

Graph Signal Processing in a Nutshell

Guilherme B. Ribeiro, *Student Member, IEEE*, Juliano B. Lima, *Member, IEEE*

Abstract—The framework of graph signal processing was conceived in the last decade with the ambition of generalizing the tools from classical digital signal processing to the case in which the signal is defined over an irregular structure modelled by a graph. Instead of *discrete time* — what one would call a *regular 1-D domain*, in which a signal sample is adjacent to only two neighbors and for any pair of contiguous samples the distance is the same — the signals here are defined over *graphs* and, therefore, the distance and relations between adjacent samples vary along the signal. For instance, one may consider the temperature signal defined from the data of a sensor mesh network. When creating the tools in such a scenario, many challenges arise even with basic concepts of the classical theory. In this paper, the core ideas of graph signal processing are presented, focusing on the two main frameworks developed along the years, and a couple of examples and applications are shown. We conclude drawing attention to a few of the many open opportunities for further studies in the field.

Index Terms—Graph signal processing, structured data, graph theory, linear algebra.

I. INTRODUCTION

Multivariate data defined over networks are nowadays ubiquitous, being constantly generated, stored and processed in the most diverse systems in engineering and technology. Measurements in a set of IoT sensors and mobile devices [1]–[4], number of citations in a scientific collaboration network or social media relations (*collaboration graph*, or *social graph*) [5] and interactions between individuals in an ecosystem (*ecological networks*) [6] are some examples of situations in which the acquired data are intimately related to the topology of the network over which they are defined.

Such multivariate network-like systems are not only present in various applications, but are also systematically growing in number, as sensors become cheaper and smaller and concepts such as cloud storage/computing and Big Data consolidate, as indicated by the 2011 report from McKinsey Global Institute [7]. This document also states that the information acquired from the adequate processing of such massive networked data is a fundamental requisite for the companies to thrive from now on.

Still another motivation that feeds the urge to study processing techniques for data defined over network-like domains, for example, is the growth of research on *smart cities*, which takes advantage of the considerable information (that are or are yet

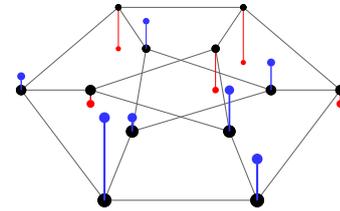


Figure 1: Example of signal defined over a graph. The height of the vertical bars indicate the value of the signal samples, which are indexed by the graph vertices. The graph edges capture similarity relations between samples. How could one define spectral analysis and processing techniques in such a signal domain?

to be) generated in cities to provide (or improve the) solutions for many urban problems [8].

All these examples share an important characteristic: the structure over which the data is defined may be modelled by a graph [9], to which vertices are assigned the variables of interest, as depicted in Fig. 1. That is the context in which the field of *graph signal processing* (GSP) was developed in the last decade, a theoretical framework aiming to generalize the classical signal processing methods and concepts to scenarios in which the signal is no more defined over a regular domain, but sits on a generally irregular structure, an arbitrary graph. The research is still very active and numerous contributions have been made, but two distinct frameworks consistently grew throughout the years and have been established as default mindsets when dealing with graph signals. The first one is based on algebraic signal processing and uses the graph adjacency matrix as elementary block. This approach imposes no restrictions regarding the graph being directed or undirected, and the edge weights are allowed to be negative or complex numbers [10]. The second framework draws ideas from spectral graph theory and analyzes signals defined only over undirected graphs with non-negative real edge weights, using the graph Laplacian matrix to build a basis for the signal space [11]. Both approaches have particular characteristics which make each more appropriate than the other for some applications. In this paper, we intend to present an overview of the basic aspects concerning each framework and provide the reader with a good understanding of their basic concepts and tools.

G. B. Ribeiro and J. B. Lima are associated to the Department of Electronics and Systems at the Federal University of Pernambuco (UFPE). E-mails: guilherme.boaviagem@gmail.com, juliano_bandeira@ieee.org.

The authors wish to thank the Brazilian research funding agencies CNPq (132103/2016-8, 56744/2014-2, 307686/2014-0) and CAPES, for the financial support to this work. Great appreciation also to the support given by the Electrical Engineering Postgraduate Program (PPGEE) at UFPE.

Digital Object Identifier: 10.14209/jcis.2018.22

A. The challenge of graph-like domains

One of the reasons why GSP has been such a fertile field, allowing the birth of so many different problems and ideas, is that the definition of a signal over a graph leads to a series of obstacles even with fundamental concepts of signal

processing. Let us use the simple but elucidating example given by Shuman *et al.* [11], and consider the unit shift to the right of a discrete-time signal $x[n]$, which is done in digital signal processing (DSP) by the simple variable substitution $x[n-1]$. Good enough, but what does it mean to right-shift the signal in Fig. 1, for example? Obviously the sense of right and left are meaningless for general graphs. On this problem, Shuman *et al.* argue that a naïve choice would be to label the N graph vertices from v_0 to v_{N-1} , so that the sample $x[n]$ is assigned to vertex v_n , for doing so would allow to define the shifted signal as the result of assigning $x[n]$ to vertex $v_{(n-1) \bmod N}$. Such an option, however, is not adequate, for its repeatability depends always on the way the vertices are labelled [11]. This example illustrates how a concept in DSP as simple as signal translation may deserve a cautious study in GSP.

B. Structure of the paper

This paper is structured as follows. Section II establishes the introductory concepts of graph signal processing, presenting the graph theory terminology and the notion of signal defined over a graph. Sections III and IV present the two main frameworks of GSP, referred to hereinafter as GSP_A and GSP_L, respectively. Section V aims to provide the reader with a diverse set of practical examples and applications of GSP, so as to allow for a broader understanding of the field. The paper closes with highlights of opportunities for future work in the area, in Section VI.

II. PRINCIPLES AND DEFINITIONS

The field of graph signal processing draws basic concepts from the classical theories of digital signal processing and graph theory, aiming to provide a cohesive and useful framework to tackle the aforementioned challenges. In this section, some of the main definitions found in this field are presented.

A. Graph theory: a brief terminology

A graph is commonly defined as the ordered pair $(\mathcal{V}, \mathcal{E})$, in which the set \mathcal{V} contains the so called graph *vertices* and the set of *edges* \mathcal{E} is a subset of \mathcal{V}^2 [12]. We will usually indicate by $|\mathcal{V}| = N^1$ and $|\mathcal{E}| = E$ the number of vertices and edges of a graph, respectively. For our purposes it is convenient to represent a graph as the structure $\mathcal{G} = \{\mathcal{V}, \mathbf{A}\}$, endowed with the (weighted) adjacency matrix \mathbf{A} which captures the vertex-to-vertex relations: if $A_{i,j} \neq 0$, then there is an edge of weight $A_{i,j}$ from the vertex v_j to v_i . It is denoted by d_i^- the *indegree* of vertex v_i , consisting of the sum of weights of all incoming edges to vertex v_i . Likewise, the *outdegree* d_i^+ is the sum of weights of edges departing from v_i .

A graph is called *undirected* if and only if its adjacency matrix is symmetric, in which case it is defined the *degree* of vertex v_i as $d_i^- = d_i^+ = d_i$. In this case, a graph is said to be *d-regular* whenever all graph vertices have degree d . If \mathbf{A} is asymmetric, however, the respective graph is *directed*

¹The set operator $|\cdot|$ means the cardinality, or amount of elements, of the set.

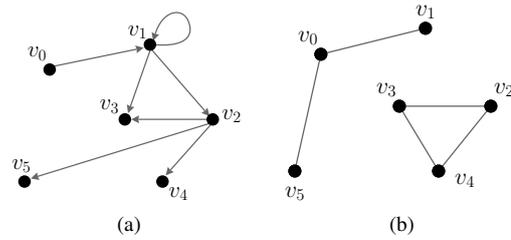


Figure 2: Examples of (a) directed and (b) undirected graphs, defined over the same vertex set.

and its pictorial representation depicts the edges as arrows, to account for the unidirectional relation between adjacent vertices. Examples of directed and undirected graphs are shown in Fig. 2.

The adjacency matrix is the building block for one of the two main frameworks of GSP, what will be covered soon, but another matrix of great importance, mainly in the branch of GSP originated from spectral graph theory, is the Laplacian matrix

$$\mathbf{L} = \mathbf{D} - \mathbf{A}, \quad (1)$$

with the *degree matrix* \mathbf{D} being a diagonal matrix with the degree d_i as its i -th entry. Depending on the context, \mathbf{D} may be taken as the indegree or outdegree matrix, although when the Laplacian matrix is used the graphs considered are more often undirected.

A *path* is a set of distinct edges (with the same orientation, if the graph is directed) linking distinct vertices. A *cycle* is a path with equal starting and end points, and if a graph has a cycle it is called *cyclic* (*acyclic*, otherwise). If the cycle has only one edge, it is called a *loop*. One refers to *multiple edges* whenever a single pair of vertices is connected by two or more edges. An undirected graph is called *simple* if it has no loops or multiple edges.

A graph is said to be *complete* if any two of its vertices are adjacent. Graph signal processing over such graphs may be extremely cumbersome, for the computational complexity of many of its techniques depends heavily on the number of graph edges. For most applications, it is desirable to have a small number of edges while keeping the graph *connected*, i. e. for any pair of vertices there exists a set of distinct edges (with the same orientation, if directed) connecting them without making a cycle.

A graph is said to be *unweighted* if all its edges have unit weight. A *subgraph* of \mathcal{G} is a graph $\mathcal{G}' = (\mathcal{V}', \mathbf{A}')$ with edge set \mathcal{E}' , in which $\mathcal{V}' \subset \mathcal{V}$ and $\mathcal{E}' \subset \mathcal{E}$. A *connected component* of \mathcal{G} is a connected subgraph $\mathcal{G}' = (\mathcal{V}', \mathbf{A}')$ in which any vertex in \mathcal{V}' is linked exclusively to another vertex also in \mathcal{V}' . This is illustrated by Fig. 2b, in which the graph has two connected components.

The *neighbourhood* of a vertex v_i is the set \mathcal{N}_i of all vertices adjacent to v_i . Sometimes it is useful as well to denote by $\mathcal{N}(i, K)$ the set of vertices connected to v_i through a path of length K or less. This notion is represented in Fig. 3.

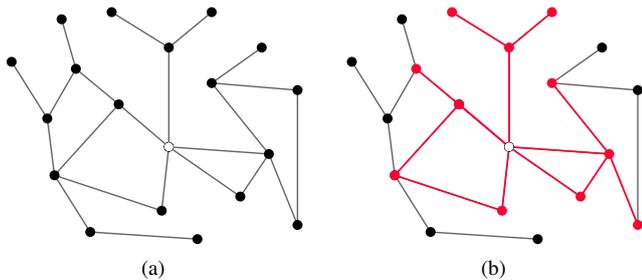


Figure 3: (a) A graph and (b) the set of vertices $\mathcal{N}(i, 2)$ shown in red, with v_i being depicted in white. The edges linking v_i to the elements in $\mathcal{N}(i, 2)$ are also highlighted in red.

The reader is encouraged to refer to this section whenever necessary. For a broader glossary with a solid introduction to graph theory, the authors recommend [13], [14].

B. Defining a graph signal

A signal \mathbf{s} defined over $\mathcal{G} = \{\mathcal{V}, \mathbf{A}\}$, with $|\mathcal{V}| = N$, is a discrete-domain function mapping the graph vertex set to a scalar set, usually the complex or real numbers,

$$s : \mathcal{V} \rightarrow \mathbb{C} \mid s(v_i) = s_i, \quad (2)$$

so that \mathbf{s} can be seen as a vector in \mathbb{C}^N indexed by the vertices of \mathcal{G} . Once the vertices $\mathcal{V} = \{v_1, \dots, v_N\}$ are clearly labelled, it is not ambiguous to represent the signal as the column vector $\mathbf{s} = (s_0 \ s_1 \ \dots \ s_{N-1})^T$, $s_i \in \mathbb{C}$, $0 \leq i \leq N-1$.

