Controller Architectures: Tradeoffs between Performance and Complexity

Mihailo Jovanović

ee.usc.edu/mihailo/

Makan
Syracuse

Fu
UTRC

Neil
U of M

Xiaofan
Siemens

Sepideh
USC

Florian
ETH Zürich

NREL Autonomous Energy Grids Workshop

Inter-area oscillations in power systems

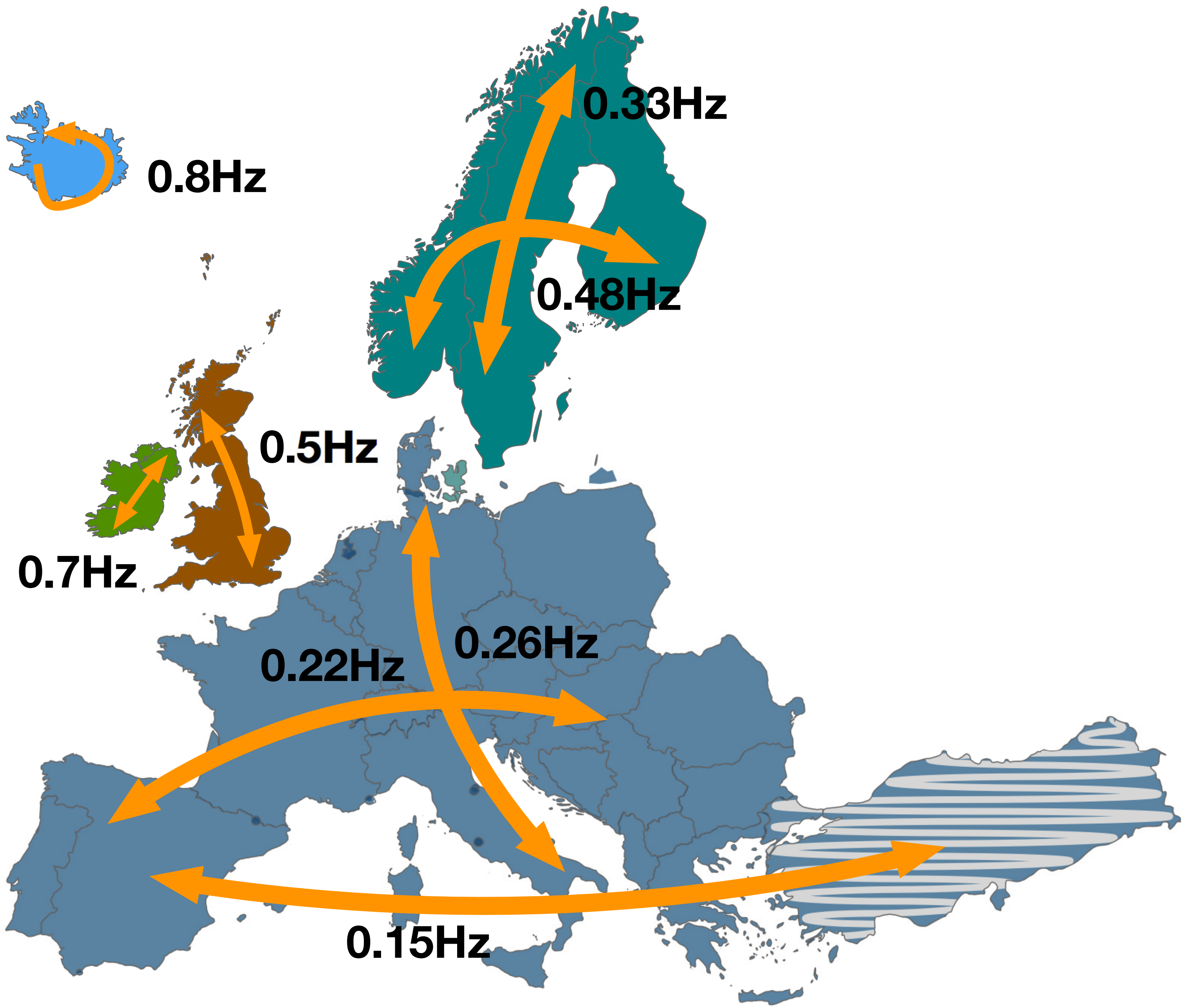
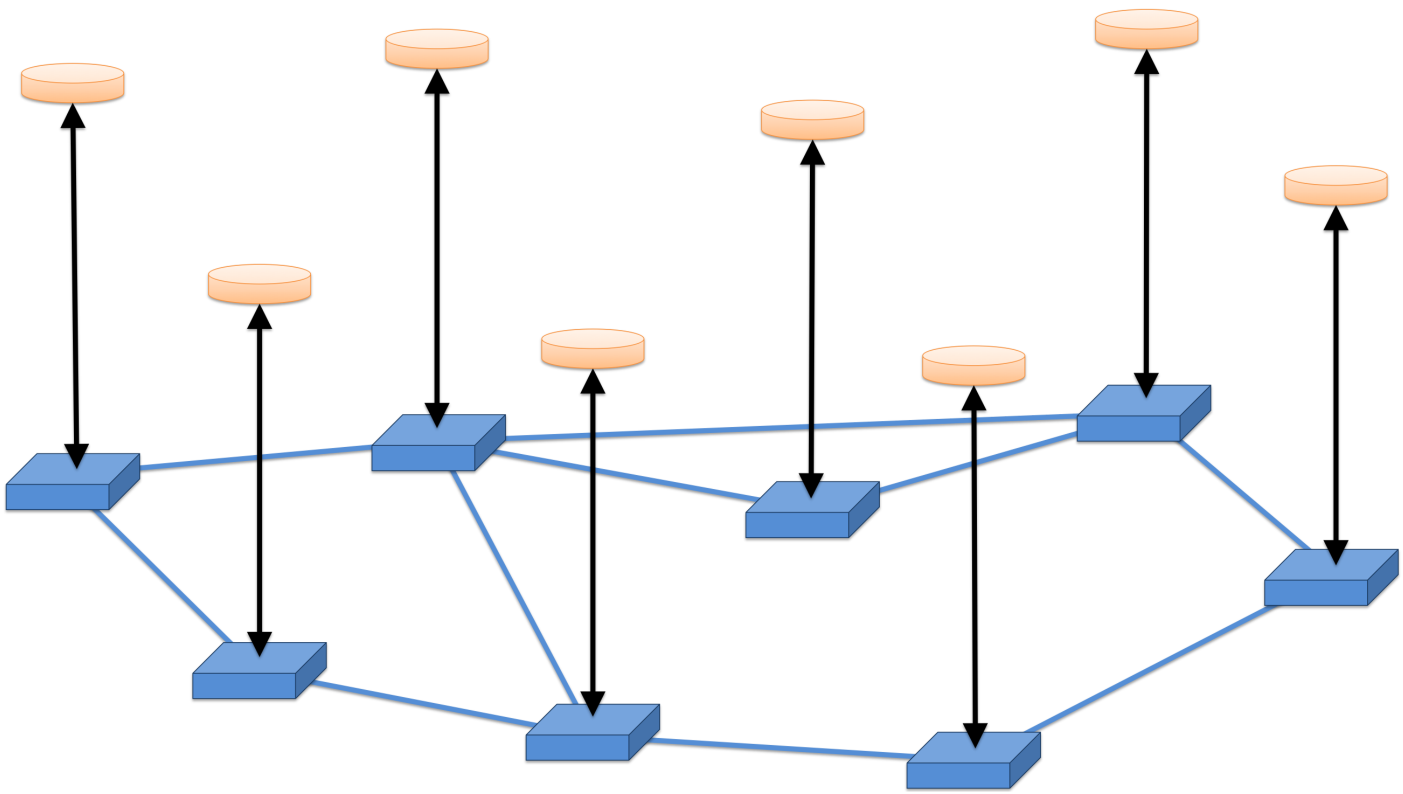


Image credit: Florian Dörfler

Conventional control of generators

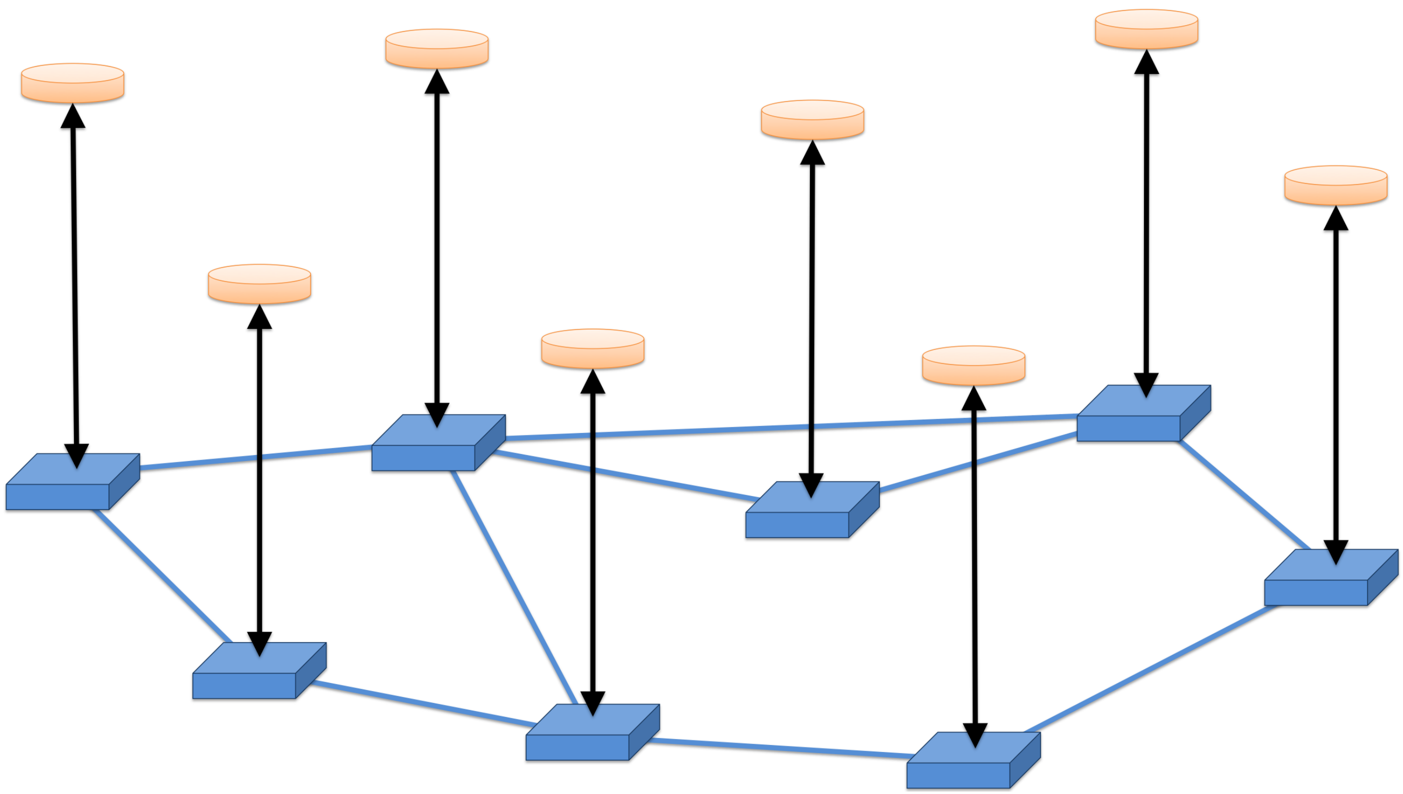
fully decentralized controller



network of generators

Conventional control of generators

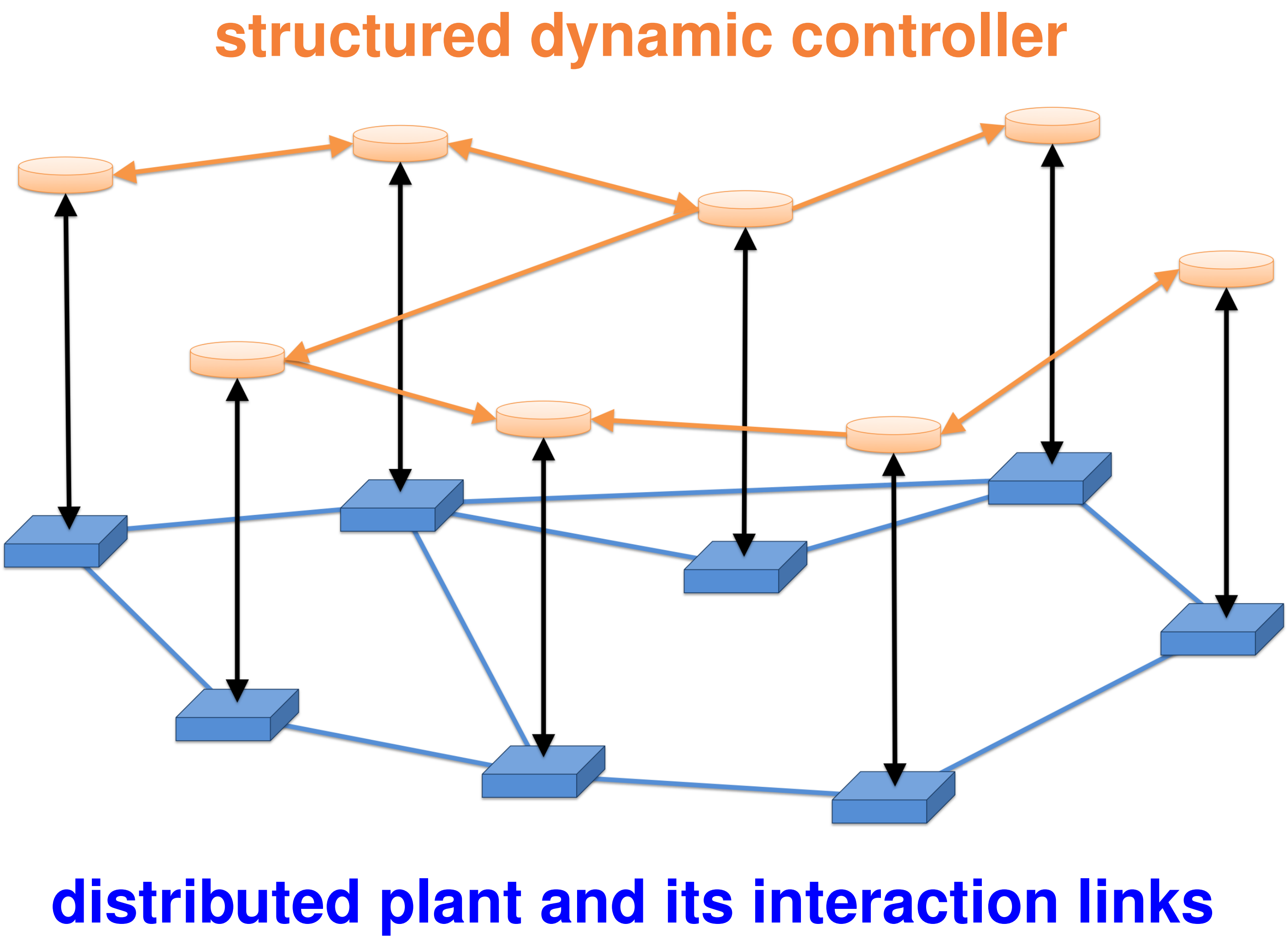
fully decentralized controller



network of generators

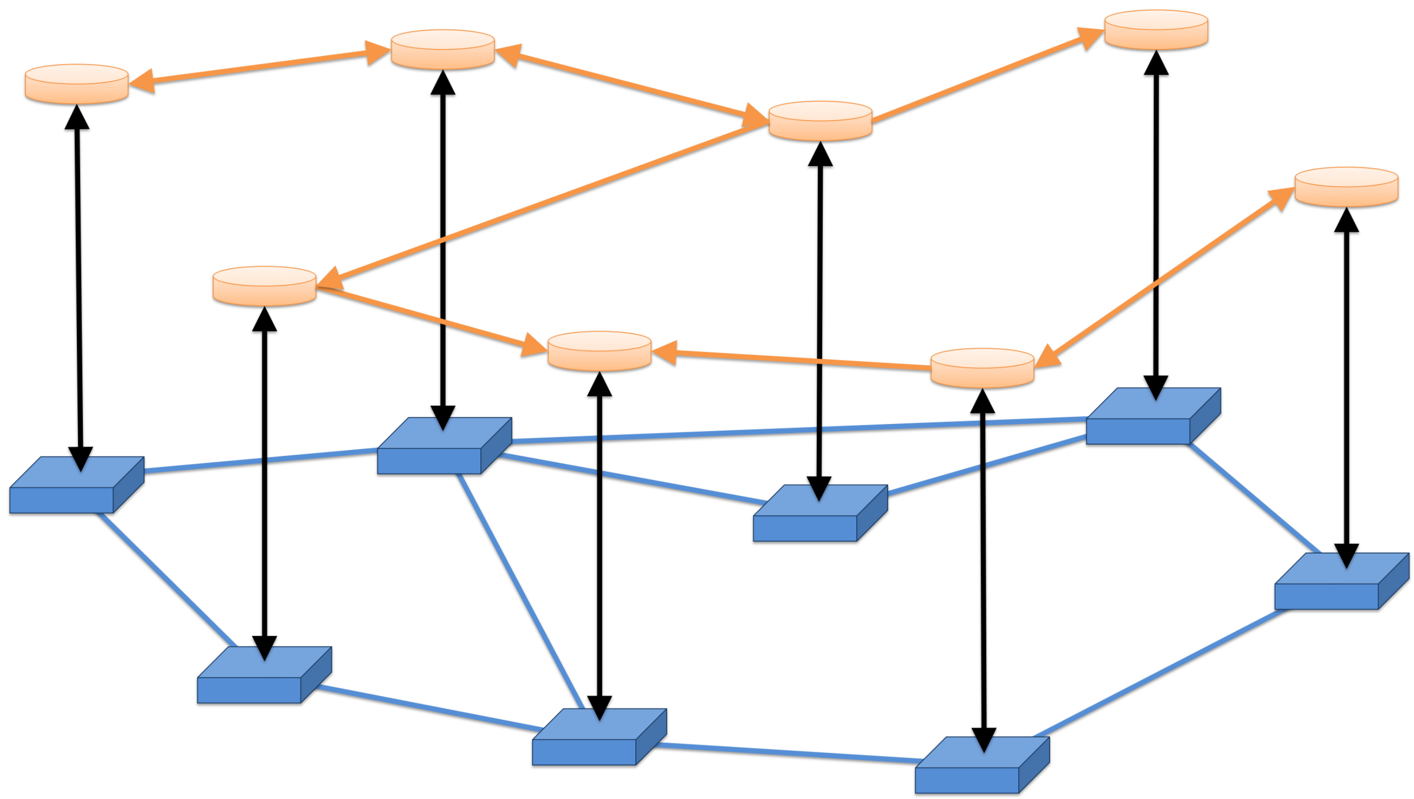
- CONVENTIONAL CONTROL
  - ⋆ local oscillations ✔
  - ⋆ inter-area oscillations ✘

