Identification of sparse representations of consensus networks

University of Minnesota

Neil Dhingra, Fu Lin, Makan Fardad, and Mihailo R. Jovanović

Syracuse University

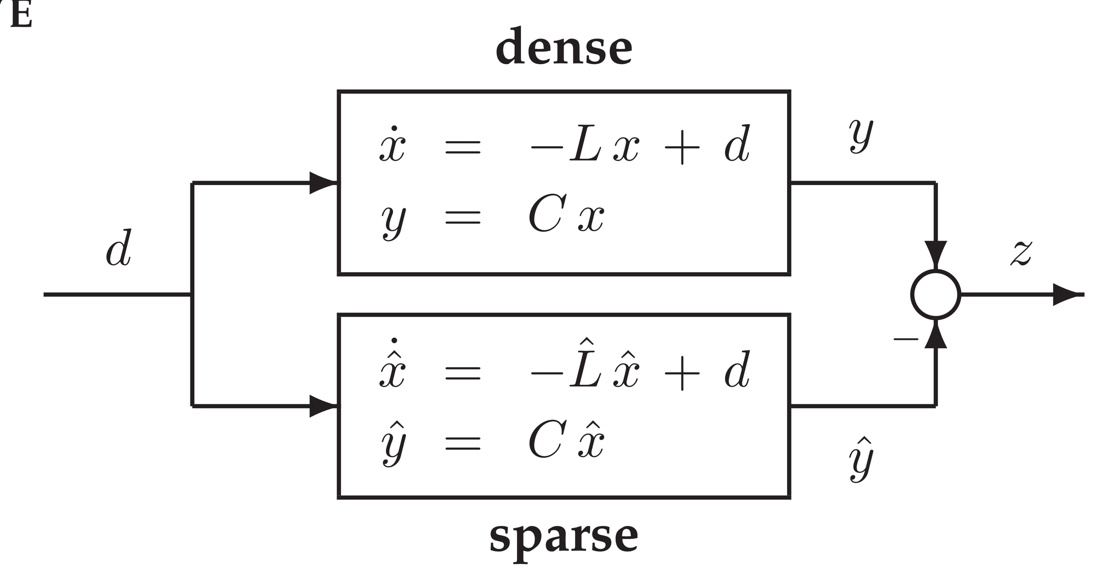
Problem Formulation

Stochastically forced undirected networks

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

$$y \;=\; \begin{bmatrix} (I - (1/n)\mathbb{1}\mathbb{1}^T)\,x \\ -L\,x \end{bmatrix} \quad \begin{array}{l}\textbf{deviation from average} \\ \textbf{control effort}\end{array}$$

Objective



- Identify **subgraph** that strikes a balance between

  - **variance amplification** $d \to z$
  - **number of edges in the subgraph**

Optimization problem

$$\underset{\hat{L}}{\text{minimize}} \quad J(\hat{L}) \quad + \quad \gamma\, \textbf{card}\,(\hat{L})$$

$J(\hat{L})$ ↓ **variance amplification**

$\gamma\,\textbf{card}\,(\hat{L})$ ↓ **sparsity-promoting penalty function**

card $(\hat{L})$ – number of non-zero elements of $\hat{L}$

$\gamma > 0$ – variance amplification vs. sparsity tradeoff

Approach

Challenges

- $J(\hat{L})$ – nonconvex; **card** $(\hat{L})$ – nonconvex and nonsmooth

Convex relaxation of cardinality

- Replace **card** with **weighted ℓ_1 norm**

$$\underset{\hat{L}}{\text{minimize}} \quad J(\hat{L}) \;+\; \gamma \sum_{i,j} W_{ij}\,|\hat{L}_{ij}|$$

