A General Convolution Theorem for Graph Data

Natali, Alberto; Leus, Geert

DOI
10.1109/IEEECONF56349.2022.10052018

Publication date
2022

Document Version
Final published version

Published in
56th Asilomar Conference on Signals, Systems and Computers, ACSSC 2022

Citation (APA)
Natali, A., & Leus, G. (2022). A General Convolution Theorem for Graph Data. In M. B. Matthews (Ed.), *56th Asilomar Conference on Signals, Systems and Computers, ACSSC 2022* (pp. 48-52). (Conference Record - Asilomar Conference on Signals, Systems and Computers; Vol. 2022-October). IEEE. https://doi.org/10.1109/IEEECONF56349.2022.10052018

A General Convolution Theorem for Graph Data

Alberto Natali and Geert Leus
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology, Delft, The Netherlands

***Abstract*—This paper focuses on the field of graph signal processing (GSP) and studies the node-varying graph filter (NV-GF) which has been proposed as a way to broaden the applicability of the classical graph filter (C-GF). In particular, we state and prove a new convolution theorem for a NV-GF which extends both the one for a C-GF and the one for a time-varying filter. The theorem relies on the definition of a so-called dual graph which characterizes the support of the frequency domain. The dual graph concept has been studied only very recently and many versions exist, yet the proposed convolution theorem is independent of the particular version. More interestingly, using non-stationary graph data on the primal graph, we can use the proposed convolution theorem to learn the dual graph and thereby introduce an innovative data-driven dual graph estimation technique.**

I. Introduction

Convolution is the central component of architectures such as digital filters [1] and (convolutional) neural networks [2], underpinning a multitude of applications including time series prediction [3], speech recognition and computer vision [4], to name a few. Even though the convolution among two functions is usually defined in the Euclidean space, graph signal processing (GSP) [5] effectively extends its shift-scale-sum principle to data residing on non Euclidean domains, modeled by a graph. This is possible through so called graph filters [6], [7], architectures which are parametric on the mathematical structure defining the *shift* operation, which brings the notion of proximity and neighborhood among samples. While in signal processing this shift operator is mathematically represented by the (lower) shift matrix, in GSP the *graph shift operator* (GSO) depends on the underlying network domain.

The notion of regularity in time and in space, which are two very well structured domains, is reflected in the definition of their frequency domain. Specifically, a signal in these domains can be decomposed into elementary building blocks (such as sine waves) which endow a physical interpretation with a well understood meaning of variability. In a less structured domain modeled by a graph, this definition is not tight and multiple interpretations are possible. Nonetheless, a convolution in one of these *primal* domains can be described as a pointwise multiplication in the corresponding frequency (or *dual*) domains.

Motivated by a modern line of research [8] attempting to model the support of the frequency domain with a so called dual graph[1], in this work we introduce a new convolution theorem which generalizes the existing (graph) convolution theorem and the one related to time-varying filters. In particular, by relying on the notion of node-varying graph filters (NV-GFs) [6], we show how a NV-GF in one domain can be expressed as a NV-GF in the other domain, remaining consistent with the already known graph convolution theorem. We then discuss some of its implications in terms of non-stationary graph signals and delineate a possible dual graph learning method which will be object of a future study.

E-mails: {a.natali; g.j.t.leus}@tudelft.nl

[1]Not to be confused with the dual graph notion in graph theory, as the graph which has a vertex for each face of the original graph.

II. Preliminaries

Graphs and Graph Signals. We consider data residing on a non Euclidean domain, which we formally model by a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{S})$, where $\mathcal{V} = \{1, \ldots, N\}$ is the set of nodes (or vertices), $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of edges, and $\mathbf{S}$ is a symmetric $N \times N$ matrix that represents the graph structure. The matrix $\mathbf{S}$ is called the graph shift operator (GSO), since it plays a role akin to the shift (delay) operator in classical signal processing. Specifically, its entries $[\mathbf{S}]_{ij}$ for $i \neq j$ are different from zero only if nodes i and j are connected by an edge; typical examples of such a matrix are the (weighted) adjacency matrix $\mathbf{W}$ [9] and the graph Laplacian $\mathbf{L}$ [5].