Possible alternative



Possible alternative

structured dynamic controller



distributed plant and its interaction links

CHALLENGE

design of controller architectures

performance vs **complexity**

Complexity via Regularization

$$\text{minimize} \quad J(K) \quad + \quad \gamma\, g(K)$$

$$\downarrow \qquad\qquad \downarrow$$

closed-loop performance — **controller complexity**

$\gamma > 0$ – performance vs complexity tradeoff

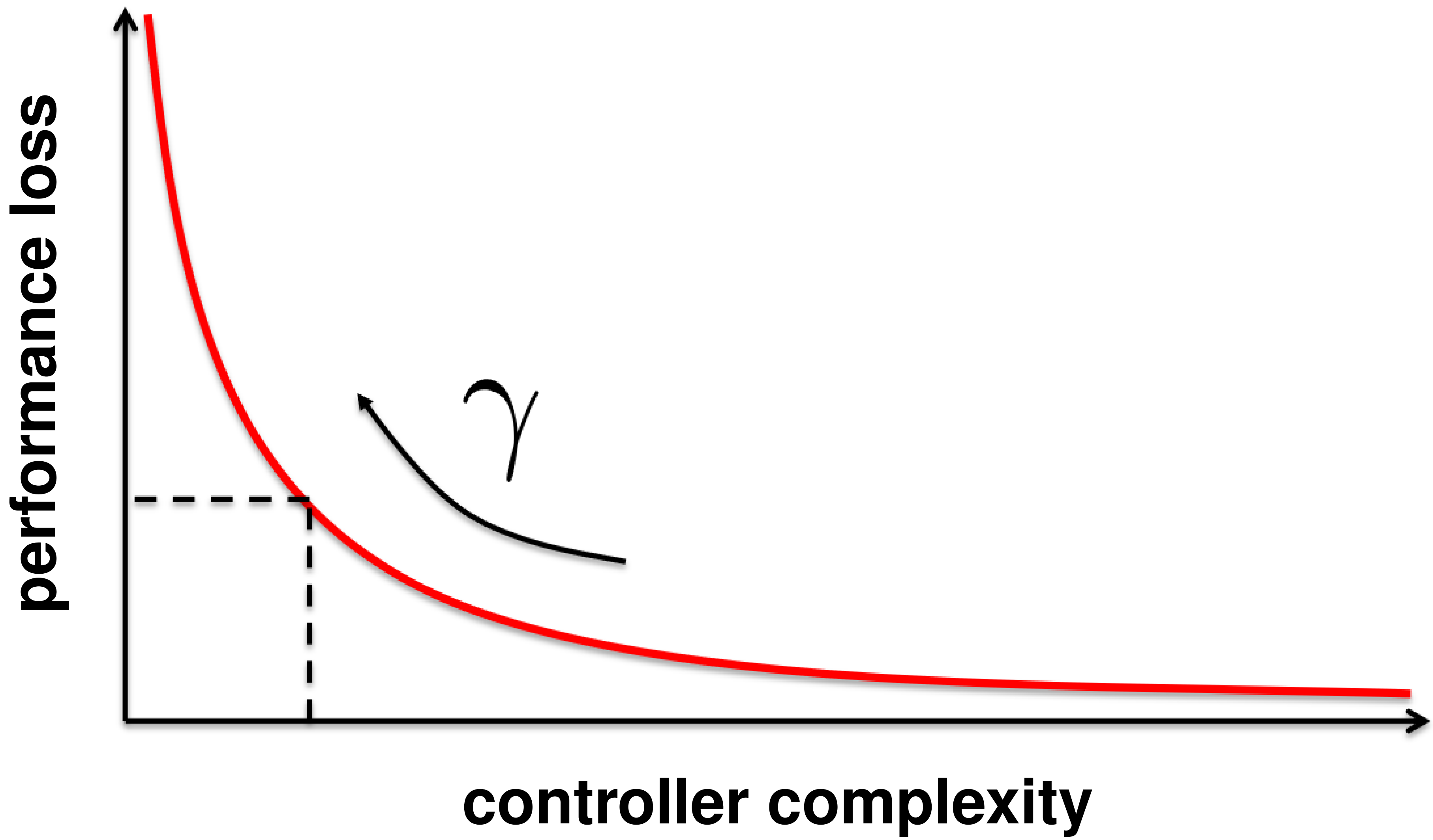
Fardad, Lin, Jovanović, ACC ’11

Lin, Fardad, Jovanović, IEEE TAC ’13

Matni & Chandrasekaran, IEEE TAC ’16

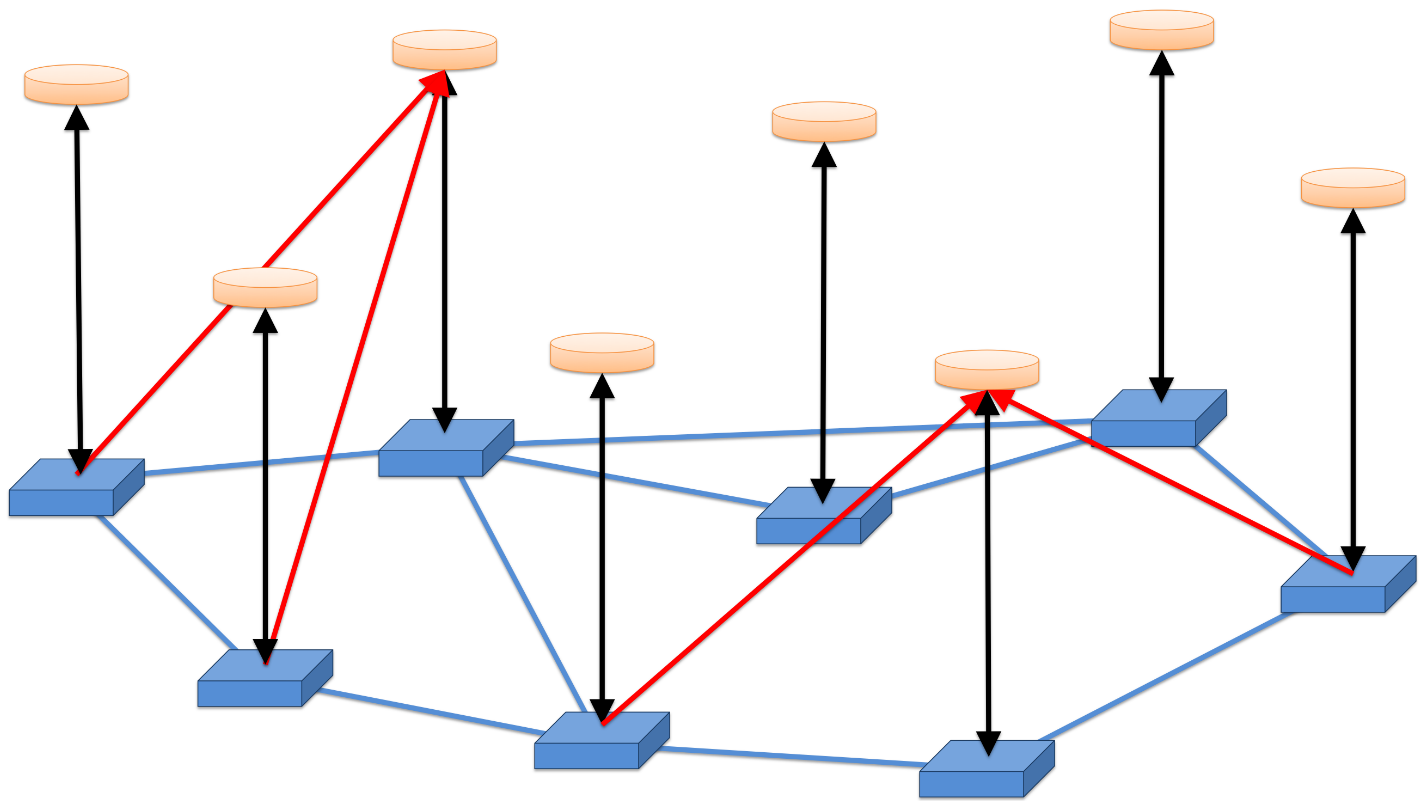
- TRADE-OFF CURVE

  ⋆ **performance vs complexity**



This talk

structured memoryless controller



distributed plant and its interaction links

OBJECTIVE

identification of a signal exchange network

performance vs **sparsity**

Control problem

Lyapunov equation

discrete-time dynamics: $x_{t+1} \;=\; A\,x_t \;+\; B\,d_t$

white-in-time forcing: $\mathbf{E}\left(d_t\,d_\tau^T\right) \;=\; W\,\delta_{t\,-\,\tau}$

- LYAPUNOV EQUATION

$$
\begin{aligned}
X_{t\,+\,1} \;&:=\; \mathbf{E}\left(x_{t\,+\,1}\,x_{t\,+\,1}^T\right) \\
&=\; \mathbf{E}\left(\left(A\,x_t \;+\; B\,d_t\right)\left(x_t^T A^T \;+\; d_t^T B^T\right)\right) \\
&=\; A\,\mathbf{E}\left(x_t\,x_t^T\right)A^T \;+\; B\,\mathbf{E}\left(d_t\,d_t^T\right)B^T \\
&=\; A\,X_t\,A^T \;+\; B\,W B^T
\end{aligned}
$$

⋆ continuous-time version

$$
\frac{\mathrm{d}\,X_t}{\mathrm{d}\,t} \;=\; A\,X_t \;+\; X_t\,A^T \;+\; B\,W B^T
$$

Minimum variance state-feedback problem

dynamics: $\dot{x} \;=\; A\,x \;+\; B_1\,d \;+\; B_2\,u$

objective function: $J \;=\; \lim_{t\,\to\,\infty} \mathbf{E}\left(x^T(t)\,Q\,x(t) \;+\; u^T(t)\,R\,u(t)\right)$

memoryless controller: $u \;=\; -F\,x$

Minimum variance state-feedback problem

dynamics: $\dot{x} \;=\; A\,x \;+\; B_1\,d \;+\; B_2\,u$

objective function: $J \;=\; \lim_{t \,\to\, \infty} \mathbf{E}\left(x^T(t)\,Q\,x(t) \;+\; u^T(t)\,R\,u(t)\right)$

memoryless controller: $u \;=\; -F\,x$

- CLOSED-LOOP VARIANCE

J – **non-convex function of** F

No structural constraints

- SDP CHARACTERIZATION

$$\begin{array}{ll} \underset{X,\,F}{\text{minimize}} & \text{trace}\left((Q + F^T R\, F)\, X\right) \\ \text{subject to} & (A - B_2 F)\, X + X\,(A - B_2 F)^T + B_1 B_1^T = 0 \\ & X \succ 0 \end{array}$$

No structural constraints

- SDP CHARACTERIZATION

$$\begin{array}{ll} \underset{X,\,F}{\text{minimize}} & \text{trace}\left((Q \,+\, F^T R\, F)\, X\right) \\ \text{subject to} & (A \,-\, B_2 F)\, X \,+\, X\, (A \,-\, B_2 F)^T \,+\, B_1 B_1^T \,=\, 0 \\ & X \,\succ\, 0 \end{array}$$

⋆ **change of variables:** $FX \,=\, Y$

$$\begin{array}{ll} \underset{X,\,Y}{\text{minimize}} & \text{trace}\,(Q\, X) \,+\, \text{trace}\left(R\, Y\, X^{-1}\, Y^T\right) \\ \text{subject to} & (A\, X \,-\, B_2\, Y) \,+\, (A\, X \,-\, B_2\, Y)^T \,+\, B_1 B_1^T \,=\, 0 \\ & X \,\succ\, 0 \end{array}$$

Schur complement $\Rightarrow$ SDP characterization

- RICCATI-BASED-CHARACTERIZATION

globally optimal controller

$$A^T P + P A - P B_2 R^{-1} B_2^T P + Q = 0$$

$$F_c = R^{-1} B_2^T P$$

- STRUCTURAL CONSTRAINTS $F \in \mathcal{S}$

centralized **fully-decentralized** **localized**

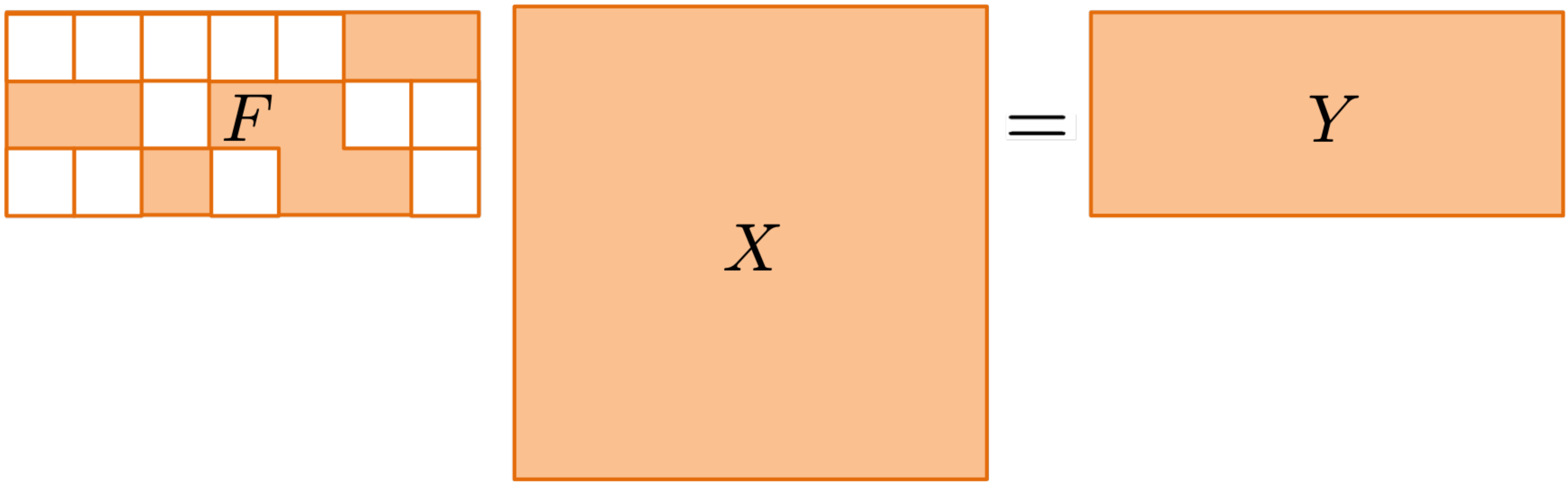
$$\begin{bmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix} \qquad \begin{bmatrix} * & & & \\ & * & & \\ & & * & \\ & & & * \end{bmatrix} \qquad \begin{bmatrix} * & * & & \\ * & * & * & \\ & * & * & * \\ & & * & * \end{bmatrix}$$

GRAND CHALLENGE

convex characterization in the face of structural constraints

difficult to establish relation between

$$\left\{ \begin{array}{c} \textbf{structural constraints} \\ \text{on } F \end{array} \right\} \text{ and } \left\{ \begin{array}{c} \textbf{structural constraints} \\ \text{on } X \text{ and } Y \end{array} \right\}$$



Classes of convex problems

- **Partially-nested systems**

  Ho & Chu, IEEE TAC '72

  Voulgaris, ACC '00; ACC '01

- **Cone- and funnel-causal systems**

  Voulgaris, Bianchini, Bamieh, SCL '03

  Bamieh & Voulgaris, SCL '05

  Fardad & Jovanović, Automatica '11

- **Quadratically-invariant systems**

  Rotkowitz & Lall, IEEE TAC '06

- **Poset-causal systems**

  Shah & Parrilo, IEEE TAC '13

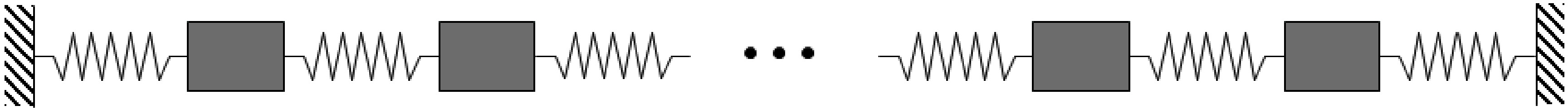
- **Positive systems**

  Tanaka & Langbort, IEEE TAC '11

  Colaneri, Middleton, Chen, Caporale, Blanchini, Automatica '14

  Rantzer, EJC '15; IEEE TAC '16

An example



$$u(t) \;=\; -\begin{bmatrix} F_p & F_v \end{bmatrix} \begin{bmatrix} p(t) \\ v(t) \end{bmatrix}$$

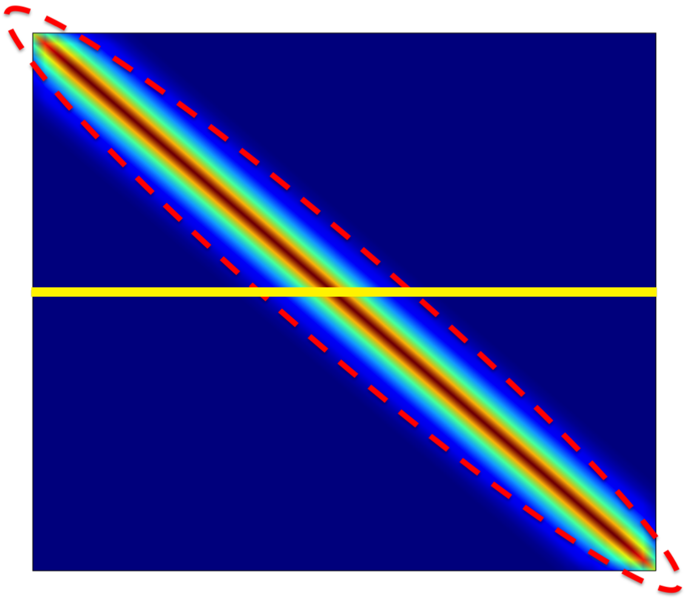
- OBJECTIVE
  - ⋆ minimize steady-state variance of p, v, u

optimal controller – Linear Quadratic Regulator

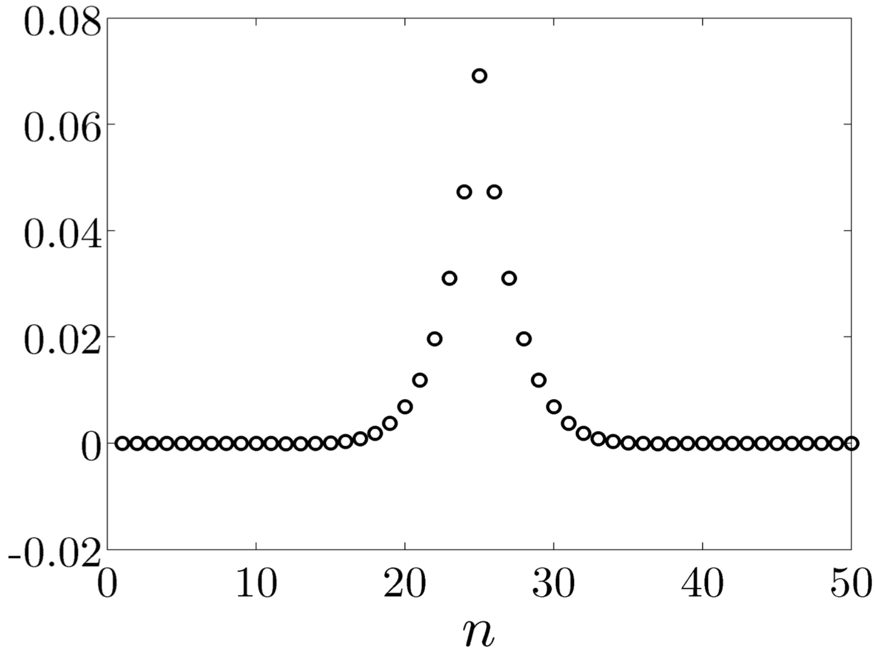
$$\begin{bmatrix} u_1(t) \\ u_2(t) \\ u_3(t) \\ u_4(t) \end{bmatrix} = -\underbrace{\begin{bmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix}}_{F_p} \begin{bmatrix} p_1(t) \\ p_2(t) \\ p_3(t) \\ p_4(t) \end{bmatrix} - \underbrace{\begin{bmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix}}_{F_v} \begin{bmatrix} v_1(t) \\ v_2(t) \\ v_3(t) \\ v_4(t) \end{bmatrix}$$

