SPARSE EIGENVECTORS OF GRAPHS

Oguzhan Teke and P. P. Vaidyanathan

Department of Electrical Engineering, MC 136-93,
California Institute of Technology, Pasadena, CA 91125, USA
E-mail: oteke@caltech.edu, ppvnath@systems.caltech.edu

ABSTRACT

In order to analyze signals defined over graphs, many concepts from the classical signal processing theory have been extended to the graph case. One of these concepts is the uncertainty principle, which studies the concentration of a signal on a graph and its graph Fourier basis (GFB). An eigenvector of a graph is the most localized signal in the GFB by definition, whereas it may not be localized in the vertex domain. However, if the eigenvector itself is sparse, then it is concentrated in both domains simultaneously. In this regard, this paper studies the necessary and sufficient conditions for the existence of 1, 2, and 3-sparse eigenvectors of the graph Laplacian. The provided conditions are purely algebraic and only use the adjacency information of the graph. Examples of both classical and real-world graphs with sparse eigenvectors are also presented.

Index Terms— Graph signals, sparsity, sparse eigenvectors.

1. INTRODUCTION

Analysis of signals defined over graphs has been of interest in recent years. For signals defined on graphs, each node (of the graph) is associated with data, and the graph is assumed to model the underlying dependency structure between the data sources. 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, 2].

The recent advancements in [3–5] studied the processing of signals defined over graphs. In these studies the analysis is based on the "graph operator," which can be selected as the adjacency matrix as in [4], or the graph Laplacian as in [5]. There are other proposals as well [6, 7]. By using the eigenvectors of the graph operator as the graph Fourier basis (GFB), sampling, reconstruction and multirate processing of graph signals are studied in [8–17]. Apart from these, another important concept in signal analysis is the uncertainty principle [18]. The studies in [19–22] extend this concept to signals defined over graphs.

This paper considers the sparse eigenvectors of the Laplacian of a given graph. By definition, an eigenvector (an element of the GFB) is the most localized signal in the graph Fourier domain. On the other hand, an eigenvector need *not* to be localized in the vertex domain in general. However, if there are sparse eigenvectors, then they are the most concentrated signals in the vertex domain and the GFB simultaneously.

In the search for sparse eigenvectors, one approach is to numerically compute all the eigenvectors of the graph Laplacian, then look for the sparse ones. However, this "brute-force" procedure has two main downsides. For large graphs numerical computation of eigenvectors is a costly operation. More importantly, some graphs have repeated eigenvalues. In this case, a repeated eigenvalue constitutes an eigenspace, hence the corresponding eigenvectors are not unique.

Even if an eigenspace has a sparse eigenvector, finding sparse vectors in a subspace is known to be an NP-hard problem [23–25]. The approach taken in this paper is therefore different: we characterize the sparse eigenvectors of a graph *algebraically*, using only the adjacency information of the graph.

In the following, we first provide a brief review of graph signal processing notation. In Sec. 2, we present the necessary and sufficient conditions for 1- and 2-sparse eigenvectors to exist. In Sec. 3, we consider the 3-sparse case. We also make the connection between 3-sparse and 2-sparse eigenvectors for unweighted graphs. In Sec. 4 we provide classical and real-world graph examples and show that they have sparse eigenvectors.

1.1. Preliminaries and Notation

Let $A \in \mathcal{M}^N$ denote the adjacency matrix of a graph of size N (i.e., N nodes or vertices). We assume the graph does not have self loops, i.e. $a_{i,i} = 0$. The weight of the edge from the jth node to the ith node is denoted by the $(i,j)^{th}$ element of A. A graph is undirected when $a_{i,j} = a_{j,i}$ for all pairs of nodes. For undirected graphs with non-negative edge weights ($a_{i,j} \geqslant 0$), the graph Laplacian is defined as $L = D - A$, where D is the diagonal degree matrix given as $(D)_{i,i} = \sum_j a_{i,j}$. The set of nodes that are adjacent to node i is denoted by $\mathcal{N}(i)$, that is, $\mathcal{N}(i) = \{j \,|\, a_{i,j} \neq 0\}$. For two sets $\mathcal{A}$ and $\mathcal{B}$, the set difference is defined as $\mathcal{A}\backslash\mathcal{B} = \{x \in \mathcal{A} \,|\, x \notin \mathcal{B}\}$. We use $\cup$ to denote the set union operator. Number of elements in a set $\mathcal{A}$ is denoted by $|\mathcal{A}|$. We use $\|x\|_1$ to denote the ℓ_1-norm of the vector x.

