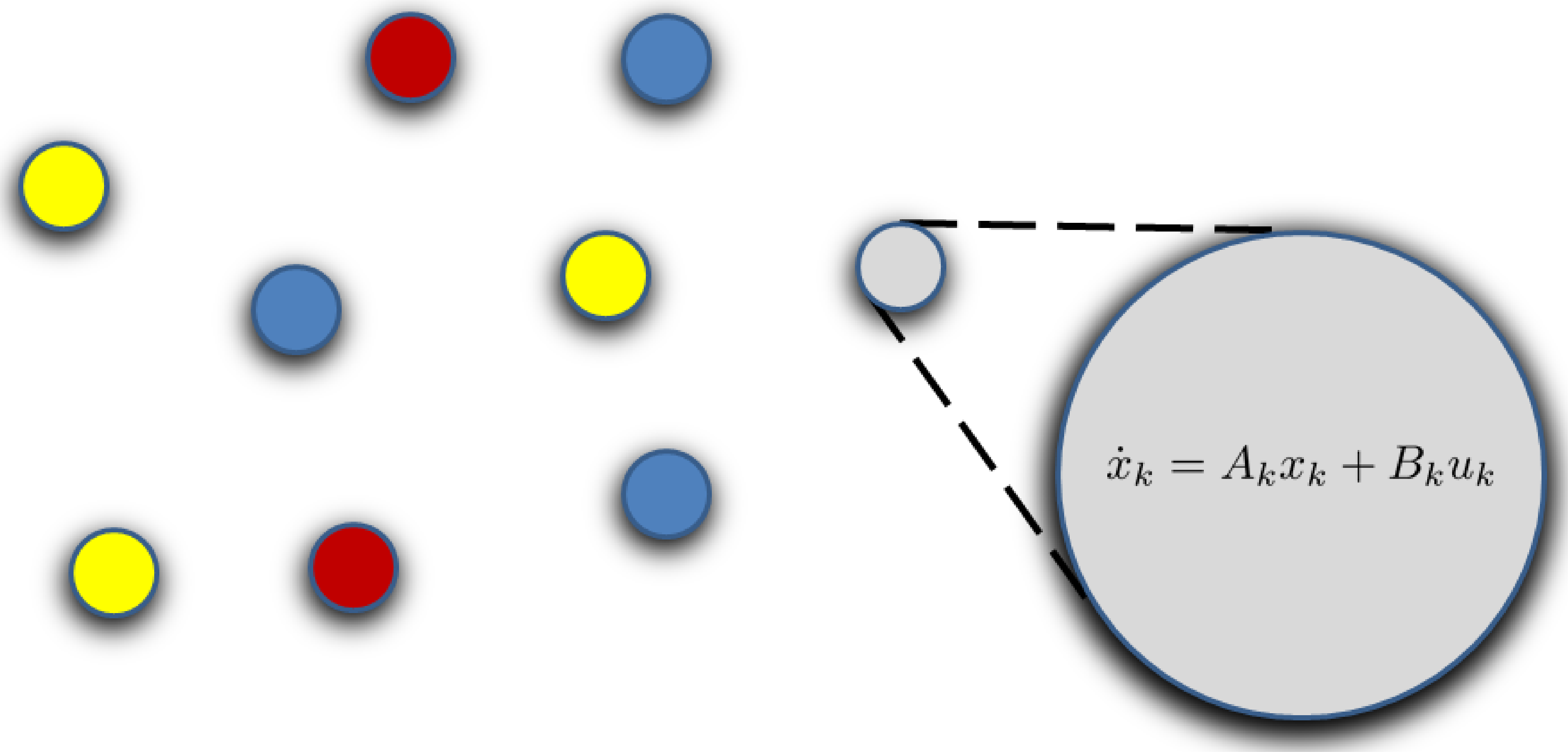
Structure identification and optimal design of networks of dynamical systems

Fu Lin
www.ece.umn.edu/∼fu

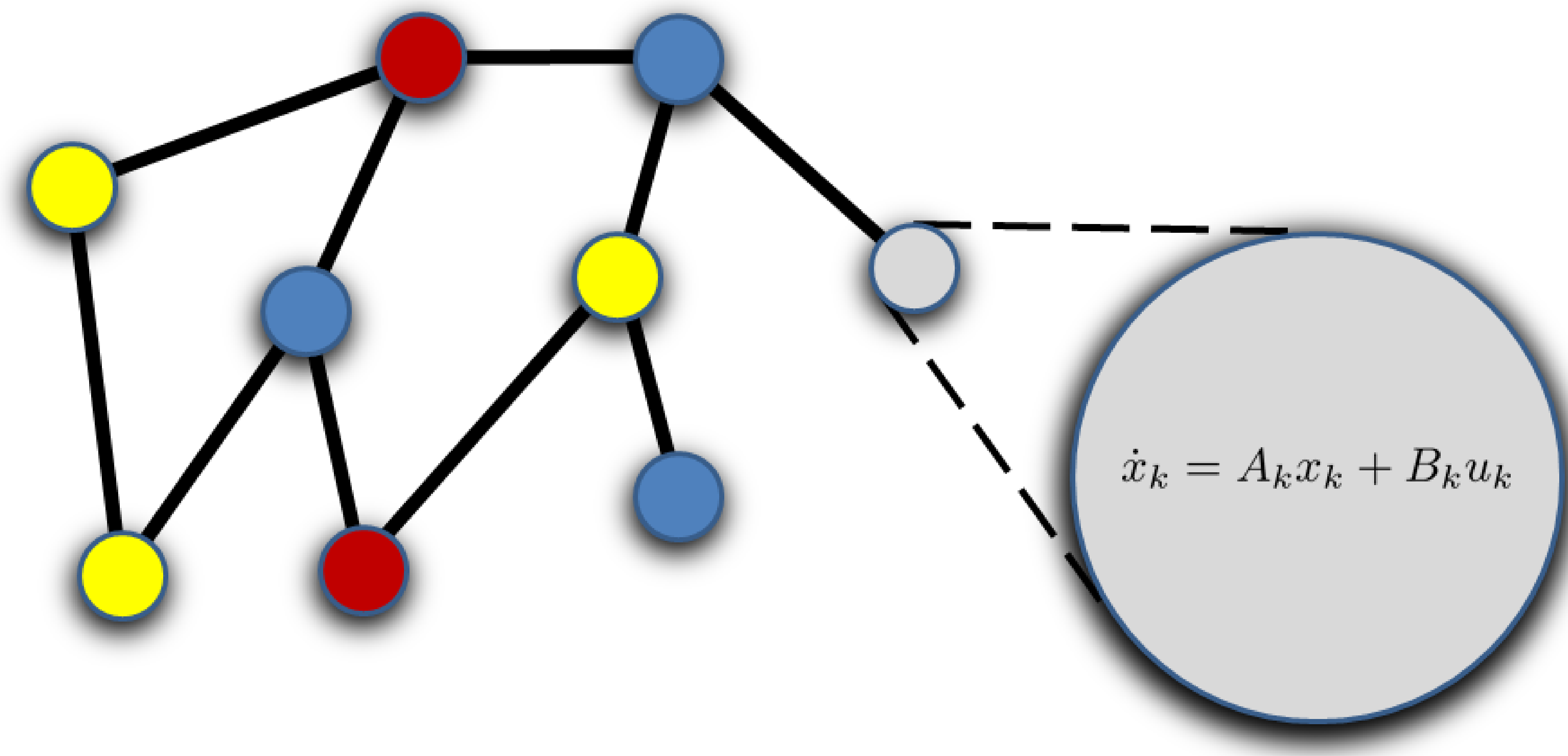
UNIVERSITY OF MINNESOTA

Argonne National Laboratory; May 6, 2013

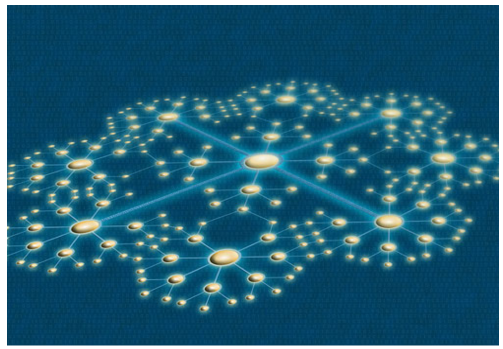
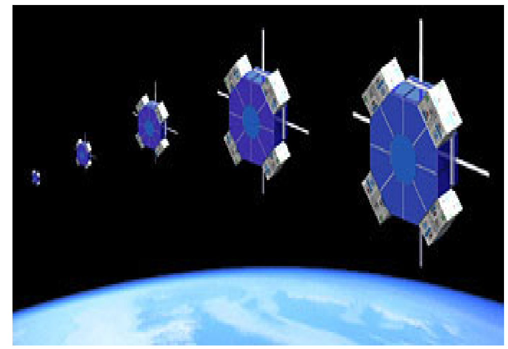
Networks of dynamical systems



