On Detecting Low-pass Graph Signals under Partial Observations

Hoang-Son Nguyen, Hoi-To Wai

The Chinese University of Hong Kong

Motivation

Modern networks are large — only a **portion** of nodes are observed for processing by GSP:

- *Partial observation* destroys smoothness/ eigenvector structure of graph signals.

GSP on partial observed graph signals require **smooth (a.k.a. low pass)** signals:

- Non-low-pass signals $\Rightarrow$ unexpected result!

$\Rightarrow$ Q: Without knowing the graph, is a dataset of partially observed graph signals low-pass?

TL;DR

To detect **low-pass property** of **partially observed graph signals**, giving a **data-driven** certificate for downstream GSP tools.

Preliminaries

Graph is undirected and connected, has N nodes, with adjacency $\mathbf{A}$, degree $\mathbf{D}$. Take as GSO the normalized Laplacian $\mathbf{L}_{\text{norm}} = \mathbf{I} - \mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^\top$, where $\mathbf{V} = [\mathbf{v}_1, ..., \mathbf{v}_N]$ are sorted by ascending eigenvalues.

Graph Filter. Let $h(\mathbf{\Lambda}) = \text{diag}(h(\lambda_1), ..., h(\lambda_n))$ be the frequency response,

$$\mathcal{H}(\mathbf{L}_{\text{norm}}) = \textstyle\sum_{t=0}^{T} h_t \mathbf{L}_{\text{norm}}^t = \mathbf{V} h(\mathbf{\Lambda}) \mathbf{V}^T,$$

Assume: $|h(\lambda_i)| \neq |h(\lambda_j)|$, $i \neq j$. By sorting $\{|h(\lambda_i)|\}_{i=1}^N$ as $|h_1| > ... > |h_N|$, $\mathcal{H}(\mathbf{L}_{\text{norm}}) = \mathbf{U}\mathbf{h}\mathbf{U}^\top$, with $\mathbf{h} = \text{diag}(h_1, ..., h_N)$, $\mathbf{U}$ is column re-ordered from $\mathbf{V}$.

Partially Observed Signals. Observed data are *filtered graph signals* with only the first n nodes retained,

$$\mathbf{y}_o = [\mathbf{I}_{n\times n}\ \mathbf{0}_{n\times(N-n)}]\mathbf{y} =: \mathbf{E}_o\mathbf{y}$$

where $\mathbf{y} = \mathcal{H}(\mathbf{L}_{\text{norm}})\mathbf{x} + \mathbf{w}$. Assume *stationary signals* s.t. $\mathbb{E}[\mathbf{x}] = \mathbf{0}, \mathbb{E}[\mathbf{x}\mathbf{x}^\top] = \mathbf{I}$, $\mathbb{E}[\mathbf{w}] = \mathbf{0}, \mathbb{E}[\mathbf{w}\mathbf{w}^\top] = \sigma^2\mathbf{I}$.

Notations. $\mathbf{R}_K \in \mathbb{R}^{K\times K}$ is an upper triangular matrix in QR factorization of $\mathbf{U}_{o,K} = \sqrt{N/n}\mathbf{Q}_K\mathbf{R}_K$. $\mathcal{T}_{\text{gnd}} \in \{\mathcal{T}_0, \mathcal{T}_1\}$ is the ground-truth hypothesis. $G \sim \text{SBM}(N, K, r, p)$ refers to SBM with N nodes, K blocks, r (resp. $r+p$) is inter (resp. intra) connectivity.

Problem Definition

Given $\{\mathbf{y}_{o,1}, ..., \mathbf{y}_{o,M}\}$, is the graph filter $\mathcal{H}(\mathbf{L}_{\text{norm}})$ **K-low-pass or not** (cf. Def. 1)?

- $\mathcal{T}_0$: $\mathcal{H}(\mathbf{L}_{\text{norm}})$ is K-low-pass
- $\mathcal{T}_1$: $\mathcal{H}(\mathbf{L}_{\text{norm}})$ is not K-low-pass

Def. 1: Low-pass Graph Filter

A graph filter $\mathcal{H}(\cdot)$ is K-low pass if

$$\eta_K = \frac{\max_{i=K+1,...,N} |h(\lambda_i)|}{\min_{i=1,...,K} |h(\lambda_i)|} < 1,$$

K is cut-off frequency, and η_K is sharpness.

