Detecting Low Pass Graph Signals via Spectral Pattern: Sampling Complexity and Applications

Chenyue Zhang, Yiran He, Hoi-To Wai

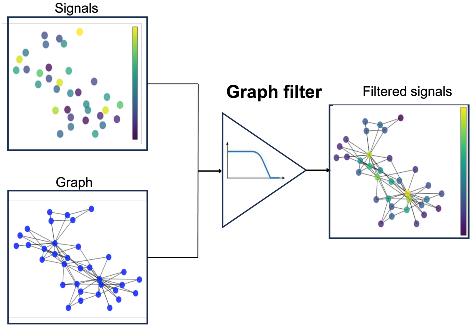
Department of Systems Engineering & Engineering Management, CUHK

June 13, 2023

GSP Workshop 2023, Oxford

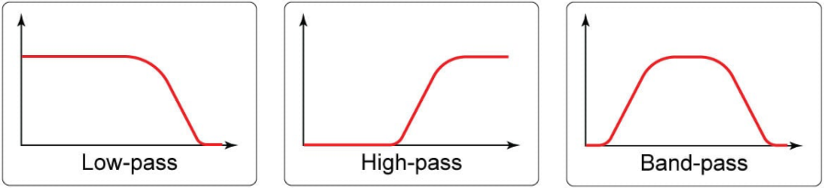
Graph Signal Processing

[Source: NETWORK MANAGEMENT]



- **Graphs:** Effectively depict data's spatial layout in diverse fields like **social**, **biology**, **transportation**, and **power networks**.
- **GSP:** A **flexible tool** that extends the concepts from **classical signal processing** to graphs and makes inference from data.
- Network Data $\rightarrow$ Filtered Graph Signals

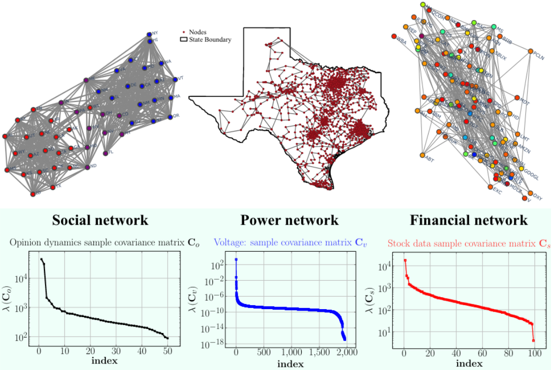
Motivation



- Like LTI filters, graph filter can be classified as **low pass**, **band pass**, or **high pass** through its graph frequency response.
- Low-pass graph signals capture a smoothing operation of input graph signals, are **prevalent** in network data, e.g., social network, financial network, power network, etc.[1]

[1] [Ramakrishna et al., 2020] R. Ramakrishna, H. -T. Wai, A. Scalgione. A user guide to low-pass graph signal processing and its applications. SPM, 2020.

Motivation



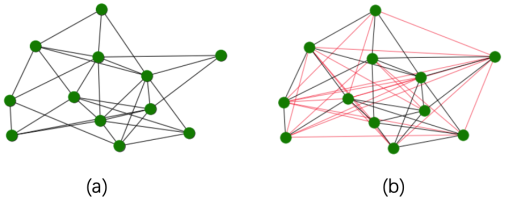
Low pass graph signals [Source: [Ramakrishna et al., 2020]]

- Many GSP algorithms rely on the **low pass assumption**. E.g. sampling [Anis et al., 2016], graph topology learning [Dong et al., 2016a], GNN [Wu et al., 2019], community detection [Schaub et al., 2020], centrality estimation [He and Wai, 2022].

Motivation

- Low pass assumption can be **dangerous** if graph signal generating model is **unknown** or data is **corrupted**.

E.g. Topology inferred from **non low pass** signals can be deceptive



(a) Ground truth. (b) Topology learnt by GL-SigRep[Dong et al., 2016b] on **non-low-pass** signals.

Can we detect if the signals are low pass before using GSP tools?