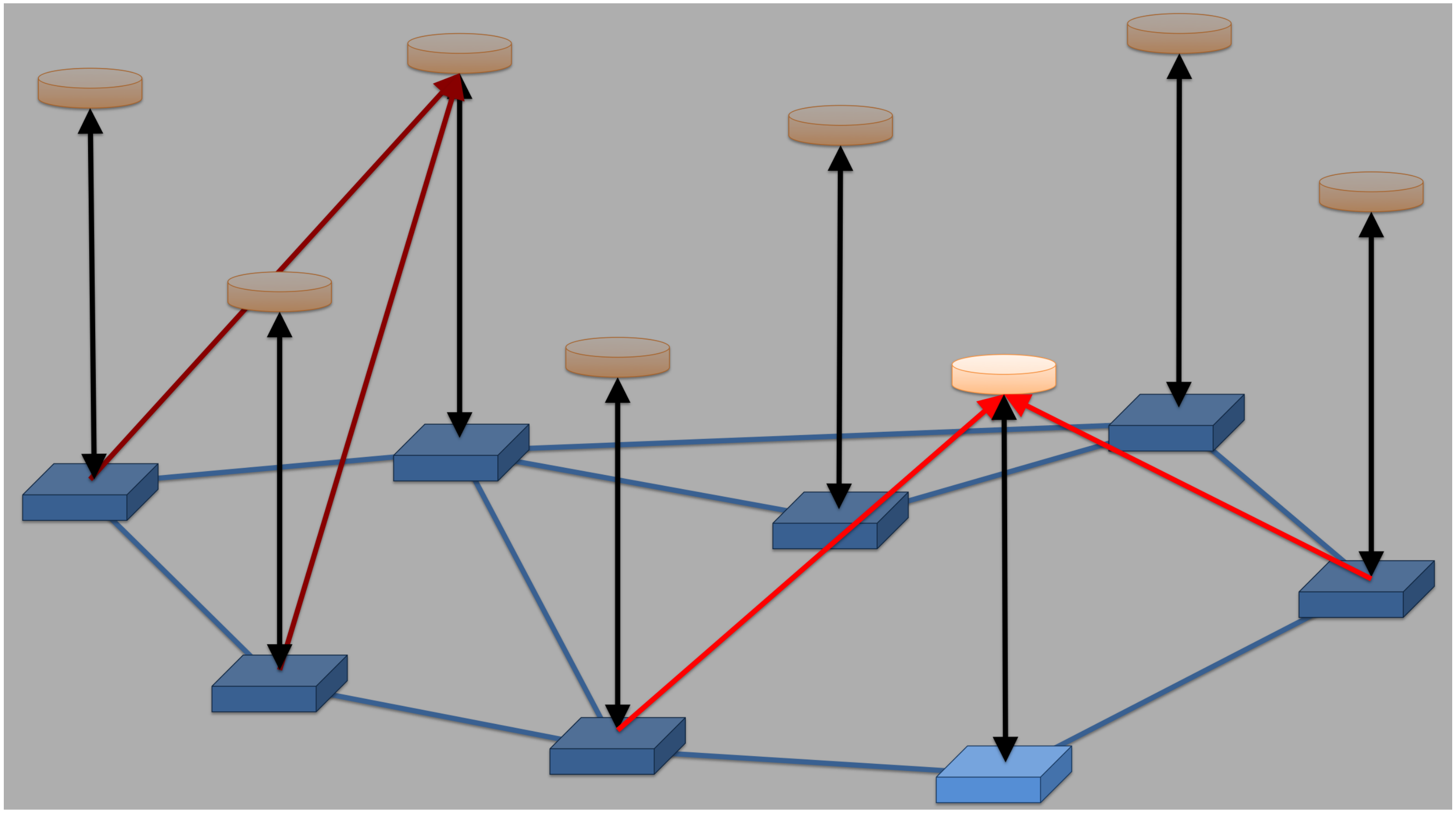
Networks of dynamical systems



- APPLICATIONS:

power grid	sensor networks	UAV formations satellite constellations
		

KEY ISSUE: Interplay between network topology and node dynamics



- Structure identification — Who should communicate with whom?
- Optimal design — How to process information from other subsystems?

- OBJECTIVE:
  - ⋆ A systematic method for structure identification and optimal design

- CHALLENGES:
  - ⋆ Networks – combinatorial objects
  - ⋆ Optimization – constrained non-convex problems

- APPROACH:
  - ⋆ Identify classes of convex problems
  - ⋆ Exploit problem structure to develop efficient algorithms
  - ⋆ Tools from optimization, control theory, compressive sensing, graph theory

Outline

- SPARSITY-PROMOTING OPTIMAL CONTROL
  - ⋆ Trade-off between performance and sparsity
  - ⋆ Cardinality minimization and convex relaxations
  - ⋆ Alternating direction method of multipliers
- SPARSE CONSENSUS NETWORKS
  - ⋆ A class of convex problems with many applications
- SUMMARY AND OUTLOOK

Sparsity-promoting optimal control

Minimum variance control

- Linear dynamical system

$$\dot{x}(t) \;=\; \underbrace{A\,x(t)}_{\text{state trajectory}} \;+\; \underbrace{B_1\,d(t)}_{\text{stochastic disturbance}} \;+\; \underbrace{B_2\,u(t)}_{\text{control input}}$$

- Steady-state variance

$$\lim_{t \to \infty} \; \mathcal{E}\{\underbrace{x^T(t)\,Q\,x(t)}_{\text{state deviation}} \;+\; \underbrace{u^T(t)\,R\,u(t)}_{\text{control effort}}\}$$

- Feedback gain design

$$u(t) \;=\; -\underbrace{F}_{\text{DESIGN VARIABLE}}\,x(t)$$

$$\text{minimize} \quad J \;=\; \lim_{t \to \infty} \mathcal{E}\{x^T Q\, x \;+\; u^T R\, u\}$$

$$\text{subject to} \quad \dot{x} \;=\; A\, x \;+\; B_1\, d \;+\; B_2\, u$$

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

- Globally optimal controller

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

centralized controller: $F_c \;=\; R^{-1} B_2^T P$

Structured minimum variance control

$$\text{minimize} \quad J = \lim_{t \to \infty} \mathcal{E}\{x^T Q x + u^T R u\}$$

$$\text{subject to} \quad \dot{x} = A x + B_1 d + B_2 u$$

$$u = -F x$$

$$F \in \mathcal{S}, \qquad F \text{ -- stabilizing}$$

centralized

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

localized

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

fully decentralized

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

How to find a stabilizing $F \in \mathcal{S}$? **(difficult problem)**

• Related work

⋆ Identify tractable/convex problems

Bamieh, Paganini, Dahleh, Voulgaris, Rotkowitz, Lall, Borrelli, Keviczky, Rantzer, . . .

⋆ Nonlinear program approach

Overton, Burke, Lewis, Apkarian, Noll, Wenk, Knapp, Toivonen, Makila, Rautert, Sachs,. . .

⋆ LMI-based approach

Geromel, El Ghaoui, Skelton, Peres, de Oliveira, Iwasaki, Gumussoy, Henrion, . . .

How to identify good sparsity structure?

Sparsity-promoting optimal control

$$\text{minimize} \quad J(F) \quad + \quad \gamma \; \mathrm{card}(F)$$

$\uparrow$ variance amplification

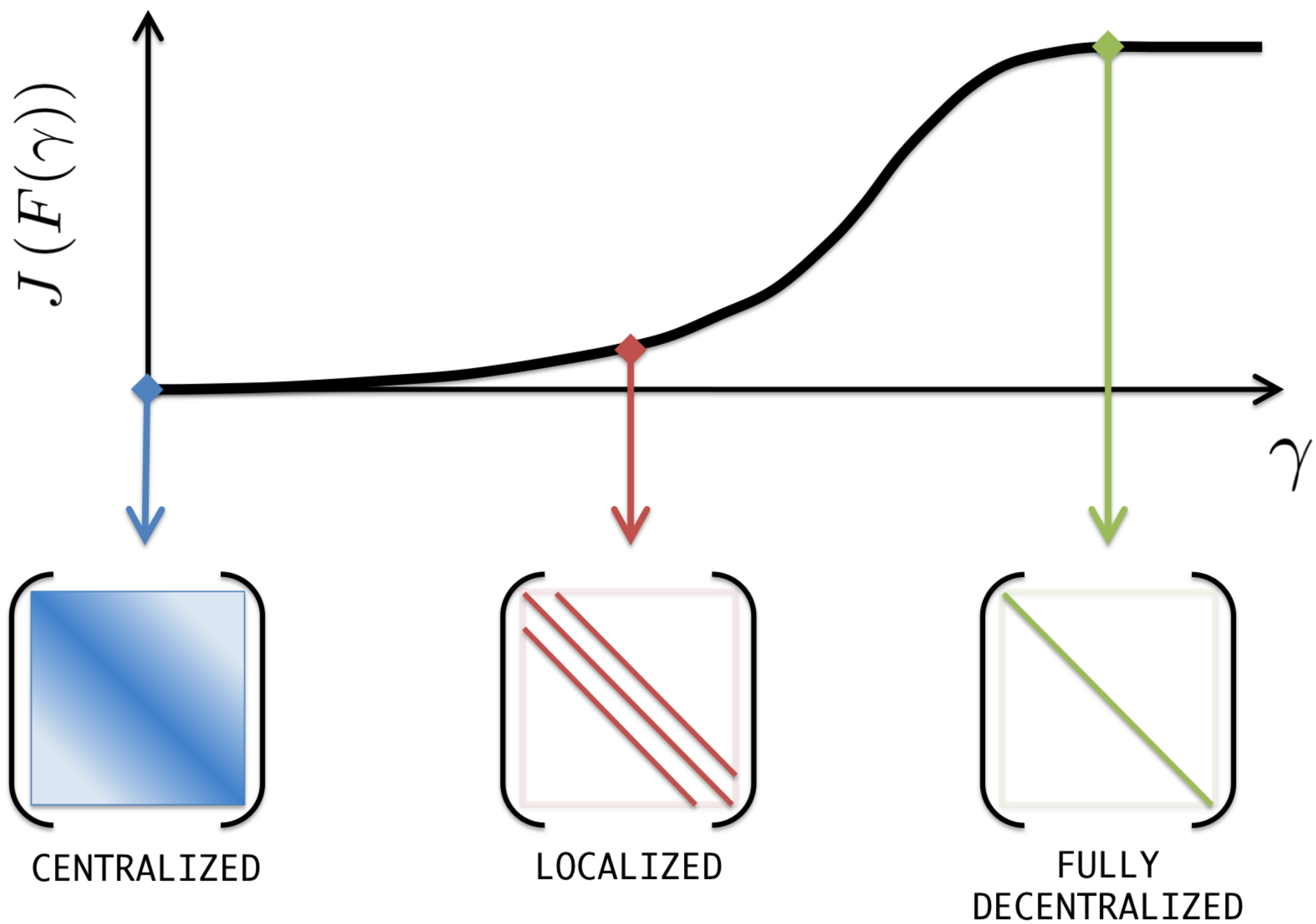
$\uparrow$ number of nonzero elements

$$F = \begin{bmatrix} 5.1 & -2.3 & 0 \\ 0 & 0 & 1.6 \end{bmatrix} \Rightarrow \mathbf{card}\,(F) = 3$$

