

A DIGRAPH FOURIER TRANSFORM WITH SPREAD FREQUENCY COMPONENTS

Rasoul Shafipour[†], Ali Khodabakhsh[‡], Gonzalo Mateos[†], and Evdokia Nikolova[‡]

[†]Dept. of Electrical and Computer Engineering, University of Rochester, Rochester, NY, USA

[‡]Dept. of Electrical and Computer Engineering, University of Texas at Austin, Austin, TX, USA

ABSTRACT

We study the problem of constructing a graph Fourier transform (GFT) for directed graphs (digraphs), which decomposes graph signals into different modes of variation with respect to the underlying network. Accordingly, to capture low, medium and high frequencies we seek a digraph (D)GFT such that the orthonormal frequency components are as spread as possible in the graph spectral domain. This specification gives rise to a challenging nonconvex optimization problem, so we resort to a simple yet efficient heuristic to construct the DGFT basis from Laplacian eigenvectors of an undirected version of the digraph. To select frequency components which are as spread as possible, we define a spectral dispersion function and show that it is supermodular. Moreover, we show that orthonormality can be enforced via a matroid basis constraint, which motivates adopting a scalable greedy algorithm to obtain an approximate solution with provable performance guarantee. The effectiveness of the novel DGFT is illustrated through numerical tests on synthetic and real-world graphs.

Index Terms— Graph signal processing, graph Fourier transform, directed graphs, graph frequencies, total variation.

1. INTRODUCTION

Network processes supported on the vertices of a graph can be viewed as graph signals, such as neural activities at different regions of the brain [1], or, infectious states of individuals in a population affected by an epidemic [2]. Under the assumption that the signal properties relate to the underlying graph, the goal of graph signal processing (GSP) is to develop algorithms that fruitfully exploit this relational structure [3, 4]. From this vantage point, generalizations of traditional signal processing tasks such as filtering [4–7], sampling and reconstruction [8, 9], spectrum estimation [10], (blind) filter identification [11, 12] as well as signal representations [13, 14] have been recently explored in the GSP literature.

An instrumental GSP tool is the graph Fourier transform (GFT), which decomposes a graph signal into orthonormal components describing different modes of variation with respect to the graph topology. Here we aim to generalize the GFT to directed graphs (digraphs); see also [15], [16]. We first propose a novel notion of signal variation (frequency) over digraphs and find an approximation of the maximum possible frequency ($f_{\max}$) that a unit-norm graph signal can achieve. We design a digraph (D)GFT such that the resulting frequencies (i.e., the directed variation of the sought orthonormal bases) distribute as evenly as possible across $[0, f_{\max}]$. To be more specific, we introduce some notations and basic GSP notions.

Notation. We consider a weighted digraph $\mathcal{G} = (\mathcal{V}, \mathbf{A})$, where $\mathcal{V}$ is the set of nodes (i.e., vertices) with cardinality $|\mathcal{V}| = N$, and $\mathbf{A} \in \mathbb{R}^{N \times N}$ is the graph adjacency matrix with entry A_{ij} denoting the edge weight from node i to node j . We assume that the graph is connected and has no self loops; i.e. $A_{ii} = 0$, and the edge weights are non-negative ($A_{ij} \geq 0$). We construct the underlying undirected graph $\mathcal{G}^u = (\mathcal{V}, \mathbf{A}^u)$ by replacing all directed

edges of $\mathcal{G}$ with undirected ones. Let $\mathbf{A}^u \in \mathbb{R}^{N \times N}$ denote the symmetric adjacency matrix of $\mathcal{G}^u$, where by convention we set $A_{ij}^u = A_{ji}^u = \max(A_{ij}, A_{ji})$. Then, the positive semi-definite Laplacian matrix takes the form $\mathbf{L} := \mathbf{D} - \mathbf{A}^u$, where $\mathbf{D}$ is the diagonal degree matrix with $D_{ii} = \sum_j A_{ji}^u$. A graph signal $\mathbf{x} : \mathcal{V} \mapsto \mathbb{R}^N$ can be represented as a vector of size N , where component x_i denotes the signal value at node $i \in \mathcal{V}$.