Related Works

- Blind identification of graph filters.
  [Zhu et al., 2020, Segarra et al., 2016, Ye et al., 2018]

  - Learn graph filter's coefficients $\rightarrow$ validate low pass
  $\rightarrow$ **Require Graph topology.**

- Network inference from spectra template. [Segarra et al., 2017]

  - Learn GSO without low pass assumption $\rightarrow$ validate low pass
  $\rightarrow$ **High computational cost, learn possibly wrong GSO.**

Low Pass Graph Filter

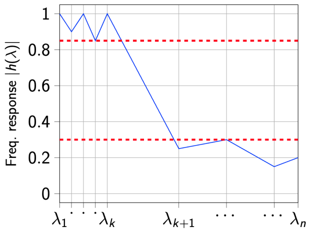
$$\text{GSO} : \boldsymbol{S}, \;\; \text{Graph Filter} : \mathcal{H}(\boldsymbol{S}) = \sum_{\ell=0}^{\infty} h_\ell \boldsymbol{S}^\ell, \;\; \text{Freq. response} : h(\lambda) = \sum_{\ell=0}^{\infty} h_\ell \lambda^\ell,$$

where GSO $=$ normalized Laplacian, with eigenvalues $0 = \lambda_1 \leq \cdots \leq \lambda_n$.

Def. For $1 \leq k \leq n-1$, define

$$\eta_k := \frac{\max\{|h(\lambda_{k+1})|, \ldots, |h(\lambda_n)|\}}{\min\{|h(\lambda_1)|, \ldots, |h(\lambda_k)|\}}.$$

If the low-pass ratio satisfies $\eta_k \in [0,1)$, then $\mathcal{H}(\boldsymbol{S})$ is k-low-pass.

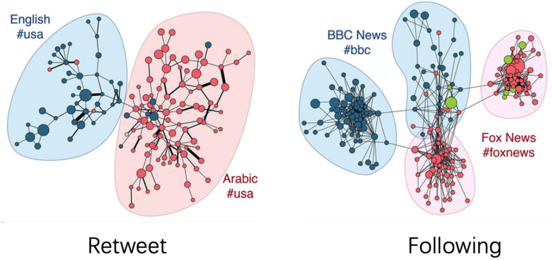


- Integer k characterizes the *bandwidth*, or the cut-off frequency.
- Let $\boldsymbol{x}$ be a white noise excitation, then k **low pass** graph signals

$$\boldsymbol{y} = \mathcal{H}(\boldsymbol{S})\boldsymbol{x}, \quad \text{where } \mathcal{H}(\boldsymbol{S}) \text{ is } k\text{-low pass}.$$

Detecting Low-pass Signals

- *Challenges*: graph topology $\boldsymbol{S}$ and filter $\mathcal{H}(\boldsymbol{S})$ are **unknown**.
- Warning: an **ill posed** problem – graph signals are *low pass* on one graph, but *non low pass* on another.
- Many real networks tend to be **modular**.

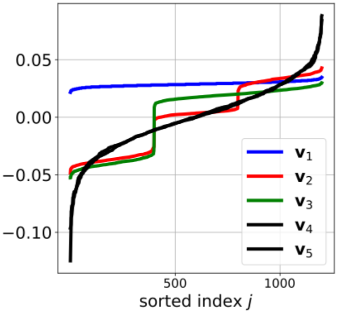


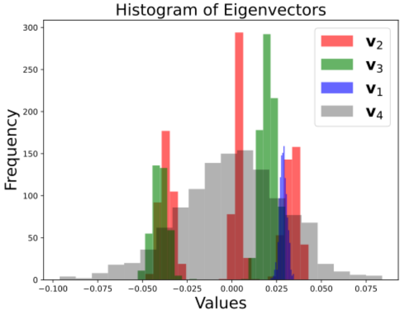
Real Social Network [Source: [Weng et al., 2013]]

- **Assume**: no. of dense clusters, K, in the graph is known a-priori. $\Longrightarrow \lambda_1, \dots, \lambda_K \approx 0 \Longrightarrow$ if filter is low pass, it will be K low pass.

Detecting Low-pass Signals

- **Observation**: graph signals from K low pass filter exhibit particular *spectral signature*. E.g., SBM graph with $K = 3$ clusters,

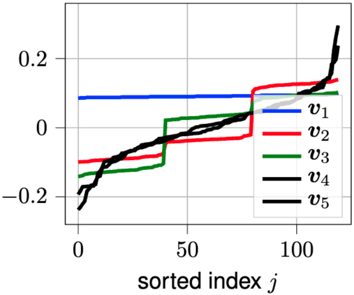


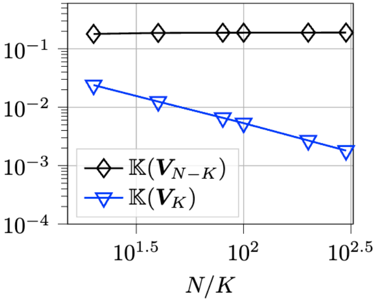


$\boldsymbol{v}_i, i \leq K$ — piecewise constant-valued [Deng et al., 2021]

