Uncertainty Principles and Sparse Eigenvectors of Graphs

Oguzhan Teke, *Student Member, IEEE*, and P. P. Vaidyanathan, *Fellow, IEEE*

***Abstract*—Analysis of signals defined over graphs has been of interest in the recent years. In this regard, many concepts from the classical signal processing theory have been extended to the graph case, including uncertainty principles that study the concentration of a signal on a graph and in its graph Fourier basis (GFB). This paper advances a new way to formulate the uncertainty principle for signals defined over graphs, by using a nonlocal measure based on the notion of sparsity. To be specific, the total number of nonzero elements of a graph signal and its corresponding graph Fourier transform (GFT) is considered. A theoretical lower bound for this total number is derived, and it is shown that a nonzero graph signal and its GFT cannot be arbitrarily sparse simultaneously. When the graph has repeated eigenvalues, the GFB is not unique. Since the derived lower bound depends on the selected GFB, a method that constructs a GFB with the minimal uncertainty bound is provided. In order to find signals that achieve the derived lower bound (i.e., the most compact on the graph and in the GFB), sparse eigenvectors of the graph are investigated. It is shown that a connected graph has a 2-sparse eigenvector (of the graph Laplacian) when there exist two nodes with the same neighbors. In this case, the uncertainty bound is very low, tight, and independent of the global structure of the graph. For several examples of classical and real-world graphs, it is shown that 2-sparse eigenvectors, in fact, exist.**

***Index Terms*—Graph signals, graph Fourier basis, sparsity, uncertainty principles, sparse eigenvectors.**

I. Introduction

SIGNALS defined over a graph are useful to express high-dimensional data where the graph models the underlying dependency structure between the data sources. In this framework a graph signal is considered as a set of data points indexed according to vertices of the graph. This type of signal structure is not limited to electrical engineering and can be found in a variety of different contexts such as social, economic, and biological networks, among others [1].

In some of the recent developments, processing of graph signals has been based on the so-called "graph operator." However, the definition of this operator is not fixed. Motivated by spectral graph theory, the study in [2] selects the graph operator to be the (normalized) graph Laplacian, whereas studies in [3], [4] focus on the adjacency matrix. There are other proposals as well [5], [6]. Once the graph operator has been selected, its eigenvectors are used to define the graph Fourier basis (GFB). Representation of a signal in the selected graph Fourier basis is then used to define the graph Fourier transform (GFT) of the signal. Inspired by the constructions in [2]–[4], sampling, reconstruction and multirate processing of graph signals are studied in [7]–[14].

An essential concept in signal analysis is the uncertainty principle, which states that a signal cannot be arbitrarily localized in both time and frequency simultaneously [15]. In classical signal processing, the uncertainty principle is useful to design filters that are maximally localized in time, for a given frequency spread, or vice versa. Due to its importance, some authors extended this principle to signals defined on graphs [16]–[19]. Details of these approaches are elaborated in Section I-C.

Similar to [16]–[19], this paper studies the concept of uncertainty for graph signals. However, unlike earlier methods, and motivated by [20]–[22], we introduce a non-local measure for uncertainty, based on the notion of sparsity of the signal in the graph domain and frequency domain. We show that a nonzero graph signal and its corresponding GFT cannot be arbitrarily sparse simultaneously, and we provide a lower bound for the total number of nonzero elements. We further provide the optimal selection of the GFB (that minimizes the uncertainty bound) when the graph operator has repeated eigenvalues. In order to find signals that achieve the derived lower bound, we consider sparse eigenvectors of the graph operator. A detailed outline and the contributions of this paper are given below.

A. Outline and Contributions

Broadly speaking, results presented here can be divided into three main parts. In the first part (Sections II and III), we propose *discrete* and *non-local* uncertainty principles that depend on the max-norm of the graph Fourier basis. These results follow from more general theory of sparse representations studied in [20]–[22]. We consider the identity matrix and the GFB as a pair of bases to represent a graph signal and interpret the methodologies in [22] in the context of graph signal processing. Similarly and independently, the study in [19] has obtained very similar interpretations based on the same theory. Apart from the overlapping results [19] generalizes these results to the case of frames. Then, it focuses on local uncertainty principles where bounds depend on a particular portion of the graph. More detailed comparison with [19] is provided in Section I-C.

Manuscript received August 2, 2016; revised January 18, 2017, May 22, 2017, and July 8, 2017; accepted July 13, 2017. Date of publication July 24, 2017; date of current version August 10, 2017. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Pierre Borgnat. This work was supported in part by the NSF under Grant CCF-1712633, in part by the ONR Grant N00014-15-1-2118, and in part by the Electrical Engineering Carver Mead Research Seed Fund of the California Institute of Technology. *(Corresponding author: Oguzhan Teke.)*

The authors are with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125 USA (e-mail: oteke@caltech.edu; ppvnath@systems.caltech.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2017.2731299

Our main contributions are presented in the remaining sections: in the second part of our results (Section IV) we discuss that given a graph, max-norm of the graph Fourier basis is not unique in the presence of the repeated eigenvalues of the selected graph operator. Since this non-uniqueness greatly affects the interpretation of the relations between the graph structure and the uncertainty, we formulate a problem to select eigenvectors (GFB) such that the max-norm of the graph Fourier basis is maximized (or, equivalently the uncertainty bounds considered in Section II are minimized). We solve this problem analytically (Theorem 6) and provide an algorithmic routine to obtain a set of eigenvectors that gives a graph Fourier basis with the maximum possible max-norm.

In the third part of our results (Section V) we focus on the sparse eigenvectors of graphs in order to find signals that achieve the proposed uncertainty bounds. We study the relation between sparsity of eigenvectors and graph topology, and their effects on the max-norm of the GFB. We provide the necessary and sufficient conditions for the existence of 1-sparse and 2-sparse eigenvectors of the graph Laplacian (Theorems 7 and 8). We then show that existence of a 2-sparse eigenvector implies *very low and attainable uncertainty bounds* (Theorem 9). Finally in Section VI we apply our results to real world graph examples. Interestingly, uncertainty bounds for these graphs are very low. We precisely explain why this is the case, and find the signals that achieve these bounds. Portions and some extensions of these results have been presented in [23], [24].

B. Graph Signal Processing and Notation

Let $\boldsymbol{x} \in \mathcal{C}^N$ be a graph signal on a graph of size N (i.e., N nodes or vertices) whose adjacency matrix is denoted as $\boldsymbol{A} \in \mathcal{R}^{N \times N}$. We assume the graph does not have self loops, i.e. $a_{i,i} = 0$. The weight of the edge from the j^{th} node to the i^{th} node is denoted by the $(i,j)^{th}$ element of $\boldsymbol{A}$. This definition of the adjacency matrix follows from [3], [4] and is the reverse of the usual definition in graph theory. Unless it is specified, we consider the general case of directed graphs (i.e., $a_{i,j} \neq a_{j,i}$). For the restricted case of undirected graphs with non-negative edge weights ($a_{i,j} = a_{j,i} \geq 0$), the graph Laplacian is defined as $\boldsymbol{L} = \boldsymbol{D} - \boldsymbol{A}$ where $\boldsymbol{D}$ is the diagonal degree matrix given as $(\boldsymbol{D})_{i,i} = \sum_j a_{i,j}$. Normalized graph Laplacian is defined as $\mathcal{L} = \boldsymbol{D}^{-1/2} \boldsymbol{L} \boldsymbol{D}^{-1/2}$. We will use $\boldsymbol{V}$ to denote the graph Fourier basis (GFB) of interest. Here the definition of $\boldsymbol{V}$ is not fixed. Eigenvectors of either the graph Laplacian or the adjacency matrix itself can be set to be the graph Fourier domain. We will use $\boldsymbol{V}_A$ and $\boldsymbol{V}_L$ to explicitly denote the eigenvectors of the adjacency matrix and the graph Laplacian, respectively. Assuming $\boldsymbol{A}$ is diagonalizable, we precisely have:

$$\boldsymbol{A} = \boldsymbol{V}_A \, \boldsymbol{\Lambda}_A \, \boldsymbol{V}_A^{-1}, \qquad \boldsymbol{L} = \boldsymbol{V}_L \, \boldsymbol{\Lambda}_L \, \boldsymbol{V}_L^{-1}, \quad (1)$$

where $\boldsymbol{\Lambda}_A$ and $\boldsymbol{\Lambda}_L$ are diagonal eigenvalue matrices. We will always assume that eigenvectors are normalized to have a unit ℓ_2-norm. When the graph is undirected $\boldsymbol{V}_A$ and $\boldsymbol{V}_L$ can be selected to be unitary. For the selected graph Fourier basis $\boldsymbol{V}$, graph Fourier transform (GFT) of a signal $\boldsymbol{x}$ is given by $\widehat{\boldsymbol{x}} = \boldsymbol{F}\,\boldsymbol{x}$ where $\boldsymbol{F} = \boldsymbol{V}^{-1}$.

In the following, the ℓ_0 pseudo-norm $\|\boldsymbol{x}\|_0$ denotes the number of nonzero elements in the signal $\boldsymbol{x}$. The ℓ_p norm of a vector will be denoted by $\|\boldsymbol{x}\|_p$. The spectral norm (largest singular value) of the matrix $\boldsymbol{A}$ will be denoted by $\|\boldsymbol{A}\|_2$. We will use $\|\boldsymbol{A}\|_{\max}$ to denote the maximum absolute value of elements in the matrix $\boldsymbol{A}$. Precisely, we have:

$$\|\boldsymbol{A}\|_2 = \max_{\boldsymbol{x} \neq \boldsymbol{0}} \|\boldsymbol{A}\boldsymbol{x}\|_2 \,/\, \|\boldsymbol{x}\|_2, \qquad \|\boldsymbol{A}\|_{\max} = \max_{i,j} |a_{i,j}|. \quad (2)$$

Notice that $\|\boldsymbol{A}\|_{\max}$ is equivalent to the mutual coherence between the matrix $\boldsymbol{A}$ and the identity matrix [22].

For a matrix $\boldsymbol{A}$ and two index sets S and K, $\boldsymbol{A}_S$ denotes the matrix with columns of $\boldsymbol{A}$ indexed by S and $\boldsymbol{A}_{K,S}$ denotes the matrix with columns indexed by S and rows indexed by K. For a vector $\boldsymbol{x}$, elements of $\boldsymbol{x}$ indexed by S will be denoted by $\boldsymbol{x}_S$. The index set $\bar{S}$ is defined as $\bar{S} = \{1, \ldots, N\} \setminus S$, where $\setminus$ stands for the set difference operator. We will use $\otimes$ to denote Kronecker product of matrices. Dimension of the right null space of a matrix $\boldsymbol{A}$ is denoted by $nullity(\boldsymbol{A})$.

C. Related Work

To the best to our knowledge, there are mainly four studies that consider uncertainty principles for signals defined over graphs [16]–[19]. The main theme in these studies (including this one) is to define a "measure" of the signal in the vertex domain and the graph Fourier domain. In particular, the study in [16] considers the following for unweighted graphs:

$$\Delta^2_{g,u_0}(\boldsymbol{x}) = \frac{\boldsymbol{x}^H \, \boldsymbol{P}^2_{u_0} \, \boldsymbol{x}}{\|\boldsymbol{x}\|_2^2}, \qquad \Delta^2_s(\boldsymbol{x}) = \frac{\boldsymbol{x}^H \, \mathcal{L} \, \boldsymbol{x}}{\|\boldsymbol{x}\|_2^2}, \quad (3)$$

where $\Delta^2_{g,u_0}(\boldsymbol{x})$ is referred to as the vertex spread (with $\boldsymbol{P}_{u_0}$ being the diagonal distance matrix with respect to the node u_0), and $\Delta^2_s(\boldsymbol{x})$ is referred to as the spectral spread of the signal $\boldsymbol{x}$. This approach is extended to weighted graphs in [25]. The study in [17] works on an alternative definition and focuses on the following:

$$\alpha = \|\boldsymbol{D}_S \, \boldsymbol{x}\|_2 \,/\, \|\boldsymbol{x}\|_2, \qquad \beta = \|\boldsymbol{B}_F \, \boldsymbol{x}\|_2 \,/\, \|\boldsymbol{x}\|_2, \quad (4)$$

where α^2 and β^2 represent the amount of energy confined in the vertex set S and the frequency set F (with $\boldsymbol{D}_S$ and $\boldsymbol{B}_F$ being corresponding projection matrices), respectively. The study in [18], [26] considers the smoothness of the signal in both domains: using the difference operator on the graph, $\boldsymbol{D}_r$, the interplay between $\|\boldsymbol{D}_r \, \boldsymbol{x}\|_2^2$ and $\|\boldsymbol{D}_r \, \widehat{\boldsymbol{x}}\|_2^2$ is studied.

The main observation of these studies is that measures in both domains cannot be arbitrarily small simultaneously: a signal "limited" in one domain cannot be "limited" in the other domain. These trade-offs are then considered as uncertainty principles. Depending on their corresponding definitions, they characterize these uncertainty curves theoretically and study the signals that achieve them with equality.

More recently, the study in [19] takes a non-local perspective where vertex and spectral measures are defined with ℓ_p norms. Using the fact that the identity matrix and the GFB form a pair of bases to represent graph signals, [19] proposes some uncertainty principles where the results are based on theory of sparse representations studied in [20]–[22], [27]. These results show

that the "graph Fourier coherence" is a fundamental quantity for the uncertainty principles of interest. Notice that we specifically consider the case of ℓ_0 and ℓ_1 in this study and obtain very similar results where we use the term "max-norm of GFB" instead of coherence. This overlap is a direct consequence of the influence of the sparse representation theory in [20]–[22]. In the rest, [19] focuses on the representation of graph signals using frames and proposes some uncertainty results based on these representations. It later considers local uncertainty principles where bounds depend on a region of the graph.

In order to find signals that achieve the proposed uncertainty bounds we consider sparse eigenvectors (2-sparse in particular) of the graph Laplacian. In this context, the study in [28] reveals a specific graph structure (referred to as motif-doubling) that results in sparse eigenvectors of both the adjacency matrix and the graph Laplacian. The structure of motif-doubling gives rise to sparse eigenvectors with even number of non-zero entries. However, this structure is only sufficient to have sparse eigenvectors, whereas our condition on 2-sparse eigenvectors is necessary and sufficient. A detailed comparison with [28] will be provided in Section V-B.

