Learning the Structure of Large Networked Systems Obeying Conservation Laws

Anirudh Rayas Rajsekhar Anguluri Gautam Dasarathy

Arizona State University

Paper:

Motivation

- Conservation laws are fundamental laws which govern the evolution of various processes in nature. Examples include conservation of charge, energy, momentum etc.
- In this paper we focus on conservation laws in networked physical systems.
- For instance in electric networks, the dynamics of flows (currents) are governed by a conservation law (Kirchoff's law).
- This phenomenon can be observed in conceptual networks such as brain and social networks.
- The dynamics in general can be described by a balance equation of the form $X = B^*Y$, where X, Y are the vectors of injected flows and node potentials and B^* is the graph Laplacian, which captures the connectivity structure of the network.

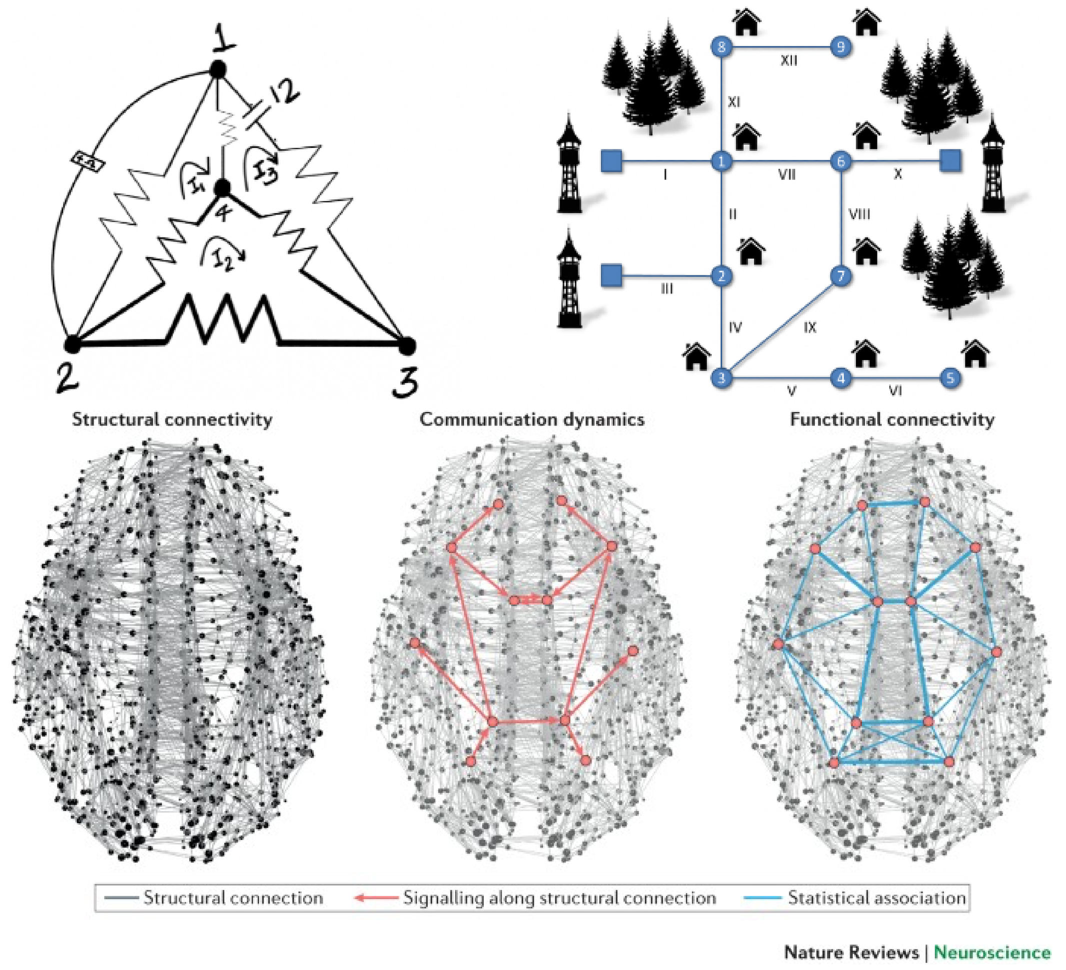


Figure 1. Electric networks, Water networks, Brain networks

Problem Setup

Let $B^* \succ 0$, then we can rewrite the balance equation as

$$Y = B^{*-1}X \tag{1}$$

where $Y, X \in \mathbb{R}^p$ are the vectors of node potentials and injected flows respectively and $B^* \in \mathbb{R}^{p\times p}$ is a sparse positive definite matrix that captures the connectivity structure of the network.

