Raksha Ramakrishna, Hoi-To Wai, and Anna Scaglione

# A User Guide to Low-Pass Graph Signal Processing and Its Applications

*Tools and applications*



*Digital Object Identifier 10.1109/MSP.2020.3014590*
*Date of current version: 28 October 2020*

The notion of graph filters can be used to define generative models for graph data. In fact, the data obtained from many examples of network dynamics may be viewed as the output of a graph filter. With this interpretation, classical signal processing tools, such as frequency analysis, have been successfully applied with analogous interpretation to graph data, generating new insights for data science. What follows is a user guide on a specific class of graph data, where the generating graph filters are low pass; i.e., the filter attenuates contents in the higher graph frequencies while retaining contents in the lower frequencies. Our choice is motivated by the prevalence of low-pass models in application domains such as social networks, financial markets, and power systems. We illustrate how to leverage properties of low-pass graph filters to learn the graph topology and identify its community structure; efficiently represent graph data through sampling; recover missing measurements; and denoise graph data. The low-pass property is also used as the baseline to detect anomalies.

## Introduction

A growing trend in signal processing and machine learning is to develop theories and models for analyzing data defined in irregular domains, such as graphs. Graphs often express relational ties, such as social, economic, and gene networks, for which several mathematical and statistical models relying on graphs have been proposed to explain trends in the data [1]. Another case is that of physical infrastructures (utility networks, including gas and water delivery, and transportation systems), where physical laws, in addition to connectivity, define the structure in signals.

For a period of time, the graphical interpretation was primarily used in statistics, with the aim of making inferences about graphical models. Meanwhile, the need for processing graph data has led to the emerging field of graph signal processing (GSP), which takes a deterministic and system theoretic approach to justify the properties of graph data and inspire the associated signal processing algorithms. A cornerstone of GSP is the formal definition of *graph filter*, which extends

the notions of the linear time-invariant (LTI) filtering of time series signals to processing data that are defined on a graph as well as graph signals.

In a similar vein as LTI filters in discrete-time signal processing, a graph filter can be classified as either low pass, bandpass, or high pass, depending on its graph frequency response. Among them, this article focuses on low-pass graph filters and the graph signals generated from them. These graph filters capture a smoothing operation applied to the input graph signals, which is a common property of processes observed in physical/social systems (see the "Models of Low-Pass Graph Signals" section). As a motivating example, in Figure 1, we illustrate a few real data sets from social networks, power systems, and the financial market and show the eigenvalue spectra of their sample covariance matrices. A salient feature observed is that these sample covariance matrices are low rank, thus displaying an important characteristic of low-pass filtered graph signals (discussed in the "User Guide to Low-Pass GSP" section).

Previous articles, such as [2] and [3], have provided a comprehensive introduction to modeling and processing graph and network data using GSP, favoring general abstractions over focusing on particular structures and concrete applications. This user guide takes a different approach, concentrating on low-pass graph filters and the corresponding low-pass graph signal outputs. Low-pass graph signals have specific properties that affect their structure and dictate how to approach, for example, sampling, denoising, and inference problems. They are worth focusing on because they are common in practice. In fact, resorting to existing underlying network dynamical models that justify different data sets, we show that low-pass graph processes are nearly ubiquitous in contexts where GSP is applicable.

## Basics of GSP

Many tools introduced in this user guide involve several fundamental concepts of GSP, including a formal definition of low-pass graph filters/signals. These ideas will be briefly reviewed in this section. For more details, the readers are referred to excellent prior overview articles, such as [2] and [3]. We denote vectors through boldface lowercase letters ($\boldsymbol{x}$) and use uppercase letters for matrices ($\mathbf{A}$). The operation $\mathtt{Diag}(\boldsymbol{x})$ creates a diagonal matrix with elements from vector $\boldsymbol{x}$.

We focus on a weighted undirected graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ with $n$ nodes such that $\mathcal{N} = \{1, \ldots, n\}$, and $\mathcal{E} \subseteq \mathcal{N} \times \mathcal{N}$ is the edge set. A graph signal is a function $x: \mathcal{N} \to \mathbb{R}$ that can be represented by an $n$-dimensional vector $\boldsymbol{x} = (x(i))_{i \in \mathcal{N}}$. A graph shift operator (GSO) is a matrix $\boldsymbol{S} \in \mathbb{R}^{n \times n}$ satisfying $[\boldsymbol{S}]_{ij} \neq 0$ if and only if $i = j$ or $(i, j) \in \mathcal{E}$. When multiplied by a graph signal $\boldsymbol{x}$, each entry of the shifted graph signal is a linear combination of the one-hop neighbors' values, therefore "shifting" the graph signal with respect to the graph topology. In this

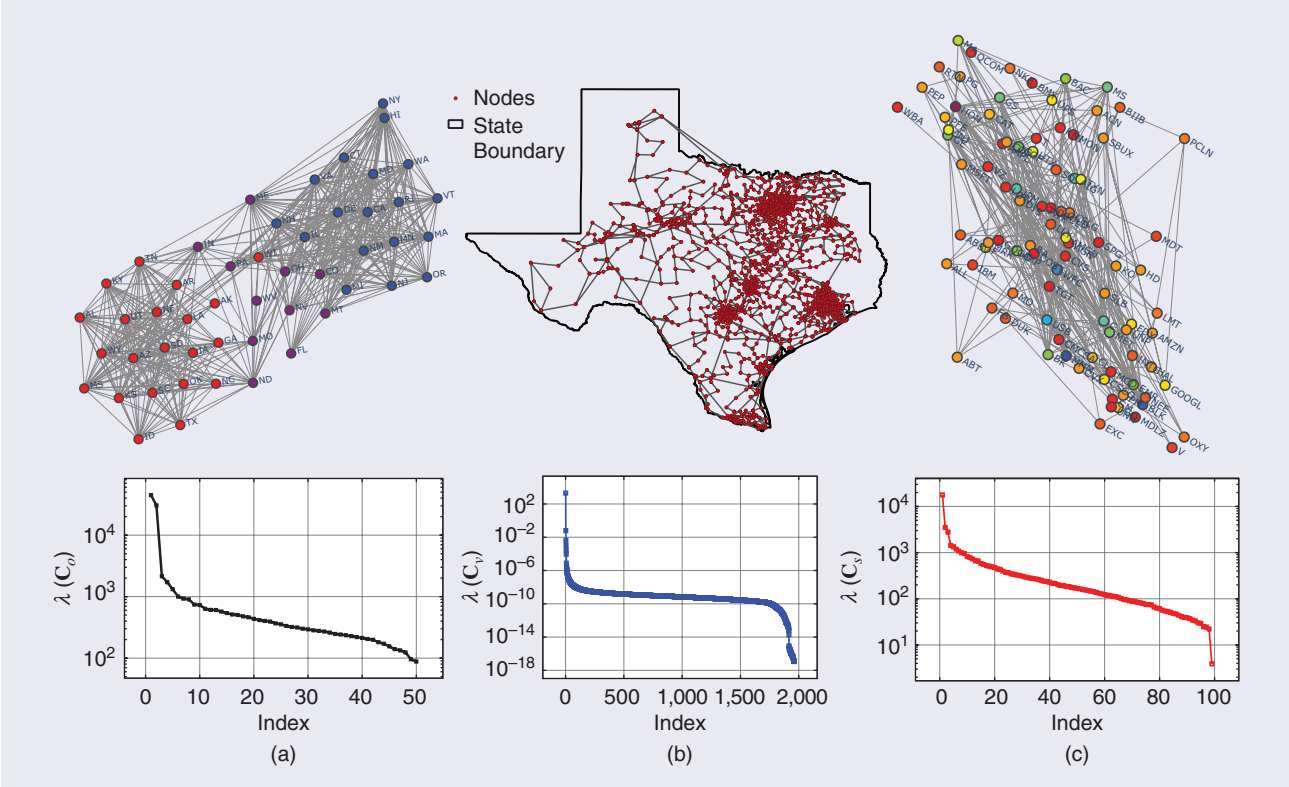


**FIGURE 1.** The eigenvalue spectra of (a) a social network, with opinion dynamics data from a roll call during the 110th U.S. Senate (sample covariance matrix $\mathbf{C}_o$); (b) a power network, including voltage data from a synthetic 2,000 bus power system test case known as ACTIVSg2000 (sample covariance matrix $\mathbf{C}_v$); and (c) a financial network, incorporating the daily returns of Standard & Poor's (S&P) 100 stocks from May 2018 to August 2019 (sample covariance matrix $\mathbf{C}_s$.). These data admit physical/social models that can be regarded as low-pass filtered graph signals. A salient feature of their low-pass nature is observed as the low-rank property of the sample covariance matrices.

article, we take the Laplacian matrix as the GSO. The Laplacian matrix is defined as $\boldsymbol{L} := \boldsymbol{D} - \boldsymbol{A}$, where $\boldsymbol{A}$ is the weighted symmetric adjacency matrix of $\mathcal{G}$ and $\mathbf{D} = \texttt{Diag}(\boldsymbol{A}\mathbf{1})$ is a diagonal matrix of the weighted degrees. It is also common to take the GSO as the normalized Laplacian matrix or the adjacency matrix [4].

Having defined the GSO, we discuss how to measure the smoothness of graph signals and analyze their content in the graph frequency domain. Recall that if a signal is smooth in time, the norm of its time derivative is small. For a graph signal $\boldsymbol{x}$, its graph derivative is defined as

$$[\nabla \boldsymbol{x}]_{ij} = \sqrt{A_{ij}}\,(x_i - x_j).$$

The squared Frobenius norm of the graph derivative, also known as the graph quadratic form [2], provides an idea of the smoothness of the graph signal $\boldsymbol{x}$:

$$\mathrm{S}_2(\boldsymbol{x}) := \frac{1}{2}\|\nabla \boldsymbol{x}\|_{\mathrm{F}}^2 = \boldsymbol{x}^\top \boldsymbol{L}\boldsymbol{x} = \sum_{i,j} A_{ij}(x_i - x_j)^2. \tag{1}$$

Observe that if $x_i \approx x_j$ for any neighboring nodes $i$, $j$, then $\mathrm{S}_2(\boldsymbol{x}) \approx 0$. Therefore, we say that a graph signal is smooth if $\mathrm{S}_2(\boldsymbol{x})/\|\boldsymbol{x}\|_2$ is small.