In this paper we always consider *undirected* graphs with *non-negative* edge weights. Notice that Theorems 1 and 2 apply to weighted graphs, whereas Theorem 3 and 4 are specific to unweighted graphs. A graph is said to be connected if there is a path between any pair of nodes. In the following, the term "eigenvector" always refers to eigenvectors of the graph Laplacian.

2. SPARSE EIGENVECTORS OF GRAPH LAPLACIAN

When the graph of interest is *disconnected*, it is straightforward to find sparse eigenvectors. To see this, let A be the adjacency matrix of a graph with D disconnected components. Then, under the proper labeling of the nodes, the adjacency matrix and the Laplacian can be written in the block diagonal form:

$$A = \begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_K \end{bmatrix}, \qquad L = \begin{bmatrix} L_1 & & \\ & \ddots & \\ & & L_K \end{bmatrix}, \quad (1)$$

where $A_i \in \mathcal{M}^{N_i}$ and $L_i \in \mathcal{M}^{N_i}$ are the adjacency matrix and the graph Laplacian of the ith component, respectively. Due to block-diagonal form of L, corresponding eigenvectors can be selected to be block sparse. Therefore, for $1 \leqslant i \leqslant D$, there exists an eigenvector that has at most N_i non-zero elements. If there is an arbitrarily small component, then we can find arbitrarily sparse eigenvectors of the

Laplacian. However, this result is valid for only one direction, and the converse is not true: if a graph has a sparse eigenvector, it does *not* imply that the graph is disconnected. This will be clear from Theorems 2 and 3, which prove that a connected graph can have a sparse eigenvector. However, 1-sparse eigenvectors are exceptions in this regard as stated next:

Theorem 1 (Isolated nodes of a graph). *Assume that the graph of interest is undirected and non-negatively weighted. Then, the graph Laplacian has a 1-sparse eigenvector if and only if the graph has an isolated node.*

  Proof: Assume that the graph has an isolated node. According to block diagonal form in (1), there exists a 1-sparse eigenvector of the graph Laplacian.

  For the converse, let v be a 1-sparse eigenvector. Without loss of generality assume that the first index is non-zero $v_1 = 1$, and the rest is zero. Therefore Lv is equivalent to the first column of L. That is, $Lv = [d_1 \ -a_{R,1}^T]^T = \lambda[1 \ \mathbf{0}^T]^T$, where $a_{R,1} \in \mathcal{R}^{N-1}$ is the vector that denotes the adjacency of node 1, and $d_1 = \|a_{R,1}\|_1$ is the degree of node 1. Notice that the eigenvalue equation yields $a_{R,1} = \mathbf{0}$, which means that no other node is adjacent to node 1, that is, node 1 is an isolated node. ∎

  Now we provide the characterization theorem for 2-sparse eigenvectors of *weighted* graphs.

Theorem 2 (2-sparse eigenvectors of a connected graph). *Let 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*

$$a_{i,k} = a_{j,k} \quad \forall \, k \in \{1, \cdots, N\} \backslash \{i, j\}, \tag{2}$$

if and only if the graph Laplacian, L, has a 2-sparse eigenvector with nonzero eigenvalue $\lambda = d_i + a_{i,j}$. The condition (2) reduces to:

$$\mathcal{N}(i) \backslash \{j\} = \mathcal{N}(j) \backslash \{i\}, \tag{3}$$

when the graph is unweighted. ◇

  Proof: Assume that the graph Laplacian of a connected graph has a two-sparse eigenvector 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) 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 A denote the adjacency matrix of the graph. We have

$$A = \begin{bmatrix} 0 & a_{1,2} & a_{R,1}^T \\ a_{2,1} & 0 & a_{R,2}^T \\ a_{R,1} & a_{R,2} & A_R \end{bmatrix}, \quad L = \begin{bmatrix} d_1 & -a_{1,2} & -a_{R,1}^T \\ -a_{2,1} & d_2 & -a_{R,2}^T \\ -a_{R,1} & -a_{R,2} & L_R \end{bmatrix}, \tag{4}$$