$\boldsymbol{v}_i, i > K$ — ‘Gaussian’-valued (an open problem [Kadavankandy et al., 2015])

Detecting Low-pass Signals





Our idea:

- Define $\widehat{\mathbf{C}}_y^m := (1/m)\sum_{\ell=1}^m \mathbf{y}^\ell \left(\mathbf{y}^\ell\right)^\top, \widehat{\mathbf{v}}_i = i$ th-EV
- Detect if $\widehat{\boldsymbol{V}}_K$ has piecewise constant columns $\Longrightarrow$ K-means:

$$\mathbb{K}(\widehat{\boldsymbol{V}}_K) = \min_{S_i \cap S_j = \emptyset, i \neq j} \sum_{i=1}^K \sum_{\ell \in S_i} \left\| \widehat{\mathbf{v}}_\ell - \frac{1}{|S_i|} \sum_{j \in \mathcal{S}_i} \widehat{\mathbf{v}}_j \right\|^2 \lessgtr \ \delta.$$

Sampling Complexity Analysis

With high probability and assumption $\mathbb{K}(\boldsymbol{v}_\ell) \geq \mathrm{c}_{\mathsf{SBM}} > 0$, $\ell = K+1, \ldots, N$, if graph signals are K low pass,

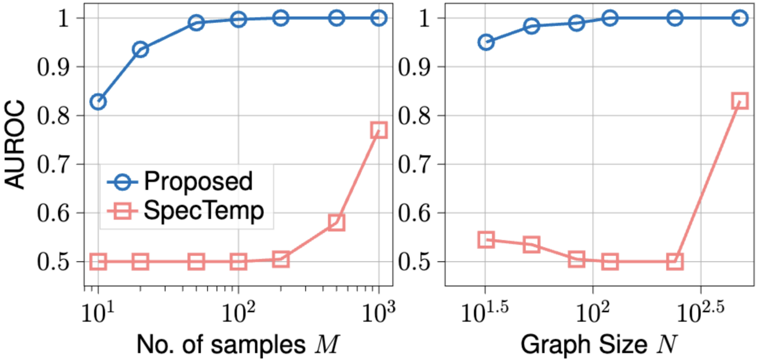
$$\mathbb{K}(\widehat{\boldsymbol{V}}_K) \leq \frac{16K}{\Delta^2}\left(2c_1\sqrt{\frac{\log M}{M}}\operatorname{tr}(\boldsymbol{C}_y) + \sigma^2\right)^2 + \frac{2450K^3 \log N}{p(N-K)}$$

if graph signals are not K low pass,

$$\mathbb{K}(\widehat{\boldsymbol{V}}_K) \geq \left(\sqrt{\mathrm{c}_{\mathsf{SBM}}} - \frac{2^{3/2}\sqrt{K}}{\Delta}2c_1\sqrt{\frac{\log M}{M}}\operatorname{tr}(\boldsymbol{C}_y) + \sigma^2\right)^2.$$

With large sample size M and graph size N, our method can provide correct detection.

Experiment



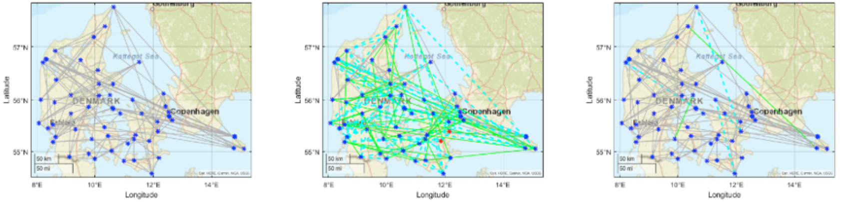
- Our blind detection outperforms SpecTemp[Segarra et al., 2017] with less computational complexity and sampling complexity.
- Our performance improves (AUROC $\to$ 1) as M, N increase.

Application: Robustifying Graph Learning.

Problem: Learning graph topology with **corrupted/non low pass data** is challenging[2].

Approach: Pre-screened graph learning procedure

1. Apply low pass detection algorithm on signal batches.
2. **Remove** batches with non low pass/**corrupted** signals.
3. Apply graph learning method on the **remaining** signals.