Fig. 4 provides examples of graph signal representations, in which the vertex labelling is omitted for the sake of simplicity, as it will be assumed that the signal sample s_i is assigned to vertex v_i . The signal values are indicated in two manners: either by writing down its numerical value next to the respective vertex, or by using a pseudocolor scale, the latter of which is the scheme adopted throughout this paper.

It is crucial to stress a certain graph which links GSP to the classical Discrete Signal Processing (DSP) theory: the directed ring graph, shown in Fig. 4a, which models the finite-length discrete-time domain. Its directed edges model the causality of time domain, whereas the feedback edge accounts for the boundary condition of periodicity imposed by the DFT analysis. Other signals that arise in practical applications have the respective graphs easily identified: the rectangular lattice in Fig. 4b, for example, models the digital image domain [15], and Fig. 4c shows an example of signal defined over a mesh network of sensors, with the edges weighted using the inverse of the euclidian distance, which arises in many scenarios such as IoT applications.

The spectral characteristics of a signal depend heavily on the domain over which it is defined, but one does not need to acknowledge this in the context of DSP, for in this case the domains are always regular and uniform². From the

²One could argue that, in the theory of *nonuniform sampling*, the signal is defined over an irregular domain, since the samples may be randomly spaced. Even in this case, however, the classical techniques still aim to *recover* the signal so as to represent it in its usual – and *uniform* – domain.

classical theory, the common understanding states that a signal has mostly low frequencies if adjacent samples have similar values, and high frequencies otherwise. When dealing with signals defined over graphs, it is clear that the adjacency relations depend on the graph topology, and therefore one may foresee that *the same signal may present different spectra when defined over different graphs*. This intuition is visually confirmed (and will soon be mathematically proved) in Fig. 5, which depicts a signal and its spectra when two different graphs are taken as a domain. The reader may notice that, in Fig. 5c, the samples with highest values are adjacent to the ones with small values, what causes bigger frequency components in this signal than in the one defined over the undirected ring graph in Fig. 5a.

C. Graph inference

Some contexts in which GSP is to be applied, to perform whatever signal processing technique is necessary, do not provide clear information on *how the underlying graph is structured*. For example, let us suppose the temperature data (or any other data in fact) of some Brazilian Northeastern cities will be treated using GSP. How is one supposed to weight the graph edges, and before this, how does one decide which vertices to connect? Is the graph shown in Fig. 4c the only option? Clearly it is not. Although generally the problem of graph inference is complex, this type of geography-based graph has an adequate method topology estimation.

The general idea is that, if there is a clear metric to evaluate the *expected* similarity between samples as a function of the available information regarding the respective vertices, then this metric may be used as the edge weight and a threshold is set so that any weight below this value causes the respective edge to be eliminated. In the case of vertices which have geodesic positions, the euclidian distance may be used as the metric because vertices that are closer together are *expected* to have similar signal samples, and therefore the adjacency matrix of the underlying graph may have entries given by

$$A_{ij} = \begin{cases} \exp\left(-\frac{\text{dist}^2(v_i, v_j)}{2\theta^2}\right) & \text{if } \text{dist}(v_i, v_j) < T \\ 0 & \text{otherwise,} \end{cases} \quad (3)$$

as used in [11]. The choice of the parameters T^3 and θ (standard deviation of the distribution), and of how to use the metric (in this case, inside a Gaussian distribution), are dictated by the application and by the analyst experience.

However, if there is an isolated vertex, far from the others, the use of (3) may lead to a compromise between keeping the graph connected and obtaining a sparse adjacency matrix, since imposing connectivity to the graph in this case implies increasing T , and therefore having many edges. To deal with this problem and still have a good representation of the underlying graph, one alternative is to connect a vertex to its K closest neighbours (setting K to an appropriate value,

³ T indicates a distance threshold above which we set the edge weight to zero, effectively leaving the vertices unlinked. This means that the distance between them is assumed to be too high for any significant interdependence to exist.

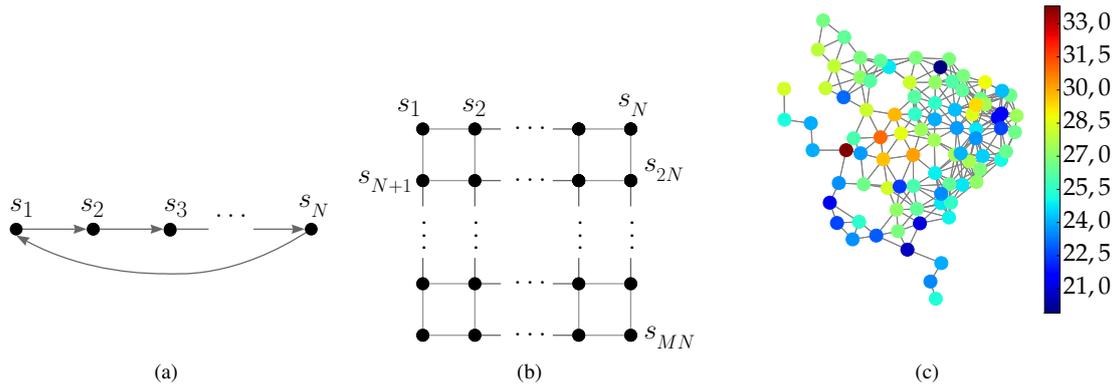


Figure 4: Examples of depictions of graph signals over (a) a directed ring graph, (b) an undirected regular grid graph and (c) a graph of cities from the Brazilian Northeastern region, over which was defined a signal of temperature measurements from February 1st of 2012, retrieved from the *Banco de Dados Meteorológicos para Ensino e Pesquisa* (BDMEP, freely translated as Meteorological Database for Teaching and Research), available at: <http://www.inmet.gov.br/porta1/index.php?r=bdmep/bdmep>.

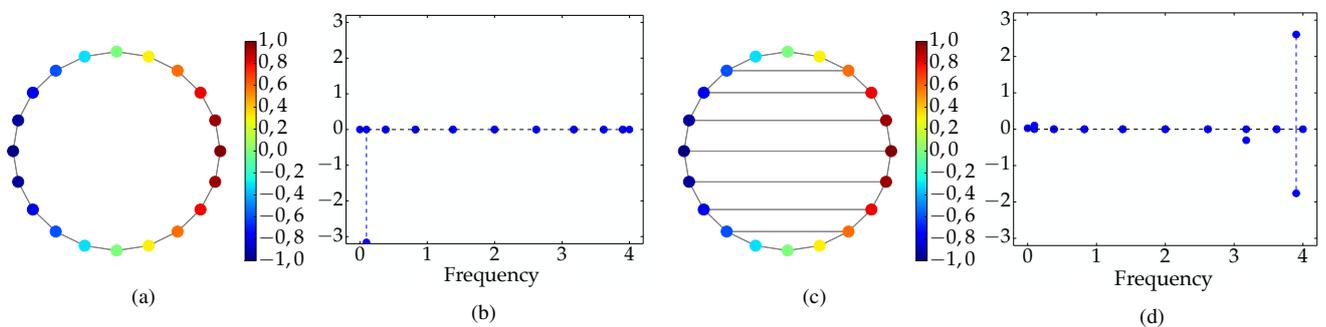


Figure 5: The same signal was defined over similar graphs, one of them being (a) the undirected ring graph. In (b) and (d) are depicted the Fourier spectra according to GSP_L (as defined in Subsection IV-A) of the signals in (a) and (c), respectively.

according to the context) and weight the edges using the Gaussian distribution in (3).

As previously discussed, these methods require an adequate metric to evaluate the expected similarity between samples in the graph vertex, but given the diverse areas in which graph signals may arise, estimating the topology of the underlying graph constitutes a challenge of its own [9], [16].

III. GSP_A : GRAPHS AND ALGEBRAIC SIGNAL PROCESSING

In 2008, Püschel and Moura published their *algebraic signal processing* (ASP) theory [17], [18], which expands DSP by moving to an algebraic point of view: each signal processing theory is studied as a triple $(\mathcal{A}, \mathcal{M}, \Phi)$ consisting of an algebra \mathcal{A} (a vector space endowed with multiplication between vectors), an \mathcal{A} -module \mathcal{M} (a vector space over the same base field as \mathcal{A} which admits left-multiplication by elements of \mathcal{A}) and a linear transformation Φ . \mathcal{A} is called the filter space, \mathcal{M} is the signal space and Φ is the Fourier transform (homomorphism over \mathcal{M}) associated to the structure.

When these authors drew inspiration from ASP to develop their GSP theory, the starting point was necessarily to find (better, to define) the unit shift operator of graph signals, the reason being that such an operator in ASP is the building block of the algebra \mathcal{A} (as, for example, the unit delay z^{-1} is

the building block for filters of discrete-time and finite-length signals $\mathcal{A} = \{\sum_{\ell=0}^{N-1} h_{\ell} z^{-\ell} | h_{\ell} \in \mathbb{C}\}$). To do so, the shift of discrete-time signals, defined over directed ring graphs, was investigated.

By inspection of the *adjacency matrix* of the directed ring graph,

$$\mathbf{C} = \begin{bmatrix} & & & 1 \\ 1 & & & \\ & \ddots & & \\ & & 1 & \end{bmatrix}, \quad (4)$$

it was noticed that the unit (circular) shift of discrete-time signals is precisely the left-multiplication by \mathbf{C} , for given a discrete-time signal $\mathbf{x} = (x_0 \ x_1 \ \dots \ x_{N-1})^T$,

$$\mathbf{C}\mathbf{x} = \begin{bmatrix} & & & 1 \\ 1 & & & \\ & \ddots & & \\ & & 1 & \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{N-1} \end{bmatrix} = \begin{bmatrix} x_{N-1} \\ x_0 \\ \vdots \\ x_{N-2} \end{bmatrix} \triangleq \mathbf{x}^{(1)}, \quad (5)$$

and the generalization followed: *the graph unit shift was defined as the left-multiplication by the graph adjacency matrix*. This is the reason why in this paper the branch of GSP developed by Sandryhaila, Moura and their peers is referred to as GSP_A , to indicate the fundamental role of the matrix \mathbf{A} .

In other words, for a signal \mathbf{x} defined over the graph $\mathcal{G} = \{\mathcal{V}, \mathbf{A}\}$, the adjacency matrix \mathbf{A} acts as a *filter* which “delays” (i. e. translates) \mathbf{x} by one unit, producing the delayed version represented hereinafter by $\mathbf{x}^{(1)} = \mathbf{A}\mathbf{x}$.

A. Graph filters

Seeing the adjacency matrix as a filter suggested the general definition of *graph filter* as any matrix $\mathbf{H} \in \mathbb{C}^{N \times N}$ [19], which preserves the necessary property that the output of a filter (i. e. the matrix-vector product) is a signal (i. e. a vector). Such a definition implies that *linearity* is always valid for graph filters, since the distributivity of matrix multiplication with respect to matrix addition guarantees that

$$\mathbf{H}(\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2) = \alpha_1 \mathbf{H}\mathbf{x}_1 + \alpha_2 \mathbf{H}\mathbf{x}_2. \quad (6)$$

The next desirable property would be *shift invariance*, analogous to the classical time invariance of DSP, and this means that filtering and shifting should commute. In other words, for a graph filter \mathbf{H} to be *linear and shift invariant* (LSI) it is required that $\mathbf{A}\mathbf{H}\mathbf{x} = \mathbf{H}\mathbf{A}\mathbf{x} \forall \mathbf{x}$, and therefore $\mathbf{A}\mathbf{H} = \mathbf{H}\mathbf{A}$. Sandryhaila and Moura have shown [10], [20] that LSI filters can be represented as polynomials $h(\cdot)$ evaluated over \mathbf{A} ,

$$h(\mathbf{A}) = \sum_{\ell=0}^{L-1} h_\ell \mathbf{A}^\ell, \quad (7)$$

with L smaller than or equal to the degree of the minimal polynomial of \mathbf{A} , i. e. filters LSI are finite power series on the shift operator, exactly as happens in DSP, in which LTI filters have polynomial representations on z^{-1} .