In this manuscript, for the sake of generality, we consider the shift operator $\mathbf{S}$ to be a normal matrix with eigenvalue decomposition (EVD) written as $\mathbf{S} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^H$, with $\mathbf{V}$ a unitary matrix collecting the eigenvectors and $\mathbf{\Lambda}$ a diagonal matrix collecting the eigenvalues of $\mathbf{S}$. A fundamental assumption in GSP is that the matrix $\mathbf{V}$ provides a basis for expressing signals living on $\mathbf{S}$, and with favourable DFT-like properties providing a notion of frequency similar to the ones in classical signal processing. For this reason, the matrix $\mathbf{V}^H$ is often referred to as the graph Fourier transform (GFT) and the projection of $\mathbf{x}$ into this basis, i.e., $\hat{\mathbf{x}} = \mathbf{V}^H\mathbf{x}$ as the GFT signal.

Filtering on Graphs. Given a graph $\mathbf{S}$, a classical graph filter (C-GF) of order $L-1$ is the matrix polynomial:

$$\mathbf{H}(\mathbf{p}, \mathbf{S}) = \sum_{l=0}^{L-1} p_l \mathbf{S}^l, \tag{1}$$

where $\mathbf{p} = [p_0, \ldots, p_{L-1}]^\top$ collects the graph filter coefficients (taps). The application of the filter $\mathbf{H}(\mathbf{p}, \mathbf{S})$ on a signal $\mathbf{x}$ to obtain a new signal $\mathbf{y}$, i.e., $\mathbf{y} = \mathbf{H}(\mathbf{p}, \mathbf{S})\mathbf{x}$, is often referred to as graph filtering or *graph convolution*, as it respects the scale-sum-shift principle of convolution. With a few simple calculations, it is easy to show that in the (graph) frequency domain, a graph convolution is expressed as a pointwise multiplication; this is the (graph) convolution theorem, which can be expressed as follows:

$$\boxed{\mathbf{y} = \sum_{l=0}^{L-1} p_l \mathbf{S}^l \mathbf{x} \qquad \hat{\mathbf{y}} = \sum_{l=0}^{L-1} p_l \mathbf{\Lambda}^l \hat{\mathbf{x}}} \tag{2}$$

with $\hat{\mathbf{y}} = \mathbf{V}^H\mathbf{y}$ the GFT of $\mathbf{y}$.

In this paper, our focus is on a more flexible and expressive version of (1), as we consider the node-variant graph filter [6], which allows a *per-node* weighting scheme of each shifted version of the input signal. Due to its relevance in this work, we distinguish among two flavours of a NV-GF, henceforth referred to as *type-I* and *type-II*, defined, for a given a graph $\mathbf{S}$ and fixed order $L-1$, respectively as:

$$\mathbf{H}_I(\mathbf{P},\mathbf{S}) = \sum_{l=0}^{L-1} \mathrm{Diag}(\mathbf{p}_l)\mathbf{S}^l, \tag{3}$$

$$\mathbf{H}_{II}(\mathbf{P},\mathbf{S}) = \sum_{l=0}^{L-1} \mathbf{S}^l\,\mathrm{Diag}(\mathbf{p}_l), \tag{4}$$

where $\mathbf{P}$ is the $N\times L$ matrix of coefficients $\mathbf{P} = [\mathbf{p}_0,\dots,\mathbf{p}_{L-1}]$ and $\mathbf{p}_l := [p_{l1},\dots,p_{lN}]^\top$ is the l-hop filter tap vector. As a short-hand notation, we will use $\mathbf{H}_I$ and $\mathbf{H}_{II}$ to refer to the NV-GF in (3) and (4), respectively; when convenient for clarity of exposition, we will explicitly write $\mathbf{H}_I(\mathbf{P},\mathbf{S})$ or $\mathbf{H}_{II}(\mathbf{P},\mathbf{S})$ concordantly. The application of a NV-GF on a signal $\mathbf{x}$ to obtain a new signal $\mathbf{y}$ will be referred to as node-variant graph convolution. From a theoretical point of view, both NV-GF types have the same expressive behavior, yet the order of shifting and weighing is reversed. Specifically, in type-I, each node performs a linear combination of the (shifted) signal value of neighboring nodes, where the weights of the linear combination are neighbor-specific; in type-II, each node performs a linear combination of the (shifted) signal value of neighboring nodes, which have been already scaled by such nodes. Nonetheless, both can be implemented with the same complexity and in a distributed manner [6].

