# Identification of Edge Disconnections in Networks Based on Graph Filter Outputs

Shlomit Shaked and Tirza Routtenberg
Ben-Gurion University of the Negev, Israel
School of Electrical and Computer Engineering

# Overview

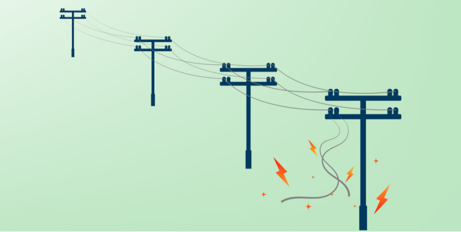
1. Motivation and Background
2. Measurements Model
3. Identification of edge disconnections
4. Greedy approaches for identifying edge disconnections
5. Simulations
6. Conclusions

# Identifying Edge Disconnections Using Graph Signal Processing

**Motivation:**

- GSP methods are based on known underlying topology
- Topology changes may degrade the performance of GSP tasks
- Edge disconnections are a common problem, especially in physical networks.

**Goal:** use graph signals to identify edge disconnections, where the original underlying network is known



**Example:** Identifying line outages in electrical networks, due to environmental factors, damages, aging, malicious attacks, etc.

Shaked, Shlomit, and Tirza Routtenberg. "Identification of Edge Disconnections in Networks Based on Graph Filter Outputs." IEEE Transactions on Signal and Information Processing over Networks 7 (2021): 578-594.

# Related works

- Numerous works in the literature have focused on (full) graph-topology learning.
  - Inefficient for identifying only a few specific disconnections
  - Suboptimal, since the nominal topology is known
- Other works detect topology changes based on graph data and not on "graph signals"
- Matched subspace detectors based on graph signals to decide which graph matches a given dataset, but does not use information on the nature of the change [1], [2]
- Edge exclusion tests for general graphical models [3]
- Laplacian learning in Gaussian Markov random field models with known connectivity [4]

---

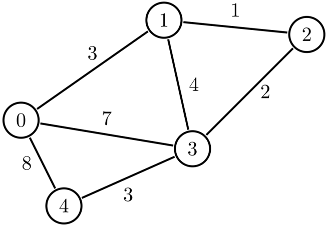
[1] C. Hu, J. Cheng, K. A. Sepulcre, G. E. Fakhri, Y. M. Lu, and K. Li, "Matched signal detection on graphs: Theory and application to brain imaging data classification", 2016.
[2] E. Isufi, A. S. U. Mahabir and G. Leus, "Blind Graph Topology Change Detection", 2018
[3] K. Tugnait, "Edge exclusion tests for graphical model selection: Complex Gaussian vectors and time series", 2019
[4] H. Egilmez, E. Pavez, and A. Ortega, "Graph Learning from Data under Structural and Laplacian Constraints", 2017

# Background: GSP Definitions

Given an undirected, connected, weighted graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}, \mathbf{W}\}$:

- $\mathcal{V}$ is set of vertices, where $N \triangleq |\mathcal{V}|$ and $\mathcal{E}$ is a set of edges.
- $\mathbf{W} \in \mathbb{R}^{N \times N}$ is the non-negative weighted *adjacency matrix* of the graph.
  If $(i,j) \in \mathcal{E}$, the entry $\mathbf{W}_{i,j} > 0$ represents the weight of the edge; otherwise, $\mathbf{W}_{i,j} = 0$.
- The *Laplacian matrix* is $\mathbf{L} \triangleq \mathrm{diag}(\mathbf{W1}) - \mathbf{W}$, where each entry satisfies

$$\mathbf{L}_{i,j} = \begin{cases} \sum_{k:\,(i,k)\in\mathcal{E}} \mathbf{W}_{i,k} \,, & i = j \,,\ i \in \mathcal{V} \\ -\mathbf{W}_{i,j} \,, & (i,j) \in \mathcal{E} \\ 0 \,, & \text{otherwise} \end{cases}$$




$$\underbrace{\begin{pmatrix} 18 & -3 & 0 & -7 & -8 \\ -3 & 8 & -1 & -4 & 0 \\ 0 & -1 & 3 & -2 & 0 \\ -7 & -4 & -2 & 16 & -3 \\ -8 & 0 & 0 & -3 & 11 \end{pmatrix}}_{\mathbf{L}}$$