Detecting Low-pass Signals with Partial Observations

Covariance Matrix. For $\mathbf{C}_o = \mathbb{E}[\mathbf{y}_o\mathbf{y}_o^\top]$,

$$\mathbf{C}_o = \mathbf{V}_o h(\mathbf{\Lambda})^2\mathbf{V}_o^\top + \sigma^2\mathbf{I} = \mathbf{U}_o\mathbf{h}^2\mathbf{U}_o^\top + \sigma^2\mathbf{I},$$

where $\mathbf{V}_o = \mathbf{E}_o\mathbf{V}, \mathbf{U}_o = \mathbf{E}_o\mathbf{U} \in \mathbb{R}^{n\times N}$. When the filter is *sharp* ($\eta_K \ll 1$ under $\mathcal{T}_0$),

$$\mathbf{C}_o - \sigma^2\mathbf{I} =: \overline{\mathbf{C}}_o \approx \mathbf{U}_{o,K}\mathbf{h}_K^2\mathbf{U}_{o,K}^\top,$$

where $\mathbf{U}_{o,K}$ takes K left columns from $\mathbf{U}_o$.

<u>**Observation 1:**</u> Let the top-K eigenvectors of the sampled covariance $\widehat{\mathbf{C}}_o$ be $\widehat{\mathbf{Q}}_K$, then

$$\widehat{\mathbf{Q}}_K \approx \mathbf{U}_{o,K} = \mathbf{E}_o\mathbf{U}_K$$

<u>**Observation 2:**</u> Is $\mathcal{H}$ low-pass? From [1],

- ($\mathcal{T}_0$) $\mathbf{U}_{o,K}$ corresponds to *row-sampled* ver. of $\mathbf{V}_K = [\mathbf{v}_1, ..., \mathbf{v}_K] \to$ <u>Clusterizable</u>
- ($\mathcal{T}_1$) $\mathbf{U}_{o,K}$ contains *row-sampled* version of $\{\mathbf{v}_{K+1}, ..., \mathbf{v}_N\} \to$ <u>Non-clusterizable</u>

<u>**Proposed Algorithm**</u>. define K-means score:

$$\mathbb{K}^*(\mathbf{N}) := \min_{\substack{\mathcal{C}_1,...,\mathcal{C}_K \\ \subseteq\{1,...,n\}}} \sum_{k=1}^{K}\sum_{i\in\mathcal{C}_k} \|\mathbf{n}_i^{\text{row}} - \frac{1}{|\mathcal{C}_k|}\sum_{j\in\mathcal{C}_k}\mathbf{n}_j^{\text{row}}\|_2^2,$$

where $\mathbf{n}_i^{\text{row}}$ is the ith row of $\mathbf{N}$. **Observe**:

$$\mathbb{K}^*(\widehat{\mathbf{Q}}_K) = \begin{cases} \text{small} & \text{, under } \mathcal{T}_0, \\ \text{large} & \text{, under } \mathcal{T}_1. \end{cases}$$

If $K \ll n$, **comp. complexity** $= \mathcal{O}(n^2(K+M))$,

$$\text{Complexity} = \mathcal{O}\left(\underbrace{n^2M}_{\text{form } \widehat{\mathbf{C}}_o} + \underbrace{n^2K}_{\text{find } \widehat{\mathbf{Q}}_K} + \underbrace{2^{K/\epsilon}Kn}_{\text{find } \mathbb{K}^*(\widehat{\mathbf{Q}}_K)}\right).$$

Sample Complexity Analysis

<u>**A1:**</u> W.h.p., there is ρ_{gap} such that $\lambda_{n-K-1}(\overline{\mathbf{C}}_o) - \lambda_{n-K}(\overline{\mathbf{C}}_o) - \|\widehat{\mathbf{C}}_o - \overline{\mathbf{C}}_o\|_2 \geq \rho_{\text{gap}} > 0$.

<u>**A2:**</u> $\mathcal{H}(\mathbf{L}_{\text{norm}})$ is at least η-sharp and γ-flat: $\frac{\max_{i=K+1,...,N}|h_i|}{\min_{i=1,...,K}|h_i|} \leq \eta < 1$, $\frac{\max_{1\leq i\leq K} h_i^2}{\min_{1\leq j\leq K} h_j^2} \leq \gamma$.

<u>**A3:**</u> $G \sim \text{SBM}(N, K, r, p)$ with $p \geq r > 0$, $p/K + r \geq (32\log N + 1)/N$.

<u>**A4:**</u> W.h.p., there is $c_{\text{SBM}} > 0$ such that $\min_{l=K+1,...,N} \mathbb{K}^*(\mathbf{v}_l) \geq c_{\text{SBM}}$.

