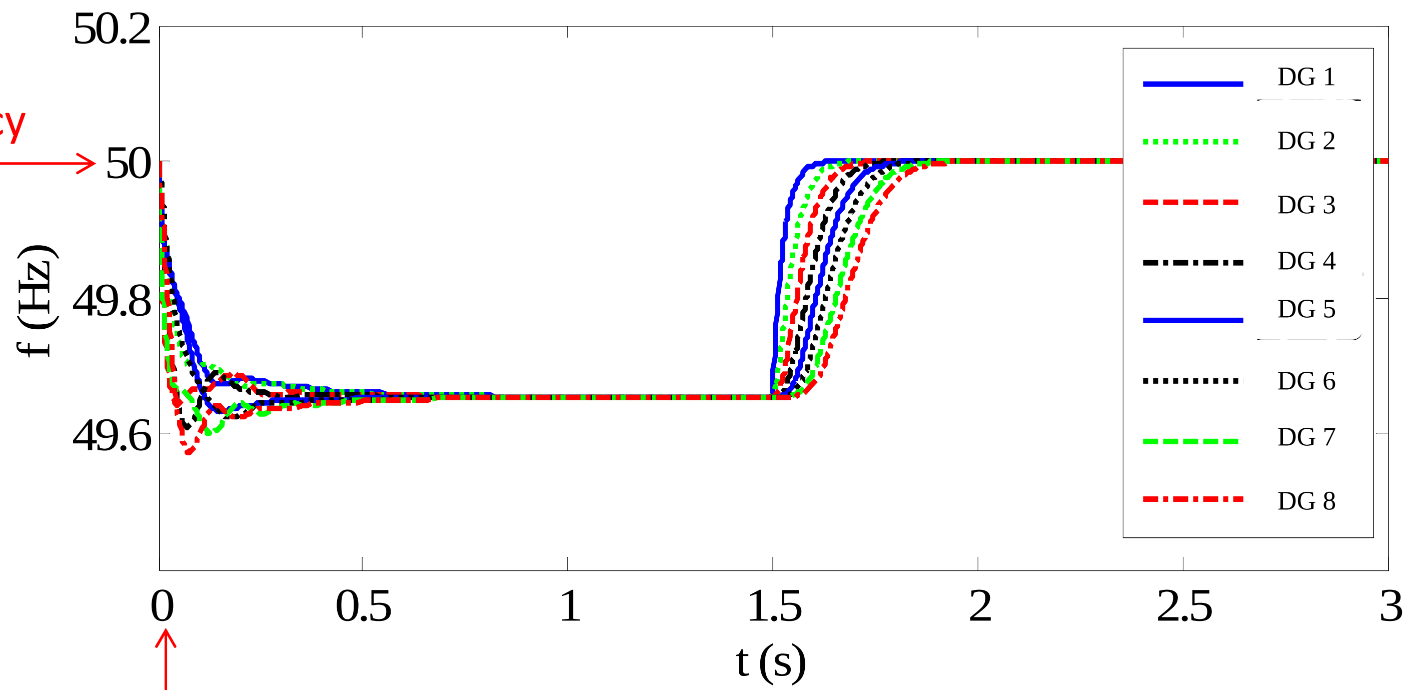


Cyber communication network- sparse



1. Secondary frequency control

Ref. Frequency
Is 50 Hz



Islanding

Turn on
Coop secondary control

Secondary frequency and power control

The local neighborhood tracking error control

$$u_i = -c\Big(\sum_{j \in N_i} a_{ij}(\omega_i - \omega_j) + g_i(\omega_i - \omega_{ref}) + \sum_{j \in N_i} a_{ij}(m_{Pi}P_i - m_{Pj}P_j)\Big)$$

Guarantees that $-c((L+G)(\omega - \underline{\omega}_{ref}) + Lm_P P) = 0.$

This does not guarantee synchronization of freq. and power separately

---

There is another relation between power and frequency in the microgrid DGs

Write the output active power as $P_i = \frac{|v_{oi}||v_{bi}|}{X_{ci}} \sin(\delta_i) \equiv h_i \sin(\delta_i),$

So that approximately $\dot{P}_i = h_i(\omega_i - \omega_{ref}),$

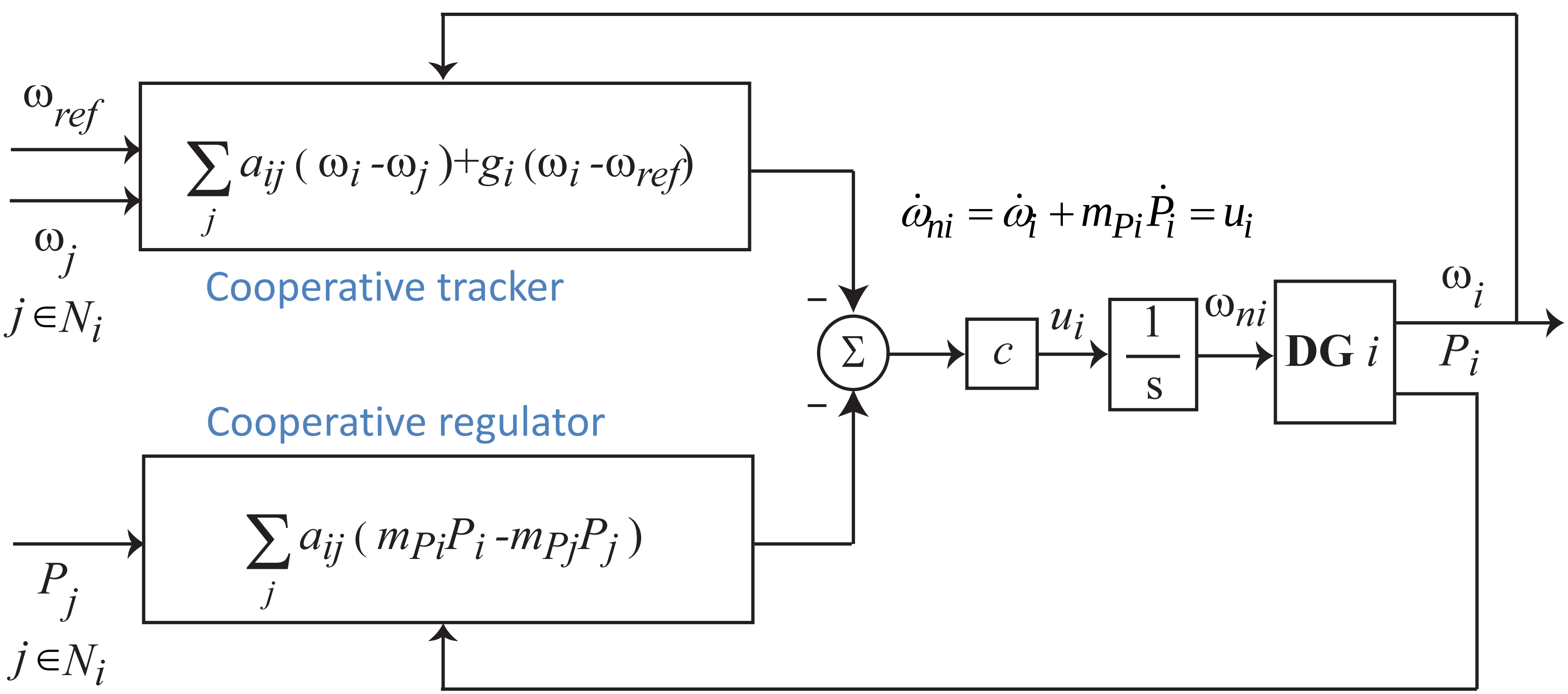
In global form $\dot{P} = h(\omega - \underline{\omega}_{ref}),$

Therefore at steady state all frequencies synchronize to the reference frequency

So that $G(\omega - \underline{\omega}_{ref}) + Lm_P P = 0$

Secondary frequency and power control

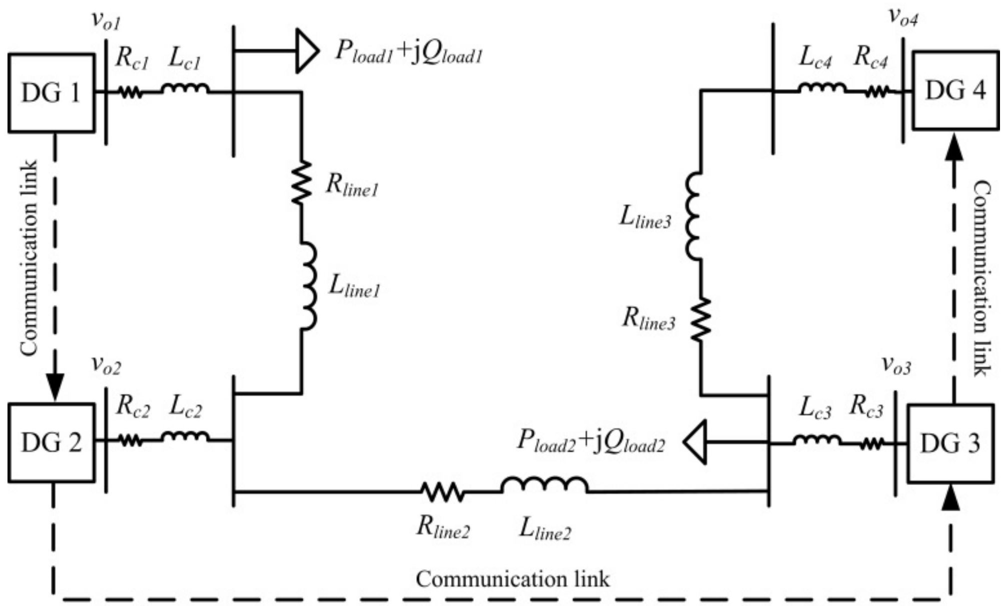
$$u_i = -c\Big(\underbrace{\sum_{j\in N_i} a_{ij}(\omega_i - \omega_j) + g_i(\omega_i - \omega_{ref})}_{\text{Frequency synchronization}} + \underbrace{\sum_{j\in N_i} a_{ij}(m_{Pi}P_i - m_{Pj}P_j)}_{\text{Power consensus}}\Big)$$



TWO CONTROL OBJECTIVES WITH ONE CONTROL INPUT

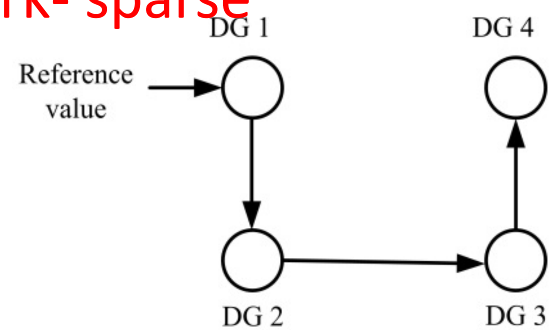
Simulation results

Physical Microgrid Network



(a)

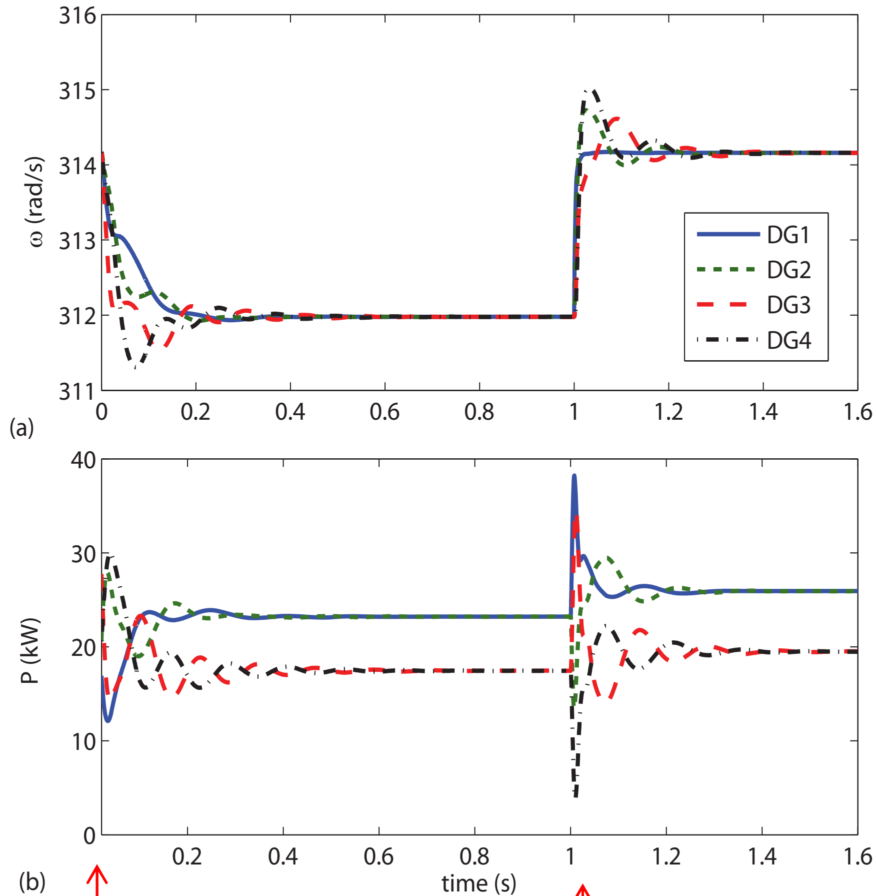
Cyber communication network- sparse



(b)