Dual Graph. The support of the GFT signal $\hat{\mathbf{x}}$ is usually assumed to be described by the eigenvalues $\boldsymbol{\lambda} := \mathrm{diag}(\boldsymbol{\Lambda})$ of $\mathbf{S}$, which correspond to a discretization/sampling of a continuous domain, either the real line $\mathbb{R}$ or the complex plane. This is consistent with the discrete signal processing notion of frequency domain: when $\mathbf{S}$ represents a cycle graph (or more generally, any circulant graph), possibly capturing the time domain, its eigenvector matrix $\mathbf{V}$ coincides with the discrete Fourier Transform (DFT) matrix, and its eigenvalues $\boldsymbol{\lambda}$ with the complex frequencies on the unit circle, i.e., $\boldsymbol{\lambda} = [1\ e^{-j2\pi/N}\dots e^{-j2\pi(N-1)/N}]$. However, a modern line of research attempts to model the (graph) frequency domain with a graph [8]. The motivation behind this line of research relies on the fact that classical signal processing tasks usually performed in the frequency domain, such as frequency-shifting, do not have their counterpart in GSP. Moreover, since a graph signal resides on a graph, it would be appealing to have also its Fourier counterpart to reside on a graph. This leads to the notion of a dual graph $\mathbf{S}_f$, which represents the support for the frequency (GFT) signal $\hat{\mathbf{x}}$. Since $\mathbf{x} = \mathbf{V}\hat{\mathbf{x}}$, i.e., $\mathbf{x}$ is expressed as a linear combination of the eigenvectors of the graph on which it resides, we expect our frequency graph signal $\hat{\mathbf{x}}$, residing on a dual GSO $\mathbf{S}_f = \mathbf{V}_f\boldsymbol{\Lambda}_f\mathbf{V}_f^H$, to be expressed as the linear combination of its eigenvectors $\mathbf{V}_f$, i.e., $\hat{\mathbf{x}} = \mathbf{V}_f\mathbf{x}$. From the definition of GFT it follows $\mathbf{V}_f = \mathbf{V}^H$. Thus, the primal graph provides *spectral templates* for the frequency domain, i.e., the eigenvectors $\mathbf{V}_f$ for the dual graph $\mathbf{S}_f$ are known by knowing those of $\mathbf{S}$. The only unknown is then the eigenvalue matrix $\boldsymbol{\Lambda}_f := \mathrm{Diag}(\boldsymbol{\lambda}_f)$, which can be found, for instance, with an axiomatic or an optimization approach [8].

III. An Encompassing Convolution Theorem

In this section we generalize the graph convolution theorem [cf. (2)] and we show how a limited order NV-GF in the primal domain can be expressed as a limited order NV-GF in the dual domain through an appropriate parametrization of the filter coefficients. This leads to a generalization of the well known convolution theorem for a C-GF and the one related to time-varying filters. The following theorem formally states this.

Theorem 1 (Node-variant convolution theorem)**.** *Consider a type-I NV-GF $\mathbf{H}_I$ defined over the graph $\mathbf{S}$ with filter taps $\{\mathbf{p}_l\}_{l=0}^{L-1}$, i.e., $\mathbf{H}_I(\mathbf{P},\mathbf{S})$, and assume that a dual graph $\mathbf{S}_f$ describing the dual domain is given. Assume also that each filter tap vector $\mathbf{p}_l$ can be expressed as a polynomial of order $K-1$ in the dual graph frequencies $\boldsymbol{\lambda}_f$. Then, there exists a set of coefficients $\{\hat{\mathbf{p}}_k\}_{k=0}^{K-1}$ for which the type-I NV-GF $\mathbf{H}_I(\mathbf{P},\mathbf{S})$ in the primal domain corresponds to a type-II NV-GF $\mathbf{H}_{II}$ on the dual graph $\mathbf{S}_f$ with filter taps $\{\hat{\mathbf{p}}_k\}_{k=0}^{K-1}$, i.e., $\mathbf{H}_{II}(\hat{\mathbf{P}},\mathbf{S}_f)$.*

Proof. By multiplying both sides of (3) with the GFT matrix $\mathbf{V}^H$, we have:

$$\hat{\mathbf{y}} = \mathbf{V}^H\sum_{l=0}^{L-1}\mathrm{Diag}(\mathbf{p}_l)\mathbf{S}^l\mathbf{x} = \mathbf{V}^H\sum_{l=0}^{L-1}\mathrm{Diag}(\mathbf{p}_l)\mathbf{V}\boldsymbol{\Lambda}^l\hat{\mathbf{x}}. \tag{5}$$