Related work. For undirected graphs, the GFT of signal $\mathbf{x}$ can be defined as $\tilde{\mathbf{x}} = \mathbf{V}^T \mathbf{x}$, where $\mathbf{V} := [\mathbf{v}_1, \dots, \mathbf{v}_N]$ comprises the eigenvectors of the Laplacian [3]. Defining the total variation of the signal $\mathbf{x}$ with respect to the Laplacian $\mathbf{L}$ as

$$\text{TV}(\mathbf{x}) = \mathbf{x}^T \mathbf{L} \mathbf{x} = \sum_{i,j=1, j>i}^N A_{ij}^u (x_i - x_j)^2 \quad (1)$$

then it follows that the total variation of eigenvector $\mathbf{v}_k$ is $\text{TV}(\mathbf{v}_k) = \lambda_k$, the k^{th} Laplacian eigenvalue. Hence, eigenvalues $0 = \lambda_1 < \lambda_2 \leq \dots \leq \lambda_N$ can be viewed as graph frequencies, indicating how the GFT bases vary over the graph. Note that there may be more than one eigenvector corresponding to a graph frequency in case of having repeated eigenvalues. A more general GFT definition is based on the Jordan decomposition of adjacency matrix $\mathbf{A} = \mathbf{V} \mathbf{J} \mathbf{V}^{-1}$, where the frequency representation of graph signal $\mathbf{x}$ is $\tilde{\mathbf{x}} = \mathbf{V}^{-1} \mathbf{x}$ [15]. While valid for digraphs, the associated notion of signal variation in [15] does not ensure that constant signals have zero variation. Moreover, $\mathbf{V}$ is not necessarily orthonormal and Parseval's identity does not hold. From a computational standpoint, obtaining the Jordan decomposition is expensive and often numerically unstable; see also [17]. Recently, a fresh look to the GFT for digraphs was put forth in [16] based on minimization of the (convex) Lovász extension of the graph cut size, subject to orthonormality constraints on the sought bases. However, the solution to such an optimization problem may not be unique. Also, the definition of cut size (and its Lovász extension which can be interpreted as a graph directed variation measure) is based on a bipartition of the graph, while the network may have multiple (more than two) clusters. While the GFT bases in [16] tend to be constant across clusters of the graph, in general they may fail to yield signal representations capturing different modes of signal variation with respect to $\mathcal{G}$; see also Remark 1.

Contributions. Here instead we introduce a novel DGFT (Section 2) that has the following desirable properties: P1) The bases offer notions of frequency and signal variation over digraphs which are also consistent with those used for subsumed undirected graphs. P2) Frequencies are designed to be (approximately) equidistributed in $[0, f_{\max}]$, to better capture low, middle, and high frequencies. P3) Bases are orthonormal so Parseval's identity holds and inner products are preserved in the vertex and graph frequency domains. Moreover, the inverse DGFT can be easily computed. To construct a DGFT basis with the aforementioned properties, a greedy algorithm is outlined in Section 3 which is simple (thus scalable to large graphs) and efficient, while it offers provable performance guarantees.

2. PRELIMINARIES AND PROBLEM STATEMENT

In this section we extend the notion of signal variation to digraphs and accordingly define graph frequencies. We then state the problem as one of finding an orthonormal basis with prescribed (approximately) equidistributed frequencies in the graph spectral domain.

