On Detecting Low-pass Graph Signals under Partial Observations

Hoang-Son Nguyen
Dept. of Computer Science & Engineering
The Chinese University of Hong Kong
sonngh.00@link.cuhk.edu.hk

Hoi-To Wai
Dept. of Systems Engineering & Engineering Management
The Chinese University of Hong Kong
htwai@se.cuhk.edu.hk

***Abstract*—The application of graph signal processing (GSP) on partially observed graph signals with missing nodes has gained attention recently. This is because processing data from large graphs are difficult, if not impossible due to the lack of availability of full observations. Many prior works have been developed using the assumption that the generated graph signals are smooth or low pass filtered. This paper treats a blind graph filter detection problem under this context. We propose a detector that certifies whether the partially observed graph signals are low pass filtered, without requiring the graph topology knowledge. As an example application, our detector leads to a pre-screening method to filter out non low pass signals and thus robustify the prior GSP algorithms. We also bound the sample complexity of our detector in terms of the class of filters, number of observed nodes, etc. Numerical experiments verify the efficacy of our method.**

***Index Terms*—graph signal processing, low pass graph filter, partial observations.**

I. Introduction

An important goal of graph signal processing (GSP) [1] is to extract insights from complex network data. Using graph shift operator (GSO), graph filters & signals as the underlying constructs, GSP has led to many theoretically justified graph learning methods [2], [3], e.g., prior works showed how to estimate the structure of weather [4] and brain networks [5]. Meanwhile, the overwhelming size of complex networks has necessitated practical methods to consider the *partial observation* setting where a fraction of nodes are never observed.

The partial observation setting may break a number of properties such as structure of eigenvectors, smoothness of graph signals, etc., that are necessary for graph learning. To this end, the early work [6] proposed to exploit the 'low-rank+sparse' structure in the precision matrix of partially observed graph signal. Subsequent work such as [7] proposed a graph learning criterion using smoothness of graph signals, [8] considered time-series data, [9] considered a linear influence model, and [10], [11] focused on identifiability of network dynamical systems. Additionally, the authors have studied graph feature learning from partial observations, such as community [12], central nodes [13]. In the above works, a common assumption made is that the graph signals are *smooth*, or more generally, generated from a network process that can be modeled as exciting a *low-pass graph filter* [14].

This work is supported in part by HKRGC Project #24203520.

While the *low-pass graph filter* assumption can be motivated by modeling network processes from social-physical aspects (e.g., [14], [15]), the latter often requires prior knowledge on the given dataset. In the absence of prior knowledge or when the dataset is corrupted, applying GSP methods may lead to unexpected results. Under this context, it is natural to ask: *Do we know if a dataset of partially observed graph signals is generated from a low-pass filter, without knowing the underlying graph beforehand?* Addressing the question gives a certificate *prior to* applying the mentioned methods on partially observed signals and guarantees reliable outcomes.

Our plan is to build on the authors' prior work [16], which tackled a similar detection problem but was focused on fully observed graph signals. Particularly, as the detection problem is ill-posed in general since smoothness/low-pass-ness are defined with respect to the graph itself, [16] focuses on a simplified case where the graph is known to be modular [17], a common feature for graphs found in networked systems. It then derives a detector based on the clusterizability of principal components, i.e., spectral pattern, for observed graph signals.

For partially observed graph signals, the challenge lies on how to account for the effect of missing nodes on the *observed* spectral pattern. To this end, our contributions are:

- We show that the K-means score detector in [16] can correctly detect the spectral pattern of *partially observed* low-pass graph signals. Though the latter also exhibit spectral pattern that distinguishes itself from any non-low-pass signals, we prove that the sampling complexity critically depends on the number of observed nodes.
- We demonstrate that the proposed detector can be used as a pre-screening procedure to robustify community detection from partially observed graph signals.

The rest of this paper is structured as follows. Section II describes the partial observation setting and formulates the blind detection problem. Section III develops the proposed method and reports its sample complexity. Finally, Section IV presents results from preliminary numerical experiments.

Notations. We use $||\cdot||_2$ to denote spectral norm for matrices and Euclidean norm for vectors, and $||\cdot||_{\rm F}$ to denote Frobenius matrix norm. For a symmetric matrix $\mathbf{X}$, $\lambda_i(\mathbf{X})$ denotes the i^{th} smallest eigenvalue of a matrix.

II. Problem Statement

Consider an undirected, connected N-node graph $G = (V, E)$ where $V = \{1, ..., N\}$ and $E \subseteq V \times V$. The graph can be represented as an adjacency matrix $\mathbf{A} \in \{0,1\}^{N\times N}$, a Laplacian matrix $\mathbf{L} = \mathbf{D} - \mathbf{A}$ where $\mathbf{D} = \text{diag}(\mathbf{A1})$, or a normalized Laplacian matrix $\mathbf{L}_{\text{norm}} = \mathbf{I} - \mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2}$. The aforementioned matrix representations of G qualify as graph shift operators (GSO), which are any symmetric matrix $\mathbf{S} \in \mathbb{R}^{N\times N}$ such that $\mathbf{S}_{ij} \neq 0$ only if $(i, j) \in E$. A GSO admits an eigendecomposition $\mathbf{S} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^\top$, where the columns of $\mathbf{V} = [\mathbf{v}_1, \cdots, \mathbf{v}_N]$ are the orthonormal eigenvectors associated with eigenvalues sorted in ascending order, and $\mathbf{\Lambda}$ is a diagonal matrix of eigenvalues, also known as the graph frequencies. We focus on the case of $\mathbf{S} = \mathbf{L}_{\text{norm}}$.

The graph filter is defined as a polynomial of the GSO:

$$\mathcal{H}(\mathbf{S}) = \textstyle\sum_{t=0}^{T} h_t \mathbf{S}^t = \mathbf{V} h(\mathbf{\Lambda}) \mathbf{V}^\top, \tag{1}$$

where $\{h_t\}_{t=0}^{T-1}$ are the filter coefficients. The latter also defines the frequency response function: $h(\lambda) = \sum_{t=0}^{T} h_t \lambda^t$ and $h(\mathbf{\Lambda}) = \text{diag}(h(\lambda_1), ..., h(\lambda_n))$. For simplicity, we assume the magnitudes of frequency responses to be distinct, i.e. $|h(\lambda_i)| \neq |h(\lambda_j)|$ for all $i \neq j$. By sorting the magnitude of frequency responses in descending order as $|h_1| > ... > |h_N|$, the graph filter operator can be written as $\mathcal{H}(\mathbf{S}) = \mathbf{U}\mathbf{h}\mathbf{U}^\top$, where $\mathbf{h} = \text{diag}(h_1, ..., h_N)$ and $\mathbf{U}$ is accordingly the column re-ordered version of $\mathbf{V}$.