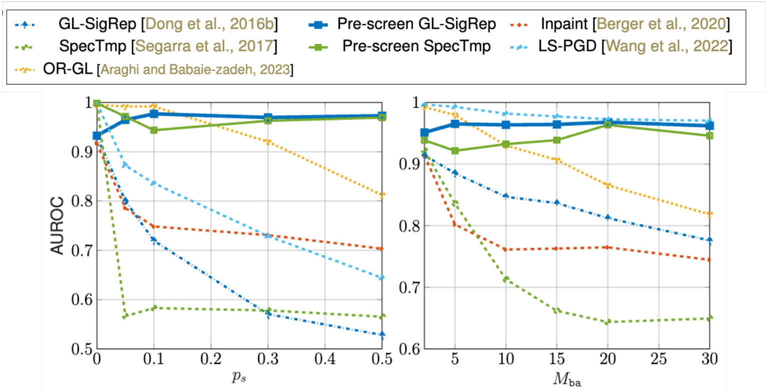


(a) (b) (c)

(a) original dataset (b) corrupted dataset (c) dataset after pre-screening.

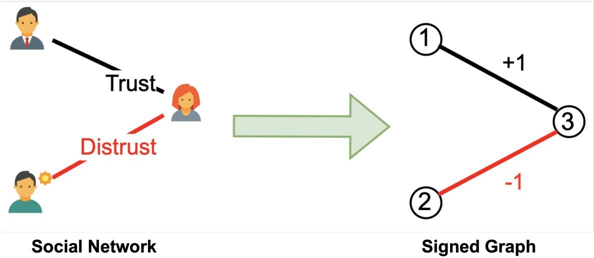
[2]Corruption data types include missing data, outliers, uncertainty, etc.

Application: Robustifying Graph Learning.



Pre-screening procedure can robustify graph learning against outliers, missing data, uncertainty corruption.

Application: Detecting Antagonistic Opinion Dynamics

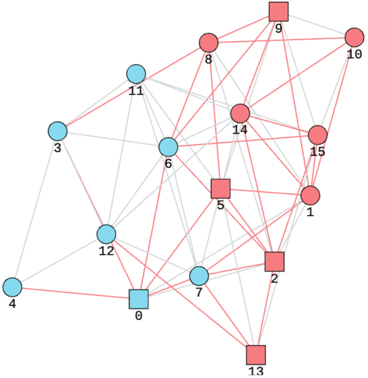


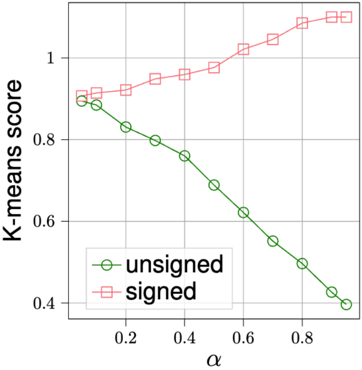
V = individuals, E^+ = friends/trust, E^- = unfriend/distrust

- Observed steady state :

$$\boldsymbol{y}_m := \lim_{\tau \to \infty} \boldsymbol{y}_m(\tau) = (1-\alpha)(\boldsymbol{I} - \alpha \boldsymbol{A}_{E^+} - \alpha \boldsymbol{A}_{E^-})^{-1} \boldsymbol{B} \boldsymbol{z}_m$$

Application: Detecting Antagonistic Opinion Dynamics





- 'signed' are negative edges in the real graph;'unsigned' means all edges being positive.
- Antagonistic/consensus behaviour are detected as distrust/trust α increase.

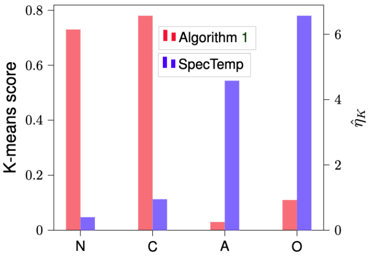
Application: Detecting Antagonistic Opinion Dynamics

On the Nomination (N): *e.g.,* "Thomas J. Vilsack, of Iowa, to be Secretary of Agriculture", "Rahm Emanuel, of Illinois, to be Ambassador to Japan", ...

On the Cloture Motion (C): *e.g.,* "Beth Robinson, of Vermont, to be United States Circuit Judge for the Second Circuit", "Douglas R. Bush, of Virginia, to be an Assistant Secretary of the Army", ...