$\gamma > 0$ – performance vs. sparsity trade-off

- A parameterized family of feedback gains

$$F(\gamma) \;:=\; \arg\min_{F} \left(J(F) \;+\; \gamma \, \mathbf{card}\,(F) \right)$$



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

$$J(F) \; = \; \mathrm{trace} \left(\int_0^\infty B_1^T \mathrm{e}^{(A-BF)^T t} \left(Q \, + \, F^T R F \right) \mathrm{e}^{(A-BF)t} \, B_1 \, \mathrm{d}t \right)$$

- DIFFICULTIES
  - ⋆ $J(F)$: non-convex function
  - ⋆ $\mathbf{card}\,(F)$: non-convex discontinuous function

- APPROACH
  - ⋆ Identify convex problems
  - ⋆ Convex relaxations of $\mathbf{card}\,(F)$

Lin, Fardad, Jovanović, IEEE TAC '13

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

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

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

- Cardinality vs. weighted ℓ_1 norm

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

RE-WEIGHTED SCHEME

⋆ **Use feedback gains from previous iteration to form weights**

$$W_{ij}^{+} \;=\; \frac{1}{|F_{ij}| \;+\; \varepsilon}, \qquad 0 \;<\; \varepsilon \;\ll\; 1$$

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

Sparsity-promoting penalty functions

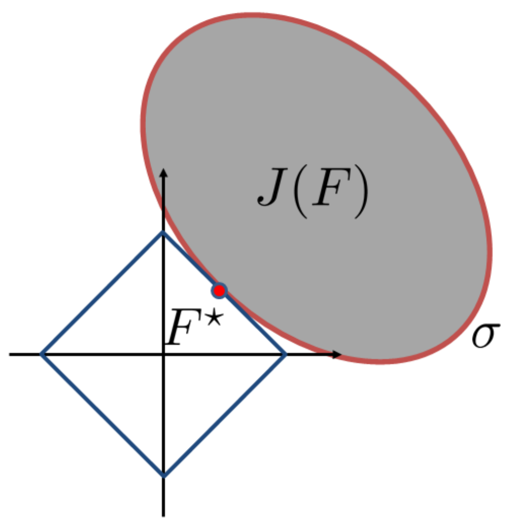
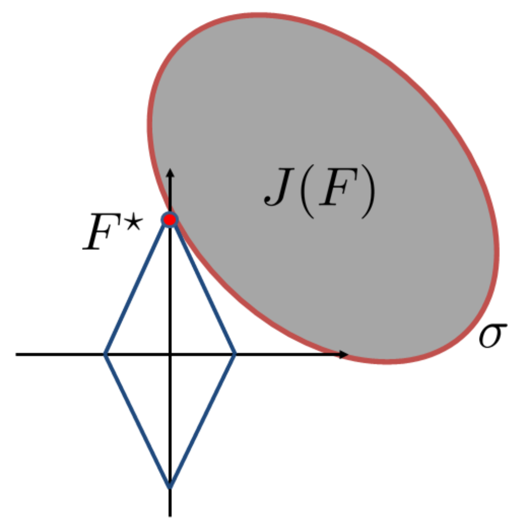
original problem:

minimize	$\mathbf{card}\,(F)$
subject to	$J(F) \le \sigma$

$\Rightarrow$

relaxation:

minimize	$g(F)$
subject to	$J(F) \le \sigma$

ℓ_1	**weighted** ℓ_1
	

A sparsity-promoting optimal control framework

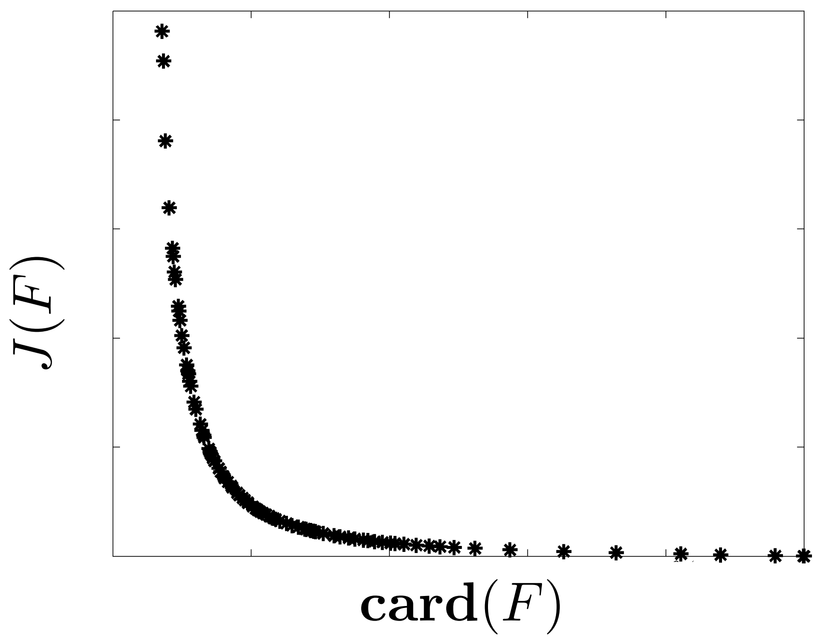
- SPARSITY STRUCTURE IDENTIFICATION

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

$J(F)$ ↑ variance amplification

$\gamma\, g(F)$ ↑ convex relaxation of cardinality

trade-off curve $\gamma > 0$



Problem structure

$$\text{minimize} \quad J(F) \; + \; \gamma \sum_{i,\,j} W_{ij} |F_{ij}|$$

- $J(F)$ – non-convex but smooth

$$J(F) \;=\; \text{trace} \left(\int_0^\infty B_1^T \mathrm{e}^{(A - B_2 F)^T t} \left(Q \;+\; F^T R F \right) \mathrm{e}^{(A - B_2 F) t} B_1 \, \mathrm{d}t \right)$$

- weighted ℓ_1 – convex but non-differentiable

$$g(F) \;=\; \sum_{i,\,j} W_{ij} |F_{ij}|$$

- $J(F) + \gamma \, g(F)$ – non-convex non-differentiable

Alternating direction method of multipliers

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

- **Step 1:** introduce 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

$$\mathcal{L}_\rho(F, G, \Lambda) \;=\; J(F) \;+\; \gamma\, g(G) \;+\; \operatorname{trace}\left(\Lambda^T (F \;-\; G)\right) \;+\; \frac{\rho}{2}\, \|F \;-\; G\|_F^2$$

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

$$\mathcal{L}_\rho(F, G, \Lambda) \; = \; J(F) \, + \, \gamma \, g(G) \, + \, \mathrm{trace} \left(\Lambda^T (F - G) \right) \, + \, \frac{\rho}{2} \, \| F - G \|_F^2$$

ADMM:

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

Many modern applications

Parallel implementation

Boyd et al., Foundations and Trends in Machine Learning '11

Separability of G-minimization problem

$$\mathcal{L}_\rho(F, G, \Lambda) \;=\; J(F) \,+\, \gamma \sum_{i,\,j} W_{ij} |G_{ij}| \,+\, \mathrm{trace}\left(\Lambda^T (F - G)\right) \,+\, \frac{\rho}{2} \|F - G\|_F^2$$

$$\underset{G}{\text{minimize}} \quad \gamma \sum_{i,\,j} W_{ij} |G_{ij}| \,+\, \frac{\rho}{2} \|G \,-\, V\|_F^2$$

$$V \;:=\; (1/\rho)\Lambda^k \,+\, F^{k+1}$$

- **Element-wise separable**

$$\underset{G_{ij}}{\text{minimize}} \quad \sum_{i,\,j} \left(\gamma\, W_{ij}\, |G_{ij}| \,+\, \frac{\rho}{2} (G_{ij} \,-\, V_{ij})^2\right)$$

Solution to F-minimization problem

$$\underset{F}{\text{minimize}} \quad J(F) + \frac{\rho}{2} \|F - U\|_F^2$$

$$U := G^k - (1/\rho)\Lambda^k$$

NECESSARY CONDITIONS FOR OPTIMALITY:

$$\begin{aligned} (A - B_2 F) L + L (A - B_2 F)^T &= - B_1 B_1^T \\ (A - B_2 F)^T P + P (A - B_2 F) &= -(Q + F^T R F) \\ F L + \rho (2R)^{-1} F &= R^{-1} B_2^T P L + \rho (2R)^{-1} U \end{aligned}$$