Our goal is to find N orthonormal bases capturing different modes of variation over the graph $\mathcal{G}$. We collect these desired bases in a matrix $\mathbf{U} := [\mathbf{u}_1, \dots, \mathbf{u}_N] \in \mathbb{R}^{N \times N}$, where $\mathbf{u}_i \in \mathbb{R}^N$ represents the i^{th} frequency component. For undirected graphs, the quantity $\text{TV}(\mathbf{x})$ in (1) measures how signal $\mathbf{x}$ varies over the network with Laplacian $\mathbf{L}$. This motivates defining a more general notion of signal variation for digraphs, called directed variation (DV), as (cf. [16, eq. (2)])

$$\text{DV}(\mathbf{x}) := \sum_{i,j=1}^N A_{ij} [\mathbf{x}_i - \mathbf{x}_j]_+^2, \quad (2)$$

where $[x]_+ := \max(0, x)$. To better appreciate (2), consider a graph signal $\mathbf{x} \in \mathbb{R}^N$ on digraph $\mathcal{G}$ and suppose a directed edge represents the direction of signal flow from a larger value to a smaller one. Thus, an edge from node i to node j (i.e., $A_{ij} > 0$) contributes to $\text{DV}(\mathbf{x})$ only if $x_i > x_j$. Notice that if $\mathcal{G}$ is undirected, then $\text{DV}(\mathbf{x}) \equiv \text{TV}(\mathbf{x})$. Analogously to the undirected case, we define the frequency $f_i := \text{DV}(\mathbf{u}_i)$ as the directed variation of the basis $\mathbf{u}_i$.

Problem statement. Similar to the discrete spectrum of periodic time-varying signals, ideally we would like to have N equidistributed graph frequencies forming an arithmetic sequence

$$f_k = \text{DV}(\mathbf{u}_k) = \frac{k-1}{N-1} f_{\max}, \quad k = 1, \dots, N \quad (3)$$

where $f_{\max}$ is the maximum variation of a unit-norm graph signal on $\mathcal{G}$. Accordingly, given $\mathcal{G}$ and for prescribed graph spectrum $\{f_k\}_{k=1}^N$, the DGFT basis $\mathbf{U}$ can be found by solving the following non-convex optimization problem ($\mathbf{I}_N$ is the $N \times N$ identity matrix)

$$\mathbf{U} = \underset{\mathbf{U} \in \mathbb{R}^{N \times N}}{\text{argmin}} \sum_{k=1}^N (\text{DV}(\mathbf{u}_k) - f_k)^2, \quad \text{s. t. } \mathbf{U}^T \mathbf{U} = \mathbf{I}_N. \quad (\mathcal{P}1)$$

Problem ($\mathcal{P}1$) can be tackled by splitting methods for orthogonality constrained problems, e.g. [18]. However those methods are computationally expensive and do not offer convergence guarantees for non-convex objectives such as in ($\mathcal{P}1$). In Section 3, we will propose a simple yet efficient heuristic to construct the DGFT basis from Laplacian eigenvectors of $\mathcal{G}^u$.

Going back to design of the graph spectrum $\{f_k\}_{k=1}^N$, a few noteworthy challenges remain. First, attaining the exact frequencies in (3) may be impossible. This can be clearly seen for undirected graphs, where one has the additional constraint that the summation of frequencies is constant, since $\sum_{k=1}^N f_k = \sum_{k=1}^N \text{TV}(\mathbf{u}_k) = \text{trace}(\mathbf{L})$. Second, one needs to determine the maximum frequency $f_{\max}$ that a unit-norm basis can attain. In $\mathcal{G}^u$, it is immediate that

$$f_{\max}^u = \max_{\|\mathbf{u}\|=1} \text{TV}(\mathbf{u}) = \max_{\|\mathbf{u}\|=1} \mathbf{u}^T \mathbf{L} \mathbf{u} = \lambda_{\max}, \quad (4)$$

where $\lambda_{\max}$ is the largest eigenvalue of the Laplacian matrix $\mathbf{L}$. However, finding the maximum directed variation is in general challenging, since one needs to solve the (non-convex) spherically-constrained problem

$$\mathbf{u}^* = \underset{\|\mathbf{u}\|=1}{\text{argmax}} \quad \text{DV}(\mathbf{u}) \quad \text{and} \quad f_{\max} := \text{DV}(\mathbf{u}^*). \quad (5)$$

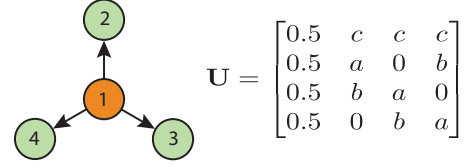


Fig. 1: A toy digraph and the GFT basis $\mathbf{U}$ from [16], where $a = (\sqrt{5} + 1)/4$, $b = (1 - \sqrt{5})/4$ and $c = -0.5$.

