



The Dual Graph Shift Operator: Identifying the Support of the Frequency Domain

Geert Leus¹ · Santiago Segarra² · Alejandro Ribeiro³ · Antonio G. Marques⁴

Received: 1 September 2020 / Revised: 15 February 2021 / Accepted: 14 April 2021 /

Published online: 13 May 2021

© The Author(s) 2021

Abstract

Contemporary data is often supported by an irregular structure, which can be conveniently captured by a graph. Accounting for this graph support is crucial to analyze the data, leading to an area known as graph signal processing (GSP). The two most important tools in GSP are the graph shift operator (GSO), which is a sparse matrix accounting for the topology of the graph, and the graph Fourier transform (GFT), which maps graph signals into a frequency domain spanned by a number of graph-related Fourier-like basis vectors. This alternative representation of a graph signal is denominated the graph frequency signal. Several attempts have been undertaken in order to interpret the support of this graph frequency signal, but they all resulted in a one-dimensional interpretation. However, if the support of the original signal is captured by a graph, why would the graph frequency signal have a simple one-dimensional support? Departing from existing work, we propose an irregular support for the graph frequency signal, which we coin dual graph. A dual GSO leads to a better interpretation of the graph frequency signal and its domain, helps to understand how the different

Work in this paper is supported by Spanish NSF Grants No. TEC2013-41604-R, TEC2016-75361-R, PID2019-105032GB-I00, and USA NSF Awards CCF-1217963 and CCF-2008555.

✉ Geert Leus
g.j.t.leus@tudelft.nl

Santiago Segarra
segarra@rice.edu

Alejandro Ribeiro
aribeiro@seas.upenn.edu

Antonio G. Marques
antonio.garcia.marques@urjc.es

¹ Department of Electrical Eng., Math. and Comp. Science, Delft University of Technology, Delft, The Netherlands

² Department of Electrical and Computer Eng., Rice University, Houston, USA

³ Department of Electrical and Systems Eng., University of Pennsylvania, Philadelphia, USA

⁴ Department of Signal Theory and Comms., King Juan Carlos University, Madrid, Spain

graph frequencies are related and clustered, enables the development of better graph filters and filter banks, and facilitates the generalization of classical SP results to the graph domain.

Keywords Graph signal processing · Dual graph shift operator · Frequency support · Graph Fourier transform

1 Introduction

Graph signal processing (GSP) has emerged as an effective solution to handle data with an irregular support. Its approach is to represent this support by a graph, view the data as signals defined on its nodes, and use algebraic and spectral properties of the graph to study these signals [22]. Such a data structure appears in many domains, including social networks, smart grids, sensor networks, and neuroscience. Instrumental to GSP are the notions of the graph shift operator (GSO), which is a matrix that accounts for the topology of the graph, and the graph Fourier transform (GFT), which transforms a graph signal to the so-called graph frequency domain leading to a graph frequency signal. These tools are the fundamental building blocks for the development of graph filters [8, 17], filter banks [14, 24], node-varying filters [20], edge-varying filters [5], graph sampling schemes [2, 11], statistical GSP [12], and other GSP techniques [23].

Motivated by the practical importance of the GFT, some efforts have been made to establish a *total ordering* of the graph frequencies [18, 22, 28], implicitly assuming a one-dimensional support for the graph frequency signal. Such an ordering translates into proximities between frequencies, which are critical for the definition of bandlimit-ness and smoothness as well as for the design of sampling and filter (bank) schemes. However, the basis vectors (or modes) associated with graph frequencies that are close in such one-dimensional domains are often dissimilar and focus on completely different parts of the graph [25], suggesting that a one-dimensional support is not descriptive enough to capture the similarity relationships between graph frequencies. To overcome that limitation, we propose a (not necessarily regular) support of a graph frequency signal by means of a graph, which we denominate as *dual graph*,¹ and its corresponding dual GSO. Note that the original graph is therefore often labeled as primal graph.

A *dual graph* helps in better describing the existing relations between graph frequencies. It tells us how similar certain graph frequencies are and this not only based on their value but also based on specific features of the related graph frequency modes. For instance, two graph frequencies that are close in value could be less similar (and hence less connected in a dual graph) than two graph frequencies that are further apart but have related graph frequency modes (e.g, localized in the same area of the graph). Therefore, clustering graph frequencies in a dual graph, as opposed to segmenting the one-dimensional frequency domain, can lead to different subbands. This will for instance have consequences when applying a subband-based compression scheme. Also, predicting a graph signal for a particular graph frequency from the signal on

¹ This is not related to the graph-theoretic notion of dual graph of a *planar* graph $\mathcal{G}$, which is a graph that has a vertex for each *face* of $\mathcal{G}$ [27].

its “neighboring” graph frequencies depends on how we define these neighboring frequencies, i.e., by merely considering the closeness of the graph frequencies or by using a dual graph.

We first proposed this idea in [9], and the current work can be considered an extension of that paper. Similar concepts were also presented in [21] as well as [3,4,10,16]. Specifically, [21] considered the same framework as [9]. However, although the dual graph presented in [21] specializes to classical temporal signal processing, it misses certain desirable properties that we discuss in the current paper. For instance re-ordering the nodes of the primal graph or re-ordering the graph frequency modes and related graph frequencies, both lead to a different dual graph whereas the primal graph basically did not change. The papers [3,4,10,16], on the other hand, take a different approach and try to embed the graph frequency modes in some lower-dimensional Euclidean space (two- or three-dimensional for instance) instead of the regular one-dimensional graph frequency domain determined only by the graph frequencies themselves. This embedding is based on computing some type of similarity between the graph frequency modes. Although this method reveals some interesting relations between the frequency modes, different similarity functions exist and it depends on the graph which one is more appropriate. Furthermore, the method is not reciprocal, i.e., we cannot go back to the original graph domain.

In this paper, after providing some basics and defining the problem statement (Sect. 2), we first derive the eigenvectors of a dual GSO (Sect. 3) and then establish how the eigenvalues of such a dual GSO can be computed (Sect. 4). For this, we consider an axiomatic approach as well as an optimization approach. According to this framework, we observe that a Laplacian primal GSO can never lead to a Laplacian dual GSO. However, we prove that this does not hold for an adjacency matrix without self-loops. In other words, if the primal GSO has zero diagonal entries, then it is always possible to find a dual GSO with zero diagonal entries. Simulation results (Sect. 5) support the claims made in this work. First, we show that graph frequencies that are close in value are not necessarily well-connected in a dual graph, and we illustrate this for several dual graph constructions. Next, the property that an adjacency shift without self-loops can be preserved going from the primal to the dual domain is corroborated and the dimension of the resulting class of potential dual adjacency shifts is illustrated for different types of graphs. Finally, we demonstrate that running a prediction graph filter on a dual graph outperforms predicting the graph signal adopting a traditional convolution in the one-dimensional graph frequency domain.

Notation Boldface capital letters are used for matrices, boldface lowercase letters for column vectors and calligraphic capital letters for sets. The entries of a matrix $\mathbf{X}$ are referred to either as X_{ij} or $[\mathbf{X}]_{ij}$. Similarly, the entries of a vector $\mathbf{x}$ are referred to either as x_i or $[\mathbf{x}]_i$. The notation T , H , and † respectively correspond to transpose, Hermitian, and complex conjugate. When applied to a vector $\mathbf{x}$, the operator $\text{diag}(\cdot)$ returns a square diagonal matrix whose diagonal elements are those in $\mathbf{x}$. When applied to a diagonal matrix, the operator $\text{diag}(\cdot)$ returns a column vector whose elements are those in the diagonal on the input matrix. Finally, $\text{vec}(\mathbf{X})$ returns a vector concatenating the columns of $\mathbf{X}$, $\mathbf{Z} = \mathbf{Y} \odot \mathbf{X}$ denotes the Khatri–Rao product ($\mathbf{Z}$, $\mathbf{Y}$, $\mathbf{X}$ have the same number of columns and the i th column of $\mathbf{Z}$ is the Kronecker product of the i th column

of $\mathbf{Y}$ with the i th column of $\mathbf{X}$), and we recall that if $\mathbf{Z} = \mathbf{Y} \text{diag}(\mathbf{w}) \mathbf{X}^T$, then it holds that $\text{vec}(\mathbf{Z}) = (\mathbf{X} \odot \mathbf{Y}) \mathbf{w}$.

2 Dual Graph

We start by reviewing fundamental concepts of GSP and then state formally the problem of identifying a dual GSO.

2.1 Fundamentals of GSP

Consider a (possibly directed) graph $\mathcal{G}$ of N nodes or vertices with node set $\mathcal{N} = \{n_1, \dots, n_N\}$ and edge set $\mathcal{E} = \{(n_i, n_j) \mid n_i \text{ is connected to } n_j\}$. The graph $\mathcal{G}$ is further characterized by the so-called GSO, which is an $N \times N$ matrix $\mathbf{S}$ whose entries $[\mathbf{S}]_{ij}$ for $i \neq j$ are zero whenever there is no edge from n_j to n_i . The diagonal entries of $\mathbf{S}$ can be selected freely and typical choices for the GSO include the Laplacian or adjacency matrices [18,22]. A graph signal defined on $\mathcal{G}$ can be conveniently represented by a vector $\mathbf{x} = [x_1, \dots, x_N]^T \in \mathbb{C}^N$, where x_i is the signal value associated with node n_i .