- The singular value decomposition (SVD) is given by $\mathbf{L} = \mathbf{U}^{(\mathbf{L})} \Lambda^{(\mathbf{L})} \left(\mathbf{U}^{(\mathbf{L})}\right)^{\top}$.

# Background: GSP Definitions

- A *graph signal* is a vector measured over the vertices, $\mathbf{a} : \mathcal{V} \to \mathbb{R}^N$.
- The *graph fourier transform (GFT)* w.r.t $\mathbf{L}$ is $\tilde{\mathbf{a}}^{(\mathbf{L})} = \left(\mathbf{U}^{(\mathbf{L})}\right)^\top \mathbf{a}$.
- The *inverse GFT (IGFT)* w.r.t $\mathbf{L}$ is $\mathbf{a} = \mathbf{U}^{(\mathbf{L})}\tilde{\mathbf{a}}^{(\mathbf{L})}$.
- The *smoothness* or *Dirichlet energy* is measured by

$$Q_{\mathbf{L}}(\mathbf{a}) \triangleq \frac{1}{2} \sum_{(i,j)\in\mathcal{E}} \mathbf{W}_{i,j}[\mathbf{a}_i - \mathbf{a}_j]^2 = \mathbf{a}^\top \mathbf{L}\mathbf{a}.$$

> *Smooth graph signals* are signals with "small" Dirichlet energy.
> Intuitively, a smooth graph signal is considered to be a "good match" with the graph if the signal values are close to their neighbors' values.

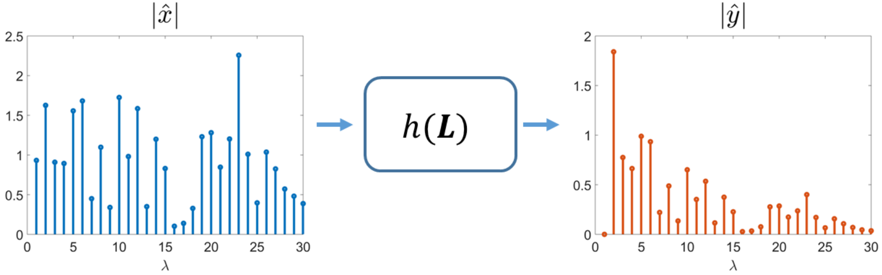
# Background: Graph Filter

Filtering in graph Fourier space can be represented by

$$\tilde{\mathbf{a}}_{\text{out}}^{(\mathbf{L})} = \text{diag}\big([h(\lambda_1^{(\mathbf{L})}), \ldots, h(\lambda_N^{(\mathbf{L})})]^T\big)\tilde{\mathbf{a}}_{\text{in}}^{(\mathbf{L})}.$$

Then,the *graph filter* is a linear operator relates to input-output $\mathbf{a}_{\text{out}} = h(\mathbf{L})\mathbf{a}_{\text{in}}$, where

$$h(\mathbf{L}) \triangleq \mathbf{U}^{(\mathbf{L})}\text{diag}\big([h(\lambda_1^{(\mathbf{L})}), \ldots, h(\lambda_N^{(\mathbf{L})})]^T\big)\big(\mathbf{U}^{(\mathbf{L})}\big)^T.$$