Still, it follows that $f_{\max}$ can be upper-bounded by $\lambda_{\max}$. This is because dropping the direction of any edge can not decrease the directed variation, hence $\lambda_{\max}$ gives an upper bound for the maximum DV. In the next section, we propose a way to find a basis with variation of at least $f_{\max}/2$. But before moving on, a remark is in order.

Remark 1 (Motivation for spread frequencies) Consider the toy graph with 4 nodes shown in Fig. 1 (left). Let $\text{DV}'(\mathbf{x}) := \sum_{i,j=1}^N A_{ij} [\mathbf{x}_i - \mathbf{x}_j]_+$ be a directed variation measure, and consider minimizing $\sum_{k=1}^N \text{DV}'(\mathbf{u}_k)$, s. t. $\mathbf{U}^T \mathbf{U} = \mathbf{I}_N$ as in [16]. The optimum basis $\mathbf{U}$ is shown in Fig. 1 (right), where $a = (\sqrt{5} + 1)/4 \approx 0.8090$, $b = (1 - \sqrt{5})/4 \approx -0.3090$, and $c = -0.5$. These values are chosen such that $a + b + c = 0$, $a^2 + b^2 + c^2 = 1$, and $c^2 + ab = 0$ which implies the orthonormality of $\mathbf{U}$. As a result, all columns of $\mathbf{U}$ satisfy $\text{DV}'(\mathbf{u}_k) = 0$, $k = 1, \dots, 4$, and the synthesis formula $\mathbf{x} = \mathbf{U} \tilde{\mathbf{x}}$ fails to offer an expansion of $\mathbf{x}$ with respect to *different* modes of variation (e.g., low and high graph frequencies).

3. A DIGRAPH FOURIER TRANSFORM HEURISTIC

As mentioned in Section 2, one challenge in finding a well distributed set of frequencies on a digraph is to calculate the maximum frequency $f_{\max}$ [cf. (5)]. Here we show how to find a basis vector $\tilde{\mathbf{u}}$ with an approximate $\tilde{f}_{\max} := \text{DV}(\tilde{\mathbf{u}})$ which is at least half of $f_{\max}$.

Proposition 1 For a digraph $\mathcal{G}$, recall $\mathcal{G}^u$ and the spectral radius $\lambda_{\max}$ of its Laplacian $\mathbf{L}$. Let $\mathbf{u}$ be the dominant eigenvector of $\mathbf{L}$. Then,

$$\tilde{f}_{\max} := \max \{ \text{DV}(\mathbf{u}), \text{DV}(-\mathbf{u}) \} \geq \frac{f_{\max}}{2}. \quad (6)$$

Proof: First recall that

$$\lambda_{\max} = \mathbf{u}^T \mathbf{L} \mathbf{u} = \sum_{i,j=1, j>i}^N A_{ij}^u (u_i - u_j)^2 = \frac{1}{2} \sum_{i,j=1}^N A_{ij}^u (u_i - u_j)^2.$$

Since $A_{ij}^u \leq A_{ij} + A_{ji}$ and $(u_i - u_j)^2 = [u_i - u_j]_+^2 + [u_j - u_i]_+^2$, then

$$\begin{aligned} \lambda_{\max} &\leq \frac{1}{2} \sum_{i,j=1}^N (A_{ij} + A_{ji}) \left([u_i - u_j]_+^2 + [u_j - u_i]_+^2 \right) \\ &= \frac{1}{2} (2 \times \text{DV}(\mathbf{u}) + 2 \times \text{DV}(-\mathbf{u})) = \text{DV}(\mathbf{u}) + \text{DV}(-\mathbf{u}). \end{aligned}$$

In conclusion, at least one of $\text{DV}(\mathbf{u})$ or $\text{DV}(-\mathbf{u})$ is larger than $\lambda_{\max}/2$, and this completes the proof since $\lambda_{\max} \geq f_{\max}$. ■

In practice, we can compute $\max \{ \text{DV}(\mathbf{u}), \text{DV}(-\mathbf{u}) \}$ for *any* eigenvector $\mathbf{u}$ of the Laplacian matrix. This will possibly give a higher frequency in $\mathcal{G}$, while preserving the 1/2-approximation.