A graph signal on G is represented as a N-dimensional vector that is the output of a graph filter (1):

$$\mathbf{y} = \mathcal{H}(\mathbf{S})\mathbf{x} + \mathbf{w}. \tag{2}$$

The ith element of $\mathbf{y}$ denotes the signal on node $i \in V$, subjected to the excitation graph signal $\mathbf{x} \in \mathbb{R}^N$, and $\mathbf{w} \in \mathbb{R}^n$ is a zero-mean white noise with $\mathbb{E}[\mathbf{w}\mathbf{w}^\top] = \sigma^2\mathbf{I}$ for $\sigma^2 \geq 0$. The observed signal $\mathbf{y}$ is assumed to be stationary, i.e., $\mathbb{E}[\mathbf{x}] = \mathbf{0}$ and $\mathbb{E}[\mathbf{x}\mathbf{x}^\top] = \mathbf{I}$ [18], [19] for simplicity; however, our analysis can be extended to the non-stationary case of $\mathbb{E}[\mathbf{x}\mathbf{x}^\top] \neq \mathbf{I}$.

The graph signal $\mathbf{y}$ in (2) can also be modeled as the output of a network process. Prior works in GSP have suggested to categorize network process according to their frequency response. Among others, an important class of graph filters is the low-pass graph filters [14], [20], which is defined by:

Definition 1. *A graph filter $\mathcal{H}(\cdot)$ is said to be K-low-pass if*

$$\eta_K := \frac{\max_{i=K+1,...,N} |h(\lambda_i)|}{\min_{i=1,...,K} |h(\lambda_i)|} < 1, \tag{3}$$

where K is cut-off frequency, and η_K is the sharpness of $\mathcal{H}$.

From the definition, a low-pass graph filter retains (resp. attenuates) the energy of the excitation graph signal at low (resp. high) frequencies. A graph signal is said to be low-pass if it is the output of a low-pass graph filter.

We further consider the scenario when the graph signals in (2) can only be partially observed. Without loss of generality, we assume that the first n nodes are observed and denote

$$\mathbf{y}_o = [\mathbf{I}_{n\times n}, \mathbf{0}_{n\times(N-n)}]\,\mathbf{y} =: \mathbf{E}_o\mathbf{y}. \tag{4}$$

As mentioned in the Introduction, the application of GSP on partial observations has gained popularity as the model arises naturally for large graphs where it is difficult to obtain observations on every nodes. Under this context, GSP applications such as graph learning [7], community detection [12] have exploited the *smoothness* property and motivate the latter by modeling the graph signal observations as *low pass signals*.

We depart from the above works and inquire if the *smoothness* property is valid for a given dataset. This leads to the blind low-pass graph filter detection problem:

Problem 1. *Given the parameter K and a set of partially observed graph signals [cf. (2), (4)], determine if the underlying graph filter is K-low-pass or not [cf. Definition 1]. We denote the null hypothesis $\mathcal{T}_0$ (resp. alternative hypothesis $\mathcal{T}_1$) as '$\mathcal{H}(\mathbf{S})$ is (resp. not) K-low-pass'.*

Notice that Problem 1 serves as a data-driven certificate to the successful applications of the prior GSP works.

There are two challenges in solving Problem 1: (i) the graph topology or the GSO $\mathbf{S}$ is unknown, (ii) the graph signals are partially observed where $\mathbf{E}_o$ is unknown. Either challenge has made it impossible to verify Definition 1 directly. Our prior work [16] proposed to narrow down the detection problem w.r.t. arbitrary graphs to the class of K-modular graphs [17] with K densely connected components. The number of densely connected components naturally determines the parameter K for the low-pass filter. It then exploits the spectral pattern of graphs to formulate a K-means score detector[1]. The detector is proven to produce accurate result under mild assumptions on the noise statistics and graph filter properties.

This paper aims to extend the aforementioned detector in [16] to the partial observation context. Interestingly, we show that the K-means score detector is still robust in this scenario, whose performance loss depends naturally with the ratio n/N.

III. Low-pass Detection with Partial Observations

This section develops a detector for Problem 1 under the partial observation setting. Our development begins by investigating the *covariance matrix* of partially observed signals:

$$\begin{aligned}\mathbf{C}_o = \mathbb{E}[\mathbf{y}_{o,m}\mathbf{y}_{o,m}^\top] &= \mathbf{V}_o h(\mathbf{\Lambda})^2 \mathbf{V}_o^\top + \sigma^2\mathbf{I} \\ &= \mathbf{U}_o \mathbf{h}^2 \mathbf{U}_o^\top + \sigma^2\mathbf{I},\end{aligned} \tag{5}$$

where we have used $\mathbf{y}_{o,m}$ to denote the mth realization of the partially observed signal in (4). We have defined the *row-sampled* eigenmatrices $\mathbf{V}_o = \mathbf{E}_o\mathbf{V}, \mathbf{U}_o = \mathbf{E}_o\mathbf{U} \in \mathbb{R}^{n\times N}$. The noiseless covariance is $\overline{\mathbf{C}}_o = \mathbf{C}_o - \sigma^2\mathbf{I} = \mathbf{U}_o\mathbf{h}^2\mathbf{U}_o^\top$.

Following the insight from [12], we note that when the graph filter has a *sharp* cut-off (e.g., $\eta_K \ll 1$ under $\mathcal{T}_0$), the following approximation holds

$$\overline{\mathbf{C}}_o \approx \mathbf{U}_{o,K}\mathbf{h}_K^2\mathbf{U}_{o,K}^\top, \tag{6}$$

where $\mathbf{U}_{o,K}$ takes the K left-most column vectors from $\mathbf{U}_o$. Under $\mathcal{T}_0$, the matrix $\mathbf{U}_{o,K}$ corresponds to the row-sampled

[1]We remark that when $K = 1$, i.e., the graph contains only one dense component, applying the Perron Frobenius theorem [21] suffices to detect the 1-low-pass graph signals. Here, we shall focus on the case of $K \geq 2$.