B. Graph Fourier transform

The topic of spectral analysis is key in signal processing, and the authors of GSP_A would certainly want to spend time reflecting upon how this would fit into their theory. The starting point was to look to the classical Fourier transform as the signal decomposition into a basis of eigenfunctions of the LTI filtering [21], as is indeed the basis of complex time exponentials. Then, the *graph Fourier transform* (GFT) could be defined as the decomposition into a basis of eigenvectors of LSI filtering.

Let us take the graph $\mathcal{G} = \{\mathcal{V}, \mathbf{A}\}$, $|\mathcal{V}| = N$. If \mathbf{A} is diagonalizable⁴, then one may write

$$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}, \quad (8)$$

in which \mathbf{V} contains the N eigenvectors of \mathbf{A} in its columns,

$$\mathbf{V} = (\mathbf{v}_0 \ \mathbf{v}_1 \ \dots \ \mathbf{v}_{N-1}). \quad (9)$$

Since LSI filters are polynomials in \mathbf{A} , and since a matrix and its powers share the same set of eigenvectors, the columns of \mathbf{V} form a basis of vectors invariant to LSI filtering. Besides, given that the subspaces generated by the linearly independent eigenvectors of a same eigenvalue of \mathbf{A} are irreducible, have null intersection and the dimensions of all subspaces add to

⁴If not, the reasoning may be replicated using the Jordan decomposition of \mathbf{A} .

N [22], \mathbf{V} provides a basis which is invariant to LSI filtering for the space of signals defined over \mathcal{G} .

Therefore, a signal \mathbf{x} may be decomposed into its components with respect to \mathbf{V} as

$$\begin{aligned} \mathbf{x} &= \hat{x}_0 \mathbf{v}_0 + \dots + \hat{x}_{N-1} \mathbf{v}_{N-1} \\ &= \mathbf{V}(\hat{x}_0 \ \hat{x}_1 \ \dots \ \hat{x}_{N-1})^T \\ &= \mathbf{V}\hat{\mathbf{x}}, \end{aligned} \quad (10)$$

and this is the synthesis equation of the GFT according to GSP_A . The analysis equation follows,

$$\hat{\mathbf{x}} = \mathbf{V}^{-1}\mathbf{x}. \quad (11)$$

It has been emphasized that the directed ring graph is the link between GSP and DSP, because it models the discrete-time domain. This provides a way of checking how consistent with the classical theory are the proposed GSP tools. When investigating how the GFT would act upon discrete-time signals, one should first diagonalize the adjacency matrix \mathbf{C} of the directed ring graph, given by (4). Since it is circulant, it is known to be diagonalized by the DFT matrix \mathbf{F} , with entries $F_{n,k} = \exp(-j\frac{2\pi}{N}nk)$, which contains in its *rows* the DFT eigenvectors. The calculation of the characteristic polynomial of \mathbf{C} ,

$$p_{\mathbf{C}}(\lambda) = \det(\lambda\mathbf{I} - \mathbf{C}) = \begin{vmatrix} \lambda & & & & -1 \\ -1 & \lambda & & & \\ & & \ddots & & \\ & & & -1 & \lambda \end{vmatrix} = \lambda^N - 1, \quad (12)$$

shows that its eigenvalues are the N complex roots of unity. Setting these eigenvalues as the entries of a diagonal matrix $\mathbf{\Lambda}_{\mathbf{C}}$, the eigendecomposition of \mathbf{C} may be written as

$$\mathbf{C} = \mathbf{F}^{-1}\mathbf{\Lambda}_{\mathbf{C}}\mathbf{F}, \quad (13)$$

and one can see that, in the case of directed ring graphs, the GFT and the DFT matrices *coincide*, since $\mathbf{V}^{-1} = \mathbf{F}$. This equivalence indicates a desirable consistency with the classical theory.

C. The frequency domain

The GFT in the sense of GSP_A naturally suggests the interpretation of the adjacency matrix eigenvectors \mathbf{v}_i as “frequency components” associated to the “graph frequencies” given by the eigenvalues λ_i , exactly as the Fourier component $e^{-j\Omega t}$, in the continuous time domain t , is associated to the frequency Ω . This subsection aims to provide the mathematical justification used by Sandryhaila and Moura [23] to support this understanding, along with some of our own comments.

The reader may have noticed a curious consequence from what was previously stated: unless the minimal and characteristic polynomials of \mathbf{A} are equal, the same frequency may be associated to two or more linearly independent frequency components, as indeed was the case in the example of Fig. 6. Furthermore, this figure shows that although the signal seems to be smooth, its frequency components are mostly associated with eigenvalues of high magnitude, what is counter-intuitive

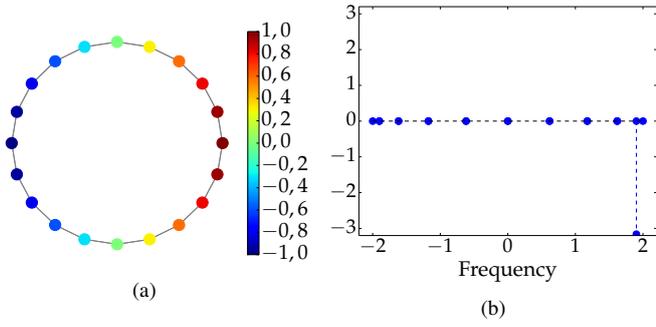


Figure 6: (a) Signal defined over an undirected ring graph and (b) its spectrum in the GSP_A sense, in which the frequency is considered to be the eigenvalues of the adjacency matrix.

and provides a motivation to define a clear criterion to distinguish high and low graph frequencies.

The following mathematical reasoning consists of taking a metric which quantifies the expected signal smoothness, and use it to propose or confirm a notion of graph frequency. The metric used by Sandryhaila and Moura was the *total variation*, taken from classical real analysis and defined for differentiable functions as [24], [25]

$$\|f\|_V = \int_{-\infty}^{\infty} |f'(t)| dt. \quad (14)$$

For discrete domain functions $f_N[n]$, the Riemman integral is replaced by first order differences,

$$\|f_N\|_V = \sum_p |f_N[n_p + 1] - f_N[n_p]|, \quad (15)$$

which clearly quantifies the dissimilarity between contiguous values of the function f_N . With this in mind, it was natural for Sandryhaila and Moura to use this metric in their mathematical formulation of frequency in GSP_A , wherein they represented the total variation of a finite-length *discrete-time signal* \mathbf{x} by

$$TV(\mathbf{x}) = \sum_n |x_n - x_{n-1 \bmod N}|. \quad (16)$$

From (5), one can see that (16) may be written in terms of the ℓ_1 -norm⁵ as $TV(\mathbf{x}) = \|\mathbf{x} - \mathbf{C}\mathbf{x}\|_1$, by using the directed ring graph adjacency matrix to perform the cyclic shift. From that point, the generalization consisted of using this expression and defining the *total variation on graphs* of a signal \mathbf{s} defined over the graph $\mathcal{G} = \{\mathcal{V}, \mathbf{A}\}$ as

$$TV_G(\mathbf{s}) \triangleq \|\mathbf{s} - \mathbf{A}^{\text{norm}}\mathbf{s}\|_1, \quad (17)$$

with $\mathbf{A}^{\text{norm}} = |\lambda_{max}|^{-1}\mathbf{A}$ and λ_{max} being the eigenvalue of \mathbf{A} having the highest absolute value. The normalization of the adjacency matrix avoids the excessive magnification of the shifted signal [23].

Let \mathbf{A} be diagonalizable as in (8) with (possibly complex) eigenvalues ordered like so

$$|\lambda_0| \leq |\lambda_1| \leq \dots \leq |\lambda_{N-1}| \triangleq |\lambda_{max}|, \quad (18)$$

⁵Throughout this paper, the concepts of ℓ_1 - and ℓ_2 -norm will be frequently used. They are particular cases of the ℓ_n -norm of a vector $\mathbf{x} \in \mathbb{C}^N$, defined as $\|\mathbf{x}\|_n \triangleq \left(\sum_{k=0}^{N-1} |x_k|^n\right)^{1/n}$.

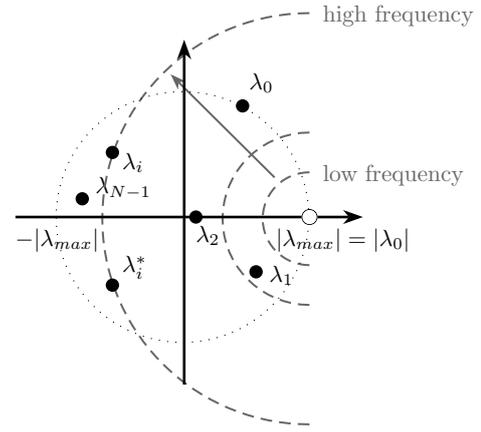


Figure 7: Frequency ordering of graph signals, from low to high frequencies, in the complex plane [23].

associated to the eigenvectors $(\mathbf{v}_i)_{i=0, \dots, N-1}$, scaled so that $\|\mathbf{v}_i\|_1 = 1 \forall i$. Taking the total variation (on graphs) of the eigenvector \mathbf{v}_k , one has

$$\begin{aligned} TV_G(\mathbf{v}_k) &= \|\mathbf{v}_k - \mathbf{A}\mathbf{v}_k\|_1 = \|\mathbf{v}_k - \frac{1}{|\lambda_{max}|} \lambda_k \mathbf{v}_k\|_1 \\ &= \left|1 - \frac{\lambda_k}{|\lambda_{max}|}\right| \|\mathbf{v}_k\|_1 = \left|\lambda_k - |\lambda_{max}|\right| \frac{\|\mathbf{v}_k\|_1}{|\lambda_{max}|} \end{aligned}$$

so that, since $\|\mathbf{v}_k\|_1 = 1$,

$$\left|\lambda_i - |\lambda_{max}|\right| \leq \left|\lambda_j - |\lambda_{max}|\right| \iff TV_G(\mathbf{v}_i) \leq TV_G(\mathbf{v}_j), \quad (19)$$

i. e. frequency components associated to eigenvalues closer to the real point $|\lambda_{max}|$ in the complex plane are *smoother* (because they have lower total variation), and therefore are said to be of *low frequency*. Fig. 7 illustrates this ordering for graph frequencies, what clarifies the spectrum of the signal in Fig. 6a (notice that since the graph is undirected, its adjacency matrix is symmetric and the eigenvalues are real-valued).