Structure of optimal controller

position feedback matrix



gains for middle mass



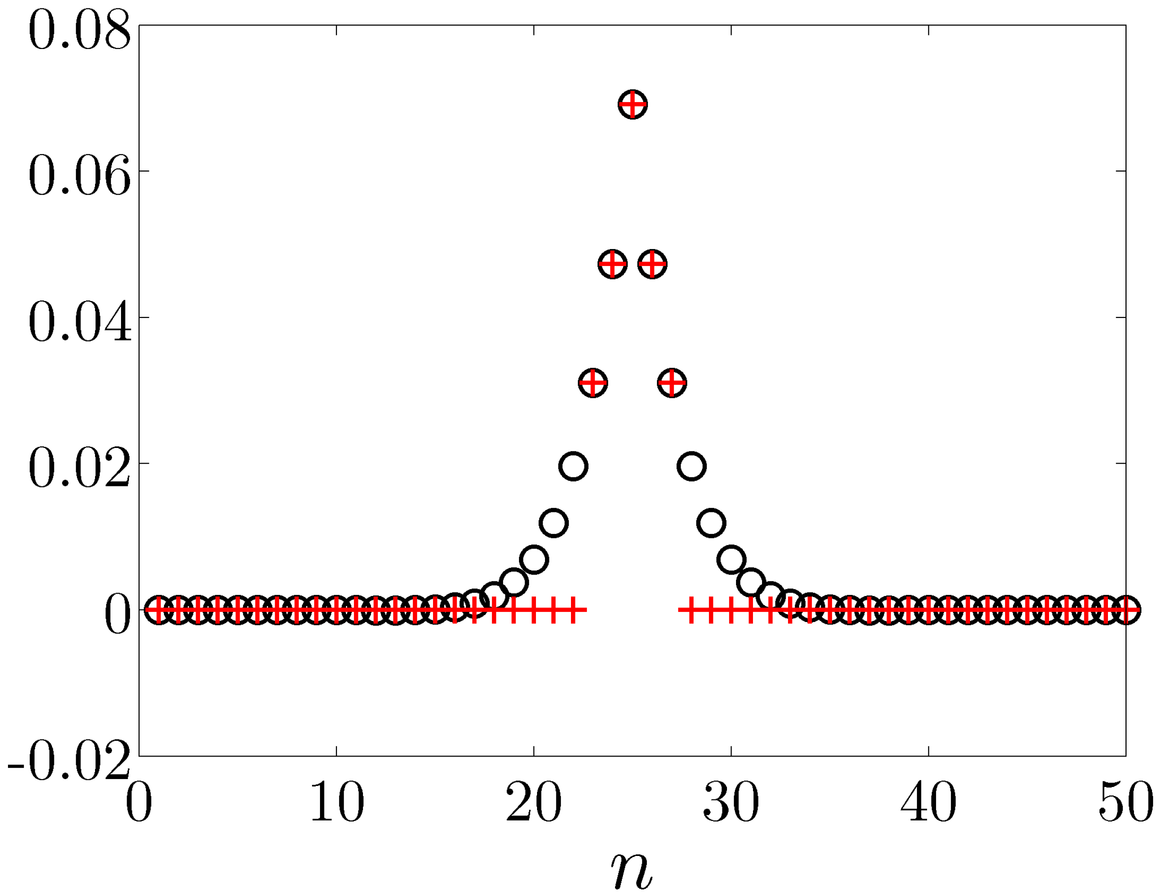
- OBSERVATIONS
  - ⋆ diagonals almost constant (modulo edges)
  - ⋆ off-diagonal decay of a centralized gain

Bamieh, Paganini, Dahleh, IEEE TAC '02

Motee & Jadbabaie, IEEE TAC '08

Enforcing sparsity?

- **One approach: truncate centralized controller**



- DANGERS
  - ⋆ significant performance degradation
  - ⋆ instability

Rest of the talk

- SPARSITY-PROMOTING OPTIMAL CONTROL
  - ⋆ identification and design of sparse feedback gains
- ALGORITHM
  - ⋆ Proximal Augmented Lagrangian Method
- CLASSES OF CONVEX PROBLEMS
  - ⋆ optimal actuator/sensor selection
  - ⋆ optimal design of consensus networks
  - ⋆ diagonal modifications of positive systems
- EXAMPLES
- SUMMARY AND OUTLOOK

Sparsity-promoting optimal control

- OBJECTIVE

  ⋆ **promote sparsity of F**

$$
\begin{bmatrix} u_1(t) \\ u_2(t) \\ u_3(t) \\ u_4(t) \\ u_5(t) \end{bmatrix}
= - \underbrace{\begin{bmatrix} * & * & & & \\ * & * & & & * \\ & * & * & * & \\ & & * & * & \\ & * & & * & * \end{bmatrix}}_{F}
\begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \\ x_5(t) \end{bmatrix}
$$

Sparsity-promoting optimal control

$$\text{minimize} \quad J(F) \quad + \quad \gamma \, \mathbf{card}\,(F)$$

$J(F)$ ↓ **variance amplification**

$\gamma \, \mathbf{card}\,(F)$ ↓ **sparsity-promoting penalty function**

$\mathbf{card}\,(F)$ – number of non-zero elements of F

$\gamma > 0$ – performance vs sparsity tradeoff

Fardad, Lin, Jovanović, ACC '11

Lin, Fardad, Jovanović, IEEE TAC '13

Convex relaxations of $\mathbf{card}\,(F)$

$$\ell_1 \textbf{ norm:} \quad \sum_{i,\,j} |F_{ij}|$$

$$\textbf{weighted } \ell_1 \textbf{ norm:} \quad \sum_{i,\,j} w_{ij}\,|F_{ij}|, \qquad w_{ij} \;\geq\; 0$$

- CARDINALITY VS WEIGHTED ℓ_1 NORM

$$\{w_{ij} \;=\; 1/|F_{ij}|, \;\; F_{ij} \;\neq\; 0\} \quad \Rightarrow \quad \mathbf{card}\,(F) \;=\; \sum_{i,\,j} w_{ij}\,|F_{ij}|$$

Convex relaxations of $\mathbf{card}\,(F)$

ℓ_1 **norm:** $$\mathop{}_{i,\,j} \; |F_{ij}|$$

weighted ℓ_1 norm: $$\mathop{}_{i,\,j} \; w_{ij}\,|F_{ij}|, \quad w_{ij} \;\geq\; 0$$

- CARDINALITY VS WEIGHTED ℓ_1 NORM

$$\{w_{ij} \;=\; 1/|F_{ij}|, \;\; F_{ij} \;=\; 0\} \quad \Rightarrow \quad \mathbf{card}\,(F) \;=\; \mathop{}_{i,\,j} \; w_{ij}\,|F_{ij}|$$

RE-WEIGHTED SCHEME

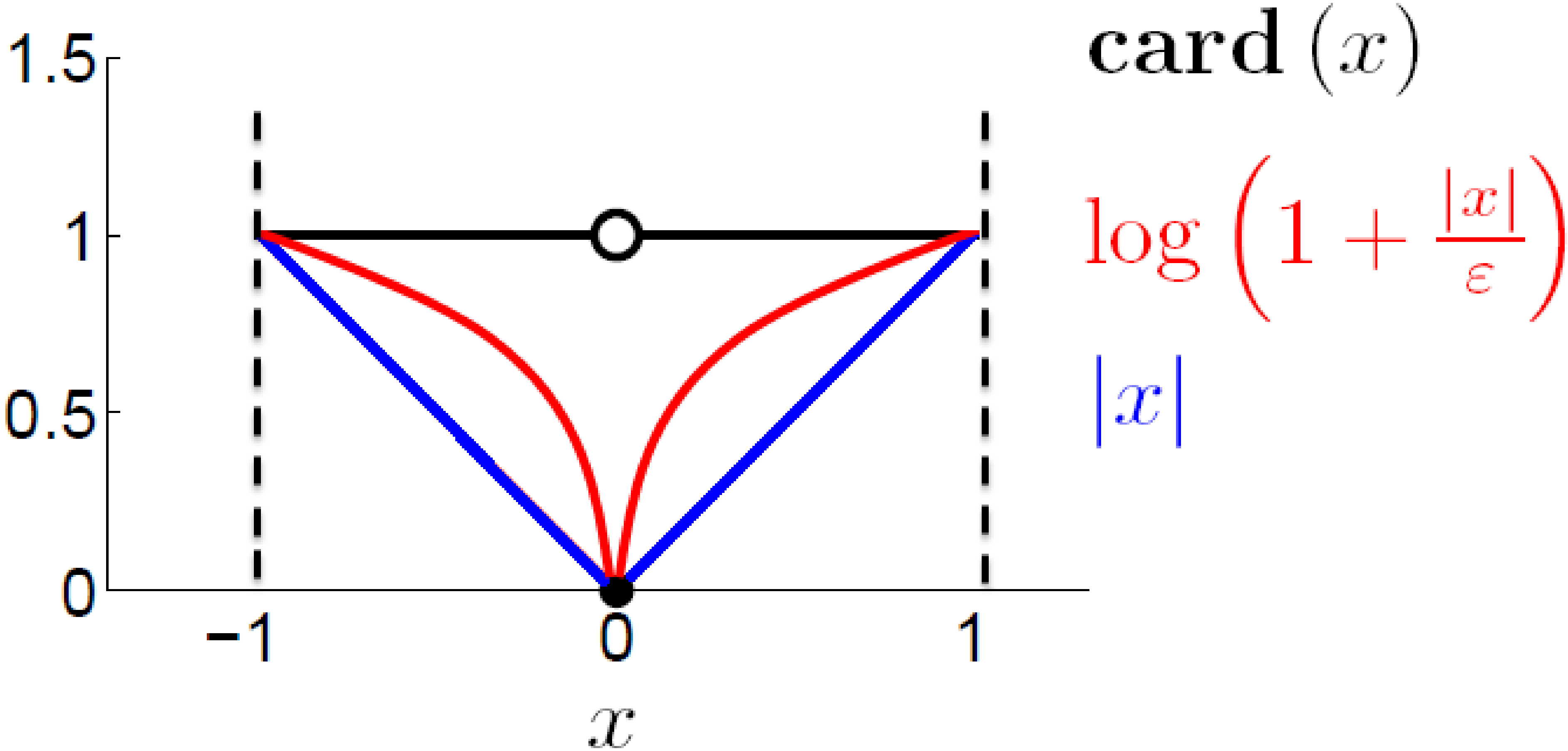
⋆ **use gains from previous iteration to form weights**

$$w_{ij}^{+} \;=\; \frac{1}{|F_{ij}| \,+\, \varepsilon}$$

Candès, Wakin, Boyd, J. Fourier Anal. Appl. '08

A non-convex relaxation of $\mathbf{card}\,(F)$

sum-of-logs: $$\sum_{i,j} \log\left(1 + \frac{|F_{ij}|}{\varepsilon}\right), \qquad 0 < \varepsilon \ll 1$$

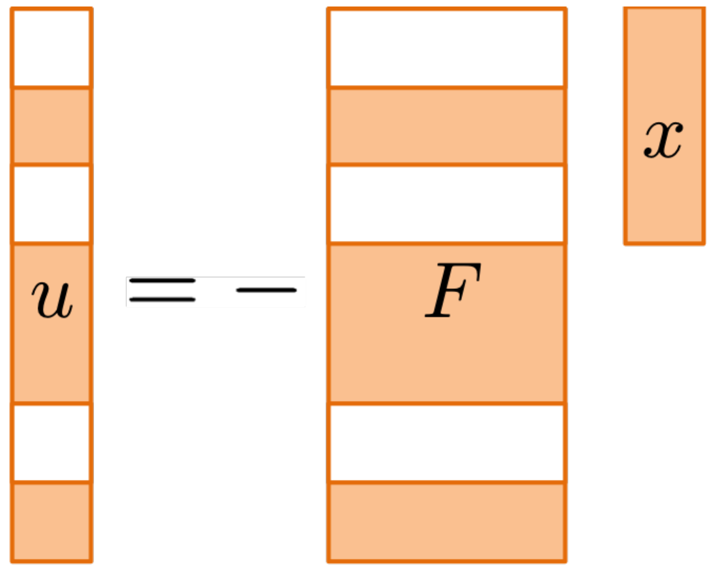


Candès, Wakin, Boyd, J. Fourier Anal. Appl. ’08

CLASSES OF CONVEX PROBLEMS

Optimal actuator/sensor selection

- OBJECTIVE: **identify row-sparse feedback gain**

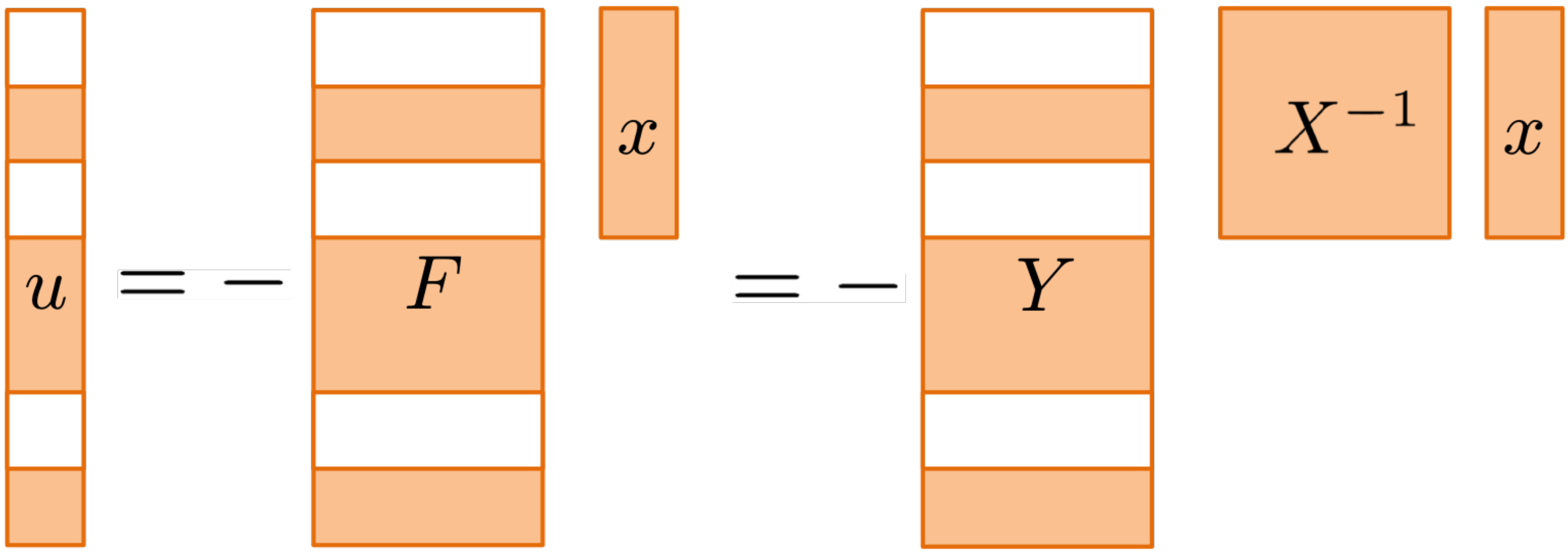


$$\text{minimize} \qquad J(F) \qquad + \qquad \gamma \quad \underset{i}{} \quad \|e_i^T F\|_2$$

$$\downarrow \qquad\qquad\qquad\qquad \downarrow$$

variance amplification — **row-sparsity-promoting penalty function**

- CHANGE OF VARIABLES: $Y := F\,X$
  - ⋆ **convex dependence of J on X and Y**
  - ⋆ **row-sparse structure preserved**



• OPTIMAL ACTUATOR SELECTION

⋆ **admits SDP characterization**

$$\text{minimize} \quad J(X,Y) \quad + \quad \gamma \underset{i}{} \|\mathrm{e}_i^T Y\|_2$$

$$\downarrow \qquad\qquad\qquad\qquad \downarrow$$

variance amplification — **row-sparsity-promoting penalty function**

Polyak, Khlebnikov, Shcherbakov, ECC '13

Münz, Pfister, Wolfrum, IEEE TAC '14

Dhingra, Jovanović, Luo, CDC '14

Design of undirected consensus networks

dynamics: $\dot{x} \; = \; - \, L \, x \, + \, d \, + \, u$

control: $u \; = \; - \, F \, x$

objective: $J \; = \; \lim_{t \, \to \, \infty} \mathbf{E} \left(x^T(t) \, Q \, x(t) \; + \; u^T(t) \, R \, u(t) \right)$

Design of undirected consensus networks

dynamics: $\dot{x} \;=\; -\,L\,x \;+\; d \;+\; u$

control: $u \;=\; -\,F\,x$

objective: $J \;=\; \lim_{t \to \infty} \mathbf{E}\left(x^T(t)\,Q\,x(t) \;+\; u^T(t)\,R\,u(t)\right)$

convex characterization

$$
\begin{array}{ll}
\text{minimize} & \operatorname{trace}(X) \;+\; \gamma\,\mathbb{1}^T\,Y\,\mathbb{1} \\[2ex]
\text{subject to} & \begin{bmatrix} X & \begin{bmatrix} Q^{1/2} \\ -R^{1/2}\,F \end{bmatrix} \\ \begin{bmatrix} Q^{1/2} & -F\,R^{1/2} \end{bmatrix} & F \;+\; L \;+\; \mathbb{1}\mathbb{1}^T/n \end{bmatrix} \;\succeq\; 0 \\[2ex]
& F\,\mathbb{1} \;=\; 0, \quad -\,Y_{ij} \;\le\; W_{ij}\,F_{ij} \;\le\; Y_{ij}
\end{array}
$$

Lin, Fardad, Jovanović, Allerton '12

Zelazo, Schuler, Allgöwer, SCL '13

Hassan-Moghaddam & Jovanović, arXiv:1506.03437

Diagonal modifications of positive systems

$$\dot{x} = \left(A + \underset{k}{} u_k D_k\right) x + d$$

A – Metzler matrix ($A_{ij} \geq 0$, $i = j$)