On the Amendment (A): *e.g.,* "To establish a deficit-neutral reserve fund relating to COVID-19 vaccine administration and a public awareness campaign", "In the nature of a substitute", "To improve the bill", ...

Others (O): *e.g.,* "A bill to provide for reconciliation pursuant to title II of S. Con. Res. 5", "A resolution impeaching Donald John Trump, President of the United States, for high crimes and misdemeanors", ...



- Senate dataset $M = 949$ votes by $N = 97$ members $K = 2$.
- Clustering into 4 groups based on vote questions.
- Antagonistic are more obvious in "Nomination" votes; while consensus are observed in "Amendment" votes.

Summary

- Propose **blind detection** method for low pass graph signals.
- Provide sampling complexity analysis for proposed algorithm.
- Applications and experiments for informing **downstream** tasks.
  - Robustify graph learning.
- Application and experiment in network **dynamics identification**.
  - Antagonistic relationship in social networks.
  - FDIA detection in power systems.

Thank you!

Any questions or comments are welcomed!

References

[Anis et al., 2016] Anis, A., Gadde, A., and Ortega, A. (2016).
Efficient sampling set selection for bandlimited graph signals using graph spectral proxies.
IEEE Transactions on Signal Processing, 64(14):3775–3789.

[Deng et al., 2021] Deng, S., Ling, S., and Strohmer, T. (2021).
Strong consistency, graph laplacians, and the stochastic block model.
The Journal of Machine Learning Research, 22(1):5210–5253.

[Dong et al., 2016a] Dong, X., Thanou, D., Frossard, P., and Vandergheynst, P. (2016a).
Learning laplacian matrix in smooth graph signal representations.
IEEE Transactions on Signal Processing, 64(23):6160–6173.

[Dong et al., 2016b] Dong, X., Thanou, D., Frossard, P., and Vandergheynst, P. (2016b).
Learning laplacian matrix in smooth graph signal representations.
IEEE Transactions on Signal Processing, 64(23):6160–6173.

[He and Wai, 2022] He, Y. and Wai, H.-T. (2022).
Detecting central nodes from low-rank excited graph signals via structured factor analysis.
IEEE Transactions on Signal Processing.

[Kadavankandy et al., 2015] Kadavankandy, A., Cottatellucci, L., and Avrachenkov, K. (2015).
Characterization of random matrix eigenvectors for stochastic block model.
In *2015 49th Asilomar Conference on Signals, Systems and Computers*, pages 861–865. IEEE.

[Ramakrishna et al., 2020] Ramakrishna, R., Wai, H. T., and Scaglione, A. (2020).
A user guide to low-pass graph signal processing and its applications: Tools and applications.
IEEE Signal Processing Magazine, 37(6):74–85.

[Schaub et al., 2020] Schaub, M. T., Segarra, S., and Tsitsiklis, J. N. (2020).
Blind identification of stochastic block models from dynamical observations.
SIAM Journal on Mathematics of Data Science, 2(2):335–367.

[Segarra et al., 2017] Segarra, S., Marques, A. G., Mateos, G., and Ribeiro, A. (2017).
Network topology inference from spectral templates.
IEEE Transactions on Signal and Information Processing over Networks, 3(3):467–483.

[Segarra et al., 2016] Segarra, S., Mateos, G., Marques, A. G., and Ribeiro, A. (2016).
Blind identification of graph filters.
IEEE Transactions on Signal Processing, 65(5):1146–1159.

[Weng et al., 2013] Weng, L., Menczer, F., and Ahn, Y.-Y. (2013).
Virality prediction and community structure in social networks.
Scientific reports, 3(1):1–6.

[Wu et al., 2019] Wu, F., Souza, A., Zhang, T., Fifty, C., Yu, T., and Weinberger, K. (2019).
Simplifying graph convolutional networks.
In *ICML*. PMLR.

[Ye et al., 2018] Ye, C., Shafipour, R., and Mateos, G. (2018).
Blind identification of invertible graph filters with multiple sparse inputs.
In *2018 26th European Signal Processing Conference (EUSIPCO)*, pages 121–125. IEEE.

[Zhu et al., 2020] Zhu, Y., Garcia, F. J. I., Marques, A. G., and Segarra, S. (2020).
Estimating network processes via blind identification of multiple graph filters.
IEEE Transactions on Signal Processing, 68:3049–3063.