Moving Bits Not Watts: Geographically Coordinated Frequency Control

Zhenhua Liu

Joint work with Joshua Comden, Tan N. Le, Yue Zhao, Bong Jun Choi

Stony Brook University, SUNY Korea

LBNL April 2018

Short bio

- Assistant Professor at Stony Brook University since 2014
  - Operations research, computer science, Smart Energy Technologies Cluster
  - On leave during 2014.6-2015.8 (ITRI-Rosenfeld Postdoctoral Fellow with Mary Ann)
- PhD in Computer Science 2014, California Institute of Technology
  - MS & BE in Computer Science and Control from Tsinghua University
  - Double degree BS in Economics from Peking University
- Research
  - Sustainable computing, Demand response, Online and distributed optimization, Scheduling and resource allocation, Big Data Systems
  - 4 NSF grants during the past 3 years
  - ~30 publications including SIGMETRICS, NSDI, ~1,600 citations
  - 4 US patents filed, 2 of which have been awarded

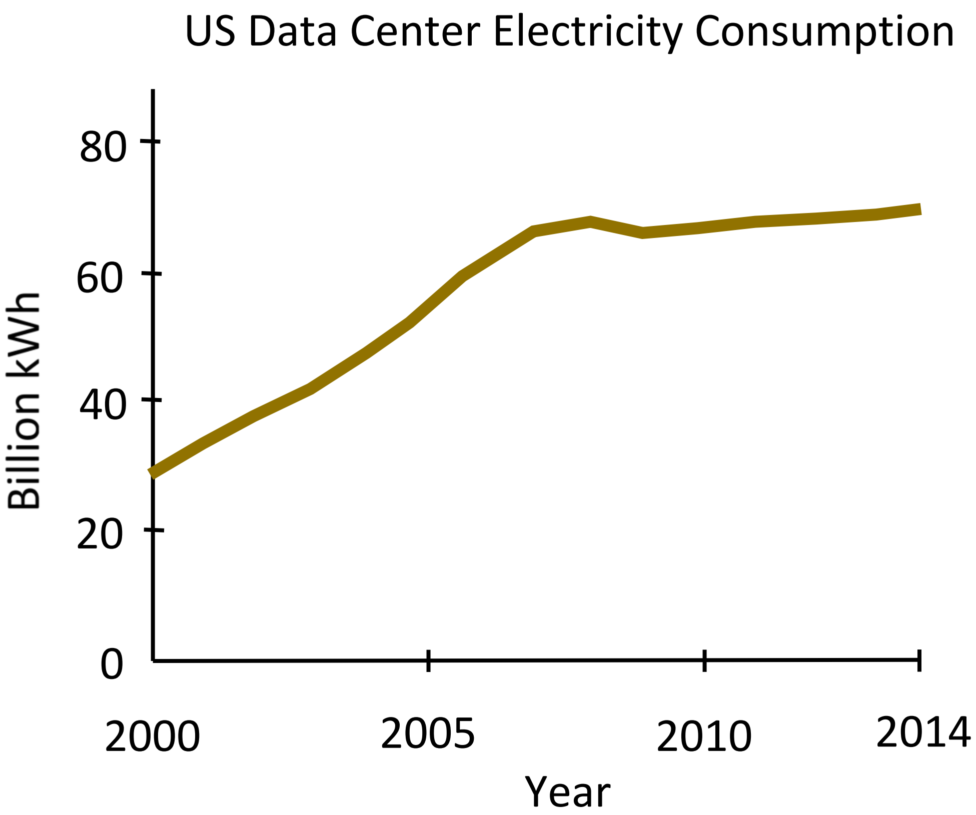
My Wishlist

- Comments/suggestions on my research
- Exploring opportunities
  - Research: real data, systems, domain knowledge, etc
  - Funding: NYSERDA, DoE, other federal agency
    - We can lead or sub
- Long-term collaborations

Research examples

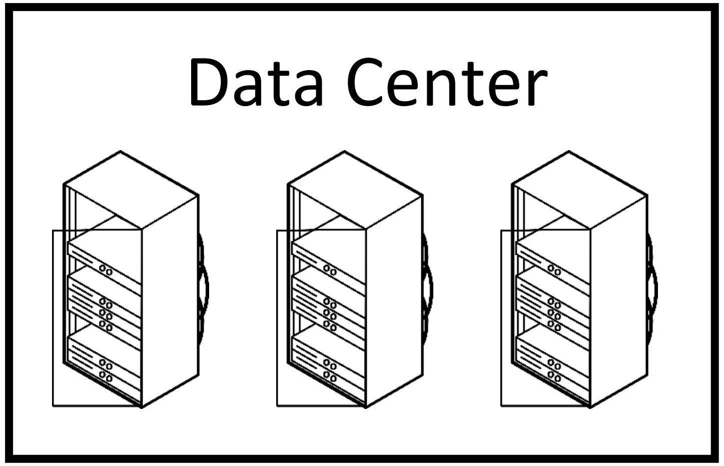
- Distributed Optimization
  - Geographical Load Balancing + Distributed Frequency Control
  - Demand Response Program Design
- Online Optimization
  - Smoothed Online Convex Optimization
  - Coincident Peak Pricing, Multi-scale electricity markets
- Big Data Systems
  - Multi-resource allocation
  - Bounded Priority Fairness, Interchangeable Resource Allocation

Data Center ability for Frequency Control



United States Data Center Energy Usage Report,
Lawrence Berkeley National Laboratory, 2016.

2% Total US Consumption

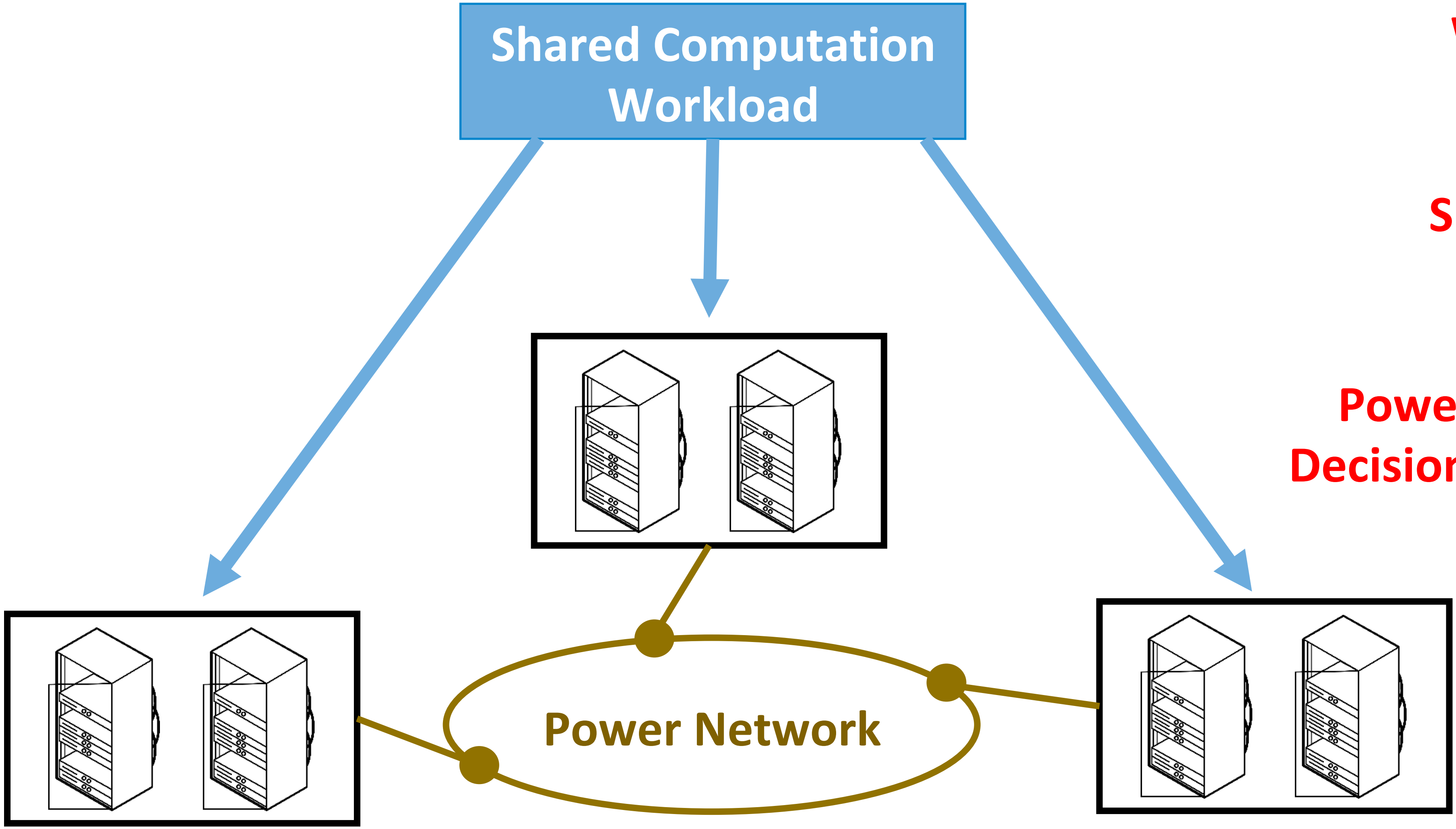


Fast & Precise Energy Control Systems

E.g. Dynamic Frequency Scaling of CPUs
Energy Storage within UPSs

Large potential for Frequency Control

Cloud Computing is Interdependent



Delayed or Unprocessed Workload

↓

Shared Cost

↓

Power Consumption Decisions are Dependent

The need for Distribution in Primary FC

Issues with Centralized Control

Communication

Infeasible for large networks

Privacy

Agents don't want to share private objectives

+

Primary Frequency Control

Speed

Requires frequency stabilization in seconds

Distributed Control for PFC

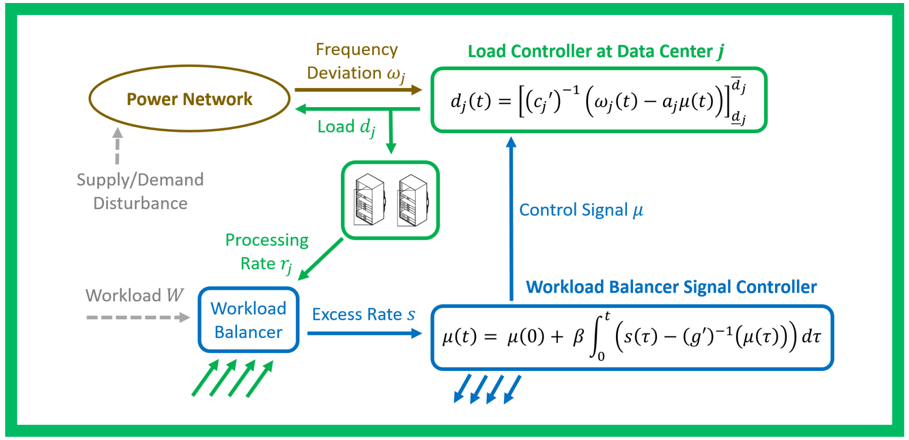
Optimal Decentralized Primary Frequency Control in Power Networks
C. Zhao and S. Low - CDC 2014