Let us take a closer look at the graph quadratic form $\mathrm{S}_2(\boldsymbol{x})$. We set the eigendecomposition of the Laplacian matrix as $\boldsymbol{L} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^\top$ and assume that it has eigenvalues of multiplicity one, ordered as $\boldsymbol{\Lambda} = \texttt{Diag}(\lambda_1, \ldots, \lambda_n)$ with $0 = \lambda_1 < \lambda_2 < \cdots < \lambda_n$, and that $\boldsymbol{U} = (\mathbf{u}_1\ \mathbf{u}_2\ \ldots\ \mathbf{u}_n)$, with $\mathbf{u}_i \in \mathbb{R}^n$, is the eigenvector for $\lambda_i$. Observe that, for any $\boldsymbol{x} \in \mathbb{R}^n$, it holds that $\mathrm{S}_2(\boldsymbol{x})/\|\boldsymbol{x}\|_2 \geq \mathrm{S}_2(\mathbf{u}_1)/\|\mathbf{u}_1\|_2 = \lambda_1$, and, for any $\boldsymbol{x}$ that is orthogonal to $\mathbf{u}_1$, it that holds $\mathrm{S}_2(\boldsymbol{x})/\|\boldsymbol{x}\|_2 \geq \mathrm{S}_2(\mathbf{u}_2)/\|\mathbf{u}_2\|_2 = \lambda_2$, and so on for the other eigenvectors. The observation indicates that the larger the eigenvalue, the more oscillatory the eigenvector is across the vertex set. In particular, the smallest eigenvalue $\lambda_1 = 0$ is associated with the flat, all-ones eigenvector $\mathbf{u}_1 = (1/\sqrt{n})\mathbf{1}$, as seen in Figure 2. The preceding motivates the definition of graph frequencies as the eigenvalues $\lambda_1, \ldots, \lambda_n$ and the description of the graph Fourier transform (GFT) basis as the set of eigenvectors $\boldsymbol{U}$ [4]. Therefore, the $i$th frequency component of $\boldsymbol{x}$ is defined as the inner product between $\mathbf{u}_i$ and $\boldsymbol{x}$,

$$\tilde{x}_i = \mathbf{u}_i^\top \boldsymbol{x},\ i = 1, \ldots, n, \tag{2}$$

and $\tilde{\boldsymbol{x}} = \boldsymbol{U}^\top \boldsymbol{x}$ is called the *GFT* of $\boldsymbol{x}$. The magnitude of the GFT vector $|\tilde{\boldsymbol{x}}|$ is the "spectrum" of the graph signal $\boldsymbol{x}$, where $|\tilde{x}_i|^2$ represents the signal power at the $\lambda_i$th frequency.

An important concept to modeling data with GSP is the graph filtering operation. To this end, a linear graph filter is described as the linear operator

$$\mathcal{H}(\boldsymbol{L}) = \sum_{p=0}^{P-1} h_p \boldsymbol{L}^p = \boldsymbol{U}\left(\sum_{p=0}^{P-1} h_p \boldsymbol{\Lambda}^p\right)\boldsymbol{U}^\top, \tag{3}$$

where $P$ is the filter order (it can be infinite) and $\{h_p\}_{p=0}^{P-1}$ are the filter's coefficients; as a convention, we use $\boldsymbol{L}^0 = \boldsymbol{I}$ and $\lambda_0^0 = 0^0 = 1$. From (3), one can see that the graph filter has an interpretation that is similar to an LTI filter in discrete-time signal processing, where the former replaces the time shifts by powers of the GSO. Meanwhile, the second expression in (3) defines the frequency response as the diagonal matrix $h(\boldsymbol{\Lambda}) := \Sigma_{p=0}^{P-1} h_p \boldsymbol{\Lambda}^p$. A graph signal $\boldsymbol{y}$ is said to be filtered by $\mathcal{H}(\boldsymbol{L})$ with the input excitation $\boldsymbol{x}$ when

$$\boldsymbol{y} = \mathcal{H}(\boldsymbol{L})\boldsymbol{x}. \tag{4}$$

To better appreciate the effects of the graph filter, note that the $i$th frequency component of $\boldsymbol{y}$ is

$$\tilde{y}_i = h(\lambda_i)\,\tilde{x}_i,\ i = 1, \ldots, n, \tag{5}$$

where $h(\lambda) := \Sigma_{p=0}^{P-1} h_p \lambda^p$ is the transfer function of the graph filter, or, equivalently, we have $\tilde{\boldsymbol{y}} = h(\boldsymbol{\lambda}) \odot \tilde{\boldsymbol{x}}$. It is similar to the convolution theorem in discrete-time signal processing.

### Low-pass graph filter

Inspired by (5), we define the ideal low-pass graph filter with a cutoff frequency $\lambda_k$ through setting the transfer function as $h(\lambda) = 1, \lambda \leq \lambda_k$ and zero otherwise [5]. Alternatively, one

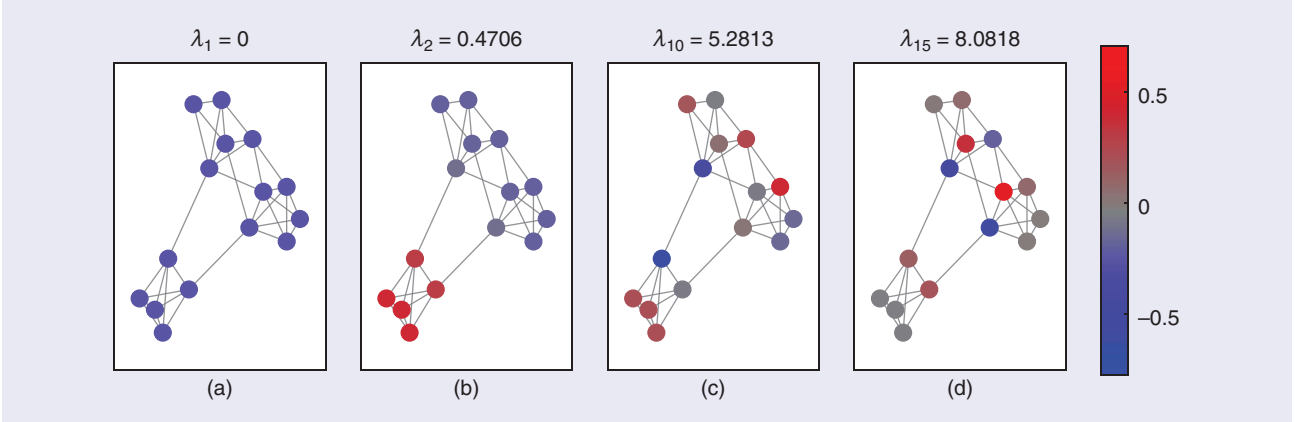


**FIGURE 2.** The GFT basis associated to the graph Laplacian of an undirected, unweighted graph with 15 nodes. As the eigenvalue increases, the eigenvectors tend to be more oscillatory. (a) $\mathbf{u}_1$. (b) $\mathbf{u}_2$. (c) $\mathbf{u}_{10}$. (d) $\mathbf{u}_{15}$.

can say that a graph filter is low pass if its frequency response is concentrated on the low graph frequencies. In this article, we adopt the following definition from [6].

Definition 1

For any $1 \le k \le n-1$, define the ratio

$$\eta_k := \frac{\max\{|h(\lambda_{k+1})|, \ldots, |h(\lambda_n)|\}}{\min\{|h(\lambda_1)|, \ldots, |h(\lambda_k)|\}}. \tag{6}$$

The graph filter $\mathcal{H}(\boldsymbol{L})$ is $k$ low pass if and only if the low-pass ratio $\eta_k$ satisfies $\eta_k \in [0,1)$.

The integer parameter $k$ characterizes the bandwidth, or the cutoff frequency of the low-pass filter is at $\lambda_k$. The ratio $\eta_k$ quantifies the "strength" of the low-pass graph filter. Upon passing a graph signal through $\mathcal{H}(\boldsymbol{L})$, the high-frequency components (greater than $\lambda_k$) are attenuated by a factor that is less than or equal to $\eta_k$. Using this definition, the ideal $k$-low-pass graph filter has the ratio $\eta_k = 0$, whose filter order has to be at least $P \ge n-k+1$ and whose transfer function $h(\lambda)$ has $\{\lambda_{k+1}, \ldots, \lambda_n\}$ as its roots. Finally, a $k$-low-pass graph signal refers to a graph signal that is the output of a $k$-low-pass filter subject to a "well-behaved" excitation (i.e., it does not possess strong high-pass components), which includes, but is not limited to, white noise.

### The impact of graph topologies

From Definition 1, one observes that the low-pass ratio $\eta_k$ of a graph filter depends on the filter's coefficients $\{h_p\}_{p=0}^{P-1}$ and the graph Laplacian matrix's spectrum $\lambda_1, \ldots, \lambda_n$. The condition $\lambda_k \ll \lambda_{k+1}$ facilitates the design of a $k$-low-pass graph filter with a favorable ratio of $\eta_k \ll 1$ and a low filter order $P$. As an example, the order-one graph filter $\mathcal{H}(\boldsymbol{L}) = \boldsymbol{I} - \lambda_n^{-1}\boldsymbol{L}$ is $k$ low pass, with the ratio $\eta_k = \lambda_n - \lambda_{k+1}/\lambda_n - \lambda_k = 1 - (\lambda_k - \lambda_{k+1}/\lambda_n - \lambda_k)$, where $\eta_k$ is small if $\lambda_k \ll \lambda_{k+1}$.

An example of graph topologies favoring the condition $\lambda_k \ll \lambda_{k+1}$ is the stochastic block model (SBM) [7] for describing random graphs with $k$ blocks/communities with nodes in $\mathcal{N}$ partitioned as $\mathcal{N}_1, \ldots, \mathcal{N}_k$. Consider a simplified SBM with $k$ equal-size blocks specified by a membership matrix $\boldsymbol{Z} \in \{0,1\}^{n \times k}$ such that $Z_{i\ell} = 1$ if and only if $i \in \mathcal{N}_\ell$ and a latent model $\boldsymbol{B} \in [0,1]^{k \times k}$, where $B_{j,\ell}$ is the probability of the edges between the nodes in block $j$ and $\ell$. We consider the homogeneous planted partition model (PPM) such that $\boldsymbol{B} = b\boldsymbol{1}\boldsymbol{1}^\top + a\boldsymbol{I}$, with $b, a > 0$. With the above specification, the adjacency matrix $\boldsymbol{A}$ is a symmetric binary matrix with independent entries satisfying $\mathbb{E}[\boldsymbol{A}] = \boldsymbol{Z}\boldsymbol{B}\boldsymbol{Z}^\top$. When the graph size grows to infinity ($n \to \infty$), the Laplacian matrix of an SBM–PPM graph converges almost surely to its expected value [7, Th. 2.1]:

$$\boldsymbol{L} \overset{a.s.}{\to} \mathbb{E}[\boldsymbol{L}] = \frac{n(a+kb)}{k}\boldsymbol{I} - \boldsymbol{Z}(b\boldsymbol{1}\boldsymbol{1}^\top + a\boldsymbol{I})\boldsymbol{Z}^\top. \tag{7}$$

From the preceding, it can be shown that $\lambda_{k+1} - \lambda_k = na/k \gg 1$ for $\mathbb{E}[\boldsymbol{L}]$, i.e., a favorable graph model for $k$-low-pass graph filters. Finally, the bottom-$k$ eigenvectors of the expected Laplacian associated with $\lambda_1, \ldots, \lambda_k$ can be collected into the matrix $\sqrt{(k/n)}\boldsymbol{Z}\boldsymbol{P}$, where $\boldsymbol{P}$ diagonalizes the matrix $\boldsymbol{B}$. In other words, the eigenvectors corresponding to the bottom-$k$ eigenvalues of $\boldsymbol{L}$ will reveal the block structure.

In contrast, Erdös-Rényi graphs have Laplacian matrices that do not generally satisfy $\lambda_k \ll \lambda_{k+1}$. In fact, asymptotically ($n \to \infty$), the empirical distribution of the eigenvalues of Laplacian matrices tends to the free convolution of the standard Gaussian distribution and Wigner's semicircular law [8]. Such a spectrum does not favor the design of a $k$-low-pass graph filter with $\eta_k \ll 1$, reflecting the fact that block structures and communities do not emerge in Erdös-Rényi graphs.