II. Discrete Non-Local Spreads

Inspired by [20]–[22], we will study the concept of uncertainty from a *discrete* and *non-local* perspective. We will consider graph spread of a signal as the total number of nonzero elements in the signal. Similarly, spectral spread of a signal will be defined as the number of nonzero elements in the GFT of the signal. Hence, we have the following definitions:

Definition 1 (ℓ_0-based spread on vertex domain): Given a nonzero signal $\boldsymbol{x}$ on a graph with GFT $\boldsymbol{F}$, the "spread" of the signal on vertex domain is defined as $\|\boldsymbol{x}\|_0$. ♢

Definition 2 (ℓ_0-based spread on Fourier domain): Given a nonzero signal $\boldsymbol{x}$ on a graph with GFT $\boldsymbol{F}$, the "spread" of the signal in the Fourier domain is defined as $\|\boldsymbol{F}\boldsymbol{x}\|_0$. ♢

The definition of the spectral spread depends on the selected GFT $\boldsymbol{F}$. Whether it is based on the adjacency matrix, the graph Laplacian, or something else, for a given graph, the GFB $\boldsymbol{V}$, hence $\boldsymbol{F} = \boldsymbol{V}^{-1}$, may *not* be unique. This is due to the fact that the selected graph operator may have *repeated eigenvalues*. In such a case, one can select different bases to span the corresponding eigenspaces resulting in different GFB matrices. In order to avoid this ambiguity we assume that, not just the graph itself, but also the associated GFT is given in Definition 2. It should be noted that in the case of repeated eigenvalues selection of the GFB is not a simple task and it requires attention. We will address this problem in Sections IV and V, where we discuss optimal selection of the GFB in order to minimize the spectral spread of signals.

For a nonzero graph signal notice that its vertex domain spread and spectral spread have to be at least 1, however, they may *not* achieve this bound simultaneously. As a simple motivational example, consider the directed cyclic graph of N vertices, whose graph Fourier Transform corresponds to DFT of size N [4]. If the signal $\boldsymbol{x}$ is an impulse, then $\|\boldsymbol{x}\|_0 = 1$, but $\|\boldsymbol{F}\boldsymbol{x}\|_0 = N$. On the contrary, if the signal is a constant, then $\|\boldsymbol{x}\|_0 = N$, but $\|\boldsymbol{F}\boldsymbol{x}\|_0 = 1$. As a result we ask the following question: *Given the GFT $\boldsymbol{F}$, what is the minimum total number of nonzero elements in a graph signal and its corresponding Fourier Transform?* For this purpose, we consider the following two definitions of uncertainty:

$$s_0(\boldsymbol{x}) = \big(\|\boldsymbol{x}\|_0 + \|\boldsymbol{F}\boldsymbol{x}\|_0\big)/2, \quad p_0(\boldsymbol{x}) = \sqrt{\|\boldsymbol{x}\|_0 \, \|\boldsymbol{F}\boldsymbol{x}\|_0}, \quad (5)$$

where $s_0(\boldsymbol{x})$ and $p_0(\boldsymbol{x})$ are referred to as *additive* and *multiplicative* uncertainty of the signal $\boldsymbol{x}$, respectively. The definitions in (5) have the following two important properties: 1) They are scale invariant, that is $s_0(\alpha\,\boldsymbol{x}) = s_0(\boldsymbol{x})$ and $p_0(\alpha\,\boldsymbol{x}) = p_0(\boldsymbol{x})$ for all $\alpha \neq 0$. This is a useful property since uncertainty of a signal is expected to be scale-invariant. 2) Both can take only a discrete, finite set of values, namely, $s_0(\boldsymbol{x})$ takes half integer values (i.e., $k/2$ for integer k) and $p_0(\boldsymbol{x})$ takes values in the form of $p_0(\boldsymbol{x}) = \sqrt{k}$ for some integer $k \geq 1$. As a final remark, notice that the AM-GM inequality dictates the following relation:

$$s_0(\boldsymbol{x}) \geq p_0(\boldsymbol{x}), \qquad \forall \boldsymbol{x} \in \mathcal{C}^N, \quad (6)$$

with equality if and only if $\|\boldsymbol{x}\|_0 = \|\boldsymbol{F}\boldsymbol{x}\|_0$.

The main purpose of this section is to find the lowest value that $s_0(\boldsymbol{x})$ can attain. More precisely the following problem will be considered:

$$s_0^\star = \min_{\boldsymbol{x} \neq \boldsymbol{0}} s_0(\boldsymbol{x}), \quad (7)$$

where $s_0^\star$ is called as the uncertainty bound since a signal and its corresponding graph Fourier Transform cannot be arbitrarily sparse simultaneously, that is, $s_0(\boldsymbol{x}) \geq s_0^\star \quad \forall \boldsymbol{x} \neq \boldsymbol{0}$.

Due to the combinatoric nature of the problem in (7), no closed form solution for $s_0^\star$ is available for an arbitrary $\boldsymbol{F}$. Nevertheless, the theory of sparse representations provides useful bounds for the problem. Motivated by Theorem 2.1 in [22], the following theorem provides a lower bound for $p_0(\boldsymbol{x})$:

Theorem 1 (Multiplicative uncertainty principle): For a graph with GFT $\boldsymbol{F}$, the multiplicative uncertainty of a nonzero signal $\boldsymbol{x}$ is lower bounded as follows:

$$p_0(\boldsymbol{x}) \geq \Big(\|\boldsymbol{F}^{-1}\|_2 \, \|\boldsymbol{F}\|_{\max}\Big)^{-1}, \quad (8)$$

where $\|\cdot\|_2$ and $\|\cdot\|_{\max}$ are defined in (2). ♢

Proof: First notice that $\boldsymbol{x} = \sum_{i=1}^N x_i \, \boldsymbol{e}_i$ where $\boldsymbol{e}_i$ is the i^{th} vector of the canonical basis, and x_i is the i^{th} element of $\boldsymbol{x}$. Let $\boldsymbol{f}_j^H$ denote the j^{th} row of $\boldsymbol{F}$. For the j^{th} element of the GFT of $\boldsymbol{x}$ we have

$$|\widehat{x}_j| = |\boldsymbol{f}_j^H \, \boldsymbol{x}| = \left| \sum_{i \in S} x_i \, \boldsymbol{f}_j^H \, \boldsymbol{e}_i \right|, \quad (9)$$

where S denotes the support (set of nonzero indices) of the signal $\boldsymbol{x}$. Notice that $|S| = \|\boldsymbol{x}\|_0$. We can upper bound $|\widehat{x}_j|$ as

$$|\widehat{x}_j| = \left| \sum_{i \in S} x_i \, \boldsymbol{f}_j^H \, \boldsymbol{e}_i \right| \leq \sum_{i \in S} \big| x_i \, \boldsymbol{f}_j^H \, \boldsymbol{e}_i \big|, \quad (10)$$

using the triangular inequality. Using Cauchy-Schwarz inequality, this can be further bounded as

$$|\widehat{x}_j| \leq \left(\sum_{i\in S}|x_i|^2\right)^{1/2}\left(\sum_{i\in S}|\boldsymbol{f}_j^H\boldsymbol{e}_i|^2\right)^{1/2}, \quad (11)$$

$$\leq \|\boldsymbol{x}\|_2\left(|S|\,\|\boldsymbol{F}\|_{\max}^2\right)^{1/2} = \|\boldsymbol{x}\|_2\,\|\boldsymbol{x}\|_0^{1/2}\,\|\boldsymbol{F}\|_{\max}, \quad (12)$$

where we use the fact that $\boldsymbol{f}_j^H\,\boldsymbol{e}_i$ is the $(j,i)^{th}$ element of $\boldsymbol{F}$ whose magnitude is upper bounded by $\|\boldsymbol{F}\|_{\max}$. Now consider the ℓ_2 norm of $\widehat{\boldsymbol{x}} = \boldsymbol{F}\boldsymbol{x}$:

$$\|\widehat{\boldsymbol{x}}\|_2^2 = \sum_{j\in K}|\widehat{x}_j|^2 \leq \|\widehat{\boldsymbol{x}}\|_0\,\|\boldsymbol{x}\|_2^2\,\|\boldsymbol{x}\|_0\,\|\boldsymbol{F}\|_{\max}^2, \quad (13)$$

where K denotes the support of $\widehat{\boldsymbol{x}}$, $\|\widehat{\boldsymbol{x}}\|_0 = |K|$. Hence we have,

$$\frac{1}{\|\widehat{\boldsymbol{x}}\|_0\,\|\boldsymbol{x}\|_0} \leq \left(\|\boldsymbol{F}\|_{\max}\frac{\|\boldsymbol{x}\|_2}{\|\widehat{\boldsymbol{x}}\|_2}\right)^2. \quad (14)$$

Notice that $\max_{\boldsymbol{x}}\|\boldsymbol{x}\|_2/\|\widehat{\boldsymbol{x}}\|_2 = \max_{\boldsymbol{y}}\|\boldsymbol{F}^{\mathrm{I}}\boldsymbol{y}\|_2/\|\boldsymbol{y}\|_2 = \|\boldsymbol{F}^{-1}\|_2$. Therefore,

$$\frac{1}{\|\widehat{\boldsymbol{x}}\|_0\,\|\boldsymbol{x}\|_0} \leq \left(\|\boldsymbol{F}\|_{\max}\frac{\|\boldsymbol{x}\|_2}{\|\widehat{\boldsymbol{x}}\|_2}\right)^2 \leq \left(\|\boldsymbol{F}\|_{\max}\,\|\boldsymbol{F}^{-1}\|_2\right)^2 \quad (15)$$

which implies $\sqrt{\|\boldsymbol{x}\|_0\,\|\boldsymbol{F}\boldsymbol{x}\|_0} \geq \left(\|\boldsymbol{F}^{-1}\|_2\,\|\boldsymbol{F}\|_{\max}\right)^{-1}$. ■

Corollary 1 (Additive uncertainty principle): For a graph with GFT $\boldsymbol{F}$, the additive uncertainty of a nonzero signal $\boldsymbol{x}$ is lower bounded as follows:

$$s_0(\boldsymbol{x}) \geq \left(\|\boldsymbol{F}^{-1}\|_2\,\|\boldsymbol{F}\|_{\max}\right)^{-1}, \quad (16)$$

where $\|\cdot\|_2$ and $\|\cdot\|_{\max}$ are defined in (2). ♢

Proof: From (6) we have $s_0(\boldsymbol{x}) \geq p_0(\boldsymbol{x})$. Therefore, any lower bound for $p_0(\boldsymbol{x})$ is also a lower bound for $s_0(\boldsymbol{x})$. ■

When the GFT of interest is unitary, Theorem 1 and Corollary 1 reduce to the following corollaries:

Corollary 2 (Weak ℓ_0 uncertainty): For any graph with *unitary* graph Fourier basis $\boldsymbol{V}$,

$$p_0(\boldsymbol{x}) \geq \|\boldsymbol{V}\|_{\max}^{-1}. \quad (17)$$

♢

Proof: Let $\boldsymbol{V}$ be unitary in Theorem 1. Then we have $\boldsymbol{F} = \boldsymbol{V}^H$, hence $\|\boldsymbol{F}\|_{\max} = \|\boldsymbol{V}\|_{\max}$. Furthermore, $\|\boldsymbol{V}\|_2 = \|\boldsymbol{F}^{-1}\|_2 = 1$. ■

This corollary is important since for undirected graphs, the adjacency matrix and the graph Laplacian are symmetric which result in a unitary graph Fourier basis. The corollary also implies the following (from AM-GM inequality).

Corollary 3 (Strong ℓ_0 uncertainty): For any graph with *unitary* graph Fourier basis $\boldsymbol{V}$,

$$s_0(\boldsymbol{x}) \geq \|\boldsymbol{V}\|_{\max}^{-1}. \quad (18)$$

♢

In Section VI, we will provide graph examples on which the inequality in (18) is satisfied with equality.

Even though (18) is a lower bound for the uncertainty, it does not say anything about the signal that achieves the bound. Furthermore, it does not say whether *there is* a signal that achieves the lower bound or not. In order to understand the existence of such signals, we provide the following result.

Theorem 2 (Existence of signals): Let $\boldsymbol{V}$ be a unitary Fourier basis of the graph. There exists a signal $\boldsymbol{x}$ on the graph that achieves the strong ℓ_0 uncertainty bound (satisfies (18) with equality) if and only if there exist index sets K and S with $|K| = |S| = \|\boldsymbol{V}\|_{\max}^{-1}$ such that $nullity\big((\boldsymbol{V}_{S,\bar{K}})^H\big) > 0$. Furthermore, a signal that achieves the bound is given as

$$\boldsymbol{x}_S \in null\Big((\boldsymbol{V}_{S,\bar{K}})^H\Big), \qquad \boldsymbol{x}_{\bar{S}} = \boldsymbol{0}. \quad (19)$$

♢

Proof: Assume that $\boldsymbol{x}$ achieves the strong ℓ_0 bound, that is,

$$s_0(\boldsymbol{x}) = \frac{\|\boldsymbol{x}\|_0 + \|\boldsymbol{F}\boldsymbol{x}\|_0}{2} = \|\boldsymbol{V}\|_{\max}^{-1}. \quad (20)$$

Then we have $s_0(\boldsymbol{x}) = p_0(\boldsymbol{x})$, which implies

$$\|\boldsymbol{x}\|_0 = \|\boldsymbol{F}\boldsymbol{x}\|_0 = \|\boldsymbol{V}\|_{\max}^{-1}, \quad (21)$$

since $s_0(\boldsymbol{x}) = p_0(\boldsymbol{x})$ if and only if $\|\boldsymbol{x}\|_0 = \|\boldsymbol{F}\boldsymbol{x}\|_0$ due to AM-GM inequality. Let S denote the support of $\boldsymbol{x}$ and K denote the support of $\widehat{\boldsymbol{x}}$. Then $\boldsymbol{F}\boldsymbol{x} = \boldsymbol{F}_S\,\boldsymbol{x}_S$. Since $\widehat{\boldsymbol{x}}$ is zero outside of its support we have $\boldsymbol{F}_{\bar{K},S}\,\boldsymbol{x}_S = \boldsymbol{0}$, which means that $nullity(\boldsymbol{F}_{\bar{K},S}) = nullity\big((\boldsymbol{V}_{S,\bar{K}})^H\big) > 0$ for some S and K with $|S| = |K| = \|\boldsymbol{V}\|_{\max}^{-1}$.

Conversely, assume that $nullity\big((\boldsymbol{V}_{S,\bar{K}})^H\big) > 0$ for some S and K with $|S| = |K| = \|\boldsymbol{V}\|_{\max}^{-1}$. Then select $\boldsymbol{x}_S$ as $\boldsymbol{x}_S \in null\big((\boldsymbol{V}_{S,\bar{K}})^H\big)$ and $\boldsymbol{x}_{\bar{S}}$ to be $\boldsymbol{0}$. Hence, $\|\boldsymbol{x}\|_0 = |S|$. Since S is the support of $\boldsymbol{x}$, we have $\boldsymbol{F}_S\,\boldsymbol{x}_S = \boldsymbol{F}\boldsymbol{x}$. Furthermore,

$$\|\boldsymbol{F}_S\,\boldsymbol{x}_S\|_0 = \|\boldsymbol{F}_{K,S}\,\boldsymbol{x}_S\|_0 + \|\boldsymbol{F}_{\bar{K},S}\,\boldsymbol{x}_S\|_0 = |K|, \quad (22)$$

since $\boldsymbol{x}_S \in null(\boldsymbol{F}_{\bar{K},S})$. As a result, $s_0(\boldsymbol{x}) = \|\boldsymbol{V}\|_{\max}^{-1}$. ■

It should be noted that $\|\boldsymbol{V}\|_{\max}^{-1}$ may not be an integer in general. In this case, we cannot find index sets of size $\|\boldsymbol{V}\|_{\max}^{-1}$ since the size of an index set is an integer. In this case, Theorem 2 tells us that there is no signal that achieves the bound in (18) with equality.