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

$$Lv = \begin{bmatrix} d_1 + a_{1,2} \\ -(a_{2,1} + d_2) \\ -a_{R,1} + a_{R,2} \end{bmatrix} = \lambda v = \lambda \begin{bmatrix} 1 \\ -1 \\ \mathbf{0} \end{bmatrix}.$$

Therefore we have $a_{R,1} = a_{R,2}$, 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 $d_1 > 0$, and λ is nonzero. Notice that the condition $a_{R,1} = a_{R,2}$ is the same as the condition in (2) (with $i=1$ and $j=2$).

  Conversely, assume that there exist two nodes with the property in (2). Without loss of generality, assume $i = 1$ and $j = 2$, and let v be a 2-sparse vector with $v_1 = -v_2 = 1$. Then partition the graph Laplacian as in (4). Due to (2), we have $a_{R,1} = a_{R,2}$, and $d_1 = d_2$. Then we have $Lv = \lambda v$ with $\lambda = d_1 + a_{1,2}$. Therefore, v is a 2-sparse eigenvector of L. Using the fact that $d_1 > 0$ for a connected graph, and the assumption that the weights are nonnegative, we conclude $\lambda > 0$. ∎

3. SPARSE EIGENVECTORS OF UNWEIGHTED GRAPHS

In this section we provide the characterization theorem for 3-sparse eigenvectors of *unweighted* graphs. We later show the connection between 3 and 2-sparse eigenvectors of unweighted graphs.

Theorem 3 (3-sparse eigenvectors of a connected graph). *Let A denote the adjacency matrix of an undirected, unweighted, and connected graph. There exist nodes i, j, and k such that*

$$\mathcal{N}(i) \backslash \{j, k\} = \mathcal{N}(j) \backslash \{i, k\} = \mathcal{N}(k) \backslash \{i, j\}, \tag{5}$$

if and only if the graph Laplacian, L, has a 3-sparse eigenvector with non-zero eigenvalue. ◇

  Proof: Assume that the graph Laplacian of a connected graph has a three-sparse eigenvector v with nonzero eigenvalue. Due to permutation invariance of the node labels, without loss of generality assume that $v_1 \neq 0$, $v_2 \neq 0$, $v_3 \neq 0$ but $v_i = 0$ for $i \geq 4$.

  For a connected graph, the all-1 vector is the only eigenvector of the graph Laplacian with the zero eigenvalue. Since the 3-sparse eigenvector (with nonzero eigenvalue) v is orthogonal to the all-1 vector we have $v_1 + v_2 + v_3 = 0$. By scaling the eigenvector, without loss of any generality, we can select $v_1 = 1$, $v_2 = \gamma$, and $v_3 = -1-\gamma$ for some γ, where $\gamma \neq 0$ and $\gamma \neq -1$ since the eigenvector v is exactly 3-sparse.

  Similar to (4), the graph Laplacian can be partitioned as follows

$$L = \begin{bmatrix} d_1 & -a_{1,2} & -a_{1,3} & -a_{R,1}^T \\ -a_{2,1} & d_2 & -a_{2,3} & -a_{R,2}^T \\ -a_{3,1} & -a_{3,2} & d_3 & -a_{R,3}^T \\ -a_{R,1} & -a_{R,2} & -a_{R,3} & L_R \end{bmatrix}, \tag{6}$$

where $L_R \in \mathcal{M}^{N-3}$ is the partition of the Laplacian, and for $1 \leq i \leq 3$, $a_{R,i} \in \mathcal{R}^{N-3}$ is the vector that denotes the adjacency of node i with the nodes $\{4, \cdots, N\}$. Since v is an eigenvector, it should satisfy the following eigenvalue equation:

$$Lv = \begin{bmatrix} d_1 - \gamma \, a_{1,2} + (1+\gamma) \, a_{1,3} \\ -a_{2,1} + \gamma \, d_2 + (1+\gamma) \, a_{2,3} \\ -a_{3,1} - \gamma \, a_{3,2} - (1+\gamma) \, d_3 \\ -a_{R,1} - \gamma \, a_{R,2} + (1+\gamma) \, a_{R,3} \end{bmatrix} = \lambda v = \lambda \begin{bmatrix} 1 \\ \gamma \\ -(1+\gamma) \\ \mathbf{0} \end{bmatrix}. \tag{7}$$