### Low-pass graph-temporal filter

When the excitation to a graph filter is time varying, and the topology is fixed, we consider a graph-temporal filter [9] with the impulse response

$$\mathcal{H}(\boldsymbol{L}, t) := \sum_{p=0}^{P-1} h_{p,t}\boldsymbol{L}^p \tag{8}$$

such that the graph filter's output is given by the time-domain convolution $\boldsymbol{y}_t = \Sigma_{s=0}^{t}\mathcal{H}(\boldsymbol{L}, t-s)\boldsymbol{x}_s$. The filter is causal, and $\boldsymbol{x}_s = 0$ for $s < 0$. We can apply a $z$ transform and the GFT to the graph signal process $\{\boldsymbol{x}_t\}_{t \ge 0}$ to obtain the $z$-GFT signal, $\widetilde{\boldsymbol{X}}(z)$, given by

$$\boldsymbol{X}(z) = \sum_{t=0}^{\infty} \boldsymbol{x}_t z^{-t}, \quad \tilde{\boldsymbol{X}}(z) = \boldsymbol{U}^\top \boldsymbol{X}(z), \tag{9}$$

which represents $\{\boldsymbol{x}_t\}_{t \ge 0}$ in the joint $z$ graph frequency domain. With that, we obtain the input-output relation $\tilde{\boldsymbol{Y}}(z) = \tilde{\boldsymbol{h}}(z) \odot \tilde{\boldsymbol{X}}(z)$ and the graph-temporal joint transfer function $\mathbb{H}(\lambda, z) := \Sigma_{t=0}^{\infty}\Sigma_{p=0}^{P-1}h_{p,t}\lambda^p z^{-t}$. A class of graph-temporal filters for modeling graph signal processes is the GF-autoregressive moving average (ARMA) ($q$, $r$) filter, whose input-output relation in the time domain and the $z$-GFT domain are described as, respectively,

$$\begin{aligned} &\boldsymbol{y}_t - \mathcal{A}_1(\boldsymbol{L})\boldsymbol{y}_{t-1} \cdots - \mathcal{A}_q(\boldsymbol{L})\boldsymbol{y}_{t-q} \\ &\quad = \mathcal{B}_0(\boldsymbol{L})\mathbf{x}_t + \cdots + \mathcal{B}_r(\boldsymbol{L})\boldsymbol{x}_{t-r}, \end{aligned} \tag{10}$$

$$\tilde{\boldsymbol{a}}(z) \odot \tilde{\boldsymbol{Y}}(z) = \tilde{\boldsymbol{b}}(z) \odot \tilde{\boldsymbol{X}}(z), \tag{11}$$

where $\tilde{\boldsymbol{a}}(z) = 1 - \Sigma_{s=1}^{q}\tilde{\boldsymbol{a}}_s z^{-s}$ and $\tilde{\boldsymbol{b}}(z) = \Sigma_{s=0}^{r}\tilde{\boldsymbol{b}}_s z^{-s}$ are the $z$ transform of the graph frequency responses of the graph filter taps $\{\mathcal{A}_s(\boldsymbol{L})\}_{s=1}^{q}, \{\mathcal{B}_s(\boldsymbol{L})\}_{s=0}^{r}$ for the GF-ARMA ($q$, $r$) filter. Note that the joint frequency response is given by $\mathbb{H}(\lambda_i, z) = [\tilde{\boldsymbol{b}}(z)]_i / [\tilde{\boldsymbol{a}}(z)]_i$, whose poles and zeros may vary depending on the graph frequencies $\lambda_1, \ldots, \lambda_n$. A relevant case is when $\mathcal{H}(\boldsymbol{L}, t)$ is a low-pass graph-temporal filter. Similar to Definition 1, we say that $\mathcal{H}(\boldsymbol{L}, t)$ is low pass with a cutoff frequency $(\lambda_k, \omega_0)$ and ratio $\eta_k$ if

$$\eta_k = \frac{\max\limits_{\lambda \in \{\lambda_{k+1}, \ldots, \lambda_n\}, \omega \in (\omega_0, 2\pi)} |\mathbb{H}(\lambda, e^{j\omega})|}{\min\limits_{\lambda \in \{\lambda_1, \ldots, \lambda_k\}, \omega \in [0, \omega_0]} |\mathbb{H}(\lambda, e^{j\omega})|} < 1. \tag{12}$$

Graph signals filtered by a low-pass graph-temporal filter are also commonly found in applications, as we will illustrate next.

## Models of low-pass graph signals

Before studying the GSP tools for low-pass graph signals, a natural question is: Where can one find such graph signals?

It turns out that many physical and social processes are naturally characterized by low-pass graph filters. In this section, we present various examples and show that their generation processes can be represented as outputs from low-pass graph filters.

## Diffusion model

The first case pertains to observations from a diffusion process, and its variants are broadly applicable in network science. As an example, we consider the heat diffusion model in [10]. In it, the relevant graph is a proximity graph where each node $i \in \mathcal{N}$ is a location (e.g., cities), and if locations $i, j$ are close to each other, then $(i, j) \in \mathcal{E}$. The graph is endowed with a symmetric weighted adjacency matrix encoding the distance between locations. The graph signal $\boldsymbol{y}_t \in \mathbb{R}^n$ encodes the temperature of $n$ locations at time $t$, and let $\boldsymbol{x}_0 \in \mathbb{R}^n$ be the initial heat distribution. The temperature of a location is diffused to its neighbors. Let $\sigma > 0$ be a constant, and we have

$$\boldsymbol{y}_t = e^{-t\sigma \boldsymbol{L}}\boldsymbol{x}_0 = (\boldsymbol{I} - t\sigma\boldsymbol{L} + \frac{(t\sigma)^2}{2}\boldsymbol{L}^2 - \cdots)\boldsymbol{x}_0, \tag{13}$$

where (13) is a discretization of the heat diffusion equation [12]. Since $\boldsymbol{L}\boldsymbol{1} = \boldsymbol{0}$, the matrix exponential $e^{-t\sigma\boldsymbol{L}} = \boldsymbol{I} - t\sigma\boldsymbol{L} + ((t\sigma)^2/2)\boldsymbol{L}^2 - \cdots$ is row stochastic. The temperature at time $t$ is thus a weighted average of neighboring locations' temperatures at $t = 0$; i.e., this is a diffusion dynamical process.

To understand (13) under the context of low-pass filtering, we observe that $\boldsymbol{y}_t$ is a filtered graph signal with the excitation $\boldsymbol{x}_0$ and the graph filter $\mathcal{H}(\boldsymbol{L}) = e^{-t\sigma\boldsymbol{L}}$. We verify that $\mathcal{H}(\boldsymbol{L})$ is $k$ low pass with Definition 1 for any $k < n$. Note that the low-pass ratio $\eta_k$ is

$$\frac{e^{-t\sigma\lambda_{k+1}}}{e^{-t\sigma\lambda_k}} = e^{-t\sigma(\lambda_{k+1}-\lambda_k)}.$$

Because $\lambda_{k+1} > \lambda_k$ and $t\sigma > 0$, we see that $\mathcal{H}(\boldsymbol{L})$ is a $k$-low-pass graph filter for any $k = 1, \ldots, n-1$.

We have assumed that $\boldsymbol{x}_0$ is an impulse excitation affecting the system only at the initial time. In practice, the excitation signal may not be an impulse, and the output graph signal $\boldsymbol{y}_t$ is expressed as the convolution $\boldsymbol{y}_t = \sum_{s=0}^{t} e^{-(t-s)\sigma\boldsymbol{L}}\boldsymbol{x}_s$. This corresponds to a low-pass graph-temporal filter with the joint transfer function $\mathbb{H}(\lambda, z) = (1 - e^{-\lambda\sigma}z^{-1})^{-1}$. Besides, the diffusion process is common in network science, as similar models arise in contagion processes and product adoption, to name a few.

## Opinion dynamics

This example pertains to opinion data mined from social networks with the influence of external excitation [6], [11]. The relevant graph $\mathcal{G}$ is the social network graph, where each node $i \in \mathcal{N}$ is an individual and $\mathcal{E}$ is the set of friendships. Similar to the previous case study, this graph is endowed with a symmetric weighted adjacency matrix $\boldsymbol{A}$, where the weights measure the trust among pairs of individuals. Let $\alpha \in (0, \lambda_n^{-1}), \beta \in (0, 1)$ be parameters of individuals' trust toward others and susceptibility to external influences, respectively. The evolution of opinions follows that of a combination of DeGroot's and Friedkin-Johnsen's models [14], which is a GF-AR(1) model [a special case of ARMA $(q, r), q = 1$ and $r = 0$]:

$$\boldsymbol{y}_{t+1} = (1 - \beta)(\boldsymbol{I} - \alpha\boldsymbol{L})\boldsymbol{y}_t + \beta\boldsymbol{x}_t, \tag{14}$$

where $\boldsymbol{y}_t \in \mathbb{R}^n$ is a graph signal of individuals' opinions at time $t$ and $\boldsymbol{x}_t \in \mathbb{R}^n$ is a graph signal of the external opinions perceived by the social network. Note that this also corresponds to a low-pass graph-temporal filter with the joint transfer function $\mathbb{H}(\lambda, z) = \beta[1 - (1 - \beta)(1 - \alpha\lambda)z^{-1}]^{-1}$.

To discuss the steady state of (14), let us assume that $\boldsymbol{x}_t \equiv \boldsymbol{x}$ for simplicity. Considering (14), we observe that $\boldsymbol{y}_{t+1}$ is a convex combination of $\boldsymbol{x}$ and the weighted average of the neighbors' opinions at time $t$ and that it is formed by taking a weighted average of neighboring signals in $\boldsymbol{y}_t$ using a diffusion operator $\boldsymbol{I} - \alpha\boldsymbol{L}$. The recursion is stable, leading to the steady-state (or equilibrium) opinions

$$\boldsymbol{y} = \lim_{t\to\infty}\boldsymbol{y}_t = (I + \tilde{\alpha}L)^{-1}\boldsymbol{x} = \mathcal{H}(\boldsymbol{L})\boldsymbol{x}, \tag{15}$$

where we have defined $\tilde{\alpha} = \beta(1 - \alpha)/\alpha > 0$ and $\boldsymbol{y}$ is a filtered graph signal excited by $\boldsymbol{x}$.

The graph filter in (13) is given by $\mathcal{H}(\boldsymbol{L}) = (\boldsymbol{I} + \tilde{\alpha}\boldsymbol{L})^{-1}$. To verify that it is a $k$-low-pass graph filter with Definition 1, we note that, for any $k < n$, the low-pass ratio $\eta_k$ is

$$\frac{1 + \tilde{\alpha}\lambda_k}{1 + \tilde{\alpha}\lambda_{k+1}} = 1 - \tilde{\alpha}\frac{\lambda_{k+1} - \lambda_k}{1 + \tilde{\alpha}\lambda_k}.$$

Again, we observe that, as $\lambda_{k+1} > \lambda_k$, the preceding graph filter is $k$ low pass for any $k = 1, \ldots, n-1$. However, we remark that this low-pass ratio may be undesirable with $\eta_k \approx 1$ when $\tilde{\alpha} \ll 1$. Interestingly, a generative model similar to (14) is found in equilibrium problems, such as quadratic games [13].