# Smooth Graph Filter: $\mathcal{E}[\mathrm{Q}_{\mathbf{L}}(\mathbf{a}_{\text{out}})] < \mathrm{E}[Q_{\mathbf{L}}(\mathbf{a}_{\text{in}})]$

| Graph Filter | $h(\lambda)$ | For $\mathbf{a}_{in} \sim \mathcal{N}(0, \mathbf{I})$ |
|---|---|---|
| Gaussian Markov random field (GMRF) with a Laplacian precision matrix | $\begin{cases} \frac{1}{\sqrt{\lambda}} & \lambda \neq 0 \\ 0 & \lambda = 0 \end{cases}$ | $\mathbf{a}_{out} \sim \mathcal{N}(0, \mathbf{L}^{\dagger})$ |
| Regularized Laplacian by Tikhonov Regularization | $\frac{1}{1+\alpha\lambda}, \ \alpha > 0$ | $\mathbf{a}_{out} \sim \mathcal{N}(0, (\mathbf{I} + \alpha\mathbf{L})^{-2})$ |
| Heat Diffusion Kernel | $\exp(-\tau\lambda), \ \tau > 0$ | $\mathbf{a}_{out} \sim \mathcal{N}(0, \exp(-2\tau\mathbf{L}))$ |

# Outline

# Measurement Model

- We consider the measurement model as an output of a smooth graph filter:

$$\mathbf{y}[m] = h(\mathbf{L})\mathbf{x}[m] + \mathbf{w}[m], \; m = 1 \dots M.$$

- The log-likelihood of the augmented output vector of $M$ time samples, $\mathbf{y} \triangleq [\mathbf{y}^T[1], \dots, \mathbf{y}^T[M]]^T$, is

$$\log f(\mathbf{y}; \mathbf{L}) = -\frac{M}{2} \log \left((2\pi)^N |\sigma_{\mathbf{x}}^2 h^2(\mathbf{L}) + \sigma_{\mathbf{w}}^2 \mathbf{I}|_+\right) - \frac{M}{2} \mathrm{Tr}\left(\left(\sigma_{\mathbf{x}}^2 h^2(\mathbf{L}) + \sigma_{\mathbf{w}}^2 \mathbf{I}\right)^\dagger \mathbf{S}_{\mathbf{y}}\right),$$

where the sample covariance matrix

$$\mathbf{S}_{\mathbf{y}} \triangleq \frac{1}{M} \sum_{m=1}^{M} \mathbf{y}[m]\mathbf{y}^T[m].$$

$\Rightarrow$ the graph filter, $h(\mathbf{L})$, "colors" the input graph signal using the network connectivity.

# Problem formulation

Identification of edge disconnections:

$$\mathcal{H}_k : \quad \mathbf{L} = \mathbf{L}^{(k)}, \quad k = 0, 1, \ldots, K$$

based on the graph signals $\mathbf{y}$, where

- $\mathbf{L}^{(0)}$ is the Laplacian of the original, known topology (set of edges: $\mathcal{E}$),
- $\mathbf{L}^{(k)}$ is the Laplacian matrix after disconnections at the edges in $\mathcal{C}^{(k)} \subset \mathcal{E}$.

$$\mathbf{L}^{(k)} \triangleq \mathbf{L}^{(0)} - \underbrace{\sum_{(i,j)\in\mathcal{C}^{(k)}} \mathbf{E}^{(i,j)}}_{\mathbf{E}^{(k)}}$$

$$\mathbf{E}^{(i,j)} \triangleq \left[\mathbf{L}^{(0)}\right]_{i,j} \left(\mathbf{e}_i\mathbf{e}_j^T + \mathbf{e}_j\mathbf{e}_i^T - \mathbf{e}_j\mathbf{e}_j^T - \mathbf{e}_i\mathbf{e}_i^T\right).$$

corresponds to a single-edge disconnection at $(i,j) \in \mathcal{E}$.