Let us take the directed graph in Fig. 8a to try to verify the consistency of the notion of frequency just derived. For this, along with the total variation on graphs, also the number of zero crossings (i. e. the number of edges connecting vertices with signal samples of different sign) will be used to *quantify* frequency. This quantity is also related to frequency in classical theory: the more a discrete signal has contiguous samples with different sign, the higher are its frequency components. These two functions, the total variation on graphs and the number of zero crossings, were calculated for each of the adjacency matrix eigenvectors, and the result is shown in Fig. 8, in which the eigenvectors \mathbf{v}_k are ordered in such a way that the respective eigenvalues λ_k appear from the closest to the farthest from the real point $|\lambda_{max}|$ in the complex plane. Both metrics have similar behaviour, but since the number of zero crossings is indifferent to graph signal variations which do not change sign, it was already expected to be less accurate as a figure of merit for frequency. It matters to highlight, however, how the adopted eigenvector ordering indeed implies an ascending frequency order, since both functions in Fig. 8b and 8c agree on the tendency of growth. More than that,

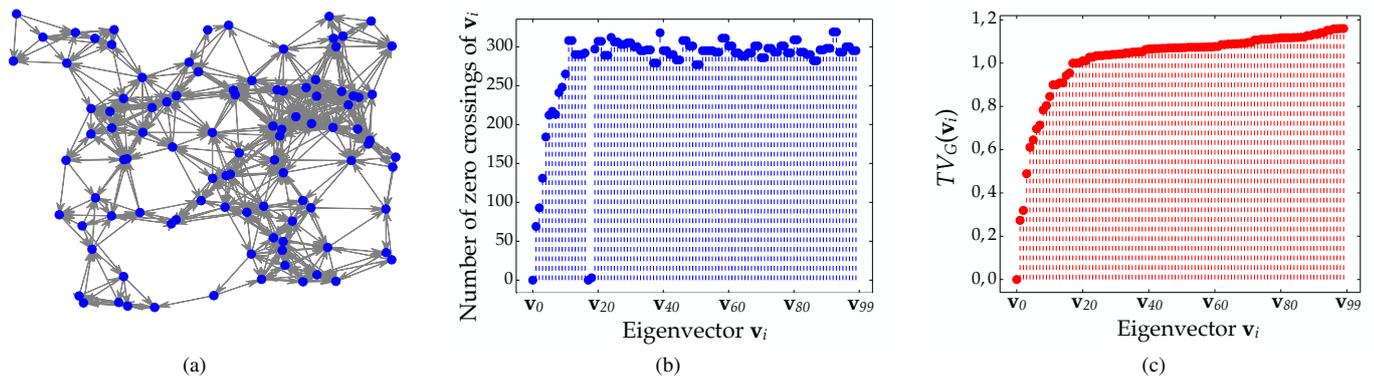


Figure 8: (a) Directed sensor graph, with $N = 100$ vertices and no loops or multiple edges. (b) Number of zero crossings and (c) total variation of the eigenvectors $(\mathbf{v}_i)_{i=0,\dots,N-1}$ of the adjacency matrix \mathbf{A} of the graph in (a), ordered so that the respective eigenvalues appear from the closest to the farthest from the real point $|\lambda_{max}|$ in the complex plane. That is, according to (19) and Fig. 7, the eigenvectors are disposed in ascending order of frequency.

$TV_G(\mathbf{v}_k)$ grows *monotonically*, as it should do according to (19).

It is convenient to conclude this discussion on the graph frequency domain by referring to the frequency response of graph filters. The definition given in Subsection III-A considers the action of a matrix on a signal \mathbf{x} in the vertex domain of the graph $\mathcal{G} = \{\mathcal{V}, \mathbf{A}\}$. In order to understand how the filter acts in the GFT domain, hereinafter called frequency domain, one may use (8) and the polynomial representation of LSI filters. Let us take the filter $\mathbf{H} = \sum_{\ell=0}^L h_{\ell} \mathbf{A}^{\ell}$ and its output to the input \mathbf{x} given by

$$\begin{aligned} \mathbf{H}\mathbf{x} &= \sum_{\ell=0}^L h_{\ell} \mathbf{A}^{\ell} \mathbf{x} = \sum_{\ell=0}^L h_{\ell} (\mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1})^{\ell} \mathbf{x} \\ &= \mathbf{V} \left(\sum_{\ell=0}^L h_{\ell} \mathbf{\Lambda}^{\ell} \right) \mathbf{V}^{-1} \mathbf{x}. \end{aligned} \quad (20)$$

Taking the GFT of both sides of the last equation yields

$$\mathbf{V}^{-1} \mathbf{H}\mathbf{x} = h(\mathbf{\Lambda}) \hat{\mathbf{x}}, \quad (21)$$

which indicates that left-multiplication by \mathbf{H} (action of the filter in the vertex domain) is equivalent to the left-multiplication, in the frequency domain, by the matrix $h(\mathbf{\Lambda})$. In other words, $h(\mathbf{\Lambda})$ represents the frequency response of \mathbf{H} .

D. Fractionalization of operators

It is worth mentioning some of the few recent efforts to provide form and meaning to *fractional* graph operators. For example, one could think of a *non-integer* graph shift \mathbf{A}^a , with $0 < a < 1$, or a fractional GFT, inspired by the classical notion of fractional Fourier transform over the complex square-integrable functions. These studies are lead by motivations mirroring those from discrete-time fractional filters and transforms, such as more suitable signal representations [26] and applications in optics, quantum mechanics and communication systems [27]–[31].

Regarding the derivation of a fractional graph Fourier transform (FrGFT), Wang *et al.* started from the Jordan

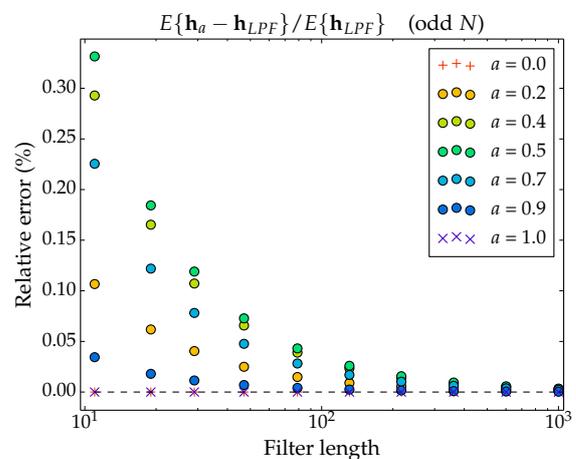


Figure 9: Percentual ℓ^2 -norm (energy) of the error $\mathbf{h}_{LPF} - \mathbf{h}_a$ (respectively ideal and graph-based fractional time shift filters) as a function of the filter length N and the delay a . We denote the ideal fractional delay (or interpolation) filter as $\mathbf{h}_{LPF} = \frac{\sin(\pi(n-a))}{\pi(n-a)}$.

decomposition of the GFT operator to arrive at an expression for the FrGFT which satisfies the desired properties of index additivity, reduction to the fractional discrete-time Fourier transform in the case of ring graphs, reduction to the DFT at integer rotation parameters and the convolution theorem [32].

The fractionalization of the *shift operator* was also investigated and studies on its meaning and applications were conducted by Ribeiro and Lima [33]. The results indicated a different behaviour of the defined operator when applied on directed and undirected graphs, and the consistency with the classical theory was verified by numerical convergence between the ideal and the graph-based fractional time shift filters, as shown in Fig. 9.

IV. GSP_L: GRAPH SPECTRAL THEORY AND GSP

Graph spectral theory is a branch of graph theory concerned with the eigendecomposition of graph characteristic matrices, such as the adjacency and Laplacian matrices, and

the properties derived from such a decomposition. Developed from matrix theory and linear algebra, it finds applications in quantum mechanics, communication theory, chemistry and quite other fields [14], and it has been used by Ortega, Shuman and many others as foundation for the proposal and growth of a theoretical framework for GSP based on the spectral properties of the Laplacian matrix, therefore being referred in this paper by GSP_L . In this branch of GSP, it is worth mentioning, the authors use to restrict the graph space in which their tools are applied to undirected graphs with non-negative real edge weights, as Sandryhaila and Moura point out [10].

The Laplacian matrix was mentioned earlier in this paper, having the expression

$$\mathbf{L} = \mathbf{D} - \mathbf{A}, \quad (22)$$

but sometimes it is used the normalized version, given by

$$\mathcal{L} = \mathbf{D}^{-1/2} \mathbf{L} \mathbf{D}^{-1/2}. \quad (23)$$

The version in (23) is largely used in graph spectral theory [14], but the framework about to be presented is known to use the non-normalized Laplacian matrix, given by (22).

An important property of the Laplacian matrix \mathbf{L} for undirected weighted graphs may be derived from its decomposition in terms of the *oriented incidence matrix*, denoted by \mathbf{B} . Although its definitions vary depending on whether the graph is directed or not, weighted or not, hereinafter the matrix \mathbf{B} is considered as follows (see [34, Chapter 2] for other definitions of the incidence matrix).

Definition 1 (Orientation of a graph). *Let us consider the undirected graph $\mathcal{G} = \{\mathcal{V}, \mathbf{A}\}$. A possible orientation of \mathcal{G} is a graph $\mathcal{G}' = \{\mathcal{V}, \mathbf{A}'\}$ so that $\mathbf{A}' + \mathbf{A}'^T = \mathbf{A}$. That is, \mathcal{G}' is a directed version of \mathcal{G} , obtained from \mathcal{G} by imposing a certain direction to each of its edges.*

Definition 2 (Oriented incidence matrix). *Let us consider the same undirected graph $\mathcal{G} = \{\mathcal{V}, \mathbf{A}\}$, with $|\mathcal{V}| = N$ vertices and $|\mathcal{E}| = E$ edges weighted according to the function $\omega : \mathcal{E} \rightarrow \mathbb{R}^+$. Let $\mathcal{G}' = \{\mathcal{V}, \mathbf{A}'\}$ be a particular orientation of \mathcal{G} . An oriented incidence matrix $\mathbf{B} \in \mathbb{R}^{N \times E}$ of \mathcal{G} is given by*

$$B_{i,j} \triangleq \begin{cases} \sqrt{\omega(e'_j)}, & \text{if } e'_j \text{ arrives at } v_i \\ -\sqrt{\omega(e'_j)}, & \text{if } e'_j \text{ departs from } v_i \\ 0, & \text{otherwise.} \end{cases} \quad (24)$$

One should notice that, in the context of GSP_L , in which the considered graphs have real non-negative edge weights, the oriented incidence matrix \mathbf{B} is always real-valued.

Theorem 1. *For a given undirected graph with oriented incidence matrix \mathbf{B} , its Laplacian matrix may be written as*

$$\mathbf{L} = \mathbf{B} \mathbf{B}^T. \quad (25)$$

Theorem 1 [34, Proposition 2.3] is a way of proving that the Laplacian matrix of an undirected graph is *positive semidefinite* (i. e. $\mathbf{x}^T \mathbf{L} \mathbf{x} \geq 0, \forall \mathbf{x}$), for any real-valued signal $\mathbf{x} = (x_0 \ x_1 \ \dots \ x_{N-1})^T$ satisfies

$$\mathbf{x}^T \mathbf{L} \mathbf{x} = \mathbf{x}^T \mathbf{B} \mathbf{B}^T \mathbf{x} = (\mathbf{B}^T \mathbf{x})^T (\mathbf{B}^T \mathbf{x}) = \|\mathbf{B}^T \mathbf{x}\|^2 \geq 0,$$

and, therefore, as occurs for all positive semidefinite matrices, the eigenvalues of \mathbf{L} are all real and non-negative. Moreover, it has been shown that zero is an eigenvalue with multiplicity equal to the number of connected components of the graph (cf. [11] and references therein), and therefore connected graphs have only one null eigenvalue of \mathbf{L} . Gathering all these facts, the eigenvalues of the Laplacian matrix, denoted by the letter γ , may be ordered as $\gamma_0 = 0 < \gamma_1 \leq \dots \leq \gamma_{N-1}$, $\gamma_i \in \mathbb{R}, \forall i$.

A. GFT and the Laplacian operator

Working on (22) it is possible to see \mathbf{L} as a difference operator acting upon a signal \mathbf{x} , updating each sample with the difference between the value on a vertex and its neighbours. That is,

$$\mathbf{L} \mathbf{x} = \mathbf{D} \mathbf{x} - \mathbf{A} \mathbf{x}, \quad (26)$$

$$\begin{aligned} \Rightarrow (\mathbf{L} \mathbf{x})_i &= d_i x_i - \sum_{k \mid v_k \in \mathcal{N}_i} A_{ik} x_k = \sum_k A_{ik} x_i - \sum_{k \mid v_k \in \mathcal{N}_i} A_{ik} x_k \\ &= \sum_{k \mid v_k \in \mathcal{N}_i} A_{ik} [x_i - x_k]. \end{aligned} \quad (27)$$

When conceiving their definition for the GFT, Shuman *et al.* started from the fact that the continuous-time Fourier transform consists of a decomposition into a basis of eigenfunctions of the Laplacian operator (second derivative) [11], [35]:

$$\Delta e^{j\omega t} = \frac{\partial^2}{\partial t^2} e^{j\omega t} = -\omega^2 e^{j\omega t}. \quad (28)$$

The generalization followed: the GFT of a signal \mathbf{x} was defined to be its expansion in terms of the eigenvectors of the Laplacian matrix (of the graph over which the signal is defined). Since \mathbf{L} is always (in GSP_L) real-valued and symmetric, therefore diagonalizable, it may be written as $\mathbf{L} = \mathbf{U} \mathbf{T} \mathbf{U}^{-1}$, and the GFT pair of equations are given by

$$\begin{aligned} \tilde{\mathbf{x}} &\triangleq \mathbf{U}^{-1} \mathbf{x} && \text{(analysis)} \\ \mathbf{x} &= \mathbf{U} \tilde{\mathbf{x}}, && \text{(synthesis)} \end{aligned} \quad (29)$$

with \mathbf{U} containing the eigenvectors in its columns and \mathbf{T} consisting of a diagonal matrix filled with the Laplacian matrix eigenvalues.