Fixing $f_1 = 0$ and $f_N = \tilde{f}_{\max}$, in lieu of (3) we will henceforth construct a disperse set of frequencies by using the eigenvectors of $\mathbf{L}$. Let $f_i := \text{DV}(\mathbf{u}_i)$ and $\tilde{f}_i := \text{DV}(-\mathbf{u}_i)$, where $\mathbf{u}_i$ is the

i^{th} eigenvector of $\mathbf{L}$. Define the set of all candidate frequencies as $F := \{f_i, \bar{f}_i : 1 < i < N\}$. The goal is to select $N - 2$ frequencies from F such that together with $\{0, \bar{f}_{\max}\}$ they form our Fourier frequencies. To preserve orthonormality, we would opt exactly one from each pair $\{f_i, \bar{f}_i\}$. We will argue later that this induces a matroid basis constraint.

Optimization scheme. To find the DGFT basis, we define a spectral dispersion function that measures how well spread the corresponding frequencies are over $[0, \bar{f}_{\max}]$. For frequency set $S \subseteq F$, let $s_1 \leq s_2 \leq \dots \leq s_m$ be the elements of S in non-decreasing order, where $m = |S|$. Then we define the dispersion of S as

$$\delta(S) = \sum_{i=0}^m (s_{i+1} - s_i)^2, \quad (7)$$

where $s_0 = 0$ and $s_{m+1} = \bar{f}_{\max}$. It can be verified that $\delta(S)$ is a monotone non-increasing function, which means that for any sets $S_1 \subseteq S_2$, we have $\delta(S_1) \geq \delta(S_2)$. For a fixed value of m , one can show that $\delta(S)$ is minimized when the s_i 's form an arithmetic sequence, consistent with our design goal. Hence, we seek to minimize $\delta(S)$ through a set function optimization procedure. In Lemma 1, we show that the dispersion function (7) has the supermodular property. First, we define submodularity/supermodularity.

Definition 1 (Submodularity) Let S be a finite ground set. A set function $f : 2^S \mapsto \mathbb{R}$ is submodular if:

$$f(T_1 \cup \{e\}) - f(T_1) \geq f(T_2 \cup \{e\}) - f(T_2), \quad (8)$$

for all subsets $T_1 \subseteq T_2 \subseteq S$ and any element $e \in S \setminus T_2$.

A set function f is said to be supermodular if $-f$ is submodular, i.e. (8) holds in the other direction. Submodularity captures the *diminishing returns* property that arises in many applications such as facility location, sensor placement, feature selection [19].

Lemma 1 The spectral dispersion function $\delta : 2^F \mapsto \mathbb{R}$ defined in (7) is a supermodular function.

Proof: Consider two subsets S_1, S_2 such that $S_1 \subseteq S_2 \subseteq F$, and a single element $e \in F \setminus S_2$. Let s_1^L and s_1^R be the largest value smaller than e and smallest value greater than e in $S_1 \cup \{0, \bar{f}_{\max}\}$, respectively (i.e., e breaks the gap between s_1^L and s_1^R). Similarly, let s_2^L and s_2^R be defined for S_2 . Since $S_1 \subseteq S_2$, then $s_1^L \leq s_2^L \leq e \leq s_2^R \leq s_1^R$. The result follows by comparing the marginal values

$$\begin{aligned} \delta(S_1 \cup \{e\}) - \delta(S_1) &= (s_1^R - e)^2 + (e - s_1^L)^2 - (s_1^R - s_1^L)^2 \\ &= -2(s_1^R - e)(e - s_1^L) \leq -2(s_2^R - e)(e - s_2^L) \\ &= \delta(S_2 \cup \{e\}) - \delta(S_2). \end{aligned}$$