Goal: Learn the structure of the network given samples from observation vector Y and access to only the statistics of X ie. Σ_X.

We propose an ℓ_1- regularized Maximum Likelihood Estimator(MLE) $\widehat{B}$ to infer the sparsity structure of B^* in the high dimensional regime.

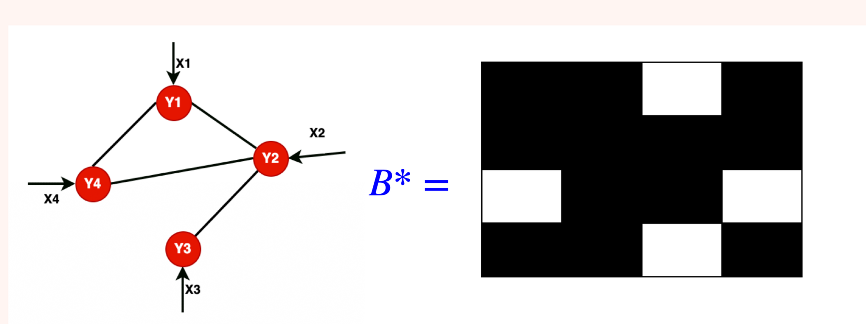


Figure 2. Sparsity structure determines edge connectivity ie. if $B^*_{ij} = 0$ then $(i,j) \notin E$

A Convex Estimator

$$\widehat{B} = \underset{B=B^\top, B\succ 0}{\arg\min} \left[\mathrm{Tr}(S\underbrace{B\Sigma_X^{-1}B}_{\Theta_Y}) - \log\det(B^2) + \lambda_n \|B\|_{1,\text{off}} \right] \tag{2}$$

where $\|B\|_{1,\text{off}} \triangleq \sum_{i\neq j}|B_{ij}|$, S is the sample covariance matrix from n i.i.d samples of observation Y and λ_n is the regularization parameter.

Lemma:

For any $\lambda_n > 0$ and $B \succ 0$, the ℓ_1-regularized MLE is strictly convex and has a unique minimizer.

Assumptions

1. **Mutual incoherence condition:**
   Let Γ^* be the Hessian of the log-determinant function,
   $$\Gamma^* \triangleq \nabla_B^2 \log\det(B)|_{B=B^*} = B^{*-1} \otimes B^{*-1}. \tag{3}$$
   The matrix B satisfies the mutual incoherence condition if there exists some $\alpha \in (0,1]$ such that, $|||\Gamma^*_{E^cE}(\Gamma^*_{EE})^{-1}|||_\infty \leq 1-\alpha$.
2. **Hessian regularity condition:**
   Let d be the maximum number of non zero entries among all the rows in B^* (i.e., the degree of the underlying graph), $\Theta^* = B^*\Sigma_X^{-1}B^*$. Then,
   $$|||\Gamma^{*-1}|||_\infty \leq \frac{1}{4d\|\Theta^{*-1}\|_\infty |||\Sigma_X^{-1}|||_\infty}. \tag{4}$$

Theoretical Results

Theorem 1:

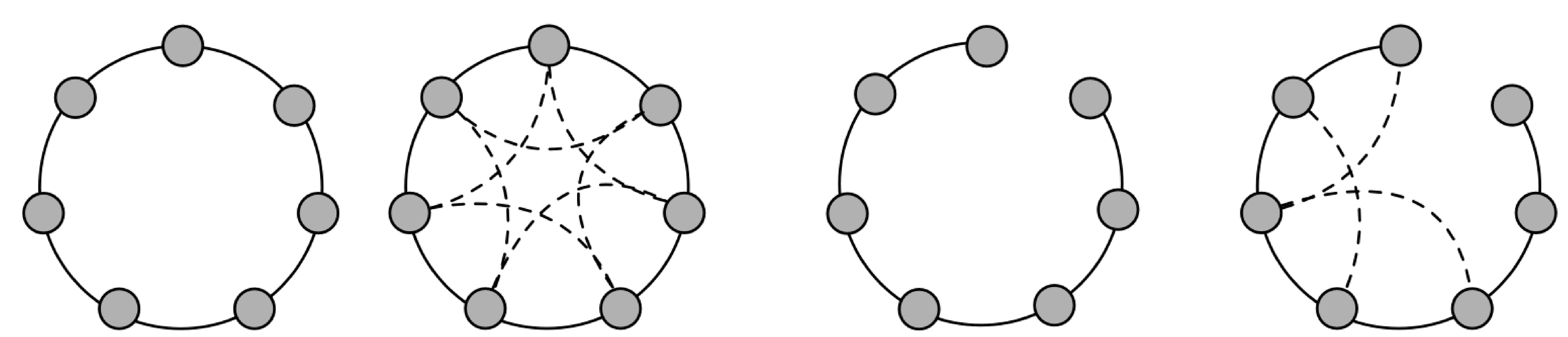
Let the vector of injected flows $X = (X_1, \ldots, X_p)$ be sub-Gaussian. Under some assumptions, if the number of samples $n = \Omega(d^2\log p)$ (high dim regime, $n \ll p$) then with high probability, the estimator $\widehat{B}$ has the following properties:

- Exact (and signed) support recovery: $\widehat{B}_{E^c} = 0$
- Element-wise $\ell_\infty-$norm consistency: $\|\widehat{B} - B^*\|_{\max} = \mathcal{O}(\sqrt{\log p/n})$

Theorem 2:

Consider the vector of injected flows whose $4k^{\text{th}}$ moments are bounded, if the number of samples $n = \Omega(d^2p^{1/k})$, then with high probability, the estimator satisfies the same properties as in Theorem 1.

Novelty



(a) true graph (b) graphical model $(c)\ \ell_1-$MLE (this paper) (d) GLASSO+2HR

Figure 3. Stylistic visualization of ℓ_1-MLE vs GLASSO+2HR. The GLASSO+2HR estimates Fig(b) and eliminates the 2-hop neighbours(spurious edges denoted by dashed lines) via thresholding, while the ℓ_1-regularized MLE estimates the structure directly.

GLASSO+SR (naive baseline)	GLASSO+2HR (Hop Refinement)	ℓ_1-regularized MLE
• If Σ_X is diagonal, $\widehat{B} = \sqrt{\widehat{\Theta}_{GL}}$ • Σ_X should be diagonal	Identify support of B^* when $\widehat{\Theta}_{GL} \leq -\tau$ Cycle length > 3	Estimates B^* directly No structural assumptions
• $n = \mathcal{O}(d^4\log p)$	$n = \mathcal{O}(d^4\log p)$	$n = \mathcal{O}(d^2\log p)$

Experiments

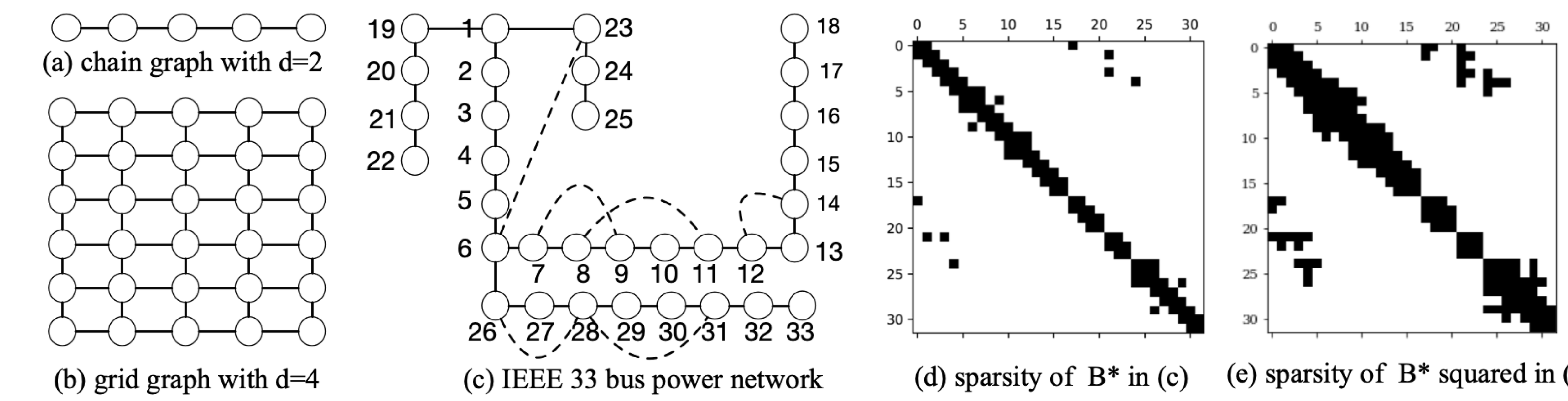


Figure 4. Graphs used in experiments. (a) Chain graph with maximum degree $d = 2$. (b) Grid graph $d = 4$. (c) IEEE 33 bus (node) distribution network with additional loops (shown in dashed lines). (d) Sparsity of B^* associated with the IEEE 33 bus network. (d) Sparsity of $(B^*)^2$. Notice that $(B^*)^2$ is denser relative to B^*. Consequently, GLASSO+2HR needs more samples than ℓ_1-regularized MLE to recover the support.