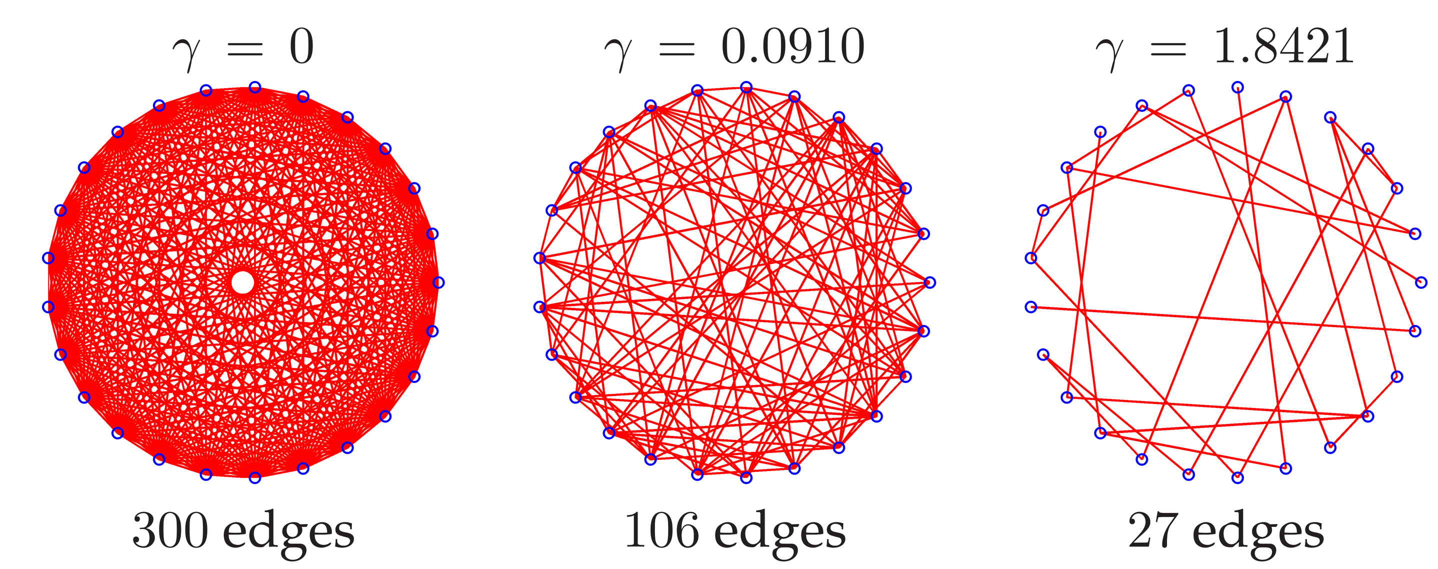
Re-weighted algorithm: $W_{ij}^{+} \;:=\; (|\hat{L}_{ij}| \,+\, \epsilon)^{-1}$

Performance vs. sparsity

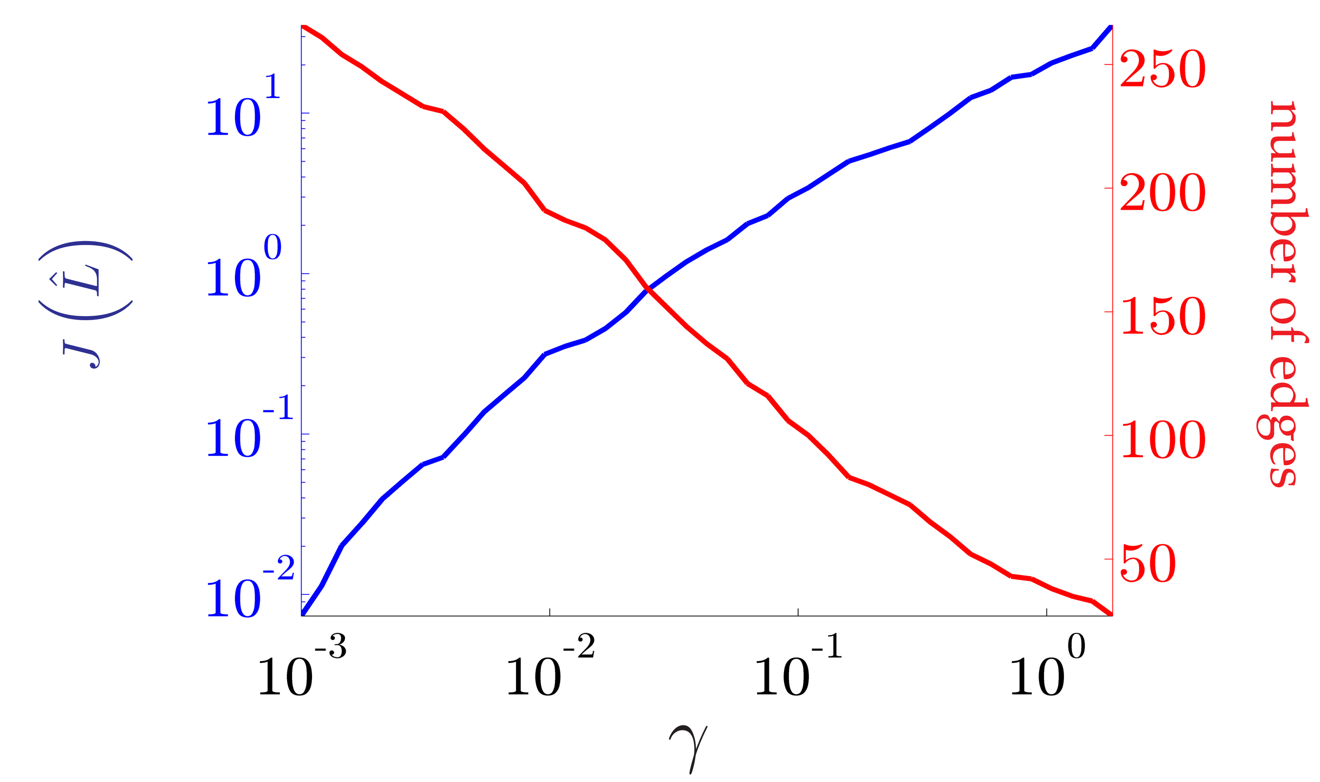
- Identify a γ-parameterized family of subgraphs
- Larger γ – subgraphs with **fewer edges** but **worse performance**