Algorithm 1 Low-pass Detection with Partial Observations

1: **Input:** Partially observed graph signals $\{\mathbf{y}_{o,m}\}_{m=1}^{M}$, no. of clusters $K \geq 2$, detection threshold $\delta > 0$.
2: Calculate $\widehat{\mathbf{C}}_o := (1/M)\sum_{m=1}^{M} \mathbf{y}_{o,m}\mathbf{y}_{o,m}^\top$.
3: Compute the top-K eigenvectors $\widehat{\mathbf{Q}}_K \in \mathbb{R}^{n\times K}$ of $\widehat{\mathbf{C}}_o$.
4: **Output:** $\widehat{\mathcal{T}} = \mathcal{T}_0$ if $\mathbb{K}^*(\widehat{\mathbf{Q}}_K) < \delta$; or $\widehat{\mathcal{T}} = \mathcal{T}_1$ otherwise.

and column permuted version of $\mathbf{V}_K = [\mathbf{v}_1, ..., \mathbf{v}_K]$. To this end, the *row vectors* of $\mathbf{V}_K$ are clusterizable when G is K-modular. Meanwhile, under $\mathcal{T}_1$ when the graph filter is *not* K-low-pass, $\mathbf{U}_{o,K}$ corresponds to the row sampled versions of the bulk eigenvectors $\{\mathbf{v}_{K+1}, ..., \mathbf{v}_N\}$ which are not clusterizable [22]. Together, they motivate the following K-means score: for any $\mathbf{N} \in \mathbb{R}^{N\times K}$, we denote

$$\begin{aligned} &\mathbb{K}^*(\mathbf{N}) := \min_{\mathcal{C}} \mathbb{K}(\mathbf{N}, \mathcal{C}), \\ &\mathbb{K}(\mathbf{N}, \mathcal{C}) := \textstyle\sum_{k=1}^{K} \sum_{i\in\mathcal{C}_k} ||\mathbf{n}_i^{\text{row}} - \frac{1}{|\mathcal{C}_k|}\sum_{j\in\mathcal{C}_k} \mathbf{n}_j^{\text{row}}||_2^2, \end{aligned} \tag{7}$$

where $\mathcal{C} = \{\mathcal{C}_1, \ldots, \mathcal{C}_K\}$ is a set of non-overlapping partition for $\{1, ..., N\}$ and $\mathbf{n}_i^{\text{row}}$ denotes the ith row vector of $\mathbf{N}$.

Define the sampled covariance matrix $\widehat{\mathbf{C}}_o := (1/M)\sum_{m=1}^{M} \mathbf{y}_{o,m}\mathbf{y}_{o,m}^\top$ and its top-K eigenvectors are stacked up as $\widehat{\mathbf{Q}}_K$. Following the insights from [12], [16] and observe that $\widehat{\mathbf{Q}}_K \approx \mathbf{U}_{o,K}$ when n is sufficiently close to N, we propose to tackle Problem 1 by detecting $\mathcal{T}_0$/$\mathcal{T}_1$ based on $\mathbb{K}^*(\widehat{\mathbf{Q}}_K)$. From the above discussions, $\mathbb{K}^*(\widehat{\mathbf{Q}}_K)$ will be small (resp. large) when the graph filter is (resp. not) K-low-pass. This motivates the proposed detector in Algorithm 1.

A. Performance Analysis and Theoretical Insights

We next present the analysis on the finite-sample performance of Algorithm 1. In addition to verifying the correctness of the detector, our analysis shall demonstrate the favorable conditions where the detector is effective. Note that there are multiple sources of error that need to be controlled carefully. For instance, the approximation in (6) is not exact, the columns of $\mathbf{U}_{o,K}$ are not orthogonal, etc.

To set up the analysis, we require the following condition on spectral gap of the covariance matrix:

H1. *With probability at least* $1 - \delta_{\text{gap}}$*, there exists* ρ_{gap} *such that* $\lambda_{n-K-1}(\overline{\mathbf{C}}_o) - \lambda_{n-K}(\overline{\mathbf{C}}_o) - ||\widehat{\mathbf{C}}_o - \overline{\mathbf{C}}_o||_2 \geq \rho_{\text{gap}} > 0$.

The above can be satisfied when $\widehat{\mathbf{C}}_o$ is sufficiently close to $\overline{\mathbf{C}}_o$, e.g., when sufficient number of samples are observed and the noise level σ^2 is small, and the noiseless covariance $\overline{\mathbf{C}}_o$ is approximately rank K. We also let:

H2. *The graph filter* $\mathcal{H}(\mathbf{S})$ *is at least* η*-sharp and* γ*-flat:*

$$\frac{\max_{i=K+1,\ldots,N} |h_i|}{\min_{i=1,\ldots,K} |h_i|} \leq \eta < 1, \quad \frac{\max_{1\leq i\leq K} h_i^2}{\min_{1\leq j\leq K} h_j^2} \leq \gamma. \tag{8}$$

The above specifies the class of graph filters that we detect. Notice if the graph filter is K-low-pass, then the above η takes the same role as η_K in (3).

As mentioned in the previous section, the proposed detector relies on the clusterizability of the top-K eigenvectors for the normalized Laplacian in K-modular graphs. To obtain theoretical insights, we assume that the full graph G is generated from the stochastic block model (SBM) with:

H3. *We have* $G \sim$ *SBM*(N, K, r, p) *with* $p \geq r > 0$, $p/K + r \geq (32\log N + 1)/N$.

By $G \sim \text{SBM}(N, K, r, p)$, we denote a random graph with N nodes equally partitioned into K blocks, described by a membership matrix $\mathbf{Z} \in \{0, 1\}^{N\times K}$ such that $\mathbf{Z}_{ij} = 1$ if and only if node i is in block j, and a connectivity matrix $\mathbf{B} \in [0, 1]^{K\times K}$, whose entries $\mathbf{B}_{ij}$ being the probability of edges between nodes in block i and block j. The parameters r, p describes the connectivity such that $\mathbf{B} = p\mathbf{I} + r\mathbf{1}\mathbf{1}^\top$. We also assume the following on the *bulk eigenvectors* of $\mathbf{L}_{\text{norm}}$:

H4. *With probability at least* $1 - \delta_{SBM}$*, there exists* $c_{SBM} > 0$ *independent of* N, r, p *with* $\min_{l=K+1,\ldots,N} \mathbb{K}^*(\mathbf{v}_l) \geq c_{SBM}$.