As an immediate example, the normalized inverse DFT matrix is a unitary graph Fourier basis for circulant graphs [29], [30]. Notice that the normalized DFT matrix of size N has $\|\boldsymbol{V}\|_{\max}^{-1} = \sqrt{N}$. Thus, circulant graphs have the following two results. 1) From Corollary 2, we have $p_0(\boldsymbol{x}) \geq \sqrt{N}$. This is a well-known uncertainty result given in [20], and the bound is known to be tight for all N. 2) From Corollary 3, we have $s_0(\boldsymbol{x}) \geq \sqrt{N}$. When N is a perfect square "picket fence" signal is known to achieve this bound [22]. Details of these results will be elaborated in Section VI-A1.

Even though Theorem 2 is useful to characterize signals that achieve the bound in (18), it has a major drawback in terms of practical usability. Finding the index sets requires a combinatorial search over all possible sets of size $\|\boldsymbol{V}\|_{\max}^{-1}$, which is *not* computationally efficient.

III. Uncertainty Based on ℓ_1 Norm

In the previous section, we defined the spread of a signal as the total number of nonzero elements in the signal. (see Definitions 1

and 2.) In this section we will consider a "smoother" measure by replacing the ℓ_0 with ℓ_1 norm.

Imitating (5), we can define an ℓ_1 based additive uncertainty as $s_1(\boldsymbol{x}) = (\|\boldsymbol{x}\|_1 + \|\boldsymbol{F}\boldsymbol{x}\|_1)/2$. With this definition we have $s_1(\alpha\,\boldsymbol{x}) = |\alpha| s_1(\boldsymbol{x})$, that is, $s_1(\boldsymbol{x})$ is *not* scale invariant. The problem is that a nonzero signal can have arbitrarily small uncertainty, which is an undesired property. One way to impose the scale invariance is to use a normalization as follows:

$$s_1(\boldsymbol{x}) = \frac{\|\boldsymbol{x}\|_1 + \|\boldsymbol{F}\boldsymbol{x}\|_1}{2\,\|\boldsymbol{x}\|_p}. \tag{23}$$

For any nonzero p, $s_1(\boldsymbol{x})$ in (23) has the property of $s_1(\alpha\boldsymbol{x}) = s_1(\boldsymbol{x})$ for $\alpha \neq 0$, hence it can be used as an uncertainty measure. However, it should be noted that characteristics of $s_1(\boldsymbol{x})$ depend on the selected ℓ_p norm. In the following, we will consider the case of $p = 2$, and define the ℓ_1 based uncertainty measures as follows:

$$s_1(\boldsymbol{x}) = \frac{\|\boldsymbol{x}\|_1 + \|\boldsymbol{F}\boldsymbol{x}\|_1}{2\,\|\boldsymbol{x}\|_2}, \quad p_1(\boldsymbol{x}) = \frac{\sqrt{\|\boldsymbol{x}\|_1\,\|\boldsymbol{F}\boldsymbol{x}\|_1}}{\|\boldsymbol{x}\|_2}, \tag{24}$$

where $s_1(\boldsymbol{x})$ and $p_1(\boldsymbol{x})$ are referred to as ℓ_1 based additive and multiplicative uncertainty of the signal $\boldsymbol{x}$, respectively.

The main reason for considering the case of $p = 2$ is that such a selection has strong connections with the ℓ_0 based uncertainty measures discussed in Section II. These relations will be elaborated at the end of this section (see Theorem 4).

As done in Section II, motivated by Theorem 2.1 in [22], a lower bound for $p_1(\boldsymbol{x})$ can be obtained as follows:

Theorem 3: For a graph with GFT $\boldsymbol{F}$, ℓ_1-based multiplicative uncertainty of a nonzero signal $\boldsymbol{x}$ is lower bounded as follows:

$$p_1(\boldsymbol{x}) \geq \left(\|\boldsymbol{F}^{-1}\|_2\,\|\boldsymbol{F}\|_{\max}^{1/2}\right)^{-1}, \tag{25}$$

where $\|\cdot\|_2$ and $\|\cdot\|_{\max}$ are defined in (2). ♢

The reader should carefully notice the presence of the square root in (25), which was not there in (8).

Proof: Notice that $\boldsymbol{F}$ is an invertible matrix, therefore $\|\widehat{\boldsymbol{x}}\|_2/\|\boldsymbol{x}\|_2$ is lower and upper bounded as follows

$$\|\boldsymbol{F}^{-1}\|_2^{-2} \leq \frac{\|\widehat{\boldsymbol{x}}\|_2^2}{\|\boldsymbol{x}\|_2^2} \leq \|\boldsymbol{F}\|_2^2. \tag{26}$$

Therefore we have:

$$\|\boldsymbol{F}^{-1}\|_2^{-2}\,\|\boldsymbol{x}\|_2^2 \leq \|\widehat{\boldsymbol{x}}\|_2^2 = \widehat{\boldsymbol{x}}^H\,\boldsymbol{F}\,\boldsymbol{x} = \sum_{i,j} \widehat{x}_i^*\,F_{i,j}\,x_j, \tag{27}$$

$$\leq \sum_{i,j} \left|\widehat{x}_i^* F_{i,j} x_j\right| \leq \sum_{i,j} |\widehat{x}_i|\|\boldsymbol{F}\|_{\max}|x_j|, \tag{28}$$

$$= \|\widehat{\boldsymbol{x}}\|_1\,\|\boldsymbol{x}\|_1\,\|\boldsymbol{F}\|_{\max}, \tag{29}$$

where we use the fact that $|F_{i,j}| \leq \|\boldsymbol{F}\|_{\max}$ for all (i,j) in (28). Notice that taking square-root of both sides and re-arranging the terms in (29) give the result in (25). ■

When the GFT of interest is unitary Theorem 3 reduces to the following corollary.

Corollary 4 (Weak ℓ_1 uncertainty): For any graph with *unitary* Fourier basis $\boldsymbol{V}$,

$$p_1(\boldsymbol{x}) \geq \|\boldsymbol{V}\|_{\max}^{-1/2}. \tag{30}$$

♢

Proof: In Theorem 3, when $\boldsymbol{V}$ is unitary, we have $\|\boldsymbol{V}\|_2 = \|\boldsymbol{F}^{-1}\|_2 = 1$. Furthermore, $\boldsymbol{F} = \boldsymbol{V}^H$, hence $\|\boldsymbol{F}\|_{\max} = \|\boldsymbol{V}\|_{\max}$. ■

We can finally provide a lower bound for the ℓ_1-based additive uncertainty as follows.

$$s_1(\boldsymbol{x}) \geq \|\boldsymbol{V}\|_{\max}^{-1/2}, \tag{31}$$

where the inequality follows from $s_1(\boldsymbol{x}) \geq p_1(\boldsymbol{x})$ (AM-GM inequality). Hence, any lower bound for $p_1(\boldsymbol{x})$ (Corollary 4) is a lower bound for $s_1(\boldsymbol{x})$.

It is quite interesting to observe that the term $\|\boldsymbol{V}\|_{\max}$ appears in the lower bound for both ℓ_0-based and ℓ_1-based uncertainty definitions. It should be noted that $1/\sqrt{N} \leq \|\boldsymbol{V}\|_{\max} \leq 1$ for any matrix $\boldsymbol{V}$ since $\boldsymbol{V}$ is assumed to have columns with unit ℓ_2 norm. Therefore, $\|\boldsymbol{V}\|_{\max}^{-1} \geq \|\boldsymbol{V}\|_{\max}^{-1/2}$ always holds true. This shows that the lower bound given by (31) is always less than the bound in (18). In fact, not just the bounds but ℓ_0 and ℓ_1 based uncertainties are also related with each other. The following theorem establishes this relation.

Theorem 4: Assume that graph of interest has a unitary GFB $\boldsymbol{V}$. Then, we have the following inequality

$$p_0(\boldsymbol{x}) \geq \Big(p_1(\boldsymbol{x})\Big)^2, \tag{32}$$

for all signals defined on the graph. ♢

Proof: We start with the following change of variables,

$$\Big(p_1(\boldsymbol{x})\Big)^2 = \frac{\|\boldsymbol{x}\|_1\,\|\boldsymbol{F}\boldsymbol{x}\|_1}{\|\boldsymbol{x}\|_2^2} = \|\bar{\boldsymbol{x}}\|_1\,\|\boldsymbol{F}\bar{\boldsymbol{x}}\|_1, \tag{33}$$

where we define $\bar{\boldsymbol{x}} = \boldsymbol{x}/\|\boldsymbol{x}\|_2$.

Given any two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ with $\|\boldsymbol{x}\|_2 = \|\boldsymbol{y}\|_2 = 1$, we have the following relation (page 20 of [22])

$$\|\boldsymbol{x}\|_1\,\|\boldsymbol{y}\|_1 \leq \sqrt{\|\boldsymbol{x}\|_0\,\|\boldsymbol{y}\|_0}. \tag{34}$$

Notice that we have $\|\bar{\boldsymbol{x}}\|_2 = 1$. Since we have assumed that the GFB is unitary, we further have $\|\boldsymbol{F}\bar{\boldsymbol{x}}\|_2 = 1$. Then (34) gives the following

$$\|\bar{\boldsymbol{x}}\|_1\,\|\boldsymbol{F}\bar{\boldsymbol{x}}\|_1 \leq \sqrt{\|\bar{\boldsymbol{x}}\|_0\,\|\boldsymbol{F}\bar{\boldsymbol{x}}\|_0}. \tag{35}$$

Notice that left-hand-side of (35) is equal to $(p_1(\boldsymbol{x}))^2$ due to (33). Remember that $p_0(\boldsymbol{x})$ is scale-invariant. As a result we have $\sqrt{\|\bar{\boldsymbol{x}}\|_0\,\|\boldsymbol{F}\bar{\boldsymbol{x}}\|_0} = \sqrt{\|\boldsymbol{x}\|_0\,\|\boldsymbol{F}\boldsymbol{x}\|_0}$, which shows that right-hand-side of (35) is equal $p_0(\boldsymbol{x})$. Hence, we conclude that $(p_1(\boldsymbol{x}))^2 \leq p_0(\boldsymbol{x})$. ■

Combining the result of Theorem 4, Corollary 4 and (6), we have the following relations between the aforementioned uncertainty definitions

$$s_0(\boldsymbol{x}) \geq p_0(\boldsymbol{x}) \geq \Big(p_1(\boldsymbol{x})\Big)^2 \geq \|\boldsymbol{V}\|_{\max}^{-1} \tag{36}$$

for a unitary graph Fourier basis. It should be noted that ℓ_0-based additive uncertainty has the strongest result. Namely, if a signal achieves the ℓ_0-based additive uncertainty bound ($s_0(\boldsymbol{x}) = \|\boldsymbol{V}\|_{\max}^{-1}$), then the signal also achieves the bounds given in Corollaries 2 and 4. (This is due to (36).) However, the converse is not true. For this reason, in the rest of the paper, we will focus on $s_0(\boldsymbol{x})$ as the uncertainty measure.

IV. Case of Repeated Eigenvalues

Definitions 1 and 2 are generic in the sense that one can choose any suitable graph Fourier basis. Most common selections are eigenvectors of the adjacency matrix [4], the graph Laplacian and normalized Laplacian [2]. Even after one decides on which of these should be used, there still is a significant point that requires attention: the possibility of *repeated eigenvalues*. This is mostly the case for unweighted graphs or graphs with integer edge weights (see Section VI). The relation between eigenvalue multiplicity of a graph and the topology of the graph is an interesting problem. Interested reader may refer to [31] for some results. More on this topic can be found in [32]–[35].

In the following we will use eigenvectors of the graph Laplacian as the graph Fourier basis. However, the discussion is also valid for the adjacency matrix and the normalized graph Laplacian. *Further, we assume that the graph Laplacian, the adjacency matrix and the normalized graph Laplacian are diagonalizable matrices*. Hence, geometric and algebraic multiplicity of an eigenvalue are the same. This justifies the use of the term "multiplicity" without specifying which one.

Let λ_i be an eigenvalue of the graph Laplacian with multiplicity N_i. The corresponding eigenspace $\mathcal{S}_i$ is then defined as $\mathcal{S}_i = null(\boldsymbol{L} - \lambda_i \boldsymbol{I})$, where $\mathcal{S}_i$ is an N_i dimensional sub-space of $\mathcal{C}^N$. When λ_i is a repeated eigenvalue ($N_i > 1$), any vector in $\mathcal{S}_i$ is an eigenvector with eigenvalue λ_i. Therefore, by selecting different set of eigenvectors, one can come up with a different graph Fourier basis. Hence, *the graph Fourier basis for the graph Laplacian is not unique*.

Selection of the graph Fourier basis may affect the proposed uncertainty principles significantly. For example, consider the complete graph of N nodes. It is possible to select $\boldsymbol{V}$ such that either $\|\boldsymbol{V}\|_{\max}^{-1} = \sqrt{N/(N\text{-}1)}$, or $\|\boldsymbol{V}\|_{\max}^{-1} = \sqrt{N}$. The former decreases with N and approaches unity for large N, whereas the latter increases with N unboundedly. Therefore, interpretations of the proposed uncertainty bounds differ greatly depending on the selected basis. It should be noted that the complete graph of size N is an extreme example since it has an $N-1$ dimensional eigenspace corresponding to eigenvalue N. Nevertheless it shows the importance of the selection of the graph Fourier basis.

Since our definition of uncertainty depends on the selected graph Fourier basis, in the following, we will mainly discuss two different approaches for the selection of the eigenvectors. This section (Section IV) will study the first approach where we select the GFB in a way that the lower bound in Corollary 3 is minimized. In the next section (Section V), we will consider the second approach where we look for the sparsest eigenvectors. We will also relate these two approaches in Section V.

A. Minimizing the Lower Bound

In Corollary 3, we showed that the average number of nonzero elements in a graph signal and its graph Fourier transform is lower bounded by $\|\boldsymbol{V}\|_{\max}^{-1}$ where $\boldsymbol{V}$ is assumed to be unitary. When the graph of interest has repeated eigenvalues, one can select different set of eigenvectors which results in different values for $\|\boldsymbol{V}\|_{\max}^{-1}$. In this section, our purpose is to select eigenvectors of the given graph Laplacian $\boldsymbol{L}$ such that lower bound in Corollary 3 is *minimized* (or equivalently $\|\boldsymbol{V}\|_{\max}$ is maximized). We precisely define this problem as follows:

$$\max_{\boldsymbol{V}} \|\boldsymbol{V}\|_{\max} \quad \text{s.t.} \quad \boldsymbol{L} = \boldsymbol{V}\,\boldsymbol{\Lambda}_L\,\boldsymbol{V}^H, \tag{37}$$

where $\boldsymbol{\Lambda}_L$ is a diagonal matrix of eigenvalues of $\boldsymbol{L}$.