This implies that $a_{R,1} + \gamma \, a_{R,2} - (1+\gamma) \, a_{R,3} = \mathbf{0}$. This vector equation holds true if and only if it is satisfied element-wise. That is

$$a_{r,1} - a_{r,3} = \gamma \, (a_{r,3} - a_{r,2}) \qquad \forall \, r \in \{4, \cdots, N\}. \tag{8}$$

Remember that the graph is assumed to be unweighted, therefore, $a_{r,1}$, $a_{r,2}$, and $a_{r,3}$ are either 1 or 0. Hence, (8) can appear in $2^3 = 8$ different variations, each of which results in a different value for γ. The following table considers each case separately and provides the solution(s) for γ.

$a_{r,1}$	$a_{r,2}$	$a_{r,3}$	γ	Validness
0	0	0	$\mathcal{R}$	✓
0	0	1	-1	✗
0	1	0	0	✗
0	1	1	$-\infty$	✗
1	0	0	∞	✗
1	0	1	0	✗
1	1	0	-1	✗
1	1	1	$\mathcal{R}$	✓

In the table, $\mathcal{R}$ means that any real number is a solution. Remember that $\gamma \neq 0$ and $\gamma \neq -1$ since the eigenvector $\boldsymbol{v}$ is assumed to be exactly 3-sparse. As a result, a solution to (8) exists only if $a_{r,1} = a_{r,2} = a_{r,3}$. Since this is necessary for all $r \in \{4, \cdots, N\}$, we get $\boldsymbol{a}_{R,1} = \boldsymbol{a}_{R,2} = \boldsymbol{a}_{R,3}$. This condition is the same as (5).

Conversely, assume that the condition (5) holds. Without loss of generality assume that $i = 1$, $j = 2$, and $k = 3$. Then we have $\boldsymbol{a}_{R,1} = \boldsymbol{a}_{R,2} = \boldsymbol{a}_{R,3}$, where $\boldsymbol{a}_{R,i}$ is the same as in (6). Define $s = \|\boldsymbol{a}_{R,1}\|_1 = |\mathcal{N}(i)\backslash\{j,k\}|$. Notice that $s > 0$, since $s = 0$ implies that the first three nodes are disconnected from the rest of the graph. Now consider the following eigenvalue equation:

$$\begin{bmatrix} a_{1,2}+a_{1,3} & -a_{1,2} & -a_{1,3} \\ -a_{1,2} & a_{1,2}+a_{2,3} & -a_{2,3} \\ -a_{1,3} & -a_{2,3} & a_{1,3}+a_{2,3} \end{bmatrix} \begin{bmatrix} 1 \\ \gamma \\ -(1+\gamma) \end{bmatrix} = (\lambda - s) \begin{bmatrix} 1 \\ \gamma \\ -(1+\gamma) \end{bmatrix}. \quad (9)$$

Notice that the matrix on the left-hand side is the Laplacian of the subgraph on the first three nodes. Since the graph is unweighted, this matrix can have $2^3 = 8$ different forms. By exhaustively considering each case, one can show that (9) can always be solved for λ and γ with $\gamma \neq 0$ and $\gamma \neq -1$. However, values of both λ and γ depend on the matrix. Eigenvalues of a graph Laplacian are always non-negative, therefore $\lambda - s \geqslant 0$. As a result $\lambda \geqslant s > 0$.

Notice that $d_1 = s + a_{1,2} + a_{1,3}$, $d_2 = s + a_{1,2} + a_{2,3}$, and $d_3 = s + a_{1,3} + a_{2,3}$. Therefore, a pair of (λ, γ) that satisfies (9) also satisfies (7). Hence, using γ solved from (9), a 3-sparse vector $\boldsymbol{v}$ constructed as $v_1 = 1$, $v_2 = \gamma$, $v_3 = -(1+\gamma)$, and $v_i = 0$ for $i \geqslant 4$ is an eigenvector of the graph Laplacian $\boldsymbol{L}$. Furthermore, the corresponding eigenvalue λ (computed via (9)) is nonzero. ∎