Such a definition brings a notion of frequency which is consistent with the classical interpretation. In (28), the frequency information is hidden in the eigenvalues $-\omega^2$ associated to each eigenfunction $e^{j\omega t}$, and the closer to zero is the eigenvalue, the smoother is its respective harmonic component. The same occurs with the eigenvalues and eigenvectors of the Laplacian matrix, and a way of verifying this is using the norm $\|\mathbf{L} \mathbf{x}\|_2$: since \mathbf{L} is a difference operator, $\|\mathbf{L} \mathbf{x}\|_2$ acts as a metric similar to the total variation on graphs of GSP_A (cf. (16)). If \mathbf{u}_i is the eigenvector of \mathbf{L} associated to the eigenvalue γ_i ,

$$\|\mathbf{L} \mathbf{u}_i\| = \|\gamma_i \mathbf{u}_i\| = \gamma_i \|\mathbf{u}_i\| \quad (\text{since } \gamma_i \geq 0), \quad (30)$$

so that, if $\|\mathbf{u}_i\|_2 = 1$ (normalized magnitude), then

$$\gamma_i \leq \gamma_j \Rightarrow \|\mathbf{L} \mathbf{u}_i\|_2 \leq \|\mathbf{L} \mathbf{u}_j\|_2, \quad (31)$$

in which case \mathbf{u}_i is smoother than \mathbf{u}_j .

Besides, looking at the unique null eigenvalue of \mathbf{L} (for connected undirected graphs), one has

$$\mathbf{L}\mathbf{u}_0 = \gamma_0\mathbf{u}_0 = \mathbf{0}, \quad (32)$$

which consists of a system of homogeneous linear equations on the entries of vector \mathbf{u}_0 . The solutions for such a system form the so called *null space* of \mathbf{L} , and since its dimension equals the geometric multiplicity of $\gamma_0 = 0$, it follows that this null space is one-dimensional. Therefore, it suffices to find a single non-trivial solution of (32) to obtain a basis for the possible eigenvectors \mathbf{u}_0 . From (27), one may see that

$$(\mathbf{L}\mathbf{u}_0)_i = 0 \Rightarrow \sum_{k \mid v_k \in \mathcal{N}_i} A_{ik}[u_{0i} - u_{0k}] = 0, \quad (33)$$

what shows that any *constant* vector is a solution for (32). This leads to a quite satisfactory conclusion: the eigenvector associated to the null eigenvalue (frequency) is constant, as it is known to occur in the classical signal processing theory. If the eigenvectors are normalized, then each entry of \mathbf{u}_0 equals $1/\sqrt{N}$.

Fig. 10 confirms visually the notion of frequency based on the eigenvalues of the Laplacian matrix, using as an example a sensor graph with 1000 vertices, with edges weighted according to (3). The component with zero frequency is the eigenvector \mathbf{u}_0 , and the smoothness of the eigenvectors decrease as the frequency (eigenvalue) rises. Fig. 10f shows the number of zero crossings of each eigenvector, meaning the number of graph edges connecting samples of the eigenvector with different sign.

B. Filtering

Filtering in GSP_L is firstly defined in the frequency domain: the frequency response of a filter \mathbf{H} is the vector $\tilde{\mathbf{h}} \in \mathbb{C}^N$ such that, for a signal \mathbf{x} with spectrum $\tilde{\mathbf{x}} = \mathbf{U}^{-1}\mathbf{x}$ defined over a certain graph, the filter output is

$$\tilde{\mathbf{y}} \triangleq \tilde{\mathbf{h}} \odot \tilde{\mathbf{x}} = \text{diag}(\tilde{\mathbf{h}})\tilde{\mathbf{x}}, \quad (34)$$

in which \odot represents the Hadamard (element-wise) product and $\text{diag}(\cdot)$ is a diagonal matrix with diagonal entries given by the argument. Taking the inverse GFT of both sides yields

$$\mathbf{U}\tilde{\mathbf{y}} = \mathbf{y} = \mathbf{U} \left[\text{diag}(\tilde{\mathbf{h}})\tilde{\mathbf{x}} \right] = \mathbf{U} \text{diag}(\tilde{\mathbf{h}})\mathbf{U}^{-1}\mathbf{x}, \quad (35)$$

that is, filtering in the vertex domain is given by

$$\mathbf{y} = \mathbf{H}\mathbf{x}, \quad (36)$$

with

$$\mathbf{H} \triangleq \mathbf{U} \text{diag}(\tilde{\mathbf{h}})\mathbf{U}^{-1}. \quad (37)$$

It is possible to write the entries of $\tilde{\mathbf{h}}$ as polynomials of degree less than or equal to K on the eigenvalues of \mathbf{L} [11], so that

$$\tilde{h}_\ell = \sum_{k=0}^K a_k \gamma_\ell^k, \quad (38)$$

and so that the filter output to the input \mathbf{x} , obtained by taking the inverse GFT of (34), has entries given by

$$y_i = \sum_{\ell=0}^{N-1} u_{\ell,i} \tilde{h}_\ell \tilde{x}_\ell, \quad (39)$$

in which $u_{\ell,i}$ is the i -th component of the eigenvector \mathbf{u}_ℓ . Substituting (38) into (39),

$$\begin{aligned} y_i &= \sum_{\ell=0}^{N-1} u_{\ell,i} \left(\sum_{k=0}^K a_k \gamma_\ell^k \right) \left(\sum_{j=0}^{N-1} u_{\ell,j}^* x_j \right) \\ &= \sum_{\ell} \sum_k \sum_j x_j a_k \gamma_\ell^k u_{\ell,i} u_{\ell,j}^* \\ &= \sum_j x_j \left[\sum_k a_k \left(\sum_{\ell} \gamma_\ell^k u_{\ell,i} u_{\ell,j}^* \right) \right], \end{aligned} \quad (40)$$

with $u_{\ell,j}^*$ representing the complex conjugate of the j -th component of the eigenvector \mathbf{u}_ℓ .

But since $L_{i,j} = (\mathbf{U}\mathbf{T}\mathbf{U}^{-1})_{i,j} = \sum_{\ell=0}^{N-1} \gamma_\ell u_{\ell,i} u_{\ell,j}^*$, then

$$\sum_{\ell=0}^{N-1} \gamma_\ell^k u_{\ell,i} u_{\ell,j}^* = (\mathbf{L}^k)_{i,j} \quad (41)$$

and therefore

$$y_i = \sum_{j=0}^{N-1} x_j \sum_{k=0}^K a_k (\mathbf{L}^k)_{i,j}. \quad (42)$$

Finally, one may use the fact that if the length of the smallest path from v_i to v_j is greater than k , then $(\mathbf{L}^k)_{i,j} = 0$ ([36, lemma 5.2], as referenced in [11]). Thus,

$$y_i = \sum_{j|v_j \in \mathcal{N}(i,K)} H_{i,j} x_j, \quad (43)$$

in which $H_{i,j} = \sum_{k=0}^K a_k (\mathbf{L}^k)_{i,j}$. As suggested by (43), filtering in the vertex domain is a localized linear operation, linearly combining samples in the neighbourhood of a vertex up to a radius of K edges.

V. APPLICATIONS

This section aims to provide examples of data handling using the concepts from GSP, scenarios in which the graph model offers valuable tools for processing the signal at hand. We present two practical applications involving processing of sensor networks with GSP_L , and close the section offering a review of other fields of science and technology which have benefited from GSP techniques. It is appropriate, however, to firstly comment on the work by Deri and Moura [37], in which GSP_A is applied to the analysis of New York City taxi data. This highly recommended paper deals with the problems of signal definition and extraction from the New York taxi database, large-scale eigenvalue computation for non-diagonalizable matrices and graph frequency interpretation.

The underlying graph vertices, in that case, consist of 6048 geo-locations in Manhattan connected by road segments (graph edges), and the graph signals (Fig. 11a) contain the number of taxi rides over a certain time interval that crosses each graph node. In order to create those signals, for each pair of taxi trip start- and end-point a shortest path through the graph had to be found, using Dijkstra's algorithm. The analysis have shown that the frequency components (Fig. 11b) provided information which was not clear from the graph signal alone.

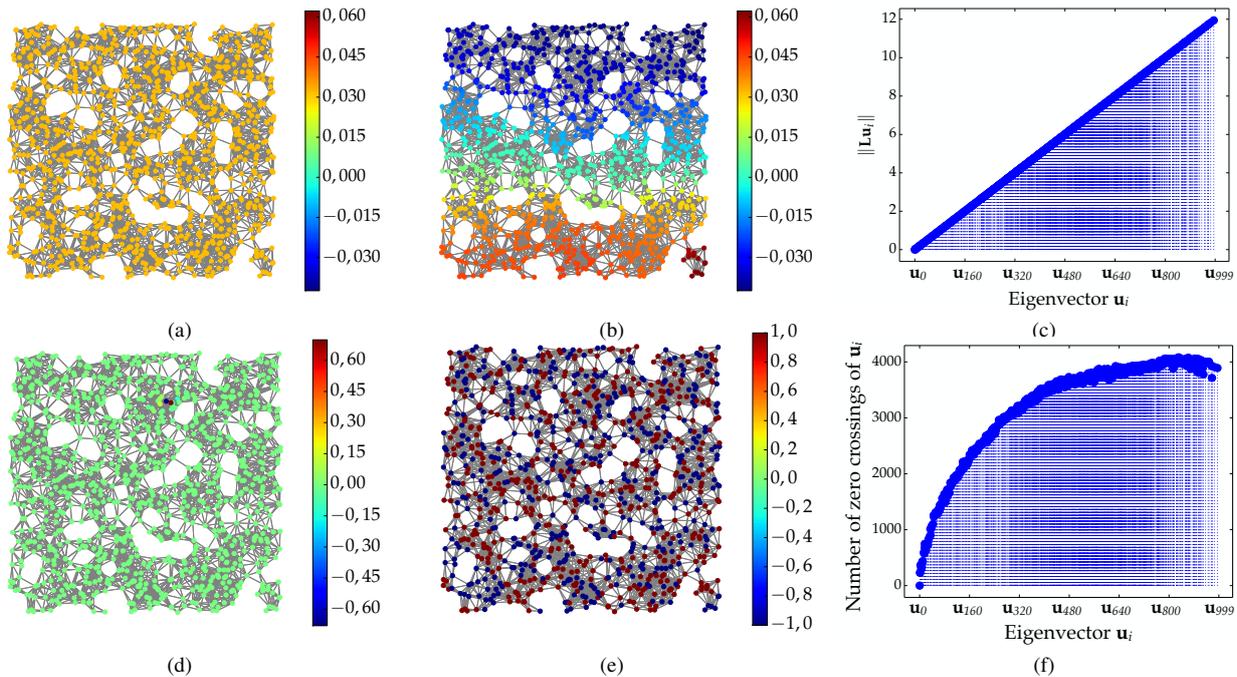


Figure 10: Some eigenvectors of the Laplacian matrix of a sensor directed graph with 1000 vertices: (a) \mathbf{u}_0 , (b) \mathbf{u}_1 and (d) \mathbf{u}_{999} . Since \mathbf{u}_{999} presented a high-intensity isolated peak, hard to discriminate visually, in (e) it is shown a binarized version $\bar{\mathbf{u}}_{999}$, with positive samples being set to 1 and negative samples set to -1 . In (c) and (f) two measures of signal smoothness are depicted: the norm $\|\mathbf{L}\mathbf{u}_i\|_2$ and the number of zero crossings of the eigenvectors. The entries of \mathbf{u}_0 equal $1/\sqrt{1000} \approx 0,03$.