- ITERATIVE SCHEME

Given F_0 solve for $\{L_1, P_1\} \rightarrow F_1 \rightarrow \{L_2, P_2\} \rightarrow F_2 \cdots$

descent direction $+$ **line search** $\Rightarrow$ **convergence**

Lin, Fardad, Jovanović, IEEE TAC '13

Structured optimal design

$\star$ ADMM $\begin{cases} \textbf{identifies sparsity patterns} \\ \textbf{provides good initial condition for structured design} \end{cases}$

$$
\boxed{
\begin{array}{rcl}
(A - B_2 F)^T P + P (A - B_2 F) & = & -\left(Q + F^T R F\right) \\
(A - B_2 F) L + L (A - B_2 F)^T & = & -B_1 B_1^T \\
\left[\left(R F - B_2^T P\right) L\right] \circ I_{\mathcal{S}} & = & 0
\end{array}
}
$$

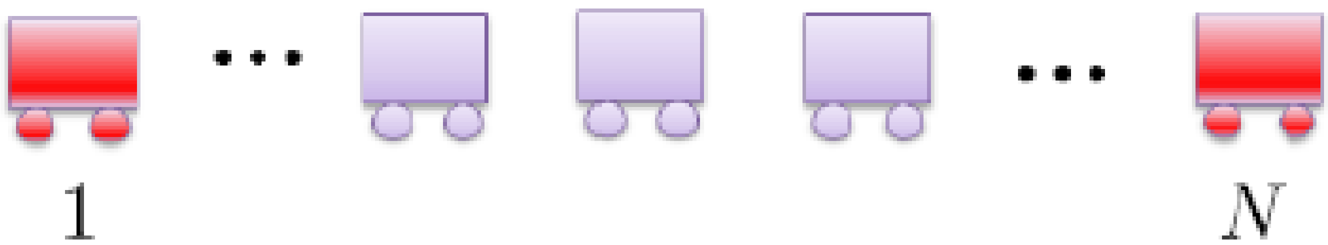
Newton's method + conjugate gradient

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

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

Examples

- FORMATION OF VEHICLES



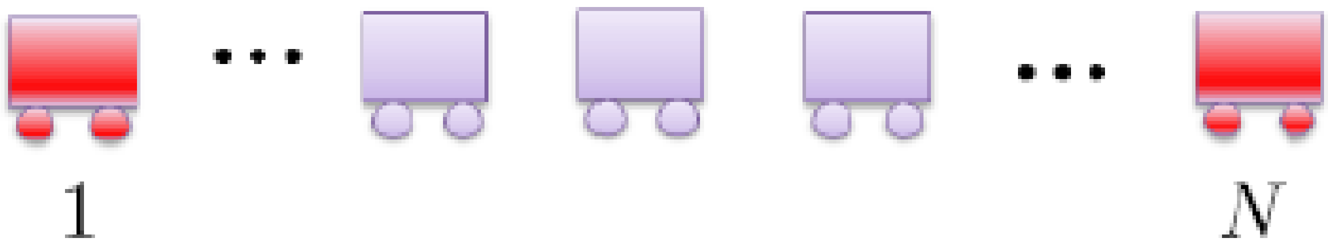
⋆ Each vehicle: relative information exchange

$$u_i = -\sum_{j \neq i} F_{ij} (x_i - x_j), \qquad i \in \{2, \ldots, N-1\}$$

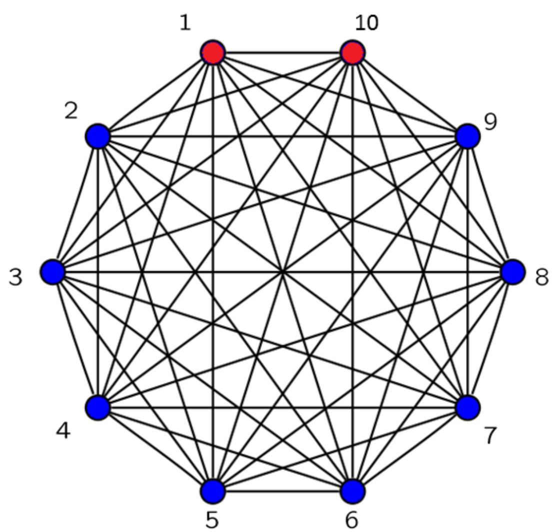
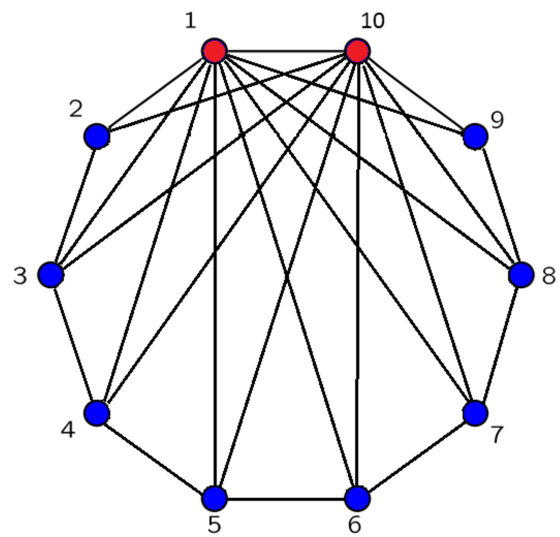
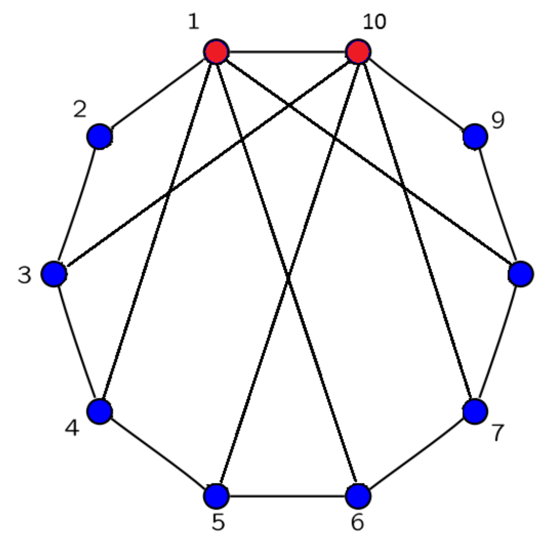
⋆ Leaders: equipped with GPS devices

$$u_1 = -\sum_{j \neq 1} F_{1j} (x_1 - x_j) - F_{11} x_1$$

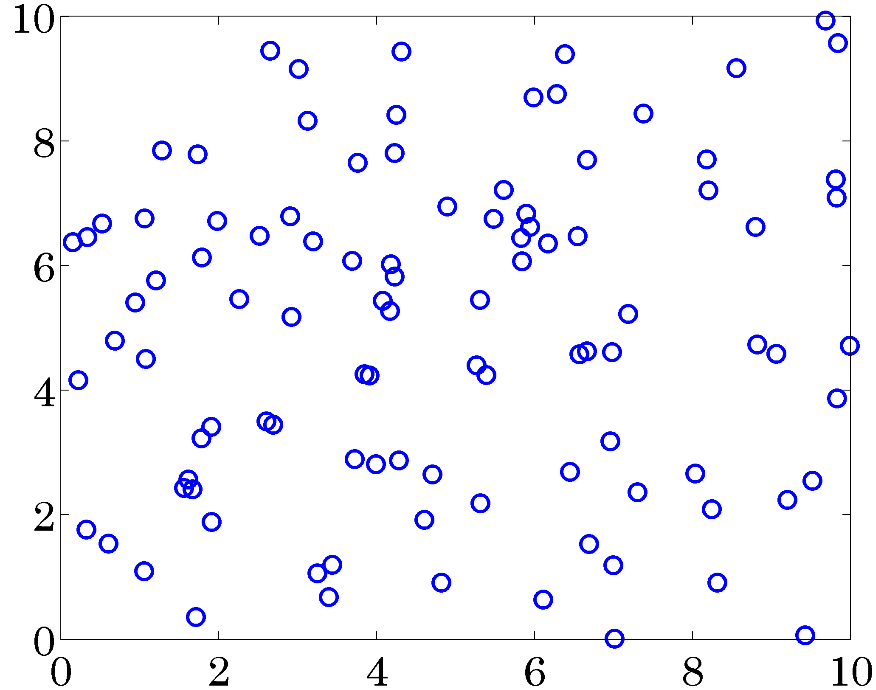
$$u_N = -\sum_{j \neq N} F_{Nj} (x_N - x_j) - F_{NN} x_N$$



IDENTIFIED COMMUNICATION ARCHITECTURES:

$\gamma = 0$	$\gamma = 0.01$	$\gamma = 0.05$
		
all-to-all	**nearest neighbors** + **leaders-to-all**	**nearest neighbors** + **leaders-to-some**