■

Recalling the orthonormality constraint, we define $\mathcal{B}$ to be the set of all subsets $S \subseteq F$ that satisfy $|S \cap \{f_i, \bar{f}_i\}| = 1, i = 2, \dots, N - 1$. Then, frequency selection from F boils down to solving

$$\min_S \delta(S), \quad \text{s. t. } S \in \mathcal{B}. \quad (\mathcal{P}2)$$

Next, in Lemma 2 we show that the constraint in $(\mathcal{P}2)$ is a matroid basis constraint. But first, we define the notion of partition matroid.

Definition 2 (Partition matroid [20]) Let S denote a finite set, and let $S_1, \dots, S_m$ be a partition of S , i.e. a collection of disjoint sets such that $S_1 \cup \dots \cup S_m = S$. Let $d_1, \dots, d_m$ be a collection of non-negative integers. Define a set $\mathcal{I}$ by $A \in \mathcal{I}$ iff $|A \cap S_i| \leq d_i$ for all $i = 1, \dots, m$. Then, $\mathcal{M} = (S, \mathcal{I})$ is called the partition matroid.

Algorithm 1 Greedy Dispersion Minimization

- 1: **Input:** Set of possible frequencies F .
 - 2: **Initialize** $S = \emptyset$.
 - 3: **repeat**
 - 4: $e \leftarrow \operatorname{argmax}_{f \in F} \{\tilde{\delta}(S \cup \{f\}) - \tilde{\delta}(S)\}$.
 - 5: $S \leftarrow S \cup \{e\}$.
 - 6: Delete from F the pair $\{f_i, \bar{f}_i\}$ that e belongs to.
 - 7: **until** $F = \emptyset$
-

Lemma 2 There exists a (partition) matroid $\mathcal{M}$ such that the set $\mathcal{B}$ in $(\mathcal{P}2)$ is the set of all bases of $\mathcal{M}$.

Proof: Recall Definition 2 and set $S = F$, $S_i = \{f_i, \bar{f}_i\}$ and $d_i = 1$ for all $i = 2, \dots, N - 1$, to get a partition matroid $\mathcal{M} = (F, \mathcal{I})$. Moreover, the bases of $\mathcal{M}$, which are defined as the maximal elements of $\mathcal{I}$, are those subsets $A \subseteq F$ that satisfy $|A \cap \{f_i, \bar{f}_i\}| = 1$ for all $i = 2, \dots, N - 1$, which are the elements of $\mathcal{B}$. ■

Lemmas 1 and 2 assert that $(\mathcal{P}2)$ is a supermodular minimization problem subject to a matroid basis constraint. Since supermodular minimization is NP-hard and hard to approximate to any factor [21, 22], we create a submodular function $\tilde{\delta}(S)$ and use the algorithms for submodular maximization to find a set of disperse bases $\mathcal{U}$. In particular, we define

$$\tilde{\delta}(S) := \bar{f}_{\max}^2 - \delta(S), \quad (9)$$

which is a non-negative (increasing) submodular function, because $\delta(\emptyset) = \bar{f}_{\max}^2$ is an upper bound for $\delta(S)$. There are several results for maximizing submodular functions under matroid constraints for both the non-monotone [23] and monotone cases [24, 25]. We adopt the greedy algorithm of [24] due to simplicity (tabulated under Algorithm 1), which offers the following performance guarantee.

Theorem 1 ([24]) Let S^* be the solution of problem $(\mathcal{P}2)$ and S^g be the output of the greedy Algorithm 1. Then,

$$\tilde{\delta}(S^g) \geq \frac{1}{2} \times \tilde{\delta}(S^*). \quad (10)$$

Notice that Theorem 1 offers a worst-case guarantee, and Algorithm 1 is usually able to find near-optimal solutions in practice.