The GSO $\mathbf{S}$ —encoding the structure of the graph—is crucial to define the GFT and graph filters. The former transforms graph signals into a frequency domain, whereas the latter represents a class of local linear operators between graph signals. Assume for simplicity that the GSO $\mathbf{S}$ is normal, such that its eigenvalue decomposition (EVD) can always be written as $\mathbf{S} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^H$, where $\mathbf{V}$ is a unitary matrix that stacks the eigenvectors and $\mathbf{\Lambda}$ is a diagonal matrix that collects the eigenvalues. To simplify the exposition, we also assume that the eigenvalues of the shift are simple (non-repeated), such that the associated eigenspaces are unidimensional. The eigenvectors $\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_N]$ correspond to the graph frequency basis vectors whereas the eigenvalues $\boldsymbol{\lambda} = \text{diag}(\mathbf{\Lambda}) = [\lambda_1, \dots, \lambda_N]^T$ can be viewed as graph frequencies. With these conventions, the definitions of the GFT and graph filters are given next.

Definition 1 Given the GSO $\mathbf{S} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^H$, the GFT of the graph signal $\mathbf{x} \in \mathbb{C}^N$ is $\tilde{\mathbf{x}} = [\tilde{x}_1, \dots, \tilde{x}_N]^T := \mathbf{V}^H \mathbf{x}$.

Definition 2 Given the GSO $\mathbf{S} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^H$, a graph filter $\mathbf{H} \in \mathbb{C}^{N \times N}$ of degree L is a graph-signal operator of the form

$$\mathbf{H} = \mathbf{H}(\mathbf{h}, \mathbf{S}) := \sum_{l=0}^L h_l \mathbf{S}^l = \mathbf{V} \text{diag}(\tilde{\mathbf{h}}) \mathbf{V}^H, \quad (1)$$

where $\mathbf{h} := [h_0, \dots, h_L]$ and $\tilde{\mathbf{h}} := \text{diag}(\sum_{l=0}^L h_l \mathbf{\Lambda}^l)$.

Definition 1 implies that the inverse GFT (iGFT) is simply $\mathbf{x} = \mathbf{V} \tilde{\mathbf{x}}$. Vector $\mathbf{h}$ in Definition 2 collects the filter coefficients and $\tilde{\mathbf{h}} \in \mathbb{C}^N$ in (2) can be deemed as the frequency response of the filter. The particular case of the filter being $\mathbf{H} = \mathbf{S}$, so that $\tilde{\mathbf{h}} = \boldsymbol{\lambda}$, will be subject of further discussion in Sect. 3. Graph filters and the GFT have been shown useful for sampling, compression, filtering, windowing, and spectral estimation of graph signals [2,8,11,12,17,20,23].

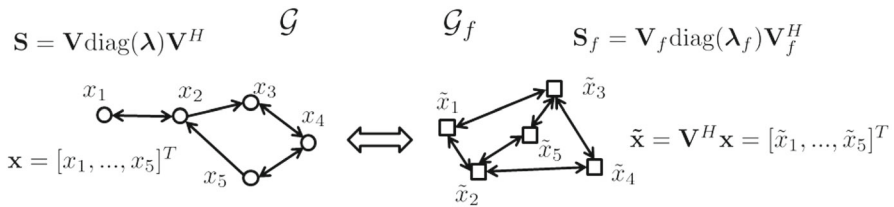


Fig. 1 The primal graph (left) represents the support of the vertex domain, while a dual graph (right) represents the support of the frequency domain

2.2 Support of the Frequency Domain

The underlying assumption in GSP is that to analyze and process the graph signal $\mathbf{x} \in \mathbb{C}^N$ one has to take into account its graph support $\mathcal{G}$ via the associated GSO $\mathbf{S}$. Moreover, according to Definition 1, the graph frequency signal $\tilde{\mathbf{x}} \in \mathbb{C}^N$ is an alternative representation of $\mathbf{x}$. Thus, a natural problem is the identification of a graph and GSO corresponding to $\tilde{\mathbf{x}}$. More precisely, we are interested in finding a dual graph $\mathcal{G}_f$ —represented via the corresponding dual GSO $\mathbf{S}_f$ —that characterizes the support of the frequency domain.

Let $\mathcal{N}_f = \{n_{f,1}, \dots, n_{f,N}\}$ denote the node set of the dual graph $\mathcal{G}_f$. Each element in $\mathcal{N}_f$ corresponds to a different frequency $(\lambda_i, \mathbf{v}_i)$, thus, the edge set $\mathcal{E}_f$ indicates pairwise relations between the different frequencies. We interpret $\tilde{\mathbf{x}}$ as a signal defined on this dual graph, where $\tilde{x}_i$ is associated with the node (frequency) $n_{f,i}$. As for the primal GSO, the EVD of the $N \times N$ matrix $\mathbf{S}_f$ associated with $\mathcal{G}_f$ will be instrumental to study $\tilde{\mathbf{x}}$. We start from the assumption that normality of $\mathbf{S}$ implies the normality of $\mathbf{S}_f$. Later on, we will see that this assumption is valid. Due to normality, we have then that $\mathbf{S}_f = \mathbf{V}_f \mathbf{\Lambda}_f \mathbf{V}_f^H$, and thus the dual graph has (dual) frequency basis vectors $\mathbf{V}_f = [\mathbf{v}_{f,1}, \dots, \mathbf{v}_{f,N}]$ and (dual) graph frequencies $\mathbf{\Lambda}_f = \text{diag}(\mathbf{\Lambda}_f) = [\lambda_{f,1}, \dots, \lambda_{f,N}]^T$ (cf. Fig. 1).

Problem statement Given the GSO $\mathbf{S} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^H$ find an appropriate dual GSO $\mathbf{S}_f = \mathbf{V}_f \mathbf{\Lambda}_f \mathbf{V}_f^H$.

To address this problem we postulate desirable properties that we want a dual GSO to satisfy. First, we start by identifying $\mathbf{V}_f$ (Sect. 3). We then proceed to determine $\mathbf{\Lambda}_f$ (Sect. 4), which is a more challenging problem.

3 Eigenvectors of a Dual Graph

We want the GFT $\mathbf{V}_f^H$ associated with a dual graph to map $\tilde{\mathbf{x}}$ back to the graph signal $\mathbf{x}$. Given that $\tilde{\mathbf{x}} = \mathbf{V}^H \mathbf{x}$ (cf. Definition 1), the ensuing result follows.

Property 1 Given the primal GSO $\mathbf{S} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^H$, the eigenvectors of a dual GSO $\mathbf{S}_f$ are $\mathbf{V}_f = \mathbf{V}^H$.

As announced in the previous section, since we have that $\mathbf{V}_f^{-1} = \mathbf{V}_f^H$, then the dual shift $\mathbf{S}_f$ is normal too. With $\mathbf{e}_i \in \mathbb{R}^N$ denoting the i th canonical basis vector

(all entries are zero except for the one corresponding to the i th node, which is one), then $\mathbf{v}_{f,i}$ can be written as $\mathbf{v}_{f,i} = \mathbf{V}^H \mathbf{e}_i = \tilde{\mathbf{e}}_i$, i.e., the GFT of the graph signal $\mathbf{e}_i$. Hence, the dual frequency vector $\mathbf{v}_{f,i}$ can be viewed as how node i expresses each of the primal graph frequencies, revealing that each frequency of the dual graph $\mathcal{G}_f$ is related to a particular node of the primal graph $\mathcal{G}$. Moreover, we can also interpret the dual eigenvalues from a primal perspective. To that end, note that λ_f is the frequency response of the dual filter $\tilde{\mathbf{H}} = \mathbf{S}_f$ (cf. discussion after Definition 2); thus, the i th entry of λ_f can be understood as how strongly the primal value at the i th node x_i is amplified when $\mathbf{S}_f$ is applied to $\tilde{\mathbf{x}}$.

One interesting implication of Property 1 is that the dual of a Laplacian shift $\mathbf{S} = \mathbf{V}\mathbf{A}\mathbf{V}^H$ is, in general, not a Laplacian. Laplacian matrices require the existence of a constant eigenvector. Hence, for $\mathbf{S}_f$ to be a Laplacian, one of the rows of $\mathbf{V}$ —corresponding to the columns of $\mathbf{V}_f$ —needs to be constant, which in general is not the case. Another implication of Property 1 is the duality of the filtering and windowing operations, as shown next.

Corollary 1 *Given the graph signal $\mathbf{x} \in \mathbb{C}^N$ and the window $\mathbf{w} \in \mathbb{C}^N$, define the windowed graph signal $\mathbf{x}_w \in \mathbb{C}^N$ as*

$$\mathbf{x}_w = \text{diag}(\mathbf{w})\mathbf{x}. \quad (2)$$

Then, recalling that $\tilde{\mathbf{x}} = \mathbf{V}^H \mathbf{x}$ and $\tilde{\mathbf{x}}_w = \mathbf{V}^H \mathbf{x}_w$, if $\mathbf{S}_f$ does not have repeated eigenvalues it holds that

$$\tilde{\mathbf{x}}_w = \mathbf{H}(\mathbf{h}_f, \mathbf{S}_f)\tilde{\mathbf{x}}, \quad \text{with } \mathbf{H}(\mathbf{h}_f, \mathbf{S}_f) = \sum_{l=0}^L h_{f,l}(\mathbf{S}_f)^l \quad (3)$$

for some $\mathbf{h}_f := [h_{f,0}, \dots, h_{f,L}]^T$ and $L \leq N - 1$.