- A network with $N = 100$ **unstable** subsystems

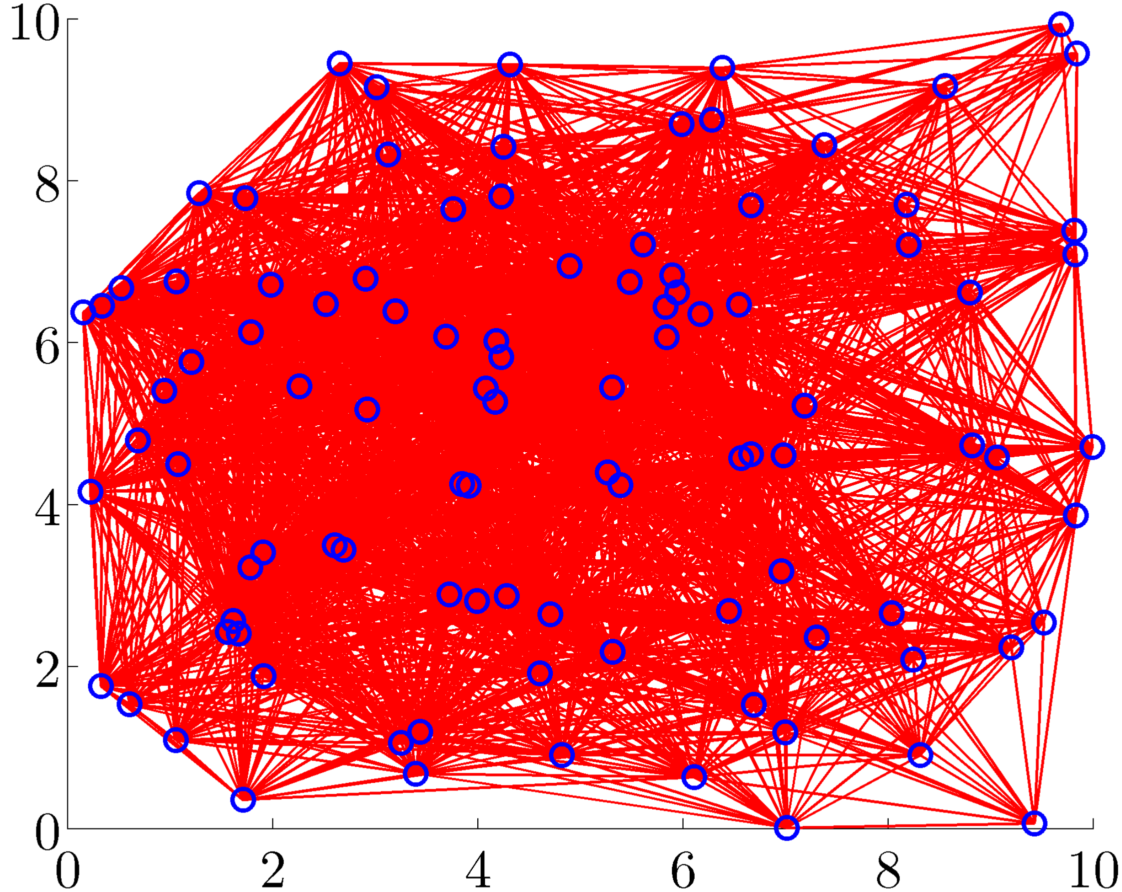


$$\begin{bmatrix} \dot{p}_i \\ \dot{v}_i \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} p_i \\ v_i \end{bmatrix}}_{\text{unstable dynamics}} + \underbrace{\sum_{j \neq i} e^{-\alpha(i,j)} \begin{bmatrix} p_j \\ v_j \end{bmatrix}}_{\text{coupling}} + \begin{bmatrix}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

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

Motee & Jadbabaie, IEEE TAC '08

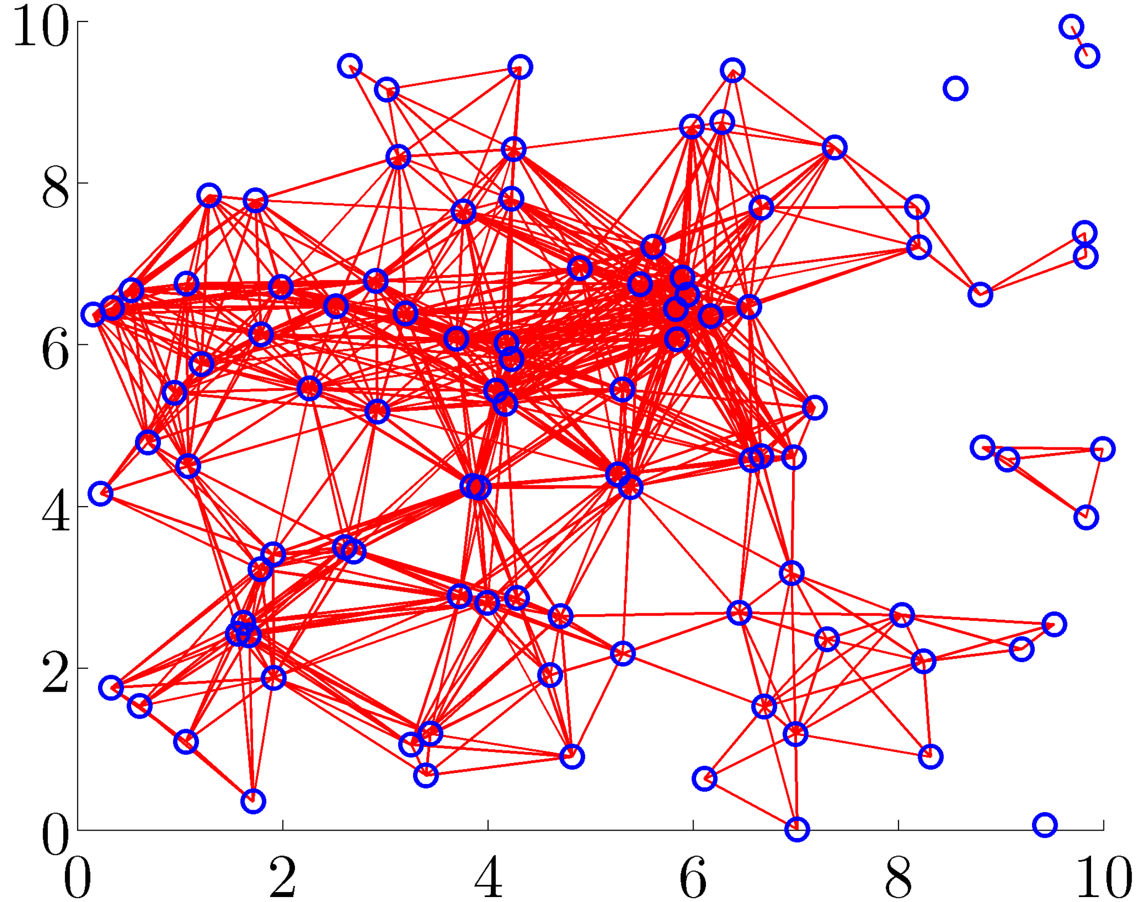
communication graph of truncated centralized gain:



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

non-stabilizing

communication graph identified via ADMM:

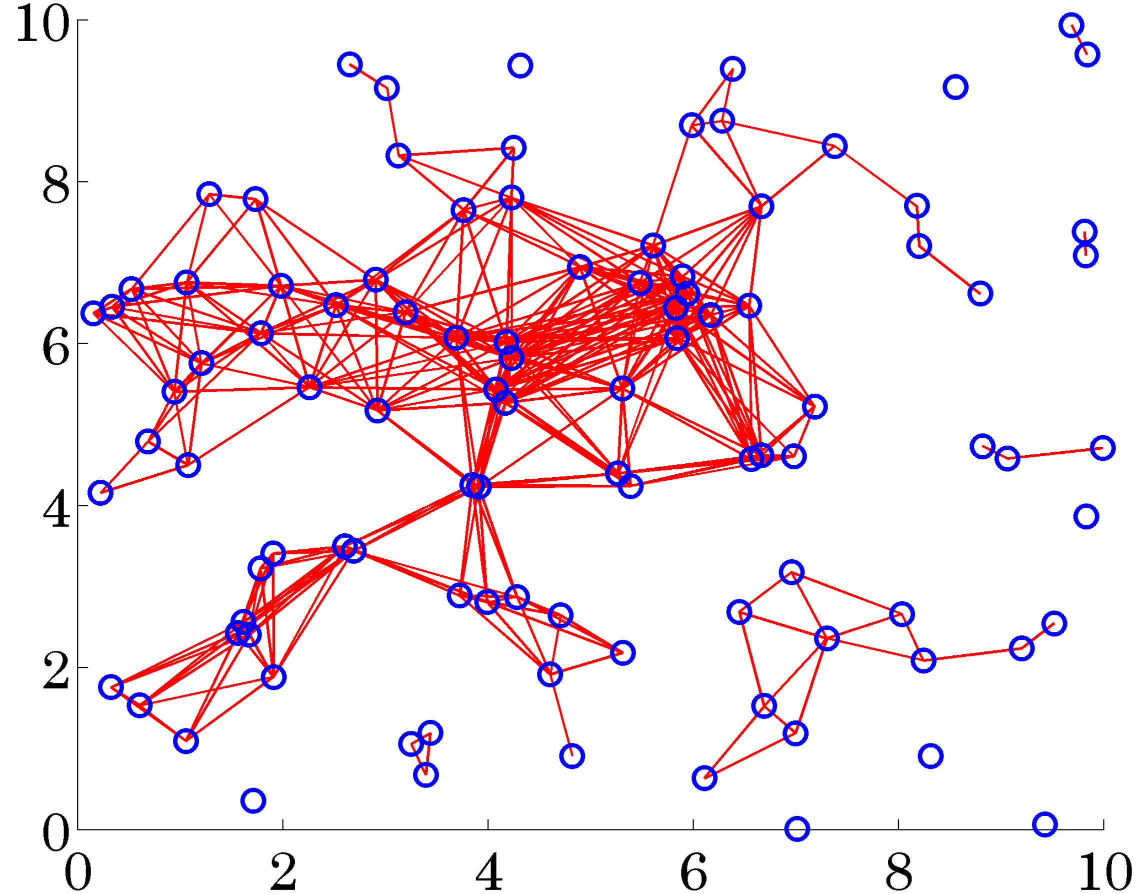


$$\gamma = 5$$

$$\mathbf{card}(F)/\mathbf{card}(F_c) = 8.8\%$$

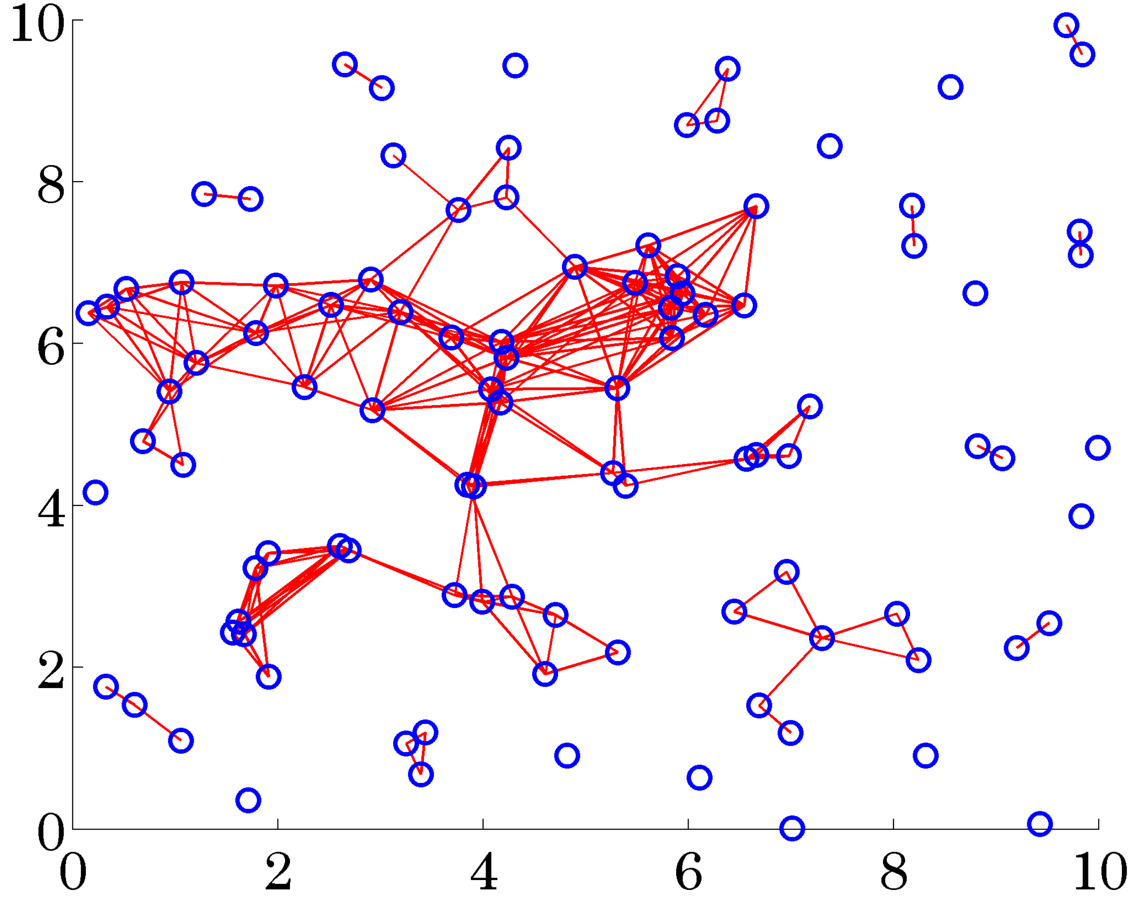
$$(J - J_c)/J_c = 24.6\%$$

communication graph identified via ADMM:



$$\begin{aligned} \gamma &= 11 \\ \mathbf{card}\,(F)\,/\mathbf{card}\,(F_c) &= 5.1\% \\ (J - J_c)\,/J_c &= 40.9\% \end{aligned}$$

communication graph identified via ADMM:

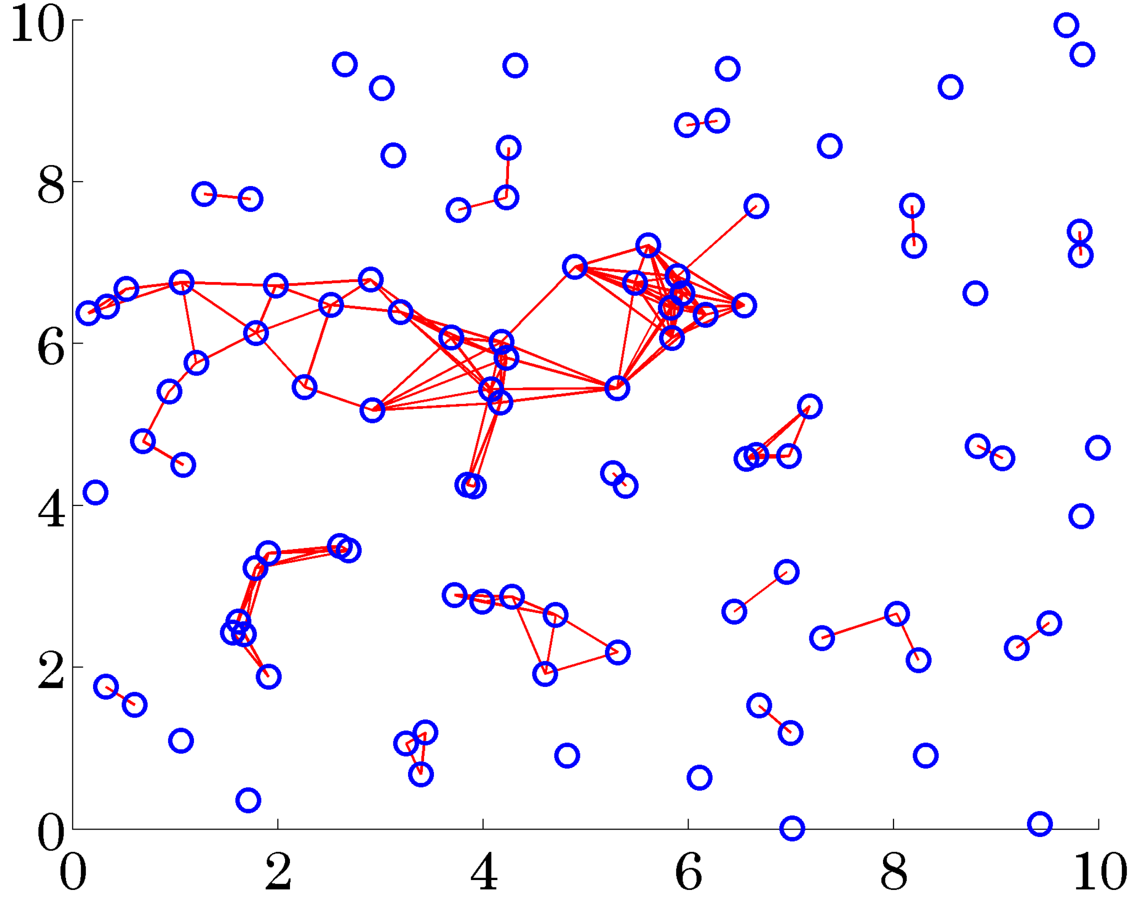


$$\gamma = 18$$

$$\mathbf{card}\,(F)\,/\mathbf{card}\,(F_c) = 3.4\%$$

$$(J - J_c)\,/J_c = 48.7\%$$

communication graph identified via ADMM:

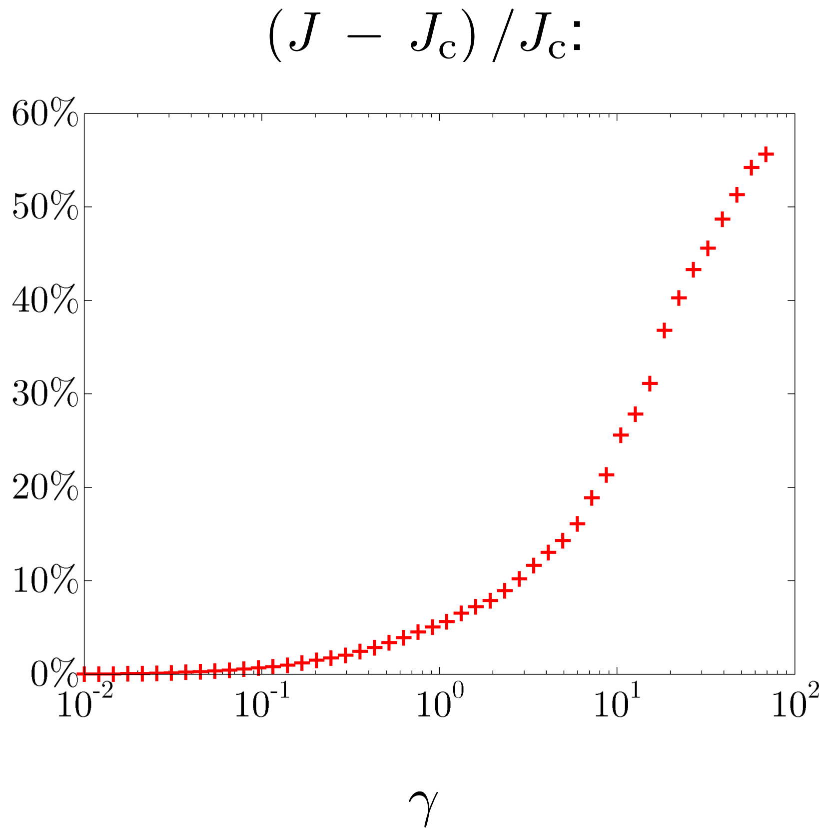


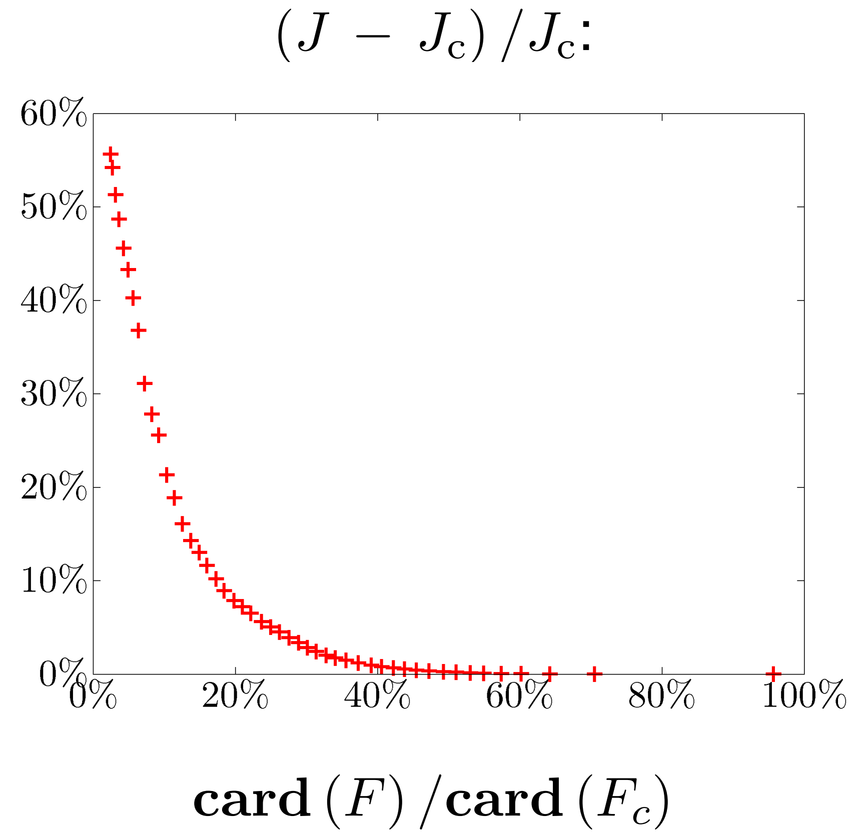
$$\gamma = 30$$

$$\mathbf{card}\,(F)\,/\mathbf{card}\,(F_c) = 2.4\%$$

$$(J - J_c)\,/J_c = 54.8\%$$

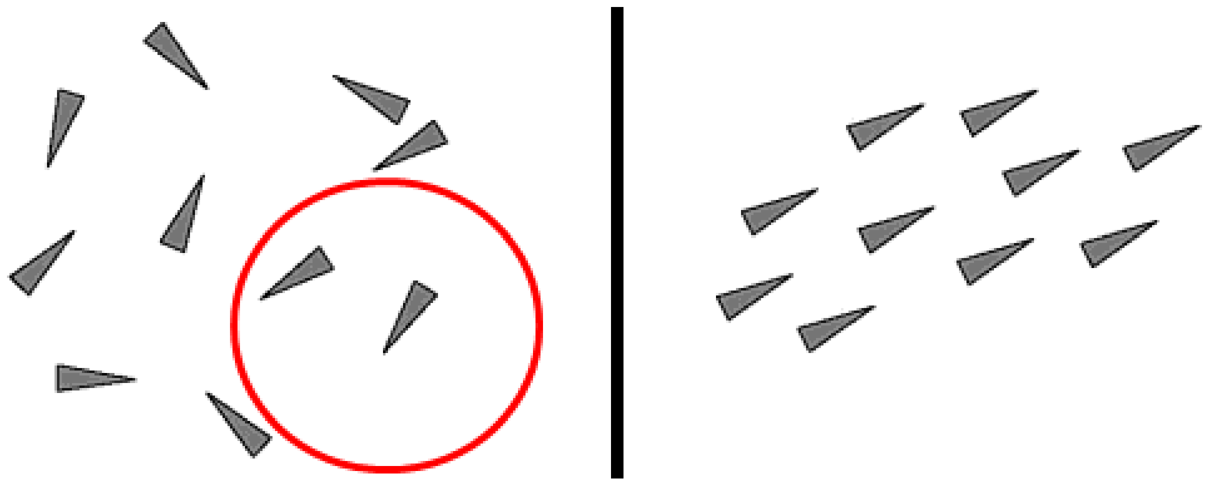
Performance comparison: sparse vs. centralized





Sparse consensus networks

Average consensus



- Update using relative differences with neighbors

$$\dot{x}_i(t) = - \sum_{j \in \mathcal{N}_i} F_{ij} \left(x_i(t) - x_j(t) \right)$$

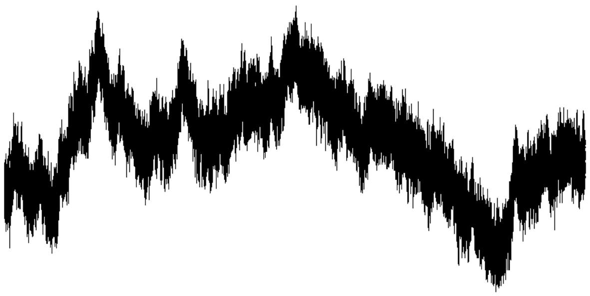
Reaching average consensus: $$\lim_{t \to \infty} x_i(t) = \frac{1}{N} \sum_{i=1}^{N} x_i(0)$$

distributed estimation, synchronization of oscillators, load balancing, . . .

Consensus with stochastic disturbances

$$\dot{x}_i(t) \;=\; - \sum_{j \,\in\, \mathcal{N}_i} F_{ij}(x_i(t) \;-\; x_j(t)) \;+\; d_i(t) \quad \leftarrow \quad \text{disturbance}$$

Average mode $\frac{1}{N} \sum_{i\,=\,1}^{N} x_i(t)$ undergoes a random walk



If other modes are stable, $x_i(t)$ fluctuates around average

- Deviation from average: $\tilde{x}(t) \;=\; x(t) \;-\; x_{\text{average}}(t)$

Variance amplification $\lim_{t \,\to\, \infty} \mathcal{E} \left\{ \tilde{x}^T(t)\, \tilde{x}(t) \right\}$ quantifies consensus performance

Structure identification step: SDP formulation

$$\text{minimize} \quad J(F) + \gamma \sum_{i,j} W_{ij}|F_{ij}|$$

$$\text{subject to} \quad F = F^T \qquad F + \frac{1}{N}\mathbb{1}\mathbb{1}^T \succ 0 \qquad F\mathbb{1} = 0$$

F **– Laplacian matrix of undirected graphs**

SDP formulation:

$$\underset{X,\,Y,\,F}{\text{minimize}} \quad \text{trace}\,(X + F) + \gamma \sum_{i,j} Y_{ij}$$

$$\text{subject to} \quad \begin{bmatrix} X & Q^{1/2} \\ Q^{1/2} & F + \frac{1}{N}\mathbb{1}\mathbb{1}^T \end{bmatrix} \succeq 0$$

$$F = F^T \qquad F\mathbb{1} = 0$$

$$-Y_{ij} \le W_{ij}F_{ij} \le Y_{ij}$$

Lin, Fardad, Jovanović, Allerton 2012

Optimal design step: SDP formulation

$$\begin{array}{ll} \text{minimize} & J(F) \\ \text{subject to} & F = F^T \quad F + \frac{1}{N}\mathbb{1}\mathbb{1}^T \succ 0 \quad F\mathbb{1} = 0 \quad F \in \mathcal{S} \end{array}$$

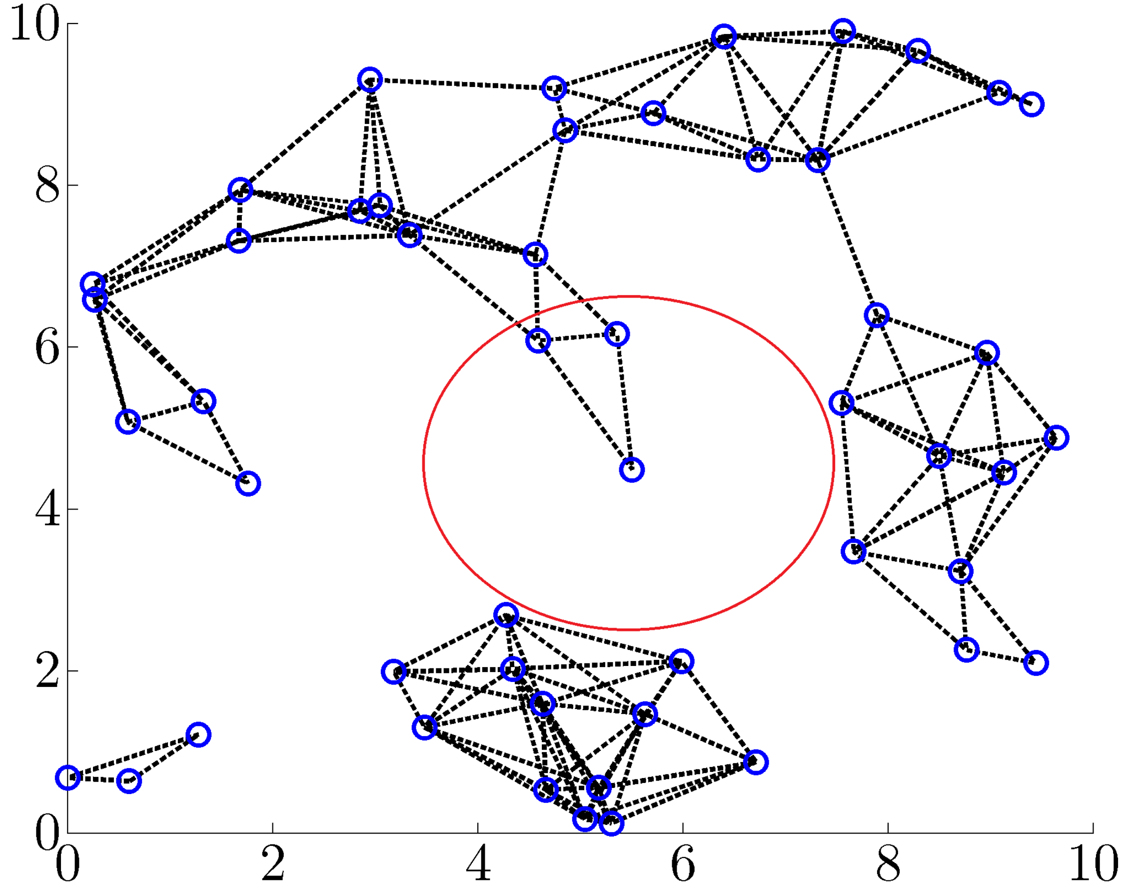
SDP formulation:

$$\begin{array}{ll} \underset{X,\,F}{\text{minimize}} & \operatorname{trace}(X + F) \\ \text{subject to} & \begin{bmatrix} X & Q^{1/2} \\ Q^{1/2} & F + \frac{1}{N}\mathbb{1}\mathbb{1}^T \end{bmatrix} \succeq 0 \\ & F = F^T \quad F\mathbb{1} = 0 \quad F \circ I_{\mathcal{S}} = F \end{array}$$