$$\underbrace{\begin{pmatrix} 11 & -3 & 0 & 0 & -8 \\ -3 & 8 & -1 & -4 & 0 \\ 0 & -1 & 3 & -2 & 0 \\ 0 & -4 & -2 & 9 & -3 \\ -8 & 0 & 0 & -3 & 11 \end{pmatrix}}_{\mathbf{L}^{(k)}} = \underbrace{\begin{pmatrix} 18 & -3 & 0 & -7 & -8 \\ -3 & 8 & -1 & -4 & 0 \\ 0 & -1 & 3 & -2 & 0 \\ -7 & -4 & -2 & 16 & -3 \\ -8 & 0 & 0 & -3 & 11 \end{pmatrix}}_{\mathbf{L}^{(0)}} - \underbrace{\begin{pmatrix} 7 & 0 & 0 & -7 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ -7 & 0 & 0 & 7 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}}_{\mathbf{E}^{(k)}}$$

# Outline

# Maximum likelihood decision rule

The maximum likelihood decision rule for this problem is given by

$$T(\mathbf{y}) = \underset{0 \le k \le K}{\operatorname{argmax}} \frac{\log f(\mathbf{y}; \mathbf{L}^{(k)})}{\log f(\mathbf{y}; \mathbf{L}^{(0)})} = \underset{0 \le k \le K}{\operatorname{argmax}} \, l(\mathbf{y}|\mathbf{L}^{(k)}) - \rho(\mathbf{L}^{(k)}).$$

where

- $l(\mathbf{y}|\mathbf{L}^{(k)}) \triangleq \mathrm{Tr}\left(\left(\left(\sigma_{\mathbf{x}}^2 h^2(\mathbf{L}^{(0)}) + \sigma_{\mathbf{w}}^2 \mathbf{I}\right)^{\dagger} - \left(\sigma_{\mathbf{x}}^2 h^2(\mathbf{L}^{(k)}) + \sigma_{\mathbf{w}}^2 \mathbf{I}\right)^{\dagger}\right)\mathbf{S}_{\mathbf{y}}\right)$ - data term
- $\rho(\mathbf{L}^{(k)}) \triangleq \log\left(\frac{|\sigma_{\mathbf{x}}^2 h^2(\mathbf{L}^{(k)}) + \sigma_{\mathbf{w}}^2 \mathbf{I}|_+}{|\sigma_{\mathbf{x}}^2 h^2(\mathbf{L}^{(0)}) + \sigma_{\mathbf{w}}^2 \mathbf{I}|_+}\right)$ - penalty term
- $\mathbf{S}_{\mathbf{y}} \triangleq \frac{1}{M} \sum_{m=1}^{M} \mathbf{y}[m]\mathbf{y}^T[m]$ - sample covariance matrix

$\Rightarrow$The problem of testing *structured* covariance matrix of random Gaussian vectors

# Remark 1: Penalty function interpretation

**proposition:** Consider two connected graphs, $\mathcal{G}^{(k_1)}$ and $\mathcal{G}^{(k_2)}$, with the Laplacian matrices, $\mathbf{L}^{(k_1)}$ and $\mathbf{L}^{(k_2)}$, respectively, and assume:

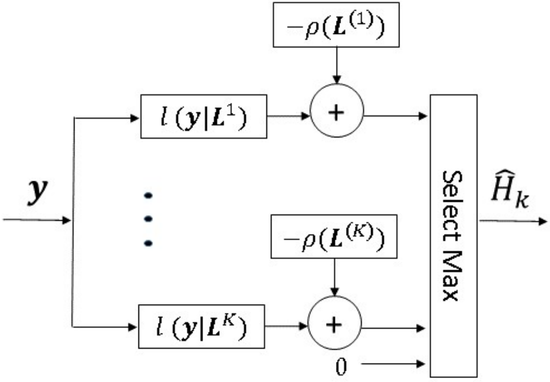
C.1) $\mathcal{C}^{(k_2)}$ is a proper subset of $\mathcal{C}^{(k_1)}$, i.e. $\mathcal{C}^{(k_2)} \subset \mathcal{C}^{(k_1)}$.

C.2) The graph filter, $h(\lambda)$, is a monotonic decreasing function of $\lambda$ for any $\lambda > 0$.

C.3) The covariance matrices are non-singular matrices.