Simulation results

Ref. Frequency
Is 50 Hz



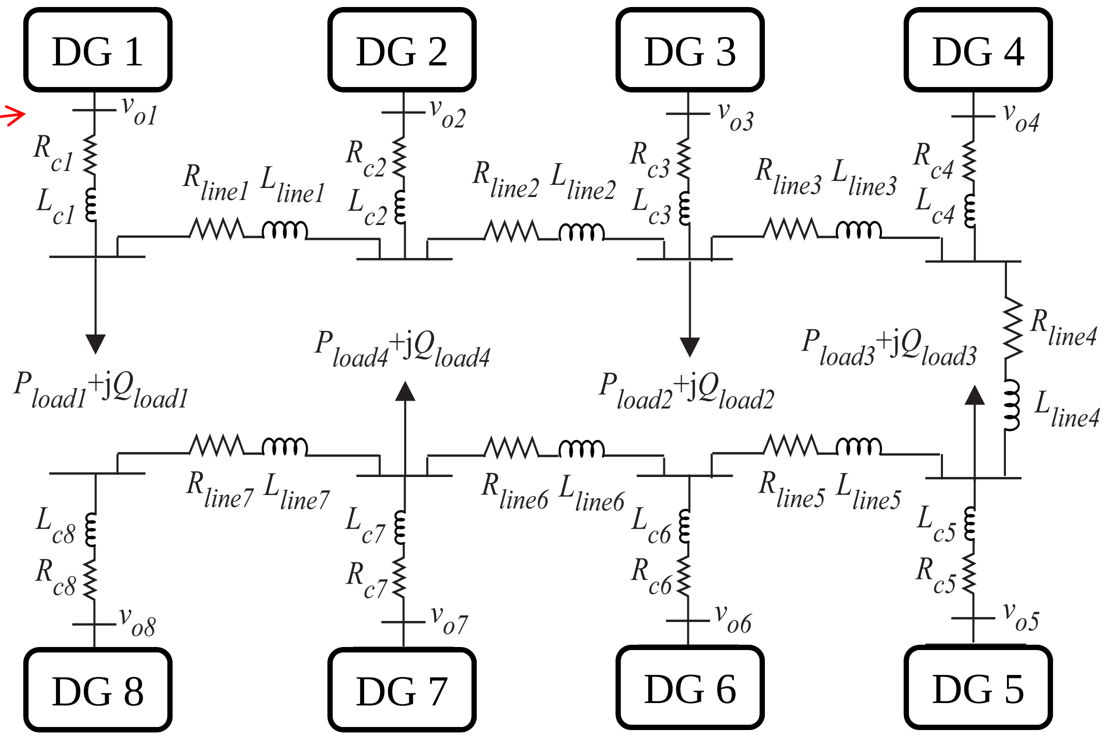
Islanding

Turn on
Coop secondary control

2. Secondary Voltage Control
Microgrid of Interconnected DG

Work of Ali Bidram
With Dr. A. Davoudi

Synchronize per-unit voltages



1. Frequency synchronization $\quad y_i = \omega_i = \omega_{ni} - m_{Pi} P_i$

2. Voltage synchronization (per unit) $\quad E \equiv v_{mag} = V_n - n_Q Q$

$$v_{o,magi} = \sqrt{v_{odi}^2 + v_{oqi}^2}$$

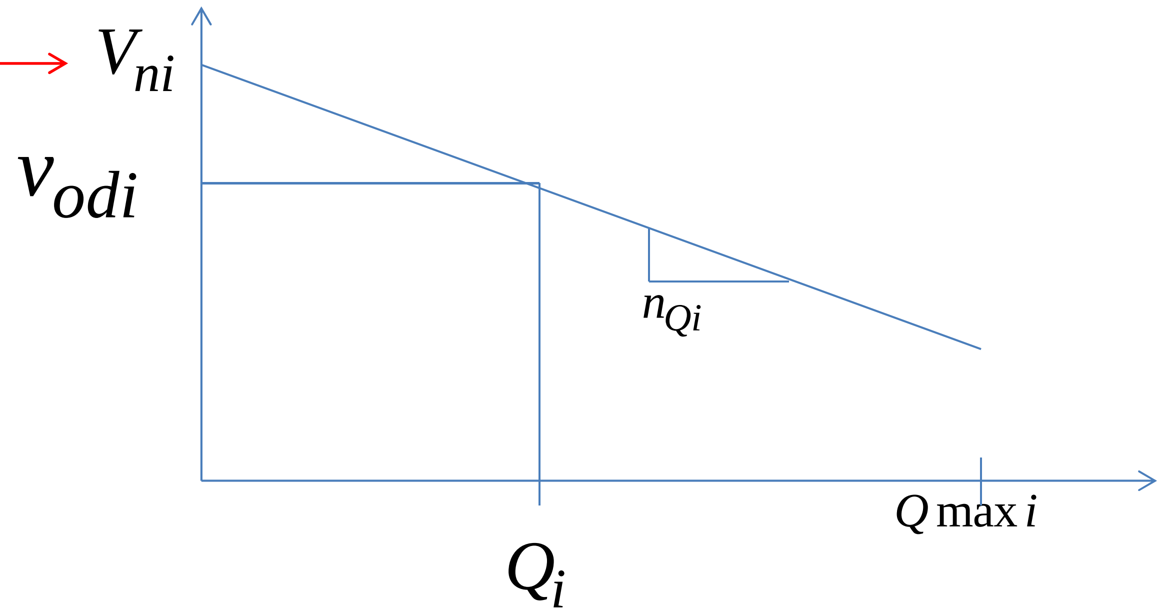
2. Secondary Voltage Control

New Secondary Control Input for Voltage Synchronization

$$E \equiv v_{odi} = V_{ni} - n_{Qi} Q_i$$

Secondary Control input

Prescribed voltage



Change V_{ni}

To change v_{odi}

To synchronize v_{oi}

Existing power conditions in the microgrid

Primary Droop Control

2. Secondary Voltage Control

If $y_i = v_{oi}$, there is no direct relationship between the output and

input $u_i = V_{ni}$.

$$d_i = 0$$

Must use Lie derivatives

Input-output feedback linearization for heterogeneous nonlinear agents

$$\begin{cases} \dot{\mathbf{x}}_i = \mathbf{f}_i(\mathbf{x}_i) + \mathbf{k}_i(\mathbf{x}_i)\mathbf{D}_i + \mathbf{g}_i(\mathbf{x}_i)u_i \\ y_i = h_i(\mathbf{x}_i) \end{cases}$$

$$\mathbf{F}_i(\mathbf{x}_i) \equiv \mathbf{f}_i(\mathbf{x}_i) + \mathbf{k}_i(\mathbf{x}_i)\mathbf{D}_i$$

$$y_i^{(r)} = L_{\mathbf{F}_i}{}^r h_i + L_{\mathbf{g}_i} L_{\mathbf{F}_i}{}^{r-1} h_i u_i$$

$$v_i \equiv L_{\mathbf{F}_i}{}^r h_i + L_{\mathbf{g}_i} L_{\mathbf{F}_i}{}^{r-1} h_i u_i \Longrightarrow u_i = (L_{\mathbf{g}_i} L_{\mathbf{F}_i}{}^{r-1} h_i)^{-1}(-L_{\mathbf{F}_i}{}^r h_i + v_i)$$

$$y_i^{(r)} = v_i, \forall i \Longrightarrow \begin{cases} \dot{y}_i \equiv y_{i,1} \\ \dot{y}_{i,1} \equiv y_{i,2} \\ \vdots \\ \dot{y}_{i,r-1} = v_i \end{cases}, \forall i \Longrightarrow \dot{\mathcal{Y}}_i = \mathbf{A}\mathcal{Y}_i + \mathbf{B}v_i, \forall i,$$

Internal Dynamics $\dot{\mu}_i = W_i(\mathcal{Y}_i, \mu_i), \forall i$

Assume relative degree *r* is the same for all agents

Zero dynamics can be different, but assume they are stable

2. Secondary voltage control

DG Agent Dynamics

$$\dot{\mathcal{Y}}_i = \mathbf{A}\mathcal{Y}_i + \mathbf{B}v_i, \forall i, \qquad \mathcal{Y}_i = [y_i \quad y_{i,1} \quad \dots \quad y_{i,r-1}]^T$$

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}_{r\times r} \qquad \mathbf{B} = [0,0,\dots,1]^T_{r\times 1}$$

Internal dynamics $\dot{\mu}_i = W_i(\mathcal{Y}_i, \mu_i), \forall i$

Leader node dynamics

$$\begin{cases} \dot{\mathbf{x}}_0 = \mathbf{f}_0(\mathbf{x}_0) \\ y_0 = h_0(\mathbf{x}_0) \end{cases} \Longrightarrow \dot{\mathbf{Y}}_0 = \mathbf{A}\mathbf{Y}_0 + \mathbf{B}y_0^{(r)}, \quad \mathbf{Y}_0 = [y_0 \quad \dot{y}_0 \quad \dots \quad y_0^{(r-1)}]^T$$

The synchronization problem is to find a distributed v_i such that $\mathcal{Y}_i \rightarrow \mathbf{Y}_0, \forall i.$

Assumption. The vector $\underline{\mathbf{y}}_0^{(r)} = \mathbf{1}_N\, y_0^{(r)}, \forall r$ is bounded so that $\left\| \underline{\mathbf{y}}_0^{(r)} \right\| \le Y_M^r$, with a finite but generally unknown bound.

2. Secondary voltage control

Theorem. Let the digraph of the multi-agent system have a spanning tree and the pinning gain be nonzero for at least one root node.

Let all agents have stable zero dynamics $\dot{\mu}_i = W_i(0, \mu_i), \forall i$

Let the auxiliary control be chosen as

$$v_i = -\underbrace{cK}\mathbf{e}_i \qquad \mathbf{e}_i = \sum_{j \in N_i} a_{ij}(\mathcal{Y}_i - \mathcal{Y}_j) + g_i(\mathcal{Y}_i - \mathbf{Y}_0)$$

where $c \in R$ is the coupling gain, and $K \in R^{1\times r}$ is the feedback control gain.

Then, $\mathcal{Y}_i$ are cooperative UUB with respect to $\mathbf{Y}_0$ and all nodes synchronize to $\mathbf{Y}_0$ if $K \in R^{1\times r}$ is chosen as

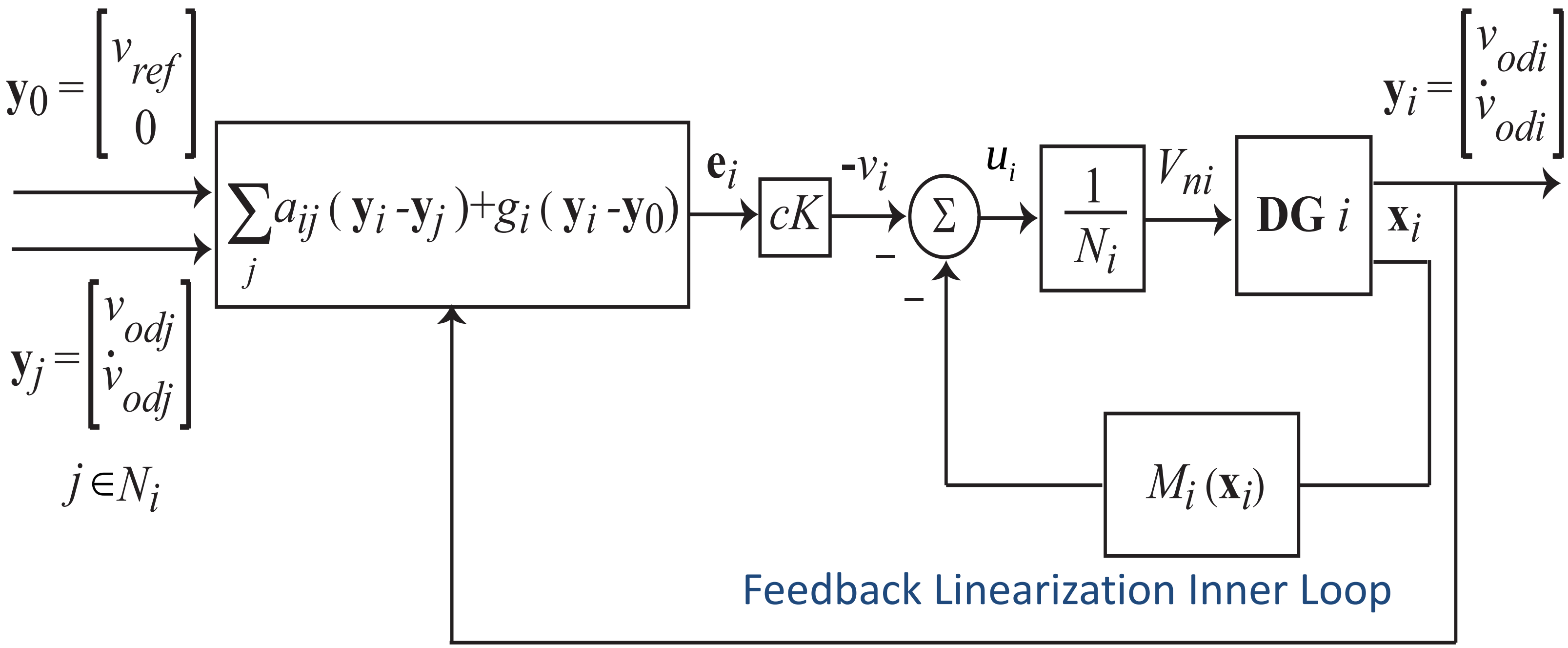
$$K = R^{-1}\mathbf{B}^T P_1, \qquad \mathbf{A}^T P_1 + P_1\mathbf{A} + Q - P_1\mathbf{B}R^{-1}\mathbf{B}^T P_1 = 0.$$

and $c \geq \dfrac{1}{2\lambda_{\min}}$, $\lambda_{\min} = \min_{i\in\mathcal{N}} Re(\lambda_i)$ λ_i is the first eigenvalue of $L+G$

Zhang, H., Lewis, F. L., & Das, A. (2011).Optimal design for synchronization of cooperative systems: State feedback, observer, and output feedback. *IEEE Transactions on Automatic Control*, *56*(8), 1948–1952.