Next, inspired by time-varying communication systems, where the time-varying filter taps are smooth over time and/or expressed through a basis expansion model (BEM) [10], we introduce a similar construction for the NV filter coefficients $\{\mathbf{p}_l\}_{l=0}^{L-1}$. In particular, we express each $\mathbf{p}_l$ through powers of the dual eigenvalues $\boldsymbol{\lambda}_f$, representing our basis expansion; that is:

$$\mathbf{p}_l = \sum_{k=0}^{K-1} c_{lk}\boldsymbol{\lambda}_f^k = \boldsymbol{\Psi}_f\mathbf{c}_l \tag{6}$$

with $\boldsymbol{\Psi}_f$ the Vandermonde matrix $\boldsymbol{\Psi}_f := [\mathbf{1}\ \boldsymbol{\lambda}_f\dots\boldsymbol{\lambda}_f^{K-1}]$ and $\mathbf{c}_l := [c_{l0},\dots,c_{l(K-1)}]^\top$ the expansion coefficients for the l-th primal filter tap vector $\mathbf{p}_l$. With this choice, substituting (6) in (5), we have:

$$\begin{aligned}\hat{\mathbf{y}} &= \mathbf{V}^H\sum_{l=0}^{L-1}\mathrm{Diag}(\sum_{k=0}^{K-1}c_{lk}\boldsymbol{\lambda}_f^k)\mathbf{V}\boldsymbol{\Lambda}^l\hat{\mathbf{x}}\\ &= \sum_{k=0}^{K-1}\mathbf{V}^H\,\mathrm{Diag}(\boldsymbol{\lambda}_f^k)\mathbf{V}\,\mathrm{Diag}(\sum_{l=0}^{L-1}c_{lk}\boldsymbol{\lambda}^l)\hat{\mathbf{x}}\\ &= \sum_{k=0}^{K-1}\mathbf{S}_f^k\,\mathrm{Diag}(\hat{\mathbf{p}}_k)\hat{\mathbf{x}}\end{aligned} \tag{7}$$

where $\hat{\mathbf{p}}_k := \sum_{l=0}^{L-1}c_{lk}\boldsymbol{\lambda}^l = \boldsymbol{\Psi}\mathbf{c}^{(k)}$ is the kth hop filter tap vector on the *dual* graph, $\boldsymbol{\Psi} := [\mathbf{1}\ \boldsymbol{\lambda}\dots\boldsymbol{\lambda}^{L-1}]$ is the Vandermonde matrix of primal eigenvalues, and $\mathbf{c}^{(k)} := [c_{0k},\dots,c_{(L-1)k}]^\top$ are the expansion coefficients for the k-th dual filter taps vector $\hat{\mathbf{p}}_k$. We denote as $\mathbf{H}_{II} = \sum_{k=0}^{K-1}\mathbf{S}_f^k\mathrm{diag}(\hat{\mathbf{p}}_k)$ a NV-GF on the dual graph and, whenever the dependency on the dual coefficients and shift operator is necessary, we use $\mathbf{H}_{II}(\hat{\mathbf{P}},\mathbf{S}_f)$, where $\hat{\mathbf{P}}$ is the $N\times K$ matrix of coefficients $\hat{\mathbf{P}} = [\hat{\mathbf{p}}_0,\dots,\hat{\mathbf{p}}_{K-1}]$. $\square$

With this notation in place, we can now delineate a general convolution theorem which relies on node-variant graph filtering, also pictorially described in Fig. 1, as follows:

$$\mathbf{y} = \sum_{l=0}^{L-1} \mathrm{Diag}(\mathbf{p}_l)\mathbf{S}^l\mathbf{x} \qquad \hat{\mathbf{y}} = \sum_{k=0}^{K-1} \mathbf{S}_f^k \,\mathrm{Diag}(\hat{\mathbf{p}}_k)\hat{\mathbf{x}} \tag{8}$$

$$\mathbf{p}_l = \mathbf{\Psi}_f\mathbf{c}_l \qquad \hat{\mathbf{p}}_k = \mathbf{\Psi}\mathbf{c}^{(k)} \tag{9}$$

The connection between the primal and the dual node-variant graph filters defined in (8) is given by the $K \times L$ expansion coefficients conveniently stored in the matrix $\mathbf{C} = [\mathbf{c}_0, \ldots, \mathbf{c}_{L-1}] = [\mathbf{c}^{(0)}, \ldots, \mathbf{c}^{(K-1)}]^\top$. This enables also to concisely express the node-variant coefficients in the primal and dual domain as $\mathbf{P} = \mathbf{\Psi}_f\mathbf{C}$ and $\hat{\mathbf{P}} = \mathbf{\Psi}\mathbf{C}^\top$, respectively.