Then,

$$\rho(\mathbf{L}^{(k_2)}) \leq \rho(\mathbf{L}^{(k_1)}).$$




⇒A larger penalty for more edge disconnections

# Remark 2: The sufficient statistics

The $k$th likelihood can be written in the graph spectral domain as

$$l(\mathbf{y}|\mathbf{L}^{(k)}) = \frac{\sigma_{\mathbf{x}}^2}{\sigma_{\mathbf{w}}^2}\left(\sum_{n=1}^{N} \frac{h^2(\lambda_n^{(\mathbf{L}^{(k)})})}{\sigma_{\mathbf{w}}^2 + \sigma_{\mathbf{x}}^2 h^2(\lambda_n^{(\mathbf{L}^{(k)})})}\psi_n^{(\mathbf{L}^{(k)})} - \frac{h^2(\lambda_n^{(\mathbf{L}^{(0)})})}{\sigma_{\mathbf{w}}^2 + \sigma_{\mathbf{x}}^2 h^2(\lambda_n^{(\mathbf{L}^{(0)})})}\psi_n^{(\mathbf{L}^{(0)})}\right),$$

where the sufficient statistics for the identification are the graph-frequency energy levels, i.e.

$$\psi_n^{(\mathbf{L}^{(l)})} \triangleq \frac{1}{M}\sum_{m=1}^{M}\left(\left[\tilde{\mathbf{y}}^{(\mathbf{L}^{(l)})}[m]\right]_n\right)^2 \quad n = 1, \ldots, N,\ l = 0, k.$$

⇒the maximum likelihood decision rule only requires the evaluation of the $NK$ scalars

⇒ it can be shown that is governed by the low-graph frequencies

# Remark 3: GMRF model with a Laplacian precision

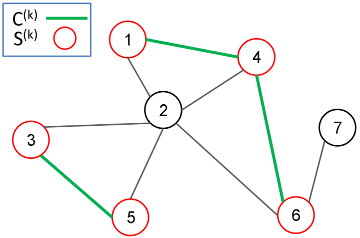
In this case, the $k$th likelihood term satisfies

$$l(\mathbf{y}|\mathbf{L}^{(k)}) = -\frac{1}{\sigma_{\mathbf{x}}^2 M} \sum_{\boxed{(i,j)\in \mathcal{C}^{(k)}}} L_{i,j}^{(0)} \sum_{m=1}^{M} (y_i[m] - y_j[m])^2$$

**proposition:** Consider a connected graph, $\mathcal{G}^{(k)}$, with a Laplacian matrix, $\mathbf{L}^{(k)}$. Then, for the noiseless GMRF model with a Laplacian precision matrix, the $k$th term

$$\underbrace{l(\mathbf{y}|\mathbf{L}^{(k)})}_{\text{measurements + 2nd-order statistics}} - \underbrace{\rho(\mathbf{L}^{(k)})}_{\text{2nd-order statistics}}$$

is only a function of the vertices in $\mathcal{S}^{(k)}$.




$\mathcal{G}^{(k)}$ is obtained by removing the edges in $\mathcal{C}^{(k)} = \{(1,4),(4,6),(3,5)\}$, associated with the vertices in $\mathcal{S}^{(k)} = \{1,3,4,5,6\}$.

# Computational complexity

- Number of hypotheses in the general case:

$$K = \sum_{r=1}^{r_{\max}} \binom{|\mathcal{E}|}{r},$$

where $r_{\max}$ is the maximum number of possible edge disconnections

- The calculation of $l(\mathbf{y}|\mathbf{L}^{(k)})$ and $\rho(\mathbf{L}^{(k)})$ requires an inversion of $N \times N$ matrix
- The computational complexity of the maximum likelihood decision rule grows exponentially with the graph size and is impractical for large networks
- We develop efficient low-complexity methods based on:
  - Greedy approach
  - Local properties of smooth graph filters

# Outline

1. Motivation and Background
2. Measurements Model
3. Identification of edge disconnections
4. Greedy approaches for identifying edge disconnections
5. Simulations
6. Conclusions

# Greedy approach

**Algorithm 1:** Greedy identification

- Input: $\mathbf{S}_{\mathbf{y}}$, $\mathbf{L}^{(0)}$, $\mathcal{E}$, $\sigma^2_{\mathbf{x}}$, $\sigma^2_{\mathbf{w}}$, $h(\cdot)$, Optional: $r_{\max}$.
- Output: Estimated edge disconnections set, $\hat{\mathcal{C}}$.