Secondary voltage control

Synchronizes Output voltages after Islanding



$$y_i^{(2)} = L_{\mathbf{F}_i}{}^2 h_i + L_{\mathbf{g}_i} L_{\mathbf{F}_i}{}^1 h_i u_i$$

$$v_i = L_{\mathbf{F}_i}{}^2 h_i + L_{\mathbf{g}_i} L_{\mathbf{F}_i}{}^1 h_i u_i \quad \Rightarrow$$

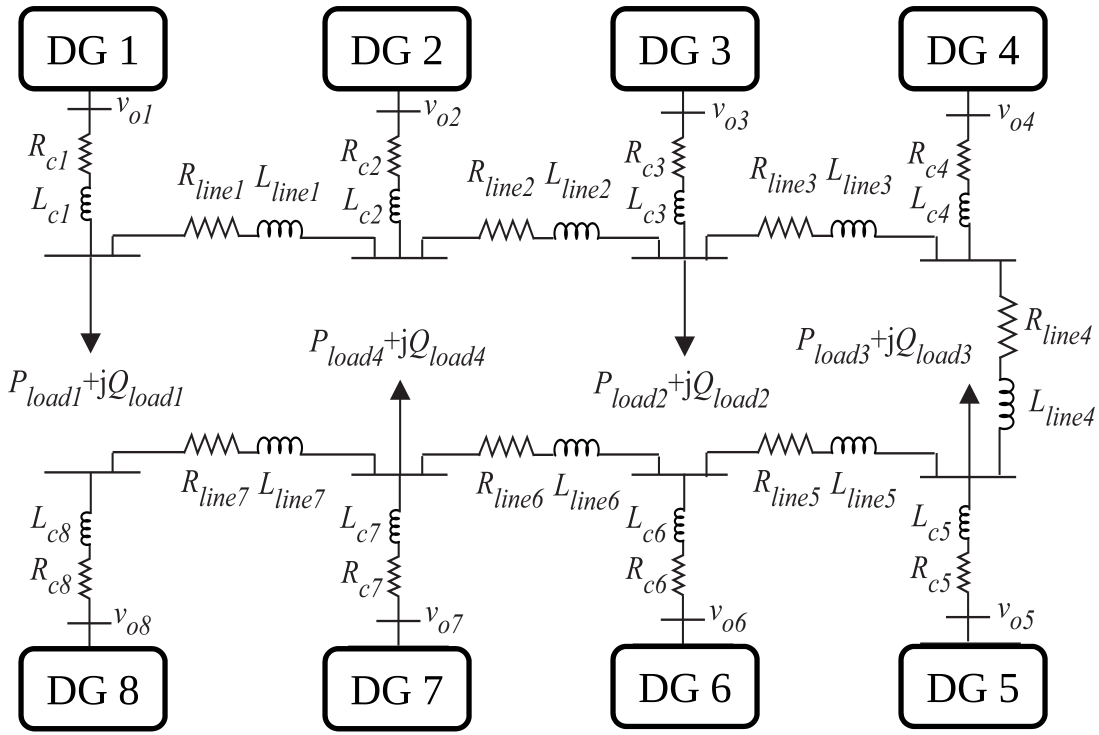
$$u_i = (L_{\mathbf{g}_i} L_{\mathbf{F}_i}{}^1 h_i)^{-1}(-L_{\mathbf{F}_i}{}^2 h_i + v_i)$$

$$V_{ni} = \frac{v_i - M_i(\mathbf{x}_i)}{N_i}, \forall i.$$

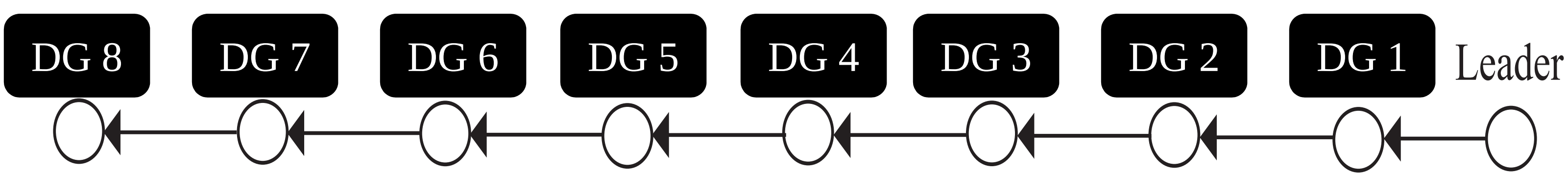
2. Secondary voltage control

Simulation Example

Physical Microgrid Network



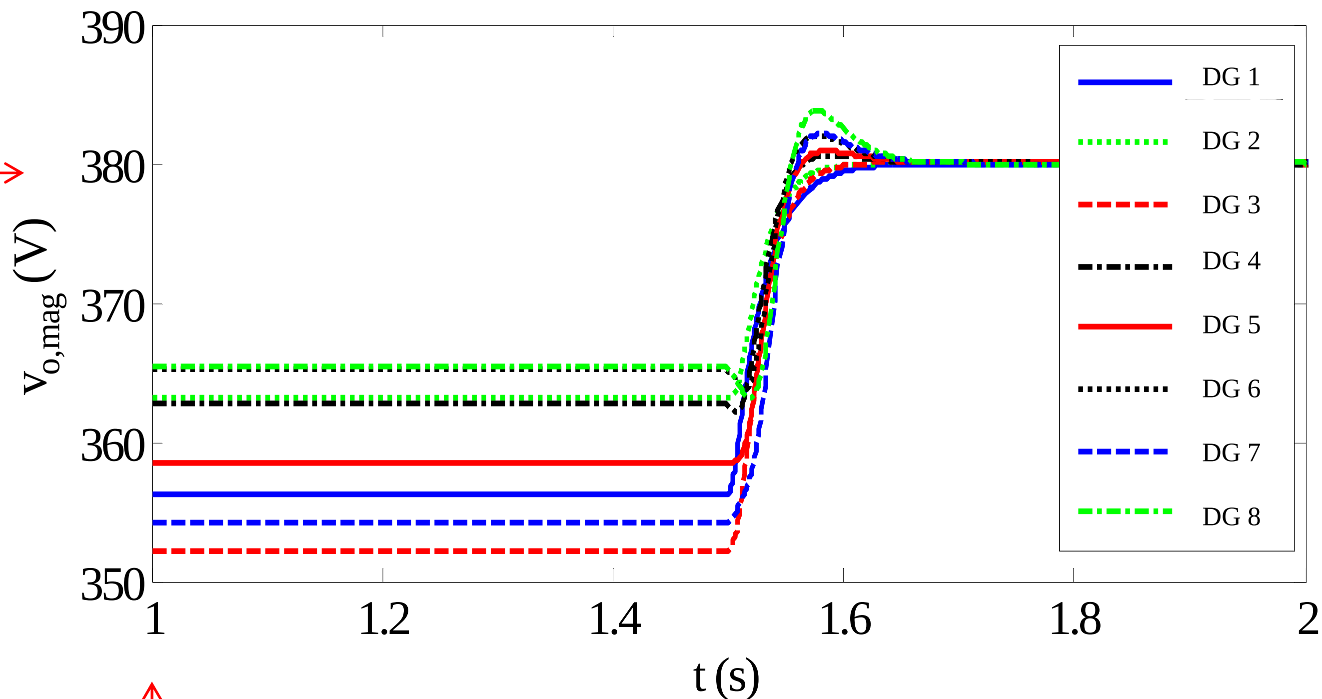
Cyber communication network- sparse



2. Secondary voltage control

Simulation results

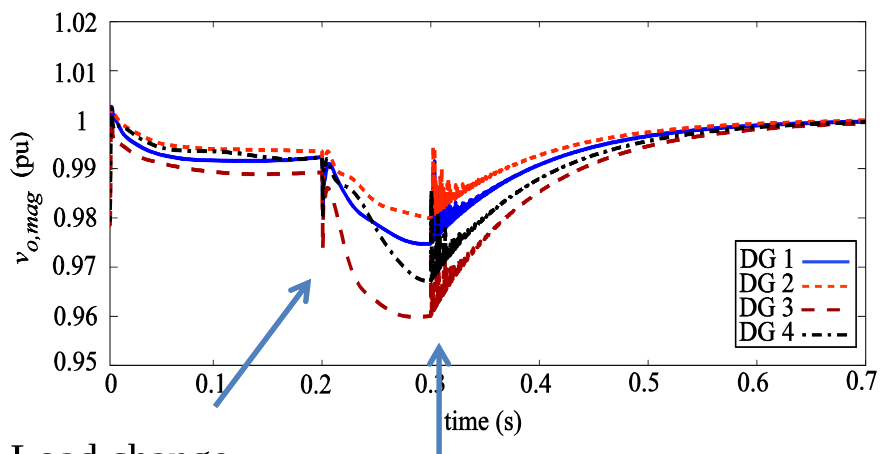
Ref. Per-unit
Voltage
Is 380 V



Islanding

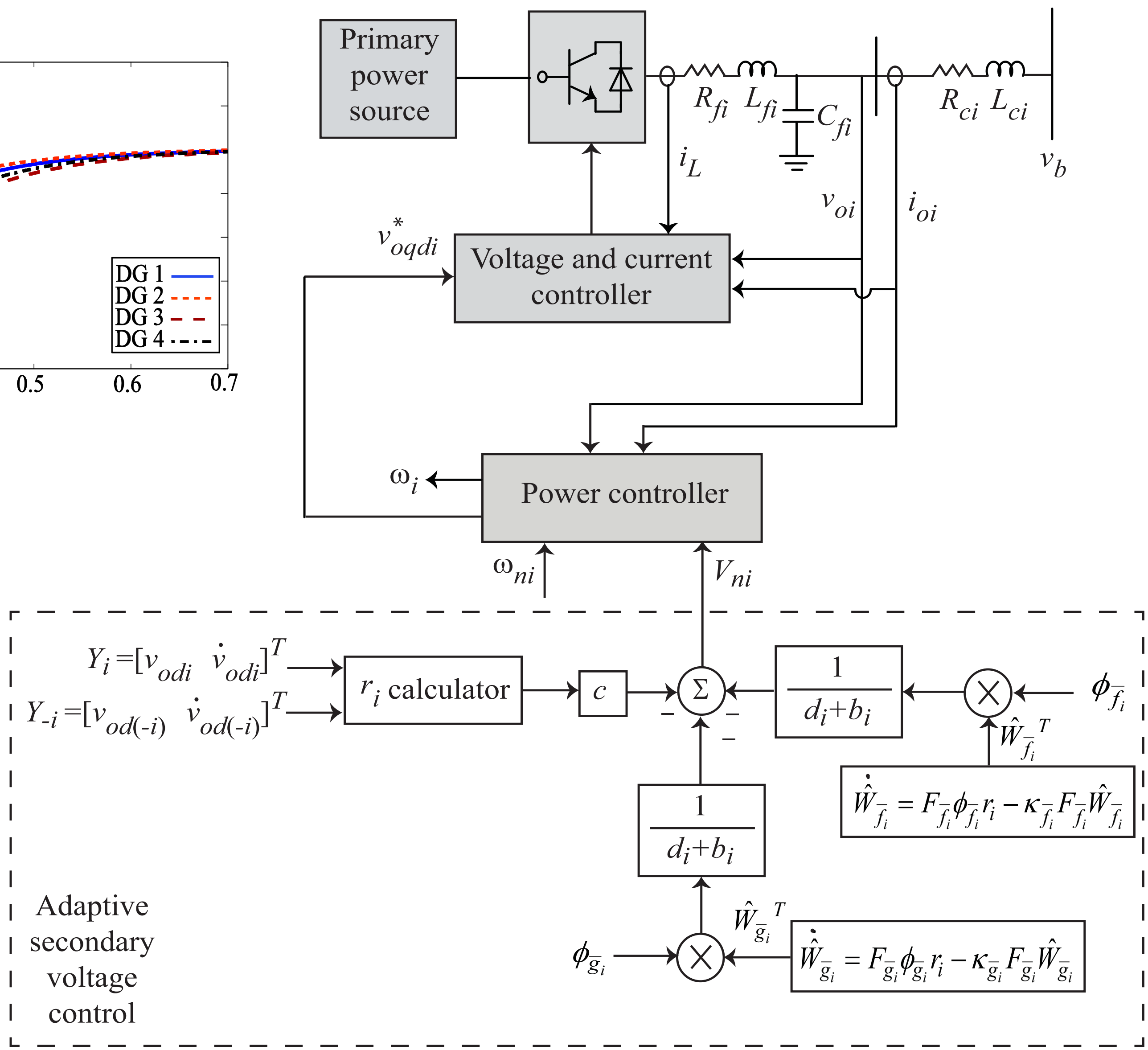
Turn on
Coop secondary control

Adaptive Voltage Control



Load change
Primary alone!