Proof Substituting $\mathbf{x}_w = \text{diag}(\mathbf{w})\mathbf{x}$ and $\mathbf{x} = \mathbf{V}\tilde{\mathbf{x}}$ into the definition of $\tilde{\mathbf{x}}_w$ yields $\tilde{\mathbf{x}}_w = \mathbf{V}^H \text{diag}(\mathbf{w})\mathbf{V}\tilde{\mathbf{x}}$. This reveals that the mapping from $\tilde{\mathbf{x}}$ to $\tilde{\mathbf{x}}_w$ is given by the matrix $\tilde{\mathbf{H}} = \mathbf{V}^H \text{diag}(\mathbf{w})\mathbf{V}$. Since $\mathbf{V}^H$ is normal and unitary, $\mathbf{V}^H$ are the eigenvectors of $\tilde{\mathbf{H}}$ and $\mathbf{w}$ are its eigenvalues. Because $\mathbf{V}^H$ are also the eigenvectors of $\mathbf{S}_f$ (cf. Property 1), to show that $\tilde{\mathbf{H}}$ is a filter on $\mathbf{S}_f$ we only need to show that there exist coefficients $\mathbf{h}_f := [h_{f,0}, \dots, h_{f,N-1}]^T$ such that $\mathbf{w} = \text{diag}(\sum_{l=0}^{N-1} h_{f,l} \mathbf{A}_f^l)$ [cf. (1)]. Defining $\Psi_f \in \mathbb{C}^{N \times N}$ as $[\Psi_f]_{i,l} = (\lambda_{f,i})^{l-1}$, the equality can be written as $\mathbf{w} = \Psi_f \mathbf{h}_f$. Since Ψ_f is Vandermonde, if all the dual eigenvalues $\{\lambda_{f,i}\}_{i=1}^N$ are distinct, a vector $\mathbf{h}_f$ solving $\mathbf{w} = \Psi_f \mathbf{h}_f$ exists. $\square$

The proof holds regardless of the particular λ_f and only requires $\mathbf{S}_f$ to have non-repeated eigenvalues. The corollary states that multiplication in the vertex domain is equivalent to filtering in the dual domain—note that the GSO of the filter in (3) is $\mathbf{S}_f$. Clearly, when the entries of $\mathbf{w}$ are binary values, multiplying $\mathbf{x}$ by $\mathbf{w}$ acts as a windowing procedure preserving the values of $\mathbf{x}$ in the support $\mathbf{w}$, while discarding the information at the remaining nodes.

4 Eigenvalues of a Dual Graph

Given $\mathbf{S} = \mathbf{V}\text{diag}(\boldsymbol{\lambda})\mathbf{V}^H$ and using Property 1 to write the dual shift as $\mathbf{S}_f = \mathbf{V}^H\text{diag}(\boldsymbol{\lambda}_f)\mathbf{V}$, the last step to identify $\mathbf{S}_f$ is to obtain $\boldsymbol{\lambda}_f$. Two different (complementary) approaches to accomplish this are discussed next.

4.1 Axiomatic Approach

Our first approach is to postulate properties that we want the dual shift $\mathbf{S}_f$ to satisfy, and then translate these properties into requirements on the dual eigenvalues $\boldsymbol{\lambda}_f$. We denominate these properties as *axioms*, which we state next. In the following, $\mathbf{P}$ denotes an arbitrary permutation matrix.

(A1) Axiom of Duality The dual of the dual graph is equal to the original graph

$$(\mathbf{S}_f)_f = \mathbf{S}. \quad (4)$$

(A2) Axiom of Reordering The dual graph is robust to reordering the nodes in the primal graph

$$(\mathbf{PSP}^T)_f = \mathbf{S}_f. \quad (5)$$

(A3) Axiom of Permutation Permutations in the EVD of the primal shift lead to permutations in the dual graph

$$(\mathbf{VPdiag}(\mathbf{P}^T\boldsymbol{\lambda})\mathbf{P}^T\mathbf{V}^H)_f = \mathbf{P}^T(\mathbf{Vdiag}(\boldsymbol{\lambda})\mathbf{V}^H)_f\mathbf{P}. \quad (6)$$

Consistency with Property 1 is encoded in the Axiom of Duality (A1). More precisely, since the GFT of the dual shift transforms a frequency signal $\tilde{\mathbf{x}}$ back into the graph domain $\mathbf{x}$, we want the associated shift to be recovered as well. The Axiom of Reordering (A2) ensures that the frequency structure encoded in the dual shift is invariant to relabelings of the nodes in the primal shift. Specifically, the frequency coefficients of a given signal $\mathbf{x}$ with respect to $\mathbf{S}$ should be the same as those of $\mathbf{x}' = \mathbf{P}\mathbf{x}$ with respect to $\mathbf{S}' = \mathbf{PSP}^T$. Finally, since the nodes of the dual graph correspond to different frequencies, the Axiom of Permutation (A3) ensures that if we permute the eigenvectors (and corresponding eigenvalues) of $\mathbf{S}$, the nodes of the dual shift are permuted accordingly.

Axioms (A1)–(A3) impose conditions on the possible choices for the dual eigenvalues $\boldsymbol{\lambda}_f$. More precisely, let us define the function $\phi : \mathbb{C}^N \times \mathbb{C}^{N \times N} \rightarrow \mathbb{C}^N$, that computes the dual eigenvalues $\boldsymbol{\lambda}_f = \phi(\boldsymbol{\lambda}, \mathbf{V})$ as a function of the eigendecomposition of $\mathbf{S}$. In terms of ϕ , axiom (A1) requires that

$$\boldsymbol{\lambda} = \phi(\boldsymbol{\lambda}_f, \mathbf{V}_f) = \phi(\phi(\boldsymbol{\lambda}, \mathbf{V}), \mathbf{V}^H). \quad (7)$$

In order to translate (5) into a condition on ϕ , notice that if the labels of the nodes are permuted we have that $\mathbf{PSP}^T = \mathbf{P}\mathbf{V}\text{diag}(\boldsymbol{\lambda})\mathbf{V}^H\mathbf{P}^T$, so that $(\mathbf{PSP}^T)_f$ from Property 1 must be equal to $\mathbf{V}^H\mathbf{P}^T\text{diag}(\boldsymbol{\lambda}')\mathbf{P}\mathbf{V}$. Thus, for $(\mathbf{PSP}^T)_f$ to coincide with $\mathbf{S}_f$ we need

$\lambda' = \mathbf{P}\lambda_f$ which ultimately requires that

$$\phi(\lambda, \mathbf{P}\mathbf{V}) = \lambda' = \mathbf{P}\lambda_f = \mathbf{P}\phi(\lambda, \mathbf{V}). \quad (8)$$

Lastly, in order to find the requirement imposed by axiom (A3) on ϕ , we again leverage Property 1 to obtain $(\mathbf{V}\mathbf{P}\text{diag}(\mathbf{P}^T\lambda)\mathbf{P}^T\mathbf{V}^H)_f = \mathbf{P}^T\mathbf{V}^H\text{diag}(\lambda')\mathbf{V}\mathbf{P}$. It readily follows that to satisfy (6) we need $\lambda' = \lambda_f$, i.e.

$$\phi(\mathbf{P}^T\lambda, \mathbf{V}\mathbf{P}) = \lambda' = \lambda_f = \phi(\lambda, \mathbf{V}). \quad (9)$$

It is possible to find a function ϕ that simultaneously satisfies (7)–(9), as shown next.

Theorem 1 *The following class of functions satisfies (7)–(9), leading to a generating method for dual graphs that abides by axioms (A1)–(A3)*

$$\lambda_f = \phi(\lambda, \mathbf{V}) = \mathbf{D}_f^{-1}\mathbf{V}\mathbf{D}\lambda, \quad (10)$$

where $\mathbf{D} = \text{diag}(g(\mathbf{v}_1), \dots, g(\mathbf{v}_N))$ and $\mathbf{D}_f = \text{diag}(g(\mathbf{v}_{f,1}), \dots, g(\mathbf{v}_{f,N}))$, with $g(\cdot)$ any permutation invariant function, i.e., $g(\mathbf{P}\mathbf{x}) = g(\mathbf{x})$.

Proof We show that (10) satisfies (7), (8), and (9). Showing that (7) holds, requires only substituting (10) into $\phi(\phi(\lambda, \mathbf{V}), \mathbf{V}^H)$, which yields

$$\phi(\phi(\lambda, \mathbf{V}), \mathbf{V}^H) = \mathbf{D}^{-1}\mathbf{V}^H\mathbf{D}_f(\mathbf{D}_f^{-1}\mathbf{V}\mathbf{D}\lambda) = \lambda.$$

In order to show (8), notice that a permutation of the rows of $\mathbf{V}$ (the columns of $\mathbf{V}_f$) does not influence $\mathbf{D}$ and only permutes the diagonal entries of $\mathbf{D}_f$. Hence, we can write $\phi(\lambda, \mathbf{P}\mathbf{V})$ as

$$\phi(\lambda, \mathbf{P}\mathbf{V}) = (\mathbf{P}\mathbf{D}_f\mathbf{P}^T)^{-1}\mathbf{P}\mathbf{V}\mathbf{D}\lambda = \mathbf{P}\mathbf{D}_f^{-1}\mathbf{V}\mathbf{D}\lambda = \mathbf{P}\phi(\lambda, \mathbf{V}).$$

Finally, since a permutation of the columns of $\mathbf{V}$ (the rows of $\mathbf{V}_f$) does not influence $\mathbf{D}_f$ and only permutes the diagonal entries of $\mathbf{D}$, we can write $\phi(\mathbf{P}^T\lambda, \mathbf{V}\mathbf{P})$ as [cf. (9)]

$$\begin{aligned} \phi(\mathbf{P}^T\lambda, \mathbf{V}\mathbf{P}) &= \mathbf{D}_f^{-1}(\mathbf{V}\mathbf{P})(\mathbf{P}^T\mathbf{D}\mathbf{P})(\mathbf{P}^T\lambda) = \mathbf{D}_f^{-1}\mathbf{V}\mathbf{D}\lambda \\ &= \phi(\lambda, \mathbf{V}). \end{aligned}$$

□

Note that Theorem 1 proves the existence of a class of eligible dual graphs, but it does not indicate that *every* dual graph falls in this class. If we restrict ourselves to the class in (10), which can be described by the function $g(\cdot)$, the simplest choice for $g(\cdot)$ is $g(\mathbf{x}) = 1$. This results in $\lambda_f = \mathbf{V}\lambda$, but, e.g., any power of any norm is also a valid function, i.e., $g(\mathbf{x}) = \|\mathbf{x}\|_p^q$. A possible policy to design a dual graph could be to select the function $g(\cdot)$ that optimizes a particular figure of merit (such as the minimization

of the number of edges in the dual graph $\mathcal{G}_f$) yet keeping faithful to (A1)–(A3). This problem is discussed in more detail at the end of the following section.