There are two remarks regarding Theorems 2 and 3. 1) The existence of sparse eigenvectors does *not* depend on the size and the global structure of the graph. Existence of nodes with the properties in (2) or (5) directly implies the claimed results. 2) The sparse eigenvectors are localized on the graph. If the nodes have the properties in (2) or (5), they must have at least one common neighbor. (This follows from the fact that the graph is connected). Hence, non-zero elements of the eigenvector are *at most* 2 *hops away from each other.*

Similarity between the conditions (3) and (5) encourages us to pursue a more general condition on a (connected) graph so that an eigenvector with an arbitrary number of non-zero elements exists. In fact, such a generalization is possible only as a sufficient condition. However, finding a necessary condition is not easy. The main reason is that it is possible to combine sparser eigenvectors in a given eigenspace in order to achieve less sparse ones. To see this, let K be an arbitrary sparsity $K \geqslant 4$. It can be written as $K = 2m + 3n$ for some integer m and n. Hence, if there exist m 2-sparse and n 3-sparse eigenvectors (with the same eigenvalue and disjoint supports), a linear combination of these 2 and 3-sparse eigenvectors yields a K-sparse eigenvector. Furthermore, m and n are not unique for a given K in general. In short, a K-sparse eigenvector might exist for various different reasons, which makes it difficult to find a necessary condition for a K-sparse eigenvector to exist. In particular, consider the Minnesota road graph (to be studied in Sec. 4.5). It has four orthogonal 2-sparse eigenvectors with eigenvalue 1. (See Fig. 3(a)-3(d).) Since these 2-sparse eigenvectors are in the same eigenspace, any linear combination of these is also an eigenvector. Furthermore, it is apparent from Fig. 3(a)-3(d) that these 2-sparse eigenvectors have disjoint supports. As a result, one can find a 6-sparse eigenvector via a linear combination of three 2-sparse eigenvectors. However, a 6-sparse eigenvector could have been the result of a combination of two 3-sparse eigenvectors (with the same eigenvalue and disjoint supports) as well. This empirically shows that a necessary condition is not easy to obtain for an arbitrary sparsity. Also notice that one can find 4, 6, and 8-sparse eigenvectors via linear combinations of the 2-sparse eigenvectors of Fig. 3(a)-3(d). Unlike the 2-sparse ones, these 4, 6, and 8-sparse eigenvectors are not localized (in terms of number of hops) on the graph. Hence, *a K-sparse eigenvector may not be localized on the graph.*

It is interesting to observe that the condition for 3-sparse eigenvectors is more strict than the condition for 2-sparse eigenvectors for unweighted graphs. We formally state this result as follows.

Theorem 4 (3-sparse implies 2-sparse). *If the Laplacian of an undirected, unweighted and connected graph has a 3-sparse eigenvector, then it has a 2-sparse eigenvector.* ◇

Proof: Assume that the Laplacian of an undirected, unweighted and connected graph has a 3-sparse eigenvector. Then, due to Theorem 3, there exist nodes i, j, k with the condition in (5). Let $\mathcal{S} = \mathcal{N}(i)\backslash\{j,k\} = \mathcal{N}(j)\backslash\{i,k\} = \mathcal{N}(k)\backslash\{i,j\}$. The relations in-between the nodes i, j, k can have 4 different forms. This follows from the fact that there are 4 non-isomorphic simple graphs on 3 nodes (page 4 of [26]). These cases are illustrated Fig. 1.
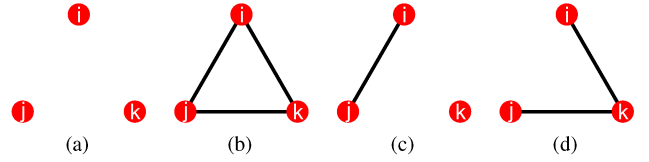

Fig. 1. All four non-isomorphic graphs on 3 nodes.

In the following table, we consider all 4 cases separately and show that there exists a pair of nodes i, j with $\mathcal{N}(i)\backslash\{j\} = \mathcal{N}(j)\backslash\{i\}$.