Remark 1. As pointed out in [6], the filter $\mathbf{H}_{II}$ defined by (7) is by all means a node-variant GF and can be implemented in a distributed and sequential fashion. Precisely, the equation on the right in (8) can be implemented through the following recursion $\mathbf{t}_k = \mathbf{S}_f\mathbf{t}_{k-1} + \mathrm{Diag}(\hat{\mathbf{p}}_{K-k})\hat{\mathbf{x}}$ for $k > 0$ and initializing with $\mathbf{t}_0 = \mathbf{0}$; finally $\hat{\mathbf{y}} = \mathbf{t}_K$.

Corollary 1. Given a graph signal $\mathbf{x}$, the application of a node-variant graph filter $\mathbf{H}_I(\mathbf{P}, \mathbf{S})$ in the primal domain followed by the GFT $\mathbf{V}^H$ is equivalent to the application of the GFT followed by a node-variant graph filter $\mathbf{H}_{II}(\hat{\mathbf{P}}, \mathbf{S}_f)$ in the dual domain. In other words, it holds (see also Fig. 1):

$$\mathbf{V}^H\mathbf{H}_I(\mathbf{P}, \mathbf{S}) = \mathbf{H}_{II}(\hat{\mathbf{P}}, \mathbf{S}_f)\mathbf{V}^H. \tag{10}$$

In classical signal processing, the frequency representation of windowing in the time domain is the convolution between the spectra of the signal and the window. Because a node-variant convolution of order $L-1$ is nothing else than the application of L windows on shifted versions of the input graph signal $\mathbf{x}$, a similar result can be derived in the graph setting; the following corollary expresses this.

Corollary 2. Given an input graph signal $\mathbf{x}$ and a type-I NV-GF $\mathbf{H}_I(\mathbf{P}, \mathbf{S})$, with each $\{\mathbf{p}_l\}_{l=0}^{L-1}$ parametrized as in (9), a node-variant graph convolution of order $L-1$ in one domain is equivalent to the sum of L classical graph convolutions in the other domain, each one with input a (modulated) version of $\hat{\mathbf{x}}$; that is:

$$\hat{\mathbf{y}} = \sum_{l=0}^{L-1} \mathbf{H}(\mathbf{c}_l, \mathbf{S}_f)(\boldsymbol{\lambda}^l \odot \hat{\mathbf{x}}) \tag{11}$$

$$= \mathbf{H}(\mathbf{c}_0, \mathbf{S}_f)\hat{\mathbf{x}} + \ldots + \mathbf{H}(\mathbf{c}_{L-1}, \mathbf{S}_f)(\boldsymbol{\lambda}^{L-1} \odot \hat{\mathbf{x}}) \tag{12}$$

Proof. From the first equality of (7), we have:

$$\begin{aligned}\hat{\mathbf{y}} &= \sum_{l=0}^{L-1}\sum_{k=0}^{K-1} c_{lk}\mathbf{V}^H \,\mathrm{Diag}(\boldsymbol{\lambda}_f^k)\mathbf{V}\mathbf{\Lambda}^l\tilde{\mathbf{x}} \\ &= \sum_{l=0}^{L-1}\left(\sum_{k=0}^{K-1} c_{lk}\mathbf{S}_f^k\right)\mathbf{\Lambda}^l\hat{\mathbf{x}} \\ &= \sum_{l=0}^{L-1} \mathbf{H}(\mathbf{c}_l, \mathbf{S}_f)(\boldsymbol{\lambda}^l \odot \hat{\mathbf{x}}) \end{aligned} \tag{13}$$

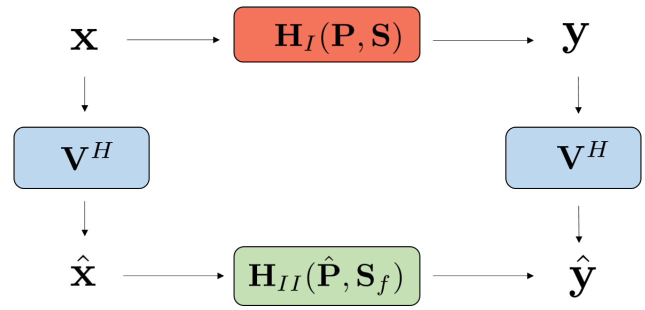

Figure 1. General convolution theorem. A node-variant graph convolution in the primal domain followed by a GFT is equivalent to a GFT followed by a node-variant graph filtering in the dual domain.