Note that H4 is observed for $G \sim \text{SBM}(N, K, r, p)$ empirically [16], yet it remains an open conjecture to be verified theoretically. With H3, H4, it is easy to deduce that $\mathbb{K}^*(\mathbf{V}_K) = \mathcal{O}(\log N/N)$ [23], while the K-means score for the bulk eigenvectors is bounded away by $c_{\text{SBM}} > 0$.

Let $\mathcal{T}_{\text{gnd}} \in \{\mathcal{T}_0, \mathcal{T}_1\}$ be the ground truth hypothesis. Our main analytical result is summarized below:

Theorem 1. *Under H1, H2, H3, H4. Suppose that the following threshold-dependent term satisfies*

$$\begin{aligned} \tilde{\delta}_{\min} := \min \Bigg\{ &\delta - \sqrt{\frac{N}{n}}\sqrt{\frac{1225K^3\log N}{p(N-K)}}, \\ &\sqrt{\frac{N}{n}}\sqrt{c_{SBM} - \frac{2450K^3\log N}{p(N-K)}} - \delta \Bigg\} > 0, \end{aligned}$$

and that

$$\tilde{\sigma} := \frac{\rho_{\text{gap}}(\tilde{\delta}_{\min} - \sqrt{K}(||\mathbf{I} - \mathbf{R}_K||_2 + 6\gamma\eta))}{2\sqrt{K}} - \sigma^2 > 0,$$

where $\mathbf{R}_K \in \mathbb{R}^{K\times K}$ *is an upper triangular matrix in the QR factorization of* $\mathbf{U}_{o,K} = \sqrt{N/n}\mathbf{Q}_K\mathbf{R}_K$*. If the number of samples* M *satisfies*

$$\sqrt{\frac{M}{\log M}} \geq \frac{\sqrt{2}c_1 \operatorname{tr}(\overline{\mathbf{C}}_o)}{\tilde{\sigma}} \tag{9}$$

where c_1 *is a constant independent of* N, M*, then we have*

$$\mathbb{P}(\widehat{\mathcal{T}} = \mathcal{T}_{\text{gnd}}) \geq 1 - 4/N - 5/M - \delta_{\text{gap}} - \delta_{SBM}. \tag{10}$$

The above result considers randomnesses in the graph signals generation (2) and the SBM graph properties in H3, H4.

The theorem asserts that when $\tilde{\delta}_{\min} > 0$, $\tilde{\sigma} > 0$, then with a sufficiently large number of samples, Algorithm 1 will return a correct detection with high probability as $N, M \to \infty$. To satisfy $\tilde{\delta}_{\min} > 0$, as $c_{\text{SBM}} = \Theta(1)$, the requirement can be fulfilled with $\delta = \Theta(\sqrt{N/n})$. Furthermore, to satisfy $\tilde{\sigma} > 0$, we require two criterion: (i) the noise level σ^2 is sufficiently small, (ii) the filter constant $\gamma\eta$, and the factor $\|\mathbf{I} - \mathbf{R}_K\|_2$ are smaller than $\mathcal{O}(\tilde{\delta}_{\min})$. Note that $\|\mathbf{I} - \mathbf{R}_K\|_2$ decreases to

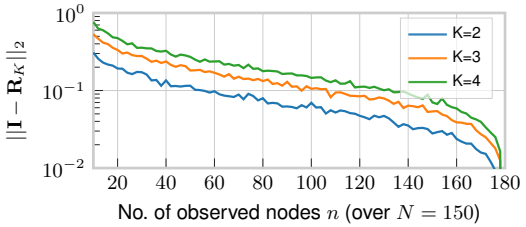

Fig. 1. Monte-Carlo simulation of $||\mathbf{I} - \mathbf{R}_K||_2$ as $n \to N$, where the corresponding $\mathbf{U}_{o,K} = \mathbf{Q}_K \mathbf{R}_K$ is from $\mathbf{L}_{\text{norm}}$ of a graph generated by $\text{SBM}(180, K, \log N/N, 4 \log N/N)$, with $K \in \{2, 3, 4\}$.

0 as $n \to N$; see Fig. 1 for illustration. Finally, we note that the sample complexity, i.e., minimal M needed to satisfy (9), is proportional to $\tilde{\sigma}^{-1}$. From the above discussions, $\tilde{\sigma}^{-1}$ is reduced when the graph filters to be detected are *sharp and flat*, i.e., $\eta \ll 1, \gamma \approx 1$, and the number of observed nodes is large enough $n \to N$. Lastly, we remark that the proof for Theorem 1 is adapted from our prior works [12], [16] which applied [23], [24]. It can be found in the online appendix[2].

IV. Numerical Experiments

This section presents numerical experiments to validate our findings. We first evaluate the direct detection performance in tackling Problem 1, then we consider an application on robustifying the blind community detection method.

A. Detecting Low-pass Signals from Partial Observations

We use synthetic data to evaluate the performance of our proposed detector in various settings. In the following experiment, the graph G with $N = 150$ nodes and $K = 3$ blocks is generated according to H3 such that $G \sim \text{SBM}(150, 3, \log N/N, 4 \log N/N)$. The full graph signals in (2) are generated with $\mathbf{x} \in \mathbb{R}^N \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ and $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I})$ where $\sigma^2 = 10^{-2}$, then we select n nodes uniformly at random to form the partial observations (4). We benchmark Algorithm 1 in distinguishing signals generated by a low-pass filter $e^{-\tau \mathbf{L}_{\text{norm}}}$ (null hypothesis $\mathcal{T}_0$) from signals by a non-low-pass filter $e^{\tau \mathbf{L}_{\text{norm}}}$ (alternative hypothesis $\mathcal{T}_1$), where the sharpness of the filter η decreases as $\tau > 0$ increases. The performance is measured by the area under ROC (AUROC) such that $\text{AUROC} = 1$ when the detection is perfect. Figure 2 reports the results from 1000 Monte-Carlo trials.