As a result of the graph Laplacian being a symmetric matrix, eigenspaces, $\mathcal{S}_i$, of $\boldsymbol{L}$ are orthogonal to each other. Hence, (37) is a decoupled problem in the sense that we can focus on individual eigenspaces rather than finding all eigenvectors of $\boldsymbol{L}$. To be more precise, assume that the graph Laplacian has K distinct eigenvalues with the corresponding eigenspaces $\mathcal{S}_i$ for $1 \le i \le K$. Then, we can write $\boldsymbol{V}$ as follows

$$\boldsymbol{V} = [\boldsymbol{V}_1\ \boldsymbol{V}_2\ \cdots\ \boldsymbol{V}_K], \tag{38}$$

where each $\boldsymbol{V}_i$ has the dimension $\boldsymbol{V}_i \in \mathcal{C}^{N\times N_i}$ (N_i is the multiplicity of the corresponding eigenvalue), spans the eigenspace $\mathcal{S}_i$, and it is unitary $\boldsymbol{V}_i^H \boldsymbol{V}_i = \boldsymbol{I}_{N_i}$ for $1 \le i \le K$. We also have $\boldsymbol{V}_i^H \boldsymbol{V}_j = \boldsymbol{0}$ for $i \ne j$ (orthogonality of eigenspaces), and $\|\boldsymbol{V}\|_{\max} = \max_{1\le i\le K} \|\boldsymbol{V}_i\|_{\max}$. Hence, we can write (37) as

$$\max_{1\le i\le K} \ \max_{\boldsymbol{V}_i} \|\boldsymbol{V}_i\|_{\max} \quad \text{s.t.} \quad \begin{array}{l} \mathcal{S}_i = null(\boldsymbol{L} - \lambda_i \boldsymbol{I}), \\ span(\boldsymbol{V}_i) = \mathcal{S}_i, \\ \boldsymbol{V}_i^H \boldsymbol{V}_i = \boldsymbol{I}_{N_i}. \end{array} \tag{39}$$

where λ_i is the i^{th} distinct eigenvalue of $\boldsymbol{L}$.

It is important to notice that inner maximization in (39) can be solved independently for each eigenspace. For this purpose, we have the following definition:

Definition 3 (Max-max norm of a subspace): Let $\mathcal{S}$ be an M-dimensional subspace of $\mathcal{C}^N$. The max-max norm of $\mathcal{S}$, $m(\mathcal{S})$, is defined as:

$$m(\mathcal{S}) = \max_{\mathbf{U}\in\mathcal{C}^{N\times M}} \|\boldsymbol{U}\|_{\max} \quad \text{s.t.} \quad span(\boldsymbol{U}) = \mathcal{S}, \quad \boldsymbol{U}^H\boldsymbol{U} = \boldsymbol{I}.$$

That is, $m(\mathcal{S})$ is the maximum of max-norm of matrices with orthonormal columns that span the given sub-space $\mathcal{S}$. ♢

Notice that *any* element of *any* unitary basis that spans $\mathcal{S}$ is always less than (or equal to) $m(\mathcal{S})$ in absolute sense. In the following theorem, we will provide a closed form solution for the max-max norm of a sub-space.

Theorem 5: Let $\mathcal{S}$ be a M-dimensional subspace of $\mathcal{C}^N$. Let $\boldsymbol{U} \in \mathcal{C}^{N\times M}$ be *any* matrix with $span(\boldsymbol{U}) = \mathcal{S}$, and $\boldsymbol{U}^H\boldsymbol{U} = \boldsymbol{I}$. Then max-max norm of sub-space $\mathcal{S}$ is

$$m(S) = \max_{1\le j\le N} \sqrt{\left(\boldsymbol{U}\,\boldsymbol{U}^H\right)_{j,j}} \tag{40}$$

where $(\cdot)_{j,j}$ denotes the j^{th} diagonal entry. ♢

Proof: Let $\boldsymbol{U}$ and $\boldsymbol{Q}$ be two matrices with orthonormal columns such that $span(\boldsymbol{U}) = span(\boldsymbol{Q}) = \mathcal{S}$. Since both span the same sub-space, we can write $\boldsymbol{Q} = \boldsymbol{U}\boldsymbol{X}$ for some unitary matrix $\boldsymbol{X}$ of size M. Then we can write $m(\mathcal{S})$ as:

$$m(\mathcal{S}) = \max_{\mathbf{X}\in\mathcal{C}^{M\times M}} \|\boldsymbol{U}\boldsymbol{X}\|_{\max} \quad \text{s.t.} \quad \boldsymbol{X}^H\boldsymbol{X} = \boldsymbol{I}, \tag{41}$$

where $\boldsymbol{U}$ is an arbitrary matrix with orthonormal columns that span $\mathcal{S}$. Let $\boldsymbol{x}_i$ be the i^{th} column of $\boldsymbol{X}$, and $\boldsymbol{u}_j^H$ be the j^{th} row

of $\boldsymbol{U}$. Then we can write (41) as:

$$m(\mathcal{S}) = \max_{\substack{1\le j\le N\\ 1\le i\le M\\ \boldsymbol{x}_i}} \left|\boldsymbol{u}_j^H \boldsymbol{x}_i\right| \quad \text{s.t.} \quad \boldsymbol{x}_i^H \boldsymbol{x}_j = \delta_{i,j} \tag{42}$$

$$= \max_{\substack{1\le j\le N\\ \boldsymbol{x}_1}} \left|\boldsymbol{u}_j^H \boldsymbol{x}_1\right| \quad \text{s.t.} \quad \|\boldsymbol{x}_1\|_2 = 1 \tag{43}$$

$$= \max_{1\le j\le N} \|\boldsymbol{u}_j^H\|_2 = \max_{1\le j\le N} \sqrt{\left(\boldsymbol{U}\,\boldsymbol{U}^H\right)_{j,j}} \tag{44}$$

where we assume (w.l.o.g.) in (42) that $\boldsymbol{x}_1$ is the vector that achieves the maximum. Furthermore, other $\boldsymbol{x}_i$'s will have the additional constraint of being orthonormal to $\boldsymbol{x}_1$. As a result, $\boldsymbol{x}_i$'s for $2 \le i \le M$ *cannot* produce a larger inner product. Further, once the optimal $\boldsymbol{x}_1$ is selected, the remaining $\boldsymbol{x}_i$'s can be selected arbitrarily as long as they are orthonormal to each other. This can be done via Gram-Schmidt process. In (43) we use the fact that inner product is maximized when the vectors are aligned with each other. Equality in (44) follows from the fact that ℓ_2-norm-square of a row is the corresponding diagonal entry of the outer product. ■

Finally, we state the maximized objective function value in (37) in the following theorem.

Theorem 6 (Maximum Max-Norm of GFB): Assume the graph Laplacian, $\boldsymbol{L}$, has K distinct eigenvalues. Let N_i denote the multiplicity, and $\mathcal{S}_i$ denote the eigenspace of the eigenvalue λ_i for $1 \le i \le K$. Let $\boldsymbol{U}_i \in \mathcal{C}^{N\times N_i}$ be *any* matrix with $\boldsymbol{U}_i^H \boldsymbol{U}_i = \boldsymbol{I}$ and $span(\boldsymbol{U}_i) = \mathcal{S}_i$. Then we have the following:

$$\max_{\substack{1\le j\le N\\ 1\le i\le K}} \sqrt{\left(\boldsymbol{U}_i\,\boldsymbol{U}_i^H\right)_{j,j}} = \max_{\boldsymbol{V}} \|\boldsymbol{V}\|_{\max} \quad \text{s.t.} \quad \boldsymbol{L} = \boldsymbol{V}\boldsymbol{\Lambda}_L\boldsymbol{V}^H,$$

where $\boldsymbol{\Lambda}_L$ is a diagonal matrix of eigenvalues of $\boldsymbol{L}$. ♢

Proof: Follows from equivalence between (37) and (39), Definition 3 and Theorem 5. ■

Theorem 6 only provides the value of the maximized objective function in (37), which is useful to find the minimum lower bound given by Corollary 3. In fact, we can explicitly construct the set of eigenvectors that result in the maximum max-norm. For this purpose, consider again the proof of Theorem 5. Notice that it is a *constructive* proof, which can be translated into an algorithm as follows. Given the graph Laplacian, one can take the eigenvalue decomposition and obtain $\boldsymbol{L} = \boldsymbol{V}\boldsymbol{\Lambda}_L\boldsymbol{V}^H$ with a proper ordering of eigenvectors such that $\boldsymbol{V} = [\boldsymbol{V}_1 \cdots \boldsymbol{V}_K]$ and columns of each $\boldsymbol{V}_i \in \mathcal{C}^{N\times N_i}$ belong to the same eigenspace. Then, we utilize the following three steps for each $\boldsymbol{V}_i$:

1) Let $\boldsymbol{v}_{i,j}^H$ denote the j^{th} row of $\boldsymbol{V}_i$, and let $j^\star$ be such that $j^\star = \arg\max_{1\le j\le N} \|\boldsymbol{v}_{i,j}^H\|_2$.
2) Let $\boldsymbol{X}_i = [\boldsymbol{x}_{i,1} \;\; \boldsymbol{x}_{i,2} \;\; \cdots \;\; \boldsymbol{x}_{i,N_i}] \in \mathcal{C}^{N_i\times N_i}$, and select $\boldsymbol{x}_{i,1} = \boldsymbol{v}_{i,j^\star}/\|\boldsymbol{v}_{i,j^\star}\|_2$. Remaining columns of $\boldsymbol{X}_i$ can be selected *arbitrarily* such that $\boldsymbol{X}_i^H\boldsymbol{X}_i = \boldsymbol{I}_{N_i}$ holds true.
3) Compute $\boldsymbol{V}_i^\star = \boldsymbol{V}_i\,\boldsymbol{X}_i$.

Then, the set of eigenvectors that has the maximum max-norm can be constructed as $\boldsymbol{V}^\star = [\boldsymbol{V}_1^\star \;\; \cdots \;\; \boldsymbol{V}_K^\star]$.

Notice that $\boldsymbol{V}_i^\star$ is just a different unitary basis for the eigenspace spanned by $\boldsymbol{V}_i$. However, unlike $\|\boldsymbol{V}_i\|_{\max}$, $\|\boldsymbol{V}_i^\star\|_{\max}$ is guaranteed to achieve the max-max norm of the corresponding eigenspace (Theorem 5). As a result, $\boldsymbol{V}^\star$ is a solution to (37) since it has the largest max-norm among all possible selections of the eigenvector matrices.

B. Numerical Problems

Even though Theorem 6 provides an efficient way to select the eigenvectors such that lower bound in Corollary 3, $\|\boldsymbol{V}\|_{\max}^{-1}$, is minimized, there is an important numerical issue. In order to utilize Theorem 6, we need to group the eigenvectors that belong to the same eigenspace, which requires an equality check between the corresponding eigenvalues. However, it is not possible to distinguish two values if they are closer than the *precision* of the numerical system. As a particular example consider a (undirected, unweighted) path graph of size N. Eigenvalues of the adjacency matrix of this graph are given as $\lambda_k = 2\cos(\pi k/(N+1))$ for $1 \le k \le N$ [36]. One can show that $|\lambda_1\text{-}\lambda_2| \le \epsilon$, when the size of the graph is $N \ge \sqrt{3}\,\pi\,\epsilon^{-1/2}$. Hence, for any numerical precision ϵ, there exists a graph of size N such that λ_1 and λ_2 cannot be distinguished from each other. Study in [37] has observed similar numerical problems as well.

V. Sparse Eigenvectors of Graphs

After the definition of additive uncertainty given in (5), the ultimate purpose of this study is to find a solution to (7) in order to find signals that are sparse *both* in the vertex domain and the Fourier domain of a given graph. Unfortunately, solution to (7) is not straightforward due to its combinatorial nature. We have provided lower bounds for the solution to (7) in Corollaries 1 and 3. In Section IV we studied the optimal selection of the graph Fourier basis in order to minimize the lower bound given by Corollary 3. However, these approaches have two downsides 1) Even though there are examples where the bound given by Corollary 3 is tight (see Section VI), it may not be the case for an arbitrary graph. 2) Even if the solution to (7) is known, aforementioned results are unable to find the signals that achieve the minima (except for Theorem 2, which requires an exhaustive search to find a bound achieving signal). In this section, in order to overcome these downsides, we will consider additive uncertainty of graph Fourier basis elements.

Without losing any generality we will use orthogonal eigenvectors of the graph Laplacian as the graph Fourier basis. That is, $\boldsymbol{V}$ is the GFB where $\boldsymbol{L} = \boldsymbol{V}\boldsymbol{\Lambda}_L\boldsymbol{V}^H$. Then, assuming a predefined ordering of eigenvectors, the i^{th} column of $\boldsymbol{V}$, denoted by $\boldsymbol{v}_i$, will be the i^{th} element of GFB. Since GFT is defined as $\boldsymbol{F} = \boldsymbol{V}^{-1}$, we have $\|\boldsymbol{F}\boldsymbol{v}_i\|_0 = 1$. Then, the additive uncertainty of a graph signal that is an element of GFB is given as

$$s_0(\boldsymbol{v}_i) = \left(\|\boldsymbol{v}_i\|_0 + 1\right)/2. \tag{45}$$

Notice that quantity in (45) is directly related to the sparsity of the GFB element. If $\boldsymbol{v}_i$ itself is a sparse vector, then we have a *direct evidence* of a signal that is sparse in the vertex domain and graph Fourier domain simultaneously. Furthermore, additive uncertainty of sparse eigenvectors may achieve (or, come close to) the bound given in Corollary 3. Therefore, our aim in this section, is to find sparse eigenvectors of graphs. However, it should be noted that a signal that achieves the minima of the additive uncertainty may *not* be an element of the GFB.

Therefore, if the GFB elements, v_i, are dense, we cannot reach any conclusion using (45). In Section VI-A2 we will provide examples in this regard.

A. Sparse Vectors in an Eigenspace

When the graph Laplacian has repeated eigenvalues, eigenvectors are not unique and they form a sub-space. In Section IV-A, we discussed how eigenvectors can be selected so that *the lower bound* for $s_0(\boldsymbol{x})$ is minimized. In this section we will try to select eigenvectors such that $s_0(\boldsymbol{v}_i)$ in (45) is minimized. To be more precise, assume that $\boldsymbol{L}$ has K ($K \leq N$) distinct eigenvalues with corresponding eigenspaces $\mathcal{S}_i$ for $1 \leq i \leq K$. Then we consider the following problem:

$$\min_{\boldsymbol{v}} \|\boldsymbol{v}\|_0 \quad \text{s.t.} \quad \boldsymbol{v} \in \mathcal{S}_i, \quad \|\boldsymbol{v}\|_2 = 1. \tag{46}$$

for each eigenspace of $\boldsymbol{L}$.