Two remarks are in order. First, social networks are typically directed, and this suggests using a nonsymmetric shift operator, as opposed to the symmetric Laplacian matrix, which we used for the simplicity of exposition. Second, many alternative models for social network interactions are nonlinear, and linear GSP is insufficient in those contexts.

## Finance data

Financial systems, such as stock markets and hedge funds, periodically produce return reports about their business performance. A collection of these reports can be studied as graph signals, where the relevant graph $\mathcal{G}$ consists of nodes $\mathcal{N}$ that are financial institutions and edges $\mathcal{E}$ that are business ties between them. The issue of whether business results are correlated according to business ties has been studied [14]. Moreover, the returns are affected by a number of common factors [15]. Inspired by [14] and [15, Ch. 12.2], let $\beta \in (0, 1)$ be the strength of external influences; a reasonable model for the transient dynamics of the graph signal $\boldsymbol{y}_t$ of business performance measures is also a GF-AR(1):

$$\boldsymbol{y}_{t+1} = (1-\beta)\mathcal{H}(\boldsymbol{L})\boldsymbol{y}_t + \beta\boldsymbol{B}\boldsymbol{x}, \tag{16}$$

where $\mathcal{H}(\boldsymbol{L})$ is an unknown but low-pass graph filter, $\boldsymbol{B} \in \mathbb{R}^{n\times r}$ represents the factor model affecting financial institutions, and $\boldsymbol{x} \in \mathbb{R}^r$ is the excitation strength. The equilibrium of (14) is

$$\boldsymbol{y} = \lim_{t\to\infty}\boldsymbol{y}_t = (\frac{1}{\beta}\boldsymbol{I} - \frac{\bar{\beta}}{\beta}\mathcal{H}(\boldsymbol{L}))^{-1}\boldsymbol{B}\boldsymbol{x} \equiv \tilde{\mathcal{H}}(\boldsymbol{L})\boldsymbol{B}\boldsymbol{x},$$

where $\bar{\beta} = 1-\beta$. We see that $\boldsymbol{B}\boldsymbol{x}$ is the excitation signal and that the equilibrium $\boldsymbol{y}$ is the filter output. Suppose that $\mathcal{H}(\boldsymbol{L})$ is a $k$-low-pass graph filter with the frequency response satisfying $h(\lambda) \geq 0$; then, for $\tilde{\mathcal{H}}(\boldsymbol{L})$, we can evaluate the low-pass ratio $\eta_k$ as

$$1 - \frac{\bar{\beta}\{\min\limits_{\ell=1,\ldots,k} h(\lambda_\ell) - \max\limits_{\ell=k+1,\ldots,n} h(\lambda_\ell)\}}{1-\bar{\beta}\max\limits_{\ell=k+1,\ldots,n} h(\lambda_\ell)}.$$

Because $\min_{\ell=1,\ldots,k} h(\lambda_\ell) - \max_{\ell=k+1,\ldots,n} h(\lambda_\ell) > 0$ since $\mathcal{H}(\boldsymbol{L})$ is a $k$-low-pass graph filter itself, we observe that $\tilde{\mathcal{H}}(\boldsymbol{L})$ is again $k$ low pass, according to Definition 1.

For $\boldsymbol{y}$ to be a $k$-low-pass graph signal, one also has to assume that $\boldsymbol{B}\boldsymbol{x}$ is not high pass (i.e., not orthogonal to a low-pass one). This is a mild assumption, as the latent factor affecting financial institutions is either independent of the network or aligned with the communities. Above all, we remark that (16) is an idealized model, where determining the exact representation is an open problem in economics; see [14] and [15].

### Power systems

In the case of power systems, the relevant graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ is the electrical transmission line network. The node (as well as the bus) set includes generator buses, $\mathcal{N}_g = \{1,\ldots,|\mathcal{N}_g|\}$, and nongenerator/load buses, $\mathcal{N}_\ell = \mathcal{N}\backslash\mathcal{N}_g$. The edge set $\mathcal{E}$ refers to the transmission lines connecting the buses. The branch admittance matrix, $\boldsymbol{Y}$, models the effect of transmission lines and is a complex symmetric matrix associated with $\mathcal{G}$, where $[\boldsymbol{Y}]_{ij}$ is the complex admittance of the branch between nodes $i$ and $j$, provided that $(i,j) \in \mathcal{E}$. The graph signals we consider are the complex voltage phasors, denoted as $\boldsymbol{v}_t \in \mathbb{C}^n$, when measured at time $t$. They can be obtained using PMUs [16] installed on each bus $i \in \mathcal{N}$. The GSO, in this case, is a diagonally perturbed branch admittance matrix,

$$\boldsymbol{S}_{\text{grid}} := \boldsymbol{Y} + \texttt{Diag}([\boldsymbol{y}_g^\top, \boldsymbol{y}_\ell^\top(0)]), \tag{17}$$

where $\boldsymbol{y}_g \in \mathbb{C}^{|\mathcal{N}_g|}$ is the generator admittance and $\boldsymbol{y}_\ell(0) \in \mathbb{C}^{|\mathcal{N}_\ell|}$ is the load admittance at $t = 0$.

Note that $\boldsymbol{S}_{\text{grid}}$ is a GSO on the grid graph, as $[\boldsymbol{S}_{\text{grid}}]_{ij} = 0$ if $(i,j) \notin \mathcal{E}$. The complex symmetric matrix $\boldsymbol{S}_{\text{grid}}$ can be decomposed as $\boldsymbol{S}_{\text{grid}} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^\top$, where $\mathbf{U}$ is a complex orthogonal matrix satisfying $\mathbf{U}^\top\mathbf{U} = \boldsymbol{I}$ and $\boldsymbol{\Lambda}$ is a diagonal matrix with diagonal elements $\{\lambda_1,\ldots,\lambda_n\}$ sorted as $0 < |\lambda_1| \leq \cdots \leq |\lambda_n|$; see [17] for modeling details. Let $\boldsymbol{i}_t^g \in \mathbb{C}^n$ be the outgoing current at each node at time $t$, given by $\boldsymbol{i}_t^g := [\boldsymbol{y}_g^\top \odot \exp(\boldsymbol{x}_t^\top), \boldsymbol{0}]^\top$, where elements in $\exp(\boldsymbol{x}_t) \in \mathbb{C}^{|\mathcal{N}_g|}$ are the internal voltage phasors at the generator buses. Applying Kirchoff's current law in a quasi-steady state, the voltage phasors $\boldsymbol{v}_t \in \mathbb{C}^n$ are

$$\boldsymbol{v}_t = \boldsymbol{S}_{\text{grid}}^{-1}\boldsymbol{i}_t^g + \boldsymbol{w}_t = \mathcal{H}(\boldsymbol{S}_{\text{grid}})\boldsymbol{i}_t^g + \boldsymbol{w}_t, \tag{18}$$

where $\boldsymbol{w}_t \in \mathbb{C}^n$ captures the slow time-varying nature of the load and other modeling approximations. In other words, $\boldsymbol{v}_t$ is a graph signal obtained by the graph filter $\mathcal{H}(\boldsymbol{S}_{\text{grid}}) = \boldsymbol{S}_{\text{grid}}^{-1}$ and the excitation signal $\boldsymbol{i}_t^g \in \mathbb{C}^n$. In particular, we observe that $\mathcal{H}(\boldsymbol{S}_{\text{grid}}) = \boldsymbol{S}_{\text{grid}}^{-1}$ is a low-pass graph filter. Consider any $k \leq n$; the low-pass ratio $\eta_k$ is

$$|\lambda_k| / |\lambda_{k+1}|.$$

Because power grids tend to be organized as communities to serve different areas that have high population densities, the system admittance matrix $\boldsymbol{Y}$ is block diagonal and sparse. In particular, with $k$ communities in the grid graph, these facts indicate that the graph filter is $k$ low pass, satisfying $\eta_k \ll 1$.

The excitation graph signal $\boldsymbol{i}_t^g$ itself has a low-rank structure, as $[\boldsymbol{i}_t^g]_i = 0$ at $i \notin \mathcal{N}_g$. The temporal dynamics of $\boldsymbol{x}_t$ can be described as an AR(2) graph filter [9] using a reduced generator-only shift operator $\boldsymbol{S}_{\text{red}} \in \mathbb{C}^{|\mathcal{N}_g|}$ with the graph-temporal transfer function $\mathbb{H}(\lambda_{\text{red}}, z)$:

$$\boldsymbol{x}_t = \sum_{s=0}^{t} \mathcal{H}(\boldsymbol{S}_{\text{red}}, t-s)\boldsymbol{p}_s,$$
$$\mathbb{H}(\lambda_{\text{red}}, z) = \sigma_p^2(1 - \sum_{p=0}^{1} a_{p,1}\lambda_{\text{red}}^p z^{-1} - \sum_{p=0}^{1} a_{p,2}\lambda_{\text{red}}^p z^{-2})^{-1}, \tag{19}$$

where $\boldsymbol{p}_s$ is the stochastic power input to the system. The graph-temporal filter is low pass in the time domain. The overall system in (18) has approximately the properties of a low-pass graph temporal filter, according to the definition in (12).

## User guide to low-pass GSP

If we observe a set of low-pass graph signals, what can we learn from these signals? Can we find efficient representations for them? Can we exploit this structure to denoise the signal and detect anomalies? To answer these questions, we begin by studying two salient features of low-pass graph signals, namely the low-rank covariance matrix and smoothness as measured by the graph quadratic form. Then, we illustrate how these features can enable low-pass GSP to sample graph signals (and therefore compress them) to infer the graph topology and detect anomalous activities. Furthermore, when the graph topology admits a clustered structure, we highlight how these clusters emerge in the low-pass graph signals and provide insights on the optimal sampling patterns.

We now consider a set of $m$ low-pass graph signals that can be modeled as outcomes of independent and identically distributed random experiments, given as

$$\boldsymbol{y}_\ell = \mathcal{H}(\boldsymbol{L})\boldsymbol{x}_\ell + \boldsymbol{w}_\ell, \quad \ell = 1,\ldots,m \tag{20}$$

such that $\mathcal{H}(\boldsymbol{L})$ is a $k$-low-pass graph filter defined on the Laplacian matrix, $\boldsymbol{x}_\ell$ is the excitation signal, and $\boldsymbol{w}_\ell$ is additive noise. For simplicity, we do not consider the more general low-pass graph-temporal processes and assume that $\boldsymbol{x}_\ell, \boldsymbol{w}_\ell$ are zero-mean white noise with $\mathbb{E}[\boldsymbol{x}_\ell \boldsymbol{x}_\ell^\top] = \boldsymbol{I}, \mathbb{E}[\boldsymbol{w}_\ell \boldsymbol{w}_\ell^\top] = \sigma^2 \boldsymbol{I}$. We remark that the following observations still hold for the general setting when $\mathbb{E}[\boldsymbol{x}_\ell \boldsymbol{x}_\ell^\top]$ is not white or even diagonal. The latter relaxation is important for the applications listed in the previous section. For instance, in opinion dynamics, the excitation signals may represent external opinions that do not affect the social network uniformly; e.g., they are news articles written in a foreign language.