D_k – diagonal matrices

Diagonal modifications of positive systems

$$\dot{x} \;=\; A \;+\; \underset{k}{} \; u_k \, D_k \;\; x \;+\; d$$

A – Metzler matrix ($A_{ij} \geq 0$, $i = j$)

D_k – diagonal matrices

- EXAMPLES

  ⋆ combination drug therapy

    x_i mutates to x_j at rate A_{ji}

    u_k kills x_i at rate $(D_k)_{ii}$

  ⋆ leader selection in directed networks

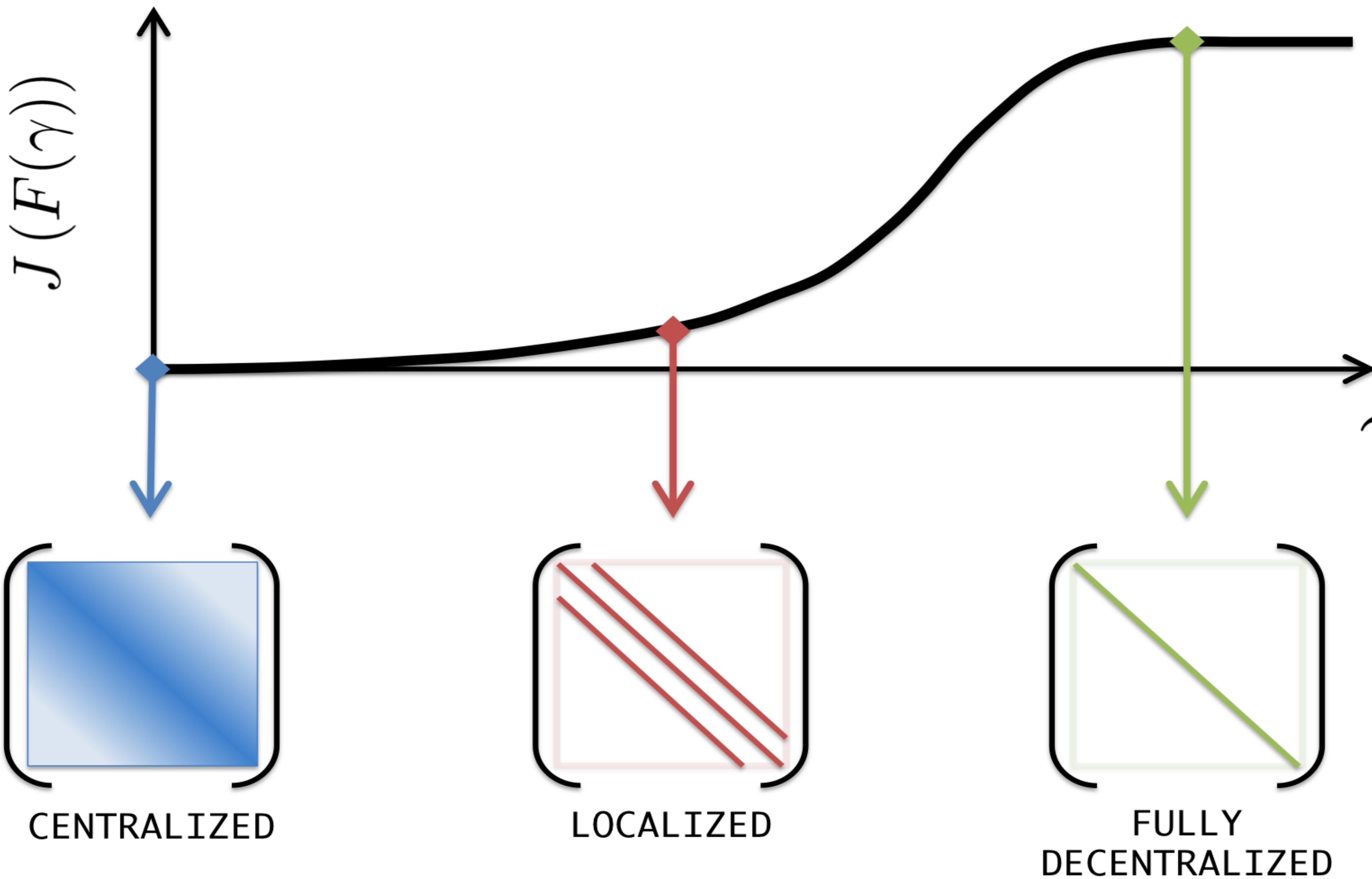
Rantzer & Bernhardsson, CDC ’14

Jonsson, Matni, Murray, CDC ’14

Dhingra, Colombino, Jovanović, ECC ’16

Parameterized family of feedback gains

$$F(\gamma) \;:=\; \underset{F}{\operatorname{argmin}}\,\big(J(F) \;+\; \gamma\, g(F)\big)$$



CASE STUDY: WIDE-AREA CONTROL

Dörfler, Jovanović, Chertkov, Bullo, IEEE TPWRS '14

Wu, Dörfler, Jovanović, IEEE TPWRS '16

http://people.ece.umn.edu/users/mihailo/software/lqrsp/wac.html

Electro-mechanical oscillations in power systems

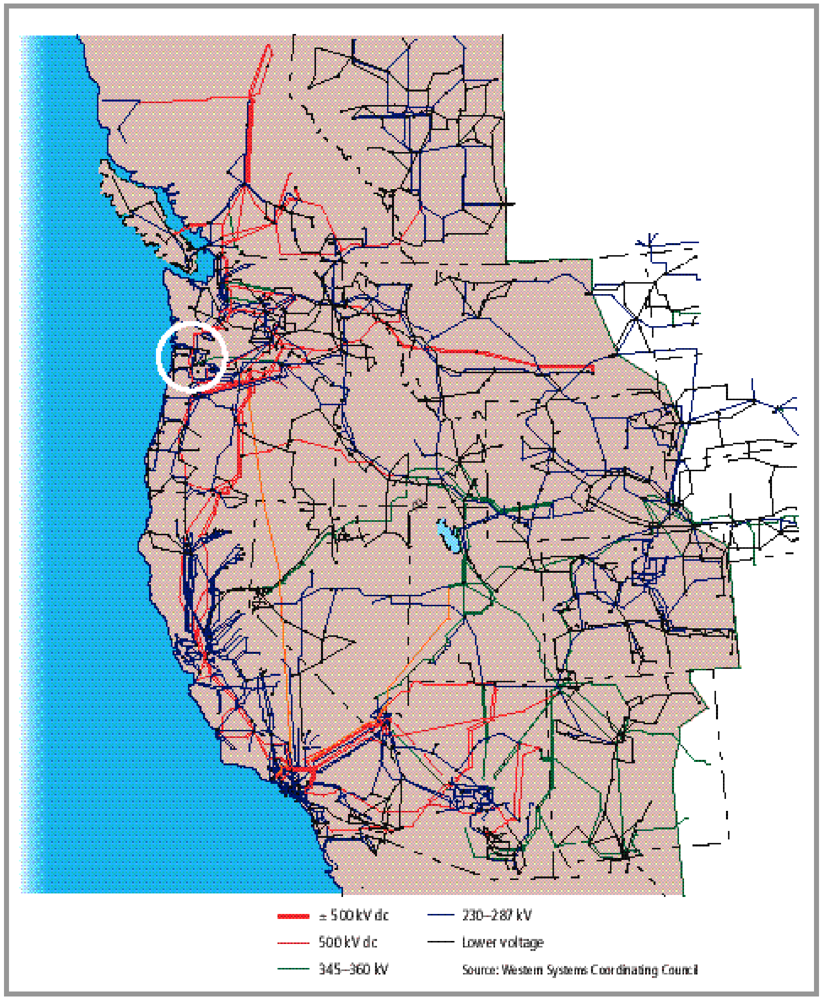
- **Local oscillations**
  - ⋆ single generators swing relative to the rest of the grid
  - ⋆ typically damped by Power System Stabilizers (PSSs)

- **Inter-area oscillations**
  - ⋆ groups of generators oscillate relative to each other
  - ⋆ associated with dynamics of power transfers

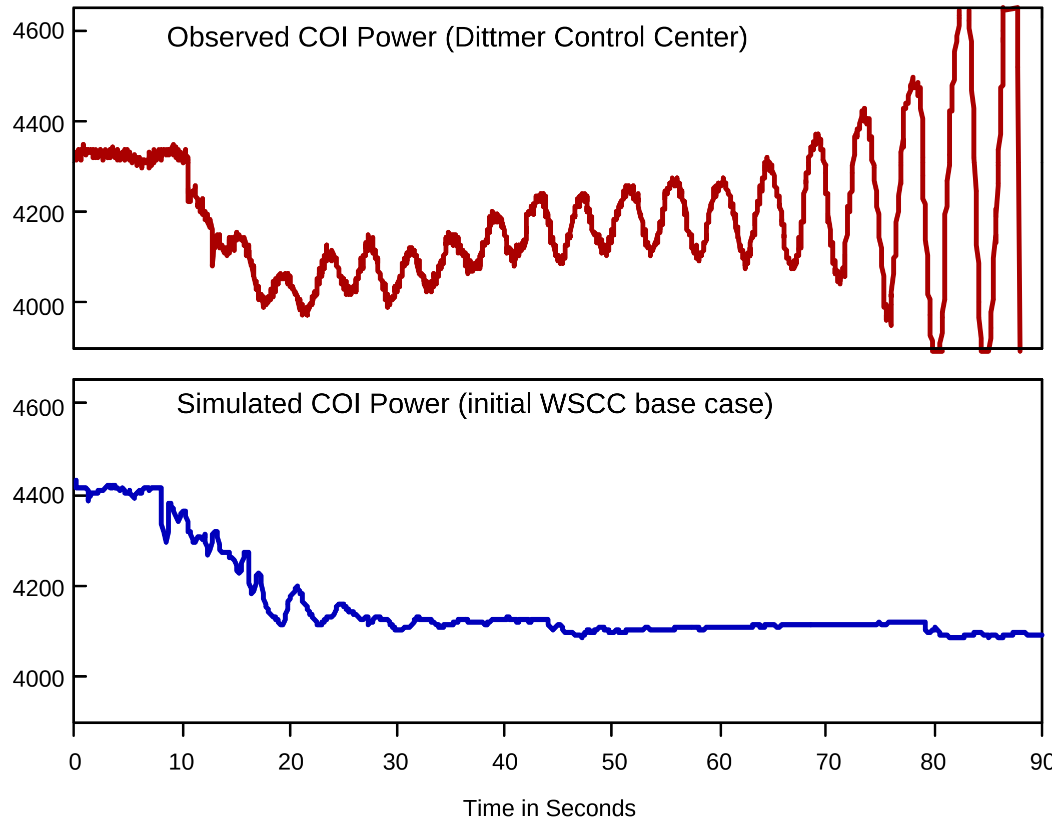
Inter-area oscillations

- **Blackout of Aug. 10, 1996**
  - ⋆ **resulted from instability of the $0.25\,\mathrm{Hz}$ mode**

western interconnected system:



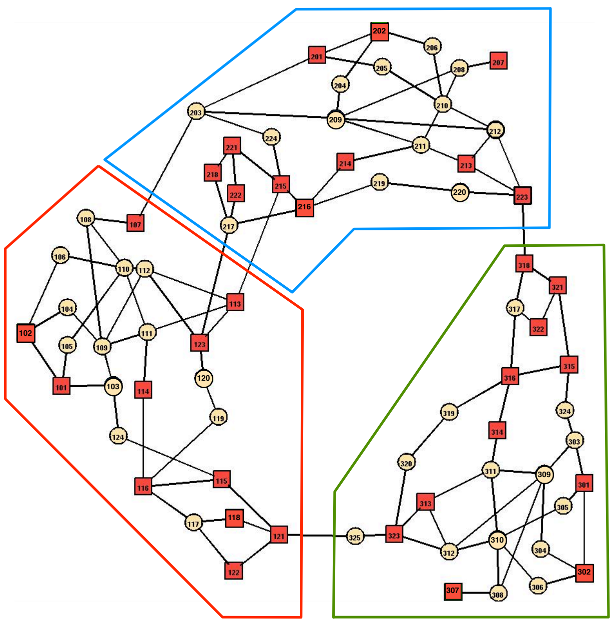
California-Oregon power transfer:



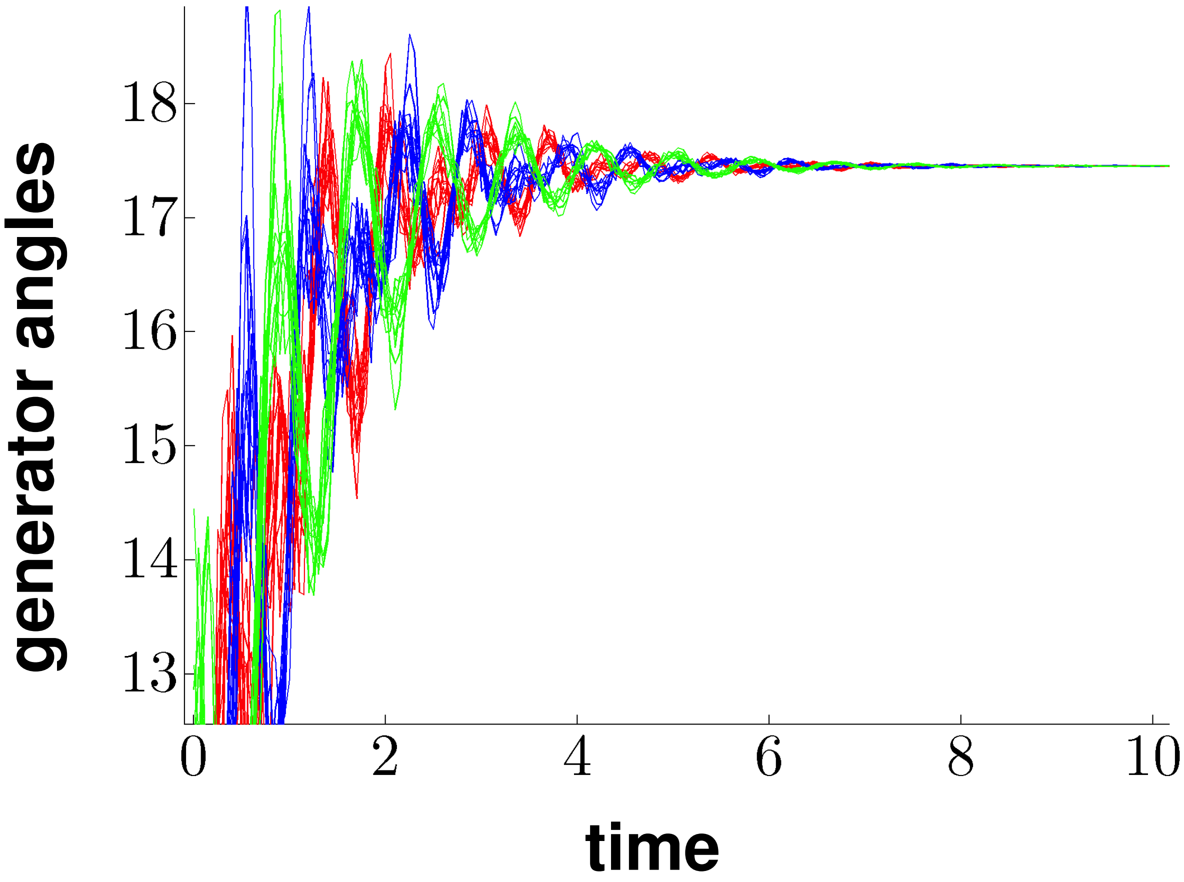
Slow coherency theory

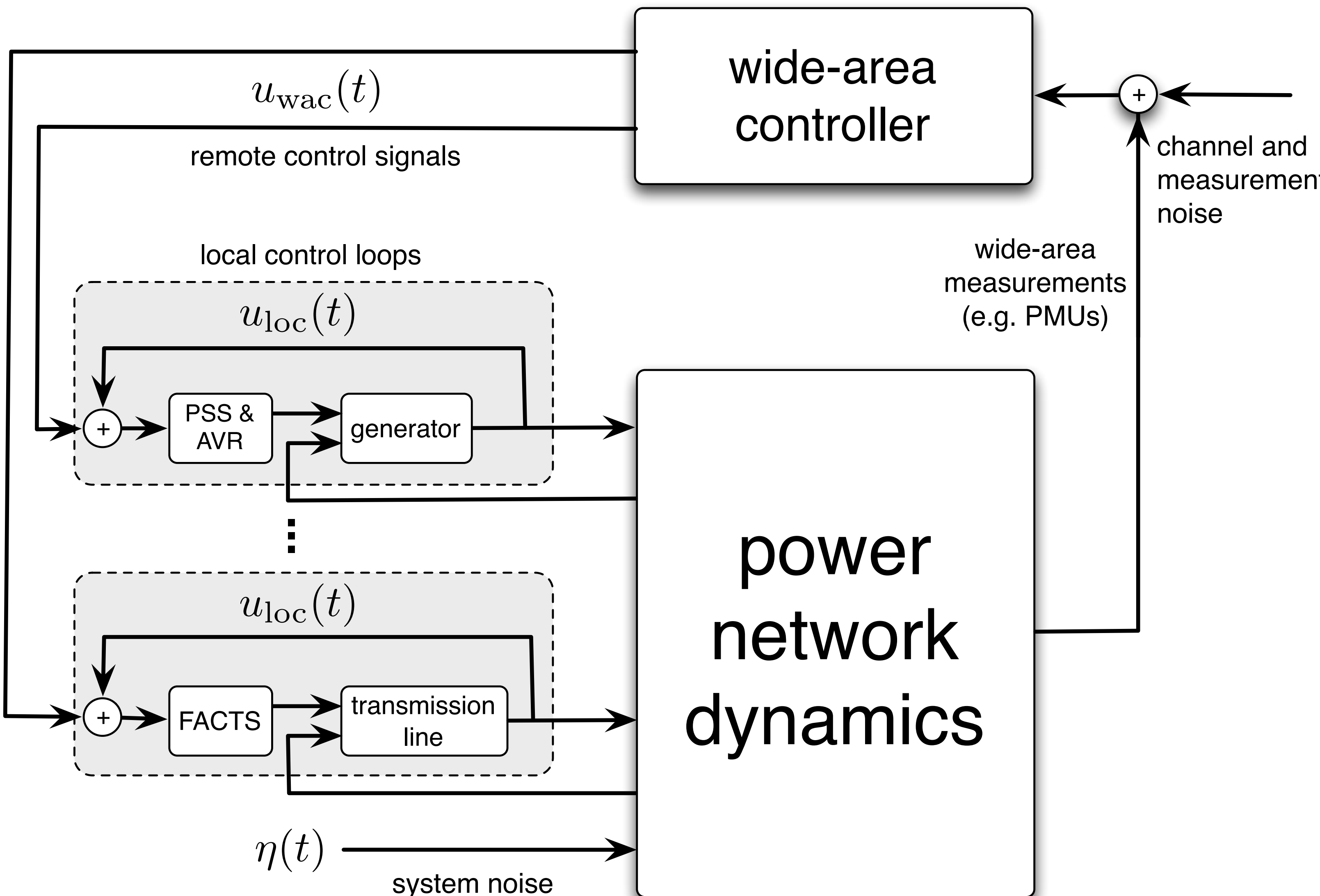
- WHERE ARE THE INTER-AREA MODES COMING FROM?
  - ⋆ **slow coherency theory** *Chow, Kokotović, et al. ’78, ’82*