For example, the generalized eigenvector with highest Fourier coefficient presented high energy density around areas such as Hell's Kitchen and Broadway, famous for concentrating many options of restaurants and theaters. Besides, the authors have shown that reconstructing the graph signal with 70% of frequency components (Fig. 11d) lead to a resulting signal preserving most of the original information, with peak signal-to-noise ratio of around 30dB. Such a work illustrates the valuable benefits, along with the non-trivial challenges, of using GSP_A to address problems with data over directed graphs.

A. Signal denoising

As a first hands-on example involving GSP_L and sensor networks, let us consider the smooth signal defined over the Minnesota road graph, as in Fig. 12a, with components given by

$$s_i = \cos(\text{coord}_x(v_i)) + \sin(\text{coord}_y(v_i)), \quad (44)$$

in which coord_x and coord_y indicate the geodesic coordinates of the vertices (street corners and crossroads) on the Minnesota state.

A Gaussian noise \mathbf{n} with zero mean and standard deviation of 20% the amplitude of \mathbf{s} was added to the signal, yielding the noisy version shown in Fig. 12b. The spectral analysis of both the original signal and the Gaussian noise suggested that, although the noise is approximately white, the signal – due to its smoothness – has its energy concentrated in the low-frequency range, in the eigenvalues $0 \leq \gamma_i \leq 0,5$, as seen in

Fig. 12c. For this reason, it was used an ideal low-pass filter of unit gain and cut-off frequency $\gamma_{\text{cut}} = 0,5$, so as to reduce the influence of the noise in the original signal.

The result after the filtering is shown in Fig. 12d. The fraction of the energy due to noise dropped from

$$\|\mathbf{s} - (\mathbf{s} + \mathbf{n})\| = \|\mathbf{n}\| \approx 44.6\% \quad (45)$$

to

$$\|\mathbf{s} - \mathbf{U}(\tilde{\mathbf{h}}_{LPF} \odot (\tilde{\mathbf{s}} + \tilde{\mathbf{n}}))\| \approx 18.9\%, \quad (46)$$

i. e. the energy of the error with respect to the original signal halved. It matters to say that although simple as this example may be, it highlights the importance of taking into account the underlying structure of signals defined over such irregular domains: the high frequency-localization of the signal energy was only possible due to its decomposition into the basis of eigenvectors of the Laplacian matrix of the graph over which the signal was defined. With such considerations, simple techniques lead to useful results, as illustrated.

B. Data inference via signal regularization

Let us consider the problem of, given a subset of graph vertices with known samples, estimate the graph signal on the other vertices so that the resulting signal is smooth. It can be viewed as a signal reconstruction scenario, and Sandryhaila and Moura used this idea to estimate a *discrete-valued signal*⁶,

⁶The graph signal is already discrete with respect to its domain, by definition, but a *discrete-valued* graph signal is also discrete with respect to the sample space, i. e. the sample values must belong to a finite set.

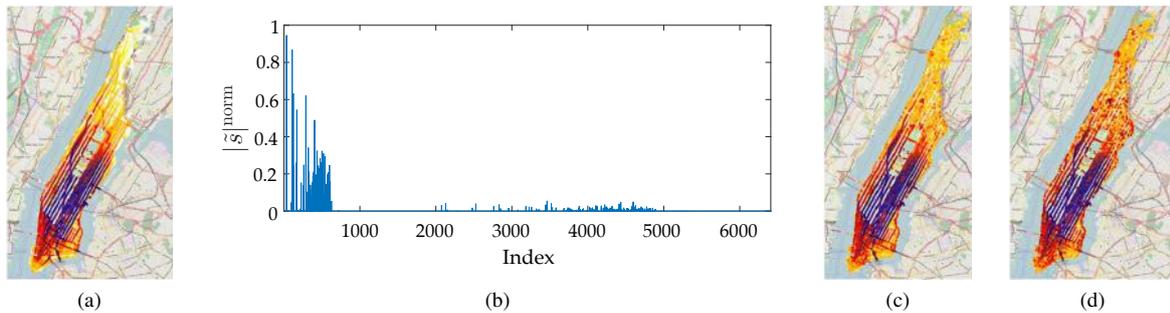


Figure 11: (a) Graph signal over the Manhattan road graph, created from the number of taxi trips from June to August on Fridays, 9:00 pm to 10:00 pm, averaged from 2010 to 2013. The pseudocolor scale ranges from white/yellow to purple/black. (b) GFT spectrum of the signal in (a), in the sense of GSP_A (the variable in the horizontal axis is the index of the eigenvalue), with frequency components ordered by the eigenvectors total variation. (c-d) Reconstructed signals using the initial 5000 and 4500 Fourier coefficients, respectively. From [37], with permission.

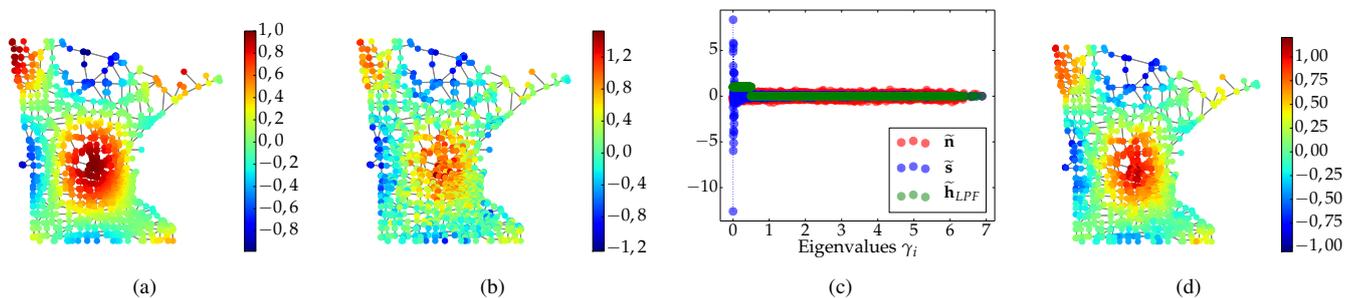


Figure 12: (a) Smooth signal defined over the Minnesota road graph. (b) Result of adding a Gaussian noise with standard deviation equal to 20% the amplitude of the original signal. (c) Spectra of the original signal, of the noise and of the ideal low-pass filter. (d) Output of the low-pass filter.

so as to apply the method in classification problems [23]. For that goal they ran an optimization algorithm to obtain the signal \mathbf{s} which minimizes the norm $\|\mathbf{s} - \mathbf{A}^{\text{norm}}\mathbf{s}\|_2$, subjected to the restriction that the estimated signal was close enough to the known values in the vertex subset mentioned previously.

A similar approach was conducted by us, but in the context of GSP_L . The following real situation was considered: given the mean rainfall intensity on January of 609 Brazilian cities, taken from the Embrapa database⁷, how could one estimate this data in all other cities of Brazil? With such problem, it is intended to illustrate the use of optimization techniques applied to GSP.

The idea consists of obtaining a signal $\mathbf{s}^{(\text{predicted})}$ with the mean precipitation indices in such a way it is smooth in the sense of GSP_L , while simultaneously preserving the known *a priori* data at the 609 cities. That translates into the optimization problem of the objective function

$$\mathbf{s}^{(\text{predicted})} = \arg \min_{\mathbf{s} \in \mathbb{R}^N} \|\mathbf{L}\mathbf{s}\|, \quad (47)$$

subjected to

$$\|\mathbf{P}(\mathbf{s} - \mathbf{s}^{(\text{known})})\| = 0, \quad (48)$$

⁷The *Banco de Dados Climáticos do Brasil* (freely translated as Brazilian Weather Database) was compiled by Embrapa and ESALQ/USP, and it is available at: <https://www.cnpm.embrapa.br/projetos/bdclima/index.html>

in which \mathbf{P} is the diagonal matrix with entries given by

$$P_{ii} = \begin{cases} 1, & \text{if } s_i^{(\text{known})} \neq 0 \\ 0, & \text{otherwise,} \end{cases} \quad (49)$$

and $\mathbf{s}^{(\text{known})}$ is a vector of size N having only 609 non-zero samples, corresponding to the vertices with known rainfall data. For a solid reference on optimization algorithms, the authors recommend [38].

In Fig. 13a all 5570 Brazilian cities are represented as vertices of an edgeless graph⁸, over which is defined a signal composed by zeros except at the 609 cities, to which were assigned the known pluviometry data. So as to run the optimization algorithm faster, it was decided to reduce the total number of vertices to $N = 1000$ (decimating only vertices with null samples), obtaining the signal $\mathbf{s}^{(\text{known})}$ shown in Fig. 13b, in which the underlying graph was created by connecting a vertex to its 5 closest neighbours. This aimed to result in a connected graph with reasonably sparse adjacency matrix (as discussed in Subsection II-C). The edges were weighted according to

$$A_{i,j} = \exp(-\text{dist}^2(v_i, v_j)), \quad (50)$$

⁸The geographic coordinates of each city were retrieved from the following database: <https://github.com/kelvins/Municipios-Brasileiros>

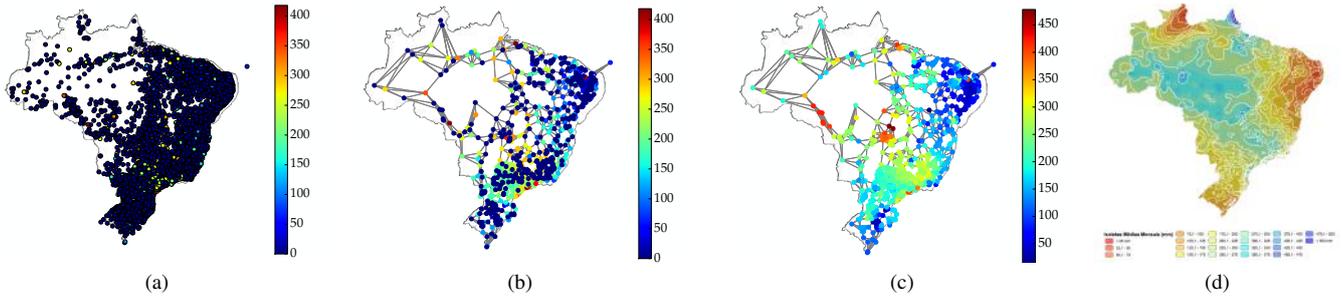


Figure 13: (a) Signal defined over the graph (without edges) of the 5570 Brazilian cities, with all samples equal zero except at the 609 vertices which correspond to the Embrapa database. (b) Signal $s^{(\text{known})}$, with 1000 samples, 609 of which are the known pluviometry data. (c) Signal $s^{(\text{predicted})}$, with the estimation of the previously null samples so as to obtain a smooth signal along the whole graph. (d) Mean rainfall intensity diagram during January, from 1977 to 2006, in Brazil. The vector image of the (approximated) profile of the Brazilian territory was made by Felipe Micaroni Lalli and is available under CC BY-SA license at https://commons.wikimedia.org/wiki/File:Contorno_do_mapa_do_Brasil.svg.

making $2\theta^2 = 1$ in (3). Since the graph was already made connected and there was no threshold T involved in the decision of which edges were to be created, there was no clear rule to indicate which value of standard deviation θ should be used. For this reason, it was made $2\theta^2 = 1$.