### *Low-rank covariance matrix*

From (20), it is straightforward to show that $\{\boldsymbol{y}_\ell\}_{\ell=1}^m$ is zero mean with the covariance matrix

$$\boldsymbol{C}_y = \boldsymbol{U} h(\boldsymbol{\Lambda})^2 \boldsymbol{U}^\top + \sigma^2 \boldsymbol{I}. \tag{21}$$

Recall that $\mathcal{H}(\boldsymbol{L})$ is a $k$-low-pass graph filter if $\eta_k \ll 1$, as defined in (6); the energy of $h(\boldsymbol{\Lambda})$ will be concentrated in the top-$k$ diagonal elements. Therefore, when the noise variance is small ($\sigma^2 \approx 0$), the low-pass graph signals lie approximately in $\text{span}(\boldsymbol{U}_k)$, a $k$-dimensional subspace of $\mathbb{R}^n$.

## Sampling graph signals

Since $k$-low-pass graph signals lie approximately in $\text{span}(\boldsymbol{U}_k)$, it is possible to map the graph signals almost losslessly onto $k$-dimensional vectors. While the $k$-dimensional representation can be obtained by projecting on the space spanned by $\boldsymbol{U}_k$, it is not necessary to do so. An alternative to generate this $k$-dimensional representation is by decimating the graph signals. To describe the setup, we let $\mathcal{N}_s = \{s_1, \ldots, s_{n_s}\} \subset \mathcal{N}$ be a sampling set with cardinality $n_s = |\mathcal{N}_s|$. A sampled version of $\boldsymbol{y}_\ell$ is constructed as (we shall omit the subscript $\ell$ for brevity).

1) Select a subset $\mathcal{N}_s \subset \mathcal{N}$;

2) Set $\boldsymbol{y}_{\text{samp}} = \boldsymbol{\Phi}\boldsymbol{y}$ where $\Phi_{q,j} = \begin{cases} 1, & \text{if } j = s_q, \\ 0, & \text{otherwise}, \end{cases}$ (22)

where $\boldsymbol{\Phi} \in \mathbb{R}^{n_s \times n}$ is a fat sampling matrix. To recover $\boldsymbol{y}$, we interpolate $\boldsymbol{y}_{\text{samp}}$ using a matching linear transformation [18], giving $\hat{\boldsymbol{y}} = \boldsymbol{\Psi}\boldsymbol{y}_{\text{samp}}$, with $\boldsymbol{\Psi} \in \mathbb{R}^{n \times n_s}$ to be designed later. Clearly, when $n_s < n$, it is not possible to exactly recover an arbitrary graph signal.

We see that ensuring an exact recovery requires some conditions on the sampling set and the graph signal. Exactly recovering $\boldsymbol{y}$ from its sampled version $\boldsymbol{y}_{\text{samp}}$ would require the sampled graph signal to be in the range space of sampling matrix $\boldsymbol{\Phi}$. We let $\bar{\boldsymbol{y}} = \boldsymbol{U}_k \boldsymbol{U}_k^\top \boldsymbol{y}$ be the projection of $\boldsymbol{y}$ onto the (low-frequency) subspace spanned by $\boldsymbol{U}_k$ and $\bar{\boldsymbol{w}} = \boldsymbol{y} - \bar{\boldsymbol{y}}$ be the projection error. The projected graph signal $\bar{\boldsymbol{y}}$ is a $k$-bandlimited (in fact, $k$-low-pass) graph signal. From [18, Th. 1], a sufficient condition for exact recovery is that if

$$\text{rank}(\boldsymbol{\Phi}\boldsymbol{U}_k) = k, \tag{23}$$

then there exists an interpolation matrix $\boldsymbol{\Psi} \in \mathbb{R}^{n \times n_s}$ such that $\boldsymbol{\Psi}\boldsymbol{\Phi}\bar{\boldsymbol{y}} = \bar{\boldsymbol{y}}$. In fact, it is possible to recover any $k$-bandlimited graph signal from its sampled version. We have

$$\hat{\boldsymbol{y}} = \boldsymbol{\Psi}\boldsymbol{y}_{\text{samp}} = \bar{\boldsymbol{y}} + \boldsymbol{\Psi}\boldsymbol{\Phi}\bar{\boldsymbol{w}} = \boldsymbol{y} + (\boldsymbol{\Psi}\boldsymbol{\Phi} - \boldsymbol{I})\bar{\boldsymbol{w}}. \tag{24}$$

Because the $k$-low-pass graph signals lie approximately in $\text{span}(\boldsymbol{U}_k)$, we have $\bar{\boldsymbol{w}} \approx 0$, provided that $\eta_k \ll 1$. Under condition (23), the sampling and interpolation procedure results in a small interpolation error.

Clearly, a necessary condition to satisfy (23) is $n_s \geq k$; i.e., we require at least the same number of samples as the bandwidth of the low-pass graph filter, which produces the graph signal $\boldsymbol{y}$. Beyond the necessary stipulation, obtaining a sufficient condition for (23) can be difficult since it is not obvious to derive conditions on the sampling set. The design of the sampling set has been the focus of the work in [18]–[20], which proposed to find $\mathcal{N}_s$ via a greedy method or the graph spectral proxies. The preceding statements are valid for any graph signal that has sparse frequency support. In the case of low-pass graph signals, we can obtain insights into what type of sampling patterns are compatible with (23). Consider the special case of SBM–PPM graphs with $k$ blocks discussed previously. Note that, as $n \to \infty$, we have $\boldsymbol{U}_k = \sqrt{k/n}\,\boldsymbol{Z}\boldsymbol{P}$ for this model, where $\boldsymbol{Z}$ is the block membership matrix. Hence, condition (23) can be easily verified if $\mathcal{N}_s$ contains at least one node from each of the $k$ blocks.

In Figure 3, we consider a power system application. We first plot the magnitude of the GFT of the voltage graph signal with respect to normalized graph frequencies in log scale. From the linear decay, it is evident that the magnitude of the GFT coefficients at lower frequencies is higher, confirming that the signal is low pass in nature. The sampling pattern (or optimal placement of sensors) for the graph signal reconstruction is also shown in the figure. The block structure in the GSO for the electric grid guides the sampling strategy. In this example, the smallest singular value of $\boldsymbol{\Phi}\boldsymbol{U}_k$ is maximized by using a greedy algorithm [20].

## Blind community detection

Another consequence of (21) relates to learning the block, or community, structure when the graph topology is unknown. When the graph topology is known, spectral clustering (SC) [7] is often the method of choice. The SC method computes the bottom-$k$ eigenvectors of the Laplacian as $\boldsymbol{U}_k$ and partitions the $n$ nodes via $k$-means:

$$\begin{aligned} F^\star &= \min_{\mathcal{N}_1, \ldots, \mathcal{N}_k} F(\mathcal{N}_1, \ldots, \mathcal{N}_k; \boldsymbol{U}_k) \\ &:= \left( \sum_{q=1}^{k} \sum_{i \in \mathcal{N}_q} \left\| \mathbf{u}_i^{\text{row}} - \frac{1}{|\mathcal{N}_q|} \sum_{j \in \mathcal{N}_q} \mathbf{u}_j^{\text{row}} \right\|_2^2 \right)^{1/2}, \end{aligned} \tag{25}$$

where $\mathbf{u}_i^{\text{row}} \in \mathbb{R}^k$ is the $i$th row vector of $\boldsymbol{U}_k$. In fact, this is an effective method for SBM–PPM graphs, where solving (25) reveals the true block membership [7].

Although only the graph signals $\{\boldsymbol{y}_\ell\}_{\ell=1}^m$ are observed, we know from (21) that the covariance matrix $\boldsymbol{C}_y$ will be

dominated by a rank-$k$ component spanned by $\boldsymbol{U}_k$ under the low-pass assumption. In fact, this is precisely what we need for community detection (CD), as hinted in (25). To this end, [6] proposed the blind CD procedure, which includes the following steps:

1) Find the top-$k$ eigenvectors $\hat{\boldsymbol{U}}_k \in \mathbb{R}^{n\times k}$ of sample covariance $\hat{\boldsymbol{C}}_y = 1/m\Sigma_{\ell=1}^{m}\boldsymbol{y}_\ell\boldsymbol{y}_\ell^\top$.
2) Apply $k$-means on the rows of $\hat{\boldsymbol{U}}_k$.

If we denote the detected communities as $\hat{\mathcal{N}}_1,\ldots,\mathcal{N}_k$, then

$$F(\hat{\mathcal{N}}_1,\ldots,\hat{\mathcal{N}}_k;\boldsymbol{U}_k) - F^\star = O(\eta_k + \sigma + m^{-1/2}). \tag{26}$$

In other words, blind CD approaches the performance of SC if the graph filter is $k$ low pass with $\eta_k \ll 1$, the observation noise $\sigma$ is small, and the number of samples $m$ is large. Notice that (26) is a general result that holds even if $\mathbb{E}[\boldsymbol{x}_\ell\boldsymbol{x}_\ell^\top]$ is nondiagonal or low rank. Moreover, blind CD is shown to outperform a two-step approach that learns the graph first and then applies SC; see [6].

In Figure 4, we illustrate results for the CD of opinion dynamics and financial data by using information about the U.S. Senate during the 115th Congress (2017–2019) and daily return data for stocks in the S&P 100 index from February 2013 to December 2016, respectively. The observed steady-state graph signals $\boldsymbol{y}_\ell$ for the opinion dynamics case are the aggregated vote records of each state, and we observe $m = 502$ voting rounds. In Figure 4(a), we apply blind CD to partition the states into $K = 2$ groups, where a close alignment between our results and the actual party memberships of this Congress is observed. The financial data set that we used contains $m = 975$ days of information for $n = 92$ stocks. In Figure 4(b), we apply blind CD to partition the stocks into $K = 10$ groups. Each of the detected communities includes companies of the same business type (for instance, Bank of America is with J.P. Morgan), showing the effectiveness of the method.

## Smooth graph signals

As discussed previously, the graph quadratic form quantifies the smoothness of a graph signal. If $S_2(\boldsymbol{y}) = \boldsymbol{y}^\top\boldsymbol{L}\boldsymbol{y} \ll \|\boldsymbol{y}\|_2$, the graph signal $\boldsymbol{y}$ is said to be smooth. For $k$-low-pass graph signals, we observe that