Case	$\mathcal{N}(i)$	$\mathcal{N}(j)$	$\mathcal{N}(i)\backslash\{j\}$	$\mathcal{N}(j)\backslash\{i\}$
Fig. 1(a)	$\mathcal{S}$	$\mathcal{S}$	$\mathcal{S}$	$\mathcal{S}$
Fig. 1(b)	$\mathcal{S} \cup \{j,k\}$	$\mathcal{S} \cup \{i,k\}$	$\mathcal{S} \cup \{k\}$	$\mathcal{S} \cup \{k\}$
Fig. 1(c)	$\mathcal{S} \cup \{j\}$	$\mathcal{S} \cup \{i\}$	$\mathcal{S}$	$\mathcal{S}$
Fig. 1(d)	$\mathcal{S} \cup \{k\}$	$\mathcal{S} \cup \{k\}$	$\mathcal{S} \cup \{k\}$	$\mathcal{S} \cup \{k\}$

As a result, due to Theorem 2, the graph Laplacian has a 2-sparse eigenvector *independent* of the relation between the nodes i, j, k. ∎

It is important to note that the result of Theorem 4 is specific to 2 and 3-sparse eigenvectors and cannot be generalized to arbitrary sparsity. As a simple counter-example, consider the Minnesota road graph (Sec. 4.5). It has 2-sparse and 4-sparse eigenvectors, but it does *not* have a 3-sparse eigenvector.

4. EXAMPLES

In the following we will provide graph examples that satisfy, and do not satisfy, the conditions in (2) and (5). Notice that the graphs in Sec. 4.1-4.5 are unweighted, whereas the one in Sec. 4.6 is weighted.

4.1. Complete Graph, K_N

A complete graph on N nodes has an edge between any two nodes. Figure 2(a) provides a visual representation of K_8. Let i, j and k be three arbitrary nodes of a complete graph. Then, we have

$$\mathcal{N}(i)\backslash\{j,k\} = \mathcal{N}(j)\backslash\{i,k\} = \mathcal{N}(k)\backslash\{i,j\} = \{1, \cdots, N\}\backslash\{i,j,k\},$$

which shows that an unweighted complete graph of an arbitrary size ($N \geqslant 3$) has a 3-sparse eigenvector, which in particular implies that it has a 2-sparse eigenvector as well (Theorem 4).

4.2. Complete Bi-Partite Graph, $K_{N,M}$

A complete bi-partite graph of size $N + M$ is a bi-partite graph (one color having N nodes, and other color having M nodes) such that every node of a color is connected to every node of the other color. Figure 2(b) provides a visual representation of $K_{4,5}$. Let i, j, and k be three nodes that belong to *the same* color. Then we have that

$$\mathcal{N}(i)\backslash\{j,k\} = \mathcal{N}(j)\backslash\{i,k\} = \mathcal{N}(k)\backslash\{i,j\} = \text{Nodes of the other color},$$

which shows that an unweighted complete bi-partite graph of an arbitrary size (given that a color has at least 3 nodes) has a 3-sparse eigenvector, which in particular implies that it has a 2-sparse eigenvector as well.

4.3. Star Graph, S_N

A star graph of size N is a complete bi-partite graph $K_{1,N\text{-}1}$. In particular, it has a center node that is connected to any other node, and all the nodes are connected only to the center node. Figure 2(c) provides a visual representation of S_9. Assume that the center node is labeled as 1. Let i, j and k be three nodes other than the center node. Then we have $\mathcal{N}(i) = \mathcal{N}(j) = \mathcal{N}(k) = \{1\}$. Therefore,

$$\mathcal{N}(i)\backslash\{j,k\} = \mathcal{N}(j)\backslash\{i,k\} = \mathcal{N}(k)\backslash\{i,j\} = \{1\},$$

which shows that an unweighted star graph of an arbitrary size ($N \geqslant 3$) has a 3-sparse eigenvector, which in particular implies that it has a 2-sparse eigenvector as well (Theorem 4).