The problem in (46) is precisely defined as "The Null Space Problem" in [38], and shown to be NP-Hard. Interested reader is referred to [39]–[41] for computational approaches to solution of (46). Apart from these, the study in [42] proposes an iterative algorithm in order to find an approximate solution to (46) via ℓ_1 relaxation.

Unlike the case of minimizing the lower bound in Corollary 3 (see Theorem 6), selection of the sparsest eigenvector is not computationally tractable due to NP completeness of the problem in (46). However, it is quite interesting that the max-max norm of a subspace (given in Definition 3) provides a lower bound for the problem in (46). In the following we will precisely establish this relation.

Let $\boldsymbol{x} \in \mathcal{C}^N$ be a vector with unit ℓ_2 norm, $\|\boldsymbol{x}\|_2 = 1$. Then the infinity norm of $\boldsymbol{x}$ can be bounded as $1 \geq \|\boldsymbol{x}\|_\infty \geq 1/\sqrt{N}$. In fact, if we further know that $\boldsymbol{x}$ has L nonzero elements ($L \leq N$), we can improve the lower bound as $\|\boldsymbol{x}\|_\infty \geq 1/\sqrt{L}$. Therefore, we have the following inequality:

$$\|\boldsymbol{x}\|_0 \geq \|\boldsymbol{x}\|_\infty^{-2} \qquad \forall \boldsymbol{x} \quad \text{s.t} \quad \|\boldsymbol{x}\|_2 = 1, \tag{47}$$

where equality is achieved when $|x_i| = 1/\sqrt{L}$ for all i's in the support of $\boldsymbol{x}$.

Using the inequality in (47) we can obtain a lower bound for (46) as follows:

$$\min_{\substack{\boldsymbol{x} \in \mathcal{S} \\ \|\boldsymbol{x}\|_2 = 1}} \|\boldsymbol{x}\|_0 \geq \min_{\substack{\boldsymbol{x} \in \mathcal{S} \\ \|\boldsymbol{x}\|_2 = 1}} \|\boldsymbol{x}\|_\infty^{-2} = \left(\max_{\substack{\boldsymbol{x} \in \mathcal{S} \\ \|\boldsymbol{x}\|_2 = 1}} \|\boldsymbol{x}\|_\infty \right)^{-2}, \tag{48}$$

where we use the fact that $\|\boldsymbol{x}\|_\infty$ is strictly positive, finite, and bounded away from zero so that $\|\boldsymbol{x}\|_\infty^{-1}$ is finite.

In the following we will show that the maximization problem in the right hand side of (48) is equivalent to definition of max-max norm of the subspace $\mathcal{S}$. Remember from Definition 3 that max-max norm is defined as

$$m(\mathcal{S}) = \max_{\boldsymbol{U} \in \mathcal{C}^{N \times M}} \|\boldsymbol{U}\|_{\max} \quad \text{s.t.} \quad span(\boldsymbol{U}) = \mathcal{S}, \; \boldsymbol{U}^H \boldsymbol{U} = \boldsymbol{I}. \tag{49}$$

Assume that $\boldsymbol{U}^\star = [\boldsymbol{u}_1^\star \cdots \boldsymbol{u}_M^\star]$ is a solution to the problem in (49). Then we have $m(\mathcal{S}) = \|\boldsymbol{U}^\star\|_{\max} = \max_i \|\boldsymbol{u}_i^\star\|_\infty$. Further, $\boldsymbol{u}_i^\star \in \mathcal{S}$, and $\|\boldsymbol{u}_i^\star\|_2 = 1$. Hence we have $m(\mathcal{S}) \leq \max_{\boldsymbol{x}} \|\boldsymbol{x}\|_\infty$, $\boldsymbol{x} \in \mathcal{S}$, $\|\boldsymbol{x}\|_2 = 1$.

Now assume that $\boldsymbol{x}^\star$ is a solution to right hand side of (48). Then consider the matrix $\boldsymbol{U} = [\boldsymbol{x}^\star \; \boldsymbol{u}_2 \cdots \boldsymbol{u}_M]$ by selecting $\boldsymbol{u}_i$'s such that $\boldsymbol{U}^H \boldsymbol{U} = \boldsymbol{I}$ and $span(\boldsymbol{U}) = \mathcal{S}$. Hence $\boldsymbol{U}$ is in the feasible set of the problem in (49). Therefore, $m(\mathcal{S}) \geq \|\boldsymbol{U}\|_{\max} \geq \|\boldsymbol{x}^\star\|_\infty = \max_{\boldsymbol{x}} \|\boldsymbol{x}\|_\infty$, $\boldsymbol{x} \in \mathcal{S}$, $\|\boldsymbol{x}\|_2 = 1$.

As a result, we conclude that maximization on the right-hand side of (48) is equivalent to $m(\mathcal{S})$, and provide the following lower bound for the solution of the problem in (46)

$$\left(m(\mathcal{S})\right)^{-2} \leq \min_{\boldsymbol{x}} \|\boldsymbol{x}\|_0 \quad \text{s.t.} \quad \boldsymbol{x} \in \mathcal{S}, \quad \|\boldsymbol{x}\|_2 = 1. \tag{50}$$

The two main points of this section can be summarized as follows:

1) The search for a sparse eigenvector in a specific eigenspace is an inherently difficult problem. Even though numerical techniques that approximate the solution exist [42], closed form solutions are not available in general. In this aspect, computation of sparse eigenvectors differs from the max-norm approach discussed in Section IV, where we provided closed form solutions by focusing on individual eigenspaces.

2) Although we do not have a closed form solution for the sparsest vector in a given eigenspace, we can provide a lower bound for the total number of nonzero elements as in (50). The inequality in (50) is especially useful when we want to show that an eigenspace does *not* have sparse vectors. We will use this inequality in Section VI-A2 to formally show that an undirected cycle graph does not have sparse eigenvectors.

B. Algebraic Characterization of Sparse Eigenvectors

In the previous section we mentioned that finding the sparsest vector in an eigenspace is a difficult problem. Therefore, when looking for sparse eigenvectors, we should consider the graph (Laplacian) as a whole rather than focusing on each eigenspace individually. Furthermore, in Section IV-B we mentioned that characterization of eigenspaces of graphs suffers from numerical precision when the graph is large (relative to the numerical precision of the system). This is a significant problem especially when a large-scale real-world data is of interest. Therefore, we need a way to characterize the sparse eigenvectors of graphs without using *numerical* techniques. The purpose of this section is to find these sparse eigenvectors *algebraically*.

In the case of disconnected graphs we have a straightforward result. Assume that the graph is *undirected* but *non-negatively weighted* and consists of D disconnected components with sizes C_i. Then the adjacency matrix and the graph Laplacian can be written in the following form

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_1 & & \\ & \ddots & \\ & & \boldsymbol{A}_D \end{bmatrix}, \qquad \boldsymbol{L} = \begin{bmatrix} \boldsymbol{L}_1 & & \\ & \ddots & \\ & & \boldsymbol{L}_D \end{bmatrix}, \tag{51}$$

where $\boldsymbol{A}_i \in \mathcal{M}^{C_i}$ and $\boldsymbol{L}_i \in \mathcal{M}^{C_i}$ are the adjacency matrix and the graph Laplacian of the i^{th} component, respectively. Due to block-diagonal form of $\boldsymbol{A}$ and $\boldsymbol{L}$, corresponding eigenvectors can be selected to be block sparse. Therefore, there exists an eigenvector that has *at most* C_i nonzero elements for each $1 \leq i \leq D$. It is important to note that the converse of this result

is not true: if a graph has a sparse eigenvector, it does *not* imply that the graph is disconnected. As a counter example consider Theorem 8, which proves that a connected graph can have a sparse eigenvector. Examples of such graphs will be provided at the end of this section. However, 1-sparse eigenvectors are exemption in this regard. That is, a graph has a 1-sparse eigenvector if and only if it has an isolated node. This result is stated as follows.

Theorem 7 (Isolated nodes of a graph): Assume that the graph of interest is *undirected* but *non-negatively weighted*. Then, the following statements are equivalent:

1) The graph has an isolated node.

2) The graph Laplacian has a 1-sparse eigenvector.

3) The GFB *can* be selected such that there exists a nonzero signal that achieves $s_0(\boldsymbol{x}) = 1$, i.e., $\|\boldsymbol{x}\|_0 = \|\boldsymbol{F}\boldsymbol{x}\|_0 = 1$. ♢

Proof: We prove (1) implies (2): Assume that the graph has an isolated node. According to (51), there exists a 1-sparse eigenvector of the graph Laplacian.

We now prove that (2) implies (1): Let $\boldsymbol{v}$ be the 1-sparse eigenvector. Without loss of generality assume that the first index is nonzero $v_1 = 1$ and the rest is zero. Therefore $\boldsymbol{L}\boldsymbol{v}$ is equivalent to the first column of $\boldsymbol{L}$. That is, $\boldsymbol{L}\boldsymbol{v} = [d_1 \quad -\boldsymbol{a}_{r,1}^T]^T = \lambda[1 \quad \boldsymbol{0}^T]^T$, where $\boldsymbol{a}_{r,1} \in \mathcal{R}^{N-1}$ is the vector that denotes the adjacency of node 1, and $d_1 = \|\boldsymbol{a}_{r,1}\|_1$ is the degree of node 1. Therefore we have $\boldsymbol{a}_{r,1} = \boldsymbol{0}$, hence $d_1 = 0$. Since edge weights are non-negative, node 1 is an isolated node.

Next we will prove that (2) implies (3): Let $\boldsymbol{v}$ be a 1-sparse eigenvector of the graph Laplacian. Then, $\boldsymbol{v}$ *can* be selected to be an element of the GFT. In this case, $\|\boldsymbol{v}\|_0 = \|\boldsymbol{F}\boldsymbol{v}\|_0 = 1$, hence $s_0(\boldsymbol{v}) = 1$.

Finally, we prove (3) implies (2): Let $s_0(\boldsymbol{x}) = 1$ for some $\boldsymbol{x} \neq 0$. Since $\|\boldsymbol{x}\|_0 \geq 1$ for any nonzero signal, we must have $\|\boldsymbol{x}\|_0 = \|\boldsymbol{F}\boldsymbol{x}\|_0 = 1$. $\|\boldsymbol{F}\boldsymbol{x}\|_0 = 1$ implies that $\boldsymbol{x}$ is an eigenvector of the graph Laplacian and $\|\boldsymbol{x}\|_0 = 1$ implies that $\boldsymbol{x}$ is 1-sparse. Hence the graph Laplacian has a 1-sparse eigenvector. ■

Now we provide the characterization theorem for 2-sparse eigenvectors of graphs. Recall that a graph is said to be connected if there is a path between any pair of nodes.

Theorem 8 (2-sparse eigenvectors of a connected graph): Let $\boldsymbol{A}$ denote the adjacency matrix of an *undirected* and *connected* graph with $a_{i,j} \geq 0$ being the weight of the edge between nodes i and j. Then, there exist nodes i and j such that

$$a_{i,k} = a_{j,k} \quad \forall\, k \in \{1, \cdots, N\} \setminus \{i, j\}, \tag{52}$$

if and only if the graph Laplacian, $\boldsymbol{L}$, has a 2-sparse eigenvector with nonzero eigenvalue $\lambda = d_i + a_{i,j}$. When the graph is unweighted, (52) can be stated as

$$\mathcal{N}(i) \setminus \{j\} = \mathcal{N}(j) \setminus \{i\}, \tag{53}$$

where $\mathcal{N}(i)$ is the set of nodes that are adjacent to node i. ♢

Note that a 2-sparse eigenvector can be assumed to have values 1 and -1 on the nodes with the property (52). (See the proof below.) This result is especially useful for Theorem 9.

Proof: Assume that the graph Laplacian of a connected graph has a two-sparse eigenvector $\boldsymbol{v}$ with nonzero eigenvalue. Due to permutation invariance of the node labels, without loss of any generality assume that the first two indices are nonzero, that is, $v_1 \neq 0$ and $v_2 \neq 0$, but $v_i = 0$ for $i \geq 3$.

For a connected graph, notice that the all-1 vector is the only eigenvector of the graph Laplacian with the zero eigenvalue. Since the graph Laplacian is a symmetric matrix the eigenspaces are orthogonal to each other. Therefore the 2-sparse eigenvector (with nonzero eigenvalue) $\boldsymbol{v}$ is orthogonal to the all-1 vector, which implies that $v_1 + v_2 = 0$. Then, we can select $v_1 = -v_2 = 1$ without loss of any generality.

Let $\boldsymbol{A}$ denote the adjacency matrix of the graph. We have

$$\boldsymbol{A} = \begin{bmatrix} 0 & a_{1,2} & \boldsymbol{a}_{r,1}^T \\ a_{2,1} & 0 & \boldsymbol{a}_{r,2}^T \\ \boldsymbol{a}_{r,1} & \boldsymbol{a}_{r,2} & \boldsymbol{A}_r \end{bmatrix}, \boldsymbol{L} = \begin{bmatrix} d_1 & -a_{1,2} & -\boldsymbol{a}_{r,1}^T \\ -a_{2,1} & d_2 & -\boldsymbol{a}_{r,2}^T \\ -\boldsymbol{a}_{r,1} & -\boldsymbol{a}_{r,2} & \boldsymbol{L}_r \end{bmatrix}, \tag{54}$$

where $\boldsymbol{A}_r \in \mathcal{M}^{N-2}$ and $\boldsymbol{L}_r \in \mathcal{M}^{N-2}$ are the partitions of the adjacency matrix and the graph Laplacian, respectively. $\boldsymbol{a}_{r,1} \in \mathcal{R}^{N-2}$ is the vector that denotes the adjacency of node 1 except node 2. $\boldsymbol{a}_{r,2}$ is the same for node 2. Notice that $d_1 = a_{2,1} + \|\boldsymbol{a}_{r,1}\|_1$ and $d_2 = a_{1,2} + \|\boldsymbol{a}_{r,2}\|_1$. Then, consider the following

$$\boldsymbol{L}\boldsymbol{v} = \begin{bmatrix} d_1 + a_{1,2} \\ -(a_{2,1} + d_2) \\ -\boldsymbol{a}_{r,1} + \boldsymbol{a}_{r,2} \end{bmatrix} = \lambda \boldsymbol{v} = \lambda \begin{bmatrix} 1 \\ -1 \\ \boldsymbol{0} \end{bmatrix}.$$

Therefore we have $\boldsymbol{a}_{r,1} = \boldsymbol{a}_{r,2}$, which in particular implies that $d_1 = d_2$ since $a_{1,2} = a_{2,1}$ (graph is undirected). Furthermore the corresponding eigenvalue is $\lambda = d_1 + a_{1,2}$. Since the graph is connected $d_1 > 0$, λ is nonzero. Notice that the condition $\boldsymbol{a}_{r,1} = \boldsymbol{a}_{r,2}$ is the same as the condition in (52).