We observe that the performance improves as n, M increases, as well as the sharpness parameter η controlled by τ. Moreover, Algorithm 1 delivers reliable performance (with $\text{AUROC} \approx 1$) when $n \geq 100, M \geq 100$. This indicates that the spectral pattern of low-pass graph signals are significant enough despite that $1/3$ of the nodes are not observed and only $M \approx n$ samples are observed. The above observations coincide with our finite-sample analysis in Theorem 1.

B. Application: Robustifying Blind Community Detection

We illustrate an application of Algorithm 1 as a pre-screening procedure before applying prior work that demands low-pass graph signals. We consider the blind community detection method [12] which directly infer communities in a graph from low-pass graph signals that are partially observed. To satisfy the low-pass graph signal requirement, we apply Algorithm 1 on *small batches* of M_{batch} graph signal observations and retain (resp. drop) the small batches that are identified as low-pass (resp. non-low-pass). The pre-screened dataset is then provided to [12] to infer the communities.

[2] https://www1.se.cuhk.edu.hk/~htwai/pdf/sam24-appendix.pdf

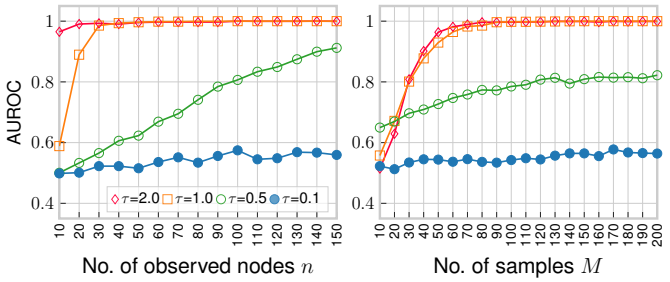

Fig. 2. Comparing low-pass detection performance against (left) no. of observed nodes n ($M = 100$), (right) no. of observed samples M ($n = 100$). The τ setting adjusts the sharpness of graph filters $e^{-\tau \mathbf{L}_{\text{norm}}}$ or $e^{\tau \mathbf{L}_{\text{norm}}}$.

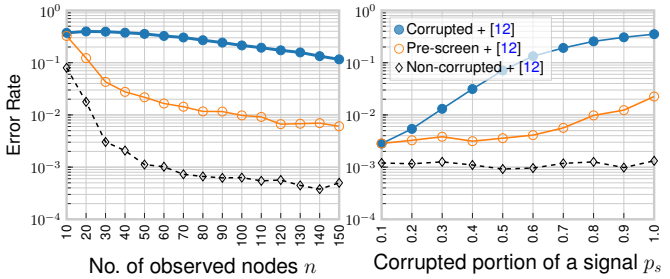

Fig. 3. Comparing blind community detection performance vs. (left) no. of observed nodes n ($p_s = 1$), (right) corrupted portion of signals p_s ($n = 50$).

We consider $G \sim \text{SBM}(150, 3, \log N/N, 7 \log N/N)$, with $N = 150$ nodes and $K = 3$ clusters. The *normal* graph signals are generated using (2), (4) with $\sigma^2 = 10^{-2}$ and the filter $\mathcal{H}(\mathbf{S}) = (\mathbf{I} - 0.5\mathbf{L}_{\text{norm}})^3$, where 10% of samples are *corrupted* in a burst of length $m_{\text{burst}} = 10$, such that p_s-fraction of nodal observations are replaced with $\mathcal{N}(0, 1)$. For the pre-screening procedure, we apply Algorithm 1 on small batches of size $M_{\text{batch}} = 50$ from $M = 10^3$ samples, with $\delta = 0.5$. Figure 3 reports the results of 1000 Monte-Carlo trials.

Observe the dataset corruption severely affects the performance of blind community detection [12]. Meanwhile, our pre-screening procedure robustifies the method in [12]. We note the effectiveness of pre-screening improves with n as it approaches the performance of non-corrupted dataset, coinciding with Theorem 1 that low-pass detection becomes more accurate as n increases. Pre-screening also delivers consistent improvement across different levels of signal corruption.

Conclusions. This paper studies the low-pass graph signal detection problem with partial observations. We showed that a simple K-means score detector can distinguish spectral pattern of the low-pass/non-low-pass signals and analyzed its sample complexity. Our work can robustify GSP on partially observed signals. Future work includes deriving an explicit bound w.r.t. no. of observed nodes n and explore other applications.

REFERENCES

[1] A. Ortega, P. Frossard, J. Kovačević, J. M. Moura, and P. Vandergheynst, "Graph signal processing: Overview, challenges, and applications," *Proceedings of the IEEE*, vol. 106, no. 5, pp. 808–828, 2018.

[2] G. Mateos, S. Segarra, A. G. Marques, and A. Ribeiro, "Connecting the dots: Identifying network structure via graph signal processing," *IEEE Signal Processing Magazine*, vol. 36, no. 3, pp. 16–43, 2019.

[3] X. Dong, D. Thanou, M. Rabbat, and P. Frossard, "Learning graphs from data: A signal representation perspective," *IEEE Signal Processing Magazine*, vol. 36, no. 3, pp. 44–63, 2019.

[4] D. Thanou, X. Dong, D. Kressner, and P. Frossard, "Learning heat diffusion graphs," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 3, no. 3, pp. 484–499, 2017.

[5] W. Huang, T. A. Bolton, J. D. Medaglia, D. S. Bassett, A. Ribeiro, and D. Van De Ville, "A graph signal processing perspective on functional brain imaging," *Proceedings of the IEEE*, vol. 106, no. 5, pp. 868–885, 2018.

[6] V. Chandrasekaran, P. A. Parrilo, and A. S. Willsky, "Latent variable graphical model selection via convex optimization," *The Annals of Statistics*, pp. 1935–1967, 2012.

[7] A. Buciulea, S. Rey, and A. G. Marques, "Learning graphs from smooth and graph-stationary signals with hidden variables," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 8, pp. 273–287, 2022.

[8] A. Jalali and S. Sanghavi, "Learning the dependence graph of time series with latent factors," *arXiv preprint arXiv:1106.1887*, 2011.

[9] V. Matta, A. Santos, and A. H. Sayed, "Graph learning under partial observability," *Proceedings of the IEEE*, vol. 108, no. 11, pp. 2049–2066, 2020.

[10] J. M. Hendrickx, M. Gevers, and A. S. Bazanella, "Identifiability of dynamical networks with partial node measurements," *IEEE Transactions on Automatic Control*, vol. 64, no. 6, pp. 2240–2253, 2018.