Secondary



$$u_i = (L_{\mathbf{g}_i} L_{\mathbf{F}_i}{}^1 h_i)^{-1}(-L_{\mathbf{F}_i}{}^2 h_i + v_i)$$

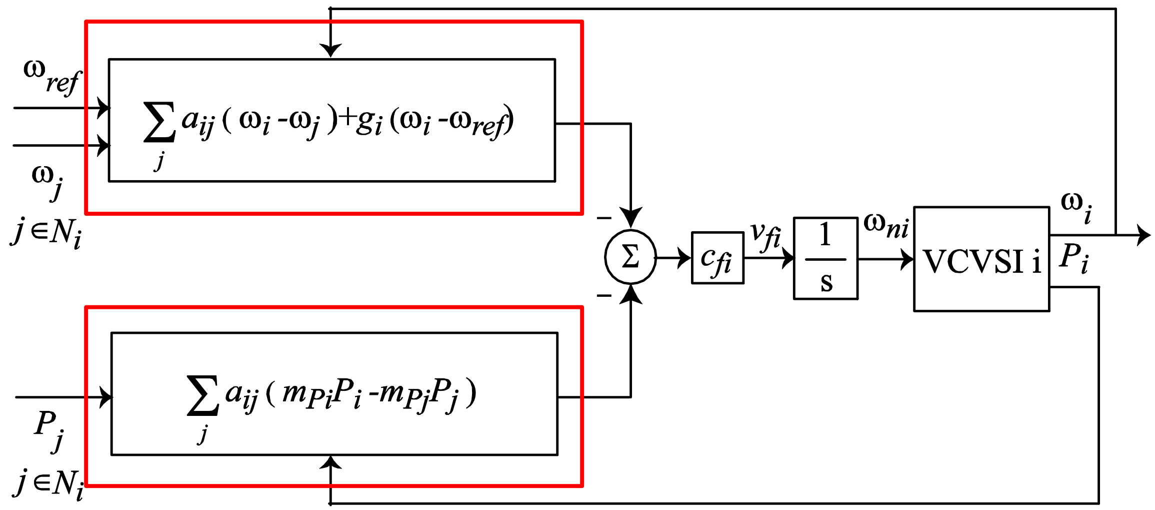
Using Neural Network to compensate for unknown nonlinear dynamics

$$M_i(x_i),\ N_i(x_i)$$

Multiobjective Distributed Secondary Control

Synchronizes frequency
Cooperative tracker

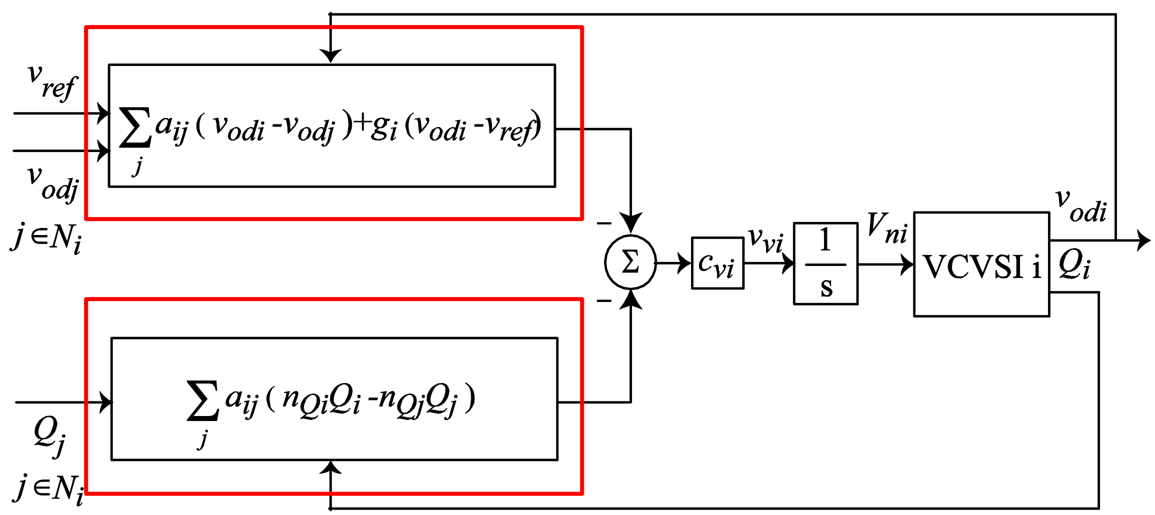
Active load sharing
Cooperative regulator



Frequency Control

Restore voltages
Cooperative tracker

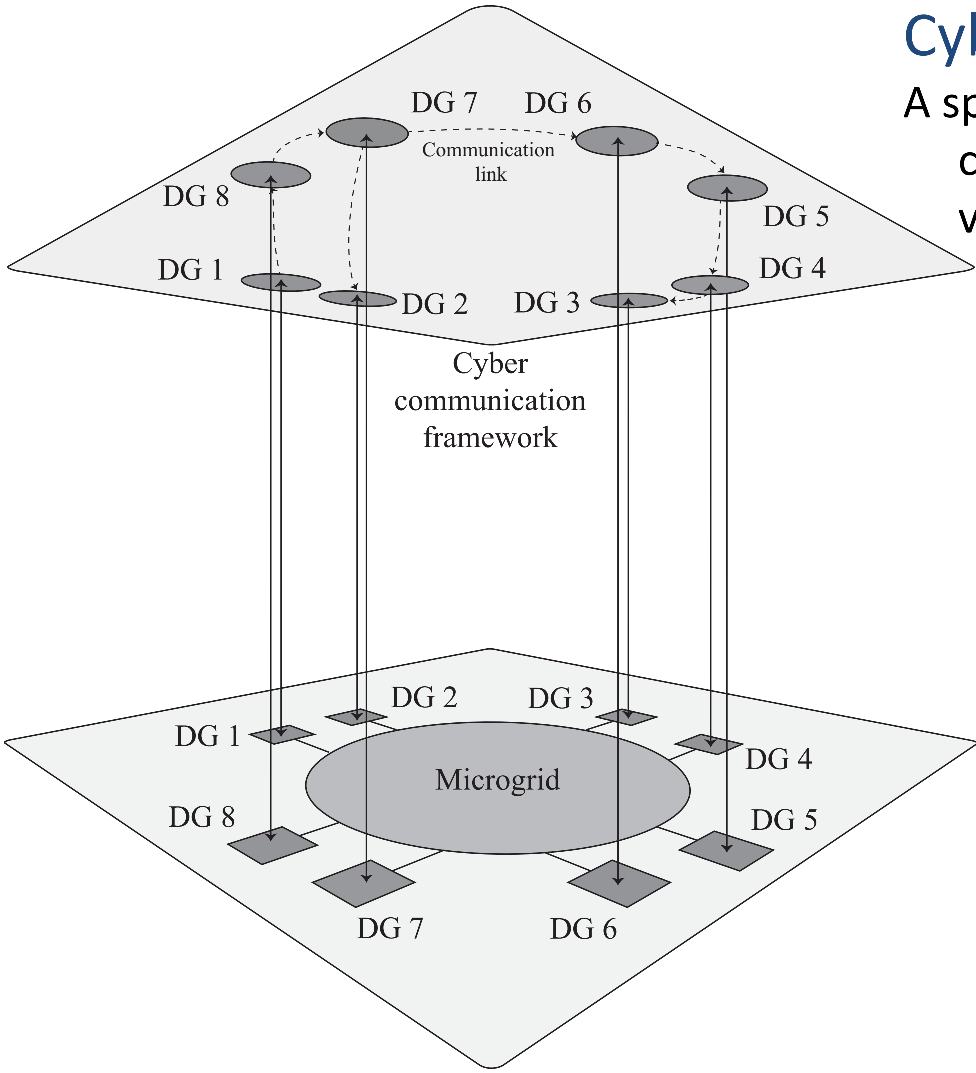
Reactive load sharing
Cooperative regulator



Voltage Control

Micro-grid secondary control: New Distributed CPS structure

Work of Ali Bidram
With Dr. A. Davoudi



Cyber layer

A sparse, efficient communication network to allow cooperative control for synchronization of voltage and frequency

Secondary Control

Cyber Physical System (CPS)

Physical Layer

The interconnect structure of the power grid

Primary Control

Game-theoretic Control for DC Microgrids

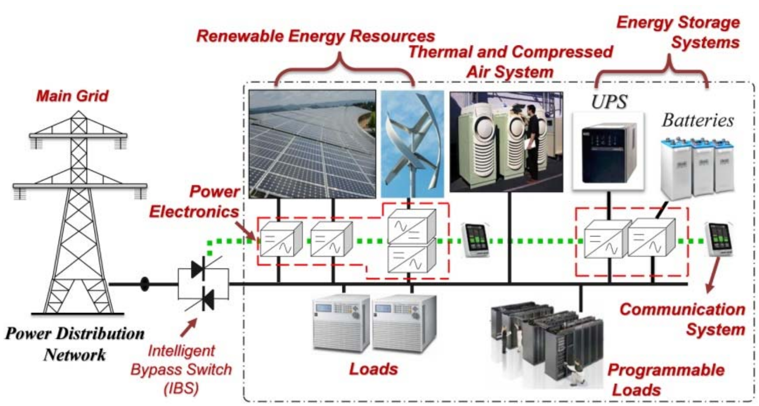
Work of Vahidreza Nasirian

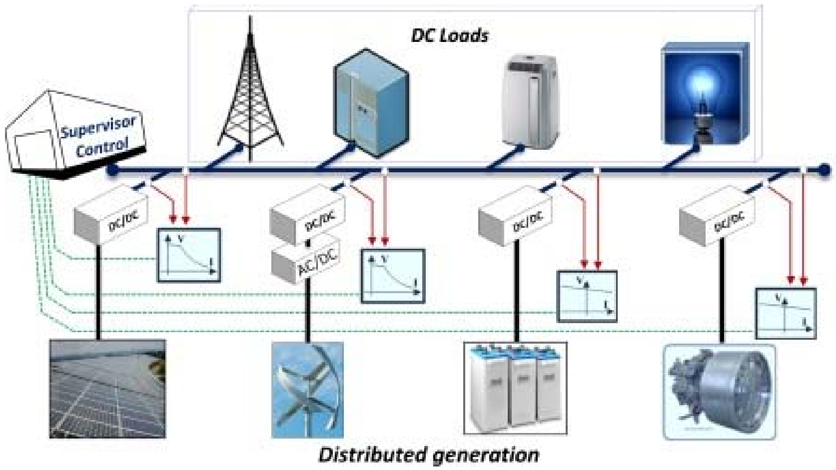
with Ali Davoudi

Advantages of DC Microgrids

AC Microgrid:

1) Complex synchronization procedure for grid-tied operation (frequency, magnitude, and phase match is required)
2) Complex control circuitry (voltage, frequency, and active/reactive power control)
3) Unwanted transmission loss due to reactive power exchange
4) Redundant dc-ac-dc conversions for integration of renewable sources, loads, and storage units
5) Harmonic current management and phase unbalances





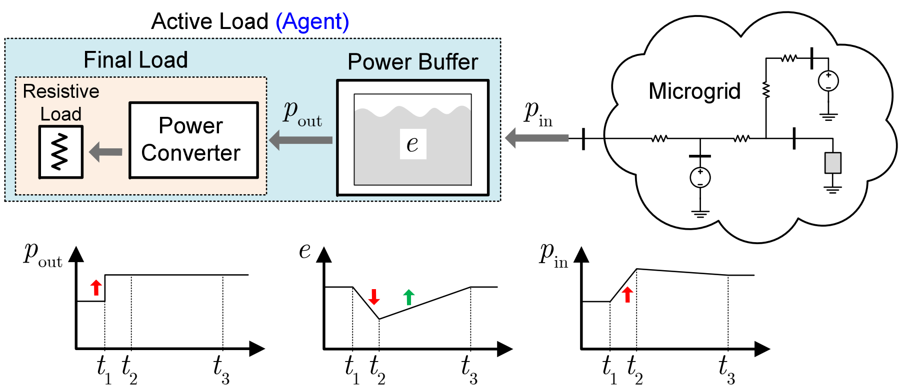
DC Microgrid:

1) Only voltage and power control is needed
2) No reactive power flow and, thus, an improved overall efficiency
3) Converted renewable energies are basically dc and, thus, a dc distribution is more effective for integration of these sources
4) No harmonic current or phase unbalance issue

Cooperative Game-theoretic Control of Active Loads in DC Microgrids

Ling-ling Fan, Vahidreza Nasirian,
Hamidreza Modares,
Frank L. Lewis, Yong-duan Song, and Ali Davoudi,

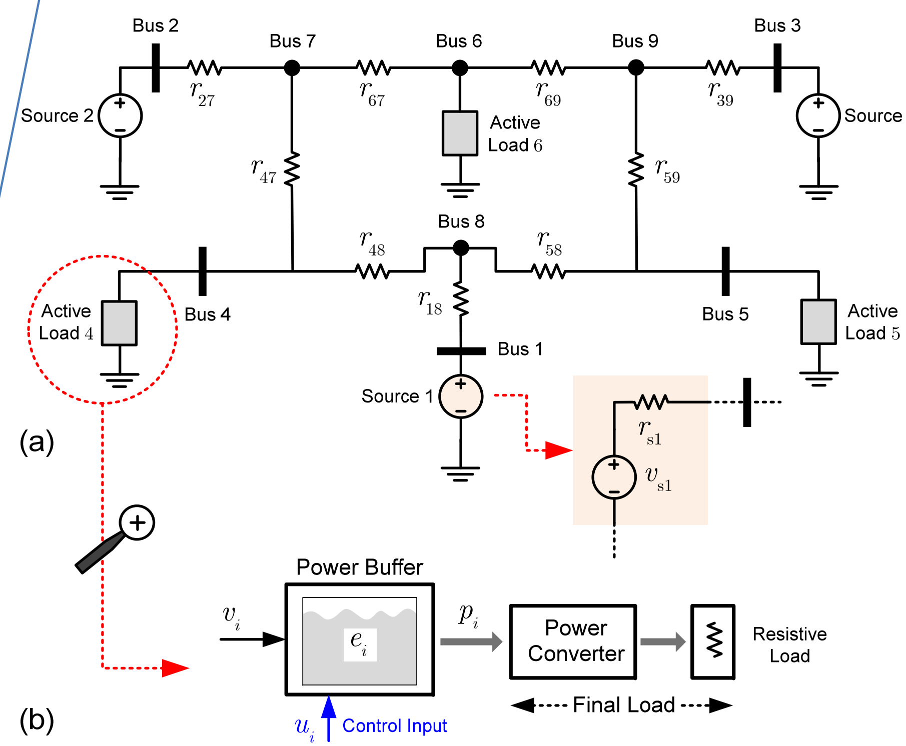
Background Work of Wayne Weaver



Power buffer operation during a step change in power demand.