Let us now Look at (10) for a few particular graph examples. First, it is clear that the cycle graph with $\mathbf{V}$ the normalized DFT matrix does not directly lead to a cycle graph but to a circular graph, which depends on how we pick $g(\cdot)$. It can also be shown that the dual graph of a Kronecker product graph is the Kronecker product graph of the related dual graphs, when $g(\mathbf{x}_1 \otimes \mathbf{x}_2) = g(\mathbf{x}_1)g(\mathbf{x}_2)$ (this is a reasonable assumption, e.g., take $g(\mathbf{x}) = 1$ or $g(\mathbf{x}) = \|\mathbf{x}\|_1$). For the Cartesian and strong product this is generally not the case.

To finalize this section, additional axioms can be imposed on $\mathbf{S}_f$ to further winnow the class of admissible functions ϕ . A possible avenue, not investigated here, is to impose a desirable behavior of $\mathbf{S}_f$ with respect to the intrinsic phase ambiguity of the primal EVD. Tackling this issue requires fixing the phases of the eigenvectors in some appropriate way. This can be dealt with by defining a canonical phase representation for every basis of eigenvectors. More precisely, for any arbitrary $\mathbf{V}$, we consider its canonical phase representation to be $\mathbf{V}\psi(\mathbf{V})$ where $\psi(\mathbf{V})$ is a phase shift matrix (diagonal matrix with unit norm diagonal elements) and $\psi(\cdot)$ computes the shift needed to turn $\mathbf{V}$ into its canonical representation $\mathbf{V}\psi(\mathbf{V})$. Under some technical requirements on ψ (here omitted) the dual shift construction can be made compatible with the notion of a canonical phase shift. For the experiments here presented, we adopt the default phase (or sign, when focusing on undirected graphs) convention in MATLAB's eigendecomposition function.

4.2 Optimization Approach

A different and complementary approach is to find a dual shift $\mathbf{S}_f$ for which certain properties of practical relevance are either enforced or promoted. For example, one may be interested in obtaining the sparsest $\mathbf{S}_f$, in recovering dual shifts without self-loops, or in both. This can be achieved by formulating judicious optimization problems where the variable to optimize is the dual shift, the constraints are designed to guarantee the desired topological properties, and (combinations of) suitable objective functions are used to promote convenient properties. To be rigorous, consider that the primal shift $\mathbf{S} = \mathbf{V}\mathbf{A}\mathbf{V}^H$ is given. Then, upon setting $\mathbf{V}_f = \mathbf{V}^H$ and $\mathbf{v}_{f,i} = \mathbf{V}^H \mathbf{e}_i$ (cf. Property 1), the dual shift $\mathbf{S}_f$ is found by solving

$$\min_{\{\mathbf{S}_f, \lambda_f\}} \ell(\mathbf{S}_f) \quad \text{s. to } \mathbf{S}_f = \sum_{i=1}^N \lambda_{f,i} \mathbf{v}_{f,i} \mathbf{v}_{f,i}^H, \quad \mathbf{S}_f \in \mathcal{S}. \quad (11)$$

In the problem above, the optimization variables are effectively the eigenvalues λ_f , since the constraint $\mathbf{S}_f = \sum_{i=1}^N \lambda_{f,i} \mathbf{v}_{f,i} \mathbf{v}_{f,i}^H$ forces the columns of $\mathbf{V}_f$ to be the eigenvectors of $\mathbf{S}_f$. Two defining features of the problem in (11) are the objective function $\ell(\cdot)$ and the constraint set $\mathcal{S}$.

The objective $\ell(\mathbf{S}_f)$ in (11) promotes desirable network structural properties on $\mathbf{S}_f$, such as sparsity or minimum-energy edge weights. The objective function can be defined as the weighted sum of multiple functions $\ell(\mathbf{S}_f) = \sum_{m=1}^M \eta_m \ell_m(\mathbf{S}_f)$, so that multiple properties are simultaneously promoted, with $\{\eta_m\}_{m=1}^M$ being nonnegative

weighting coefficients (hyper-parameters) that must be selected based on either prior knowledge or numerical search. While using $\|\cdot\|_0$ (or a surrogate) as one of the functions to minimize is a well-motivated approach (with the goal being minimizing the number of pairwise relationships between the frequencies), other interesting choices include the Frobenius norm of $\mathbf{S}_f$, the spectral norm, as well as smoothness metrics that minimize the variability (maximizes the smoothness) of a given set of signals in the dual domain [7,13].

The constraint set $\mathcal{S}$ imposes requirements on the dual shift, such as each entry being non-negative, each node having at least one neighbor, or the dual graph having no self-loops. Since the *effective* number of optimization variables in (11) is N (the size of the vector of eigenvalues λ_f), imposing a high number of (equality) constraints in $\mathcal{S}$ may render the problem infeasible.

Let us now come back to some earlier graph examples. Consider for instance the cycle graph with $\mathbf{V}$ the normalized DFT matrix and let us solve (11) using $\ell(\cdot) = \|\cdot\|_0$ and $\mathcal{S}$ the set of fully connected graphs. Then it is easy to show that the cycle graph is one of the solutions. For product graphs on the other hand, we can always make sure that the dual graph is a product graph of the same type by constraining $\mathcal{S}$ accordingly.

Finally, since the constraint $\mathbf{S}_f = \sum_{i=1}^N \lambda_{f,i} \mathbf{v}_{f,i} \mathbf{v}_{f,i}^H$ is linear, the tractability of the problem in (11) depends on the selection of costs and topological constraints in $\mathcal{S}$. If both $\ell(\cdot)$ and $\mathcal{S}$ are convex, then the resultant optimization problem can be efficiently handled. Furthermore, if strict convexity is present, the solution will be unique. Variations of the formulation in (11) have been analyzed in the literature for problems different from the one considered in this paper and, in particular, in the context of network topology inference from nodal observations [13,19].

4.2.1 Consistency with the Axiomatic Approach

An important question when implementing the approach in (11) is to investigate whether the dual shift obtained from the optimization satisfies axioms (A1)–(A3), already deemed as desirable properties. To analyze this, we will assume for simplicity that the solution to (11), denoted as $\{\mathbf{S}_f^*, \lambda_f^*\}$, is unique. Under this assumption, the following result holds.

Theorem 2 *If $\ell(\cdot)$ and $\mathcal{S}$ in (11) are invariant to permutations, then the (unique) solution $\{\mathbf{S}_f^*, \lambda_f^*\}$ satisfies the Axioms of Reordering (A2) and Permutation (A3).*

Proof We begin by showing that (A2) is satisfied, where it is required that $\mathbf{S}_f^*$, the dual shift obtained for $\mathbf{S}$, is the same than that for $\tilde{\mathbf{S}} = \mathbf{PSP}^T$, denoted by $\tilde{\mathbf{S}}_f^*$. For the permuted primal graph $\tilde{\mathbf{S}}$ we have that $\tilde{\mathbf{V}} = \mathbf{PV}$ and, hence, $\tilde{\mathbf{V}}_f = \tilde{\mathbf{V}}^H = \mathbf{V}^H \mathbf{P}^T$. This implies that, when solving (11) for the permuted primal graph, the linear constraint $\tilde{\mathbf{S}}_f = \sum_{i=1}^N \tilde{\lambda}_{f,i} \tilde{\mathbf{v}}_{f,i} \tilde{\mathbf{v}}_{f,i}^H = \tilde{\mathbf{V}}_f \text{diag}(\tilde{\lambda}_f) \tilde{\mathbf{V}}_f^T$ can be written as $\tilde{\mathbf{S}}_f = \mathbf{V}^H \mathbf{P}^T \text{diag}(\tilde{\lambda}_f) \mathbf{PV}$. Hence, if we set $\tilde{\lambda}_f = \mathbf{P} \lambda_f^*$ we have that $\tilde{\mathbf{S}}_f = \mathbf{V}^H \mathbf{P}^T \text{diag}(\mathbf{P} \lambda_f^*) \mathbf{PV} = \mathbf{V}^H \text{diag}(\mathbf{P}^T \mathbf{P} \lambda_f^*) \mathbf{V} = \mathbf{V}^H \text{diag}(\lambda_f^*) \mathbf{V} = \mathbf{S}_f^*$. Since the pair $\{\mathbf{S}_f^*, \mathbf{P} \lambda_f^*\}$ satisfies the linear constraint and both $\mathcal{S}$ and $\ell(\cdot)$ only depend on the

GSO, we have that $\{\mathbf{S}_f^*, \mathbf{P}\boldsymbol{\lambda}_f^*\}$ is the global minimizer of (11) for $\bar{\mathbf{V}}_f = \bar{\mathbf{V}}^H = \mathbf{V}^H \mathbf{P}^T$. In other words, the dual shifts $\mathbf{S}_f^*$ and $\bar{\mathbf{S}}_f^*$ are the same, so that (A2) is satisfied.