All in all, the DGFT basis construction algorithm entails the following steps. First, we form $\mathcal{G}^u$ and find the eigenvectors of $\mathbf{L}$. Second, the set F is formed by calculating DV for each eigenvector $\mathbf{u}_i$ and its negative $-\mathbf{u}_i, i = 2, \dots, N - 1$. Finally, the greedy Algorithm 1 is run on the set F , and the output determines the set of frequencies as well as the orthonormal set of DGFT bases.

4. NUMERICAL RESULTS

Here we study the performance of Algorithm 1 to construct the DGFT basis, via simulations on two graphs.

Synthetic graph. First, we construct the DGFT for a digraph with $N = 15$ nodes and compare it with the GFT in [16]. Fig. 3 shows all the bases and their frequencies derived from Algorithm 1 and (2), respectively. The DGFT can better capture low, middle, and high frequencies in comparison with the variations exhibited by the bases learned via algorithm in [16]; see Table 1. For fairness of comparison, we calculate the directed variation of the proposed bases using DV' ; see Remark 1. Apparently, the DGFT bases span a wider range of variations. Moreover, we rescale DV' values in Table 1 to the

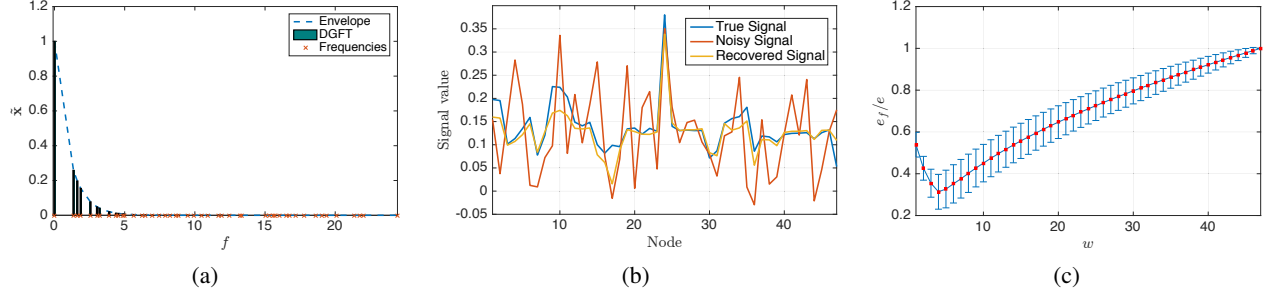


Fig. 2: (a) A synthetic signal defined in the graph spectral domain by $\tilde{x}_f = ce^{-f}$. (b) True, noisy, and recovered signal values at different nodes for a sample realization. (c) Average recovery error using windowed filter over recovery error without filtering versus the window size.

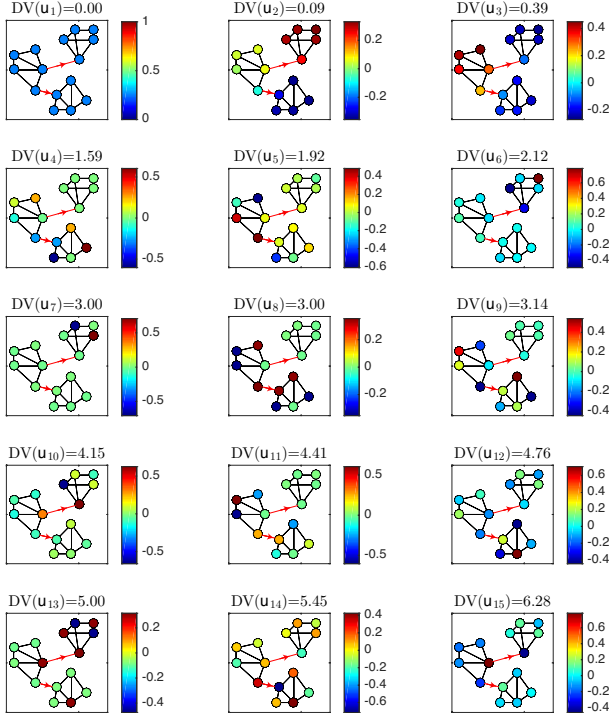


Fig. 3: DGFT bases obtained using Algorithm 1, along with their respective DV values (frequencies).