Initialize $\hat{\mathcal{C}}^0 = \emptyset$, $\hat{\mathcal{E}}^0 = \mathcal{E}$, $\hat{\mathbf{L}}^0 = \mathbf{L}^{(0)}$, and $l = 0$.
Find the maximal edge, $\hat{k} \in \hat{\mathcal{E}}^l$, by

$$\hat{k} = \underset{k=(i,j)\in\hat{\mathcal{E}}^l}{\operatorname{argmax}} \; l(\mathbf{y}|\hat{\mathbf{L}}^l - \mathbf{E}^{(k)}) - \rho(\hat{\mathbf{L}}^l - \mathbf{E}^{(k)}),$$

**if** $l(\mathbf{y}|\hat{\mathbf{L}}^l - \mathbf{E}^{(\hat{k})}) - \rho(\hat{\mathbf{L}}^l - \mathbf{E}^{(\hat{k})}) > 0$ **then**

  Update $\hat{\mathcal{C}}^{l+1} = \hat{\mathcal{C}}^l \cup \{\hat{k}\}$, $\hat{\mathcal{E}}^{l+1} = \hat{\mathcal{E}}^l \setminus \{\hat{k}\}$, $\hat{\mathbf{L}}^{l+1} = \hat{\mathbf{L}}^l - \mathbf{E}^{(\hat{k})}$, and $l \leftarrow l + 1$

  **if** $|\hat{\mathcal{C}}^l| = r_{\max}$ **then**

    **Return:** $\hat{\mathcal{C}}^l$.

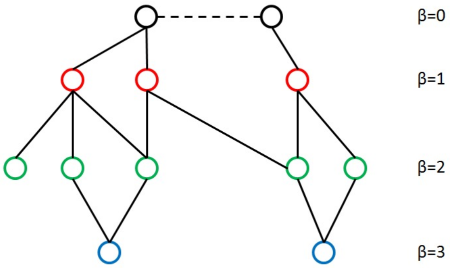
  **Repeat** to step 1.

**Return:** $\hat{\mathcal{C}}^l$.

# Neighboring strategy

**$\beta$-local maximum likelihood decision rule**

- For a given candidate edge, $(i, j)$, we calculate the likelihood ratio of the measurements in the $\beta$-neighborhood of $i$ and $j$, $\mathcal{N}(i,\beta)\bigcup\mathcal{N}(j,\beta)$, where $\mathcal{N}(i,\beta)$ is the set of vertices connected to vertex $i$ by a path of at most $\beta$ edges.
- For each iteration, building a new set of the suspicious edges for the next iteration.




$\Rightarrow$
The tunable parameter $\beta$ provides a trade-off between the identification accuracy and the computation cost.

# Computational complexity

| | ML rule | Greedy | Greedy + neighboring strategy |
|---|---|---|---|
| # Likelihood ratio calculations | $\sum_{r=1}^{r_{\max}} \binom{\lvert\mathcal{E}\rvert}{r}$ | $r_{\max} \times \lvert\mathcal{E}\rvert$ | $r_{\max} \times \lvert\mathcal{E}\rvert$ |
| Matrix inversion* | $\mathcal{O}(N^3)$ | $\mathcal{O}(N^3)$ | $\mathcal{O}(\lvert\mathcal{N}((i,j),\beta)\rvert^3)$ |
| Searching edge set size* | $\sum_{r=1}^{r_{\max}} \binom{\lvert\mathcal{E}\rvert}{r}$ | $\lvert\mathcal{E}\rvert$ | $\lvert\hat{\mathcal{E}}^{(l)}\rvert \ll \lvert\mathcal{E}\rvert$ |

*For sparse graphs, where $|\mathcal{E}| \ll \frac{N(N-1)}{2}$.