Examples

Complete graph with 25 nodes and **random edge weights**



performance vs. sparsity:



Algorithm

Alternating Direction Method of Multipliers (ADMM)

1. **Sparse Structure Identification**
   - Introduce **additional variable/constraint**

$$\text{minimize} \quad J(\hat{L}) \,+\, \gamma \sum_{i,j} W_{ij}\,|\hat{L}_{ij}|$$

$$\text{subject to} \quad \hat{L} \,-\, M \,=\, 0$$

   - Form **augmented Lagrangian**

$$\mathcal{L}_\rho \;=\; J(\hat{L}) + \gamma \sum_{i,j} W_{ij}\,|\hat{L}_{ij}| + \text{trace}\,(\Lambda^T(\hat{L} - M)) + \frac{\rho}{2}\,\|\hat{L} - M\,\|_F^2$$

   - Use **ADMM iterations** to minimize $\mathcal{L}_\rho$

$$\hat{L}^{k+1} \;=\; \underset{\hat{L}}{\text{argmin}}\; \mathcal{L}_\rho(\hat{L}, M^k, \Lambda^k) \qquad \textit{Differentiable; BFGS}$$

$$M^{k+1} \;=\; \underset{M}{\text{argmin}}\; \mathcal{L}_\rho(\hat{L}^{k+1}, M, \Lambda^k) \qquad \textit{Separable; soft-thresholding}$$

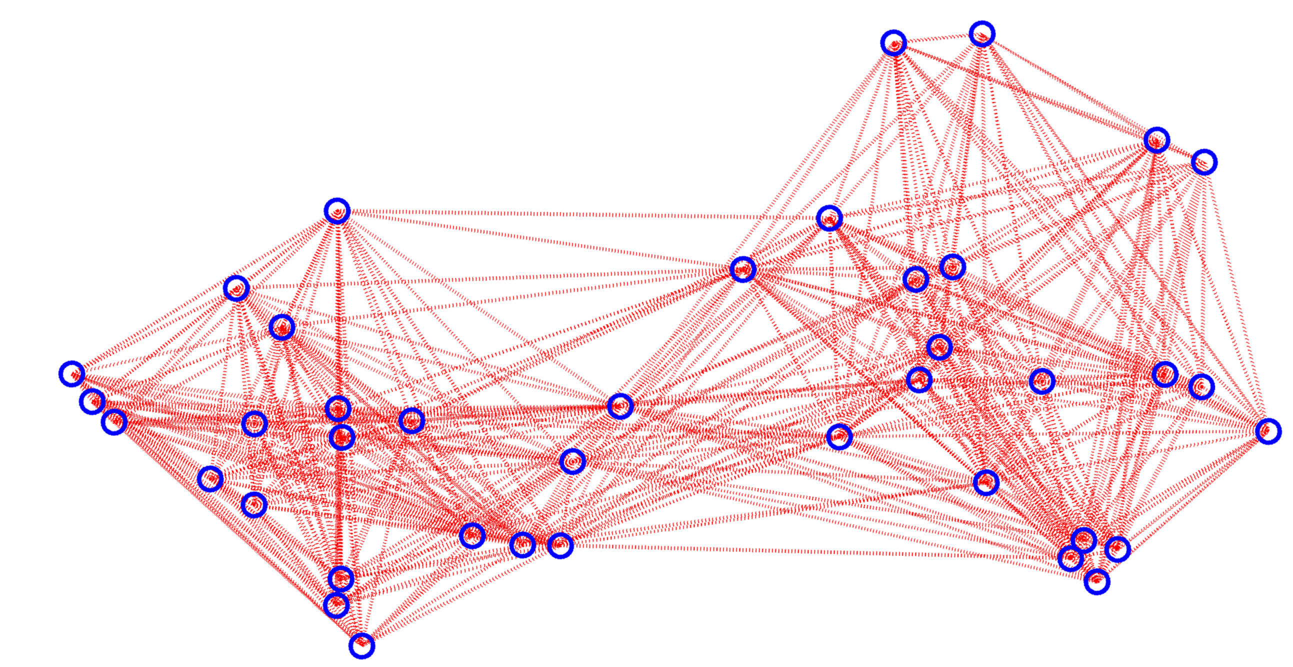
$$\Lambda^{k+1} \;=\; \Lambda^k \,+\, \rho\,(\hat{L}^{k+1} \,-\, M^{k+1})$$