[11] A. Santos, D. Rente, R. Seabra, and J. M. Moura, "Learning the causal structure of networked dynamical systems under latent nodes and structured noise," *arXiv preprint arXiv:2312.05974*, 2023.

[12] H.-T. Wai, Y. C. Eldar, A. E. Ozdaglar, and A. Scaglione, "Community inference from partially observed graph signals: Algorithms and analysis," *IEEE Transactions on Signal Processing*, vol. 70, pp. 2136–2151, 2022.

[13] Y. He and H.-T. Wai, "Central nodes detection from partially observed graph signals," in *2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2023, pp. 1–5.

[14] R. Ramakrishna, H.-T. Wai, and A. Scaglione, "A user guide to low-pass graph signal processing and its applications: Tools and applications," *IEEE Signal Processing Magazine*, vol. 37, no. 6, pp. 74–85, 2020.

[15] M. H. DeGroot, "Reaching a consensus," *Journal of the American Statistical association*, vol. 69, no. 345, pp. 118–121, 1974.

[16] C. Zhang, Y. He, and H.-T. Wai, "Detecting low pass graph signals via spectral pattern: Sampling complexity and applications," *arXiv preprint arXiv::2306.01553*, 2023.

[17] M. Girvan and M. E. J. Newman, "Community structure in social and biological networks," *Proceedings of the National Academy of Sciences*, vol. 99, no. 12, p. 7821–7826, Jun. 2002.

[18] N. Perraudin and P. Vandergheynst, "Stationary signal processing on graphs," *IEEE Transactions on Signal Processing*, vol. 65, no. 13, pp. 3462–3477, 2017.

[19] A. G. Marques, S. Segarra, G. Leus, and A. Ribeiro, "Stationary graph processes and spectral estimation," *IEEE Transactions on Signal Processing*, vol. 65, no. 22, pp. 5911–5926, 2017.

[20] A. Sandryhaila and J. M. Moura, "Discrete signal processing on graphs," *IEEE transactions on signal processing*, vol. 61, no. 7, pp. 1644–1656, 2013.

[21] Y. He and H.-T. Wai, "Identifying first-order lowpass graph signals using perron frobenius theorem," in *2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2021, pp. 5285–5289.

[22] U. von Luxburg, "A tutorial on spectral clustering," *Statistics and Computing*, vol. 17, no. 4, p. 395–416, Aug. 2007.

[23] K. Rohe, S. Chatterjee, and B. Yu, "Spectral clustering and the high-dimensional stochastic blockmodel," *The Annals of Statistics*, vol. 39, no. 4, Aug. 2011.

[24] F. Bunea and L. Xiao, "On the sample covariance matrix estimator of reduced effective rank population matrices, with applications to fPCA," *Bernoulli*, vol. 21, no. 2, pp. 1200 – 1230, 2015.

APPENDIX: PROOF OF THEOREM 1

Additional Notations. In the following analysis, we define the QR decomposition of $\mathbf{U}_{o,K}$ as $\mathbf{U}_{o,K} = c_0\mathbf{Q}_K\mathbf{R}_K$, in which $c_0 = \sqrt{n/N}$ is a normalization parameter, $\mathbf{Q}_K \in \mathbb{R}^{n\times K}$ is an orthogonal matrix spanning the range of $\mathbf{U}_{o,K}$ and $\mathbf{R}_K$ is upper-triangular. We also set the diagonal of $\mathbf{R}_K$ to be non-negative. Moreover, under H3, $\mathbf{A}$ satisfies $\mathbb{E}[\mathbf{A}] = \mathbf{ZBZ}^\top =: \mathcal{A}$; the population normalized Laplacian matrix of the SBM is then $\mathcal{L}_{\text{norm}} = \mathbf{I} - \mathcal{D}^{-1/2}\mathcal{A}\mathcal{D}^{-1/2}$, where $\mathcal{D} = \text{diag}(\sum_{j=1}^N \mathcal{A}_{1j}, ..., \sum_{j=1}^N \mathcal{A}_{Nj})$. Lastly, the set $\mathcal{R}_K^{m\times n}$ consists of $m \times n$ matrices having at most K unique rows.

Our proof is adapted from [12], [16]. First, consider the ground truth as $\mathcal{T}_{\text{gnd}} = \mathcal{T}_0$, which implies $\mathbf{U}_K = \mathbf{V}_K\mathbf{\Pi}$ for some permutation matrix $\mathbf{\Pi}$. Define the indicator matrix $\mathbf{X}^*$ be associated with the partition $\mathcal{C}^* \in \arg\min_{\mathcal{C}} \mathbb{K}(\mathbf{Q}_K, \mathcal{C})$:

$$\mathbf{X}^*_{ik} := \begin{cases} 1/\sqrt{|\mathcal{C}_i^*|} & \text{if } i \in \mathcal{C}_i^*, \\ 0 & \text{otherwise.} \end{cases}$$

We observe that

$$\begin{aligned}
&\sqrt{\mathbb{K}^*(\widehat{\mathbf{Q}}_K)} \le ||(\mathbf{I} - \mathbf{X}^*(\mathbf{X}^*)^\top)\widehat{\mathbf{Q}}_K||_{\text{F}} \\
&= ||(\mathbf{I} - \mathbf{X}^*(\mathbf{X}^*)^\top)\widehat{\mathbf{Q}}_K\widehat{\mathbf{Q}}_K^\top||_{\text{F}} \\
&\le ||(\mathbf{I} - \mathbf{X}^*(\mathbf{X}^*)^\top)\mathbf{Q}_K\mathbf{Q}_K^\top||_{\text{F}} + ||\mathbf{Q}_K\mathbf{Q}_K^\top - \widehat{\mathbf{Q}}_K\widehat{\mathbf{Q}}_K||_{\text{F}} \\
&= \sqrt{\mathbb{K}^*(\mathbf{Q}_K)} + ||\mathbf{Q}_K\mathbf{Q}_K^\top - \widehat{\mathbf{Q}}_K\widehat{\mathbf{Q}}_K||_{\text{F}}.
\end{aligned}$$