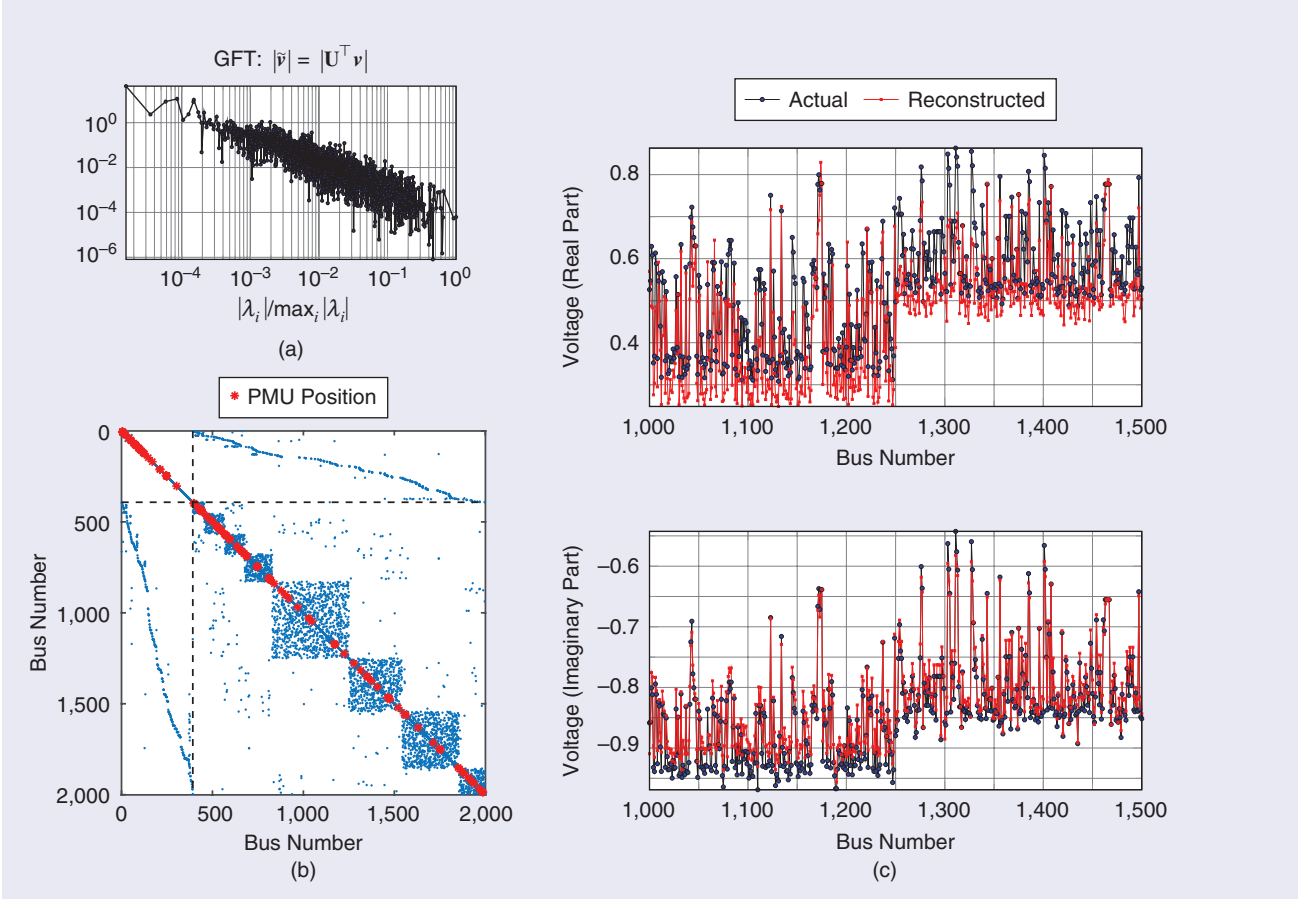


**FIGURE 3.** (a) The magnitude of the GFT of the voltage graph signal plotted with respect to normalized graph frequencies. (b) The sampling pattern overlaid on the support of GSO $\mathbf{S}_{\text{grid}}$ for placing synchrophasor sensors on the synthetic, 2,000-bus ACTIVSg power grid system by employing a greedy method to find $k = 100$ rows of $\boldsymbol{U}_k$ so that the smallest singular value of $\boldsymbol{\Phi}\boldsymbol{U}_k$ is maximized (the first entries correspond to generator buses). This case emulates the grid for the state of Texas, where there are eight areas, which matches the number of communities evident from the GSO. (c) The reconstructed voltage graph signal using optimally placed sensors at a subset of buses (1,000–1,500). PMU: phasor measurement unit.

$$\mathbb{E}[\mathrm{S}_2(\boldsymbol{y}_\ell)] \approx \sum_{i=1}^{k} \lambda_i |h(\lambda_i)|^2 + \sigma^2 \mathrm{Tr}(\boldsymbol{L}), \tag{27}$$

where $\mathcal{H}(\boldsymbol{L})$ is $k$ low pass, with $\eta_k \ll 1$, to derive the approximations. In the cases when $\lambda_i \approx 0, i = 1, \ldots, k$, such as large SBM–PPM graphs with parameters $(a, b)$ satisfying $b \ll 1, a \approx 1$, we expect the $k$-low-pass graph signal to be smooth; i.e., $\mathbb{E}[\mathrm{S}_2(\boldsymbol{y}_\ell)] \approx 0$.

## Graph topology learning

The smoothness property can be used to learn the graph topology by fitting a Laplacian matrix that best smoothens the graph signals. This is exemplified by the estimator

$$\begin{aligned} \min_{z_\ell, \ell=1,\ldots,m, \hat{\boldsymbol{L}}} \quad & \frac{1}{m} \sum_{\ell=1}^{m} \left\{ \frac{1}{\sigma^2} \| \boldsymbol{z}_\ell - \boldsymbol{y}_\ell \|_2^2 + \boldsymbol{z}_\ell^\top \hat{\boldsymbol{L}} \boldsymbol{z}_\ell \right\} \\ \text{s.t.} \quad & \mathrm{Tr}(\hat{\boldsymbol{L}}) = n,\ \hat{L}_{ji} = \hat{L}_{ij} \le 0,\ \forall\, i \ne j,\ \hat{\boldsymbol{L}}\mathbf{1} = \mathbf{0}, \end{aligned} \tag{28}$$

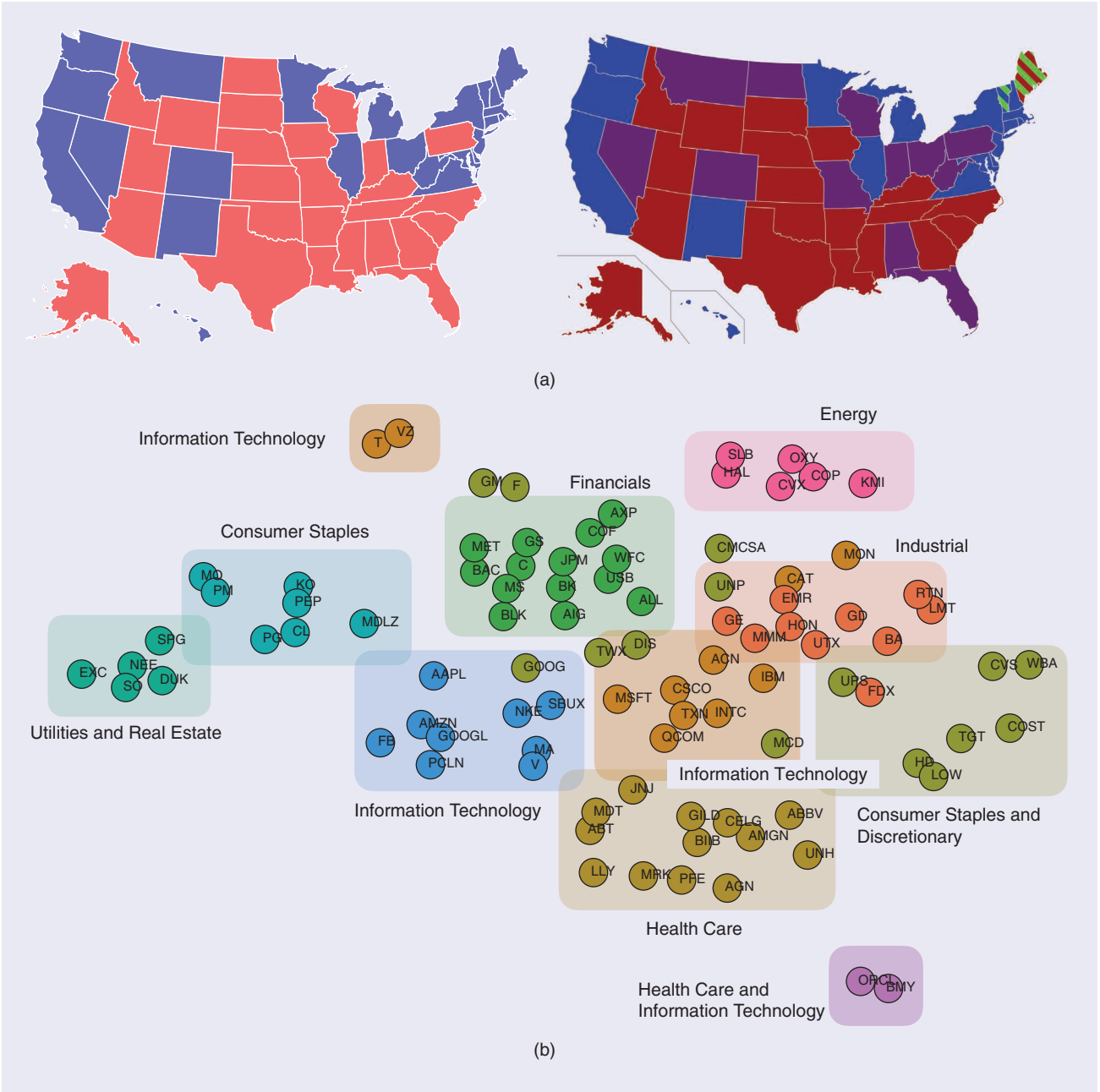


**FIGURE 4.** Communities detected from opinion dynamics and stock data. (a) U.S. Senate voting records, with (top) the inferred membership via blind CD and (bottom) the actual party membership of the senators (Source: Wikipedia, https://en.wikipedia.org/wiki/115th_United_States_Congress). Note that the purple color in the bottom indicates that the state has a Democratic senator and a Republican senator. (b) Daily returns of S&P 100 stocks (Source: Kaggle, https://www.kaggle.com/camnugent/sandp500). The colors on the nodes represent different inferred communities using blind CD, which are manually labeled according to business types.

where we have used the graph quadratic form, $\boldsymbol{z}_\ell^\top \hat{\boldsymbol{L}} \boldsymbol{z}_\ell$, to regulate the smoothness of $\boldsymbol{z}_\ell \approx \boldsymbol{y}_\ell$ with respect to the fitted $\hat{\boldsymbol{L}}$. Dong et al. [21] motivated (28) as a maximum a posteriori estimator for the Laplacian matrix, where $\boldsymbol{y}_\ell \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{L}^\dagger + \sigma^2 \boldsymbol{I})$ and $L^\dagger$ is the pseudo-inverse of the Laplacian matrix. This amounts to interpreting the data as outcomes of a Gaussian Markov random field, with a precision matrix chosen as the Laplacian, effectively connecting statistical graphical models to GSP models. Note that methods following insights similar to (28) can be found in [24] and [25].

For graph signals that are output from low-pass graph temporal filters, a smoothness property similar to (27) can also be exploited to interpolate missing data. Let $\boldsymbol{Y} \in \mathbb{R}^{n \times m}$ be a matrix whose columns are $\boldsymbol{y}_t, t = 1, 2, \ldots m$, where $\boldsymbol{y}_t$ is the graph signal at time $t$. Let $\boldsymbol{Y}^{\text{samp}}$ be the sampled version of $\boldsymbol{Y}$, where some values are missing at different node/time indices. The key for interpolating the data is to regularize via the graph quadratic form and the $\ell_2$ norm of the time derivative in addition to minimizing the $\ell_2$ misfit between the available samples $\boldsymbol{Y}^{\text{samp}}$ and the reconstructed samples at known locations, $\mathcal{M}(\boldsymbol{Y})$; i.e.,

$$\min_{\boldsymbol{Y} \in \mathbb{R}^{n \times m}} \| \mathcal{M}(\boldsymbol{Y}) - \boldsymbol{Y}^{\text{samp}} \|_F^2 + \gamma \left\{ \sum_{\ell=1}^{m} \boldsymbol{y}_\ell^\top \boldsymbol{L} \boldsymbol{y}_\ell + \sum_{\ell=2}^{m} \| \boldsymbol{y}_\ell - \boldsymbol{y}_{\ell-1} \|_2^2 \right\}.$$

See [24] and the references therein for a detailed discussion.