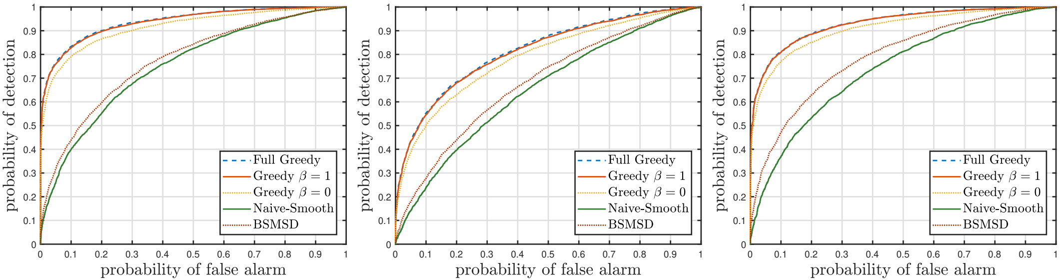
# Outline

# Simulations #1: synthetic data

- Smooth graph filters
- The initial graph was generated by using the Watts-Strogatz small-world graph model, with $N = 50$ vertices, mean degree of $d = 2$, and $|\mathcal{E}| = 100$
- The elements of $\mathbf{W}^{(0)}$ are independent, uniformly distributed weights in $[0.1, 5]$
- Topology change is obtained by removing an arbitrary set of $r$ edges from $\mathcal{E}$
- Comparison with
  - Blind simple-MSD (BSMD) detector [1, 2]
  - Smoothness detector
  - GGM-GLRT: uses the sample covariance matrix of the 1st-order neighbors of the edges [3]
  - Combinatorial graph Laplacian (CGL) method [4]
  - Constrained CGL (CCGL) method: CGL + information on the initial Laplacian matrix, $\mathbf{L}^{(0)}$

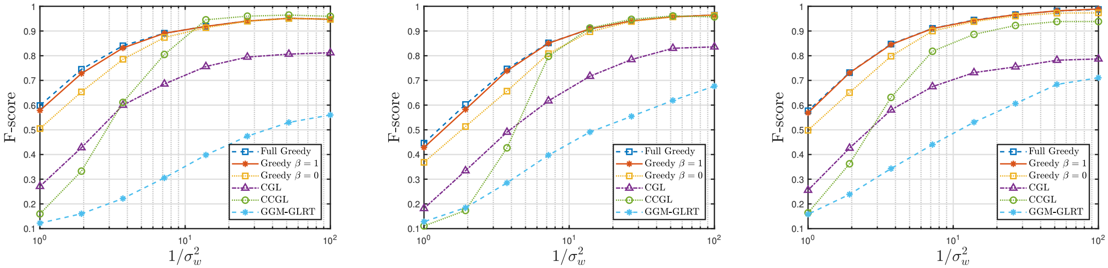
[1] C. Hu, J. Cheng, K. A. Sepulcre, G. E. Fakhri, Y. M. Lu, and K. Li, "Matched signal detection on graphs: Theory and application to brain imaging data classification", 2016.
[2] E. Isufi, A. S. U. Mahabir and G. Leus, "Blind Graph Topology Change Detection", 2018
[3] K. Tugnait, "Edge exclusion tests for graphical model selection: Complex Gaussian vectors and time series", 2019
[4] H. Egilmez, E. Pavez, and A. Ortega, "Graph Learning from Data under Structural and Laplacian Constraints", 2017

# Detection performance




Receiver operating characteristic (ROC) curves of edge disconnection detection by the greedy algorithm, $\beta = 0, 1$-neighbors greedy algorithm, smoothness detector, and BMSD for: the GMRF (left), the regularized Laplacian (middle), and the heat diffusion (right) filters, with noise variance $\sigma_{\mathbf{w}}^2 = 0.5$, $M = 100$ edges, and $r = 5$ potential disconnections.