Supplies excess power needed during load changes until sources can respond

Power buffers in Microgrid Network



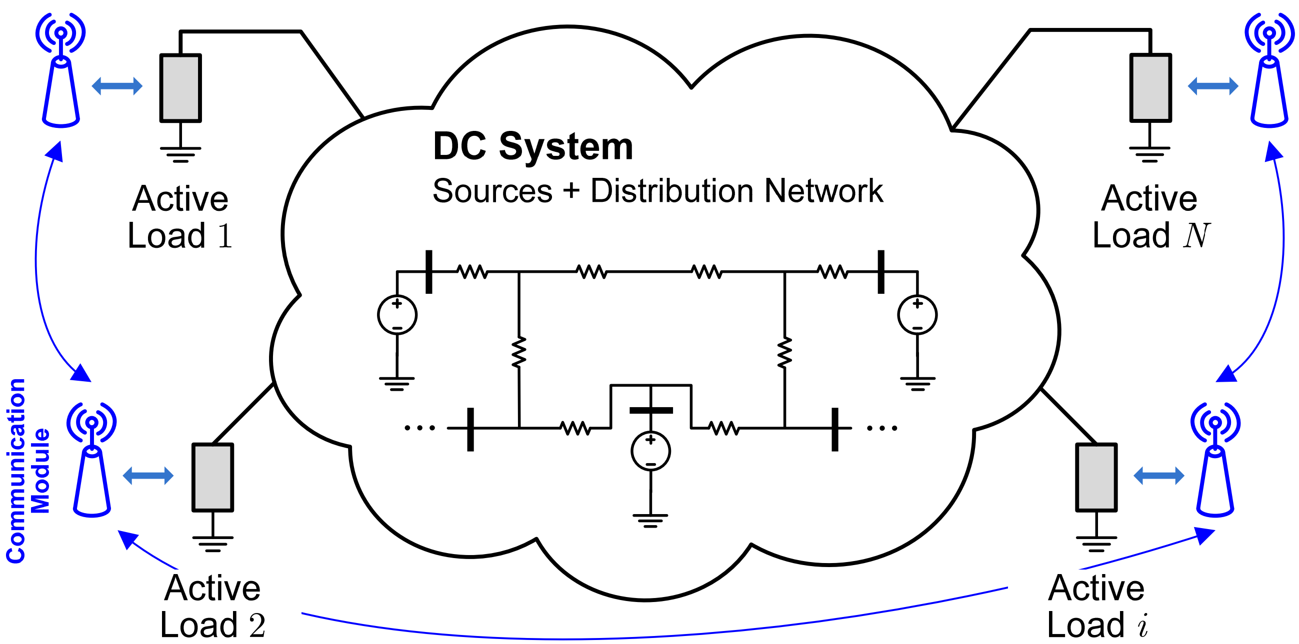
Active Load Power Buffer

Vahid Nasirian

$$\begin{cases} \dot{e}_i = \dfrac{v_i^2}{r_i} - p_i, \\ \dot{r}_i = u_i \end{cases}$$

Nonlinear dynamics
Not obvious how to handle p_i

Stored energy e_i
Input impedance r_i
Bus voltage v_i
Control input u_i
Output power = a disturbance p_i



Vahid Nasirian
Reza Modares

Dr. Ali Davoudi

Solve for bus voltage to get coupled agent dynamics

Linearize.
Add p_i as a state.
Formulate as H-infinity Problem.

$$\underbrace{\begin{bmatrix}\dot{e}_i\\ \dot{r}_i\\ \dot{p}_i\end{bmatrix}}_{\dot{\mathbf{x}}_i} = \underbrace{\begin{bmatrix}0 & 2i_i^{\mathrm{q}}\gamma_{ii} - \left(i_i^{\mathrm{q}}\right)^2 & -1\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}}_{\mathbf{A}_i}\underbrace{\begin{bmatrix}e_i\\ r_i\\ p_i\end{bmatrix}}_{\mathbf{x}_i} + \underbrace{\begin{bmatrix}0\\ 1\\ 0\end{bmatrix}}_{\mathbf{B}_i}u_i + \underbrace{\begin{bmatrix}0\\ 0\\ 1\end{bmatrix}}_{\mathbf{D}_i}w_i + 2i_i^{\mathrm{q}}\begin{bmatrix}\displaystyle\sum_{j=M+1(\neq i)}^{M+N}\gamma_{ij}r_j\\ 0\\ 0\end{bmatrix}, \qquad i = M+1,\cdots,M+N,$$

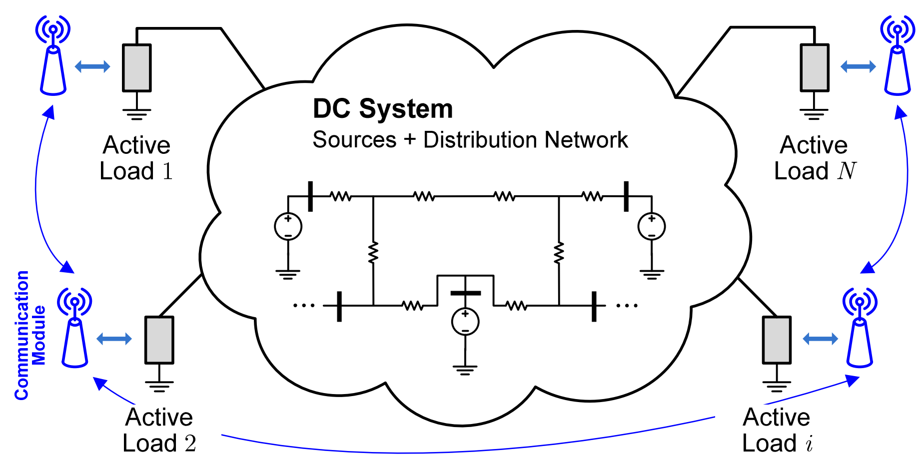
Coupling terms

Define coupled performance indices

$$J_i = \int_0^\infty\left(\sum_{j\in\bar{N}_i}\mathbf{x}_j^{\mathrm{T}}\mathbf{Q}_{ij}\mathbf{x}_j + \rho_i u_i^2\right)\mathrm{d}t, \qquad i = M+1,\cdots,M+N,$$

Define Communication Graph

Sparse efficient topology
Optimal design provides Resilience
and disturbance rejection



Optimal Cooperative Control as a Dynamic Graphical Game

$$\dot{\mathbf{x}}_i = \mathbf{A}_i\mathbf{x}_i + \mathbf{B}_i u_i + \mathbf{D}_i w_i + \begin{bmatrix} 2i_i^{q} & 0 & 0 \end{bmatrix}^{\mathrm{T}} \sum_{j \in N_i} \Gamma_{ij}\mathbf{x}_j$$

Graphical Game

Minimize the performance function for active loads

$$J_i = \int_0^{\infty} \left(\sum_{j \in \bar{N}_i} \mathbf{x}_j^{\mathrm{T}} \mathbf{Q}_{ij} \mathbf{x}_j + \rho_i u_i^2 \right) \mathrm{d}t$$

Let's define the neighborhood state vector as $\bar{\mathbf{x}}_i = \left(\mathbf{x}_i^{\mathrm{T}}, \left\{ \mathbf{x}_j^{\mathrm{T}} \right\}_{j \in N_i} \right)^{\mathrm{T}}$

The optimal solution is in a general form of $u_i = -\mathbf{k}_i \bar{\mathbf{x}}_i$

With such solutions, the performance function $\boldsymbol{J_i}$ is quadratic in $\mathbf{x}$: $J_i(\bar{\mathbf{x}}_i) = \bar{\mathbf{x}}_i^{\mathrm{T}} \mathbf{P}_i \bar{\mathbf{x}}_i$

which helps to find the optimal solution by solving an algebraic Riccati equation

$$u_i^* = -\left(\bar{\mathbf{B}}_{ii}^{\mathrm{T}} \mathbf{P}_i \bar{\mathbf{x}}_i \right) \rho_i^{-1}$$

82

Optimal Cooperative Control: Policy Iteration finds Optimal Solutions

- Substituting the optimal solution in Bellman equations leads to the following coupled Algebraic Riccati Equations (ARE)

$$H_i = \bar{\mathbf{x}}_i^{\mathrm{T}} \mathbf{Q}_i \mathbf{x}_i^{\mathrm{T}} + \rho_i \left(u_i^*\right)^2 + \\ + \bar{\mathbf{x}}_i^{\mathrm{T}} \mathbf{P}_i \left(\bar{\mathbf{A}}_i \bar{\mathbf{x}}_i + \bar{\mathbf{B}}_i \bar{\mathbf{u}}_i^* + \bar{\mathbf{D}}_i \bar{\mathbf{w}}_i + \Psi(\bar{\mathbf{x}}_i) \right) \\ + \left(\bar{\mathbf{A}}_i \bar{\mathbf{x}}_i + \bar{\mathbf{B}}_i \bar{\mathbf{u}}_i^* + \bar{\mathbf{D}}_i \bar{\mathbf{w}}_i + \Psi(\bar{\mathbf{x}}_i) \right)^{\mathrm{T}} \mathbf{P}_i \bar{\mathbf{x}}_i = \mathbf{0}$$

$$\bar{\mathbf{u}}_i^* = \left(u_i^*, \left\{ u_j^* \right\}_{j \in N_i} \right)^{\mathrm{T}}$$

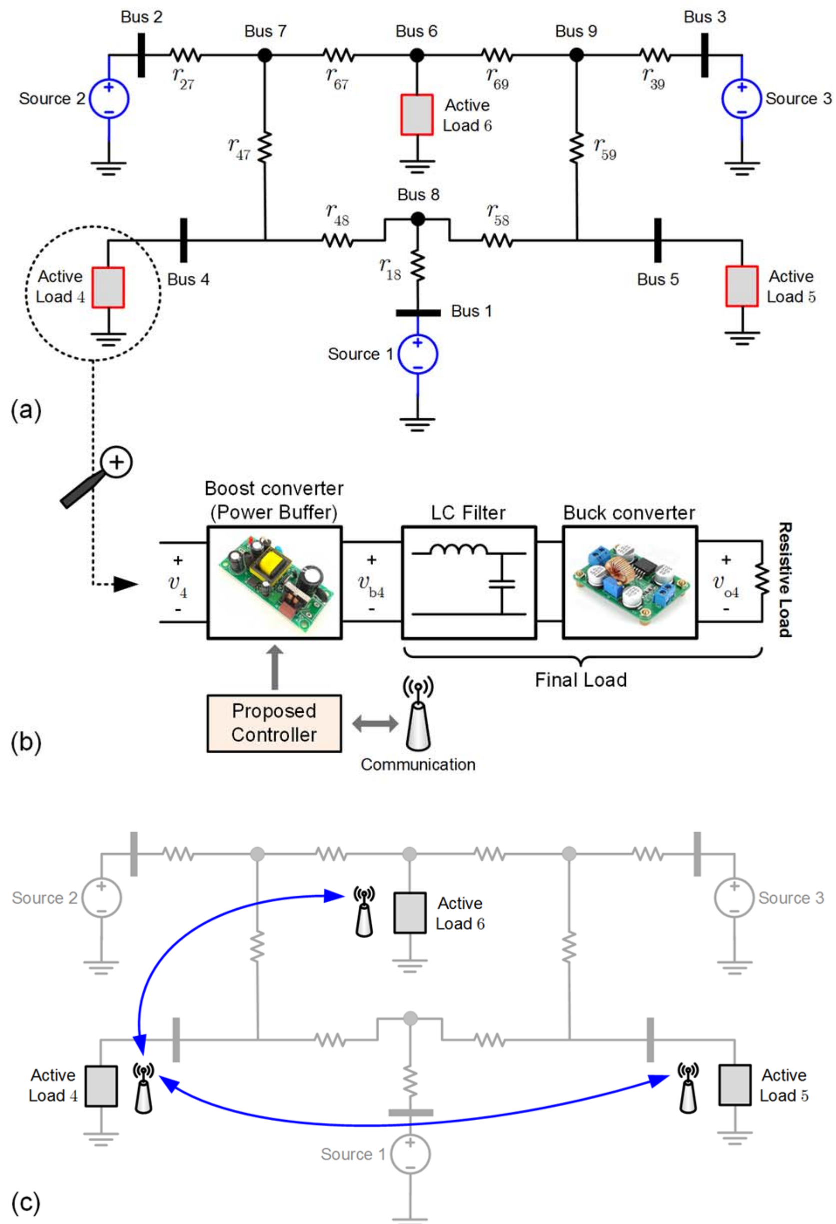
- Policy iteration (a class of reinforcement learning) is used to solve ARE and find $\mathbf{P}_i$ and the optimal control input

$$u_i^* = -\left(\bar{\mathbf{B}}_{ii}^{\mathrm{T}} \mathbf{P}_i \bar{\mathbf{x}}_i \right) \rho_i^{-1}$$

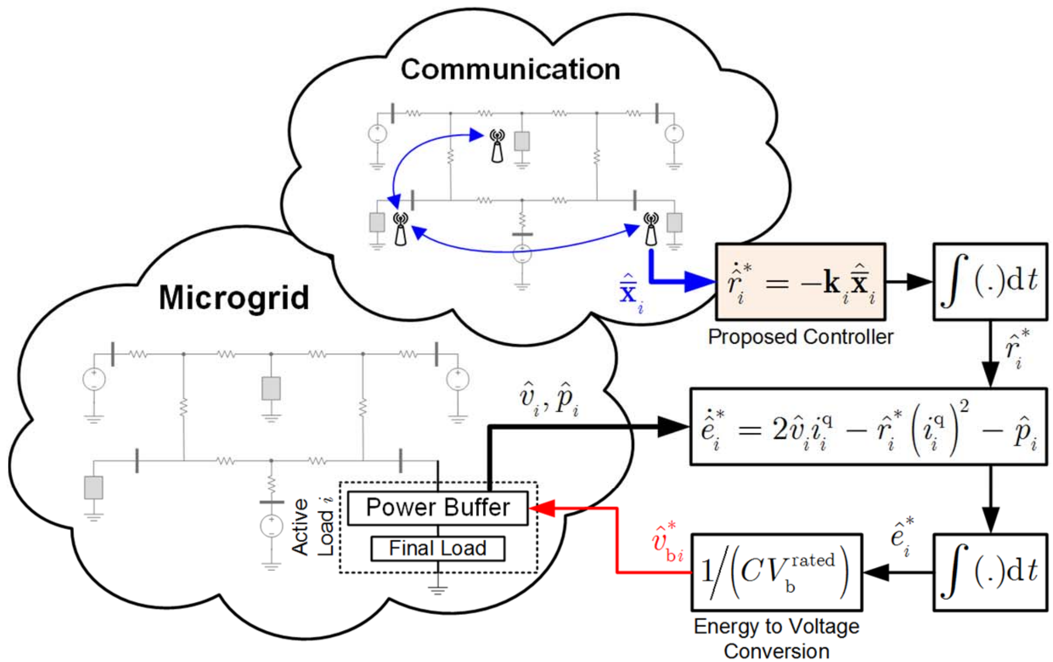
- Policy evaluation: the performance of a given control policy, $\boldsymbol{u}_i$, is evaluated using the Bellman equation, and $\mathbf{P}_{\mathrm{i}}$ are found.
- Policy improvement: an improved control policy, $\boldsymbol{u}_i$, is found for each agent, using $\mathbf{P}_{\mathrm{i}}$ found in the first step.
- Policy evaluation and improvement are repeated until no improvement in control policies, u_i, of any agent is observed.