Notice how the filter coefficients in (13) are the expansion coefficients $\mathbf{c}_l$ associated to the primal filter coefficients $\mathbf{p}_l$. □

A. Consistency with the graph convolution theorem

Because a C-GF is a NV-GF with constant filter taps, we expect that our introduced theory encompasses the existing one. Indeed, we can formally show that the graph convolution theorem (2) falls within the introduced theory. To see this, consider $\mathbf{p}_l = p_l\mathbf{1}\ \forall\ l$, i.e., the case in which each vector of filter taps $\mathbf{p}_l$ is constant over the nodes, thus corresponding to (1). From (9), we then have that $\mathbf{c}_l$ necessarily needs to be $\mathbf{c}_l = [p_l, \mathbf{0}^\top]^\top$, and overall:

$$[p_1\mathbf{1}, \ldots, p_{L-1}\mathbf{1}] = \begin{bmatrix} 1 & \cdots & \lambda_{0,f}^{K-1} \\ \vdots & \ddots & \vdots \\ 1 & \cdots & \lambda_{f,N-1}^{(K-1)} \end{bmatrix} \begin{bmatrix} p_1 & \cdots & p_{L-1} \\ 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \tag{14}$$

meaning that only the first row $\mathbf{c}^{(0)}$ of $\mathbf{C}$ is different from zero. In particular, from the right equation in (9) this implies that:

$$\hat{\mathbf{y}} = \mathbf{S}_f^0 \mathrm{diag}(\hat{\mathbf{p}}_0)\hat{\mathbf{x}} = \mathrm{diag}(\mathbf{\Psi}\mathbf{p})\hat{\mathbf{x}}. \tag{15}$$

This shows how the proposed theory fits within the principle that a classical graph convolution (node-invariant GF) is a pointwise multiplication in the frequency domain.

B. Relationship with time-varying channel propagation

The proposed theory generalizes, to the graph setting, concepts which are familiar in the context of time-varying channel propagation [11], arising for instance in mobile communication scenarios. In that case, the received signal y at time n, i.e., $y[n]$[2], is modeled as:

$$y[n] = \sum_{l=0}^{L-1} p[n, l]x[n-l], \tag{16}$$

where $p[n, l]$ denotes the channel impulse response of the l-th path at the n-th time instant, and $x[n-l]$ is the transmitted signal at the $(n-l)$-th time instant. The gains associated to the

[2]We use square brackets to indicate that the argument is a time index and not a graph node.

different paths are assumed to be time-varying and approximated by a basis expansion model [10]; specifically:

$$\mathbf{p}_l = \sum_{k=0}^{K-1} c_{lk}\mathbf{b}_k, \tag{17}$$

where $\mathbf{p}_l = [p[0,l],\ldots,p[N-1,l]]^\top$ stores the evolution of the filter impulse response over the N time instants, $\mathbf{b}_k \in \mathbb{R}^N$ is the k-th basis function, and c_{lk} is the coefficient associated to the l-th path and the k-th basis function. This alleviates the effort of having to deal with NL channel coefficients (usually a very high number), and converts the model into a simpler one with only LK BEM coefficients.