To show the result for (A3), we begin with the primal graph $\mathbf{S}$ and consider that its EVD is given by $\mathbf{S} = \mathbf{V} \text{diag}(\boldsymbol{\lambda}) \mathbf{V}^H$. Using this decomposition, we solve problem (11) with $\mathbf{V}_f = \mathbf{V}^H$ to obtain $\{\mathbf{S}_f^*, \boldsymbol{\lambda}_f^*\}$. Then, we consider $\mathbf{S} = \mathbf{V} \mathbf{P} \text{diag}(\mathbf{P}\boldsymbol{\lambda}) \mathbf{P}^T \mathbf{V}^H$ as an alternative (equally valid) EVD for the primal graph $\mathbf{S}$, use $\mathbf{V}_f = \mathbf{P}^T \mathbf{V}^H$ as the input to (11), and denote the obtained solution by $\{\bar{\mathbf{S}}_f^*, \bar{\boldsymbol{\lambda}}_f^*\}$. For (A3) to be satisfied, we need to prove that $\bar{\mathbf{S}}_f^* = \mathbf{P}^T \mathbf{S}_f^* \mathbf{P}$. The first step is to show that, when replacing $\bar{\mathbf{V}}_f$ with $\mathbf{V}_f \mathbf{P}$ and $\bar{\mathbf{S}}_f^*$ with $\mathbf{P}^T \mathbf{S}_f^* \mathbf{P}$, the linear constraint is satisfied. Specifically, we have that

$$\bar{\mathbf{S}}_f = \sum_{i=1}^N \bar{\lambda}_{f,i} \bar{\mathbf{v}}_{f,i} \bar{\mathbf{v}}_{f,i}^H = \bar{\mathbf{V}}_f \text{diag}(\bar{\boldsymbol{\lambda}}_f) \bar{\mathbf{V}}_f^H = \mathbf{P}^T \mathbf{V}_f \text{diag}(\bar{\boldsymbol{\lambda}}_f) \mathbf{V}_f^H \mathbf{P}.$$

Upon setting $\bar{\boldsymbol{\lambda}}_f = \boldsymbol{\lambda}_f^*$, we have that $\bar{\mathbf{S}}_f = \mathbf{P}^T \mathbf{V}_f \text{diag}(\boldsymbol{\lambda}_f^*) \mathbf{V}_f^H \mathbf{P} = \mathbf{P}^T \mathbf{S}_f^* \mathbf{P}$. In other words, we have that the solution $\{\mathbf{P}^T \mathbf{S}_f^* \mathbf{P}, \boldsymbol{\lambda}_f^*\}$ is feasible. Finally, leveraging the assumptions that $\ell(\cdot)$ and $\mathcal{S}$ are invariant to permutations, it follows that $\{\mathbf{P}^T \mathbf{S}_f^* \mathbf{P}, \boldsymbol{\lambda}_f^*\}$ is also optimal. $\square$

Notice the permutation invariance of the constraints and the objective function assumed in Theorem 2. In other words, when encoding the topological properties to be enforced in the set $\mathcal{S}$ and those to be promoted in the function $\ell(\cdot)$, we need to focus on encodings that depend on the properties of the underlying graph, but not on the specific ordering selected for the nodes. As a result, formulations where the objective promotes sparsity by setting $\ell(\cdot) = \|\cdot\|_0$, or those where the set $\mathcal{S}$ guarantees that the entries of the dual shift are all non-negative, will satisfy (A3). On the other hand, if either the objective or the topological constraint set is sensitive to the ordering of the nodes (e.g., by enforcing that there must exist a link between nodes 1 and 2), then (A3) will not be satisfied.

Finally, we shift focus to the Axiom of Duality (A1) and highlight that the solution to (11) will, in general, not abide by this axiom. In particular, for axiom (A1) to hold, it is necessary for the original shift $\mathbf{S}$ itself to be optimal in the sense encoded by (11). To elaborate on this, consider the unitary matrix $\mathbf{U}$ and the associated shift set $\mathcal{S}_{\mathbf{U}} := \{\mathbf{S} = \mathbf{V} \text{diag}(\boldsymbol{\lambda}) \mathbf{V}^H \mid \mathbf{V} = \mathbf{U} \text{ and } \boldsymbol{\lambda} \in \mathbb{C}^N\}$, where each selection of $\boldsymbol{\lambda}$ gives rise to a different element of $\mathcal{S}_{\mathbf{U}}$. Moreover, let $\mathbf{S}^*$ denote the solution to (11) when $\mathbf{V}_f = \mathbf{U}$ and $\mathbf{S}_f^*$ the solution when $\mathbf{V}_f = \mathbf{U}^H$. Then, it holds that: (i) the dual shift for any $\mathbf{S} \in \mathcal{S}_{\mathbf{U}}$ is given by $\mathbf{S}_f^*$, and (ii) the dual of $\mathbf{S}_f^*$ is $\mathbf{S}^*$. Hence, $\mathbf{S}^*$ is the only element of $\mathcal{S}_{\mathbf{U}}$ that guarantees that the dual of the dual is the original graph and, therefore, that (A1) holds. Alternatively, one can see $\mathcal{S}_{\mathbf{U}}$ as a shift class whose (canonical) representative is $\mathbf{S}^*$. With this interpretation any $\mathbf{S} \in \mathcal{S}_{\mathbf{U}}$ is first mapped to $\mathbf{S}^*$ and then $\mathbf{S}^*$ serves as input for (11). Under this assumption, the invertibility of the dual mapping is achieved.

4.3 The Particular Case of Adjacency Shifts

The most widely used matrix representations of a graph are the adjacency and Laplacian matrices. Section 3 discussed the case where the primal shift $\mathbf{S}$ was set to the Laplacian, concluding that, in general, the associated dual shift will not have a Laplacian form. Assuming that the primal shift $\mathbf{S}$ is set to an adjacency matrix with no self-loops, the question discussed here is whether the dual shift can also have the form of an adjacency matrix without self-loops. We formally answer this question in the form of the following theorem.

Theorem 3 *Given a primal graph shift $\mathbf{S} = \mathbf{V} \text{diag}(\boldsymbol{\lambda}) \mathbf{V}^H$ with $[\mathbf{S}]_{i,i} = 0$ for all i , there always exists a dual graph shift $\mathbf{S}_f$ that can be diagonalized by the eigenvectors $\mathbf{V}_f = \mathbf{V}^H$ and whose entries satisfy $[\mathbf{S}_f]_{i,i} = 0$ for all i .*

Proof In showing this, the $N \times N$ matrix $\mathbf{W}$ defined as $[\mathbf{W}]_{i,j} = |[\mathbf{V}]_{i,j}|^2$, with $|\cdot|$ denoting the absolute value, plays a critical role.

First notice that having $[\mathbf{S}_f]_{i,i} = 0$ requires $\mathbf{W}^T \boldsymbol{\lambda}_f = \mathbf{0}$; that is, matrix $\mathbf{W}^T$ needs to be singular and $\boldsymbol{\lambda}_f$ must be in the nullspace of $\mathbf{W}^T$. To see why this is the case, recall that $\mathbf{V}_f = \mathbf{V}^H$ implies that $\mathbf{S}_f$ can be written as $\mathbf{S}_f = \mathbf{V}^H \text{diag}(\boldsymbol{\lambda}_f) \mathbf{V}$. Vectorizing matrix $\mathbf{S}_f$, we have that $\text{vec}(\mathbf{S}_f) = (\mathbf{V}^T \odot \mathbf{V}^H) \boldsymbol{\lambda}_f$, where $\odot$ denotes the Khatri–Rao (column-wise Kronecker) product. We can now focus on the N rows of the $N^2 \times N$ matrix $\mathbf{V}^T \odot \mathbf{V}^H$ associated with the diagonal elements of $\mathbf{S}_f$. In particular, with † denoting complex conjugate, the i th diagonal element of the dual shift can be written as

$$\begin{aligned} [\mathbf{S}_f]_{i,i} &= \sum_{j=1}^N [\mathbf{V}^T]_{i,j} [\mathbf{V}^H]_{i,j} \lambda_{f,j} = \sum_{j=1}^N [\mathbf{V}]_{j,i} [\mathbf{V}]_{j,i}^\dagger \lambda_{f,j} \\ &= \sum_{j=1}^N |[\mathbf{V}]_{j,i}|^2 \lambda_{f,j} = \sum_{j=1}^N [\mathbf{W}^T]_{i,j} \lambda_{f,j}. \end{aligned}$$

This readily implies that the vector collecting the N diagonal entries of $\mathbf{S}_f$ can be obtained as $\mathbf{W}^T \boldsymbol{\lambda}_f$ and, as a result, $[\mathbf{S}_f]_{i,i} = 0$ requires $\mathbf{W}^T \boldsymbol{\lambda}_f = \mathbf{0}$.

A similar argument can be followed to show that having $[\mathbf{S}]_{i,i} = 0$ for all i implies that $\mathbf{W} \boldsymbol{\lambda} = \mathbf{0}$, so that matrix $\mathbf{W}$ is singular and $\boldsymbol{\lambda}$ belongs to the null space of $\mathbf{W}$.