Theorem 1: Sample Complexity

(A1-4) Let $\tilde{\delta}_{\min} := \min\Big\{\delta - \sqrt{\frac{N}{n}}\sqrt{\frac{1225K^3\log N}{p(N-K)}}, \sqrt{\frac{N}{n}}\sqrt{c_{\text{SBM}} - \frac{2450K^3\log N}{p(N-K)}} - \delta\Big\} > 0$,

$$\tilde{\sigma} := \frac{\rho_{\text{gap}}(\tilde{\delta}_{\min} - \sqrt{K}(\|\mathbf{I} - \mathbf{R}_K\|_2 + 6\gamma\eta))}{2\sqrt{K}} > \sigma^2.$$

If the number of samples M satisfies

$$\sqrt{M/\log M} = \Omega(1/\tilde{\sigma}),$$

then $\mathbb{P}(\widehat{\mathcal{T}} = \mathcal{T}_{\text{gnd}}) \geq 1 - 4/N - 5/M$.

Interpretations. When $\tilde{\delta}_{\min} > 0$ and $\tilde{\sigma} > 0$, then with a sufficiently large M, Algorithm 1 will return a correct detection w.h.p. as $N, M \to \infty$.

- To satisfy $\tilde{\delta}_{\min} > 0$, as $c_{\text{SBM}} = \Theta(1)$, we need $\delta = \mathcal{O}(\sqrt{\frac{N}{n}})$.
- To satisfy $\tilde{\sigma} > 0$, (i) noise level σ^2 is sufficiently small, and (ii) filter constant $\gamma\eta$, and $\|\mathbf{I} - \mathbf{R}_K\|_2$ are smaller than $\mathcal{O}(\tilde{\delta}_{\min})$ ($\|\mathbf{I} - \mathbf{R}_K\|_2 \searrow 0$ as $n \to N$; see Fig. 1).

The sample complexity is proportional to $\tilde{\sigma}^{-1}$, which is small when

- the no. of blocks K is small $\Rightarrow \sqrt{K} = \mathcal{O}(1)$,
- $\mathcal{H}(\mathbf{S})$ is *sharp and flat* $\Rightarrow \eta \ll 1, \gamma \approx 1$,
- no. of observed nodes $n \to N \Rightarrow \|\mathbf{I} - \mathbf{R}_K\|_2 \approx 0$.

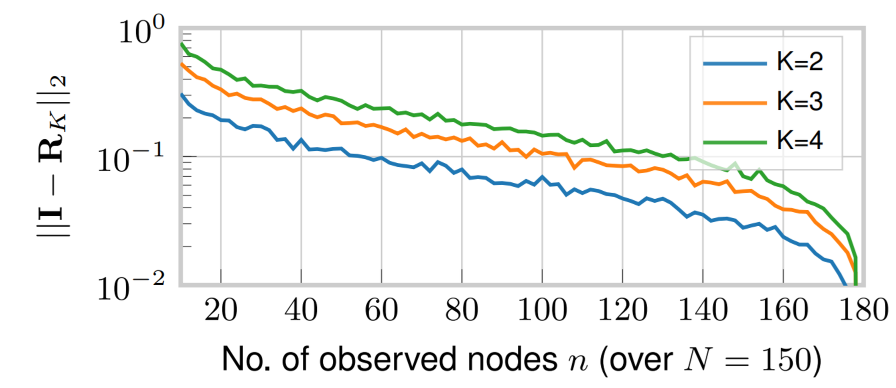

Fig. 1. Monte-Carlo simulation of $\|\mathbf{I} - \mathbf{R}_K\|_2$ as $n \to N$, where the corresponding $\mathbf{U}_{o,K} = \mathbf{Q}_K\mathbf{R}_K$ is from $\mathbf{L}_{\text{norm}}$ of a graph generated by $\text{SBM}(180, K, \log N/N, 4\log N/N)$, with $K \in \{2, 3, 4\}$.

Numerical Experiments

Graphs of $N = 150$ nodes and $K = 3$ blocks are generated from SBM, with n nodes selected uniformly at random for partial observations.

Detecting Low-pass Graph Signals. We test Alg. 1 in distinguishing signals by a low-pass filter $e^{-\tau\mathbf{L}_{\text{norm}}}$ vs. signals by a non-low-pass filter $e^{\tau\mathbf{L}_{\text{norm}}}$ (η decreases as $\tau > 0$ increases) $\Rightarrow$ when filter is **sharp**, performance is **insensitive** to n/N.