Conversely, assume that there exist two nodes with the property in (52). Without loss of generality, assume $i = 1$ and $j = 2$, and let $\boldsymbol{v}$ be a 2-sparse vector with $v_1 = -v_2 = 1$. Then partition the graph Laplacian as in (54). Due to (52), we have $\boldsymbol{a}_{r,1} = \boldsymbol{a}_{r,2}$, and $d_1 = d_2$. Then we have $\boldsymbol{L}\boldsymbol{v} = \lambda \boldsymbol{v}$ with $\lambda = d_1 + a_{r,1}$. Therefore, $\boldsymbol{v}$ is a 2-sparse eigenvector of $\boldsymbol{L}$. Further, the graph is connected $d_1 > 0$, hence $\lambda > 0$. ■

Similar to Theorem 8, the study in [28] also reveals a specific graph structure that results in sparse eigenvectors of both the adjacency matrix and the graph Laplacian. In particular, it considers the case when a graph has two copies of the same subgraph (referred to as "motif doubling" in [28]). That is, there are two disjoint subsets (of size K) of nodes, S_1 and S_2, such that the induced sub-graphs on S_1 and S_2 are the same, there is no edge between S_1 and S_2, and S_2 is connected to the rest of the graph in the same way S_1 is. In this case the adjacency matrix (and the graph Laplacian) can be shown to have $(2K)$-sparse eigenvectors (Theorem 2.2 of [28]). In the case of $K = 1$ (each subset has only one node), this motif doubling property reduces to the condition in (53) with $a_{i,j} = 0$.

However, it is important to note that the condition in Theorem 8 is more general than the one in [28] due to the following two reasons: 1) the construction in [28] provides only a *sufficient* condition, whereas Theorem 8 gives the necessary and sufficient condition to have a 2-sparse eigenvector. 2) The motif doubling idea in [28] specifically considers the case when

$a_{i,j} = 0$, whereas Theorem 8 is applicable to the case of $a_{i,j} \neq 0$ as well. In the general case of K, it is straightforward to find sufficient conditions for a K-sparse eigenvector to exist: the motif doubling in [28] and "the same neighborhood structure" in [24] are two such examples. On the other hand, it is difficult to reveal the necessary conditions for K-sparse eigenvectors to exist.

As discussed in Section II, the additive uncertainty of a signal $\boldsymbol{x}$ in (5) can take only half integer values for nonzero signals, that is $s_0(\boldsymbol{x}) \in \{1, 3/2, 2, \cdots, N\}$. It should be noted that Theorem 7 precisely characterizes the case when $s_0(\boldsymbol{x})$ takes its possible minimum value. A nonzero signal has $s_0(\boldsymbol{x}) = 1$ *if and only if* the graph has an isolated node. This result is especially useful to conclude that a nonzero signal on a *connected graph*, which does not have any isolated node, *cannot* achieve $s_0(\boldsymbol{x}) = 1$. Therefore, we consider the next attainable case for connected graphs, that is, $s_0(\boldsymbol{x}) = 3/2$. This happens under two circumstances

1) $\|\boldsymbol{x}\|_0 = 1$, $\|\boldsymbol{F}\boldsymbol{x}\|_0 = 2$: the signal $\boldsymbol{x}$ is an impulse on the vertex domain, and it has a 2-sparse GFT.
2) $\|\boldsymbol{x}\|_0 = 2$, $\|\boldsymbol{F}\boldsymbol{x}\|_0 = 1$: the signal $\boldsymbol{x}$ is a 2-sparse eigenvector of the graph Laplacian.

We note that Theorem 8 precisely characterizes the second case. Therefore, given a connected graph, existence of a pair of nodes that satisfy (52) implies that $s_0(\boldsymbol{x}) \geq 3/2$ for all nonzero signals on the graph. Furthermore, the bound is tight, and the signal that achieves the bound is known. This result is formally stated as follows.

Theorem 9 (Uncertainty bound for connected graphs): For an *undirected*, *connected*, and *non-negatively weighted* graph, assume that there exist nodes i and j satisfying the condition in (52). Then, the GFB with respect to the graph Laplacian *can be* selected such that

$$s_0(\boldsymbol{x}) \geq 3/2 \qquad \forall \boldsymbol{x} \neq \boldsymbol{0}. \tag{55}$$

Furthermore, the signal achieving this bound, $s_0(\boldsymbol{x}^\star) = 3/2$, is given as $x_i^\star = -x_j^\star = 1$ and zero everywhere else. ♢

Proof: For a simple and connected graph, Theorem 7 says that there is no signal such that $s_0(\boldsymbol{x}) = 1$. Since $s_0(\boldsymbol{x})$ can take only half integers in $[1, N]$, $s_0(\boldsymbol{x}) \neq 1$ implies that $s_0(\boldsymbol{x}) \geq 3/2$ for any nonzero signal $\boldsymbol{x}$.

Furthermore, if there is a pair of nodes satisfying (52), Theorem 8 says that the graph Laplacian has a 2-sparse eigenvector. Let $\boldsymbol{v}$ denote this eigenvector. Then we have $v_i = -v_j = 1$ and zero everywhere else. Notice that GFB with respect to the graph Laplacian *can be* selected such that $\boldsymbol{v}$ is an element of GFB. In this case we have $\|\boldsymbol{v}\|_0 = 2$ and $\|\boldsymbol{F}\boldsymbol{v}\|_0 = 1$, hence $s_0(\boldsymbol{v}) = 3/2$. This shows that the lower bound in (55) is tight and attainable. ■

There are four remarks regarding Theorem 9.

1) The tightness of the bound given in (55) does *not* depend on the size and the global structure of the graph. Existence of a pair of nodes with (52) directly implies this result.

2) The signal that achieves the bound is localized on the graph. Notice that if two nodes have the property in (52), they must have at least one common neighbor. (This follows from the fact that the graph is connected.) As a result, nonzero elements of the signal that achieves the bound in (55) are at most 2 hops away from each other. However, localization property is unique to 2-sparse eigenvectors. An eigenvector with an arbitrary level of sparsity may not be localized on the graph. Details of these are discussed in [24].

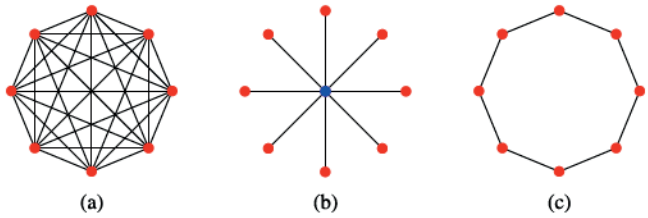


Fig. 1. (a) K_8, complete graph of size 8, (b) S_9, star graph of size 9, (c) C_8, cycle graph of size of 8.

3) Due to Corollary 3, the inequality $s_0(\boldsymbol{x}) \geq \|\boldsymbol{V}\|_{\max}^{-1}$ is always true. When $3/2$ is the smallest attainable value for $s_0(\boldsymbol{x})$, we have $3/2 \geq \|\boldsymbol{V}\|_{\max}^{-1}$. Therefore, existence of a pair of nodes with (52) proves that there exists a GFB $\boldsymbol{V}$ such that $\|\boldsymbol{V}\|_{\max} \geq 2/3$. In fact, using Corollary 2 this result can be slightly improved to $\|\boldsymbol{V}\|_{\max} \geq 1/\sqrt{2}$.

4) In the case of repeated eigenvalues of the graph Laplacian, the GFB is not unique. However, uncertainty bounds *depend* on the selection of the GFB. When there are repeated eigenvalues GFB should be selected properly (sparse eigenvector should be an element of GFB) in order to have (55). This point will be numerically demonstrated in Section VI-A3.

In the following we will provide three classical graph examples that satisfy, or do not satisfy, the condition in (53). Notice that these graphs are simple (undirected, unweighted, free from self-loops) and connected.

1) Complete Graph, K_N: A complete graph of N nodes has an edge between any two nodes. Figure 1(a) provides a visual representation of K_8. Let i and j be two arbitrary nodes of a complete graph. Then we have $\mathcal{N}(i) = \{1, \cdots, N\} \setminus \{i\}$, and $\mathcal{N}(j) = \{1, \cdots, N\} \setminus \{j\}$. Therefore, we have

$$\mathcal{N}(i) \setminus \{j\} = \mathcal{N}(j) \setminus \{i\} = \mathcal{N}(i) = \{1, \cdots, N\} \setminus \{i, j\}, \tag{56}$$

which shows that a complete graph of an arbitrary size has a 2-sparse eigenvector.

2) Star Graph, S_N: A star graph of size N has a center node that is connected to any other node, and all the nodes are connected only to the center node. Figure 1(b) provides a visual representation of S_9. Assume that the center node is labeled as 1. Let i and j be two nodes other than the center node. Then we have $\mathcal{N}(i) = \mathcal{N}(j) = \{1\}$. Therefore,

$$\mathcal{N}(i) \setminus \{j\} = \mathcal{N}(j) \setminus \{i\} = \{1\}, \tag{57}$$

which shows that a star graph of an arbitrary size has a 2-sparse eigenvector.

3) Cycle Graph, C_N: A cycle graph of size N contains a single cycle through all nodes. Figure 1(c) provides a visual representation of C_8. Notice that $C_2 = K_2$, $C_3 = K_3$, $C_4 = K_{2,2}$, hence they have 2-sparse eigenvectors as shown above. For

$N \geq 5$, C_N does *not* have a pair of nodes that satisfy (53). Therefore, a cycle graph for $N \geq 5$ does not have a 2-sparse eigenvector. In fact, an eigenvector of a cycle graph of size N has at least $N/2$ nonzero values (see Section VI-A2).

Above examples are carefully selected to point out an important observation: *sparsity of the graph and existence of sparse eigenvectors do not imply each other*. This follows from the following three facts: 1) A complete graph is dense, yet it has a sparse eigenvector. 2) A cycle graph is sparse, yet it does not have a sparse eigenvector. 3) A star graph is sparse, and it has a sparse eigenvector.

One can also use Theorem 8 to find a sparse GFB of a given graph. Existence of a pair of nodes that satisfy (53) guarantees the existence of a 2-sparse eigenvector. When there is more than one pair, it is possible to find various 2-sparse eigenvectors. Even though those eigenvectors may not be orthonormal to each other they provide a sparse GFB. In fact, $N-1$ eigenvectors of the graph Laplacian of a complete graph of size N can be selected to be 2-sparse. These eigenvectors will be linearly independent, but not orthonormal. In this case, GFB has only $3N-2$ nonzero entries. Details of these will not be elaborated here, and deserve an independent study.

It is important to notice that the condition in (52) is purely algebraic, and does not require any numerical computation. Therefore, Theorems 8 and 9 are *not* subject to the problems discussed in Section IV-B. In order to find a pair of nodes with the property in (52), one can check every pair in a brute-force manner, which results in $\binom{N}{2}$ tests in total. Therefore, complexity of verifying that a graph has a pair of nodes with the property (52) is at most $O(N^2)$. However, there may exist more efficient search algorithms for this purpose.

VI. Examples of Uncertainty Bounds

A. Standard Examples from Graph Theory

1) Circulant Graphs: A graph is said to be circulant when its adjacency matrix is a circulant matrix under suitable permutation of the node numbering [29]. This is a broad family including cyclic graphs (directed or undirected), complete graphs, complete bi-partite graphs and more. The directed cyclic graph of size N, whose adjacency matrix is given as

$$\boldsymbol{C}_N = \begin{bmatrix} & & & 1 \\ 1 & & & \\ & \ddots & & \\ & & 1 & \end{bmatrix} \in \mathcal{M}^N, \tag{58}$$

is particularly important since it relates the graph signal processing to classical signal processing [3], [4].

The adjacency matrix of a circulant graph is a circulant matrix (with suitable permutation of vertices), and can be diagonalized by the DFT matrix:

$$\boldsymbol{A} = \boldsymbol{W}_N^H \, \boldsymbol{\Lambda} \, \boldsymbol{W}_N, \tag{59}$$

for some diagonal $\boldsymbol{\Lambda}$, where $\boldsymbol{W}_N$ is the normalized DFT matrix of size N. Hence, the graph Fourier transform based on the adjacency matrix is $\boldsymbol{F} = \boldsymbol{W}_N$, and we have $\|\boldsymbol{V}\|_{\max}^{-1} = \sqrt{N}$. As a result, the strong uncertainty principle for circulant graphs of size N is $s_0(\boldsymbol{x}) \geq \sqrt{N}$.

As shown in [22], this is a tight bound when N is a perfect square. Consider the "picket fence" signal which has support $S = \{1, 1+\sqrt{N}, 1+2\sqrt{N}, \cdots, 1+N-\sqrt{N}\}$ with

$$\boldsymbol{x}_S = 1, \qquad \boldsymbol{x}_{\bar{S}} = 0. \tag{60}$$

Then we have $\widehat{\boldsymbol{x}} = \boldsymbol{F}\boldsymbol{x} = \boldsymbol{x}$. Notice that $|S| = \sqrt{N}$. As a result we have $s_0(\boldsymbol{x}) = \sqrt{N} = \|\boldsymbol{V}\|_{\max}^{-1}$, that is, strong ℓ_0 uncertainty is achieved (Corollary 3).

For the weak uncertainty we have $p_0(\boldsymbol{x}) \geq \sqrt{N}$, that is, $\|\boldsymbol{x}\|_0 \, \|\widehat{\boldsymbol{x}}\|_0 \geq N$, where $\widehat{\boldsymbol{x}}$ corresponds to DFT of $\boldsymbol{x}$. This is a well-known uncertainty result given in [20]. Unlike the strong uncertainty, weak uncertainty bound can be achieved for any N. Let $\boldsymbol{x}$ be an impulse, then $\widehat{\boldsymbol{x}}$ will have no zero elements, resulting in $\|\boldsymbol{x}\|_0 \, \|\widehat{\boldsymbol{x}}\|_0 = N$.

If the graph is unweighted, then circulant graphs are regular (each node has the same degree). In this case the graph Laplacian can be written as $\boldsymbol{L} = d\boldsymbol{I} - \boldsymbol{A}$, where d is the degree of each node. Therefore, $\boldsymbol{L}$ is also a circulant matrix, and diagonalizable by $\boldsymbol{W}_N$. As a result, strong uncertainty bound based on the graph Laplacian is also tight.