Finally, the square matrix $\mathbf{W}$ being singular implies that $\mathbf{W}^T$ is singular as well. Hence, if $[\mathbf{S}]_{i,i} = [\mathbf{V} \text{diag}(\boldsymbol{\lambda}) \mathbf{V}^H]_{i,i} = 0$ for all i , then both $\mathbf{W}$ and $\mathbf{W}^T$ are singular and, therefore, there exists a dual shift $\mathbf{S}_f = \mathbf{V}^H \text{diag}(\boldsymbol{\lambda}_f) \mathbf{V}$ for which $[\mathbf{S}_f]_{i,i} = 0$ for all i . $\square$

Theorem 3 shows that, given a primal shift with no self-loops, we can always find a dual shift that shares that same feature. Moreover, from the proof technique of the theorem, it also follows that when $\text{rank}(\mathbf{W}) = N - 1$, the dual shift is unique up to a scaling ambiguity. Specifically, let $\mathbf{W} = \mathbf{U}_L \boldsymbol{\Sigma} \mathbf{U}_R^T$ denote the singular value decomposition of $\mathbf{W}$ with the columns of $\mathbf{U}_L$ denoting the left singular vectors and those of $\mathbf{U}_R$ their corresponding right singular vectors and suppose, without loss of generality, that $\mathbf{u}_{L,1}$ and $\mathbf{u}_{R,1}$ are the singular vectors associated with the (unique) zero singular value. It then follows that $\boldsymbol{\lambda}$ is a scaled version of $\mathbf{u}_{R,1}$ and $\boldsymbol{\lambda}_f$ is a scaled version of $\mathbf{u}_{L,1}$ and, as a result, the dual shift can be written as $\mathbf{S}_f = \alpha \mathbf{V}^H \text{diag}(\mathbf{u}_{L,1}) \mathbf{V}$, with parameter α representing the scaling ambiguity.

For (primal) adjacency matrices leading to $\text{rank}(\mathbf{W}) = N - 1$, there is no guarantee that the dual adjacency $\mathbf{S}_f = \alpha \mathbf{V}^H \text{diag}(\mathbf{u}_{L,1}) \mathbf{V}$ is sparse or that it satisfies any property other than being diagonalizable by $\mathbf{V}^H$ and its diagonal elements being zero. An interesting point is, therefore, to identify families of primal graphs for which $\text{rank}(\mathbf{W}) < N - 1$. This would increase the degrees of freedom of λ_f , which lies in a subspace of dimension $D = N - \text{rank}(\mathbf{W})$, enlarging the feasibility set of the optimization in (11) and opening the door to find dual shifts that achieve a lower cost in the associated optimization.

5 Illustrative Simulations

We provide a few simple examples illustrating how representations of the frequency domain that go beyond one dimension can be of interest. Throughout the simulations, we consider three methods to find a dual graph. As explained in the paper, in all three cases the eigenvectors are set to $\mathbf{V}^H$, with the difference being on how the eigenvalues are obtained. The first approach (dual graph A) corresponds to that in Sect. 4.1 and sets the dual eigenvalues to $\lambda_f = \mathbf{V}\lambda$ (cf. discussion after Theorem 1). The second approach (dual graph B) follows Sect. 4.2 and sets the objective in (11) to $\ell(\cdot) = \|\cdot\|_1$. That is, we aim to obtain a sparse dual graph but replacing the non-convex 0-norm with its convex surrogate. Finally, the third approach (dual graph C) forces the dual graph to be an adjacency matrix as discussed in Sect. 4.3. If multiple graphs can be obtained (i.e., if the dimension of the null space of matrix $\mathbf{W}$ is greater than one), then we set the dual graph as the one with minimum 1-norm. Note that there can be cases where the dual graphs B and C coincide.

5.1 Examples of Primal and Dual Graphs

We start by generating primal graphs as realizations of an Erdős-Rényi (ER) random graph model [1] with $N = 10$ and edge probability $p = 0.15$. The results are shown in Fig. 2. The top row depicts one example of a primal graph along with its three associated dual graphs as described above. The nodes in the dual graphs are sorted according to a (decreasing) ordering of the entries of the eigenvalues $\lambda = [\lambda_1, \dots, \lambda_N]^T$ of the primal graph. In particular, node 1 of $\mathbf{S}_f$ represents the frequency (primal eigenvector) associated with the largest positive primal eigenvalue whereas node N represents the frequency associated with the most negative primal eigenvalue, so that $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$. Notice that dual graphs are not necessarily sparse, specially those associated with method A where sparsity is not explicitly promoted. The obtained representations suggest that one-dimensional frequency representations where a frequency k is considered to be close to a frequency k' if $|\lambda_k - \lambda_{k'}|$ is small may not be able to capture the more complex relationships among frequencies. Indeed, the plots of $\mathbf{S}_f$ reveal that the strongest connections in the dual graphs are not necessarily between adjacent nodes, which are the closest in terms of the distance between the associated primal eigenvalues.

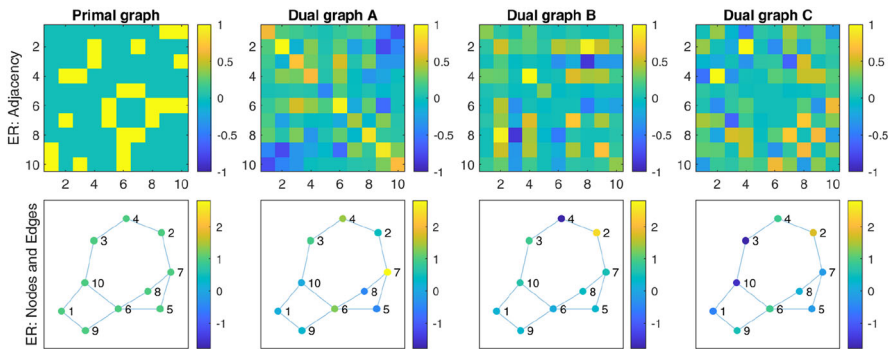


Fig. 2 The top row provides a (heat-map) representation of the weighted adjacency matrices of (left to right): the primal ER graph, the dual shift obtained using the approach in Sect. 4.1 (Dual Shift A), the dual shift obtained using the approach in Sect. 4.2 (Dual Shift B), the dual shift obtained using the approach in Sect. 4.3 (Dual Shift C). The bottom row provides a node-edge representation of the primal graph (left-most figure) along with representations of the eigenvalues of the dual graph as a signal over the primal graph for each of the three methods considered in the simulations

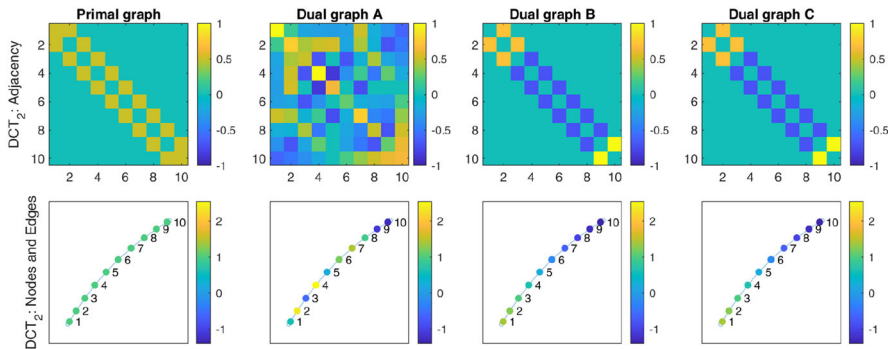


Fig. 3 The top row provides a (heat-map) representation of the weighted adjacency matrices of (left to right): the primal DCT graph, the dual shift obtained using the approach in Sect. 4.1 (Dual Shift A), the dual shift obtained using the approach in Sect. 4.2 (Dual Shift B), the dual shift obtained using the approach in Sect. 4.3 (Dual Shift C). The bottom row provides a node-edge representation of the primal graph (left-most figure) along with representations of the eigenvalues of the dual graph as a signal over the primal graph for each of the three methods considered in the simulations

To gain further insights, the bottom row of Fig. 2 represents the eigenvalues of the dual graph. From duality it follows that each of the frequencies of the dual graph is associated with a particular node and, hence, the N eigenvalues of $\mathbf{S}_f$ can be viewed as a signal defined over the original primal graph $\mathbf{S}$. As a result, we represent the obtained eigenvalues over the original primal graph. The figure confirms that the methods give rise to different estimations of the dual eigenvalues. From the eigendecomposition of $\mathbf{S}_f$ and in accordance with the discussion following Property 1, the larger the absolute value of $[\lambda_f]_n$, the more important the frequency pattern captured by node n is to describe the dual graph $\mathbf{S}_f$. To see that this is the case, we may write $\mathbf{S}_f = \sum_{n=1}^N [\lambda_f]_n \mathbf{v}_{f,n} \mathbf{v}_{f,n}^H = \sum_{n=1}^N [\lambda_f]_n \tilde{\mathbf{e}}_n \tilde{\mathbf{e}}_n^H$, where $\mathbf{e}_n$ is the n th canonical vector and $\tilde{\mathbf{e}}_n = \mathbf{V}^H \mathbf{e}_n$ is the frequency pattern associated with node n . This can be relevant, for

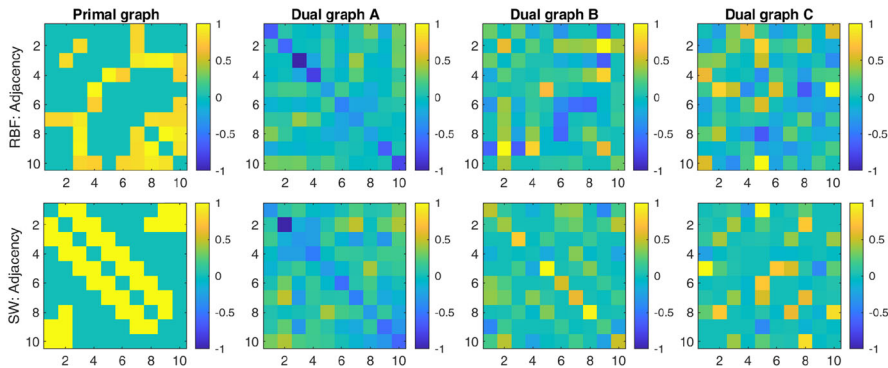


Fig. 4 Primal and dual shifts for two additional random graph models: RBF (top row) and small-world (bottom row)

example, in scenarios where one is forced to operate with only a subset of nodes of the primal graph and the goal is to select the nodes that better preserve the interaction between frequencies.