83

Microgrid Setup and Cooperative Controller

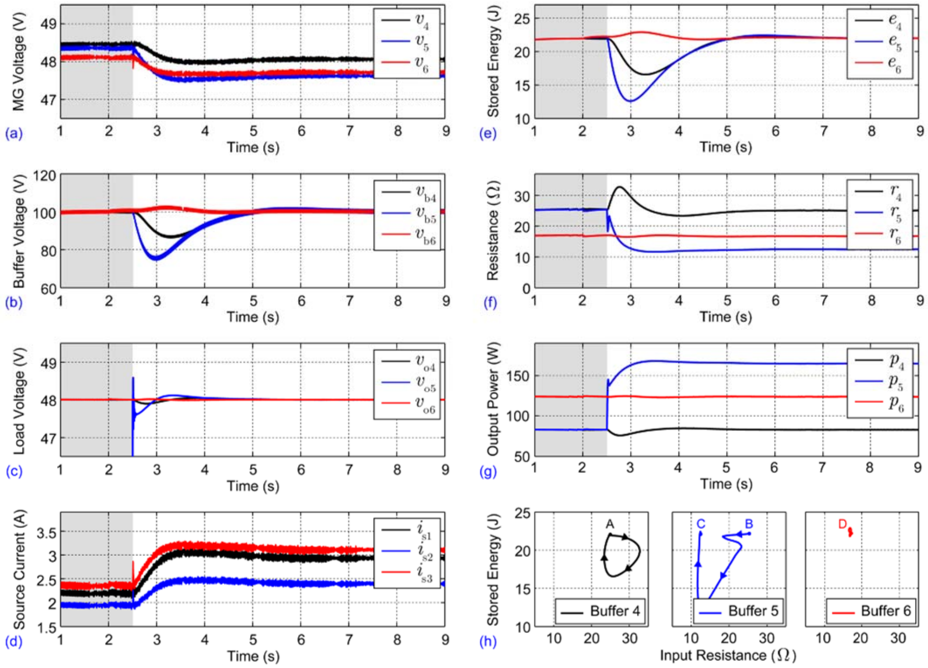


(a) DC microgrid system
(b) Active load
(c) Communication network

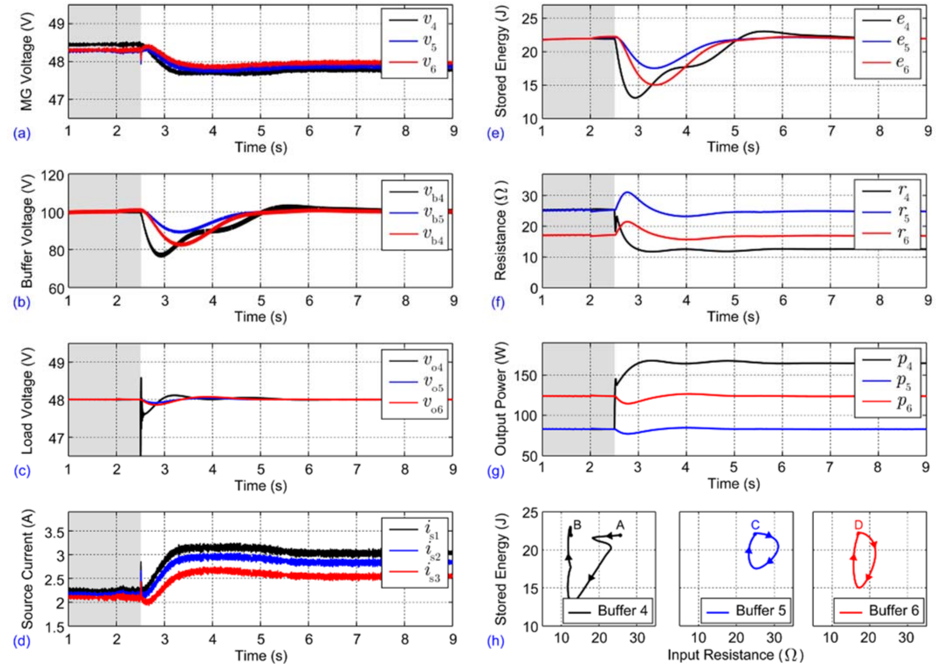


Controller Implementation

Controller Performance with Load Change



Load change in bus 5; Buffers 4 & 5 assisting



Load change in bus 4; Multiple assistive buffers

(a) microgrid bus voltages at the load terminals, (b) Output voltage of the power buffers, (c) output voltage across the resistive loads, (d) Source currents, (e) Stored energies in power buffers, (f) Input impedance of the power buffers, (g) Output of the active loads, (h) energy-impedance trajectory of power buffers during the load transient.

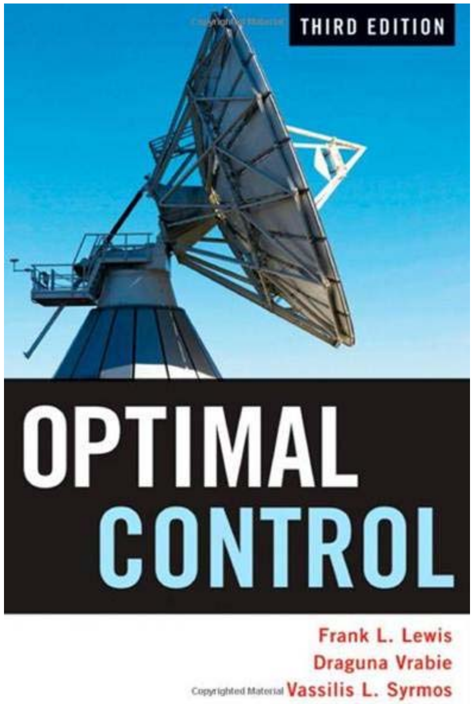
Real-Time Optimal Cooperative Control: Reinforcement Learning

Optimality and Games

Optimal Control is Effective for:
- Aircraft Autopilots
- Vehicle engine control
- Aerospace Vehicles
- Ship Control
- Industrial Process Control

Multi-player Games Occur in:
- Networked Systems Bandwidth Assignment
- Economics
- Control Theory disturbance rejection
- Team games
- International politics
- Sports strategy

But, optimal control and game solutions are found by
- Offline solution of Matrix Design equations
- A full dynamical model of the system is needed



The Importance of Optimal Control

Formulate an Optimal Control Problem

Nonlinear System dynamics $\dot{x} = f(x,u) = f(x) + g(x)u$

Cost/value $V(x(t)) = \int_t^\infty r(x,u)\, dt = \int_t^\infty (Q(x) + u^T Ru)\, dt$

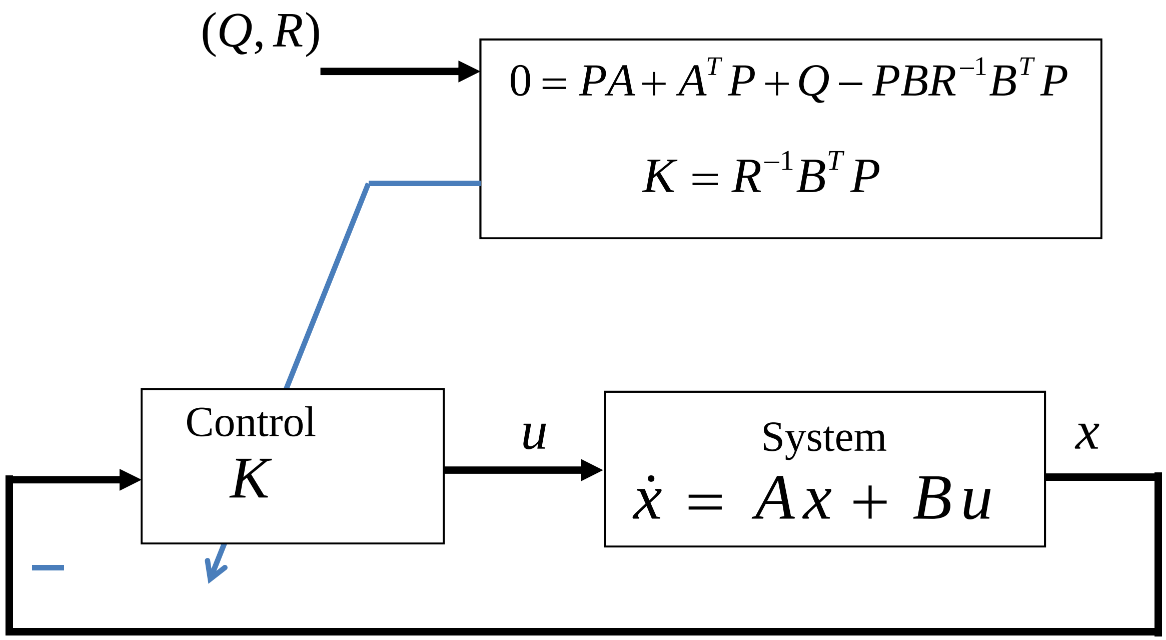
Then you can always learn the optimal solution online using data measured in real time

By using Integral Reinforcement Learning

DDO- Data-Driven Optimization

Optimal Control- The Linear Quadratic Regulator (LQR)

User prescribed optimization criterion $$V(x(t)) = \int_t^{\infty} (x^T Qx + u^T Ru)\, d\tau$$



Off-line Design Loop
Using ARE

On-line real-time
Control Loop

An Offline Design Procedure
that requires Knowledge of system dynamics model (A,B)

System modeling is expensive, time consuming, and inaccurate

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.

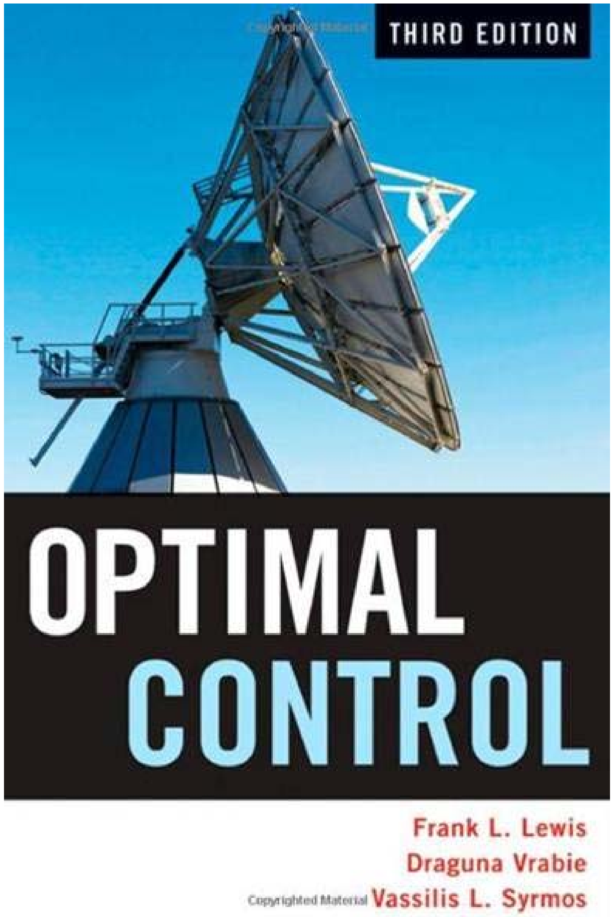
We want to find optimal control solutions
- Online in real-time
- Using adaptive control techniques
- Without knowing the full dynamics

For nonlinear systems and general performance indices

Bring together Optimal Control and Adaptive Control

Reinforcement Learning turns out to be the key to this!

Books



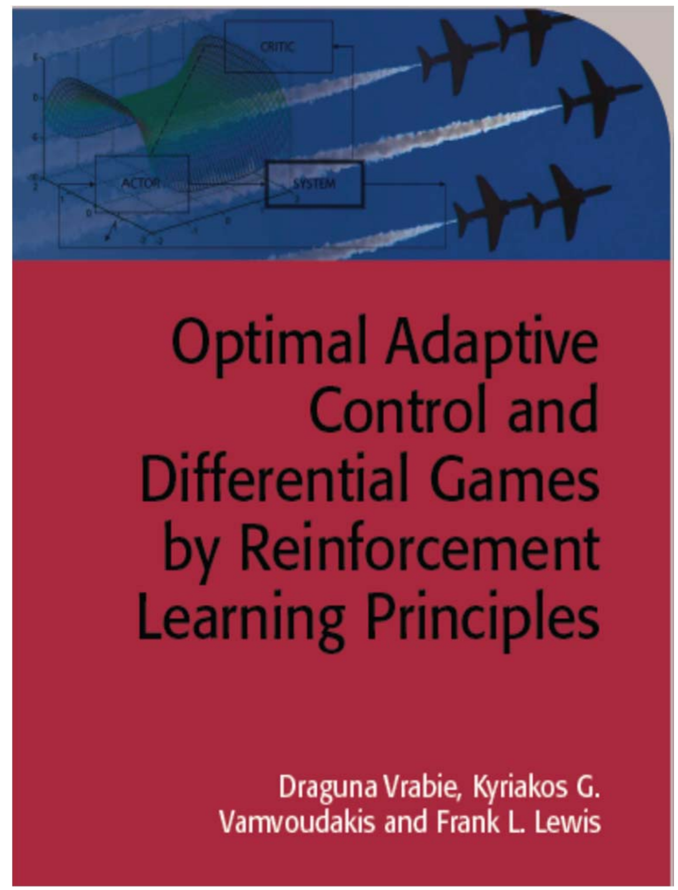
F.L. Lewis, D. Vrabie, and V. Syrmos, *Optimal Control*, third edition, John Wiley and Sons, New York, 2012.