RTS 96 power system:



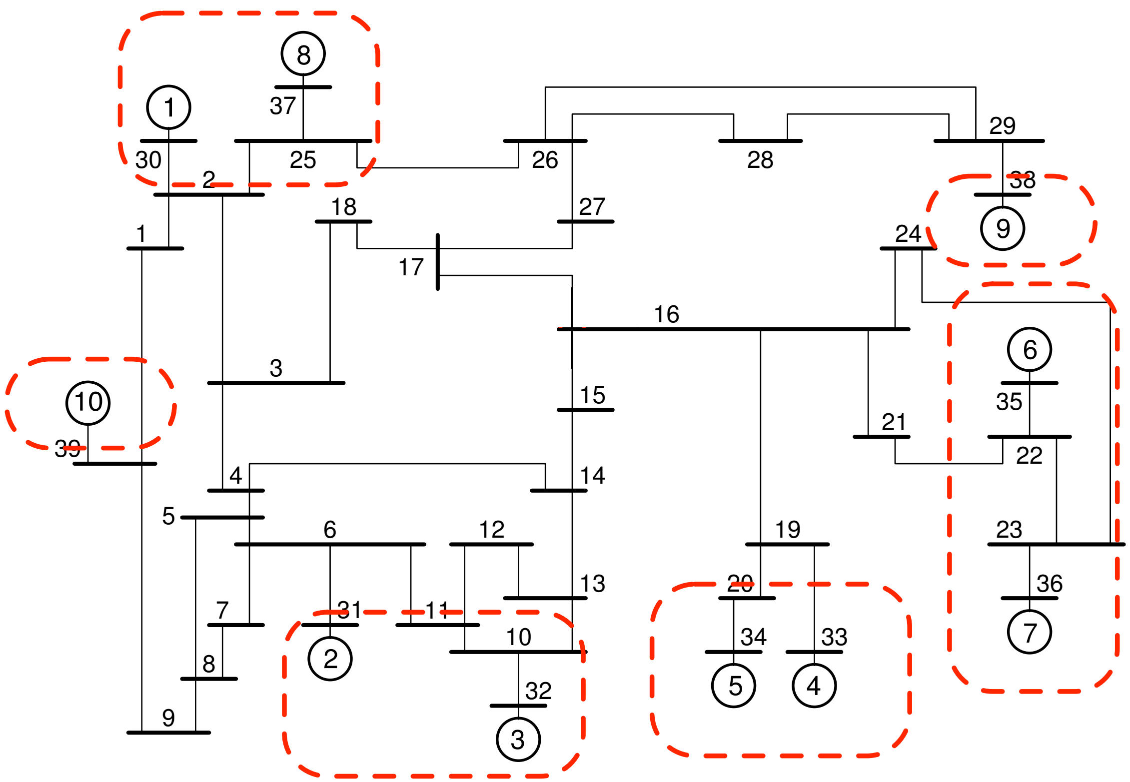
linearized swing equation:




wide-area controller
$u_{\text{wac}}(t)$
remote control signals
channel and measurement noise
wide-area measurements (e.g. PMUs)
local control loops
$u_{\text{loc}}(t)$
+
PSS & AVR
generator
$u_{\text{loc}}(t)$
+
FACTS
transmission line
power network dynamics
$\eta(t)$
system noise

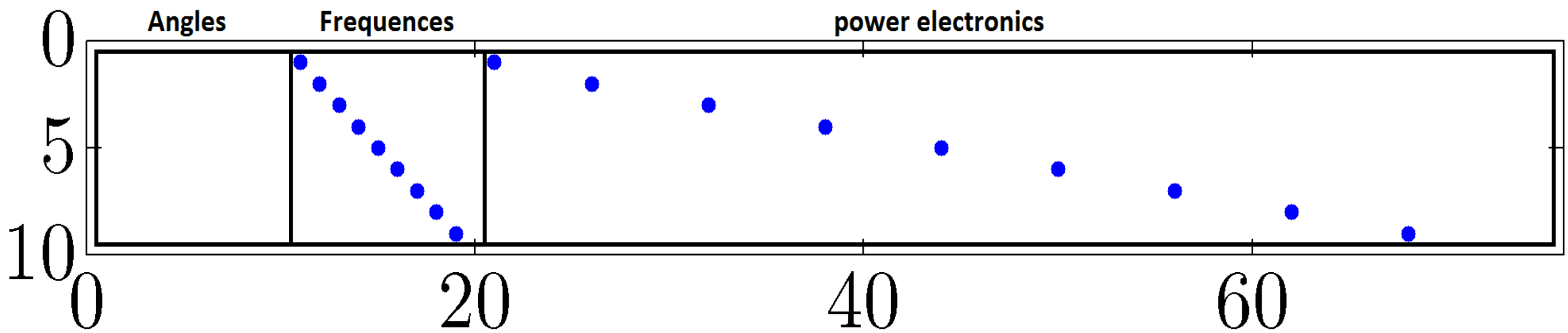
Case study: IEEE New England Power Grid

- MODEL FEATURES
  - ⋆ detailed sub-transient generator models
  - ⋆ exciters
  - ⋆ carefully tuned PSS data



Preview of a key result

- FEEDBACK GAIN STRUCTURE



fully decentralized controller $\Rightarrow$ **nearly centralized performance**

$\star$ 10% degradation relative to the optimal centralized controller

$\star$ optimal retuning of the decentralized PSS gains

An example: swing equation

$$M\ddot{\theta} + D\dot{\theta} + L\theta = d + u$$

L – Laplacian matrix

$$\Downarrow$$

only relative angle differences enter into dynamics

Performance index

- ENERGY OF POWER NETWORK

  ⋆ inspired by slow coherency theory

$$J \;:=\; \lim_{t \to \infty} \mathbf{E}\Big(\theta^T(t)\, Q_\theta\, \theta(t) \;+\; \dot{\theta}^T(t)\, M\, \dot{\theta}(t) \;+\; u^T(t)\, u(t)\Big)$$

$$Q_\theta \;:=\; I \;-\; (1/N)\, \mathbb{1}\mathbb{1}^T$$

  ⋆ Q_θ – penalizes deviation from average

$$\bar{\theta} \;:=\; (1/N)\, \mathbb{1}^T \theta$$

$$\downarrow$$

not detectable from Q_θ

Structural constraints

- ZERO E-VALUE ASSOCIATED WITH THE AVERAGE MODE

open-loop: $$A \begin{bmatrix} \mathbb{1} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}$$

closed-loop: $$(A - B_2 K) \begin{bmatrix} \mathbb{1} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}$$

Coordinate transformation

- ELIMINATE THE AVERAGE-MODE

$$\begin{bmatrix} \theta \\ \dot{\theta} \end{bmatrix} = \underbrace{\begin{bmatrix} U & 0 \\ 0 & I \end{bmatrix}}_{T} \begin{bmatrix} \psi \\ \dot{\theta} \end{bmatrix} + \begin{bmatrix} \mathbb{1} \\ 0 \end{bmatrix} \bar{\theta}$$

BY PROJECTING STATES ONTO $\begin{bmatrix} \mathbb{1} \\ 0 \end{bmatrix}^{\perp}$

columns of U – **form an orthonormal basis of** $\mathbb{1}^{\perp}$

Sparsity-promoting optimal control

$$\text{minimize} \quad J(F) \quad + \quad \gamma \, \| F \, T^T \|_1$$

$$\downarrow \qquad\qquad\qquad \downarrow$$

new coordinates **original coordinates**

(nonconvex, smooth) **(convex, nonsmooth)**

⋆ $F \;=\; K\,T$ $\;-\;$ to eliminate the average-mode

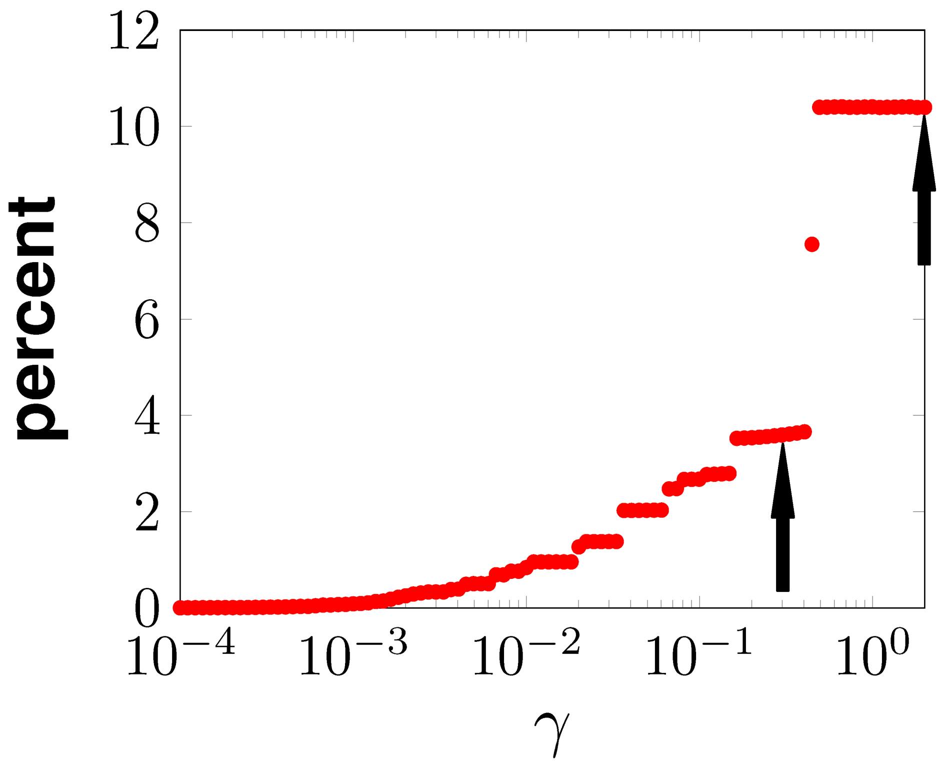
⋆ $\| F \, T^T \|_1$ $\;-\;$ **not separable in the elements of** F

- OPTIMAL CONTROL PROBLEM

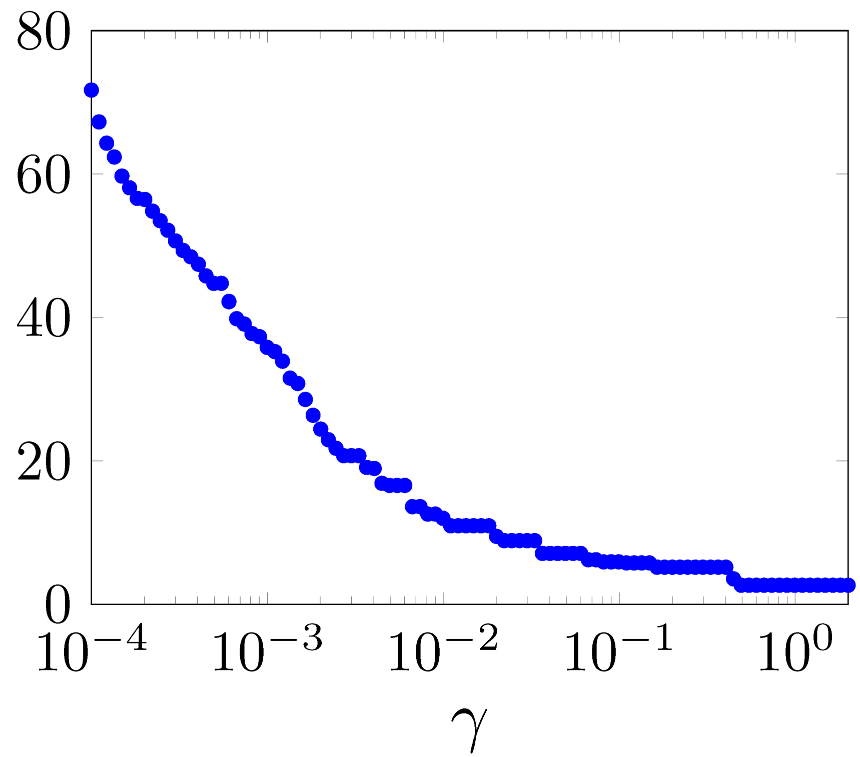
$$\begin{array}{ll} \underset{F,\ K}{\text{minimize}} & J(F) \ + \ \gamma \, \| \, K \, \|_1 \\ \text{subject to} & F \, T^T \ - \ K \ = \ 0 \end{array}$$

Performance vs sparsity

performance loss:



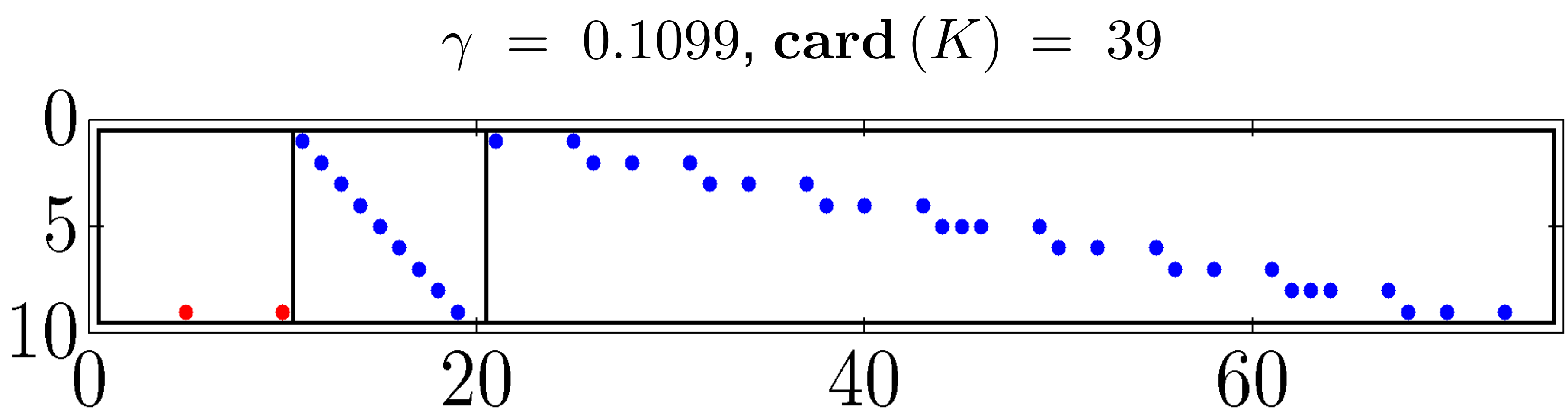
sparsity of K:



Information exchange network

SPARSITY PATTERN OF K

- **local**
- **long-range interactions**

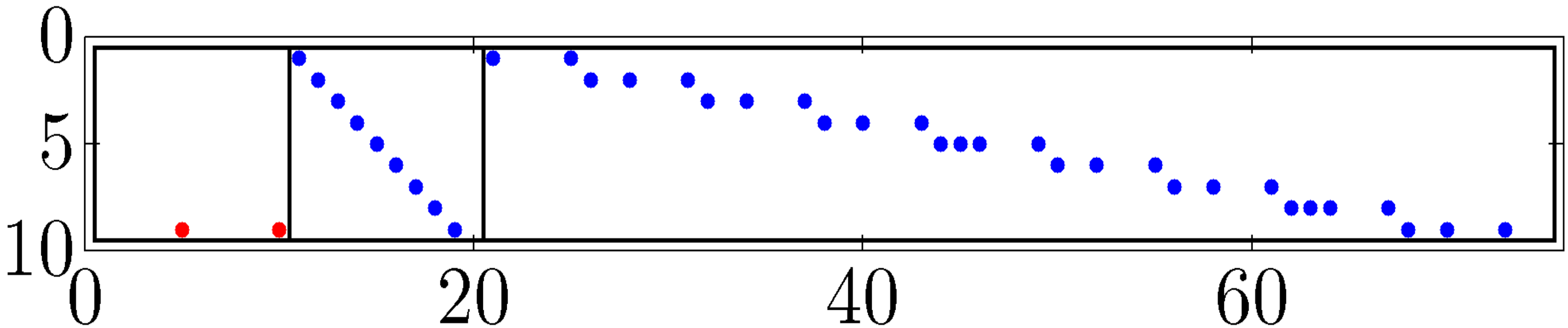


Information exchange network

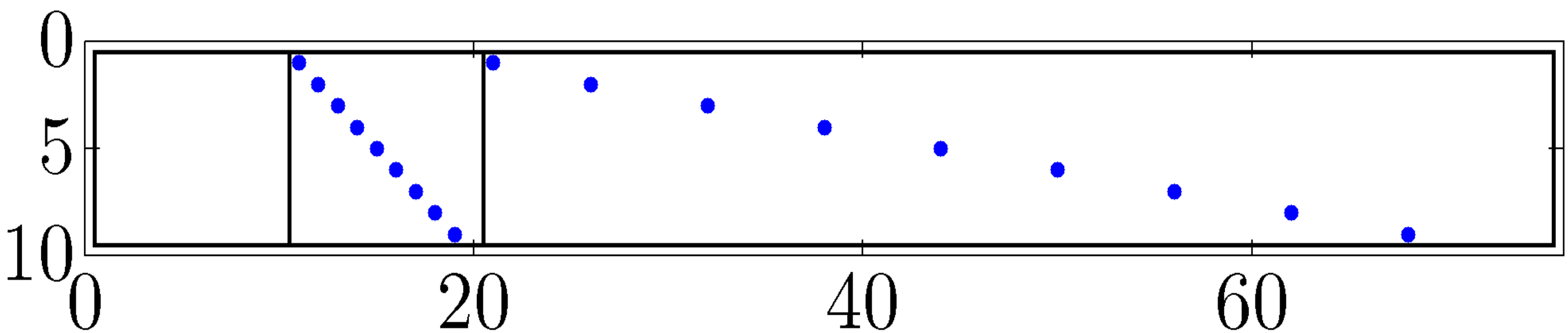
SPARSITY PATTERN OF K

- local
- long-range interactions

$\gamma = 0.1099$, **card** $(K) = 39$



$\gamma = 2$, **card** $(K) = 18$



Response to stochastic forcing

- WHITE-IN-TIME FORCING

$$\mathbf{E}\,(d(t_1)\,d^*(t_2)) \;=\; I\,\delta(t_1 \,-\, t_2)$$

⋆ **Hilbert-Schmidt norm**

power spectral density:

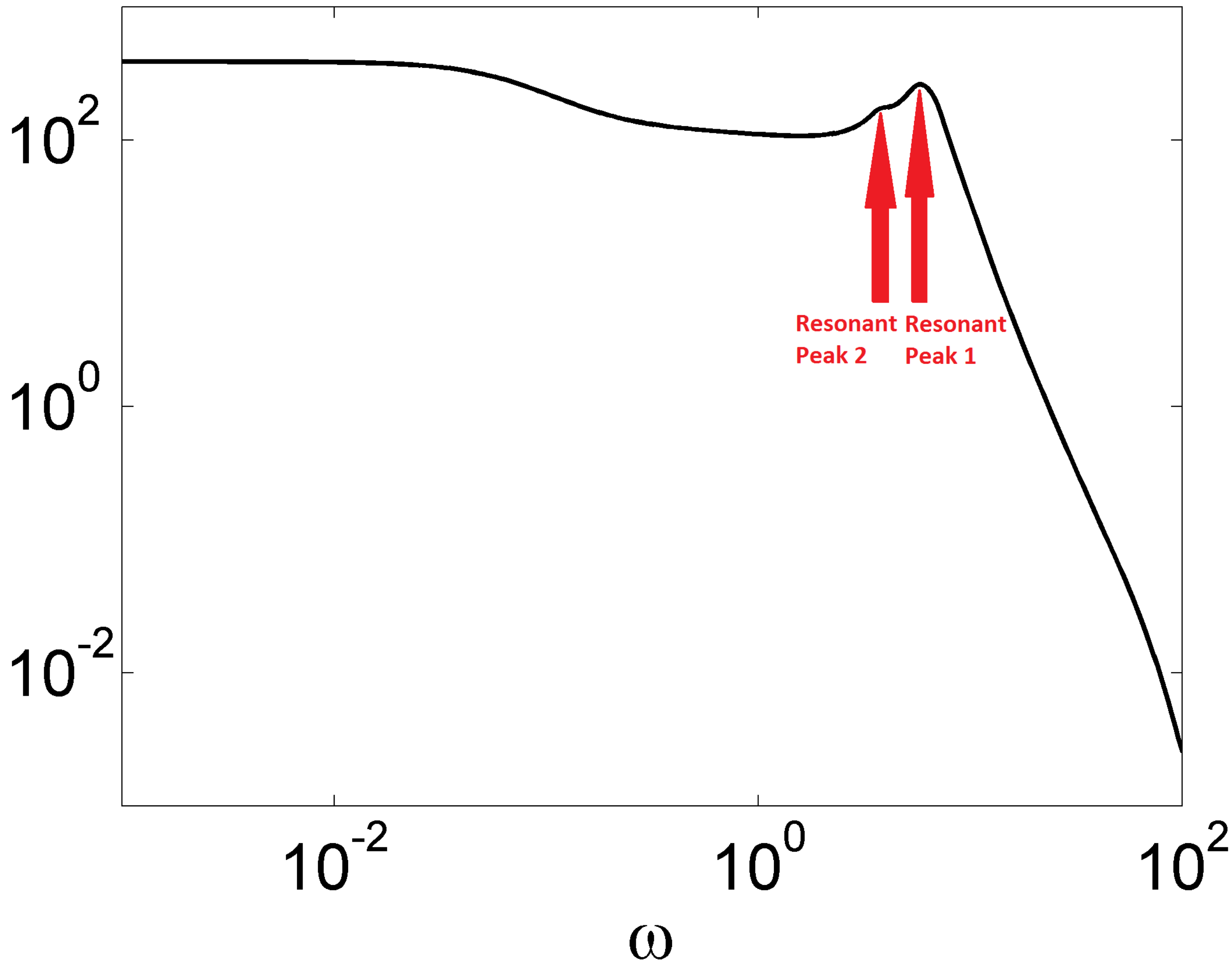
$$\|H(\omega)\|_{\mathrm{HS}}^2 \;=\; \mathrm{trace}\,(H(\omega)\,H^*(\omega)) \;=\; \sum_i \sigma_i^2\,(\omega)$$

⋆ H_2 **norm**

variance amplification:

$$\|H\|_2^2 \;=\; \frac{1}{2\,\pi}\int_{-\infty}^{\infty} \|H(\omega)\|_{\mathrm{HS}}^2\,\mathrm{d}\omega$$

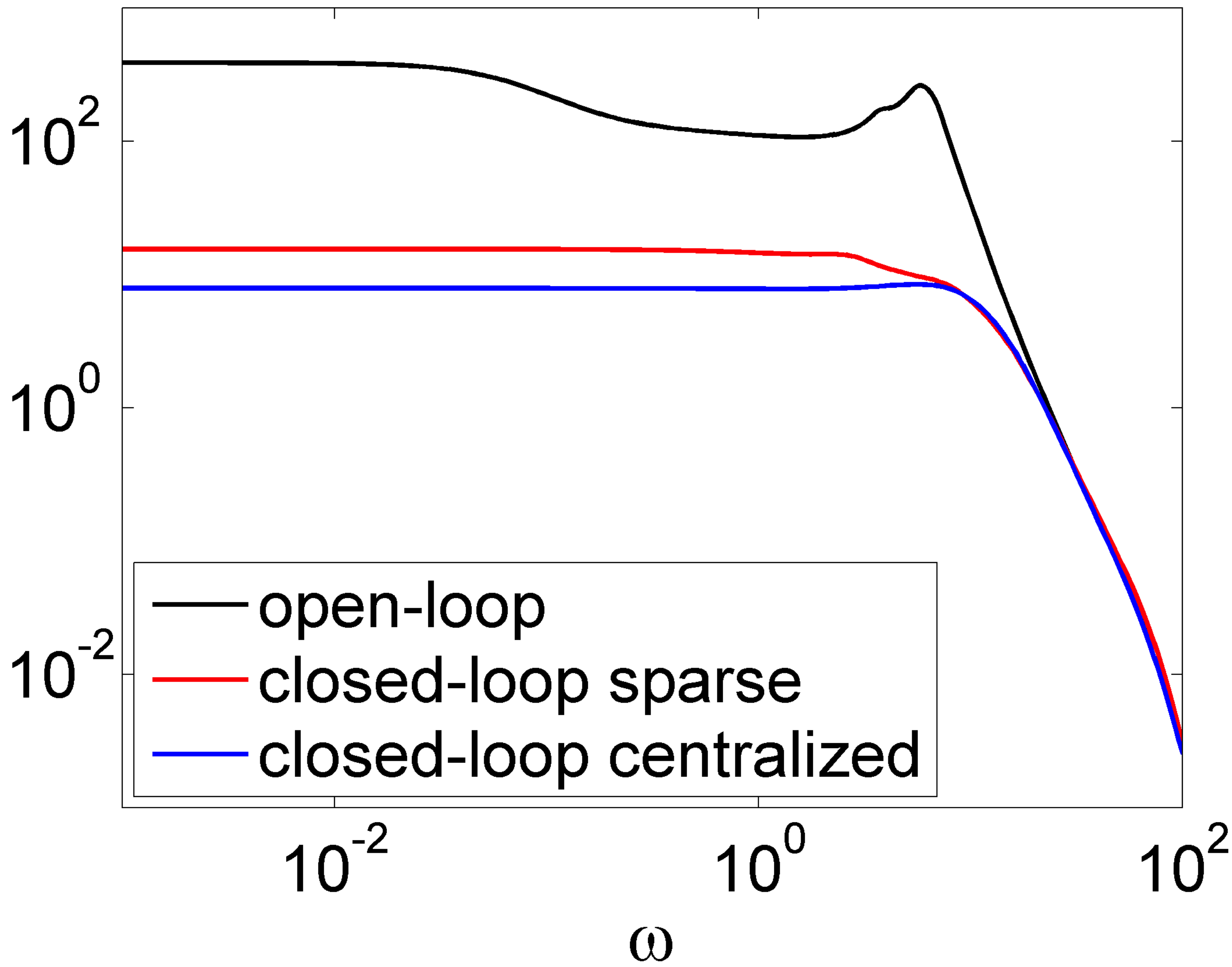
Open-loop dynamics: power spectral density



⋆ resonant peak 1: inter-area modes 2, 3, 4, 5

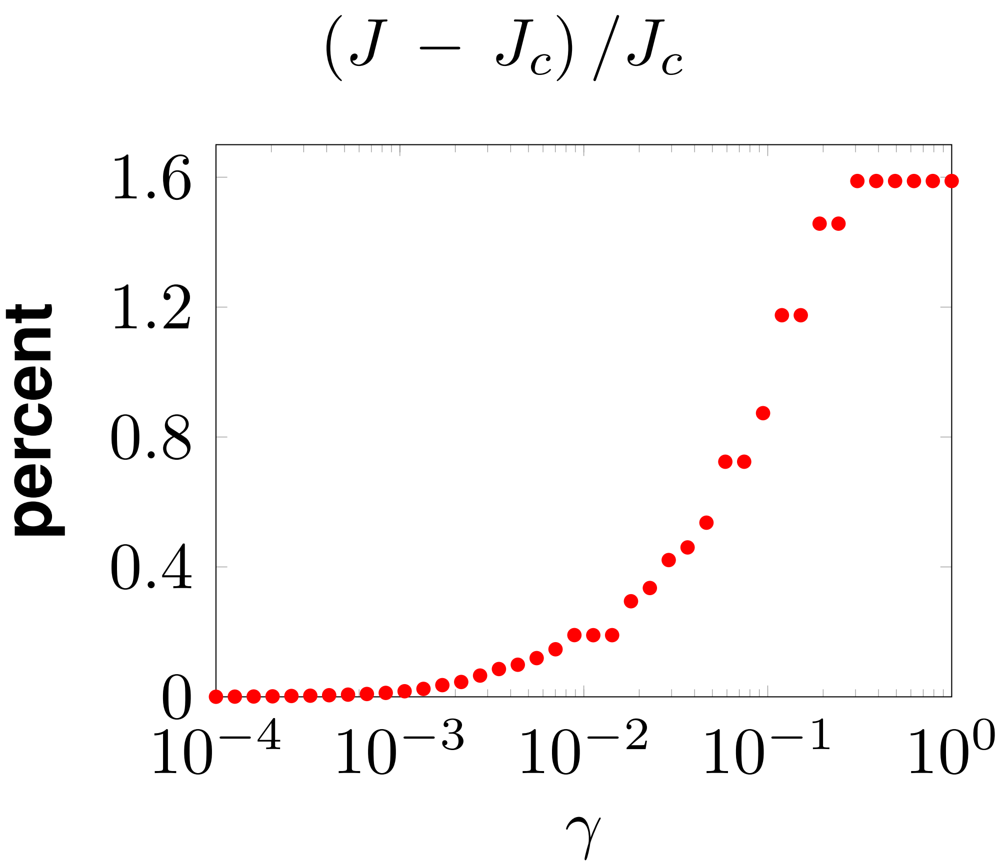
⋆ resonant peak 2: inter-area mode 1

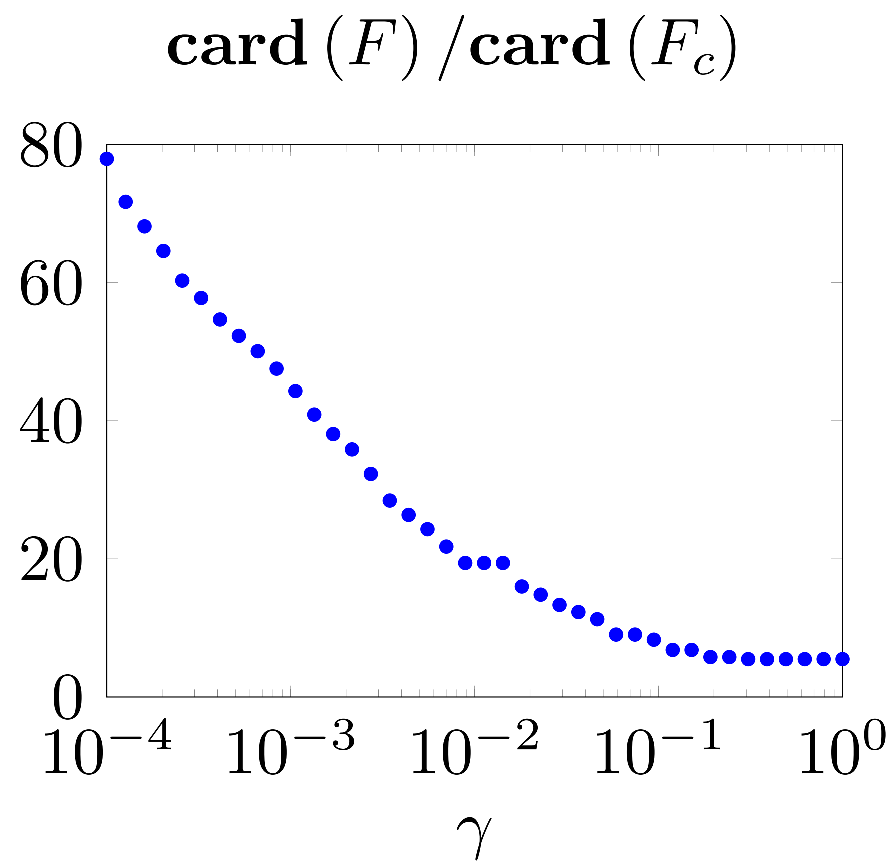
Open-loop vs closed-loop systems



⋆ low frequencies: 10% performance degradation

- **Performance comparison: block-sparse vs centralized**





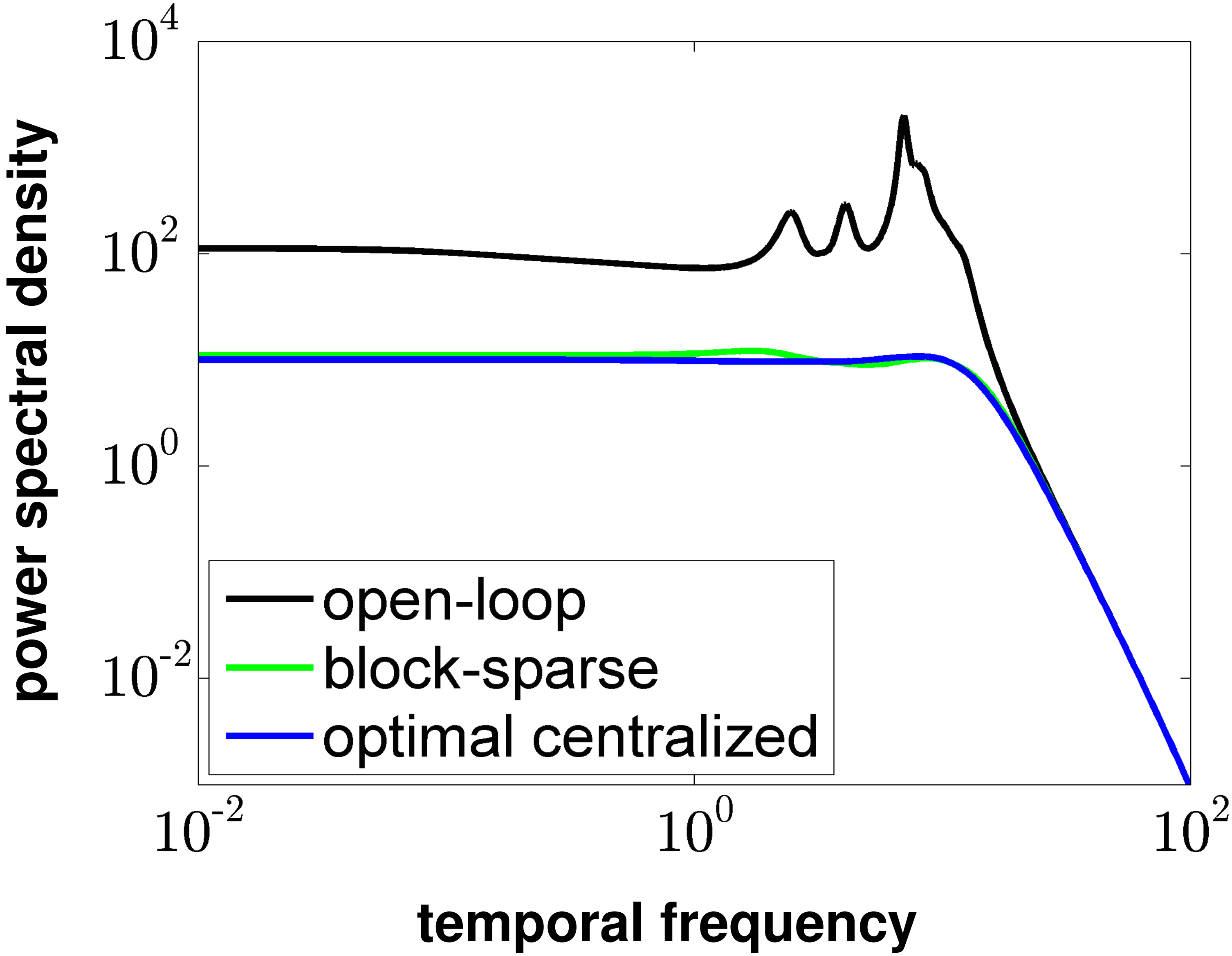
$\gamma = 1 \xrightarrow{\text{relative to } F_c}$