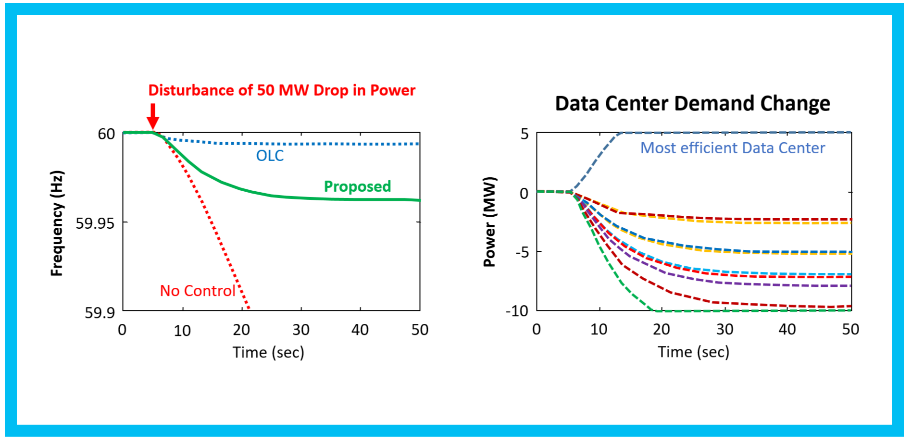
Assumes Independent Costs → Separable Objective Function (valid for some applications)

Cloud Computing Costs are Interdependent → Inseparable Objective Function

Goal: Design Primary FC that uses a **Network of Data Centers**

Approach: Incorporate **Interdependent** Costs





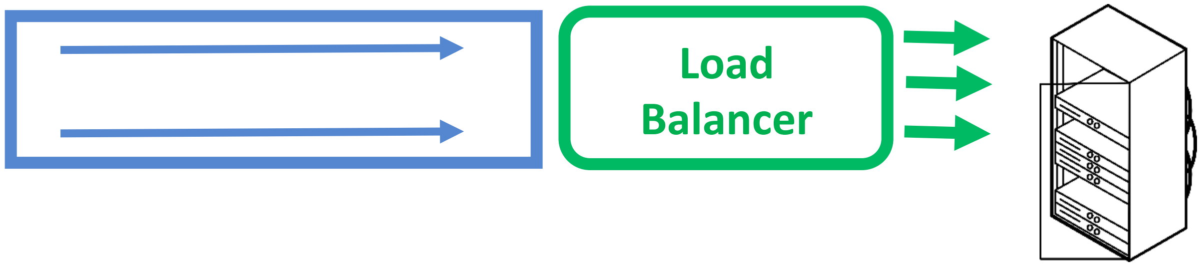
Distributed Control Laws

Simulation Results

Cloud Computing Model

Workload Incoming Rate

Data Center
Processing Rates



r_1

r_2

⋮

Excess Processing Rate
of the Network

$s := \sum_j r_j - W$

$s < 0$

Workload
unprocessed or **delayed**

$g(s)$

Network-wide
(Interdependent) Cost

Single Data Center
Power Consumption

$d_j := d_j + 1/a_j \; r_j$

$c_j(d_j)$

Cost to convert
electrical → computing
power

Linear power usage profile:
d_j constant overhead
a_j conversion coefficient

Power Network Model

Non-controllable loads

Controllable loads (Data Center)

Bus j Real Power Injection

$$P_j := p_j - D_j\,\omega_j - d_j$$

Frequency-sensitive

Voltage phase angle (w.r.t. nominal θ_0)

Frequency Deviation

$$\omega_j = d\theta_j / dt$$

Maximum power flow across line $(j,k)\epsilon E$

Real Power Flows

$$F_j(\boldsymbol{\theta}) := \sum_{k:(j,k)\epsilon E} Y_{jk} \sin(\theta_j - \theta_k) - \sum_{i:(i,j)\epsilon E} Y_{ij} \sin(\theta_i - \theta_j)$$

System Dynamics Model

Swing Equation

$$M_j \frac{d\omega_j}{dt} = p_j - D_j \omega_j - d_j - F_j(\boldsymbol{\theta})$$

Physical Inertia

Equilibrium $(\boldsymbol{\theta}^*, \boldsymbol{\omega}^*, \boldsymbol{P}^*, \boldsymbol{d}^*, s)$ $\forall j$:

$$\frac{d\omega_j^*}{dt} = 0$$

(no change in frequency)

$$\frac{dP_j^*}{dt} = 0$$

(no change in power injection)

$$\omega_j^* = \omega$$

(all buses at same frequency)

Geographic Frequency Control Problem

Minimize Cost of PFC to Data Centers + Cost of Frequency Deviations

$s, \boldsymbol{d}, \boldsymbol{\omega}$

$$g(s) + \sum_j c_j(d_j)$$

$$\sum_j \frac{D_j}{2} \omega_j^2$$

s.t. Computational Processing Power Balance

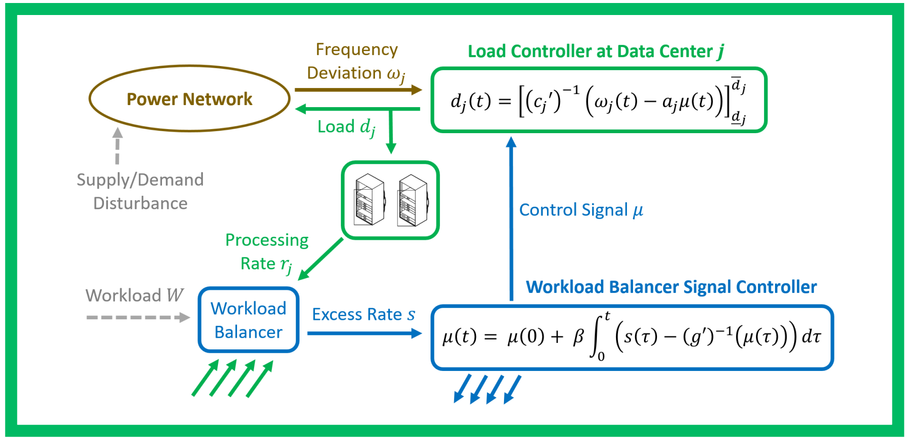
$$s = \sum_j a_j d_j - W - \sum_j a_j d_j$$

Electrical Power Network Balance

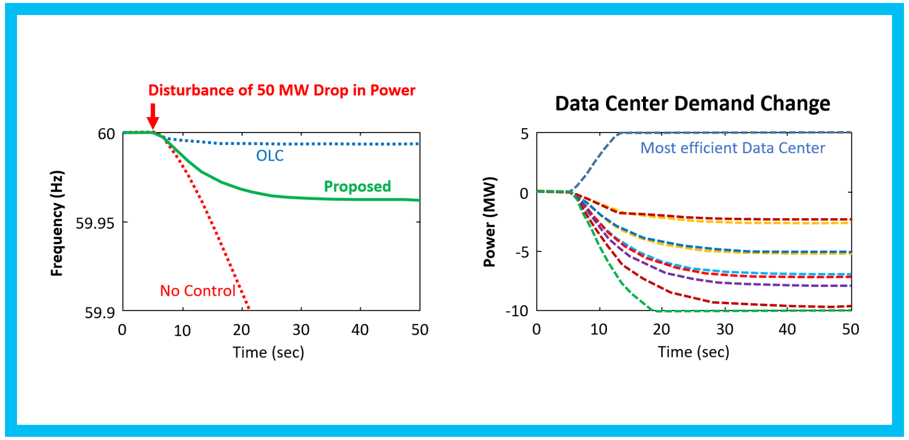
$$0 = \sum_j (p_j - D_j \omega_j - d_j)$$

Goal: Design Primary FC that uses a **Network of Data Centers**

Approach: Incorporate **Interdependent** Costs

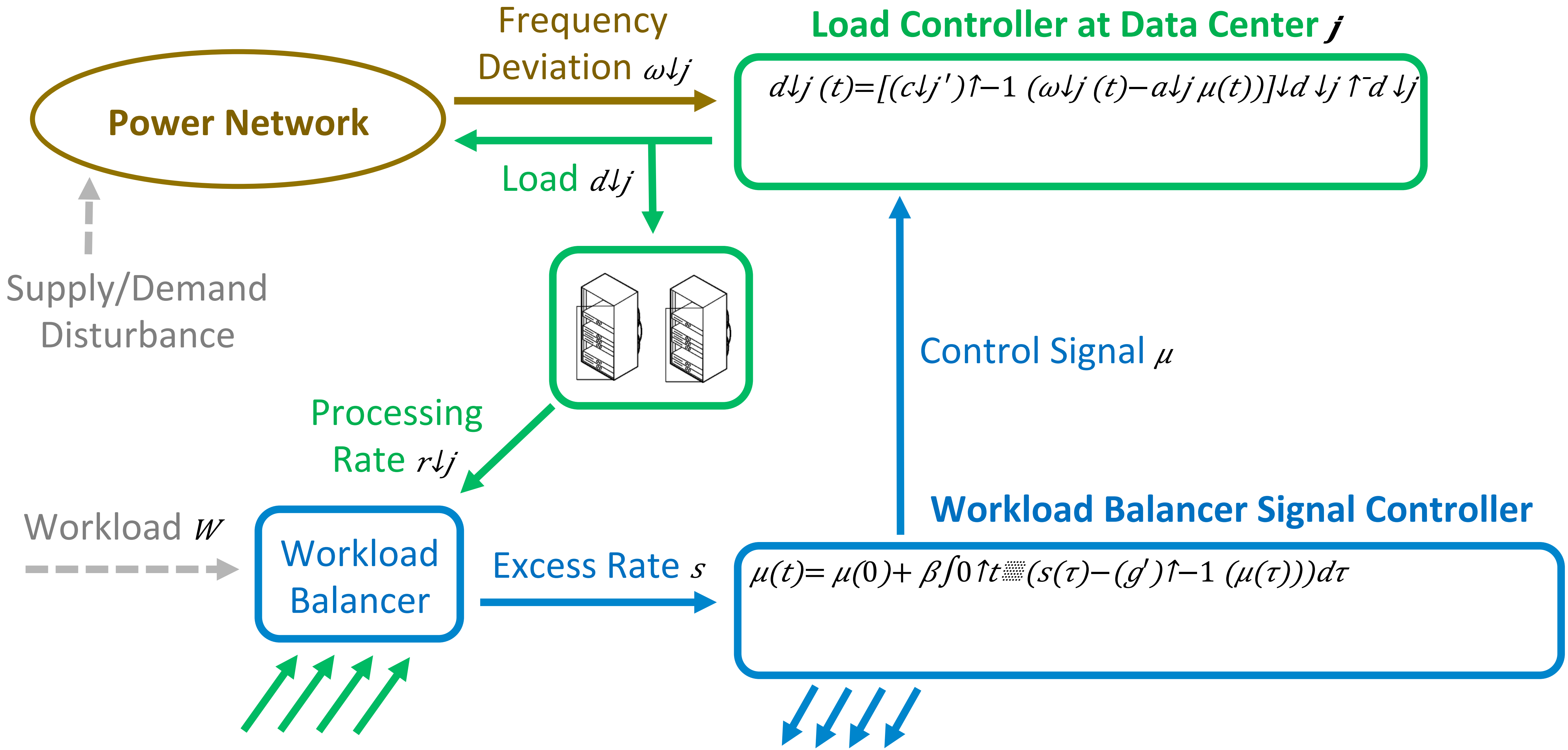


Distributed Control Laws



Simulation Results

Distributed Control Laws



Control Laws Converge to an Optimal Point

Stability

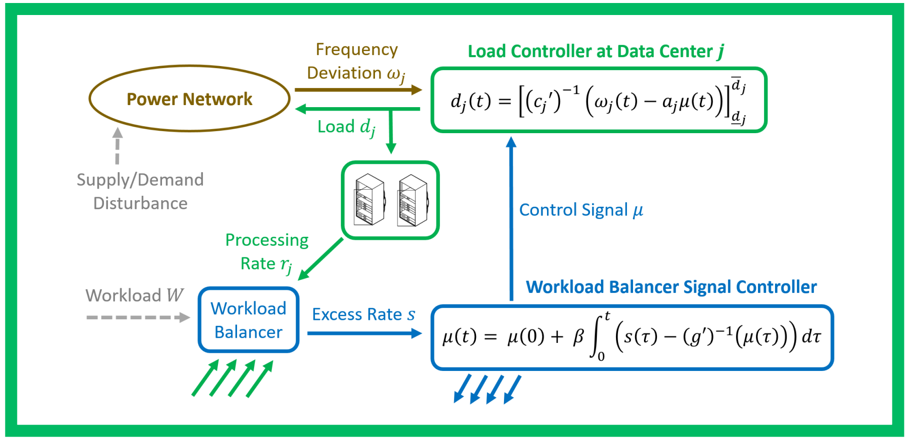
Theorem: The trajectory of $(\boldsymbol{\omega}, \boldsymbol{P}, \boldsymbol{d}, s, \mu)$ **asymptotically converges** to an equilibrium point $(\boldsymbol{\omega}^*, \boldsymbol{P}^*, \boldsymbol{d}^*, s^*, \mu^*)$.