### Anomaly detection with low-pass GSP

Consider a model consistent with (20). The fact that the low-pass graph process is dominated by low graph frequency components can be considered the null hypothesis characterized by the low-pass properties, such as the low-rank covariance matrix and smoothness. On the other hand, many anomalies can be modeled as an additive sparse noise signal $\boldsymbol{w}_i$ or a high-frequency graph signal. Such noise signals arise in several scenarios, including changes in the graph connectivity and parameters, contingencies in infrastructures, the result of malicious activities in social networks, and the sudden fall in the market value of a financial entity. High-frequency noise signals are also produced by a perturbation that is inconsistent with the generative model. For instance, in infrastructure networks, this could be symptomatic of malfunctioning sensors and even a false data injection attack (FDIA) [25].

Such anomalies cause a surge in the high-frequency spectral components of a low-pass graph signal, a fact that can be leveraged in a manner similar to the classical array processing problem of detecting a source embedded in noise. Formally, the observed signal under null and alternative hypotheses is described as

$$\boldsymbol{y}_\ell = \begin{cases} \mathcal{H}(\boldsymbol{L}) \boldsymbol{x}_\ell, & \text{under } \mathcal{A}_0, \\ \mathcal{H}(\boldsymbol{L}) \boldsymbol{x}_\ell + \boldsymbol{w}_\ell & \text{under } \mathcal{A}_1, \end{cases}$$

where $\boldsymbol{w}_\ell$ is a high-frequency graph signal. Our task amounts to testing the hypothesis $\mathcal{A}_0, \mathcal{A}_1$ and estimating the locations of nonzeros in $\boldsymbol{w}_i$ when the latter is a sparse signal and under $\mathcal{A}_1$.

Intuitively, we can apply a high-pass graph filter to distinguish between $\mathcal{A}_0$ and $\mathcal{A}_1$. Let $\mathcal{H}_{\text{HPF}}(\boldsymbol{L})$ be an ideal high-pass graph filter with the frequency response $h_{\text{HPF}}(\lambda) = 1, \lambda \geq \lambda_{k+1}$ and zero otherwise. Consider the test statistics as $\Gamma_\ell = \| \mathcal{H}_{\text{HPF}}(\boldsymbol{L}) \boldsymbol{y}_\ell \|$. Under $\mathcal{A}_0$ and the $k$-low-pass assumption, we have $\Gamma_\ell \leq \| \mathcal{H}_{\text{HPF}}(\boldsymbol{L}) \mathcal{H}(\boldsymbol{L}) \|_2 \, \| x_\ell \| = \| h_{\text{HPF}}(\boldsymbol{\lambda}) \odot h(\boldsymbol{\lambda}) \|_\infty \, \| x_\ell \| = O(\eta_k)$, and thus the test statistics $\Gamma_\ell$ will be small. On the other hand, under $\mathcal{A}_1$, we obtain $\mathcal{H}_{\text{HPF}}(\boldsymbol{L}) \boldsymbol{y}_\ell \approx \boldsymbol{w}_\ell$ since the anomalies consist of high graph frequency components. Thus, the test statistics $\Gamma_\ell$ will be large. Imposing a threshold of $\delta = \Theta(\eta_k)$, we can consider the detector

$$\Gamma_\ell = \left\| \mathcal{H}_{\text{HPF}}(\boldsymbol{L}) \boldsymbol{y}_\ell \right\| \underset{\mathcal{A}_1}{\overset{\mathcal{A}_0}{\lessgtr}} \delta.$$

Furthermore, if $\mathcal{A}_1$ holds, these anomaly events can be located from the support of $\mathcal{H}_{\text{HPF}}(\boldsymbol{L}) \boldsymbol{y}_\ell$.

As a demonstration, Figure 5(a) presents the magnitude of the GFT of the voltage graph signal after filtering using an ideal high-pass graph filter, $\boldsymbol{U}^\top \mathcal{H}_{\text{HPF}}(\boldsymbol{L}) \boldsymbol{y}_\ell$. The voltage graph signal under the hypothesis of no anomaly is the output of a low-pass graph filter. When there is an FDIA, we observe an increase in the energy of the high-frequency components. For a simple implementation of high-pass graph filters, we may consider $\mathcal{H}_{\text{SD}}(\boldsymbol{L}) = \boldsymbol{L} = \boldsymbol{D} - \boldsymbol{A}$, whose frequency response is given by $h(\lambda) = \lambda$. When applied on a graph signal $\boldsymbol{y}_\ell$, we will observe the difference between $\boldsymbol{D}\boldsymbol{y}_\ell$ and $\boldsymbol{A}\boldsymbol{y}_\ell$; the latter is a one-hop averaged version of $\boldsymbol{y}_\ell$. We call this operation the *spatial difference*, which is similar to the method proposed in [26] for anomaly detection in social networks. See the illustration in Figure 5(b).

## Concluding remarks

In this user guide, we highlighted the key elements of low-pass GSP in several applications, such as graph parameter inference and graph signal sampling, while emphasizing the intuition from time series analysis. We also discussed several physical models where low-pass GSP can be effectively used. However, the tools available for low-pass GSP are ever expanding, and they aid the discovery of new physical models where low-pass GSP can be applied. Additionally, there are several open research directions, as discussed in the following.

### Directed graphs

Throughout this article, we have assumed that the observed data are supported on a graph topology that is undirected and that the shift operator (the Laplacian matrix) is symmetric. This is clearly not a truthful model for a lot of real systems, such as social and economic networks. The challenge of extending the existing GSP tools to directed graphs lies in defining the appropriate GFT basis. For instance, the properties of a circular shift matrix are what a directed shift operator should emulate.

Much of the prior research has focused on finding the appropriate GFT basis on directed graphs. The definition of *frequency* is again variational and the shift operator $\boldsymbol{S}$ of the corresponding graph does not have to be symmetric [4]. More

formally, the idea of smoothness is defined as $\|\boldsymbol{x} - \lambda_n^{-1}\boldsymbol{S}\boldsymbol{x}\|_2^2$, where $\lambda_n$ is the maximum eigenvalue of $\boldsymbol{S}$. This is the definition used in [4] for GFT on directed graphs, where the GSO is set as the adjacency matrix $\boldsymbol{A}$ and the GFT is defined as $\tilde{\boldsymbol{y}} = \boldsymbol{U}^{-1}\boldsymbol{y}$ such that $\mathbf{U}$ is obtained from the Jordan decomposition of the adjacency matrix $\boldsymbol{A} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^{-1}$. Although $\mathbf{U}$ is a basis, it is not orthogonal, so the Parseval's identity does not hold since $\|\tilde{\boldsymbol{y}}\|_2 \neq \|\boldsymbol{y}\|_2$. That is not surprising since the norm of $\tilde{\boldsymbol{y}}$ does not have the same physical interpretation of the power spectral density that applies to signals whose support is time. A potential fix is studied in [27], which searches for the GFT basis that minimizes the directed total variation; also see [28]. Unfortunately, the GFT basis does not admit a closed-form solution.

The tools discussed in this article, including sampling theory [18] and anomaly detection, may still work for low-pass graph signals on directed graphs, with minor adjustments. The challenge lies in the graph inference/learning methods since second-order statistics, such as correlations, are difficult to justify in directed graphs, where the notion of community is also ambiguous. A useful definition of *community* must first be studied before GSP tools can be applied for community inference in directed graphs.

### Low-pass graph signals in the edge space

An alternative form of graph signals concerns those that are defined on the edges. They can be defined as the function $f : \mathcal{E} \to \mathbb{R}$ and the equivalent vector $\boldsymbol{f} \in \mathbb{R}^{|\mathcal{E}|}$, which are useful for describing flows on graphs, such as traffic in transportation networks. As suggested in [29], the shift operator can be taken as the edge Laplacian $\mathbf{L}_e = \mathbf{B}^\top\mathbf{B}$, where $\mathbf{B} \in \mathbb{R}^{n\times|\mathcal{E}|}$ is the node-to-edge incidence matrix. The null space of the edge Laplacian $\mathbf{L}e$ corresponds to the cyclic flow vector, which is also the eigenvector for the lowest graph frequency $\lambda_1 = 0$. It is anticipated that a low-pass edge graph signal, whose energy is focused in the low graph frequencies, will consist mostly of cyclic flows within communities. An interesting direction is to develop a sampling theory for low-pass edge graph signals as well as the inference of the edge Laplacian matrix.

### Identifying low-pass graph signals

So far, we have relied on domain knowledge about the data models to help justify various graph data as low-pass graph signals. For graph signals taken from an unknown system, one has to be cautious before applying this low-pass GSP user guide. Even though non-low-pass graph processes are rarely found in a natural setting, there is a crucial need to design tools for identifying low-pass graph signals. With known GSO, this can be done by inspecting the GFT spectrum; with unknown GSO, the problem is related to the joint estimation of the graph process and the topology; readers are referred to [30] for recent works in this direction.

## Acknowledgments

We thank the anonymous reviewers and the guest editors for their useful feedback. This material is based on work supported, in part, by the U.S. Army Research Laboratory and the U.S. Army Research Office, under contract/grant W911NF2010153, and the National Science Foundation, under grant CCF-BSF: CIF: 1714672. Hoi-To Wai's work is supported by the Chinese University of Hong Kong, under direct grant 4055135.

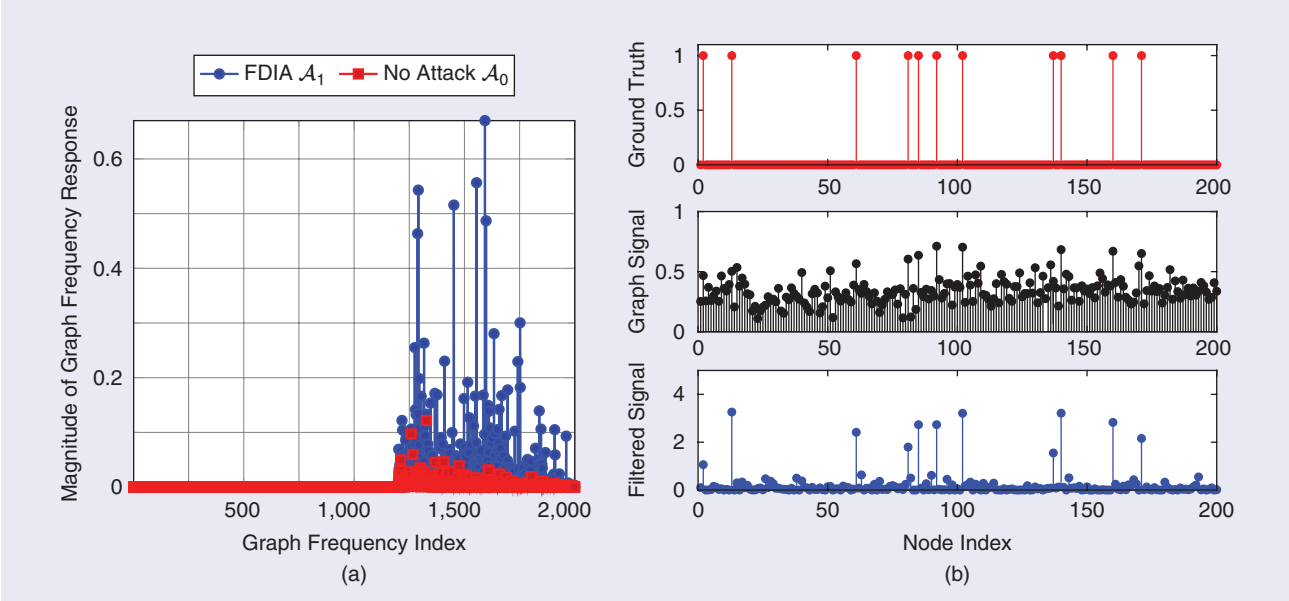


**FIGURE 5.** (a) The magnitude of the GFT of the output after ideal high-pass filtering, $\left|\boldsymbol{U}^T\mathcal{H}_{\text{HPF}}(L)\boldsymbol{y}_\ell\right|$, $k = 1, 200$, under the hypotheses of anomaly $\mathcal{A}_1$ and no anomaly $\mathcal{A}_0$. A particular example of FDIA on the voltage graph signal $\boldsymbol{y}_\ell$ from the ACTIVSg 2,000 case is shown. We observe a surge in the high-frequency components when there is an attack. (b) Spatial difference filtering of the graph signal under a diffusion model with abnormal activities on 11 nodes; the top graph shows the ground-truth locations of anomalies, while the middle and bottom graphs show the graph signal $\boldsymbol{y}_\ell$ and the filtered graph signal, $\mathcal{H}_{\text{SD}}(L)\boldsymbol{y}_\ell$, respectively.