$1.6\,\%$ performance loss

$5.5\,\%$ non-zero elements in F

• RE-DESIGN OF FULLY-DECENTRALIZED CONTROLLERS

⋆ **preserves rotational symmetry**

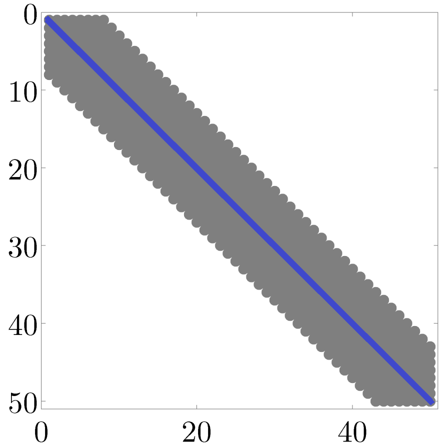


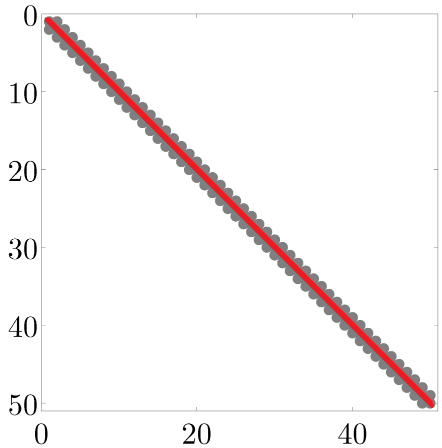
Wu, Dörfler, Jovanović, IEEE TPWRS '16

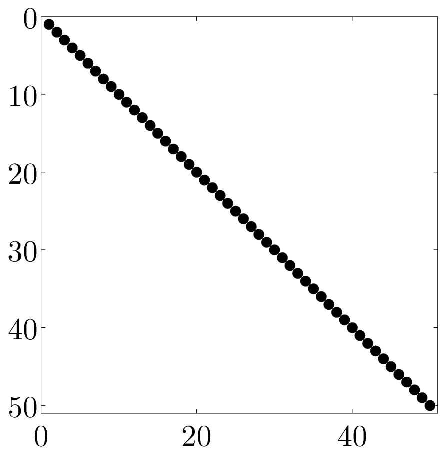
ADDITIONAL EXAMPLES

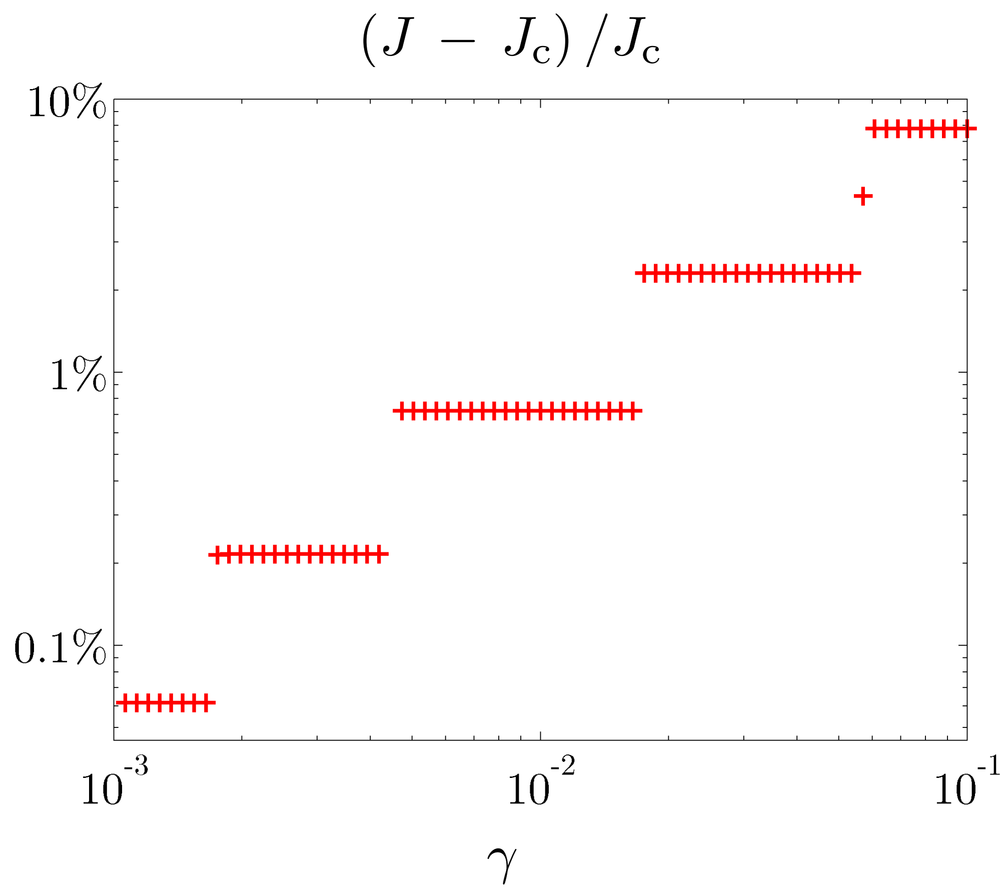
Mass-spring system

Performance comparison: sparse vs centralized



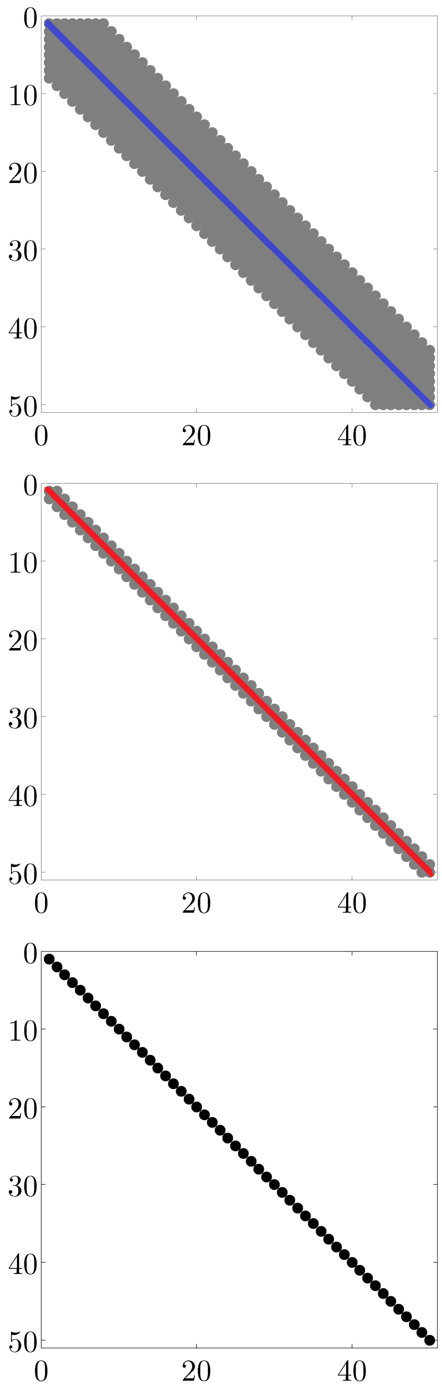


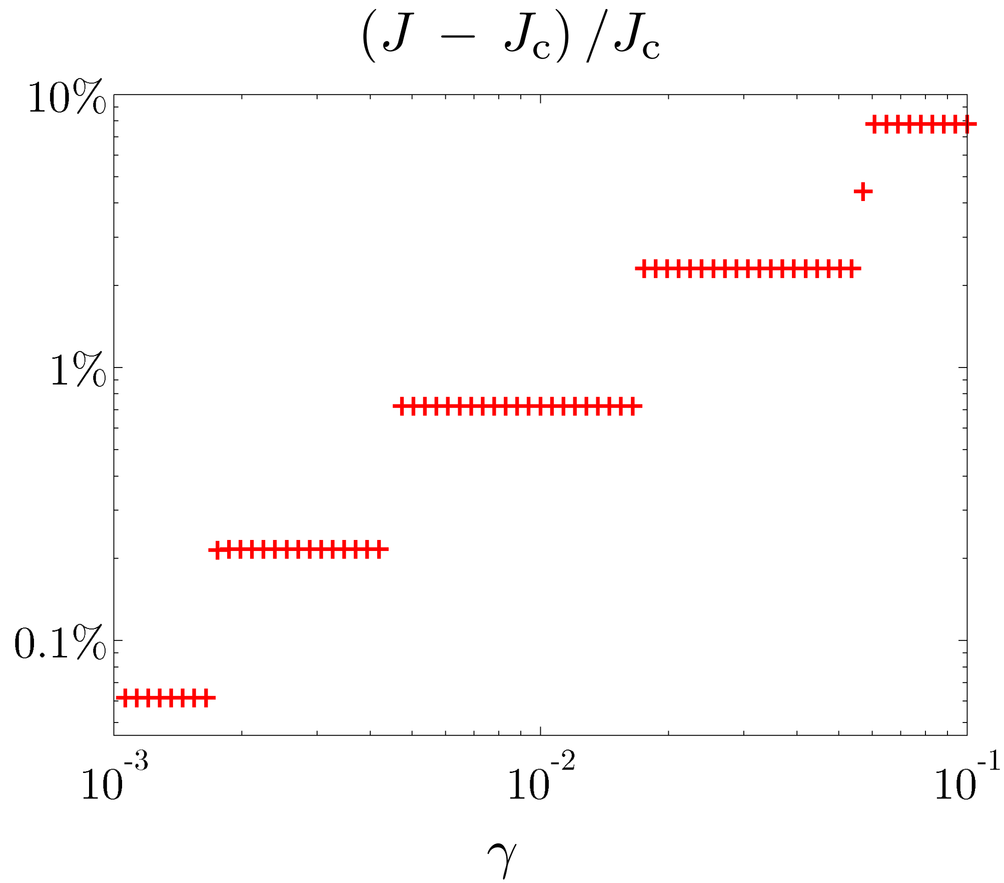




Mass-spring system

Performance comparison: sparse vs centralized

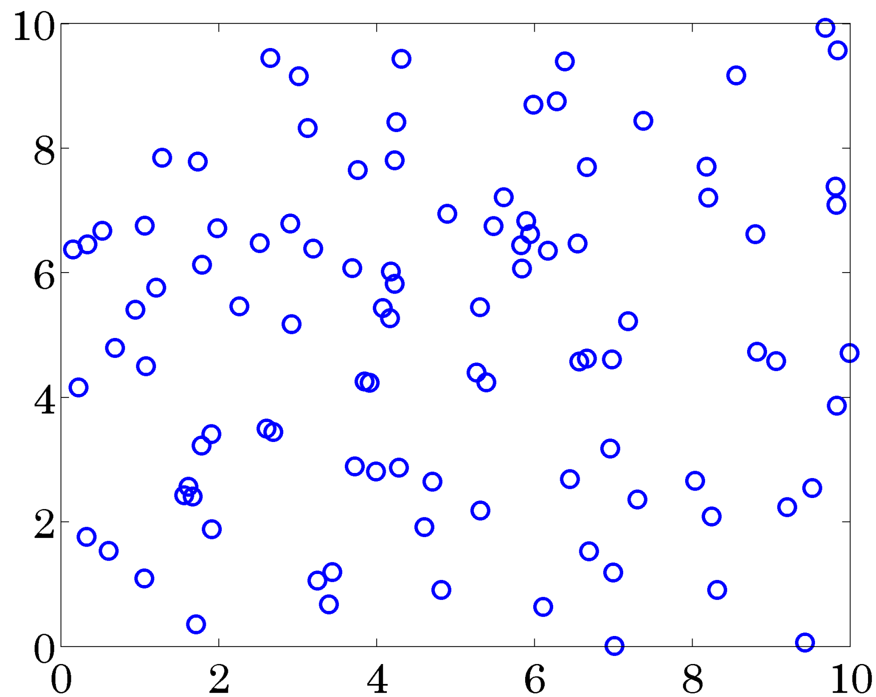




$\mathbf{card}(F)/\mathbf{card}(F_c)$	$(J - J_c)/J_c$
10%	0.75%
6%	2.4%
2%	7.8%

fully-decentralized

Network with 100 nodes

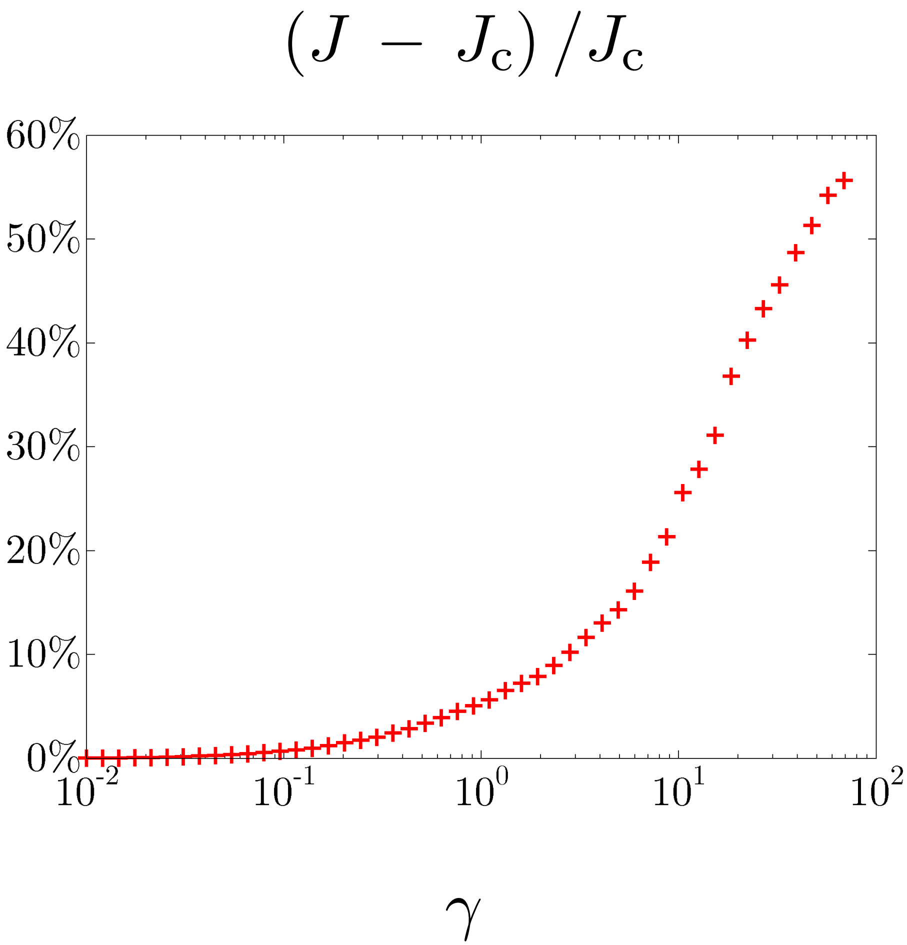


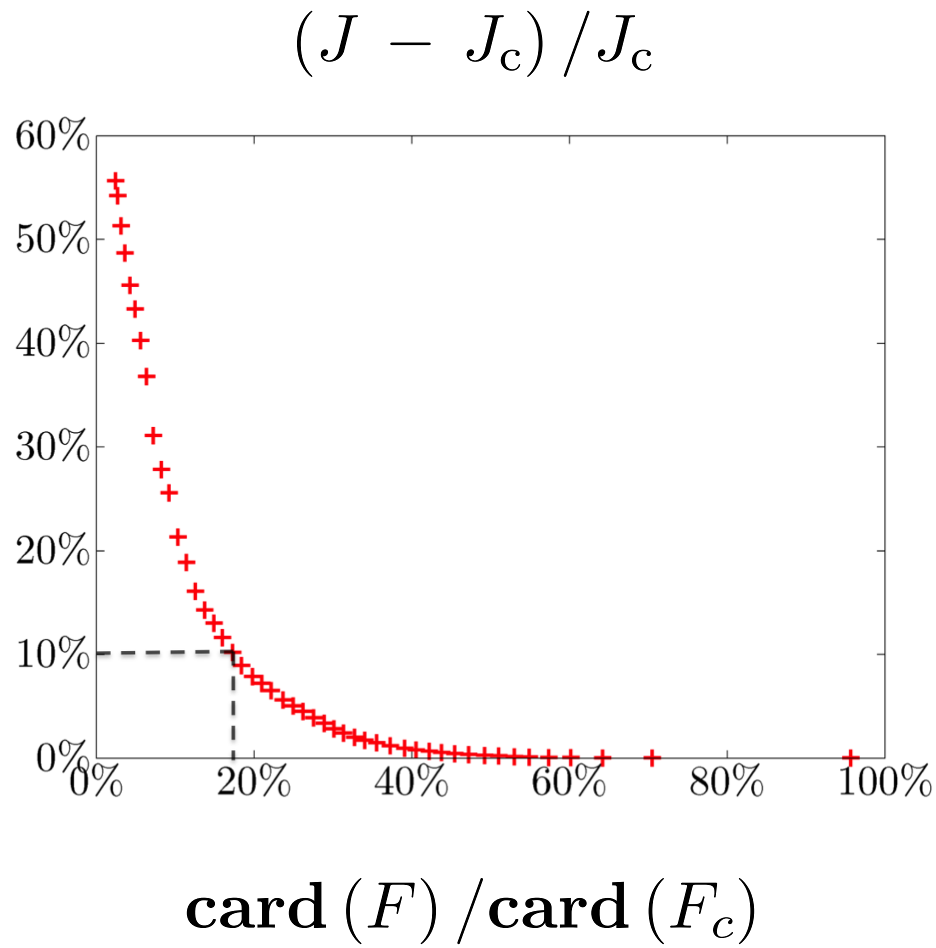
$$\begin{matrix}\dot{p}_i\\ \dot{v}_i\end{matrix} = \underbrace{\begin{matrix}1 & 1\\ 1 & 2\end{matrix} \quad \begin{matrix}p_i\\ v_i\end{matrix}}_{\textbf{unstable dynamics}} + \underbrace{{}_{j \neq i} \; e^{-\alpha(i,j)} \quad \begin{matrix}p_j\\ v_j\end{matrix}}_{\textbf{coupling}} + \begin{matrix}0\\ 1\end{matrix} \; (d_i + u_i)$$