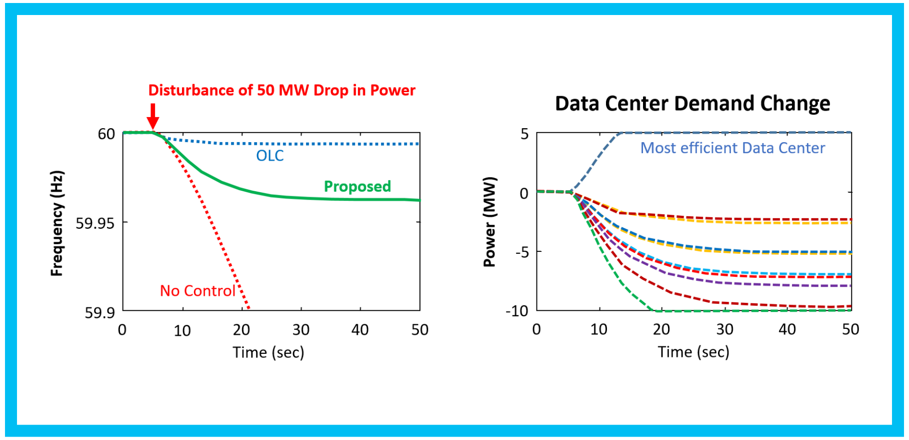
Optimality

Theorem: An equilibrium point $(\boldsymbol{\omega}^*, \boldsymbol{P}^*, \boldsymbol{d}^*, s^*, \mu^*)$ is **optimal** to the Geographic Frequency Control Problem.

Goal: Design Primary FC that uses a **Network of Data Centers**

Approach: Incorporate **Interdependent** Costs



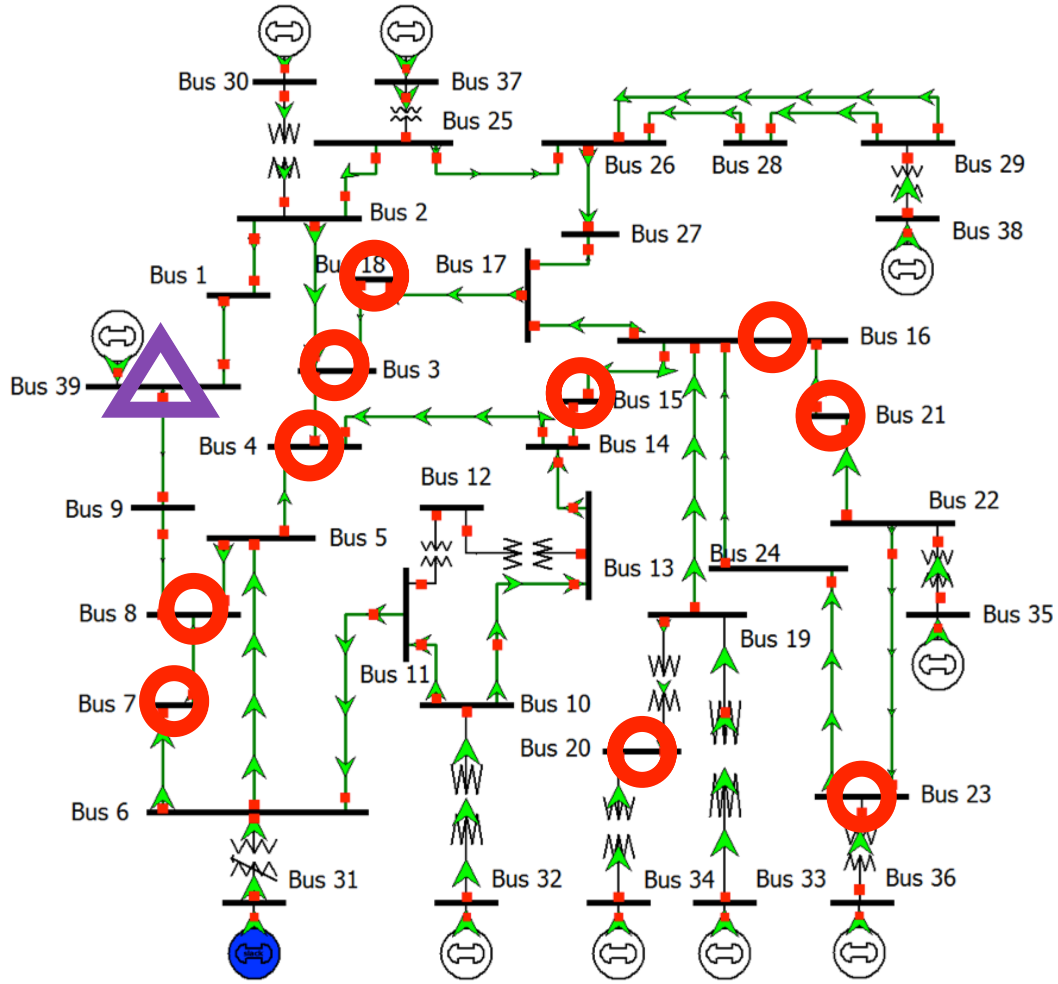


Distributed Control Laws

Simulation Results

Simulation Setup

New England IEEE 39-bus



http://icseg.iti.illinois.edu/ieee-39-bus-system/

Power System Toolbox (Matlab)

○ = 25 MW Data Center

Each with different efficiency

10 Data Centers = 1.8% Total Demand

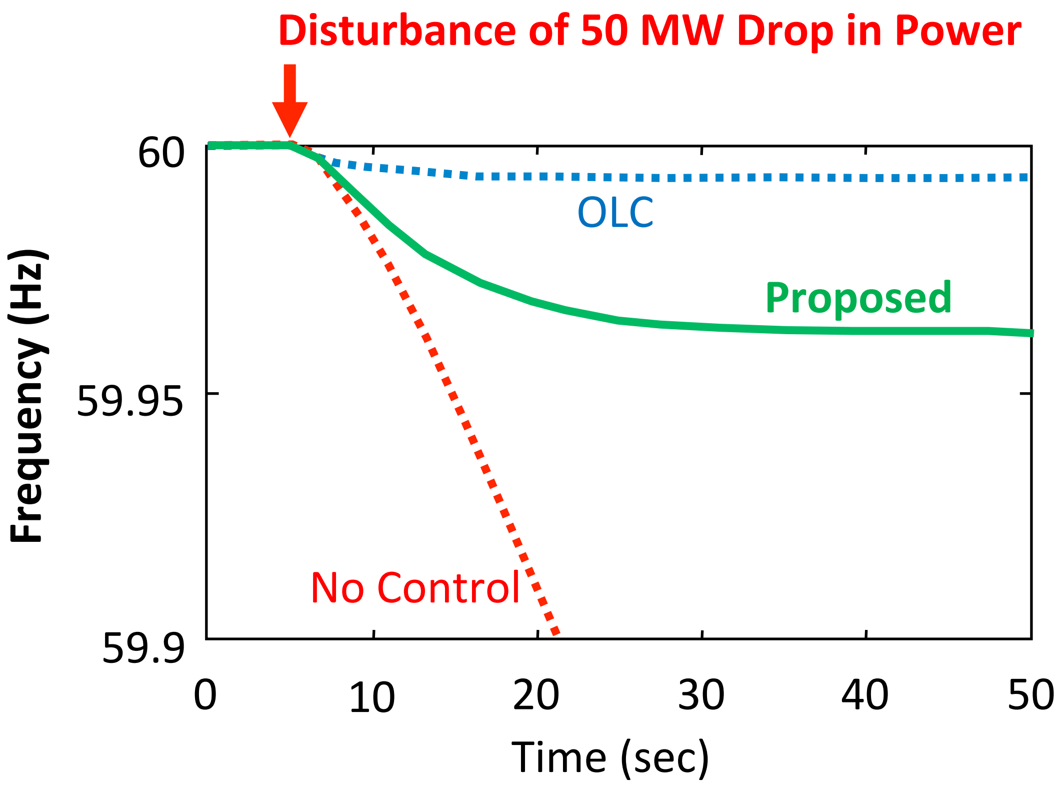
Each Data Center shares a fraction of a 100 MW equivalent computational workload.