4.4. Cycle Graph, C_N

A cycle graph of size N contains a single cycle through all nodes. Figure 2(d) 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 $N \geqslant 5$, C_N does *not* have a pair of nodes that satisfy (3). Therefore, a cycle graph for $N \geqslant 5$ does *not* have a 2-sparse eigenvector, which, in particular, implies that it does *not* have a 3-sparse eigenvector as well (Theorem 4). In fact, it can be formally shown that an eigenvector of a cycle graph of size N has at least $N/2$ non-zero values [22].
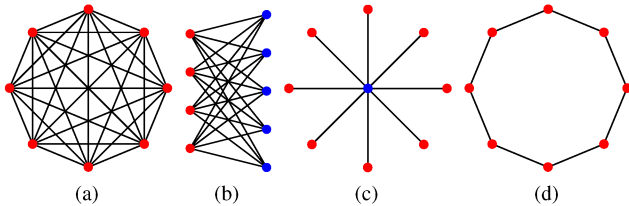

Fig. 2. a) K_8, complete graph of size 8, b) $K_{4,5}$, complete bi-partite graph of size 4-5, c) S_9, star graph of size 9, d) C_8, cycle graph of size of 8.

Above examples are carefully selected to point out an important observation: *sparsity of the graph is not related to the existence of sparse eigenvectors*. 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.

4.5. Minnesota Road Graph

In this example, we consider the Minnesota road graph [8, 17]. We use the data publicly available in [27]. 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 [8, 17] for the visual representation of the graph. This is an unweighted graph where nodes represent intersections, and edges represent roads connecting the intersections. There are total of 3302 undirected unweighted edges.

By using a brute-force search over the graph, we have found that the graph does *not* have a triplet of nodes with the property in (5), hence the graph does not have a 3-sparse eigenvector. However, it *does* have 6 different pairs of nodes with the property in (3). These pairs are visualized in Fig. 3. Hence, the graph has 2-sparse eigenvectors. Notice that the eigenvectors generated by the nodes in Fig. 3(a)-3(d) are orthogonal to each other and have eigenvalue 1.

Using linear combinations of 2-sparse eigenvectors, we can verify that the graph has 4, 6, and 8-sparse eigenvectors as well.
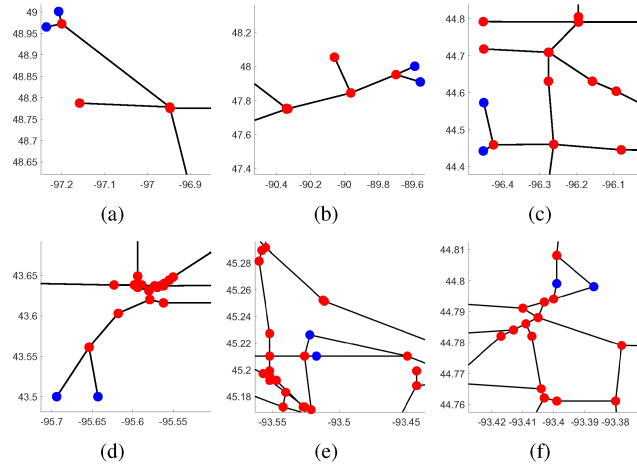

Fig. 3. Pairs of nodes in the Minnesota road graph that result in 2-sparse eigenvectors. The pairs that satisfy the condition in (3) are colored in blue. Notice that the pairs in (a)-(d) generate eigenvectors with eigenvalue 1, and the pairs in (e)-(f) generate eigenvectors with eigenvalue 2. (See Theorem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

4.6. Co-appearance Network

In this example, we consider the co-appearance network of characters in the famous novel Les Misérables by Victor Hugo [28, 29]. 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.

In a co-appearance graph, pairs of nodes with the condition in (2) have 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 (2) 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 (2), which, in turn, implies that the graph Laplacian has a 2-sparse eigenvector.

Since the graph is weighted, we can *not* utilize Theorem 3 in order to find 3-sparse eigenvectors of the Laplacian. Nevertheless, we have experimentally observed that the nodes that correspond to the above-mentioned three characters of the novel constitute a 3-sparse eigenvector!

5. CONCLUSIONS

In this paper, we studied the necessary and sufficient conditions for the existence of 1, 2, and 3-sparse eigenvectors of the Laplacian of an undirected graph. These sparse eigenvectors are important due to their simultaneous localization in the vertex domain and the graph Fourier domain. The presented results for 1 and 2-sparse eigenvectors are valid for weighted graphs, whereas the results on 3-sparse case are specific to unweighted graphs. We presented examples of both classical and real-world graphs with sparse eigenvectors. We further showed that, for unweighted graphs, the existence of a 3-sparse eigenvector implies the existence of a 2-sparse eigenvector. We also provided counter-examples to show that this result does not extend to arbitrary sparsity.