Similarly, we further have

$$\begin{aligned}
&\sqrt{\mathbb{K}^*(\mathbf{Q}_K)} \le c_0^{-1}\sqrt{\mathbb{K}^*(\mathbf{U}_{o,K})} + ||\mathbf{Q}_K - c_0^{-1}\mathbf{U}_{o,K}||_{\text{F}} \\
&\le c_0^{-1}\sqrt{\mathbb{K}^*(\mathbf{U}_{o,K})} + ||\mathbf{Q}_K||_{\text{F}}||\mathbf{I} - \mathbf{R}_K||_2 \\
&= c_0^{-1}\sqrt{\mathbb{K}^*(\mathbf{U}_{o,K})} + \sqrt{K}||\mathbf{I} - \mathbf{R}_K||_2.
\end{aligned}$$

Define the orthogonal matrix $\overline{\mathcal{O}}_K = \mathcal{O}_K\mathbf{\Pi}$, where $\mathcal{O}_K$ is from Lemma 4. Since $\mathcal{V}_K\overline{\mathcal{O}}_K \in \mathcal{R}_K^{N\times K}$ [23], we have $\mathbf{E}_o\mathcal{V}_K\overline{\mathcal{O}}_K \in \mathcal{R}_K^{n\times K}$. Consequently, by H3 and Lemma 4, with probability at least $1 - 2/N$,

$$\begin{aligned}
&\sqrt{\mathbb{K}^*(\mathbf{U}_{o,K})} = \min_{\overline{\mathbf{U}}\in\mathcal{R}_K^{n\times K}} ||\mathbf{U}_{o,K} - \overline{\mathbf{U}}||_{\text{F}} \\
&\le ||\mathbf{U}_{o,K} - \mathbf{E}_o\mathcal{V}_K\overline{\mathcal{O}}_K||_{\text{F}} \le ||\mathbf{E}_o||_2||\mathbf{U}_K - \mathcal{V}_K\overline{\mathcal{O}}_K||_{\text{F}} \\
&= ||\mathbf{U}_K - \mathcal{V}_K\overline{\mathcal{O}}_K||_{\text{F}} \le \frac{35\sqrt{K^3\log N}}{\sqrt{p(N-K)}}.
\end{aligned}$$

Combining the upper-bound of $\sqrt{\mathbb{K}^*(\mathbf{Q}_K)}$ with Lemma 2 as well as H1, H2, we conclude that when the null hypothesis holds, with probability at least $1 - 2/N - 5/M - \delta_{\text{gap}}$,

$$\begin{aligned}
\mathbb{K}^*(\widehat{\mathbf{Q}}_K) \le &\left[\sqrt{\frac{N}{n}}\frac{35\sqrt{K^3\log N}}{\sqrt{p(N-K)}} + \sqrt{K}||\mathbf{I} - \mathbf{R}_K||_2\right. \\
&\left. + 2\sqrt{K}\left(3\gamma\eta + \frac{c_1\,\text{tr}(\overline{\mathbf{C}}_o)\sqrt{2\log M/M} + \sigma^2}{\rho_{\text{gap}}}\right)\right]^2. \quad (11)
\end{aligned}$$

The next case is to consider the ground truth as $\mathcal{T}_{gnd} = \mathcal{T}_1$. Define $\widehat{\mathbf{X}}$ associated with $\widehat{\mathcal{C}} \in \arg\min_{\mathcal{C}} \mathbb{K}(\widehat{\mathbf{Q}}_K, \mathcal{C})$. Similar to the previous case, we have

$$\begin{aligned}
&\sqrt{\mathbb{K}^*(\mathbf{Q}_K)} \le ||(\mathbf{I} - \widehat{\mathbf{X}}\widehat{\mathbf{X}}^\top)\mathbf{Q}_K||_{\text{F}} \\
&= ||(\mathbf{I} - \widehat{\mathbf{X}}\widehat{\mathbf{X}}^\top)\mathbf{Q}_K\mathbf{Q}_K^\top||_{\text{F}} \\
&\le ||(\mathbf{I} - \widehat{\mathbf{X}}\widehat{\mathbf{X}}^\top)\widehat{\mathbf{Q}}_K\widehat{\mathbf{Q}}_K^\top||_{\text{F}} + ||\mathbf{Q}_K\mathbf{Q}_K^\top - \widehat{\mathbf{Q}}_K\widehat{\mathbf{Q}}_K||_{\text{F}} \\
&= \sqrt{\mathbb{K}^*(\widehat{\mathbf{Q}}_K)} + ||\mathbf{Q}_K\mathbf{Q}_K^\top - \widehat{\mathbf{Q}}_K\widehat{\mathbf{Q}}_K||_{\text{F}},
\end{aligned}$$

which implies $\sqrt{\mathbb{K}^*(\widehat{\mathbf{Q}}_K)} \ge \sqrt{\mathbb{K}^*(\mathbf{Q}_K)} - ||\mathbf{Q}_K\mathbf{Q}_K^\top - \widehat{\mathbf{Q}}_K\widehat{\mathbf{Q}}_K||_{\text{F}}$. By the same technique, we obtain $\sqrt{\mathbb{K}^*(\mathbf{Q}_K)} \ge c_0^{-1}\sqrt{\mathbb{K}^*(\mathbf{U}_{o,K})} - \sqrt{K}||\mathbf{I} - \mathbf{R}_K||_2$.

Our remaining task is to lower bound $\sqrt{\mathbb{K}^*(\mathbf{U}_{o,K})}$ using H4. Let $\mathbf{U}_{r,s} = [\mathbf{u}_r, ..., \mathbf{u}_s]$ consist of column vectors from $\mathbf{U}$, with $r \le s$. Also, let π be a permutation function on $\{1, ..., N\}$, satisfying $|h_i| = |h(\lambda_{\pi(i)})|$. We can see that the set $\mathcal{P} := \{i : 1 \le i \le K, K+1 \le \pi(i) \le N\}$ is non-empty under $\mathcal{T}_{gnd} = \mathcal{T}_1$. Then, for any $r \le s$ such that $[r, s] \in \mathcal{P}$, by Lemma 1,

$$\begin{aligned}
\mathbb{K}^*(\mathbf{U}_{o,K}) &\ge \mathbb{K}^*(\mathbf{U}_K) - |\mathbb{K}^*(\mathbf{U}_{o,K}) - \mathbb{K}^*(\mathbf{U}_K)| \\
&\ge \mathbb{K}^*(\mathbf{U}_K) - \frac{2450K^3\log N}{p(N-K)} \\
&\ge \mathbb{K}^*(\mathbf{U}_{r,s}) - \frac{2450K^3\log N}{p(N-K)}.
\end{aligned}$$