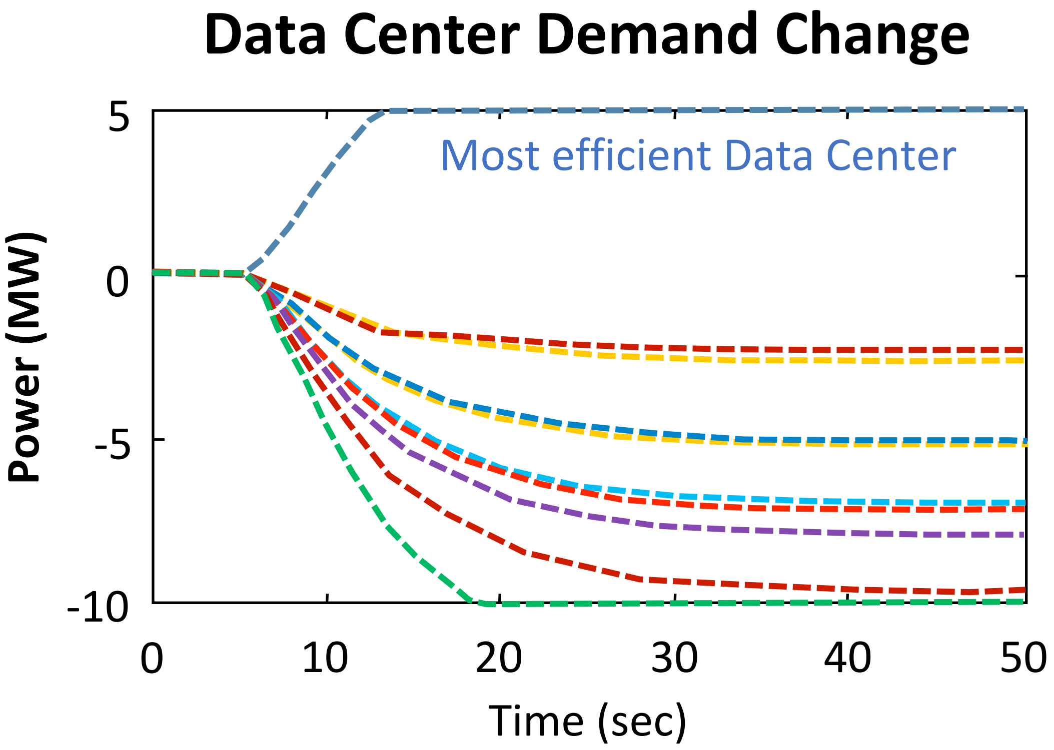
Interdependent cost: $g(s)=\gamma s^2$

△ = Disturbance of 50 MW Drop in Power

Proposed Control Laws Stabilize the System

Converges to equilibrium within 30 seconds



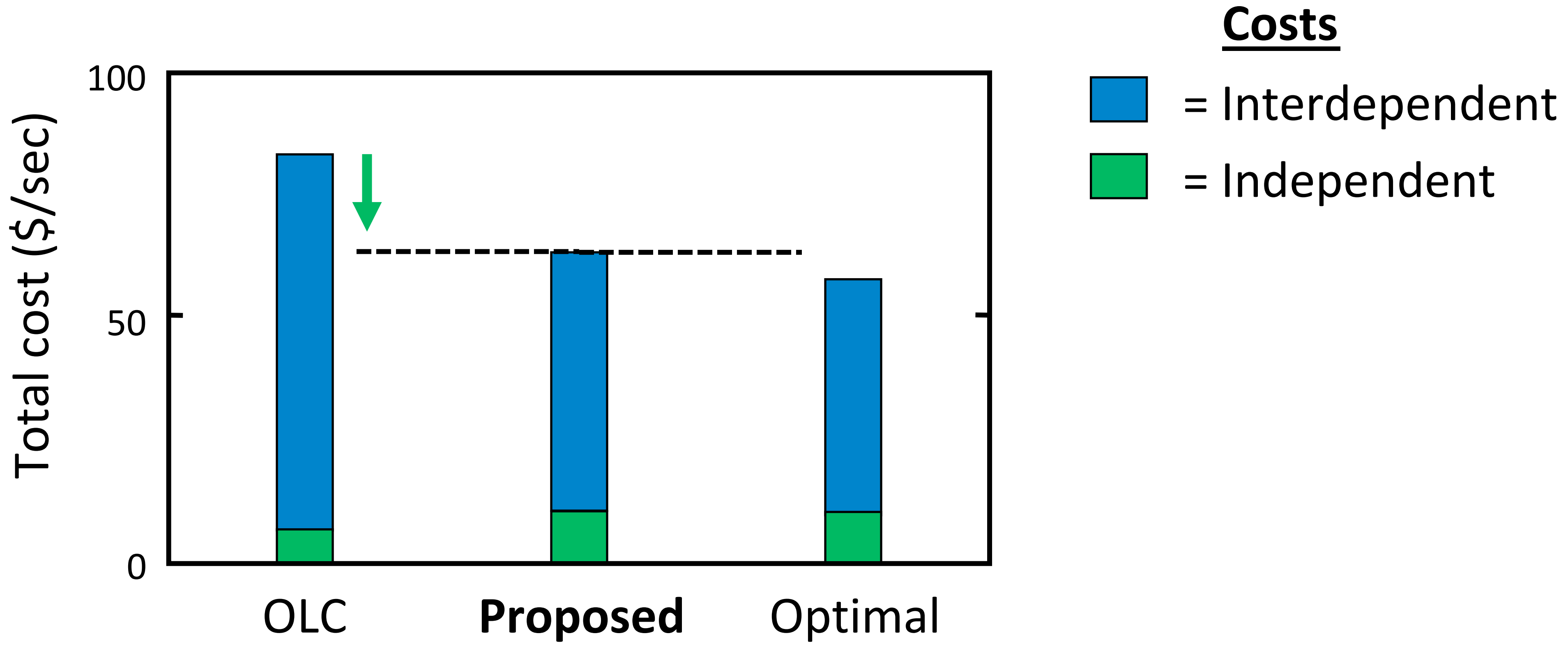


OLC (No Interdependent Costs)

Optimal Decentralized Primary Frequency Control in Power Networks

C. Zhao and S. Low - CDC 2014

Cost of Proposed Control Laws is near Optimal



Proposed incorporates **Interdependent costs**, whereas OLC does not.

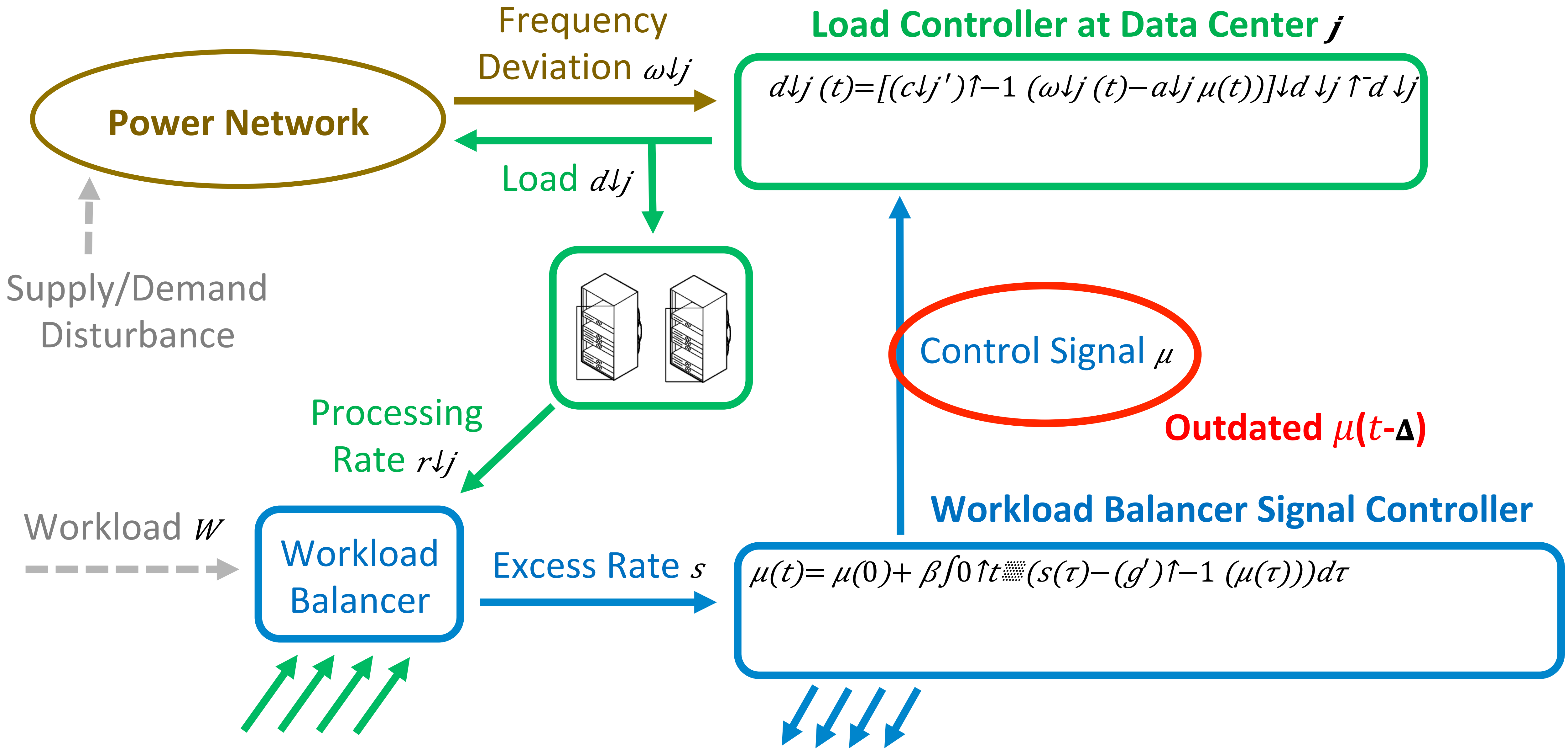
Goal: Design Primary FC that uses a **Network of Data Centers**

Approach: Incorporate **Interdependent** Costs

Asymptotically Converges toward Optimal Cost	Converges to an Equilibrium that is near Optimal Cost
Distributed Control Laws	Simulation Results

Bonus: Communication and implementation Delay

Distributed Control Laws



System with delay

Equilibrium
$(\boldsymbol{\theta}^*, \boldsymbol{\omega}^*, \boldsymbol{P}^*, \boldsymbol{d}^*, s)$

$\forall j$:

$\frac{d\omega_j^*}{dt} = 0$ (no change in frequency)

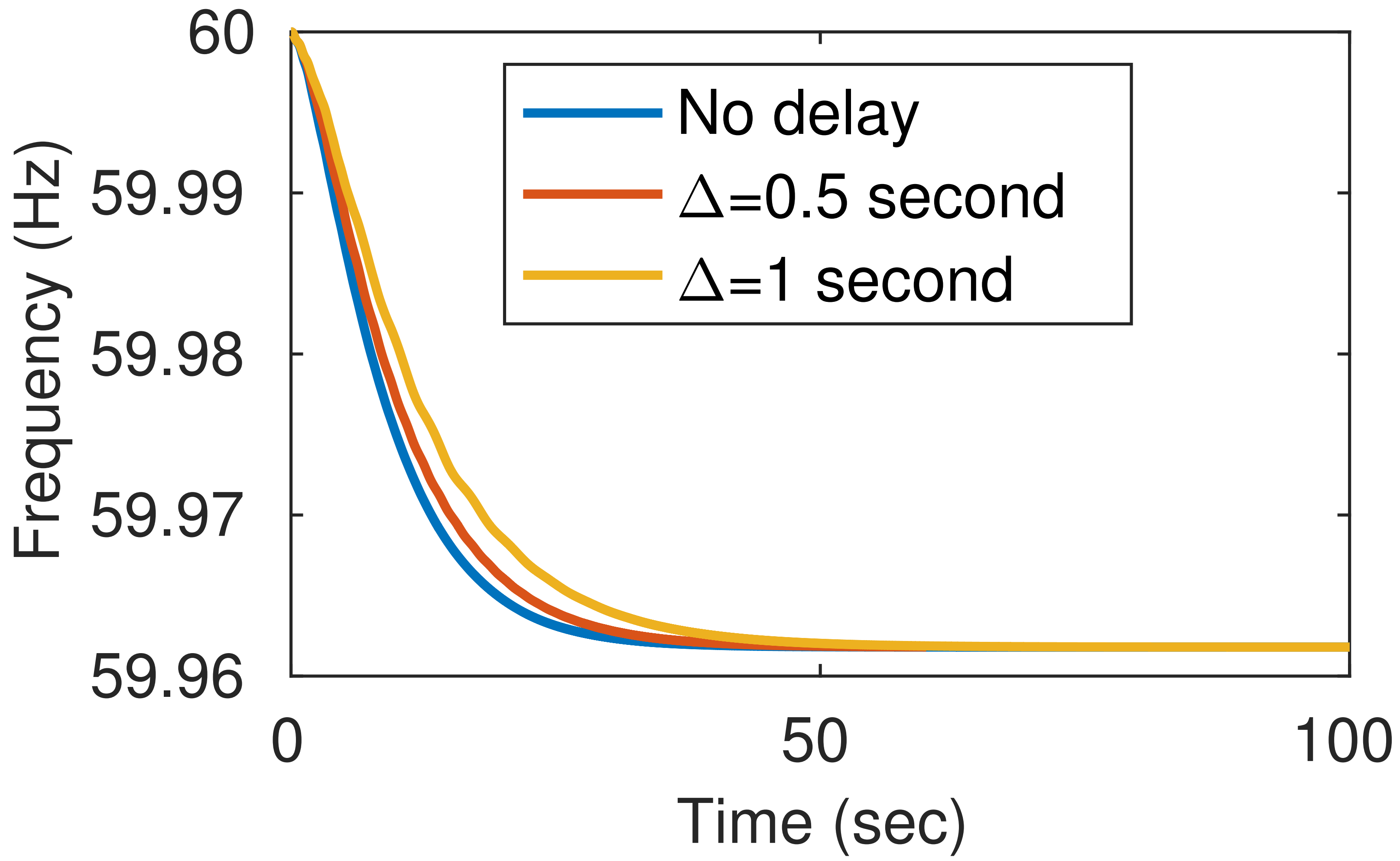
$\frac{dP_j^*}{dt} = 0$ (no change in power injection)

$\omega_j^* = \omega$ (all buses at same frequency)

And hold for at least Δ of time

Corollary: An equilibrium point $(\boldsymbol{\omega}^*, \boldsymbol{P}^*, \boldsymbol{d}^*, s^*, \mu^*)$ is **optimal** to the Geographic Frequency Control Problem.