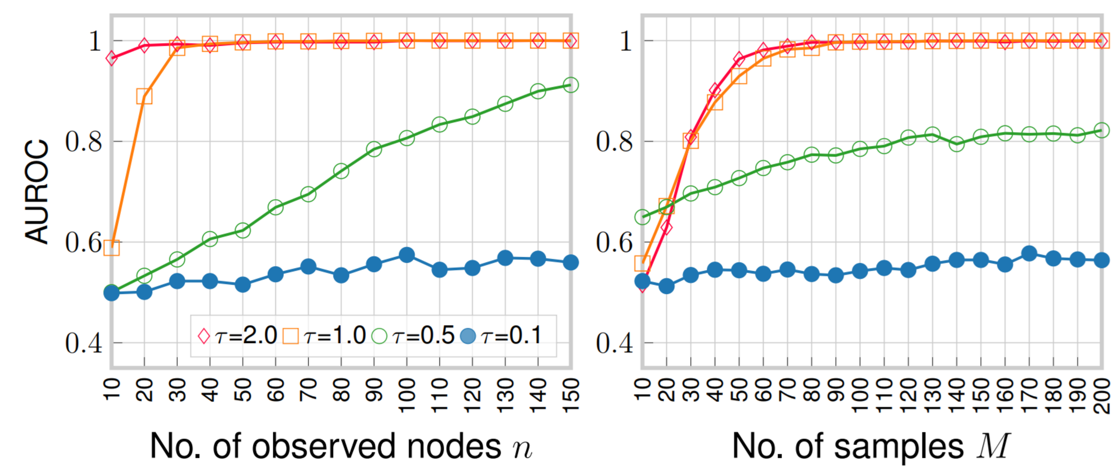

Fig. 2. Comparing low-pass detection performance against (left) no. of observed nodes n ($M = 100$), (right) no. of observed samples M ($n = 100$). The τ setting adjusts the sharpness of graph filters $e^{-\tau\mathbf{L}_{\text{norm}}}$ or $e^{\tau\mathbf{L}_{\text{norm}}}$.

Application: Robustifying Blind Community Detection. We use Alg. 1 to pre-screen potentially corrupted graph signals before applying [2] to detect communities from partial observations. Normal graph signals are generated from $\mathcal{H}(\mathbf{L}_{\text{norm}}) = (\mathbf{I} - 0.5\mathbf{L}_{\text{norm}})^3$, where 10% of samples are *corrupted* in bursts, such that p_s-fraction of nodal observations are replaced with Gaussian noise. To pre-screen, we apply Alg. 1 on small batches from $M = 10^3$ samples, and delete the batches detected as non-low-pass.

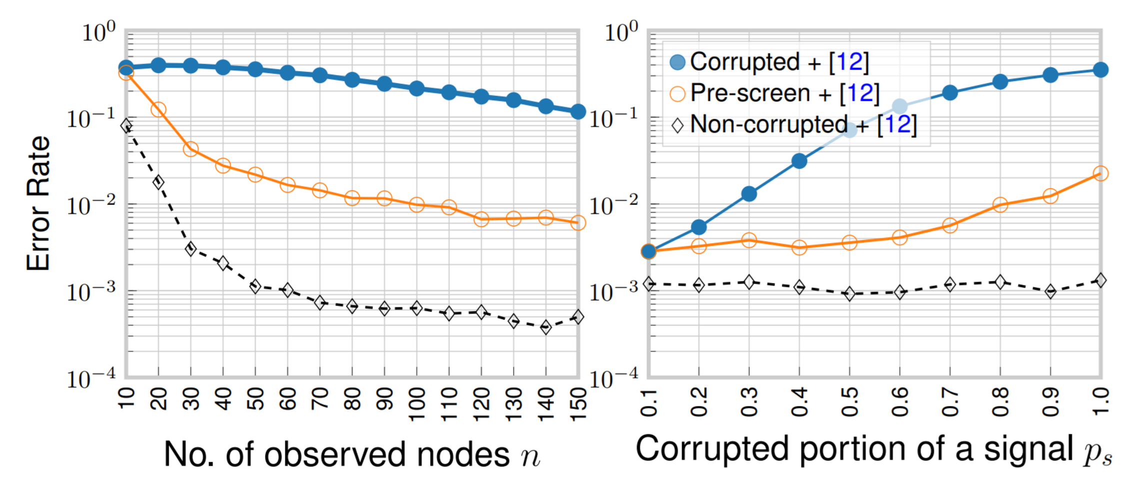

Fig. 3. Comparing blind community detection performance vs. (left) no. of observed nodes n ($p_s = 1$), (right) corrupted portion of signals p_s ($n = 50$).

References

[1] Zhang et al., "Detecting Low Pass Graph Signals via Spectral Pattern: Sampling Complexity and Applications", IEEE TSP, 2024.

[2] Wai et al., "Community inference from partially observed graph signals: Algorithms and analysis," IEEE TSP, 2022.