The estimated signal should ideally approximate the map in Fig. 13d, which depicts areas of constant rainfall index during the month of January from 1977 to 2006⁹. *It is important to notice the difference between the color pattern adopted in the map and the one used in this paper: the warmer the color in the map, the smaller is the represented value.*

So as to reduce the number of variables in the objective function, the restriction $\|\mathbf{P}(s - s^{(\text{known})})\| = 0$ was inserted into the declaration of the objective function itself, so that only $1000 - 609 = 391$ independent variables were left. As a consequence, it consists of an *unconstrained optimization* problem with 391 variables, and to handle it we used the BFGS algorithm and the function `optimize.minimize()`¹⁰ from the package Scipy, in Python [39]. The result is shown in Fig. 13c. Although the algorithm did not converge after successively increasing the iteration limit (from 100 up to 2000), the final result reasonably approximated the map in Fig. 13d, keeping in mind the opposite color pattern as previously commented. It matters to highlight shortly that it is not possible to quantify the quality of this signal estimation, since the database did not provide the mean rainfall intensity in the 391 Brazilian cities which had the data estimated, hence we are restricted to a visual comparison between the resulting graph signal and the reference map. However, as a way of illustrating the notion of frequency in GSP_L and the use of optimization to estimate the information on a graph taking by hypothesis the smoothness of the signal, the result was satisfactory, although further studies are needed to verify which objective function would be more appropriate for such an application, or which

minimization algorithm is the most adequate for the problem.

C. Other areas of applications

The previous examples illustrated the classical use of GSP in sensor networks, which naturally arises from the definition of the signal domain as a graph. However, less obvious scenarios also greatly benefited from the tools of GSP.

Firstly, the reader may remind from a few pages ago that an *image* may be modelled as an undirected graph (Fig. 4b) [15], and setting the weights of a grid graph according to the image signal defined upon it leads to great dimensionality reduction in the GFT domain [22]. Indeed, the DCT analysis, which is known as a useful image compression tool, is related to the graph processing framework because, when dealing with simple 1D and 2D grids (undirected line or grid graphs), the DCT diagonalizes the graphs Laplacian matrices and therefore matches the GFT [40]. GSP has also been used in light-field images, which capture 3D information with a single exposure, as for instance in coding [41], compression [42] and super-resolution schemes [43]. For an extensive review on GSP on image compression, restoration, filtering and segmentation, the authors recommend [44].

Another important field of work for GSP involves *biological networks*, i. e. the treatment of human or natural networked systems for data/structure inference and processing. For example, with respect to gene regulatory networks, the use of graph-based methods lead to the improvement of the three state-of-the-art schemes of network inference [45], [46]. A work by Nguyen *et al.* represented the human body as a three-dimensional dynamic mesh and applied graph wavelet filter banks to compress information of position and color, outperforming usual methods for coding of the human body [47]. The processing of brain signals may be, however, the most intriguing and prolific of such applications, arising through the assignment of signals such as fMRI readings to graphs defined by functional brain networks [48], [49], which have for example demonstrated a close relation between the signals' lowest and highest frequency components and the learning of a motor task [50].

⁹Data from the Geological Service of Brazil, also known as Companhia de Pesquisa de Recursos Minerais (CPRM), public company vinculated to the Brazilian Ministry of Mines and Energy. This and other pluviometry maps are found in: <http://www.cprm.gov.br/publique/Hidrologia/Mapas-e-Publicacoes/Atlas-Pluviometrico-do-Brasil-1351.html>

¹⁰Documentation available at: <https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.minimize.html>

As mentioned in the Subsection V-B, the regularization of a discrete-valued graph signal was used for data classification in [23]. This was an example of the large application of GSP in machine learning, from what we could also mention semi-supervised learning through the use of adaptive graph filters [51], recommendation systems consisting in collaborative and content-based filtering, the latter via regularization with total variation on graphs [52], and community detection in networks via fine graph wavelet design, what allows for multiscale community mining [53].

VI. OPPORTUNITIES FOR FURTHER INVESTIGATION

As previously stated, GSP_L was developed having in mind undirected graphs with non-negative weights. In such cases, \mathbf{L} is diagonalizable and positive semidefinite, what brings convenient properties to the application of theoretical concepts, such as having only positive frequencies and a DC component associated to the zero frequency. The GSP_A framework presents the advantage of considering more general graphs, directed and with complex weights (provided that no loops or multiple edges are included), but some disadvantages do exist as well. For example, as mentioned in [54], the use of the adjacency matrix as elementary block and the work with directed graphs lead to a series of obstacles when $\mathbf{A} = \mathbf{V}_G \mathbf{J} \mathbf{V}_G^{-1}$ is not diagonalizable, as for instance the fact that the generalized eigenvector basis \mathbf{V}_G is generally not orthogonal, and therefore the GFT does not preserve the scalar product, and the problems derived from numerical instability that arise in the computation of the Jordan normal form of practical-size matrices.

Beside this, Deri and Moura [55] show that both frameworks still present the problem of non-unicity of the GFT, whenever there are repeated frequencies (λ_i or γ_i) — even if the eigenvalues have equal geometric and algebraic multiplicities. This problem can be illustrated using a numerical example with a complete 3-vertex unweighted graph (a triangle). Complete graphs with n vertices have *normalized* Laplacian matrix with eigenvalues equal to 0 with multiplicity 1, and equal to $n/(n-1)$ with multiplicity $n-1$ [14, Example 1.1]. Therefore, its non-normalized Laplacian matrix (in this case, it is equal to the normalized version multiplied by 2) may be expressed as

$$\begin{aligned} \mathbf{L} &= \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} = \mathbf{U}_1 \mathbf{\Gamma} \mathbf{U}_1^{-1} \\ &= \begin{bmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix}, \end{aligned}$$

but another unitary eigenvector matrix is possible. Replacing the eigenvectors associated to $\gamma = 3$ by the normalized sum and difference of $(-1/\sqrt{2}, 0, 1/\sqrt{2})^T$ and $(-1/\sqrt{2}, 1/\sqrt{2}, 0)^T$ yields

$$\mathbf{U}_2 = \begin{bmatrix} \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{bmatrix}, \quad (51)$$

what leads to two different definitions of the GFT, and as a consequence a signal may have distinct spectrum depending on which matrix (\mathbf{U}_1 or \mathbf{U}_2 in this example, or, in fact, any of the infinitely many linear combination possibilities) is used. In this case, for example, the signal $\mathbf{x} = (-4 \ 0 \ 4)^T$ would have spectra given by $\mathbf{U}_1^{-1} \mathbf{x} \approx (0 \ 5.66 \ 0)^T$ and $\mathbf{U}_2^{-1} \mathbf{x} \approx (0 \ 4.9 \ 2.83)^T$. In order to approach this problem, Deri and Moura suggest the use of oblique spectral projectors to obtain a unique GFT representation [55].

It matters to point out that, although GSP_L considers only undirected graphs, the framework does not forbid *a priori* the application to directed graphs, since the Laplacian matrix is still defined by $\mathbf{L} = \mathbf{D} - \mathbf{A}$, provided that \mathbf{D} is well defined as the *indegree* or *outdegree* matrix. For example, the notion of frequency as being the eigenvalues of \mathbf{L} could still be valid, for the norm $\|\mathbf{L}\mathbf{u}_i\|$ varies linearly with the *magnitude* of each eigenvalue,

$$\|\mathbf{L}\mathbf{u}_i\| = \|\gamma_i \mathbf{u}_i\| = |\gamma_i|, \quad (52)$$

as illustrated in Fig. 14. However, the Laplacian matrix for directed graphs is no more symmetric, and therefore it is not guaranteed to be diagonalizable or positive semidefinite.

Many other problems in GSP remain open to investigation, whether it is to develop schemes to properly apply GSP in the context of large networks, or to better understand and improve techniques applied to scenarios such as the ones mentioned in Section V. For example, the Uncertainty and Sampling theorems for graph signals (which relate to signal compression and recovery) have been extensively studied [56]–[58], but there is certainly room for further analysis, for instance on the robustness of non-perfectly bandlimited signal sampling [59].

If it is true the statement of Hilbert, that “as long as a branch of science offers an abundance of problems, so long is it alive”, then we may certainly expect the field of GSP to be living still for many years to come.

REFERENCES

- [1] K. M. Alam, M. Saini, and A. El Saddik, “Toward social Internet of vehicles: Concept, architecture, and applications,” *IEEE Access*, vol. 3, pp. 343–357, 2015, DOI: 10.1109/ACCESS.2015.2416657.
- [2] L. Guo, Z. Ning, Q. Song, L. Zhang, and A. Jamalipour, “A QoS-oriented high-efficiency resource allocation scheme in wireless multimedia sensor networks,” *IEEE Sensors J.*, 2016, DOI: 10.1109/JSEN.2016.2645709.
- [3] F. Ma, B. Yao, and M. Yao, “Non-planar unclustered peterson graphs as scale-free models of the Internet of things,” in *Information Technology, Networking, Electronic and Automation Control Conference, IEEE*. IEEE, 2016, pp. 1040–1043.
- [4] L. Yu, L. Feng, C. Chen, T. Qiu, L. Li, and J. Wu, “A novel multi-feature representation of images for heterogeneous IoTs,” *IEEE Access*, vol. 4, pp. 6204–6215, 2016, DOI: 10.1109/ACCESS.2016.2607841.
- [5] F. Chung, “Graph theory in the information age,” *Notices of the AMS*, vol. 57, no. 6, pp. 726–732, 2010.
- [6] A. J. Golubski, E. E. Westlund, J. Vandermeer, and M. Pascual, “Ecological networks over the edge: Hypergraph trait-mediated indirect interaction (TMII) structure,” vol. 31, no. 5, pp. 344–354, may 2016, DOI: 10.1016/j.tree.2016.02.006.
- [7] G. I. McKinsey *et al.*, “Big data: the next frontier for innovation, competition and productivity,” *McKinsey Global Institute*, 2011.
- [8] R. K. Jain, J. M. F. Moura, and C. E. Kontokosta, “Big data+big cities: Graph signals of urban air pollution [exploratory SP],” *IEEE Signal Process. Mag.*, vol. 31, no. 5, pp. 130–136, 2014, DOI: 10.1109/MSP.2014.2330357.
- [9] J. Mei and J. M. F. Moura, “Signal processing on graphs: Causal modeling of unstructured data,” *IEEE Trans. Signal Process.*, 2016, DOI: 10.1109/TSP.2016.2634543.

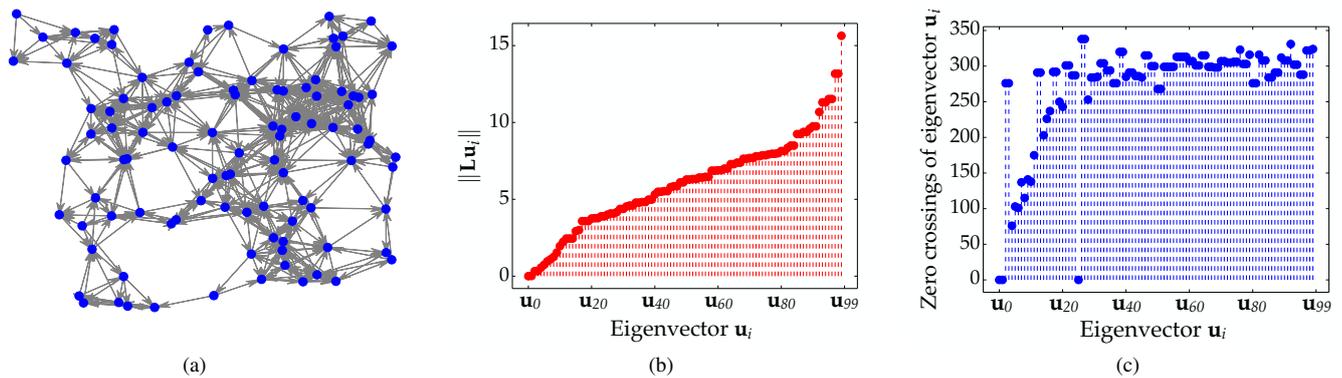


Figure 14: (a) Directed sensor graph and (b) its norm $\|Lu_i\|$ and (c) number of zero crossings of its eigenvectors, ordered so that the respective eigenvalues have ascending magnitude.