Figure 3 is the counterpart of Fig. 2 when the primal graph is given by the graph associated with the discrete cosine transform (DCT) of type II [15]. As in the previous case, the left-most panel in the top row depicts the primal graph, which is an undirected path with a self-loop in the two extremal (i.e., first and last) nodes, guaranteeing that the degree is two for all the nodes in the graph. The other three graphs in the top row correspond to the dual graphs obtained using the algorithms presented in this paper. The bottom row provides the representation of the dual eigenvalues as a graph signal on the (primal) DCT graph. Focusing first on the top row, we observe that the dual graphs B and C are the same, meaning that the sparsest graph does not have self-loops. We also observe that while the dual graphs B and C are sparse, once again, the dual graph A is not. Finally, it is also worth noticing that the dual graphs B and C, on top of being sparse, are very regular. Indeed, these two graphs are also undirected paths, but with positive and negative edges and without self-loops. Regarding the bottom row, the most striking observation is that the eigenvalues of the dual graphs B and C are perfectly ordered in the primal graph. This adds to the idea that the regularity of the DCT graph in the primal domain leads to a regularly-structured dual graph.

Figure 4 replicates the analysis in Figs. 2 and 3 (top row) but for two new types of graphs, namely, (i) a geometric radial basis function (RBF) graph where nodes are randomly dropped in a unit square, edges are formed between nodes that are at distance less than 0.75, and edge weights are given by a Gaussian kernel with standard deviation 0.5 [6]; and (ii) a small-world graph with $K = 2$ and a rewiring probability of 0.15 [26]. We first notice that, as already mentioned for the ER case, for none of the primal graphs the axiomatic approach (Dual Graph A) gives rise to a sparse dual graph. It is also evident that even the other two methods, which explicitly promote sparsity, are in some cases unable to recover *very* sparse dual graphs, reinforcing the notion that the interactions between frequencies are more complex than what can be represented by a one-dimensional structure. These complex interactions could be relevant in a

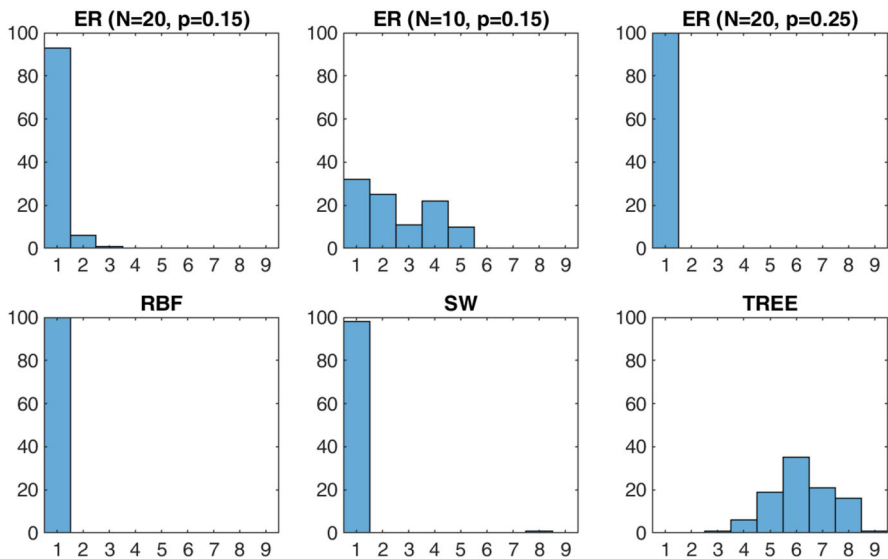


Fig. 5 For each model, 100 graph realizations are drawn and the dimension of the null space of matrix $\mathbf{W}$ (see Sect. 4.3) is computed. Each panel represents the histogram of that dimension. The top row corresponds to ER graphs with a different number of nodes N and edge probability p . The bottom row considers non-ER models, namely, RBF, SW, and a random tree, all with $N = 20$ nodes

number of problems including, for example, scenarios where one needs to estimate the value (or the power) that a signal has in a particular frequency band using values from other frequencies; see Sect. 5.3.

Finally, when observing Figs. 2, 3 and 4 jointly, it is worth noting that graphs with a very strong structure in the primal domain can be associated with strong and regular structures in the dual domain. This indicates that for markedly regular primal graphs, such as the one of the DCT graph, a one-dimensional frequency representation could be argued to be sufficient.

5.2 Uniqueness of the Adjacency Dual Shift

An interesting observation from Figs. 2, 3 and 4 is that it is often the case that the dual method C, which forces the diagonal elements of $\mathbf{S}_f$ to be zero, yields a matrix that is sparse. Based on Theorem 3 and the subsequent discussion, this indicates that the dimension of the null space of $\mathbf{W}$ must be larger than one and, hence, the set of feasible dual graphs is sufficiently large so that a sparse one can be found. To confirm this, we select different types of random graph models, for each of them we draw 100 realizations, compute the dimension of the null space of their $\mathbf{W}$ matrices, and plot the corresponding histograms in Fig. 5. The top row corresponds to ER graphs with different parameters (number of nodes N and edge probability), while the second row corresponds to three other graph types. Recall that the null space of $\mathbf{W}$ being of dimension 1 implies that the dual graph is unique (up to a trivial scalar ambiguity)

whereas larger dimensions indicate that the set of feasible dual shifts with the form of an adjacency matrix is large. Our experiments reveal that the specific graph type and parameters defining the graph model strongly influence the dimension of the null space of $\mathbf{W}$ and, thus, the size of the space of adjacency dual shifts. In particular, for none of the realizations of tree graphs, the adjacency dual shift was uniquely determined by the constraint of having no self-loops.

5.3 Frequency Estimation Using Dual Graphs

We now illustrate how the dual graph can be leveraged to address problems related to the frequency domain. In particular, we consider a graph signal $\mathbf{x} \in \mathbb{R}^N$ defined on the primal graph $\mathbf{S} = \mathbf{V}\text{diag}(\boldsymbol{\lambda})\mathbf{V}^H$ and whose frequency representation is given by $\tilde{\mathbf{x}} = \mathbf{V}^H \mathbf{x}$. In this context, we want to estimate the value of $\tilde{\mathbf{x}}$ in a particular frequency k by using a (graph) filter in the dual domain that relies on the values of $\tilde{\mathbf{x}}$ at frequencies $k' \neq k$. To be more specific, we want to estimate $\tilde{\mathbf{x}}$ via the following dual graph filter operation

$$\tilde{\mathbf{x}} = \sum_{l=0}^L h_l \mathbf{S}_f^l \tilde{\mathbf{x}} = h_0 \tilde{\mathbf{x}} + h_1 \mathbf{S}_f \tilde{\mathbf{x}} + \cdots + h_L \mathbf{S}_f^L \tilde{\mathbf{x}}, \quad (12)$$

where $L < N$. Clearly, to avoid the trivial solution of setting $h_0 = 1$ and $h_l = 0$ for all $l \geq 1$, we force h_0 to be zero so that the problem to solve is

$$\mathbf{h}^* = \underset{\{\mathbf{h} \in \mathbb{R}^L\}}{\text{argmin}} \left\| \tilde{\mathbf{x}} - \sum_{l=1}^L [\mathbf{h}]_l \mathbf{S}_f^l \tilde{\mathbf{x}} \right\|^2. \quad (13)$$

The frequency estimate is then obtained as $\tilde{\mathbf{x}}^* = \sum_{l=1}^L [\mathbf{h}^*]_l \mathbf{S}_f^l \tilde{\mathbf{x}}$ and the associated (normalized) estimation error as $\|\tilde{\mathbf{x}}^* - \tilde{\mathbf{x}}\|^2 / \|\tilde{\mathbf{x}}\|^2$.

We solve (13) for 4 different types of graphs: the dual graphs A-C as explained at the beginning of Sect. 5 and the directed cycle graph, which implements the regular convolution that tries to estimate the frequency coefficients using the values of adjacent frequencies.

In Fig. 6, we present the results for two different types of primal graphs (ER and RBF) and two different types of graph signals, namely, diffused sparse signals and bandlimited signals. To generate the diffused sparse signals, we select uniformly at random 3 seeding nodes, generate a non-zero signal value at each of the seeds from a uniform distribution in $[0, 1]$, and then diffuse those values using a low-pass filter of degree 2 with an exponential response $h_l = \beta^l$ for $\beta = 0.8$. The bandlimited signals $\mathbf{x}$ are generated as a linear combination of the top 4 eigenvectors (those associated with the largest eigenvalues), where the combination weights are drawn from a uniform distribution in $[0, 1]$.

The results reveal that the approaches that leverage the relation between frequencies provided by the dual shifts presented in this paper do a better job than the classical convolution. As expected, as the number of filter taps increases, the error decreases.