## Authors

***Raksha Ramakrishna*** (raksha.ramakrishna@asu.edu) received her B.Eng. degree in electronics and communications engineering from Rashtreeya Vidyalaya College of Engineering, Bangalore, India, in 2014 and her M.S. degree in electrical engineering from Arizona State University (ASU), Tempe, in 2017. She is currently a Ph.D. student at ASU. Her research interests include statistical signal processing and data analytics for power systems.

***Hoi-To Wai*** (htwai@cuhk.edu.hk) received his B.Eng. degree and his M.Phil. degree in electronic engineering from the Chinese University of Hong Kong (CUHK) and his Ph.D. degree in electrical engineering from Arizona State University (ASU), Tempe. He is an assistant professor in the Department of Systems Engineering and Engineering Management, CUHK, and previously held research positions at ASU; the University of California, Davis; Telecom ParisTech; Ecole Polytechnique; and the Massachusetts Institute of Technology. His research interests include signal processing, machine learning, and distributed optimization. His dissertation received the Dean's Dissertation Award from ASU, and he received a Best Student Paper Award at the IEEE International Conference on Acoustics, Speech, and Signal Processing. He is a Member of IEEE.

***Anna Scaglione*** (ascaglio@asu.edu) received her M.Sc. degree in 1995 and her Ph.D. degree in 1999. She is a professor of electrical and computer engineering at Arizona State University, Tempe. Her research interests include statistical signal processing for information networks and intelligent cyberphysical infrastructures for secure and resilient energy delivery systems. She serves as deputy editor-in-chief (EIC) of *IEEE Transactions on Control* and is one of the IEEE Signal Processing Society (SPS) Distinguished Lecturers for 2019–2020. She was EIC of *IEEE Signal Processing Letters*, a member of the SPS Board of Governors from 2011 to 2014, and a member of the SPS Awards Board in 2016–2017. She received the 2000 IEEE Signal Processing Transactions Best Paper Award and the 2013 IEEE Donald G. Fink Prize Paper Award for the best review paper in IEEE publications. She is a Fellow of IEEE.

## References


[1] E. D. Kolaczyk and G. Csárdi, *Statistical Analysis of Network Data with R*, vol. 65. New York: Springer-Verlag, 2014.

[2] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. 83–98, 2013. doi: 10.1109/MSP.2012.2235192.

[3] A. Ortega, P. Frossard, J. Kovačević, J. M. Moura, and P. Vandergheynst, "Graph signal processing: Overview, challenges, and applications," *Proc. IEEE*, vol. 106, no. 5, pp. 808–828, 2018. doi: 10.1109/JPROC.2018.2820126.

[4] A. Sandryhaila and J. M. Moura, "Discrete signal processing on graphs," *IEEE Trans. Signal Process.*, vol. 61, no. 7, pp. 1644–1656, 2013. doi: 10.1109/TSP.2013.2238935.

[5] N. Tremblay, P. Gonçalves, and P. Borgnat, "Design of graph filters and filterbanks," in *Cooperative and Graph Signal Processing*, P. Djuric and C. Richard, Eds. London: Amsterdam, The Netherlands: Elsevier, 2018, pp. 299–324.

[6] H.-T. Wai, S. Segarra, A. E. Ozdaglar, A. Scaglione, and A. Jadbabaie, "Blind community detection from low-rank excitations of a graph filter," *IEEE Trans. Signal Process.*, vol. 68, pp. 436–451, Dec. 2020. doi: 10.1109/TSP.2019.2961296.

[7] K. Rohe, S. Chatterjee, and B. Yu, "Spectral clustering and the high-dimensional stochastic blockmodel," *Ann. Statist.*, vol. 39, no. 4, pp. 1878–1915, 2011. doi: 10.1214/11-AOS887.

[8] X. Ding and T. Jiang, "Spectral distributions of adjacency and Laplacian matrices of random graphs," *Ann. Appl. Probab.*, vol. 20, no. 6, pp. 2086–2117, 2010. doi: 10.1214/10-AAP677.

[9] E. Isufi, A. Loukas, A. Simonetto, and G. Leus, "Separable autoregressive moving average graph-temporal filters," in *Proc. IEEE 2016 24th European Signal Processing Conf. (EUSIPCO)*, pp. 200–204. doi: 10.1109/EUSIPCO.2016.7760238.

[10] D. Thanou, X. Dong, D. Kressner, and P. Frossard, "Learning heat diffusion graphs," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 3, no. 3, pp. 484–499, 2017. doi: 10.1109/TSIPN.2017.2731164.

[11] C. Ravazzi, R. Tempo, and F. Dabbene, "Learning influence structure in sparse social networks," *IEEE Trans. Control Netw. Syst.*, vol. 5, no. 4, pp. 1976–1986, 2017. doi: 10.1109/TCNS.2017.2781367.

[12] N. E. Friedkin, "A formal theory of reflected appraisals in the evolution of power," *Admin. Sci. Quart.*, vol. 56, no. 4, pp. 501–529, 2011. doi: 10.1177/0001839212441349.

[13] O. Candogan, K. Bimpikis, and A. Ozdaglar, "Optimal pricing in networks with externalities," *Oper. Res.*, vol. 60, no. 4, pp. 883–905, 2012. doi: 10.1287/opre.1120.1066.

[14] M. Billio, M. Getmansky, A. W. Lo, and L. Pelizzon, "Econometric measures of connectedness and systemic risk in the finance and insurance sectors," *J. Financ. Econ.*, vol. 104, no. 3, pp. 535–559, 2012. doi: 10.1016/j.jfineco.2011.12.010.

[15] R. N. Mantegna and H. E. Stanley, *Introduction to Econophysics: Correlations and Complexity in Finance*. Cambridge, U.K.: Cambridge Univ. Press, 1999.

[16] J. D. Glover, T. Overbye, and M. Sarma, *Power System Analysis and Design*, 6th ed., Boston: Cengage Learning, 2016.

[17] R. Ramakrishna and A. Scaglione, "On modeling voltage phasor measurements as graph signals," in *Proc. 2019 IEEE Data Science Workshop (DSW 2019)*, pp. 275–279. doi: 10.1109/DSW.2019.8755588.

[18] S. Chen, R. Varma, A. Sandryhaila, and J. Kovačević, "Discrete signal processing on graphs: Sampling theory," *IEEE Trans. Signal Process.*, vol. 63, no. 24, pp. 6510–6523, 2015. doi: 10.1109/TSP.2015.2469645.

[19] A. Anis, A. Gadde, and A. Ortega, "Efficient sampling set selection for bandlimited graph signals using graph spectral proxies," *IEEE Trans. Signal Process.*, vol. 64, no. 14, pp. 3775–3789, 2016. doi: 10.1109/TSP.2016.2546233.

[20] M. Tsitsvero, S. Barbarossa, and P. Di Lorenzo, "Signals on graphs: Uncertainty principle and sampling," *IEEE Trans. Signal Process.*, vol. 64, no. 18, pp. 4845–4860, 2016. doi: 10.1109/TSP.2016.2573748.

[21] X. Dong, D. Thanou, P. Frossard, and P. Vandergheynst, "Learning Laplacian matrix in smooth graph signal representations," *IEEE Trans. Signal Process.*, vol. 64, no. 23, pp. 6160–6173, 2016. doi: 10.1109/TSP.2016.2602809.

[22] V. Kalofolias, "How to learn a graph from smooth signals," in *Proc. Artificial Intelligence and Statistics*, 2016, pp. 920–929.

[23] B. Pasdeloup, V. Gripon, G. Mercier, D. Pastor, and M. G. Rabbat, "Characterization and inference of graph diffusion processes from observations of stationary signals," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 4, no. 3, pp. 481–496, 2017. doi: 10.1109/TSIPN.2017.2742940.

[24] F. Grassi, A. Loukas, N. Perraudin, and B. Ricaud, "A time-vertex signal processing framework: Scalable processing and meaningful representations for time-series on graphs," *IEEE Trans. Signal Process.*, vol. 66, no. 3, pp. 817–829, 2017. doi: 10.1109/TSP.2017.2775589.

[25] R. Ramakrishna and A. Scaglione, "Detection of false data injection attack using graph signal processing for the power grid," in *Proc. 2019 IEEE Global Conf. Signal and Information Processing (GlobalSIP)*, pp. 1–5. doi: 10.1109/GlobalSIP45357.2019.8969373.

[26] H.-T. Wai, A. E. Ozdaglar, and A. Scaglione, "Identifying susceptible agents in time varying opinion dynamics through compressive measurements," in *Proc. 2018 IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4114–4118. doi: 10.1109/ICASSP.2018.8462377.

[27] S. Sardellitti, S. Barbarossa, and P. Di Lorenzo, "On the graph Fourier transform for directed graphs," *IEEE J. Sel. Topics Signal Process.*, vol. 11, no. 6, pp. 796–811, 2017. doi: 10.1109/JSTSP.2017.2726979.

[28] R. Shafipour, A. Khodabakhsh, G. Mateos, and E. Nikolova, "A directed graph Fourier transform with spread frequency components," *IEEE Trans. Signal Process.*, vol. 67, no. 4, pp. 946–960, 2018. doi: 10.1109/TSP.2018.2886151.

[29] M. T. Schaub and S. Segarra, "Flow smoothing and denoising: Graph signal processing in the edge-space," in *Proc. 2018 IEEE Global Conf. Signal and Information Processing (GlobalSIP)*, pp. 735–739. doi: 10.1109/GlobalSIP.2018.8646701.

[30] V. N. Ioannidis, Y. Shen, and G. B. Giannakis, "Semi-blind inference of topologies and dynamical processes over dynamic graphs," *IEEE Trans. Signal Process.*, vol. 67, no. 9, pp. 2263–2274, 2019. doi: 10.1109/TSP.2019.2903025.


SP