$\alpha(i,j)$: Euclidean distance between nodes i and j

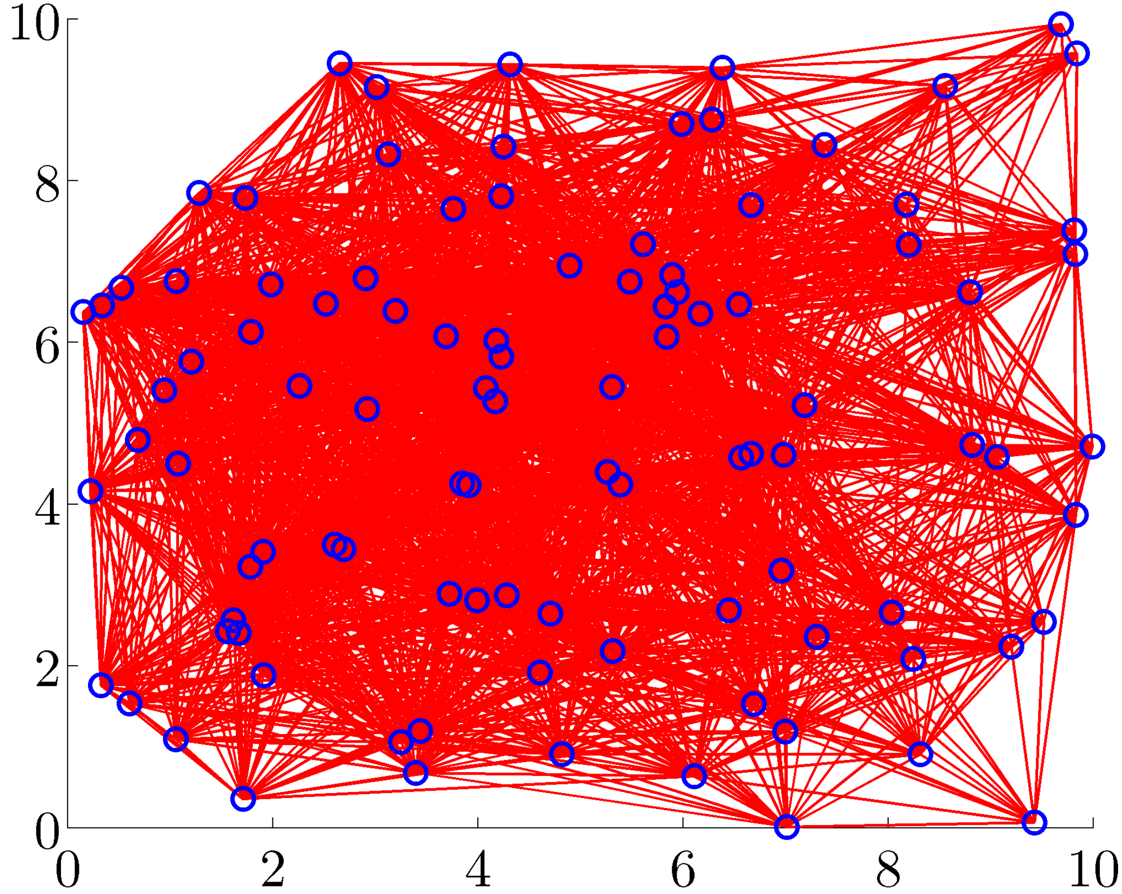
Motee & Jadbabaie, IEEE TAC '08

- **Performance comparison: sparse vs centralized**





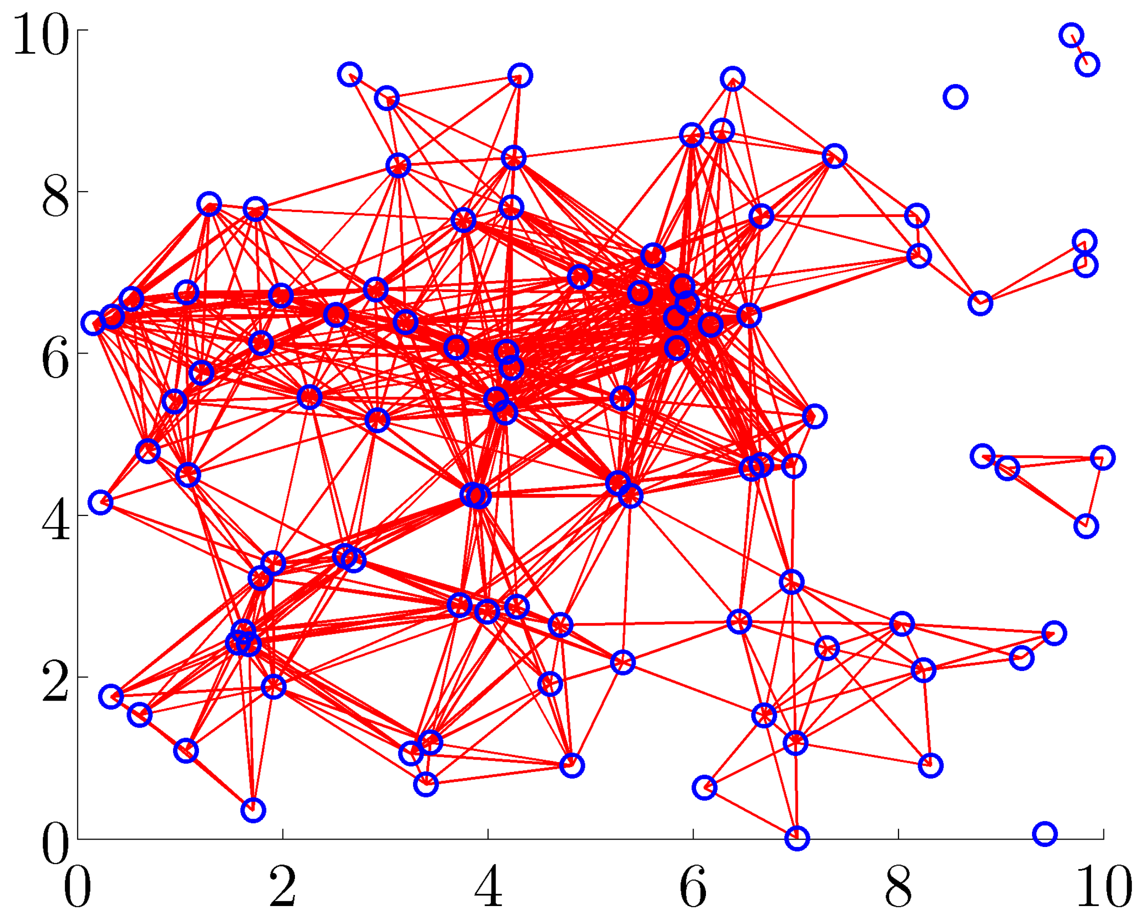
communication graph of a truncated centralized gain



$$\mathbf{card}\,(F) \;=\; 7380 \;\; (36.9\%)$$

non-stabilizing

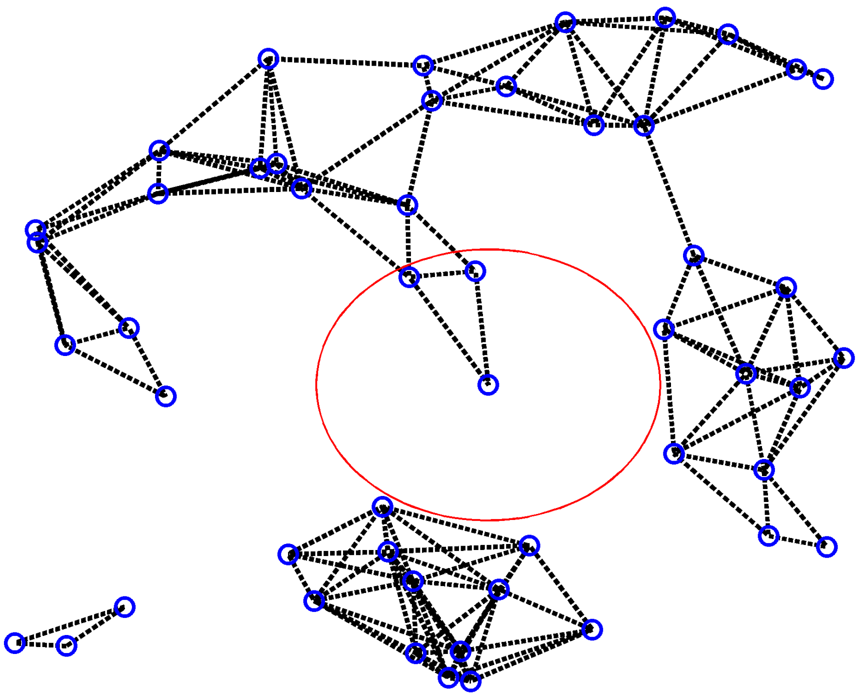
identified communication graph

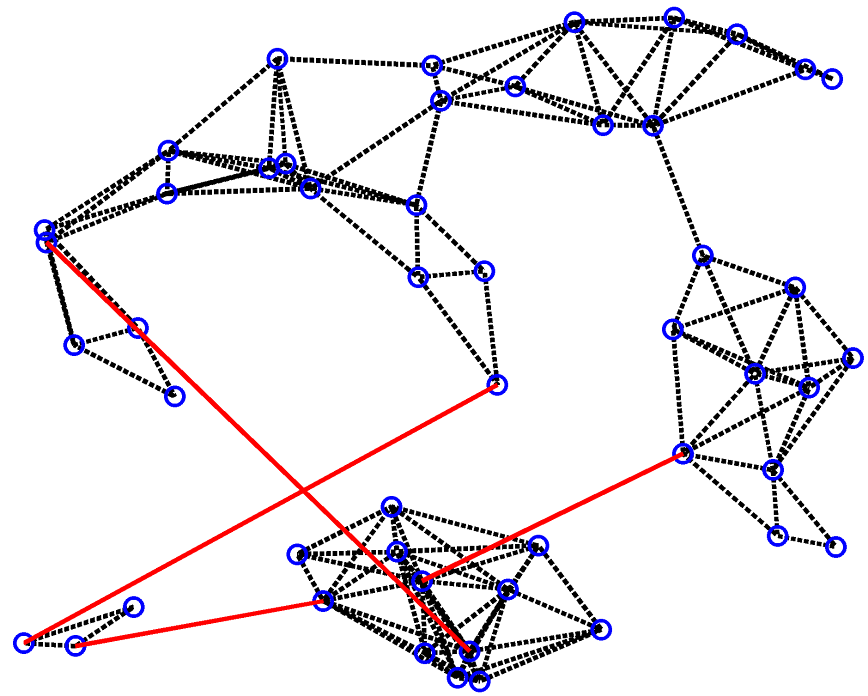


$$\begin{aligned} \gamma &= 5 \\ \mathbf{card}\,(F)\,/\,\mathbf{card}\,(F_\mathrm{c}) &= 8.8\% \\ (J - J_\mathrm{c})\,/\,J_\mathrm{c} &= 24.6\% \end{aligned}$$

Sparsity-promoting consensus algorithm

plant graph **identified communication graph**





Q := deviation from average

$$\frac{J - J_{\text{all-to-all}}}{J_{\text{all-to-all}}} \approx 82\%$$

Hassan-Moghaddam & Jovanović, arXiv:1506.03437

ALGORITHM

Method of multipliers

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

- **Step 1: introduce an additional variable/constraint**

$$\begin{array}{ll} \text{minimize} & J(F) \; + \; \gamma \, g(G) \\ \text{subject to} & F \; - \; G \; = \; 0 \end{array}$$

benefit: decouples J and g

Method of multipliers

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

- **Step 1: introduce an additional variable/constraint**

$$\begin{array}{ll} \text{minimize} & J(F) \; + \; \gamma\, g(G) \\ \text{subject to} & F \; - \; G \; = \; 0 \end{array}$$

benefit: decouples J and g

- **Step 2: introduce augmented Lagrangian**

$$\mathcal{L}_\rho(F, G; \Lambda) \; = \; J(F) \; + \; \gamma\, g(G) \; + \; \langle \Lambda, F \; - \; G \rangle \; + \; \frac{\rho}{2} \, \| F \; - \; G \|_F^2$$

- **Step 3: use MM for augmented Lagrangian minimization**

$$\mathcal{L}_\rho(F, G; \Lambda) \;=\; J(F) \;+\; \gamma\, g(G) \;+\; \langle \Lambda, F \,-\, G \rangle \;+\; \frac{\rho}{2}\, \| F \,-\, G \|_F^2$$

METHOD OF MULTIPLIERS

$$\begin{aligned} \left(F^{k+1}, G^{k+1}\right) \;&:=\; \underset{F,\,G}{\operatorname{argmin}}\; \mathcal{L}_{\rho^k}(F, G; \Lambda^k) \\ \Lambda^{k+1} \;&:=\; \Lambda^k \;+\; \rho^k \left(F^{k+1} \,-\, G^{k+1}\right) \end{aligned}$$

- **Step 4: Polishing – back to structured optimal design**

  - ⋆ MM
    - identifies sparsity patterns
    - provides good initial condition for structured design

- **Step 4: Polishing – back to structured optimal design**

⋆ MM identifies sparsity patterns

provides good initial condition for structured design

⋆ **optimality conditions for the structured problem**

$$
\begin{aligned}
(A - B_2 F)^T P + P (A - B_2 F) &= -\left(Q + F^T R F\right) \\
(A - B_2 F) X + X (A - B_2 F)^T &= -B_1 B_1^T \\
\left[\left(R F - B_2^T P\right) X \circ I_{\mathcal{S}}\right. &= 0
\end{aligned}
$$

$I_{\mathcal{S}}$ - structural identity

$$
F = \begin{bmatrix} * & * & & \\ * & * & * & \\ & * & * & * \\ & & * & * \end{bmatrix} \quad \Rightarrow \quad I_{\mathcal{S}} = \begin{bmatrix} 1 & 1 & & \\ 1 & 1 & 1 & \\ & 1 & 1 & 1 \\ & & 1 & 1 \end{bmatrix}
$$

Lin, Fardad, Jovanović, IEEE TAC ’11

Proximal operator and Moreau envelope

- PROXIMAL OPERATOR

$$\mathbf{prox}_{\mu g}(V) \;:=\; \underset{G}{\operatorname{argmin}}\; g(G) \;+\; \frac{1}{2\mu}\|G \;-\; V\|_F^2$$

MOREAU ENVELOPE

$$M_{\mu g}(V) \;:=\; \inf_{G}\; g(G) \;+\; \frac{1}{2\mu}\|G \;-\; V\|_F^2$$

Proximal operator and Moreau envelope

- PROXIMAL OPERATOR

$$\mathbf{prox}_{\mu g}(V) \ := \ \underset{G}{\text{argmin}} \ \ g(G) \ + \ \frac{1}{2\mu} \|G \ - \ V\|_F^2$$

MOREAU ENVELOPE

$$M_{\mu g}(V) \ := \ \inf_{G} \ g(G) \ + \ \frac{1}{2\mu} \|G \ - \ V\|_F^2$$

⋆ **continuously differentiable**
even when g is not

$$\nabla M_{\mu g}(V) \ = \ \frac{1}{\mu} \big(V \ - \ \mathbf{prox}_{\mu g}(V) \big)$$

Parikh & Boyd, FnT in Optimization '14

Proximal augmented Lagrangian

$$\mathcal{L}_\rho(F, G; \Lambda) \;=\; J(F) + \underbrace{\gamma\, g(G) \;+\; \frac{\rho}{2}\,\|G \,-\, (F \,+\, (1/\rho)\Lambda)\|_F^2}_{} \;-\; \frac{1}{2\rho}\,\|\Lambda\|_F^2$$

- ⋆ minimize over G

$$G^\star \;=\; \mathbf{prox}_{(\gamma/\rho)g}(F \,+\, (1/\rho)\Lambda)$$

- ⋆ evaluate $\mathcal{L}_\rho$ at $G^\star$

$$\begin{aligned} \mathcal{L}_\rho(F; \Lambda) \;&:=\; \mathcal{L}_\rho(F, G^\star(F, \Lambda); \Lambda) \\ &=\; J(F) \,+\, \gamma\, M_{(\gamma/\rho)g}(F \,+\, (1/\rho)\Lambda) \,-\, \frac{1}{2\rho}\,\|\Lambda\|_F^2 \end{aligned}$$

continuously differentiable

Method of multipliers

$$\begin{aligned} F^{k+1} &= \underset{F}{\text{argmin}} \; \mathcal{L}_{\rho^k}(F; \Lambda^k) \\ \Lambda^{k+1} &= \gamma \, \nabla M_{(\gamma/\rho^k)g}(F^{k+1} + (1/\rho^k)\Lambda^k) \end{aligned}$$

- FEATURES
  - ⋆ outstanding practical performance
  - ⋆ nonconvex J: convergence to a local minimum
  - ⋆ F-minimization: differentiable problem
  - ⋆ adaptive ρ-update

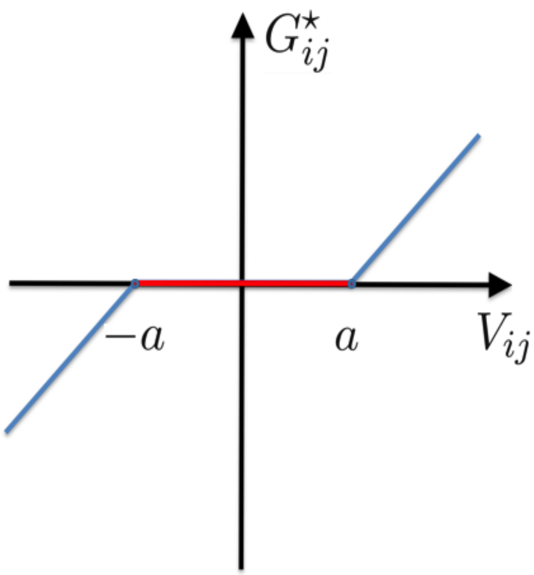
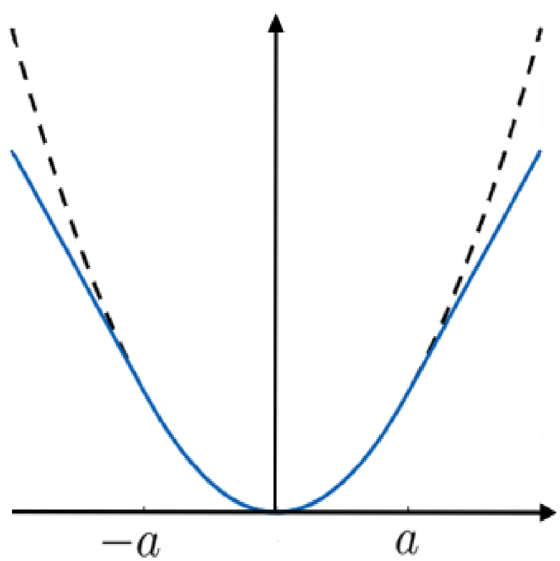
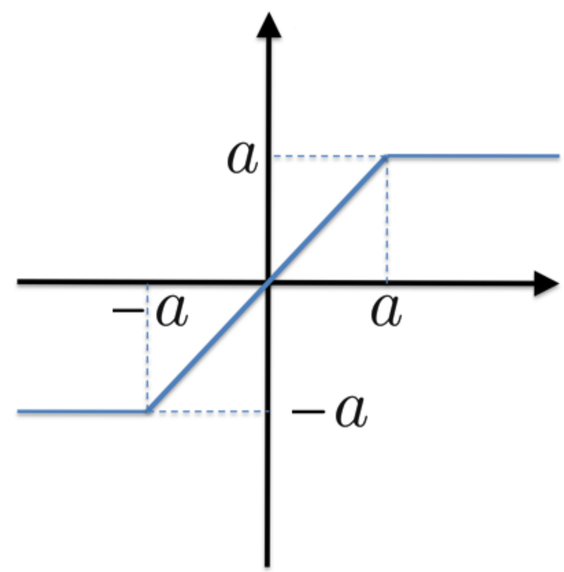
Dhingra & Jovanović, ACC '16

Dhingra & Jovanović, arXiv:1610.04514

• G-UPDATE IN SPARSITY-PROMOTING PROBLEM

$$\underset{G_{ij}}{\text{minimize}} \qquad {}_{i,\,j} \quad \gamma\, w_{ij}\, |G_{ij}| \;+\; \frac{\rho}{2}\,(G_{ij} \,-\, V_{ij})^2$$

separability $\Rightarrow$ **element-wise analytical solution**

prox operator soft-thresholding	**Moreau envelope** Huber function	∇M saturation
		
	$a \;=\; (\gamma/\rho)\, w_{ij}$	

Related effort

- SPARSITY-PROMOTING H_∞ CONTROL

  Schuler, Li, Lam, Allgöwer, IJC '11

  Schuler, Münz, Allgöwer, IFAC '12

- SYSTEMS WITH SYMMETRIES

  Dhingra & Jovanović, ACC '15

  Wu & Jovanović, SCL '17

- CONVEX RELAXATIONS

  Lavaei, Allerton '13

  Fazelnia, Madani, Lavaei, CDC '14

  Fardad & Jovanović, ACC '14

- ATOMIC NORM REGULARIZATION

  Matni, CDC '13; *IEEE TCNS '17*; *Matni & Chandrasekaran, IEEE TAC '16*

- SYSTEM-LEVEL SYNTHESIS

  Wang, Matni, Doyle, IEEE TAC '17 (submitted)

Summary

- SPARSITY-PROMOTING OPTIMAL CONTROL
  - ⋆ Performance vs sparsity tradeoff

    Lin, Fardad, Jovanović, IEEE TAC '13

    Jovanović & Dhingra, EJC '16
  - ⋆ Software

    www.umn.edu/∼mihailo/software/lqrsp/

- ONGOING EFFORT
  - ⋆ Leader selection in large dynamic networks

    Lin, Fardad, Jovanović, IEEE TAC '14
  - ⋆ Optimal synchronization of sparse oscillator networks

    Fardad, Lin, Jovanović, IEEE TAC '14
  - ⋆ Optimal design of distributed integral action

    Wu, Dörfler, Jovanović, ACC '16

Acknowledgments

Makan
Syracuse

Fu
UTRC

Neil
U of M

Xiaofan
Siemens

Sepideh
USC

Florian
ETH Zürich

Support

NSF Award ECCS-14-07958 (Program manager: Kishan Baheti)

AFOSR Award FA9550-16-1-0009 (Program manager: Frederick Leve)