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.



CT Systems- Derivation of Nonlinear Optimal Regulator

To find online methods for optimal control

Focus on these two equations

Nonlinear System dynamics $\dot{x} = f(x,u) = f(x) + g(x)u$

Cost/value $V(x(t)) = \int_t^\infty r(x,u)\, dt = \int_t^\infty (Q(x) + u^T Ru)\, dt$

Leibniz gives
Differential equivalent

Bellman Equation, in terms of the Hamiltonian function

$$\boxed{H(x,\frac{\partial V}{\partial x},u) = \dot{V} + r(x,u) = \left(\frac{\partial V}{\partial x}\right)^T \dot{x} + r(x,u) = \left(\frac{\partial V}{\partial x}\right)^T \big(f(x) + g(x)u\big) + r(x,u) = 0}$$

Stationarity condition $\frac{\partial H}{\partial u} = 0$

Problem- System dynamics
shows up in Hamiltonian

Stationary Control Policy $\boxed{u = h(x) = -\tfrac{1}{2} R^{-1} g^T(x) \frac{\partial V}{\partial x}}$

HJB equation $0 = \left(\frac{dV^*}{dx}\right)^T f + Q(x) - \frac{1}{4}\left(\frac{dV^*}{dx}\right)^T g R^{-1} g^T \frac{dV^*}{dx} \quad , \quad V(0) = 0$

Off-line solution
HJB hard to solve. May not have smooth solution.
Dynamics must be known

CT Policy Iteration – a Reinforcement Learning Technique

Given any admissible *policy* $u(x) = h(x)$

The cost is given by solving the CT Bellman equation

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

Scalar equation

Utility $r(x,u) = Q(x) + u^T R u$

Policy Iteration Solution

Pick stabilizing initial control policy $h_0(x)$

Policy Evaluation - Find cost, Bellman eq.

$$0 = \left(\frac{\partial V_j}{\partial x}\right)^T f(x, h_j(x)) + r(x, h_j(x))$$

$$V_j(0) = 0$$

Policy improvement - Update control

$$h_{j+1}(x) = -\tfrac{1}{2} R^{-1} g^T(x) \frac{\partial V_j}{\partial x}$$

- Convergence proved by Leake and Liu 1967, Saridis 1979 if Lyapunov eq. solved exactly
- Beard & Saridis used Galerkin Integrals to solve Lyapunov eq.
- Abu Khalaf & Lewis used NN to approx. V for nonlinear systems and proved convergence

Full system dynamics must be known

Off-line solution

Converges to solution of HJB

$$0 = \left(\frac{dV^*}{dx}\right)^T f + Q(x) - \tfrac{1}{4}\left(\frac{dV^*}{dx}\right)^T g R^{-1} g^T \frac{dV^*}{dx}$$

M. Abu-Khalaf, F.L. Lewis, and J. Huang, "Policy iterations on the Hamilton-Jacobi-Isaacs equation for H-infinity state feedback control with input saturation," IEEE Trans. Automatic Control, vol. 51, no. 12, pp. 1989-1995, Dec. 2006.

LQR Policy iteration = Kleinman algorithm

1. For a given control policy $u = -K_j x$ solve for the cost:

$$0 = A_j^T P_j + P_j A_j + Q + K_j^T R K_j$$

Bellman eq. = Lyapunov eq.

Matrix equation

$$A_j = A - BK_j$$

2. Improve policy:

$$K_{j+1} = R^{-1} B^T P_j$$

- If started with a stabilizing control policy K_0 the matrix P_j monotonically converges to the unique positive definite solution of the Riccati equation.
- Every iteration step will return a stabilizing controller.
- The system has to be known.

OFF-LINE DESIGN
MUST SOLVE LYAPUNOV EQUATION AT EACH STEP.

Kleinman 1968

Integral Reinforcement Learning

Work of Draguna Vrabie 2009

value $$V(x(t)) = \int_t^{\infty} r(x,u)\, d\tau = \int_t^{t+T} r(x,u)\, d\tau + \int_{t+T}^{\infty} r(x,u)\, d\tau$$

Key Idea- US Patent

Lemma 1 – Draguna Vrabie

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

Bad Bellman Equation

Is equivalent to Integral reinf. form (IRL) for the CT Bellman eq.

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

Good Bellman Equation

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

Allows definition of temporal difference error for CT systems

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

Integral Reinforcement Learning (IRL)- Draguna Vrabie

IRL Policy iteration

Policy evaluation- IRL Bellman Equation

Cost update $$V_k(x(t)) = \int_t^{t+T} r(x,u_k)\,dt \quad + \quad V_k(x(t+T))$$

CT Bellman eq.

f(x) and *g(x)* do not appear

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

Solves Bellman eq. (nonlinear Lyapunov eq.) without knowing system dynamics

Policy improvement

Control gain update $$u_{k+1} = h_{k+1}(x) = -\tfrac{1}{2} R^{-1} g^T(x) \frac{\partial V_k}{\partial x}$$

g(x) needed for control update

Initial stabilizing control is needed

Converges to solution to HJB eq. $$0 = \left(\frac{dV^*}{dx}\right)^T f + Q(x) - \tfrac{1}{4}\left(\frac{dV^*}{dx}\right)^T g R^{-1} g^T \frac{dV^*}{dx}$$

D. Vrabie proved convergence to the optimal value and control
Automatica 2009, Neural Networks 2009

Nonlinear Case- Approximate Dynamic Programming

Value Function Approximation (VFA) to Solve Bellman Equation
– Paul Werbos (ADP), Dimitri Bertsekas (NDP)

Optimal Control and Adaptive Control come together On this slide. Because of RL

$$V_k(x(t)) = \int_t^{t+T} \left(Q(x) + u_k{}^T R u_k\right) dt + V_k(x(t+T))$$

Approximate value by Weierstrass Approximator Network $V = W^T \phi(x)$

$$W_k^T \phi(x(t)) = \int_t^{t+T} \left(Q(x) + u_k{}^T R u_k\right) dt + W_k^T \phi(x(t+T))$$

$$W_k^T \underbrace{\left[\phi(x(t)) - \phi(x(t+T))\right]}_{\text{regression vector}} = \int_t^{t+T} \underbrace{\left(Q(x) + u_k{}^T R u_k\right) dt}_{\text{Reinforcement on time interval } [t,\, t+T]}$$

Scalar algebraic equation with vector unknowns

Same form as standard System ID problems in Adaptive Control

Now use RLS along the trajectory to get new weights W_k

Then find updated FB

$$u_{k+1} = h_{k+1}(x) = -\tfrac{1}{2} R^{-1} g^T(x) \frac{\partial V_k}{\partial x} = -\tfrac{1}{2} R^{-1} g^T(x) \left[\frac{\partial \phi(x(t))}{\partial x(t)}\right]^T W_k$$

Direct Optimal Adaptive Control for Partially Unknown CT Systems

Synchronous Online Solution of Optimal Control for Nonlinear Systems

K.G. Vamvoudakis and F.L. Lewis, “Online actor-critic algorithm to solve the continuous-time infinite horizon optimal control problem,” Automatica, vol. 46, no. 5, pp. 878-888, May 2010.

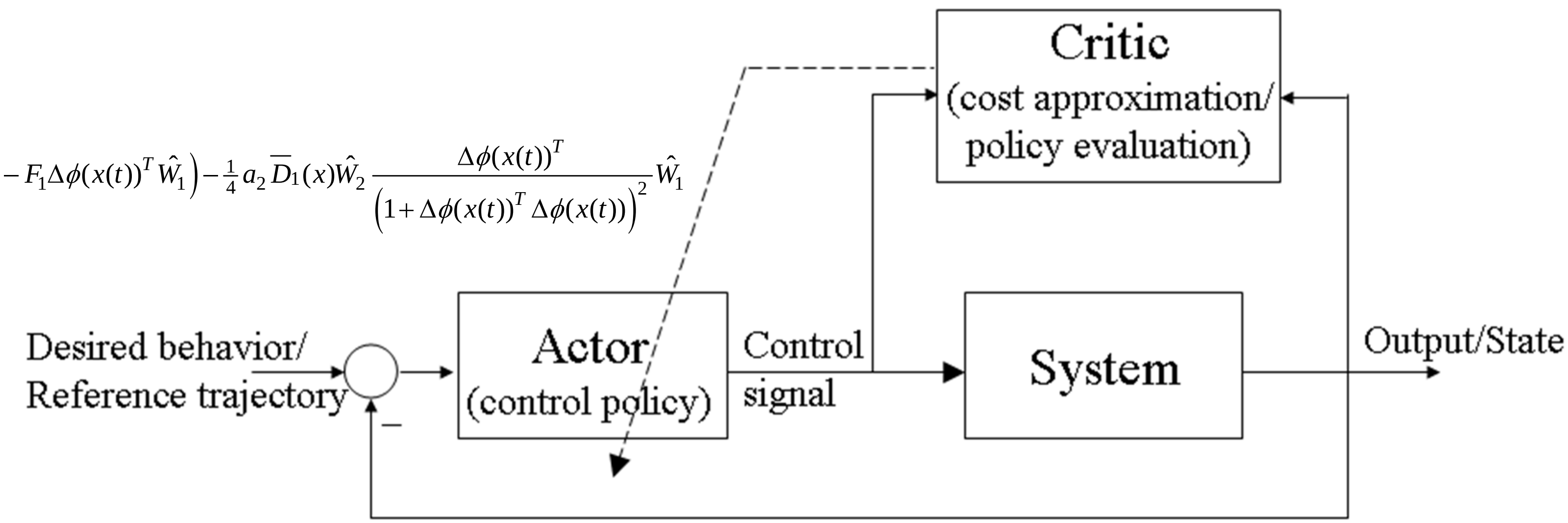
A new form of Adaptive Control with TWO tunable networks

Adaptive Critic structure

Reinforcement learning

$$\dot{\hat{W}}_1 = -a_1 \frac{\Delta\phi(x(t))}{\left(1+\Delta\phi(x(t))^T\Delta\phi(x(t))\right)^2}\left(\Delta\phi(x(t))^T\hat{W}_1 + \int_{t-T}^{t}\left(Q(x)+\frac{1}{4}\hat{W}_2^T\bar{D}_1\hat{W}_2\right)d\tau\right)$$

$$\dot{\hat{W}}_2 = -a_2\left(F_2\hat{W}_2 - F_1\Delta\phi(x(t))^T\hat{W}_1\right) - \tfrac{1}{4}a_2\bar{D}_1(x)\hat{W}_2\frac{\Delta\phi(x(t))^T}{\left(1+\Delta\phi(x(t))^T\Delta\phi(x(t))\right)^2}\hat{W}_1$$