2) Cycle Graph: In this part we will focus on the *undirected* cyclic graph as visually shown in Figure 1(c). Eigenvalues of the Laplacian of a cyclic graph are given as $\lambda_k = 2 - 2\cos(2\pi k/N)$ for $0 \leq k \leq N-1$ [36]. Notice that $\lambda_0 = 0$ is not a repeated eigenvalue. However, other eigenvalues have the property $\lambda_k = \lambda_{N-k}$ for $k \geq 1$. Therefore, the Laplacian of a cycle graph has 2-dimensional eigenspaces. Let $\mathcal{S}_k$ denote the 2-dimensional eigenspace of the Laplacian corresponding to eigenvalue λ_k. Let $\boldsymbol{w}_k$ denote the k^{th} column of $\boldsymbol{W}_N^H$. Then $\boldsymbol{U}_k = [\boldsymbol{w}_k \;\; \boldsymbol{w}_{N-k}]$ spans the eigenspace $\mathcal{S}_k$ since the Laplacian is diagonalized by $\boldsymbol{W}_N$. Notice that $|(\boldsymbol{W}_N)_{i,j}| = 1/\sqrt{N}$ for all pairs of (i,j). As a result, each row of $\boldsymbol{U}_k$ has ℓ_2 norm of $\sqrt{2/N}$. Then, Theorem 5 gives that $m(\mathcal{S}_k) = \sqrt{2/N}$. Using (50) we conclude that the total number of nonzero elements of an eigenvector in $\mathcal{S}_k$ can be at least $(m(\mathcal{S}_k))^{-2} = N/2$. Since this is true for any eigenspace, *any eigenvector of the Laplacian of a cycle graph (of size N) has at least $N/2$ nonzero values.*

As discussed in the previous sub-section, we have $s_0(\boldsymbol{x}) \geq \sqrt{N}$ for all nonzero signals on a cycle graph when GFB selected as $\boldsymbol{W}_N^H$. When we consider the additive uncertainty of elements of GFB, as in (45), we get $s_0(\boldsymbol{v}_i) \geq (N+2)/4$ since each eigenvector has at least $N/2$ nonzeros. As a result, elements of GFB are not useful candidates to achieve the bound $s_0(\boldsymbol{x}) \geq \sqrt{N}$ with equality. This is because eigenvectors of the cycle graph are not sparse.

3) Complete Graph: Being a circulant graph, GFB of a complete graph *can be* selected as $\boldsymbol{W}_N^H$. With this selection, the additive uncertainty bound is given as $s_0(\boldsymbol{x}) \geq \sqrt{N}$. It should be noted that the Laplacian of a complete graph (of size N) has only two distinct eigenvalues: 0 with multiplicity 1, N with multiplicity $N-1$. One can select different set of vectors to span the $N-1$ dimensional eigenspace. In fact, as discussed in

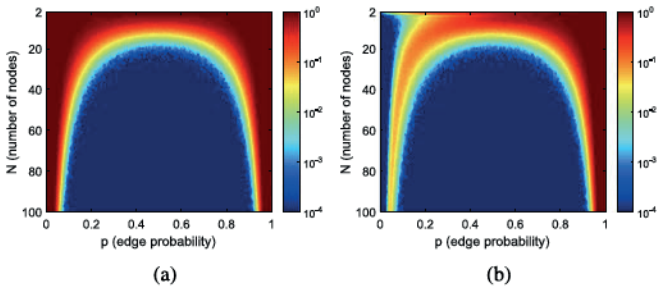


Fig. 2. Probability of $G(N,p)$ (a) having a pair of nodes satisfying (53), (b) being connected *and* having a pair of nodes satisfying (53). Probabilities are obtained via averaging over 10^4 experiments, hence the lowest observed probability is 10^{-4}.

Section V-B1, one of these vectors can be selected to be 2-sparse. With such a selection, by virtue of Theorem 9, the uncertainty bound is given as $s_0(\boldsymbol{x}) \geq 3/2$, which is significantly different from the one when $\boldsymbol{W}_N^H$ is used as GFB. At this point we are not favoring one selection of GFB over another. The sole purpose of this example is to show that selection of the GFB is an important issue in the presence of repeated eigenvalues.

B. M-Block Cyclic Graphs

In [8]–[11], it is shown that M-Block cyclic graphs play an important role in the development of multirate processing of graph signals. When we assume that all the edges have unit weights, the adjacency matrix of an M-Block cyclic graph of size N can be written as: $\boldsymbol{A} = \boldsymbol{C}_M \otimes \left(\mathbb{1}_{N/M}\, \mathbb{1}_{N/M}^T\right)$, where $\boldsymbol{C}_M$ is given in (58) and 1_N is a vector of size N with all 1 entries. Notice that both $\boldsymbol{C}_M$ and $\mathbb{1}_{N/M}\, \mathbb{1}_{N/M}^T$ are circulant matrices, hence they are diagonalizable by normalized DFT matrices of respective sizes. As a result, GFT based on the adjacency matrix (and the graph Laplacian since all the nodes have the same degree) can be selected as $\boldsymbol{F} = \boldsymbol{W}_M \otimes \boldsymbol{W}_{N/M}$. Notice that $\boldsymbol{V} = \boldsymbol{F}^H$ is unitary and $\|\boldsymbol{V}\|_{\max}^{-1} = \sqrt{N}$.

As an example consider the case $N = 9$ and $M = 3$, and consider the following signal $\boldsymbol{x} = [1\ 1\ 1]^T \otimes [1\ 0\ 0]^T$. The GFT of this signal, $\widehat{\boldsymbol{x}}$, is given as

$$\widehat{\boldsymbol{x}} = \boldsymbol{W}_3[1\ 1\ 1]^T \otimes \boldsymbol{W}_3[1\ 0\ 0]^T = [1\ 0\ 0]^T \otimes [1\ 1\ 1]^T. \quad (61)$$

Hence, we have $\|\boldsymbol{x}\|_0 = \|\widehat{\boldsymbol{x}}\|_0 = 3$, and $s_0(\boldsymbol{x}) = 3 = \|\boldsymbol{V}\|_{\max}^{-1}$. Therefore, strong uncertainty bound is achieved. In general, let N be a perfect square and $M = \sqrt{N}$. *Then for any M-Block cyclic graph of size N with unit weights, we can find a signal that achieves the strong uncertainty bound.*

C. Erdős-Rényi Graphs

An Erdős-Rényi graph $G(N,p)$ is a simple graph of N nodes where an edge between a pair of nodes appears randomly and independently with probability p [1], [43]. In Figure 2(a) we empirically compute the probability of a $G(N,p)$ having a pair of nodes satisfying the condition in (53). Notice that if a graph has such a pair of nodes then the same pair satisfies (53) on the complement of the graph as well. This is due to the equality condition in (53) that remains satisfied when all the edges are complemented. Further notice that complement of a $G(N,p)$ is a $G(N,1-p)$ graph. As a result $G(N,p)$ and $G(N,1-p)$ have the same probability of having a pair of nodes satisfying (53). This explains the symmetry of Figure 2(a) around $p = 1/2$.

It is important to note that Theorem 8 specifically considers the case of connected graphs since it is trivial to find sparse eigenvectors in disconnected graphs (see (51)). However, a $G(N,p)$ tends to be disconnected when p is small. In fact $p < log(N)/N$ results in (almost surely) isolated vertices, and $p > log(N)/N$ guarantees (almost surely) $G(N,p)$ to be connected [43]. In order to get rid of the trivial cases we need to consider the probability of $G(N,p)$ having a pair of nodes with (53) *and* being connected. Experimental computation of this probability is given in Figure 2(b). Notice that connectivity is not preserved under complementation, hence Figure 2(b) is not symmetric around $p = 1/2$.

Figure 2(b) shows the existence of connected Erdős-Rényi graphs with 2-sparse eigenvectors. Figure 2(b) also suggests that as N gets larger it is less likely to find such graphs, which can be explained as follows. The study in [44] states that the ℓ_∞-norm of any unit eigenvector of $G(N,p)$ is almost surely $o(1)$ for $p = \omega(log(N)/N)$, where $o(\cdot)$ and $\omega(\cdot)$ denotes little-o and little-omega notations, respectively. That is, as $N \to \infty$ we have $\|v\|_\infty \to 0$ for all v. Therefore, $\|\boldsymbol{V}\|_{\max} \to 0$, hence, $\|\boldsymbol{V}\|_{\max}^{-1} \to \infty$. Since the uncertainty bound in (18) goes to infinity we do not expect to find 2-sparse eigenvectors in connected Erdős-Rényi graphs for large values of N.

D. Real World Examples

In the following we will use the term $\lambda^{(k)}$ to denote that the eigenvalue λ has multiplicity k.

1) Minnesota Road Graph: In this part, we will consider the Minnesota road graph [7], [13]. We use the data publicly available in [45]. This graph has 2642 nodes in total where 2 nodes are disconnected to the rest of the graph. Since a road graph is expected to be connected, we disregard those two nodes. See [7], [13] for the visual representation of the graph. Here each node is an intersection, and $a_{i,j} = 1$ if there is a road connecting the intersections, otherwise $a_{i,j} = 0$. There are total of 3302 undirected unweighted edges. The graph is simple and connected, $\boldsymbol{A}$ and $\boldsymbol{L}$ are symmetric matrices (in particular, diagonalizable), hence $\boldsymbol{V}_A$ and $\boldsymbol{V}_L$ can be selected to be unitary.

For the Minnesota road graph, both $\boldsymbol{A}$ and $\boldsymbol{L}$ have repeated eigenvalues. As a result, $\boldsymbol{V}_A$ and $\boldsymbol{V}_L$ are *not* unique. In fact, the adjacency matrix has repeated eigenvalues of $-1^{(15)}$, $0^{(44)}$, and $1^{(13)}$. The graph Laplacian has repeated eigenvalues of $0.3820^{(2)}$, $1^{(10)}$, $2^{(7)}$, $2.6180^{(2)}$, and $3^{(6)}$. In order to minimize the uncertainty bound given by Corollary 3, we can select $\boldsymbol{V}_A$ and $\boldsymbol{V}_L$ such that $\|\boldsymbol{V}_A\|_{\max}$ and $\|\boldsymbol{V}_L\|_{\max}$ are maximized. This idea is discussed in Section IV, and the closed-form solution for such a selection is provided by Theorem 6. Via numerical

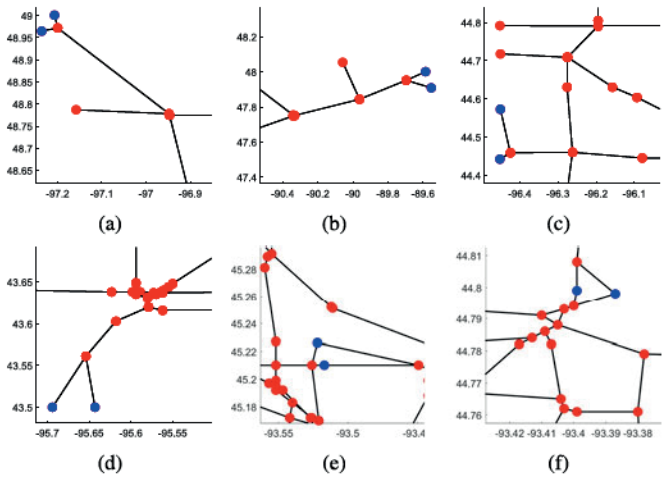


Fig. 3. Pair of nodes in Minnesota road graph that result in 2-sparse eigenvectors. Axes represent the geographical coordinates. The pairs that satisfy the condition in (52) are colored in blue. Notice that the pairs in (a)–(d) generate eigenvectors with eigenvalue 1, and the pairs in (e)–(f) generate eigenvectors with eigenvalue 2. (See Theorem 8.)

evaluation of Theorem 6 on the Minnesota road graph, we obtain the following:

$$0.7071 = \max_{\boldsymbol{V}} \|\boldsymbol{V}\|_{\max} \quad \text{s.t.} \quad \boldsymbol{A} = \boldsymbol{V}\boldsymbol{\Lambda}_A\boldsymbol{V}^H, \tag{62}$$

$$0.8343 = \max_{\boldsymbol{V}} \|\boldsymbol{V}\|_{\max} \quad \text{s.t.} \quad \boldsymbol{L} = \boldsymbol{V}\boldsymbol{\Lambda}_L\boldsymbol{V}^H. \tag{63}$$

Due to Corollaries 2 and 3, when $\boldsymbol{V}_A$ is selected to be the GFB, (62) gives the following uncertainty bounds $s_0(\boldsymbol{x}) \geq p_0(\boldsymbol{x}) \geq \|\boldsymbol{V}_A\|_{\max}^{-1} = 1.4142$. When $\boldsymbol{V}_L$ is selected to be the GFB, (63) gives the following uncertainty bounds $s_0(\boldsymbol{x}) \geq p_0(\boldsymbol{x}) \geq \|\boldsymbol{V}_L\|_{\max}^{-1} = 1.1987$.

It is important to remember that $s_0(\boldsymbol{x})$ can have values only on a discrete set, namely, $s_0(\boldsymbol{x}) = k/2$ for some integer $k \geq 2$. As a result, for both selection of GFB, signals on the Minnesota road graph *cannot* attain the uncertainty bound in Corollary 3 in a strict sense. However, by rounding-off the value of both $\|\boldsymbol{V}_A\|_{\max}^{-1}$ and $\|\boldsymbol{V}_L\|_{\max}^{-1}$ to the next attainable value of $s_0(\boldsymbol{x})$, we get

$$s_0(\boldsymbol{x}) \geq 3/2, \tag{64}$$

for the strong uncertainty bound for both selection of GFB.

Even though (64) is a valid bound, Corollary 3 and Theorem 6 gives no further information about existence and characterization of a signal that achieves the bound. At this point it is quite interesting to observe that the bound in (64) is the same as the bound provided by Theorem 9 (for $\boldsymbol{V}_L$ as GFB), which requires existence of a pair of nodes with the property in (53). In fact, the Minnesota road graph *does* have 6 different pairs of nodes with the property in (53). These pairs are visualized in Figure 3. As a result, the bound in (64) is *tight*, and the signals that achieve the bound are defined by the pairs of nodes in Figure 3 (see Theorem 9). It should be noted that tightness of (64) is valid when $\boldsymbol{V}_L$ is *selected* to include at least one 2-sparse eigenvector generated by the pairs in Figure 3.

2) Co-appearance Network: In this example, we will consider the co-appearance network of characters in the famous novel Les Misérables by Victor Hugo [46]. Data is publicly available in [47]. This is an undirected but *weighted* graph, where two characters are connected if they appear in the same scene, and the weight of an edge is the total number of co-appearances through the novel.

The graph has 77 nodes and 254 (weighted) edges in total. The Laplacian of the graph has repeated eigenvalues of $1^{(9)}$, $13^{(2)}$, and $28^{(2)}$. As a result, GFB with respect to the graph Laplacian, $\boldsymbol{V}_L$, is not unique. In order to minimize the bound given in Corollary 3, we use Theorem 6 and obtain the following result

$$0.9398 = \max_{\boldsymbol{V}} \|\boldsymbol{V}\|_{\max} \quad \text{s.t.} \quad \boldsymbol{L} = \boldsymbol{V}\boldsymbol{\Lambda}_L\boldsymbol{V}^H. \tag{65}$$

When $\boldsymbol{V}_L$ is selected to be the GFB, (65) gives the following uncertainty bounds (Corollaries 2 and 3) $s_0(\boldsymbol{x}) \geq p_0(\boldsymbol{x}) \geq \|\boldsymbol{V}_L\|_{\max}^{-1} = 1.0641$. Due to discrete nature of $s_0(\boldsymbol{x})$ it is clear that this bound cannot be satisfied with equality. When $\|\boldsymbol{V}_L\|_{\max}^{-1}$ is rounded-off to the next attainable value of $s_0(\boldsymbol{x})$, we get the same bounds as in (64). At this point we can use Theorem 9 to find signals (if there is any) that achieve the bound in (64).