Stability: if β is small enough, the trajectory asymptotically approaches an equilibrium point.



(c) Impact of communication delay

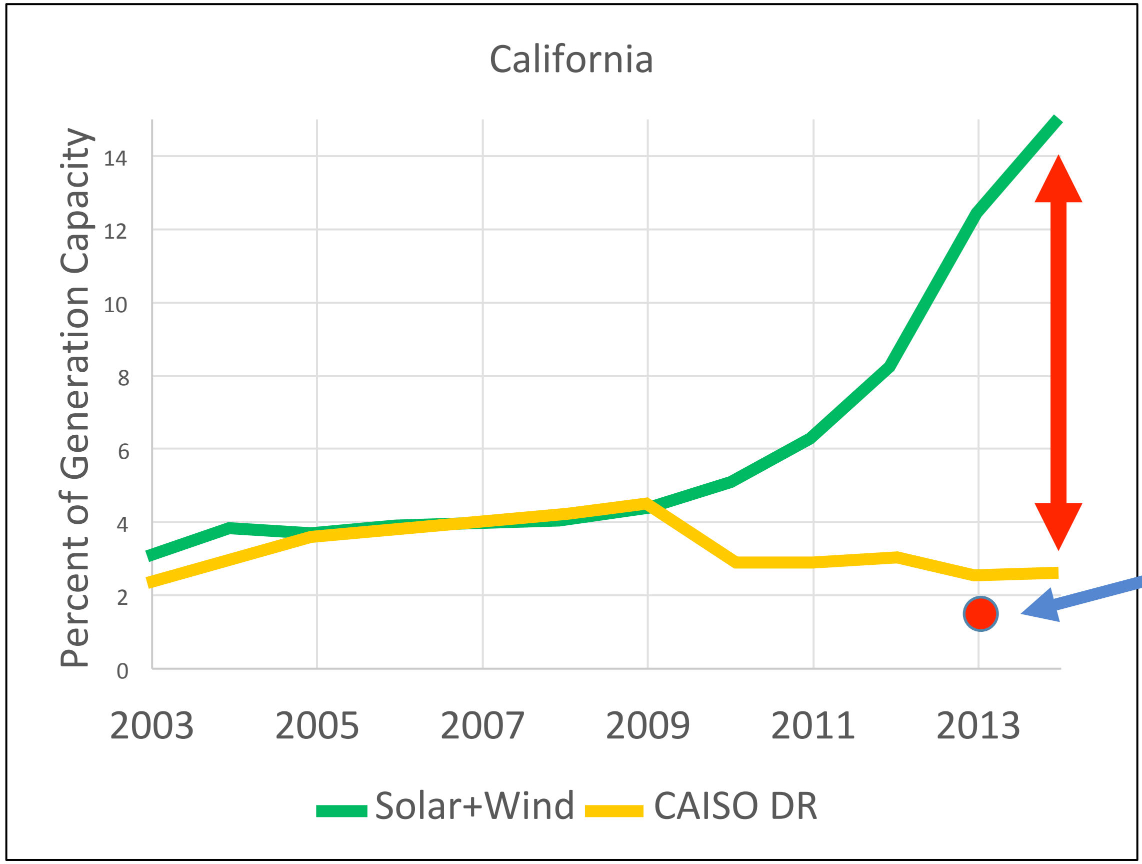
Future work

Include other geographic interdependent systems,
e.g. thermal grids, electric mass transit, natural gas.

Research examples

- Distributed Optimization
  - Geographical Load Balancing + Distributed Frequency Control
  - Demand Response Program Design
- Online Optimization
  - Smoothed Online Convex Optimization
  - Coincident Peak Pricing, Multi-scale electricity markets
- Big Data Systems
  - Multi-resource allocation
  - Bounded Priority Fairness, Interchangeable Resource Allocation

Solar + Wind energy outpaces DR adoption



California State Legislature solution (2013):
1,325 MW of grid storage by 2020

Expensive

http://energyalmanac.ca.gov/electricity/electric_generation_capacity.html
FERC Assessment of Demand Response and Advanced Metering Staff Reports: 2010-2015.
CAISO Demand Response Barriers Study 2009.

DR programs allocate customer uncertainties

Because DR Programs have various Levels of Commitment

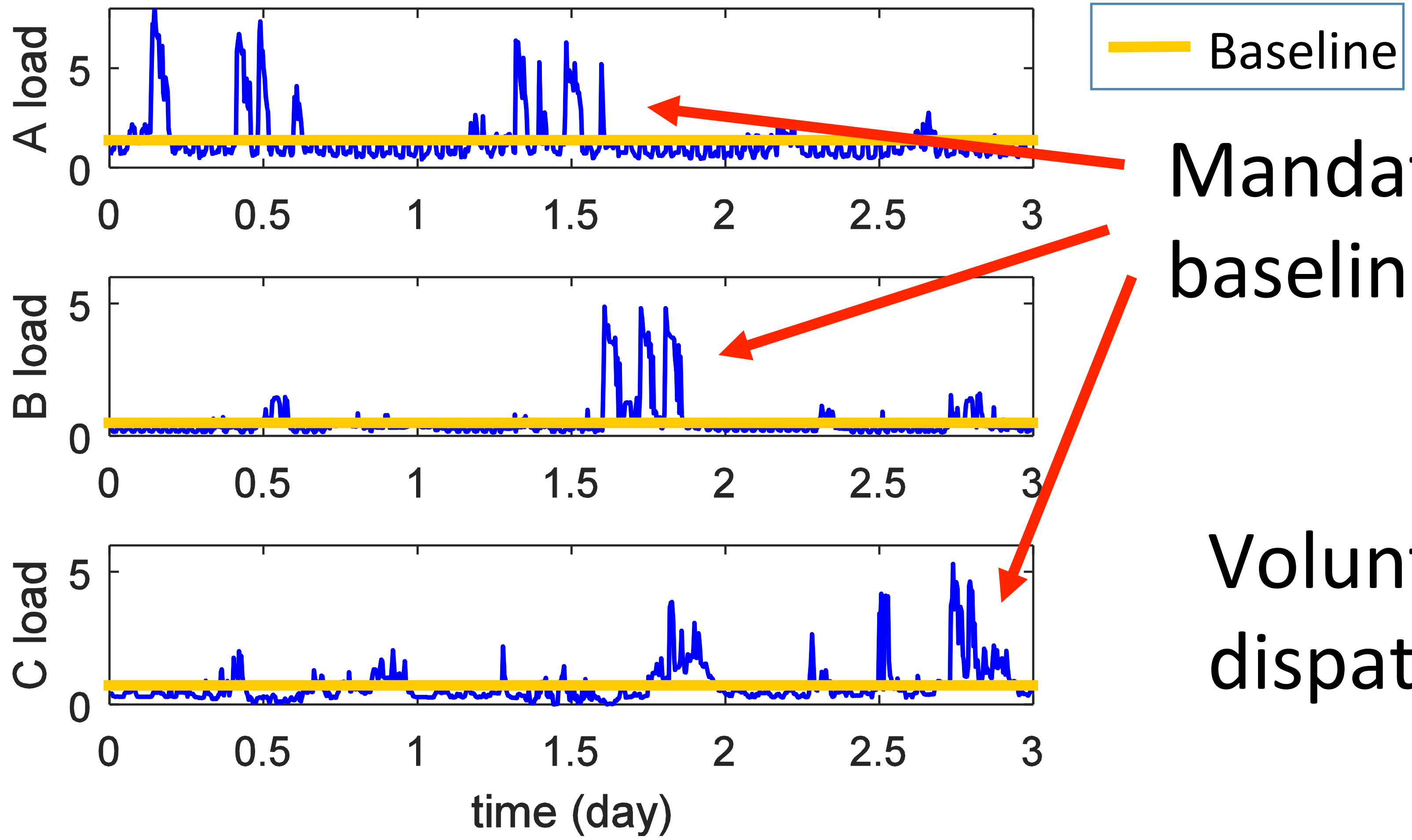
Customer Load Reductions

Customer is responsible

LSE is responsible

Mandatory		Voluntary
Base Interruptible Program (BIP)	← CAISO →	Demand Bidding Program (DBP)
Emergency Response Service (ERS)	← ERCOT →	Voluntary Load Reduction (VLR)
Special Case Resources (SCR)	← NYISO →	Emergency DR Program (EDRP)

Customer uncertainties



Mandatory DR to baseline may be costly

vs.

Voluntary DR is not dispatchable for LSE.

UMass Trace Repository: Smart* Data Set

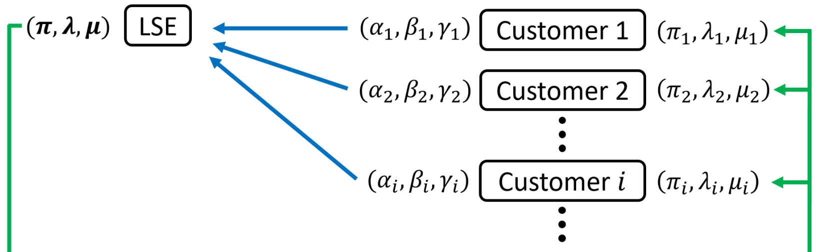
Challenges in handling **customer uncertainties**

LSE does **not know** each of the customer's uncertainties.

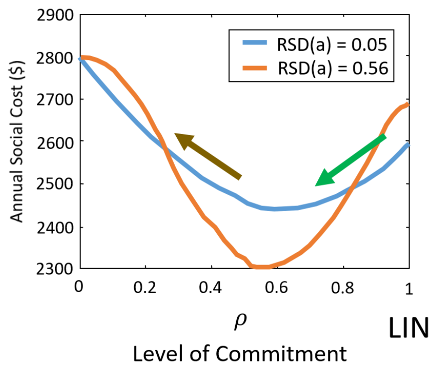
DR programs that have customers take some responsibility are **mandatory**.