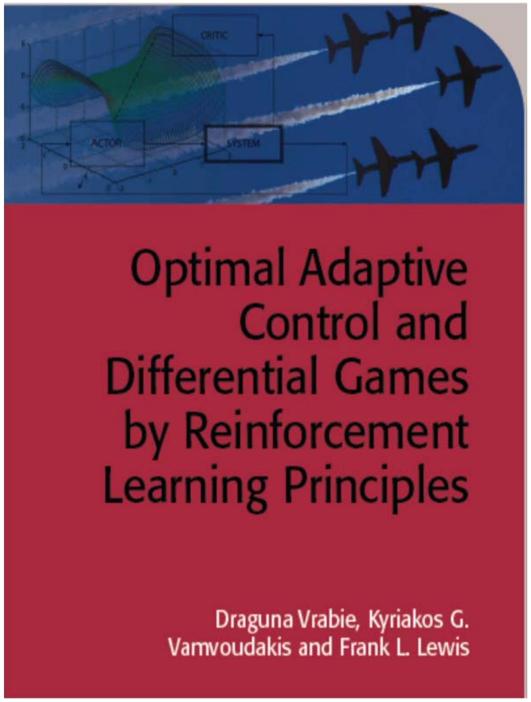


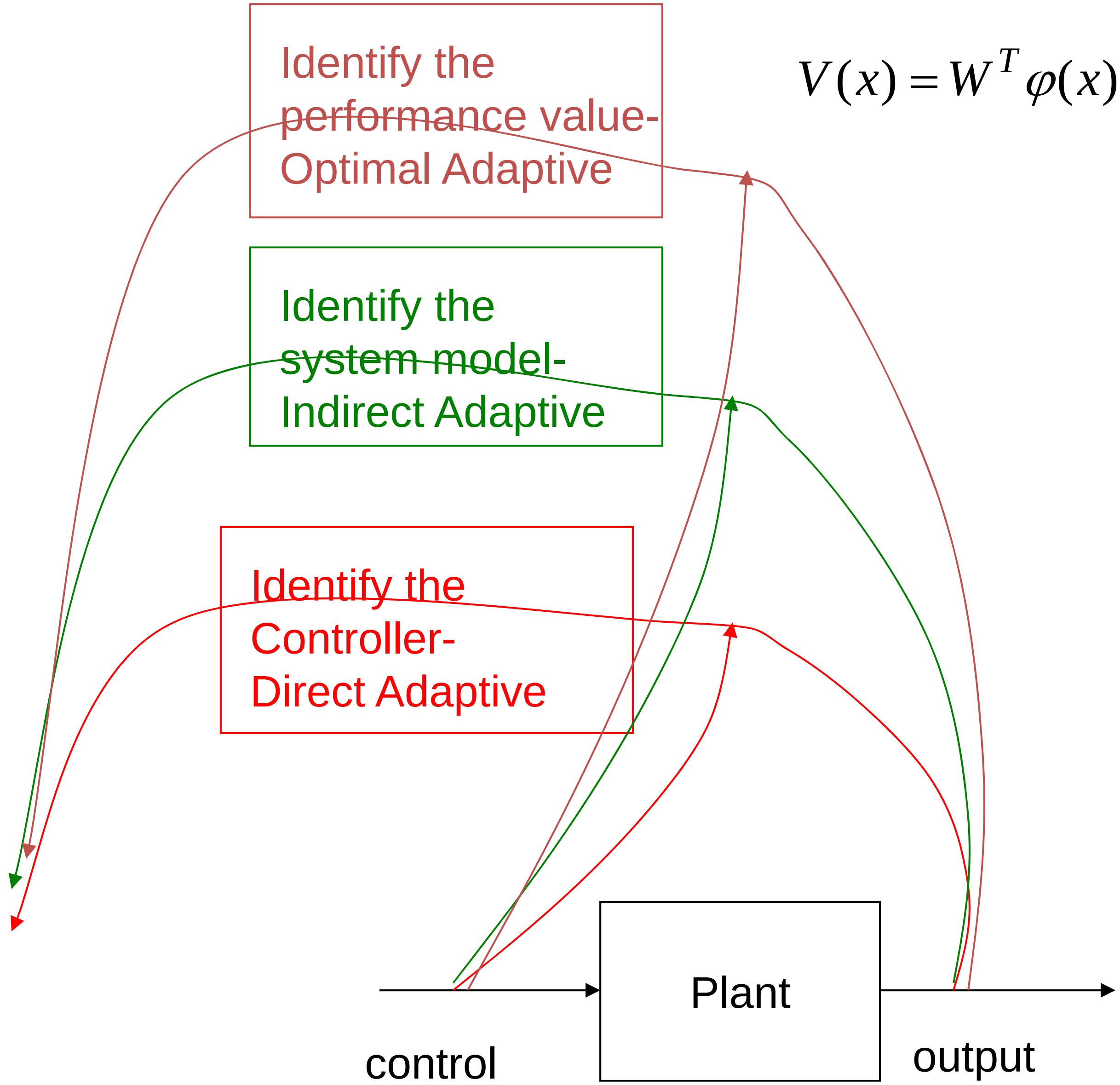
Two Learning Networks

Tune them Simultaneously

A new structure of adaptive controllers

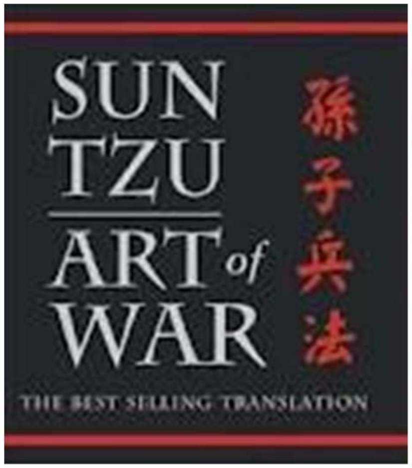
A New Class of Adaptive Control





$$V(x) = W^T \varphi(x)$$

Games on Communication Graphs



500 BC

孙子兵法

Sun Tz bin fa

Multi-player Differential Games

Game Theory-Based Control System Algorithms with Real-Time Reinforcement Learning

HOW TO SOLVE MULTIPLAYER GAMES ONLINE

KYRIAKOS G. VAMVOUDAKIS, HAMIDREZA MODARES, BAHARE KIUMARSI, and FRANK L. LEWIS

Complex human-engineered systems involve an interconnection of multiple decision makers (or agents) whose collective behavior depends on a compilation of local decisions that are based on partial information about each other and the state of the environment [1]–[4]. Strategic interactions among agents in these systems can be modeled as a multiplayer simultaneous-move game [5]–[8]. The agents involved can have conflicting objectives, and it is natural to make decisions based upon optimizing individual payoffs or costs.

Game theory has been mostly pioneered in the field of economics; [9] considered a finite win-loss game with perfect information between two players, and this classic example of computable economics stands in the long and distinguished tradition of game theory that goes back to [10] and [11]. Reference [12] discusses game theory in algorithmic modes but not in what is today referred to as *algorithmic game theory* after realizing the futility of

IMAGE COURTESY OF KYRIAKOS G. VAMVOUDAKIS.

Digital Object Identifier 10.1109/MCS.2016.2621461
Date of publication: 19 January 2017

1066-033X/17©2017IEEE

Multi-player Game Solutions
IEEE Control Systems Magazine, February 2017

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.

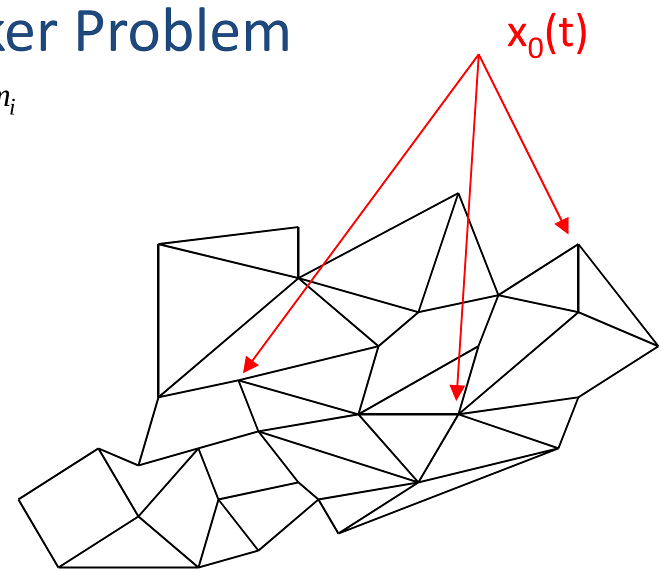
Hongwei Zhang, F.L. Lewis, and Abhijit Das
"Optimal design for synchronization of cooperative systems: state feedback, observer and output feedback," IEEE Trans. Automatic Control, vol. 56, no. 8, pp. 1948-1952, August 2011.

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)

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

Local nbhd. tracking error dynamics

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

Local agent dynamics driven by neighbors' controls

Define Local nbhd. Optimal performance index

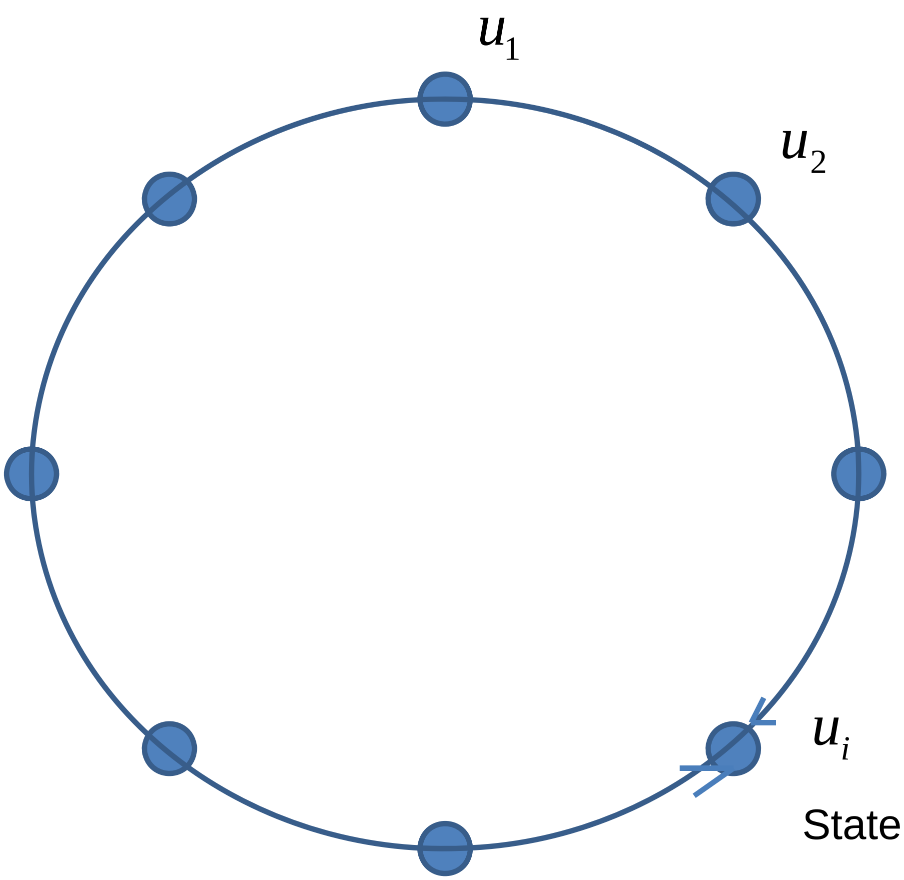
Values driven by neighbors' controls

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

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.

M. Abouheaf, K. Vamvoudakis, F.L. Lewis, S. Haesaert, and R. Babuska, "Multi-Agent Discrete-Time Graphical Games and Reinforcement Learning Solutions," Automatica, Vol. 50, no. 12, pp. 3038-3053, 2014.

New Differential Graphical Game
DISTRIBUTED ALGORITHMS- SCALABLE



Local Dynamics
Local Value Function
Only depends on
graph neighbors

Control action of player *i*

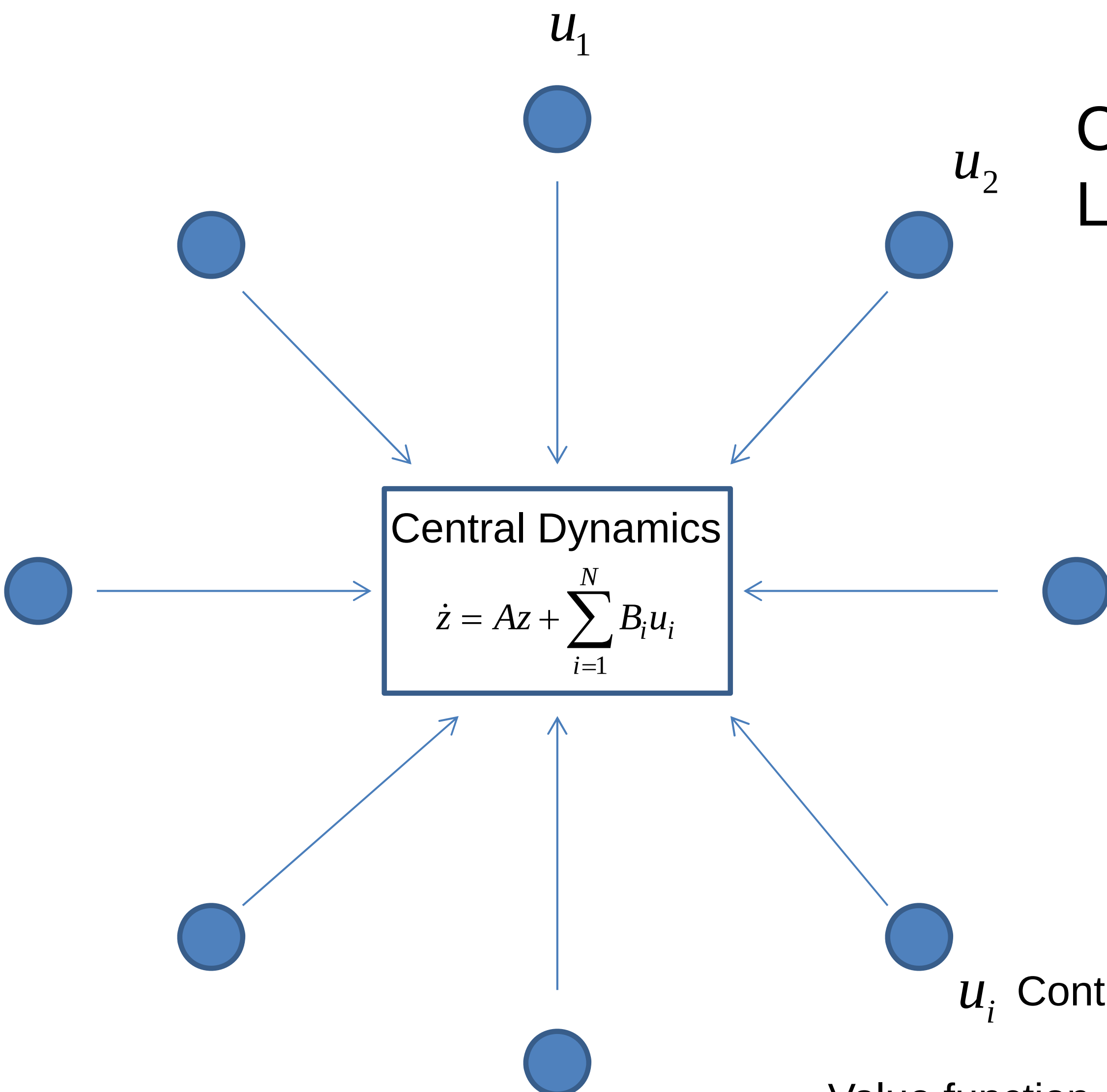
State dynamics of agent *i*

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

Value function of player *i*

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

Standard Multi-Agent Differential Game



Central Dynamics
Local Value Function
depends on ALL
other control actions

u_i Control action of player *i*

Value function of player *i*

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

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

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

Def: Interactive Nash equilibrium

$\{u_1^*, u_2^*, ..., u_{\mathrm{N}}^*\}$ are in Interactive Nash equilibrium if

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

2. There exists a policy u_j such that

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

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.

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

Now use Synchronous PI to learn optimal Nash policies online in real-time as players interact

Distributed Multi-Agent Learning Proofs

Online Solution of Graphical Games

Kyriakos Vamvoudakis

Multi-agent Learning Convergence proofs

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 $V_i^k \to V_i^*$.