6. REFERENCES

[1] M. E. J. Newman, *Networks: An Introduction.* Oxford University Press, 2010.

[2] M. O. Jackson, *Social and Economic Networks.* Princeton University Press, 2008.

[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 Processing Magazine*, vol. 31, no. 5, pp. 80–90, Sept 2014.

[4] ——, "Discrete signal processing on graphs," *IEEE Trans. Signal Process.*, vol. 61, no. 7, pp. 1644–1656, April 2013.

[5] 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 Processing Magazine*, vol. 30, no. 3, pp. 83–98, May 2013.

[6] A. Gavili and X.-P. Zhang, "On the shift operator and optimal filtering in graph signal processing," *arXiv: 1511.03512v3*, June 2016.

[7] S. Sardellitti and S. Barbarossa, "On the graph fourier transform for directed graphs," *arXiv: 1601.05972v1*, Jan. 2016.

[8] 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, June 2012.

[9] ——, "Compact support biorthogonal wavelet filterbanks for arbitrary undirected graphs," *IEEE Trans. Signal Process.*, vol. 61, no. 19, pp. 4673–4685, Oct 2013.

[10] A. Anis, A. Gadde, and A. Ortega, "Towards a sampling theorem for signals on arbitrary graphs," in *Proc. Int. Conf. Acoust. Speech, Signal Process. (ICASSP)*, May 2014, pp. 3864–3868.

[11] A. Gadde and A. Ortega, "A probabilistic interpretation of sampling theory of graph signals," in *Proc. Int. Conf. Acoust. Speech, Signal Process. (ICASSP)*, April 2015, pp. 3257–3261.

[12] 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.

[13] ——, "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.

[14] ——, "Fundamentals of multirate graph signal processing," in *Asilomar Conf. on Signals, Systems and Computers*, Nov 2015, pp. 1791–1795.

[15] ——, "Graph filter banks with m-channels, maximal decimation, and perfect reconstruction," in *Proc. Int. Conf. Acoust. Speech, Signal Process. (ICASSP)*, March 2016, pp. 4089–4093.

[16] 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.

[17] D. K. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Applied and Computational Harmonic Analysis*, vol. 30, no. 2, pp. 129 – 150, 2011.

[18] G. B. Folland and A. Sitaram, "The uncertainty principle: A mathematical survey," *J. Fourier Anal. Appl.*, vol. 3, no. 3, pp. 207–238, 1997.

[19] A. Agaskar and Y. M. Lu, "A spectral graph uncertainty principle," *IEEE Trans. on Inf. Theory*, vol. 59, no. 7, pp. 4338–4356, July 2013.

[20] P. J. Koprowski, "Graph theoretic uncertainty and feasibility," *arXiv: 1603.02059*, Feb. 2016.

[21] 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.

[22] O. Teke and P. P. Vaidyanathan, "Uncertainty principles and sparse eigenvectors of graphs," *IEEE Trans. Signal Process.*, submitted.

[23] T. F. Coleman and A. Pothen, "The null space problem i. complexity," *SIAM J. on Algebraic Discrete Methods*, vol. 7, no. 4, pp. 527–537, 1986.

[24] ——, "The null space problem ii. algorithms," *SIAM J. on Algebraic Discrete Methods*, vol. 8, no. 4, pp. 544–563, 1987.

[25] Q. Qu, J. Sun, and J. Wright, "Finding a sparse vector in a subspace: Linear sparsity using alternating directions," in *Advances in Neural Information Processing Systems 27*, 2014, pp. 3401–3409.

[26] F. Harary and E. M. Palmer, *Graphical Enumeration.* Academic Press, 1973.

[27] S. Narang and A. Ortega. (2013) Graph bior wavelet toolbox. [Online]. Available: http://biron.usc.edu/wiki/index.php/Graph_Filterbanks

[28] D. E. Knuth, *The Stanford GraphBase: A Platform for Combinatorial Computing.* Addison-Wesley Professional, 1993.

[29] M. E. J. Newman. (2013) Network data. [Online]. Available: http://www-personal.umich.edu/~mejn/netdata/