Goal: Increase **reliable** DR adoption

Approach: Incorporate Customer Uncertainties



Distributed Algorithm



Leverage Randomness

Linear contract

Load Reduction = **Aggregate mismatch** + **Individual mismatch** + **Constant reduction**

$$x_i(D, \delta_i) = \alpha_i D + \beta_i \delta_i + \gamma_i$$

- ✓ Simple and easy to implement
- ✓ Within 10% Offline Optimal solution
- ✓ Real-time Optimal: Quadratic cost functions

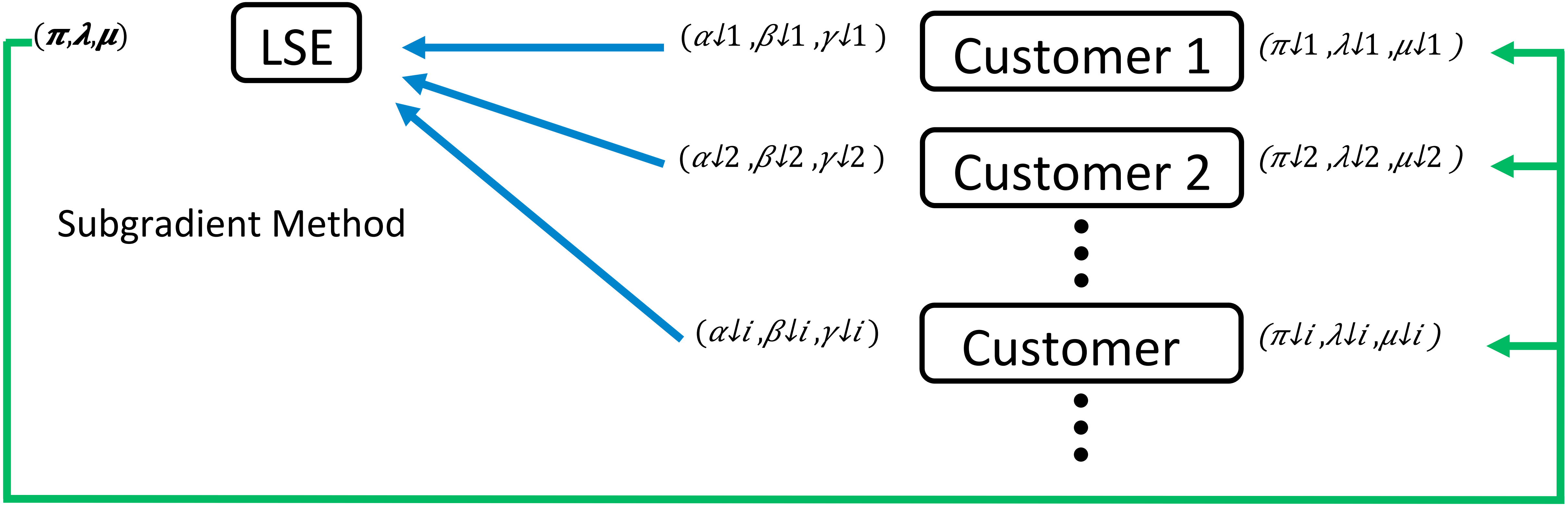
Customer cost functions raise challenges

Accuracy: LSE does not know their cost functions

Privacy: Customers do not want to give up their information

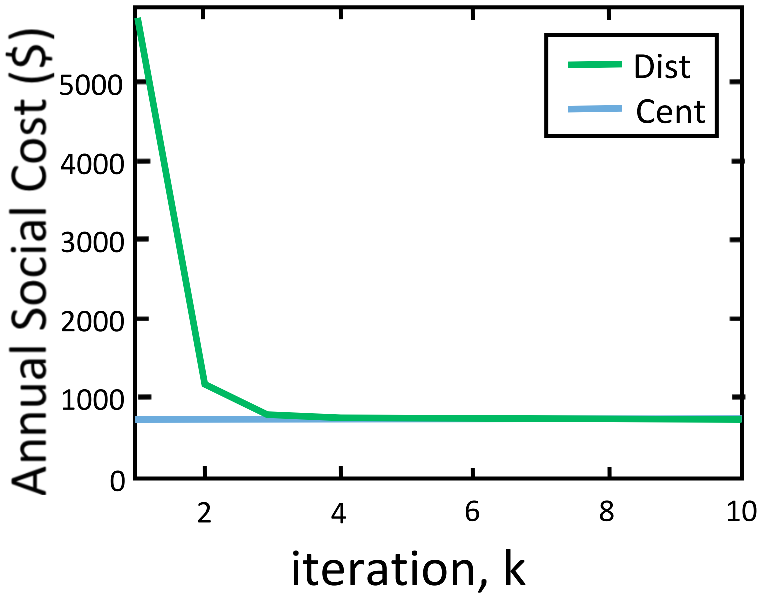
→ **Separate the DR decision problem**

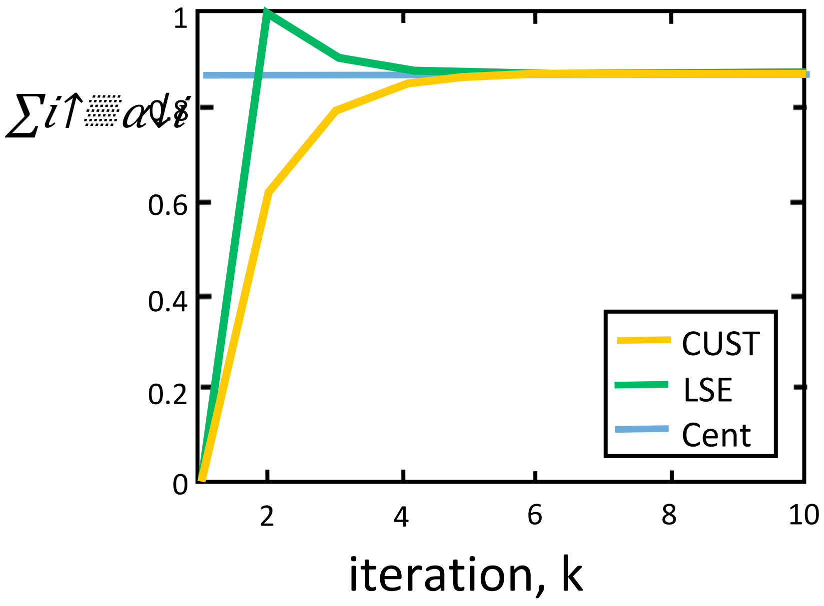
Distributed Algorithm

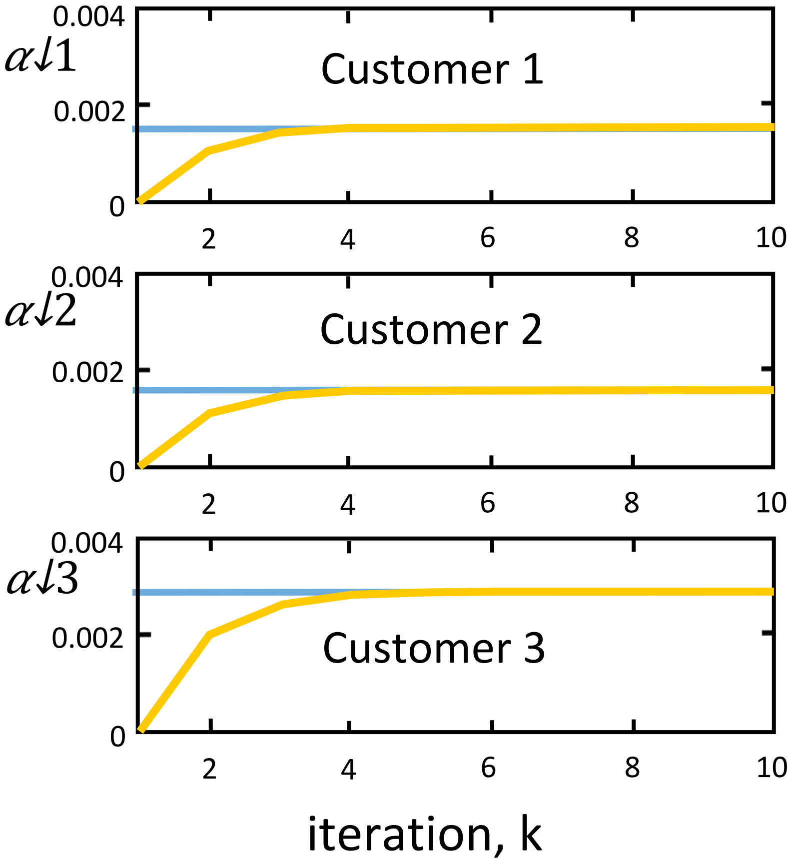


Distributed Algorithm converges to Centralized

<u>Theorem</u>: The distributed algorithm's trajectory of prices **converges to the optimal centralized prices**.

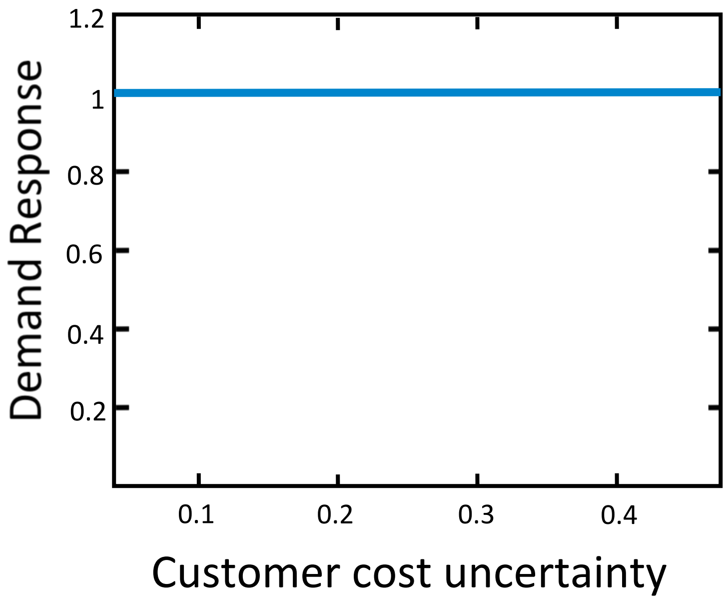


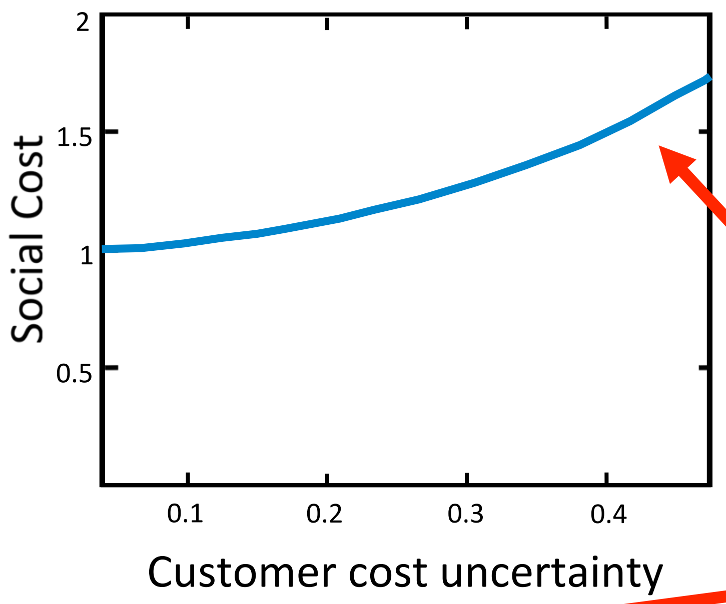




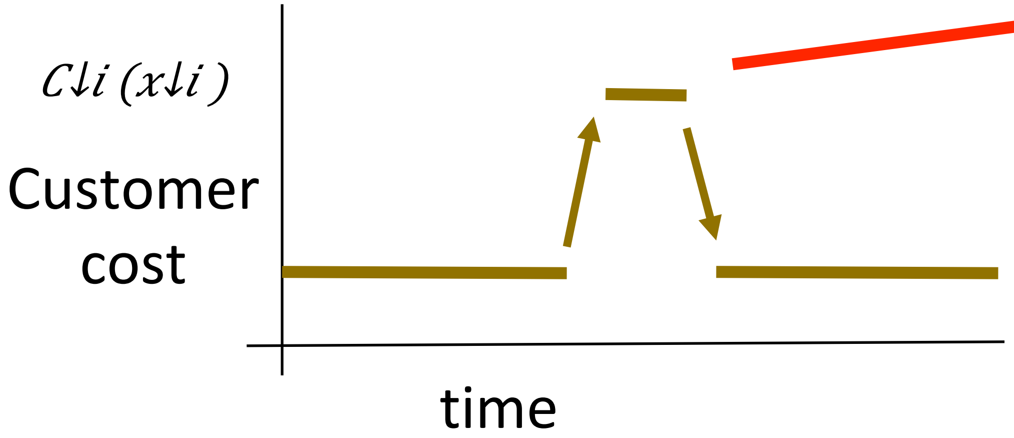
(β_i, γ_i remain at 0 for centralized and distributed solutions)

Linear contract drawback: It is **mandatory**.