It is easy to show that we can write (16) in matrix-vector form, by taking into account (17), as:

$$\mathbf{y} = \sum_{k=0}^{K-1} \mathrm{Diag}(\mathbf{b}_k)\left(\sum_{l=0}^{L-1} c_{lk}\mathbf{D}^l\right)\mathbf{x} \tag{18}$$

where $\mathbf{x} = [x[0],\ldots,x[N-1]]^\top$ and $\mathbf{D}$ is the $N\times N$ lower shift matrix; notice how the matrix $\sum_{l=0}^{L-1} c_{lk}\mathbf{D}^l$ implements a standard convolutional filter in time and observe its similarities with the left equation in (8). Next, denote with $\mathbf{F} \in \mathbb{C}^{N\times N}$ the normalized DFT matrix, and with $\mathbf{f}_k$ its kth column; the classical complex exponential BEM uses the Fourier basis as the basis functions in (17), i.e., $\mathbf{b}_k = \mathbf{f}_k^*$, with * the complex conjugate. With these choices, (18) can be expressed in the frequency domain as:

$$\begin{aligned}\hat{\mathbf{y}} = \mathbf{F}\mathbf{y} &= \sum_{k=0}^{K-1} \mathbf{F}\,\mathrm{Diag}(\mathbf{f}_k^*) \sum_{l=0}^{L-1} c_{lk}\mathbf{F}^H\mathbf{D}^l\mathbf{F}\mathbf{x} \\ &= \sum_{l=0}^{L-1}\left(\sum_{k=0}^{K-1} c_{lk}\mathbf{D}^k\right)\mathrm{Diag}(\mathbf{f}_l)\hat{\mathbf{x}}. \end{aligned}\tag{19}$$

While in (18) the matrix $\mathbf{D}$ shifts in the time domain, in (19) shifts in the frequency domain; however, such shift matrix is the same in both domains. This is different from the graph counterpart in (7), where the two shifts matrices might be different. All in all (19) is the time domain counterpart of (7), by choosing the primal eigenvector matrix $\mathbf{V}$ to be $\mathbf{V} = \mathbf{F}^H$ and the basis functions $\mathbf{b}_k$ to be $\mathbf{b}_k = \mathbf{f}_k^*$.

IV. Outlook and Future research

In the previous section, we laid down a duality theory which is consistent with the existing body of knowledge. In this section we delineate possible future research directions.

Non-Stationarity. The dual graph concept is closely related to the one of (non-)stationarity. From [12] we know that a process $\mathbf{y}$ is said to be weakly stationary on a GSO $\mathbf{S}$ if the covariance matrix $\mathbf{\Sigma}_y := \mathbb{E}[\mathbf{y}\mathbf{y}^H]$ commutes with $\mathbf{S}$ or, equivalently, if $\mathbf{y}$ can be written as the output of a C-GF $\mathbf{H}$ [cf. (1)] when excited with a white input $\mathbf{x}$, i.e., $\mathbf{y} = \mathbf{H}\mathbf{x}$. However, if $\mathbf{H} = \mathbf{H}_I$, i.e., it is a type-I NV-GF as in (3), and the covariance matrix:

$$\mathbf{\Sigma}_y = \mathbb{E}[(\mathbf{H}_I\mathbf{x})(\mathbf{H}_I\mathbf{x})^H] = \mathbf{H}_I\mathbb{E}[\mathbf{x}\mathbf{x}^H]\mathbf{H}_I^H = \mathbf{H}_I\mathbf{H}_I^H, \tag{20}$$

does not commute in general with the GSO $\mathbf{S}$. This implies that the GFT signal $\hat{\mathbf{y}}$ is correlated, and its covariance matrix

$$\mathbf{\Sigma}_{\hat{y}} = \mathbb{E}[\hat{\mathbf{y}}\hat{\mathbf{y}}^H] = \mathbf{H}_{II}\mathbf{H}_{II}^H = \mathbf{V}_f\mathbf{\Sigma}_y\mathbf{V}_f^H \tag{21}$$

is not diagonal. These observations reveal that filtering a white input with a NV-GF gives us a non-stationary signal. The inverse problem is also true. If we expect our signal to be non-stationary, we should be able to find a NV-GF that transforms a white input into our non-stationary signal.

Dual Graph Learning. Another interesting line of research entails learning in a data driven manner the dual graph $\mathbf{S}_f$ (which is tantamount to learning its eigenvalues $\boldsymbol{\lambda}_f$) in such a way that it is consistent with the theorem we proposed . Specifically, by assuming that observed graph signals $\mathbf{Y} = [\mathbf{y}_1,\ldots,\mathbf{y}_T]$ can be modeled as the output of a NV-GF $\mathbf{H}_I(\mathbf{P},\mathbf{S})$ when excited with some input data $\mathbf{X} = [\mathbf{x}_1,\ldots,\mathbf{x}_T]$, i.e., $\mathbf{Y} = \mathbf{H}_I(\mathbf{P},\mathbf{S})\mathbf{X}$, the following two-step approach can be put forth:

1) first, we learn the graph filter coefficient matrix $\mathbf{P}$ in a data driven manner, for instance as the solution of the optimization problem:

$$\hat{\mathbf{P}} := \underset{\mathbf{P}}{\mathrm{argmin}} \quad \|\mathbf{Y} - \mathbf{H}_I(\mathbf{P},\mathbf{S})\mathbf{X}\|_F^2, \tag{22}$$

which can be solved with standard convex optimization machinery;

2) then, we find the dual eigenvalues $\boldsymbol{\lambda}_f$ by exploiting (9), i.e., fit the model $\hat{\mathbf{P}} \approx \mathbf{\Psi}_f\mathbf{C}$, which is a specific structured matrix factorization with a Vandermonde factor, where only its second column is needed. A possible way to solve this problem is by following the recently proposed approach in [13], which relies on a subspace method. In there, it is shown how each solution of the problem is identifiable up to a shift and scale of the unknown vector (here representing the graph eigenvalues), thus maintaining the same topological structure of the original graph (removing the self loops caused by the shift).

V. Conclusions

In this manuscript, by relying on the concepts of node-varying graph filters, which allow a node-specific filtering scheme, and the dual graph, which allows to represent the support of the frequency domain as a graph, we generalize the standard (graph) convolution theorem, as well as the convolution theorem for time-varying filters. Specifically, we show how a node-varying graph filter in one domain can be expressed as another node-varying graph filter in the other domain. We show how such theory is consistent with classical graph filtering in graph signal processing as well as with the time-varying channel propagation literature, where each path is modeled with a basis expansion model. Finally, we delineate research directions worth to investigate, namely the role of node-varying graph filters and the dual graph to model non-stationary graph signals, as well as data driven procedures to learn the (eigenvalues of the) dual graph.

References

[1] Alan V Oppenheim, John R Buck, and Ronald W Schafer, *Discrete-time signal processing. Vol. 2*, Upper Saddle River, NJ: Prentice Hall, 2001.

[2] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[3] Helmut Lütkepohl, *Introduction to multiple time series analysis*, Springer Science & Business Media, 2013.

[4] Yann LeCun, Yoshua Bengio, et al., "Convolutional networks for images, speech, and time series," .

[5] David I Shuman, Sunil K Narang, Pascal Frossard, Antonio Ortega, and Pierre Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE signal processing magazine*, vol. 30, no. 3, pp. 83–98, 2013.
[6] Santiago Segarra, Antonio G Marques, and Alejandro Ribeiro, "Optimal graph-filter design and applications to distributed linear network operators," *IEEE Transactions on Signal Processing*, vol. 65, no. 15, pp. 4117–4131, 2017.
[7] David I Shuman, Pierre Vandergheynst, Daniel Kressner, and Pascal Frossard, "Distributed signal processing via chebyshev polynomial approximation," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 4, no. 4, pp. 736–751, 2018.
[8] Geert Leus, Santiago Segarra, Alejandro Ribeiro, and Antonio G Marques, "The dual graph shift operator: Identifying the support of the frequency domain," *Journal of Fourier Analysis and Applications*, vol. 27, no. 3, pp. 1–20, 2021.
[9] Aliaksei Sandryhaila and José MF Moura, "Discrete signal processing on graphs," *IEEE transactions on signal processing*, vol. 61, no. 7, pp. 1644–1656, 2013.
[10] Georgios B Giannakis and Cihan Tepedelenlioglu, "Basis expansion models and diversity techniques for blind identification and equalization of time-varying channels," *Proceedings of the IEEE*, vol. 86, no. 10, pp. 1969–1986, 1998.
[11] Zijian Tang, Rocco Claudio Cannizzaro, Geert Leus, and Paolo Banelli, "Pilot-assisted time-varying channel estimation for ofdm systems," *IEEE Transactions on Signal Processing*, vol. 55, no. 5, pp. 2226–2238, 2007.
[12] Antonio G Marques, Santiago Segarra, Geert Leus, and Alejandro Ribeiro, "Stationary graph processes and spectral estimation," *IEEE Transactions on Signal Processing*, vol. 65, no. 22, pp. 5911–5926, 2017.
[13] Alberto Natali and Geert Leus, "Blind polynomial regression," *arXiv preprint arXiv:2210.11874*, 2022.