In a co-appearance graph, pair of nodes with the condition in (52) has a meaningful interpretation. If two characters always appear simultaneously, they will have the same number of co-appearances with other characters, which implies the condition in (52) mathematically. As an example, consider characters "Brevet", "Chenildieu", and "Cochepaille" of the novel Les Misérables. They are three witnesses in Champmathieu's trial, and appear simultaneously through the court scenes. Nodes (of the graph) that correspond to any two of these three characters satisfy the condition in (52), which, in turn, implies that the graph Laplacian has a 2-sparse eigenvector, and $s_0(\boldsymbol{x}) \geq 3/2$ is a tight uncertainty bound when $\boldsymbol{V}_L$ is selected as GFB.

VII. Conclusions

In this paper, we studied the concept of uncertainty principle for signals defined over graphs. Unlike existing studies we took a non-local and discrete approach, where the vertex and the spectral domain spreads of a signal are defined as the number of nonzero elements of the signal and its GFT, respectively. We derived a lower bound for the total number of nonzero elements in both domains (on the graph and in the GFB) and showed that a signal and its corresponding GFT cannot be arbitrarily sparse simultaneously. Based on this, we obtained a new form of uncertainty principle for graph signals. When the graph has repeated eigenvalues we explained that GFB is not unique, and the derived lower bound can have different values depending on the selected GFB. We provided a constructive method to find a GFB that yields the smallest uncertainty bound. In order to find the signals that achieve the derived lower bound we considered sparse eigenvectors of the graph. We showed that the graph Laplacian has a 2-sparse eigenvector if and only if there exists a pair of nodes with the same neighbors. When this happens,

the uncertainty bound is very low and the 2-sparse eigenvectors achieve this bound. We presented examples of both classical and real-world graphs with 2-sparse eigenvectors. We also discussed that, in some examples, the neighborhood structure has a meaningful interpretation.

ACKNOWLEDGMENT

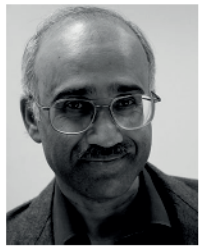
The authors would like to thank the anonymous reviewers for pointing out the study in [19], for thoughtful questions about random graphs and for other very useful suggestions.

REFERENCES

[1] M. E. J. Newman, *Networks: An Introduction*. Oxford, U.K.: Oxford Univ. Press, 2010.
[2] D. Shuman, S. 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, May 2013.
[3] 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, Sep. 2014.
[4] A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs," *IEEE Trans. Signal Process.*, vol. 61, no. 7, pp. 1644–1656, Apr. 2013.
[5] A. Gavili and X.-P. Zhang, "On the shift operator and optimal filtering in graph signal processing," arXiv: 1511.03512v3, Jun. 2016.
[6] S. Sardellitti, S. Barbarossa, and P. Di Lorenzo, "On the graph Fourier transform for directed graphs," *IEEE J. Sel. Topics Signal Process.*, 2017, preprint, doi: 10.1109/JSTSP.2017.2726979.
[7] S. Narang and A. Ortega, "Perfect reconstruction two-channel wavelet filter banks for graph structured data," *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 2786–2799, Jun. 2012.
[8] O. Teke and P. P. Vaidyanathan, "Extending classical multirate signal processing theory to graphs—Part I: Fundamentals," *IEEE Trans. Signal Process.*, vol. 65, no. 2, pp. 409–422, Jan. 2017.
[9] O. Teke and P. P. Vaidyanathan, "Extending classical multirate signal processing theory to graphs—Part II: M-Channel filter banks," *IEEE Trans. Signal Process.*, vol. 65, no. 2, pp. 423–437, Jan. 2017.
[10] O. Teke and P. P. Vaidyanathan, "Fundamentals of multirate graph signal processing," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Nov. 2015, pp. 1791–1795.
[11] O. Teke and P. P. Vaidyanathan, "Graph filter banks with m-channels, maximal decimation, and perfect reconstruction," in *Proc. Int. Conf. Acoust. Speech, Signal Process.*, Mar. 2016, pp. 4089–4093.
[12] X. Wang, P. Liu, and Y. Gu, "Local-set-based graph signal reconstruction," *IEEE Trans. Signal Process.*, vol. 63, no. 9, pp. 2432–2444, May 2015.
[13] D. K. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Appl. Comput. Harmonic Anal.*, vol. 30, no. 2, pp. 129–150, 2011.
[14] A. Anis, A. Gadde, and A. Ortega, "Towards a sampling theorem for signals on arbitrary graphs," in *Proc. Int. Conf. Acoust. Speech, Signal Process.*, May 2014, pp. 3864–3868.
[15] G. B. Folland and A. Sitaram, "The uncertainty principle: A mathematical survey," *J. Fourier Anal. Appl.*, vol. 3, no. 3, pp. 207–238, 1997.
[16] A. Agaskar and Y. M. Lu, "A spectral graph uncertainty principle," *IEEE Trans. Inf. Theory*, vol. 59, no. 7, pp. 4338–4356, Jul. 2013.
[17] M. Tsitsvero, S. Barbarossa, and P. D. Lorenzo, "Signals on graphs: Uncertainty principle and sampling," *IEEE Trans. Signal Process.*, vol. 64, no. 18, pp. 4845–4860, 2016.
[18] P. J. Koprowski, "Graph theoretic uncertainty and feasibility," arXiv: 1603.02059, Feb. 2016.
[19] N. Perraudin, B. Ricaud, D. Shuman, and P. Vandergheynst, "Global and local uncertainty principles for signals on graphs," arXiv: 1603.03030, Mar. 2016.
[20] D. L. Donoho and P. B. Stark, "Uncertainty principles and signal recovery," *SIAM J. Appl. Math*, vol. 49, no. 3, pp. 906–931, 1989.
[21] M. Elad and A. M. Bruckstein, "A generalized uncertainty principle and sparse representation in pairs of bases," *IEEE Trans. Inf. Theory*, vol. 48, no. 9, pp. 2558–2567, Sep. 2002.
[22] M. Elad, *Sparse and Redundant Representations*. New York, NY, USA: Springer, 2013.
[23] O. Teke and P. P. Vaidyanathan, "Discrete uncertainty principles on graphs," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Nov. 2016, pp. 1475–1479.
[24] O. Teke and P. P. Vaidyanathan, "Sparse eigenvectors of graphs," in *Proc. Int. Conf. Acoust. Speech, Signal Process.*, Mar. 2017, pp. 3904–3908.
[25] B. Pasdeloup, R. Alami, V. Gripon, and M. Rabbat, "Toward an uncertainty principle for weighted graphs," in *Proc. Eur. Signal Process. Conf.*, Aug. 2015, pp. 1496–1500.
[26] J. J. Benedetto and P. J. Koprowski, "Graph theoretic uncertainty principles," in *Proc. Int. Conf. Sampling Theory Appl.*, May 2015, pp. 357–361.
[27] B. Ricaud and B. Torrsani, "Refined support and entropic uncertainty inequalities," *IEEE Trans. Inf. Theory*, vol. 59, no. 7, pp. 4272–4279, Jul. 2013.
[28] A. Banerjee and J. Jost, "On the spectrum of the normalized graph Laplacian," *Linear Algorithm Appl.*, vol. 428, pp. 3015–3022, 2008.
[29] V. Ekambaram, G. Fanti, B. Ayazifar, and K. Ramchandran, "Circulant structures and graph signal processing," in *Proc. Int. Conf. Image Process.*, Sep. 2013, pp. 834–838.
[30] V. Ekambaram, G. Fanti, B. Ayazifar, and K. Ramchandran, "Spline-like wavelet filterbanks for multiresolution analysis of graph-structured data," *IEEE Trans. Signal Inf. Process. Over Netw.*, vol. 1, no. 4, pp. 268–278, Dec. 2015.
[31] R. Merris, "Laplacian matrices of graphs: A survey," *Linear Algebra Appl.*, vol. 197, pp. 143–176, 1994.
[32] R. Grone, "On the geometry and Laplacian of a graph," *Linear Algebra Appl.*, vol. 150, pp. 167–178, 1991.
[33] F. K. Bell and P. Rowlinson, "On the multiplicities of graph eigenvalues," *Bull. London Math. Soc.*, vol. 35, no. 3, pp. 401–408, May 2003.
[34] D. Cvetkovic, P. Rowlinson, and S. Simic, *Eigenspaces of Graphs*. Cambridge U.K.: Cambridge Univ. Press, 2008.
[35] M. Muzychuk and M. Klin, "On graphs with three eigenvalues," *Discrete Math.*, vol. 189, no. 1–3, pp. 191–207, 1998.
[36] A. E. Brouwer and W. H. Haemers, *Spectra of Graphs*. New York, NY, USA: Springer, 2011.
[37] J. A. Deri and J. M. F. Moura, "Taxi data in New York city: A network perspective," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Nov 2015, pp. 1829–1833.
[38] T. F. Coleman and A. Pothen, "The null space problem I. Complexity," *SIAM J. Algebr. Discrete Methods*, vol. 7, no. 4, pp. 527–537, 1986.
[39] T. F. Coleman and A. Pothen, "The null space problem II. Algorithms," *SIAM J. Algebr. Discrete Methods*, vol. 8, no. 4, pp. 544–563, 1987.
[40] M. W. Berry, M. T. Heath, I. Kaneko, M. Lawo, R. J. Plemmons, and R. C. Ward, "An algorithm to compute a sparse basis of the null space," *Numer. Math.*, vol. 47, no. 4, pp. 483–504, 1985.
[41] J. R. Gilbert and M. T. Heath, "Computing a sparse basis for the null space," *SIAM J. Algebr. Discrete Methods*, vol. 8, no. 3, pp. 446–459, 1987.
[42] Q. Qu, J. Sun, and J. Wright, "Finding a sparse vector in a subspace: Linear sparsity using alternating directions," *Adv. Neural Inform. Process. Syst.*, vol. 27, pp. 3401–3409, 2014.
[43] R. van der Hofstad, *Random Graphs and Complex Networks: Volume 1*. Cambridge U.K.: Cambridge Univ. Press, 2016.
[44] L. V. Tran, V. H. Vu, and K. Wang, "Sparse random graphs: Eigenvalues and eigenvectors," *Random Structures Algorithms*, vol. 42, no. 1, pp. 110–134, 2013.
[45] S. Narang and A. Ortega, "Graph Bior Wavelet Toolbox," 2013. [Online]. Available: http://biron.usc.edu/wiki/index.php/Graph_Filterbanks
[46] D. E. Knuth, *The Stanford GraphBase: A Platform for Combinatorial Computing*. Reading, MA, USA: Addison-Wesley, 1993.
[47] M. E. J. Newman., "Network data," 2013. [Online]. Available: http://www-personal.umich.edu/~mejn/netdata/

Oguzhan Teke (S'12) received the B.S. and the M.S. degrees in electrical and electronics engineering both from Bilkent University, Ankara, Turkey, in 2012, and 2014, respectively. He is currently working toward the Ph.D. degree in electrical engineering at the California Institute of Technology, Pasadena, CA, USA.

His research interests include signal processing, graph signal processing, and convex optimization.



P. P. Vaidyanathan (S'80–M'83–SM'88–F'91) was born in Calcutta, India, on October 16, 1954. He received the B.Sc. (Hons.) degree in physics and the B.Tech. and M.Tech. degrees in radiophysics and electronics, all from the University of Calcutta, Kolkata, India, in 1974, 1977, and 1979, respectively, and the Ph.D. degree in electrical and computer engineering from the University of California at Santa Barbara, Santa Barbara, CA, USA, in 1982. He was a postdoctoral fellow at the University of California, Santa Barbara, from September 1982 to March 1983. In March 1983, he joined the Electrical Engineering Department, Calfornia Institute of Technology as an Assistant Professor, and since 1993 has been a Professor of electrical engineering there. His main research interests include digital signal processing, multirate systems, wavelet transforms, signal processing for digital communications, genomic signal processing, radar signal processing, and sparse array signal processing.

Dr. Vaidyanathan served as the Vice-Chairman of the Technical Program committee for the 1983 IEEE International Symposium on Circuits and Systems, and as the Technical Program Chairman for the 1992 IEEE International Symposium on Circuits and Systems. He was an Associate Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS for the period 1985–1987, and is currently an Associate Editor for the JOURNAL IEEE SIGNAL PROCESSING LETTERS, and a Consulting Editor for the Journal Applied and Computational Harmonic Analysis. He has been a Guest Editor in 1998 for special issues of the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the IEEE TRANSACTIONS ON SIGNAL PROCESSING ON CIRCUITS AND SYSTEMS II, on the topics of filter banks, wavelets and subband coders. He has authored more than 500 papers in journals and conferences, and is the author/coauthor of the four books Multirate systems and filter banks, Prentice Hall, 1993, Linear Prediction Theory, Morgan and Claypool, 2008, and (with Phoong and Lin) Signal Processing and Optimization for Transceiver Systems, Cambridge University Press, 2010, and Filter Bank Transceivers for OFDM and DMT Systems, Cambridge University Press, 2010. He has written several chapters for various signal processing handbooks. He received the Award for excellence in teaching at the California Institute of Technology for the years 1983–1984, 1992–1993, and 1993–1994. He also received the NSF's Presidential Young Investigator Award in 1986. In 1989, he received the IEEE ASSP Senior Award for his paper on multirate perfect-reconstruction filter banks. In 1990, he received the S. K. Mitra Memorial Award from the Institute of Electronics and Telecommuncations Engineers, India, for his joint paper in the IETE journal. In 2009, he was chosen to receive the IETE students' Journal Award for his tutorial paper in the *IETE Journal of Education*. He was also the coauthor of a paper on linear-phase perfect reconstruction filter banks in the IEEE SP Transactions, for which the first author (Truong Nguyen) received the Young Outstanding Author Award in 1993. He received the 1995 F. E. Terman Award of the American Society for Engineering Education, sponsored by Hewlett Packard Co., for his contributions to engineering education. He has given several plenary talks including at the IEEE ISCAS-04, Sampta-01, Eusipco-98, SPCOM-95, and Asilomar-88 conferences on signal processing. He has been chosen a Distinguished Lecturer for the IEEE Signal Processing Society for the year 1996–1997. In 1999, he was chosen to receive the IEEE CAS Society's Golden Jubilee Medal. He received the IEEE Gustav Kirchhoff Award (an IEEE Technical Field Award) in 2016, for "Fundamental contributions to digital signal processing." In 2016, he received the Northrup Grumman Prize for excellence in Teaching at Caltech. He received the IEEE Signal Processing Society's Technical Achievement Award (2002), Education Award (2012), and the Society Award (2016).