Required compliance during high customer cost periods



Ex. Important tasks that need to be completed

Aim to avoid high customer cost periods

Level of Commitment

0 LIN+(ρ) 1

Customer commits to ρ fraction of timeslots to follow LIN ← LIN

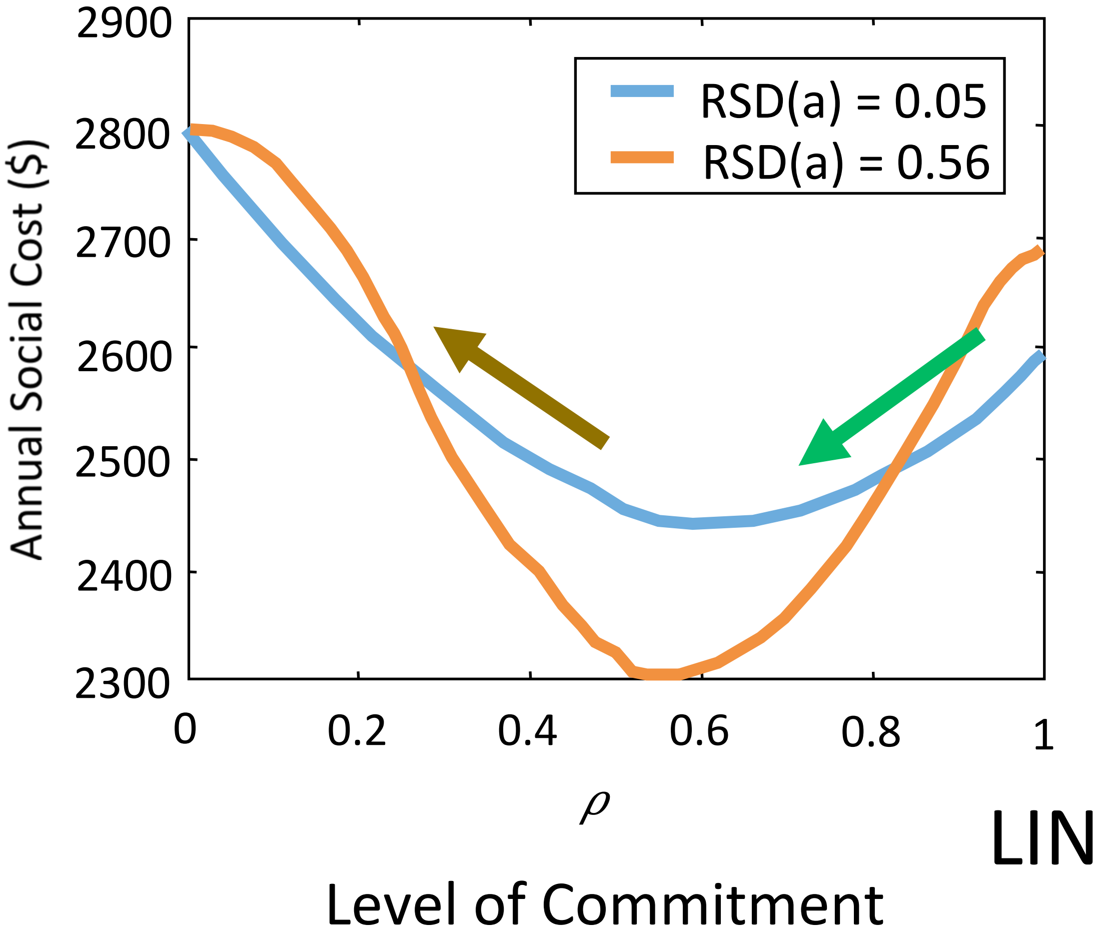
Commitment decision based **only** on realized cost function

Ex. Quadratic
$C_i(x_i) = a_i x_i^2$

Customer only knows a_i before deciding to **commit** to DR

LIN+$_{(\rho)}$ reduces cost further than LIN only

$C_i(x_i(t);t)=a_i(t)(x_i(t))^2$



Decrease from = 1
→**avoids high customer costs**

Further decrease of
→**larger mismatch for LSE**

Larger customer cost uncertainty
→ larger savings from **LIN+**$_{(\rho)}$

Goal: Increase **reliable** DR adoption

Approach: Incorporate Customer Uncertainties

Converges to centralized solution

Lower Social Cost closer to Offline Optimal

Distributed Algorithm

Leverage Randomness

Future work:

Incorporate power network constraints

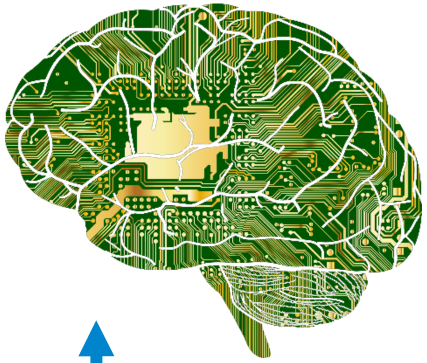
Ex. Line constraints

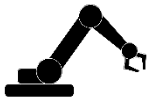
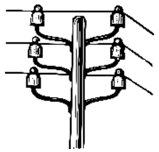
Congested lines
→Necessary to control local mismatches δ_i locally (with β_i)

Research examples

- Distributed Optimization
  - Geographical Load Balancing + Distributed Frequency Control
  - Demand Response Program Design
- Online Optimization
  - Smoothed Online Convex Optimization
  - Coincident Peak Pricing, Multi-scale electricity markets
- Big Data Systems
  - Multi-resource allocation
  - Bounded Priority Fairness, Interchangeable Resource Allocation

Artificial Intelligence



- Autonomous Vehicles 
- Robotics 
- Smart Grid 

Continual Learning

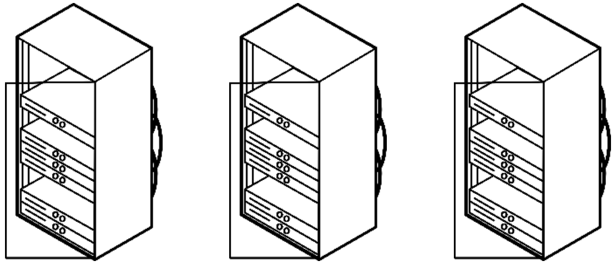
Stoica, et al. (2017)

1. Real-time decisions
2. Changing environment
3. Switching Costs
4. Utilize past information

Example: Dynamic Capacity Provisioning

Demand
y_t

Data Center



x_t = # servers on

Operating Cost + **Switching Cost**

$h(x_t, y_t)$ $\beta\|x_t - x_{t-1}\|$

How many servers should be turned on/off right now?

Future demand predicted at t

$y_{t+1|t}, y_{t+2|t}, \dots, y_{t+w-1|t}$

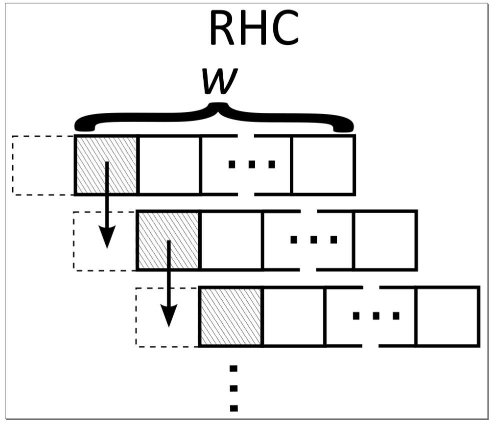
Goal: Design an Online Algorithm with Performance Guarantees.

1-Dimensional

1-Step Prediction:

- Lin, et al. (2011): **3-competitive** Lazy Capacity Provisioning
- Bansal, et al. (2015): **3-competitive** Memoryless
- Bansal, et al. (2015): **2-competitive**

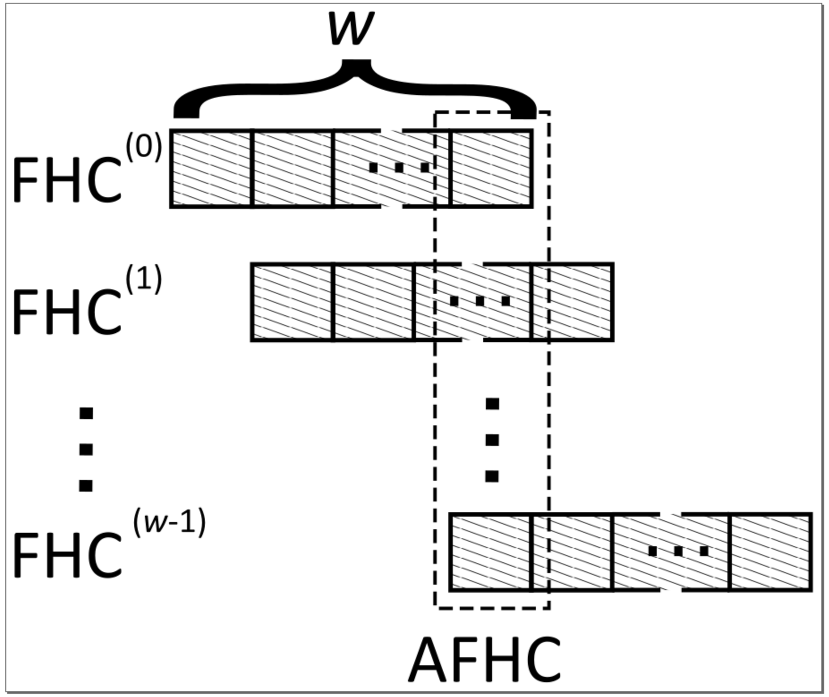
IGCC12: **1+$O(1/w)$-competitive** Receding Horizon Control



Multi-Dimensional

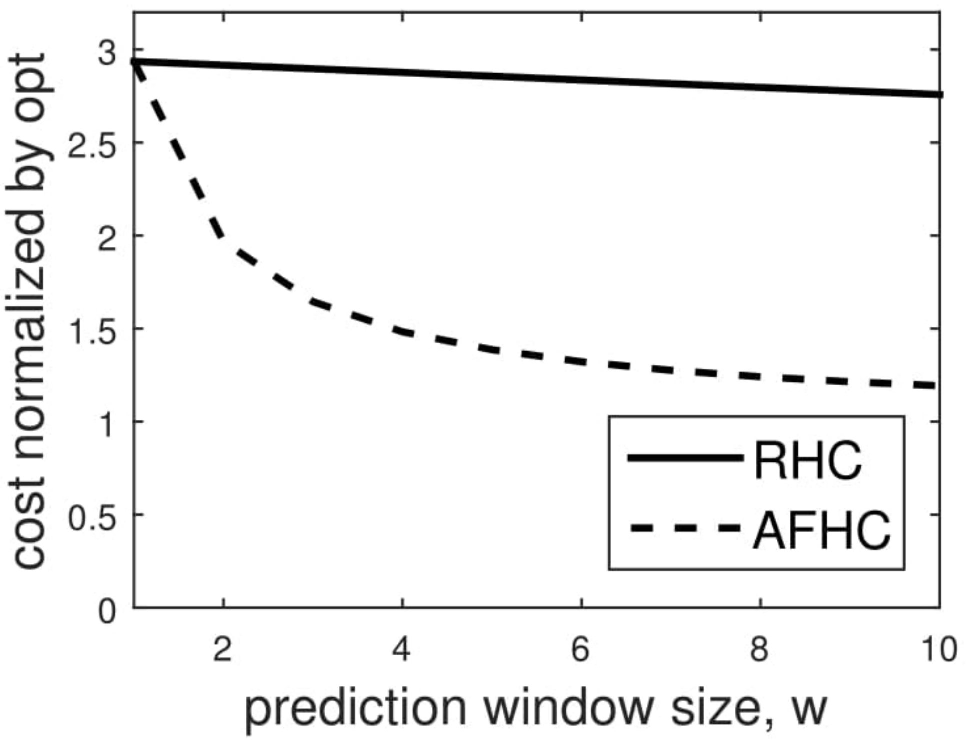
IGCC12: **1+$\Omega(1)$-competitive** Receding Horizon Control