- [10] A. Sandryhaila and J. M. F. Moura, "Big data analysis with signal processing on graphs: Representation and processing of massive data sets with irregular structure," *IEEE Signal Process. Mag.*, vol. 31, no. 5, pp. 80–90, 2014, DOI: 10.1109/MSP.2014.2329213.
- [11] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. 83–98, 2013, DOI: 10.1109/MSP.2012.2235192.
- [12] P. Feofiloff, Y. Kohayakawa, and Y. Wakabayashi, "Uma Introdução Sucinta à Teoria dos Grafos," 2011, <https://www.ime.usp.br/~pf/teoriadosgrafos/>.
- [13] J. A. Bondy and U. S. R. Murty, "Graph theory, volume 244 of graduate texts in mathematics," 2008.
- [14] F. R. K. Chung, *Spectral graph theory*. American Mathematical Soc., 1997, no. 92.
- [15] A. Sandryhaila and J. M. F. Moura, "Nearest-neighbor image model," in *Image Processing (ICIP), 2012 19th IEEE International Conference on*. IEEE, 2012, pp. 2521–2524.
- [16] S. Sardellitti, S. Barbarossa, and P. D. Lorenzo, "Graph topology inference based on transform learning," in *2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*. IEEE, dec 2016, DOI: 10.1109/globalsip.2016.7905863.
- [17] M. Püschel and J. M. F. Moura, "Algebraic signal processing theory: Foundation and 1-D time," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3572–3585, 2008, DOI: 10.1109/TSP.2008.925261.
- [18] —, "Algebraic signal processing theory: 1-D space," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3586–3599, 2008, DOI: 10.1109/TSP.2008.925259.
- [19] A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs: Graph filters," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. Institute of Electrical and Electronics Engineers (IEEE), may 2013, DOI: 10.1109%2Ficassp.2013.6638849.
- [20] —, "Discrete signal processing on graphs," *IEEE Trans. Signal Process.*, vol. 61, no. 7, pp. 1644–1656, 2013, DOI: 10.1109/TSP.2013.2238935.
- [21] A. V. Oppenheim, A. S. Willsky, and S. H. Nawab, *Signals and Systems*, ser. Prentice-Hall signal processing series. Prentice Hall, 1997. [Online]. Available: <https://books.google.com.br/books?id=LwQqAQAAMAAJ>
- [22] A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs: Graph Fourier transform," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. Institute of Electrical and Electronics Engineers (IEEE), may 2013, DOI: 10.1109%2Ficassp.2013.6638850.
- [23] —, "Discrete signal processing on graphs: Frequency analysis," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3042–3054, jun 2014, DOI: 10.1109%2Ftsp.2014.2321121.
- [24] W. Rudin, *Real and complex analysis*. Tata McGraw-Hill Education, 1987.
- [25] S. Mallat, *A wavelet tour of signal processing*. Academic press, 1999.
- [26] L. B. Almeida, "The fractional Fourier transform and time-frequency representations," *IEEE Trans. Signal Process.*, vol. 42, no. 11, pp. 3084–3091, 1994, DOI: 10.1109/78.330368.
- [27] O. Akay and G. F. Boudreaux-Bartels, "Fractional convolution and correlation via operator methods and an application to detection of linear FM signals," *IEEE Trans. Signal Process.*, vol. 49, no. 5, pp. 979–993, 2001, DOI: 10.1109/78.917802.
- [28] J. B. Lima, R. M. C. de Souza, and D. C. Cunha, "Multiuser communication based on the discrete fractional Fourier transform," in *Communications (ICC), 2012 IEEE International Conference on*. IEEE, 2012, pp. 3569–3570.
- [29] H. M. Ozaktas, Z. Zalevsky, and M. A. Kutay, *The Fractional Fourier Transform: with Applications in Optics and Signal Processing*. Wiley, 2001.
- [30] T. Wang, H. Huan, R. Tao, and Y. Wang, "Security-coded OFDM system based on multiorder fractional Fourier transform," *IEEE Commun. Lett.*, vol. 20, no. 12, pp. 2474–2477, 2016, DOI: 10.1109/LCOMM.2016.2611498.
- [31] V. Namias, "The fractional order Fourier transform and its application to quantum mechanics," *IMA Journal of Applied Mathematics*, vol. 25, no. 3, pp. 241–265, 1980, DOI: 10.1093/imamat/25.3.241.
- [32] Y.-q. Wang, B.-z. Li, and Q.-y. Cheng, "The fractional Fourier transform on graphs," in *Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), 2017*. IEEE, 2017, pp. 105–110.
- [33] G. B. Ribeiro and J. B. Lima, "Deslocamento fracionário de sinais sobre grafos," in *XXXV Simpósio Brasileiro de Telecomunicações e Processamento de Sinais, SBTr17*. SBTr, 2017, pp. 871–875.
- [34] J. Gallier, "Spectral theory of unsigned and signed graphs – applications to graph clustering: a survey," December 2017. [Online]. Available: <http://www.cis.upenn.edu/~jean/spectral-graph-notes.pdf>
- [35] D. I. Shuman, B. Ricaud, and P. Vandergheynst, "A windowed graph Fourier transform," in *Statistical Signal Processing Workshop (SSP), 2012 IEEE*. Ieee, 2012, pp. 133–136.
- [36] D. K. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Applied and Computational Harmonic Analysis*, vol. 30, no. 2, pp. 129–150, 2011, DOI: 10.1016/j.acha.2010.04.005.
- [37] J. A. Deri and J. M. Moura, "New york city taxi analysis with graph signal processing," in *Signal and Information Processing (GlobalSIP), 2016 IEEE Global Conference on*. IEEE, 2016, pp. 1275–1279.
- [38] J. Nocedal and S. J. Wright, "Numerical optimization 2nd," 2006.
- [39] E. Jones, T. Oliphant, P. Peterson *et al.*, "SciPy: Open source scientific tools for Python," 2001–, [Online]. Available: <http://www.scipy.org/>
- [40] G. Fracastoro, S. M. Fosson, and E. Magli, "Steerable discrete cosine transform," *IEEE Trans. Image Process.*, vol. 26, no. 1, pp. 303–314, 2017, DOI: 10.1109/TIP.2016.2623489.
- [41] X. Su, M. Rizkallah, T. Maugey, and C. Guillemot, "Graph-based light fields representation and coding using geometry information," in *IEEE International Conference on Image Processing (ICIP), 2017*.
- [42] V. R. M. Elias and W. A. Martins, "Graph Fourier transform for light field compression," in *XXXV Simpósio Brasileiro de Telecomunicações e Processamento de Sinais, SBTr17*. SBTr, 2017, pp. 881–885.

- [43] M. Rossi and P. Frossard, "Graph-based light field super-resolution," in *Multimedia Signal Processing (MMSP), 2017 IEEE 19th International Workshop on*. IEEE, 2017, pp. 1–6.
- [44] G. Cheung, E. Magli, Y. Tanaka, and M. K. Ng, "Graph spectral image processing," *Proc. IEEE*, pp. 1–24, 2018, DOI: 10.1109/jproc.2018.2799702.
- [45] A. Pirayre, C. Couprie, F. Bidard, L. Duval, and J.-C. Pesquet, "Brane cut: biologically-related a priori network enhancement with graph cuts for gene regulatory network inference," *BMC Bioinformatics*, vol. 16, no. 1, p. 368, 2015, DOI: 10.1186/s12859-015-0754-2.
- [46] A. Pirayre, C. Couprie, L. Duval, and J.-C. Pesquet, "Brane clust: Cluster-assisted gene regulatory network inference refinement," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2017, DOI: 10.1109/TCBB.2017.2688355.
- [47] H. Q. Nguyen, P. A. Chou, and Y. Chen, "Compression of human body sequences using graph wavelet filter banks," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 6152–6156.
- [48] L. Goldsberry, W. Huang, N. F. Wymbs, S. T. Grafton, D. S. Bassett, and A. Ribeiro, "Brain signal analytics from graph signal processing perspective," in *Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on*. IEEE, 2017, pp. 851–855.
- [49] N. Leonardi and D. Van De Ville, "Tight wavelet frames on multislice graphs," *IEEE Trans. Signal Process.*, vol. 61, no. 13, pp. 3357–3367, 2013, DOI: 10.1109/TSP.2013.2259825.
- [50] W. Huang, L. Goldsberry, N. F. Wymbs, S. T. Grafton, D. S. Bassett, and A. Ribeiro, "Graph frequency analysis of brain signals," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 7, pp. 1189–1203, 2016, DOI: 10.1109/JSTSP.2016.2600859.
- [51] S. Chen, F. Cerda, P. Rizzo, J. Bielak, J. H. Garrett, and J. Kovačević, "Semi-supervised multiresolution classification using adaptive graph filtering with application to indirect bridge structural health monitoring," *IEEE Trans. Signal Process.*, vol. 62, no. 11, pp. 2879–2893, 2014, DOI: 10.1109/TSP.2014.2313528.
- [52] K. Benzi, V. Kalofolias, X. Bresson, and P. Vandergheynst, "Song recommendation with non-negative matrix factorization and graph total variation," in *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on*. IEEE, 2016, pp. 2439–2443.
- [53] N. Tremblay and P. Borgnat, "Graph wavelets for multiscale community mining," *IEEE Trans. Signal Process.*, vol. 62, no. 20, pp. 5227–5239, 2014, DOI: 10.1109/TSP.2014.2345355.
- [54] S. Sardellitti, S. Barbarossa, and P. Di Lorenzo, "On the graph Fourier transform for directed graphs," *IEEE J. Sel. Topics Signal Process.*, 2017, DOI: 10.1109/JSTSP.2017.2726979.
- [55] J. A. Deri and J. M. Moura, "Spectral projector-based graph Fourier transforms," *IEEE J. Sel. Topics Signal Process.*, vol. 11, no. 6, pp. 785–795, 2017, DOI: 10.1109/JSTSP.2017.2731599.
- [56] L. F. O. Chamon and A. Ribeiro, "Greedy sampling of graph signals," *IEEE Trans. Signal Process.*, vol. 66, no. 1, pp. 34–47, jan 2018, DOI: 10.1109/tsp.2017.2755586.
- [57] S. Chen, R. Varma, A. Singh, and J. Kovačević, "Signal recovery on graphs: Fundamental limits of sampling strategies," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 2, no. 4, pp. 539–554, 2016, DOI: 10.1109/TSIPN.2016.2614903.
- [58] X. Wang, J. Chen, and Y. Gu, "Generalized graph signal sampling and reconstruction," in *Signal and Information Processing (GlobalSIP), 2015 IEEE Global Conference on*. IEEE, 2015, pp. 567–571.
- [59] M. Tsitsvero, S. Barbarossa, and P. Di Lorenzo, "Signals on graphs: Uncertainty principle and sampling," *IEEE Trans. Signal Process.*, vol. 64, no. 18, pp. 4845–4860, 2016, DOI: 10.1109/TSP.2016.2573748.



Guilherme Boaviagem Ribeiro received the B. Sc. degree in Electronic Engineering and the M. Sc. degree in Electrical Engineering both from the Federal University of Pernambuco (UFPE), in 2016 and 2018, respectively. He is currently a doctoral student at UFPE in the field of graph signal processing and his interests also include algebra in finite fields, bioinformatics and language processing.



Juliano Bandeira Lima was born in Brazil where he studied Electrical Engineering. He received the M.Sc. degree and the Ph.D. degree from the Federal University of Pernambuco, Brazil, in 2004 and 2008, respectively. He is currently an Assistant Professor of the Department of Electronics and Systems at Federal University of Pernambuco. His main research interests are in the field of discrete and number-theoretic transforms, and their applications in digital signal processing, communications and cryptography.