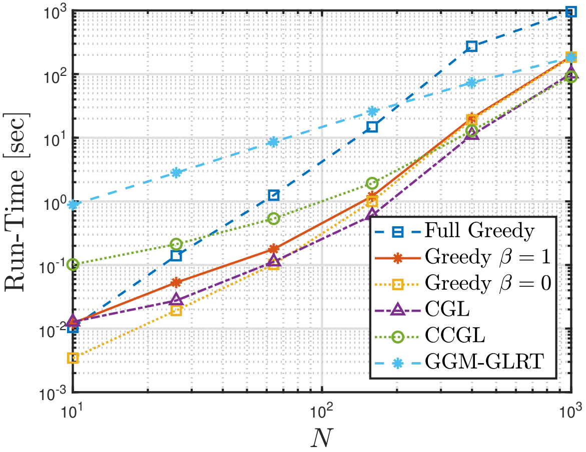
# Identification performance




The F-score measure for the GMRF (left), the regularized Laplacian (middle), and the heat diffusion (right) filters versus SNR for the greedy algorithm, $\beta = 0, 1$-neighbors greedy algorithm, CGL method, CCGL method, and the GGM-GLRT method with $M = 1,000$ and $r = 5$.

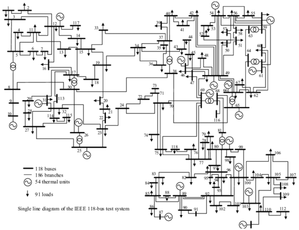
The F-score metric takes values between $0$ and $1$, where $1$ means perfect identification.

# Run-time



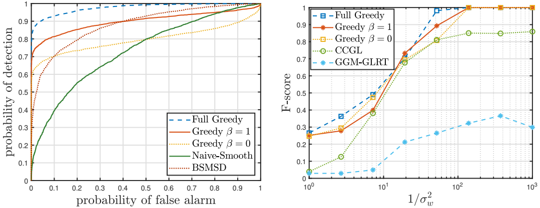

Run-time of the different methods for $\sigma_{\mathbf{w}}^2 = 0.1$, $M = 100,000$, $r = 2$, and $L = 100$

# Simulations #2: Identifying outages in power system dataset




- The vertices and the edges denote the buses (generators or loads), and the transmission lines between these buses, respectively. The branch susceptances determine the weights of the graph edges.
- We assume Phasor measurement units PMUs in the considered system that acquire noisy measurements of the voltage phases at all buses, and $|v_n| = 1$
- We tested random combinations of outages at the transmission lines,

Identifying outages in power system: ROC curves (left) of edge disconnections detector and the F-score measure (right) versus SNR by assuming the GMRF filter.



# Outline

1. Motivation and Background
2. Measurements Model
3. Identification of edge disconnections
4. Greedy approaches for identifying edge disconnections
5. Simulations
6. Conclusions

# Conclusions

- We propose identifying edge disconnections in networks based on a graph filter model.
- Interpretations of the developed maximum likelihood decision rule:
  - Based on graph energy levels in the graph spectral domain
  - Has a penalty on models with a larger number of disconnections
  - A local smoothness detector for the noiseless GMRF filter with a Laplacian precision
- We propose two greedy algorithms that
  - converge to maximum likelihood decision rule for the noiseless GMRF filter
  - outperform state-of-the-art methods on the tested scenarios in terms of detection and identification performance, and computational complexity
  - The neighboring strategy is based on localization and smoothness properties
  - provide a good trade-off between performance and complexity
- Future research directions include
  1. Extension for “blind” scenarios with unknown graph filters
  2. Dynamically change detection
  3. Other typical topology changes