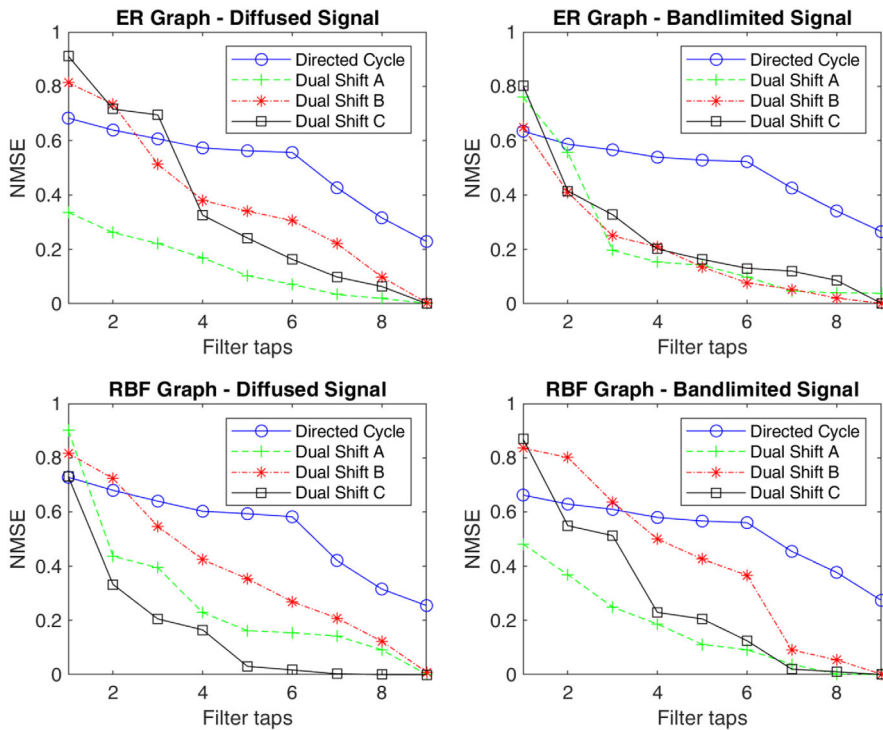


Fig. 6 Normalized Mean Squared Error (NMSE) as a function of the order (number of taps) of the filter for the estimation of the frequency content of a signal using (low-order) graph filters defined over different dual graphs. The errors are averaged over 100 (signal and graph) realizations. Each curve corresponds to a different type of dual shift: the directed cycle plus the three dual graphs described at the beginning of Sect. 5. The top row correspond to signals defined over a primal ER graph and the bottom row to signals defined over a primal RBF graph. The left column considers signals that adhere to a diffused sparse signal model while, for the right column, the signals were generated using a low-pass bandlimited model

However, in none of the scenarios considered, the approach based on regular convolutions is able to attain zero error, while many of the approaches based on the dual shift are able to achieve perfect estimation at all the frequency bands for filters with around 9 coefficients. While the particular filter that achieves the smallest error depends on the configuration at hand, we observe that the “Dual Shift A” based on the approach described in Sect. 4.1 yields in general good results. The poor performance achieved by the estimate based on the standard convolution suggests, once again, that similarity among frequencies goes beyond measuring proximity in terms of their associated primal eigenvalues.

6 Conclusions and Open Questions

This paper investigated the problem of identifying the support associated with the frequency representation of graph signals. Given the (primal) graph shift operator

supporting graph signals of interest, the problem was formulated as that of finding a compatible *dual* graph shift operator that serves as a domain for the frequency representation of these signals. We first identified the eigenvectors of the dual shift, showing that those correspond to how each of the nodes expresses the different graph frequencies. We then proposed different alternatives to find the dual eigenvalues and characterized relevant properties that those eigenvalues must satisfy. Future work includes considering additional properties for the dual eigenvalues so that the size of feasible dual shift operators is reduced, and identifying additional results connecting the vertex domain with the frequency domain. The results in this paper constitute a first step towards understanding the structure of the signals in the frequency domain as well as developing enhanced GSP algorithms for signal compression, frequency grouping, filtering, and spectral estimation schemes.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Bollobás, B.: Random Graphs. Springer, New York (1998)
2. Chen, S., Varma, R., Sandryhaila, A., Kovačević, J.: Discrete signal processing on graphs: sampling theory. *IEEE Trans. Signal Process.* **63**(24), 6510–6523 (2015). <https://doi.org/10.1109/TSP.2015.2469645>
3. Cloninger, A., Li, H., Saito, N.: Natural graph wavelet packet dictionaries. arXiv preprint [arXiv: 2009.09020](https://arxiv.org/abs/2009.09020) (2020)
4. Cloninger, A., Steinerberger, S.: On the dual geometry of Laplacian eigenfunctions. *Exp. Math.* (2020). To appear
5. Coutino, M., Isufi, E., Leus, G.: Advances in distributed graph filtering. *IEEE Trans. Signal Process.* **67**(9), 2320–2333 (2019)
6. Dong, X., Thanou, D., Frossard, P., Vandergheynst, P.: Learning Laplacian matrix in smooth graph signal representations. *IEEE Trans. Signal Process.* **64**(23), 6160–6173 (2016)
7. Dong, X., Thanou, D., Rabbat, M., Frossard, P.: Learning graphs from data: a signal representation perspective. *IEEE Signal Process. Mag.* **36**(3), 44–63 (2019)
8. Isufi, E., Loukas, A., Simonetto, A., Leus, G.: Autoregressive moving average graph filtering. *IEEE Trans. Signal Process.* **65**(2), 274–288 (2017). <https://doi.org/10.1109/TSP.2016.2614793>
9. Leus, G., Segarra, S., Ribeiro, A., Marques, A.G.: The dual graph shift operator: Identifying the support of the frequency domain. arXiv preprint [arXiv: 1705.08987](https://arxiv.org/abs/1705.08987) (2017)
10. Li, H., Saito, N.: Metrics of graph Laplacian eigenvectors. *Proc. SPIE 11138, Wavelets and Sparsity XVIII*, 111381K (2019)
11. Marques, A.G., Segarra, S., Leus, G., Ribeiro, A.: Sampling of graph signals with successive local aggregations. *IEEE Trans. Signal Process.* **64**(7), 1832–1843 (2016). <https://doi.org/10.1109/TSP.2015.2507546>
12. Marques, A.G., Segarra, S., Leus, G., Ribeiro, A.: Stationary graph processes and spectral estimation. *IEEE Trans. Signal Process.* **65**(22), 5911–5926 (2017). <https://doi.org/10.1109/TSP.2017.2739099>
13. Mateos, G., Segarra, S., Marques, A.G., Ribeiro, A.: Connecting the dots: identifying network structure via graph signal processing. *IEEE Signal Process. Mag.* **36**(3), 16–43 (2019)

14. Narang, S.K., Ortega, A.: Perfect reconstruction two-channel wavelet filter banks for graph structured data. *IEEE Trans. Signal Process.* **60**(6), 2786–2799 (2012)
15. Puschel, M., Moura, J.M.F.: Algebraic signal processing theory: 1-d space. *IEEE Trans. Signal Process.* **56**(8), 3586–3599 (2008). <https://doi.org/10.1109/TSP.2008.925259>
16. Saito, N.: How can we naturally order and organize graph Laplacian eigenvectors? In: 2018 IEEE Statistical Signal Processing Workshop (SSP), pp. 483–487. Freiburg, Germany (2018). <https://doi.org/10.1109/SSP.2018.8450808>
17. Sandryhaila, A., Moura, J.: Discrete signal processing on graphs. *IEEE Trans. Signal Process.* **61**(7), 1644–1656 (2013). <https://doi.org/10.1109/TSP.2013.2238935>
18. Sandryhaila, A., Moura, J.: Discrete signal processing on graphs: frequency analysis. *IEEE Trans. Signal Process.* **62**(12), 3042–3054 (2014). <https://doi.org/10.1109/TSP.2014.2321121>
19. Segarra, S., Marques, A.G., Mateos, G., Ribeiro, A.: Network topology inference from spectral templates. *IEEE Trans. Signal Inf. Process. Netw.* **3**(3), 467–483 (2017). <https://doi.org/10.1109/TSIPN.2017.2731051>
20. Segarra, S., Marques, A.G., Ribeiro, A.: Optimal graph-filter design and applications to distributed linear network operators. *IEEE Trans. Signal Process.* **65**(15), 4117–4131 (2017). <https://doi.org/10.1109/TSP.2017.2703660>
21. Shi, J., Moura, J.M.F.: Graph signal processing: Modulation, convolution, and sampling. *arXiv preprint arXiv: 1912.06762* (2019)
22. Shuman, D., Narang, S., Frossard, P., Ortega, A., Vandergheynst, P.: The emerging field of signal processing on graphs: extending high-dimensional data analysis to networks and other irregular domains. *IEEE Signal Process. Mag.* **30**(3), 83–98 (2013). <https://doi.org/10.1109/MSP.2012.2235192>
23. Shuman, D.I., Ricaud, B., Vandergheynst, P.: Vertex-frequency analysis on graphs. *Appl. Comput. Harmonic Anal.* **40**(2), 260–291 (2016). <https://doi.org/10.1016/J.ACHA.2015.02.005>
24. Tay, D.B.H., Zhang, J.: Techniques for constructing biorthogonal bipartite graph filter banks. *IEEE Trans. Signal Process.* **63**(21), 5772–5783 (2015)
25. Teke, O., Vaidyanathan, P.P.: Discrete uncertainty principles on graphs. In: *Asilomar Conf. on Signals, Systems, and Computers*, pp. 1475–1479. Pacific Grove (2016). <https://doi.org/10.1109/ACSSC.2016.7869622>
26. Watts, D., Strogatz, S.: Collective dynamics of small-world networks. *Nature* **393**, 440–442 (1998)
27. Yellen, J., Gross, J.L.: *Handbook of Graph Theory (Discrete Mathematics and Its Applications)*. CRC Press, Boca Raton (2003)
28. Zhu, X., Rabbat, M.: Approximating signals supported on graphs. In: *IEEE Intl. Conf. Acoust., Speech and Signal Process. (ICASSP)*, pp. 3921–3924 (2012). <https://doi.org/10.1109/ICASSP.2012.6288775>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.