2. **Polishing**
   - Find **optimal** $\hat{L}$ for **identified sparsity structure** $\hat{\mathcal{S}}$

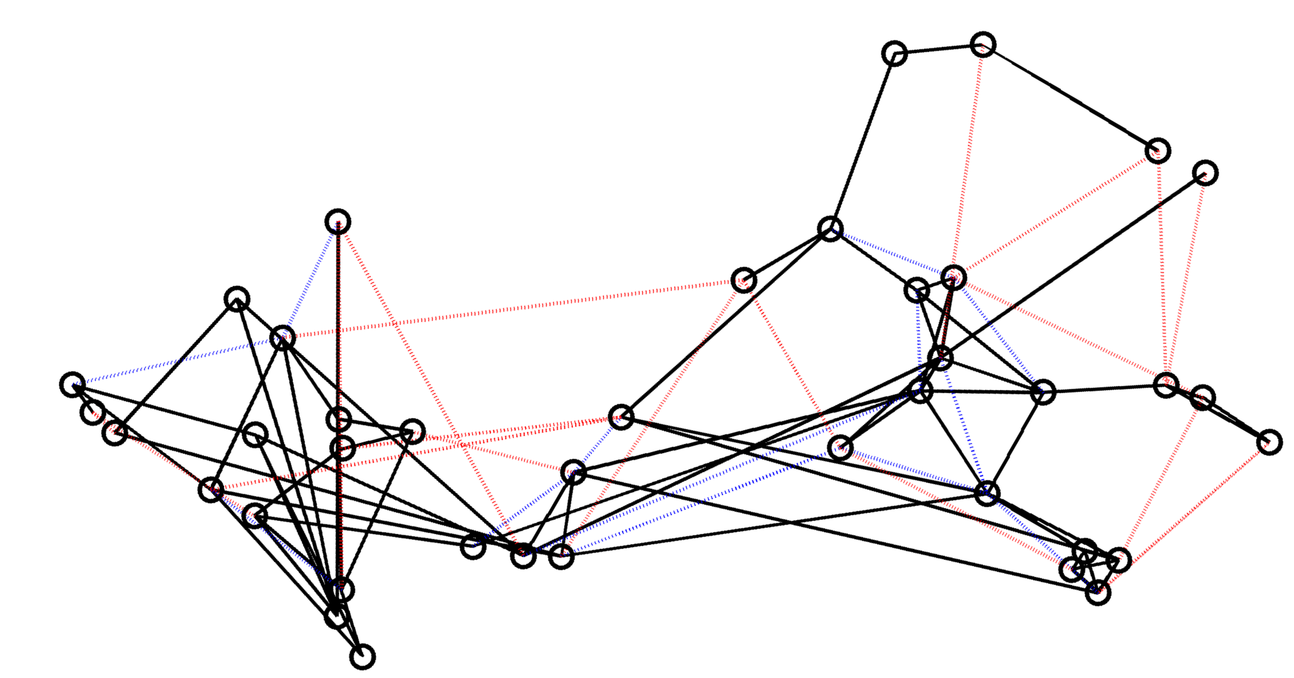
$$\text{minimize} \quad J(\hat{L}) \quad \text{subject to} \quad \hat{L} \in \hat{\mathcal{S}}$$

ADMM vs. Truncation

- Truncation – discard the smallest edge weights and polish
- 40-node, 425-edge graph with edge weights inversely related to the Euclidean distance between connected nodes



sparse subgraphs with 79 edges:



- **ADMM** and **truncation** identify { **65 common edges**; 14 different edges }

	Original	Truncation	ADMM
Number of Edges	425	79	79
Graph Diameter	3	9	6
Average Path Length	1.53	3.49	2.88
$J(\hat{L})$	0	10.39	4.10
Algebraic Connectivity	0.307	0.079	0.225

Comments

- Truncation ignores effects of increasing path lengths
- ADMM preserves long-range interactions

Acknowledgements

- NASA Harriett G. Jenkins Predoctoral Fellowship
- NSF CMMI-09-27720
- NSF CMMI-09-27509