Customized algorithm – exploiting structure of graph Laplacian matrix

Lin, Fardad, Jovanović, IEEE CDC 2010

An example



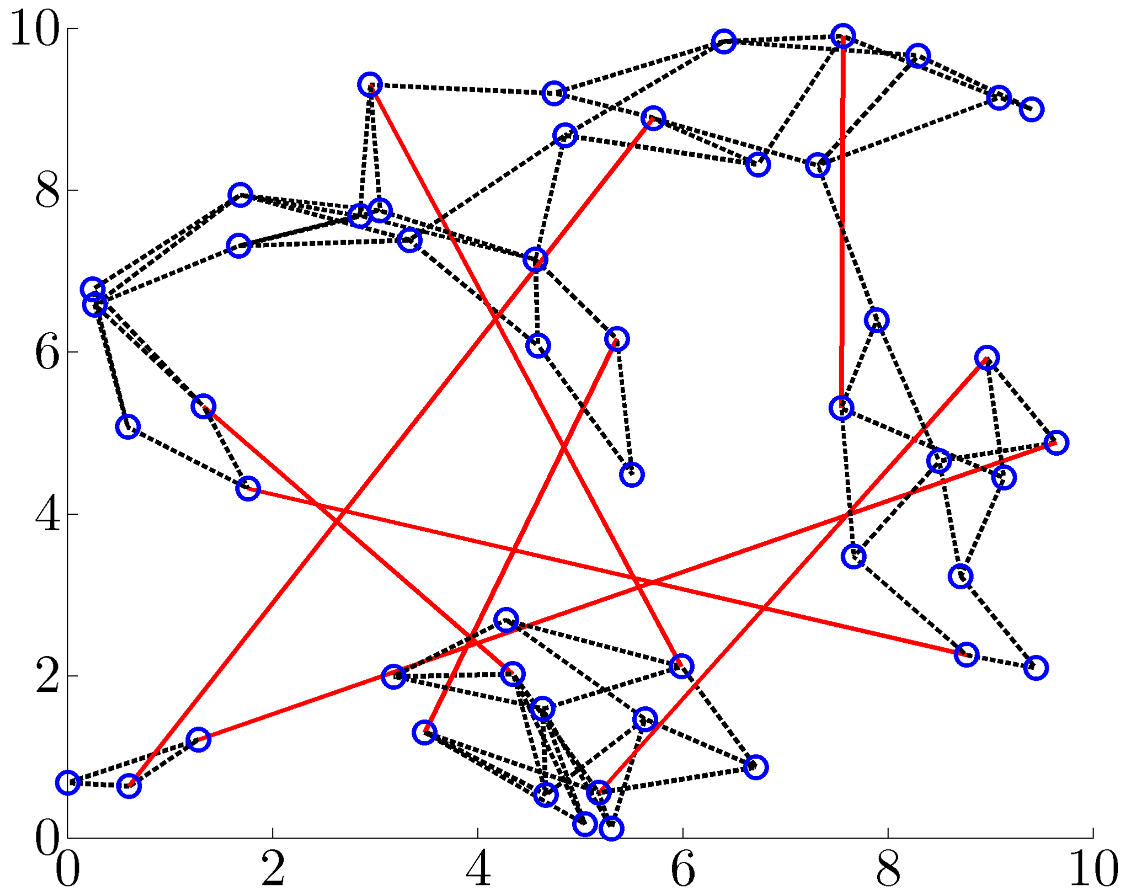
$$J = \lim_{t \to \infty} \mathcal{E}\{ \tilde{x}^T(t)\,\tilde{x}(t) + \tilde{u}^T(t)\,\tilde{u}(t) + x^T(t)\,Q_{\mathrm{local}}\,x(t) \}$$

$\uparrow$ deviation from average

$\uparrow$ control effort

$\uparrow$ deviation from local neighbors

- An identified communication network



local interactions + long-range links

$$\begin{aligned} \mathbf{card}\,(F)\,/\mathbf{card}\,(F_{\rm c}) &= 7\% \\ (J - J_{\rm c})\,/J_{\rm c} &= 14\% \end{aligned}$$

SUMMARY AND OUTLOOK

Summary

- SPARSITY-PROMOTING OPTIMAL CONTROL PROBLEM
  - ⋆ Performance vs. sparsity trade-off
  - ⋆ Cardinality minimization and convex relaxations
  - ⋆ Alternating direction method of multipliers

Lin, Fardad, Jovanović, IEEE TAC '13

- SOFTWARE
  - ⋆ www.ece.umn.edu/∼mihailo/software/lqrsp/

```
>> solpath = lqrsp(A,B1,B2,Q,R,options);
```

- SPARSE CONSENSUS NETWORKS
  - ⋆ A class of convex problems – SDP formulations

- RELATED WORK

  ⋆ Optimal synchronization of sparse oscillator networks

  Fardad, Lin, Jovanović, IEEE TAC '13 (submitted)

  ⋆ Optimal dissemination of information in social networks

  Fardad, Zhang, Lin, Jovanović, CDC '12

  ⋆ Sparse and optimal wide-area damping control in power networks

  Dorfler, Jovanović, Chertkov, and F. Bullo, ACC '13 (to appear)

- FUTURE WORK

  ⋆ Distributed implementation of ADMM

Northeast blackout 2003

BEFORE:

AFTER:

- Caused by a SINGLE power plant (at Cleveland) going offline

Removal of key nodes can affect performance and survival of networks

Leader selection in dynamic networks

Choose leaders to minimize deviation from operating point

- **Challenge:**
  - ⋆ Combinatorial optimization problem

- **Approach:**
  - ⋆ Convex relaxation $\Rightarrow$ lower bound
  - ⋆ Greedy algorithm $\Rightarrow$ upper bound

- **Contributions:**
  - ⋆ Developing customized algorithm
  - ⋆ Exploiting structure of low-rank modifications

Lin, Fardad, Jovanović, IEEE TAC '13 (conditionally accepted)

Characterization of social influence

- How to characterize social influence?

- How to identify individuals with the maximum social influence?

Fardad, Lin, Zhang, Jovanović, ACC '13 (to appear)

Matrix completion of partially known state covariances

$$
\begin{array}{ll}
\underset{Q,\, X}{\text{minimize}} & \mathbf{rank}\,(Q) \\
\text{subject to} & AX + XA^* + Q = 0 \\
& \text{trace}\,(T_i\, X) = g_i, \;\; i = 1, \ldots, N \\
& X \succeq 0
\end{array}
$$

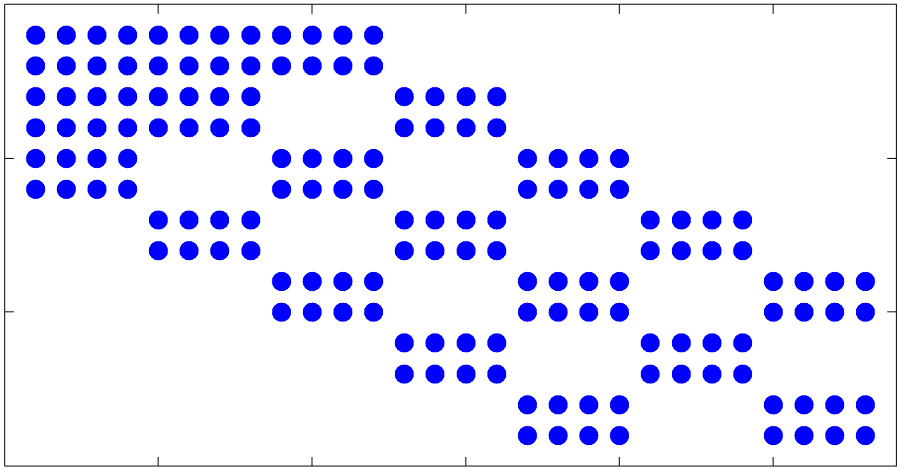
- Nuclear norm relaxation
- Efficient customized algorithm based on ADMM

Lin, Jovanović, Georgiou, IEEE CDC '13 (submitted)

THANK YOU!

ADDITIONAL SLIDES

Extension: Block sparsity



- $\mathbf{card}_{\mathrm{b}}\,(F)$ – number of non-zero blocks of F

$$\mathbf{card}_{\mathrm{b}}\,(F) \;=\; \sum_{i,\,j} \mathbf{card}\,(\|F_{ij}\|_F)$$

- PENALTY FUNCTIONS THAT PROMOTE BLOCK SPARSITY
  - ⋆ generalized ℓ_1, weighted ℓ_1, sum-of-logs