By H4, we have $\mathbb{K}^*(\mathbf{U}_{r,s}) \ge c_{\text{SBM}}$. Together with H1, H2, with probability at least $1 - 4/N - 5/M - \delta_{\text{gap}} - \delta_{\text{SBM}}$, the following lower bound holds

$$\begin{aligned}
\mathbb{K}^*(\widehat{\mathbf{Q}}_K) \ge &\left[\sqrt{\frac{N}{n}}\sqrt{c_{\text{SBM}} - \frac{2450K^3\log N}{p(N-K)}} - \sqrt{K}||\mathbf{I} - \mathbf{R}_K||_2\right. \\
&\left. - 2\sqrt{K}\left(3\gamma\eta + \frac{c_1\,\text{tr}(\overline{\mathbf{C}}_o)\sqrt{2\log M/M} + \sigma^2}{\rho_{\text{gap}}}\right)\right]^2. \quad (12)
\end{aligned}$$

Finally, we can conclude the proof by noting that $\widehat{\mathcal{T}} = \mathcal{T}_{\text{gnd}}$ holds when δ upper bounds the right-hand side of (11) and also lower bounds the right-hand side of (12).

Technical Lemmas

Lemma 1. *Under H3. Let $\mathbf{U}_K$ denote the columns of the first K eigenvectors of $\mathbf{L}_{\text{norm}}$, and $\mathbf{U}_{o,K} = \mathbf{E}_o\mathbf{U}_K$. With probability at least $1 - 4/N$,*

$$|\mathbb{K}^*(\mathbf{U}_{o,K}) - \mathbb{K}^*(\mathbf{U}_K)| \le \frac{2450K^3\log N}{p(N-K)}.$$

Proof. By the triangular inequality,

$$|\mathbb{K}^*(\mathbf{U}_{o,K}) - \mathbb{K}^*(\mathbf{U}_K)| \le \mathbb{K}^*(\mathbf{U}_{o,K}) + \mathbb{K}^*(\mathbf{U}_K).$$

Applying Lemma 3 yields $\mathbb{K}^*(\mathbf{U}_K) \le \frac{35^2K^3\log N}{pN(N-K)}$ with probability at least $1 - 2/N$.

We now derive an upper-bound for $\mathbb{K}^*(\mathbf{U}_{o,K})$. As $\mathcal{V}_K\mathcal{O}_K \in \mathcal{R}_K^{N\times K}$, we have $\mathbf{E}_o\mathcal{V}_K\mathcal{O}_K \in \mathcal{R}_K^{n\times K}$. Then, with probability at least $1-2/N$,

$$\mathbb{K}^*(\mathbf{U}_{o,K}) \leq ||\mathbf{U}_{o,K} - \mathbf{E}_o\mathcal{V}_K\mathcal{O}_K||_{\mathrm{F}}^2 \leq ||\mathbf{E}_o||_2^2||\mathbf{U}_K - \mathcal{V}_K\mathcal{O}_K||_{\mathrm{F}}^2$$
$$= ||\mathbf{U}_K - \mathcal{V}_K\mathcal{O}_K||_{\mathrm{F}}^2 = \frac{35^2K^3\log N}{p(N-K)}.$$

This concludes the proof. □

Lemma 2. *Under H1, H2, the following inequality holds with probability at least* $1-5/M$

$$||\mathbf{Q}_K\mathbf{Q}_K^\top - \widehat{\mathbf{Q}}_K\widehat{\mathbf{Q}}_K||_{\mathrm{F}} \leq$$
$$2\sqrt{K}\left(3\gamma\eta + \frac{c_1\,\mathrm{tr}(\overline{\mathbf{C}}_o)\sqrt{2\log M/M}+\sigma^2}{\rho_{\mathrm{gap}}}\right).$$

where c_1 *is a constant independent of* N, M *[24].*

Proof. By [12, Proposition 1], we have a deterministic upper-bound:

$$||\mathbf{Q}_K\mathbf{Q}_K^\top - \widehat{\mathbf{Q}}_K\widehat{\mathbf{Q}}_K||_{\mathrm{F}} \leq$$
$$\sqrt{2K}\left(\sqrt{2}\gamma(2||\mathbf{U}_K||_2 + ||\mathbf{U}_{N-K}||_2)\eta_K + \frac{||\widehat{\mathbf{C}}_o - \overline{\mathbf{C}}_o||_2}{\rho_{\mathrm{gap}}}\right)$$

where we further have $||\mathbf{U}_K||_2 \leq 1$ and $||\mathbf{U}_{N-K}||_2 \leq 1$ due to their orthogonality. In addition, applying [24, Theorem 2.1] on $\{\mathbf{y}_{o,m}\}_{m=1}^M$ yields the following inequality: with probability at least $1-5/M$,

$$||\widehat{\mathbf{C}}_o - \overline{\mathbf{C}}_o||_2 \leq 2c_1\,\mathrm{tr}(\overline{\mathbf{C}}_o)\sqrt{\frac{\log M}{M}} + \sigma^2.$$

This concludes the proof. □

The last two auxiliary lemmas are borrowed from [16], which have been inspired by [23], [24]:

Lemma 3 ([16, Proposition 2])**.** *Under H3. For* $\mathbf{V}_K$ *consisting of* K *bottom eigenvectors of* $\mathbf{L}_{\mathrm{norm}}$*, with probability at least* $1-2/N$*,*

$$\mathbb{K}^*(\mathbf{V}_K) \leq \frac{35^2K^3\log N}{p(N-K)}.$$

Lemma 4 ([16, Lemma 2])**.** *Under H3. Let* $\mathbf{V}_K, \mathcal{V}_K$ *denote the columns of the first* K *eigenvectors of* $\mathbf{L}_{\mathrm{norm}}, \mathcal{L}_{\mathrm{norm}}$*. With probability at least* $1-2/N$*, there exists an orthogonal matrix* $\mathcal{O}_K \in \mathbb{R}^{K\times K}$ *such that*

$$||\mathbf{V}_K - \mathcal{V}_K\mathcal{O}_K||_{\mathrm{F}} \leq \frac{35\sqrt{K^3\log N}}{\sqrt{p(N-K)}}.$$