IGCC12: **1+$O(1/w)$-competitive** Averaging Fixed Horizon Control

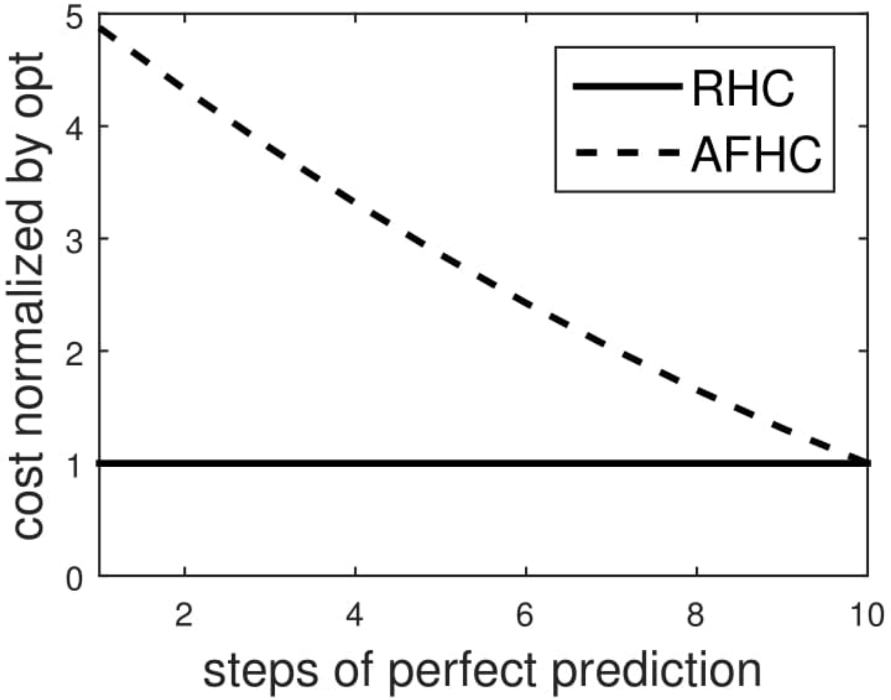


Competitive Ratio: $\sup_{\{y_t, y_{t|t}, \ldots, y_{T|t}\}_{t=1}^{T}} \text{cost}(\text{ALG})/\text{cost}(\text{OPT})$

RHC vs AFHC



3
2.5
2
1.5
1
0.5
0
2
4
6
8
10
cost normalized by opt
prediction window size, w
RHC
AFHC



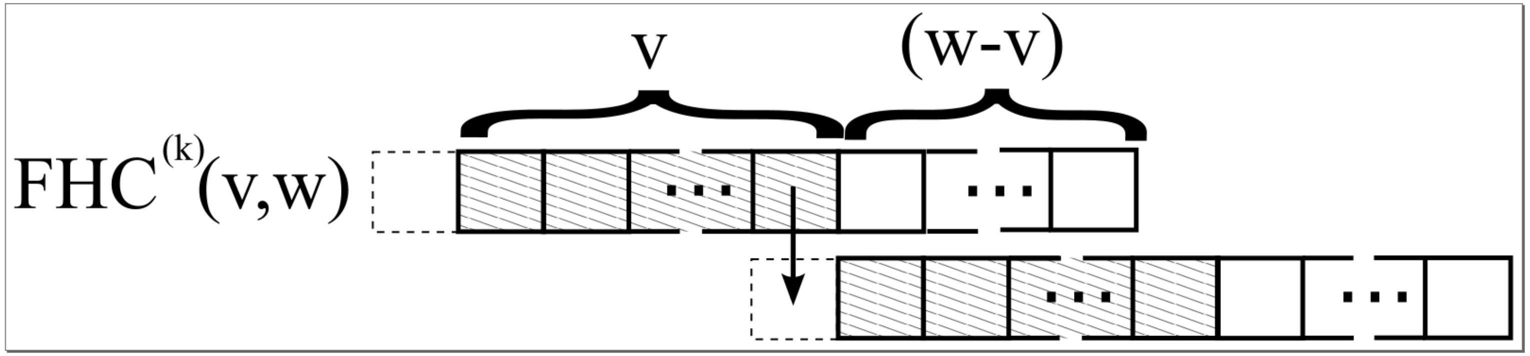
5
4
3
2
1
0
2
4
6
8
10
cost normalized by opt
steps of perfect prediction
RHC
AFHC

FHC with limited commitment v

Level of Commitment, v

Use v of the calculated actions before using new predictions.

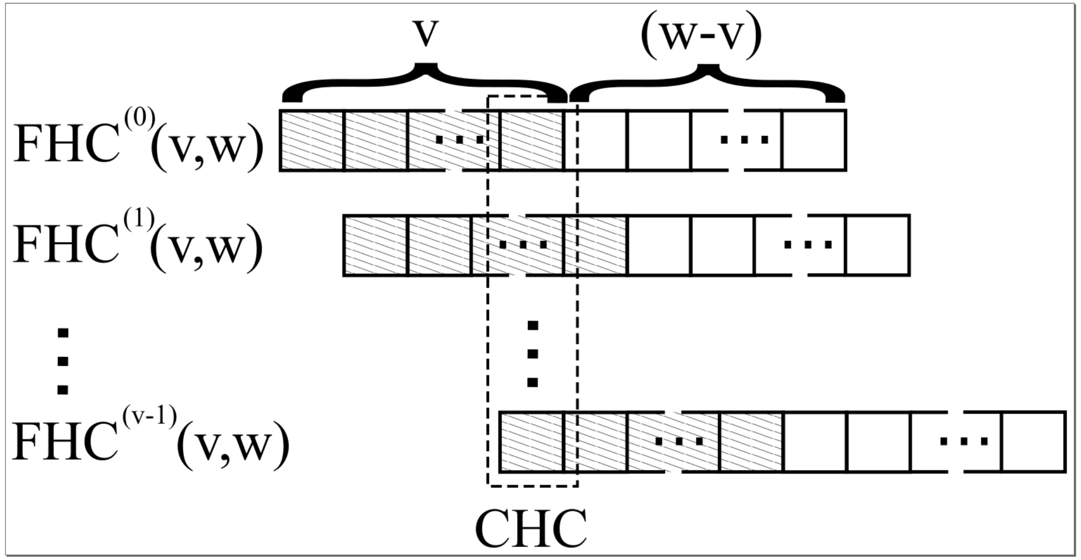
1. **Every $v \le w$** rounds, receive the predictions $y_{t|t}, \ldots, y_{t|t+w-1}$
2. Solve $\min_{x_t, \ldots, x_{t+w-1}} \sum_{\tau=t}^{t+w-1} [c(x_\tau, y_{\tau|t}) + \|x_\tau - x_{\tau-1}\|]$
3. Implement $x_t, \ldots, x$



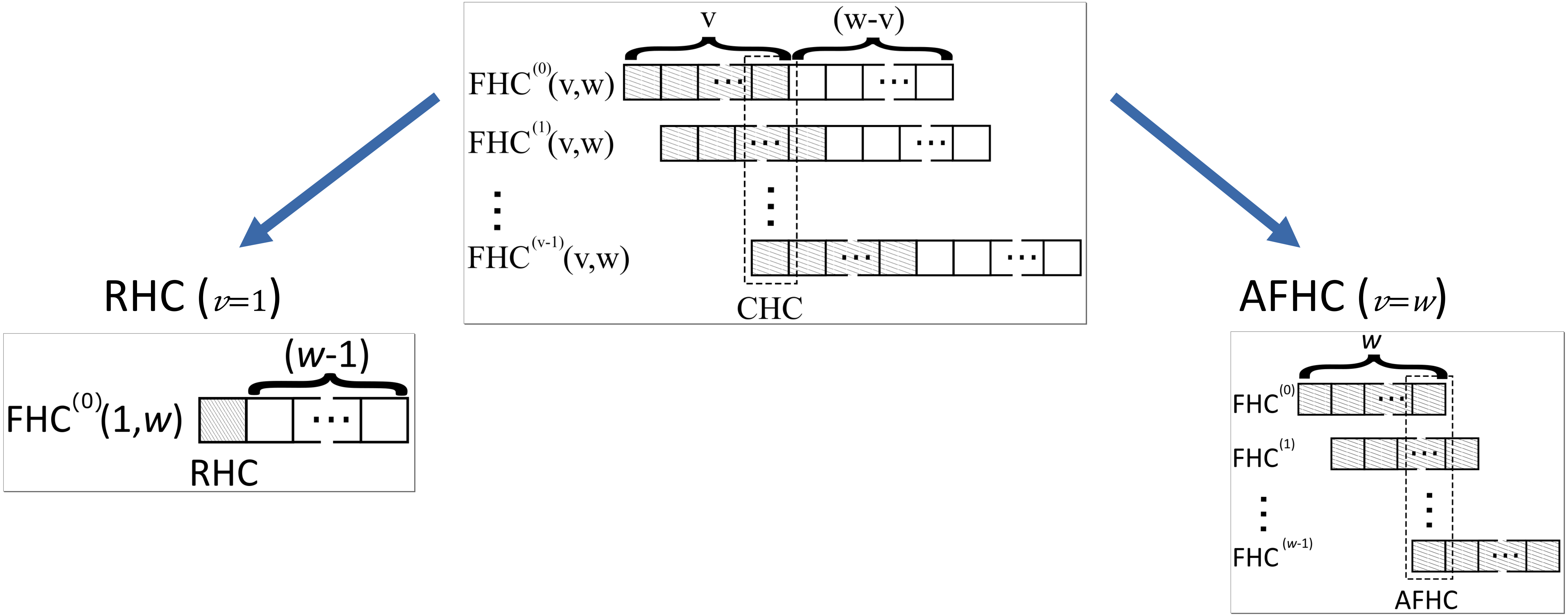
Committed Horizon Control

Average the decisions between a set of different v FHC algorithms.

1. Run v FHC algorithms with limited commitment v, each starting at a different round.
2. Use $\text{FHC}^{(k)}$ to determine $x_t^{(k)}, \ldots, x_{t+v-1}^{(k)}$
3. Implement $x_t = 1/v \sum_{k=0}^{v-1} x_t$



CHC generalizes RHC and AFHC (Sigmetrics16)



Average-case Analysis

Theorem: Let $\|f_k\|$ be a measure of the prediction error for k steps into the future, then

$$E[\text{cost}(CHC) - \text{cost}(OPT)] \leq \frac{2T}{v}\left(D + G\sum_{k=0}^{v-1} \|f_k\|^{\alpha}\right)$$

Optimal $\boldsymbol{v}$ depends on how $\|\boldsymbol{f_k}\|$ grows with $\boldsymbol{k}$

Different properties give different optimal v

Illustration of Theorem

α-Hölder continuity

$$|c(x, y_1) - c(x, y_2)| \le G \|y_1 - y_2\|_2^{\alpha}$$

White noise variance

$$\mathrm{trace}(\mathrm{Cov}(e)) = \sigma^2$$

Range limiting correlation error

$$\|f(s)\|_F = c, \quad L \ge s > 0$$
$$\|f(s)\|_F = 0, \quad s > L$$

$v^* = 1$
(RHC)