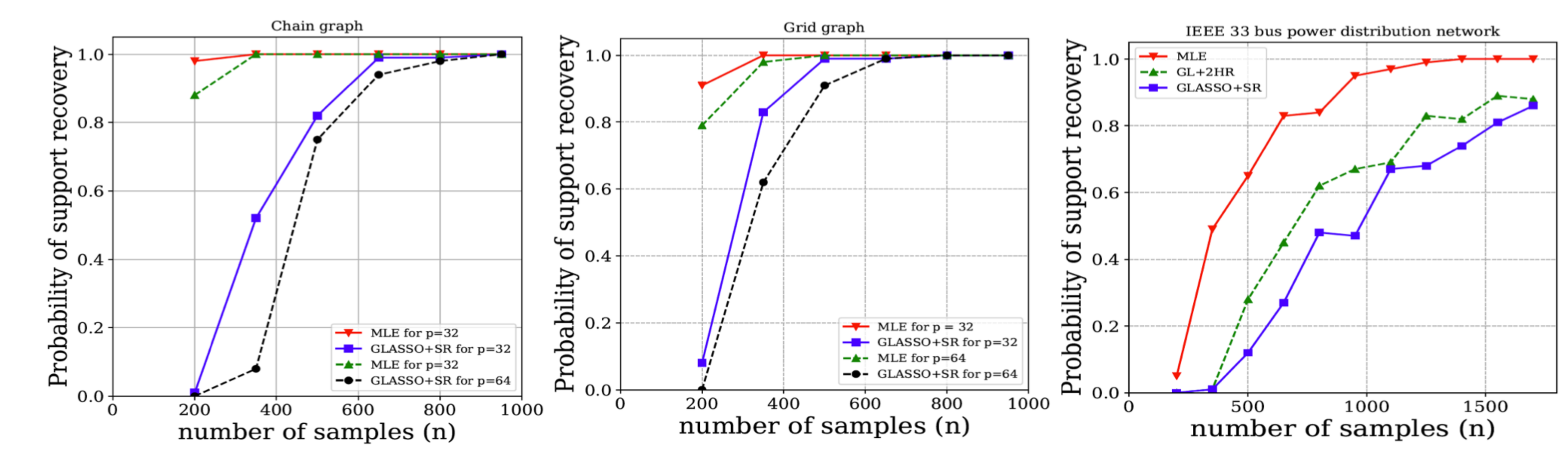


Figure 5. Empirical probability of success of various estimators versus the sample size n, for chain graph (left), grid graph (middle), and IEEE 33 bus network (right). For IEEE 33 bus network, we compare ℓ_1-regularized MLE with GLASSO+SR and GLASSO+2HR.

Conclusion

- We propose a novel ℓ_1-regularized MLE for B^* from samples of Y. Our first result shows that, the ℓ_1-regularized MLE is convex in B and it has a unique minimum even in the high-dimensional regime ($n \ll p$).
- Under a new mutual incoherence condition and a hessian regularity assumption, we provide sample complexity guarantees for exact support recovery and norm consistency in the high dimensional regime.
- We complement our theoretical results with experimental results both on the synthetic data sets and data from a benchmark power distribution system. Our experiments demonstrate the clear benefit of the proposed estimator over baseline and competing methods.

References

[1] Deepjyoti Deka, Saurav Talukdar, Michael Chertkov, and Murti V Salapaka. Graphical models in meshed distribution grids: Topology estimation, change detection & limitations. *IEEE Transactions on Smart Grid*, 11(5):4299–4310, 2020.

[2] Pradeep Ravikumar, Martin J Wainwright, Garvesh Raskutti, and Bin Yu. High-dimensional covariance estimation by minimizing ℓ_1-penalized log-determinant divergence. *Electronic Journal of Statistics*, 5:935–980, 2011.

This work was supported in part by the National Science Foundation (NSF) under the grants CCF-2048223 and OAC-1934766, and by the National Institutes of Health (NIH) under the grant 1R01GM140468-01.