$[0, 1]$ interval and calculate their dispersion using (7). The dispersion of the DGFT bases is 0.16 and the counterpart in [16] gives a dispersion of 0.24, which confirms the proposed method yields a better (i.e. smaller) frequency spread.

Brain graph. Next, we consider a real brain graph to demonstrate the effectiveness of the DGFT in a denoising task. This graph represents an anatomical connection network of the macaque cortex. It consists of 47 nodes and 505 edges (among which 121 links are directed), which is used e.g. in [1, 26]. The vertices represent different hubs in the brain, and the edges capture directed information flow among them. Let $\mathbf{U}$ be the orthonormal DGFT basis obtained via Algorithm 1. We construct a smooth graph signal in the spectral domain as $\tilde{x}_f = ce^{-f}$, where c is calculated such that the graph signal $\mathbf{x} = \mathbf{U}\tilde{\mathbf{x}}$ has unit norm. The DGFT of signal $\mathbf{x}$ is shown in Fig. 2 (a). We aim to recover the signal from noisy measurements $\mathbf{y} = \mathbf{x} + \mathbf{n}$, where the additive noise $\mathbf{n}$ is a zero-mean, Gaussian random vector with covariance matrix $10^{-2}\mathbf{I}_N$. To that end, we use a low-pass filter $\tilde{\mathbf{H}} := \text{diag}(\tilde{\mathbf{h}})$, where $\tilde{h}_i = \mathbb{1}[i \leq w]$ and w is the

Basis	DV' Alg. 1	DV' [16]	Basis	DV' Alg. 1	DV' [16]
1	0	0	9	6.01	3.53
2	0.97	0	10	6.53	3.67
3	2.19	0	11	6.55	4.08
4	4.12	2.24	12	7.78	4.25
5	4.25	2.46	13	7.90	4.62
6	4.35	3	14	8.09	4.62
7	4.90	3.37	15	8.70	4.98
8	5.66	3.46			

Table 1: Directed variations (frequencies) of obtained bases.

spectral window size. The filter selects the first w components of the signal's DGFT, and we approximate the noisy signal by

$$\hat{\mathbf{x}} = \mathbf{U}\tilde{\mathbf{H}}\tilde{\mathbf{y}} = \mathbf{U}\tilde{\mathbf{H}}\mathbf{U}^T\mathbf{y}. \quad (11)$$

Fig. 2(b) shows a sample noisy signal ($\mathbf{y}$) along with the recovered signal ($\hat{\mathbf{x}}$) using (11) with $w = 4$, and the original signal ($\mathbf{x}$). Filter design and choice of w is out of scope of this paper, but as we can see $\hat{\mathbf{x}}$ closely approximates $\mathbf{x}$. Moreover, we compute the relative recovery error with and without low-pass filtering as $e_f = \|\hat{\mathbf{x}} - \mathbf{x}\|/\|\mathbf{x}\|$ and $e = \|\mathbf{n}\|/\|\mathbf{x}\|$, respectively. Fig. 2(c) depicts e_f/e versus w averaged over 1000 Monte-Carlo simulations, and demonstrates the effectiveness of adopting filters along with the proposed DGFT.

5. CONCLUSION

In this paper, we introduced a new method to find an orthonormal set of graph Fourier bases for digraphs. To that end, we proposed a measure of directed variation to capture the notion of frequency. Our approach is to find a basis that spans the entire frequency range and for which frequency components are as evenly distributed as possible. We showed that the maximum directed variation can be approximately achieved, and then used the eigenvectors of the Laplacian matrix of the underlying undirected graph to find the Fourier basis. We defined a dispersion function to quantify the quality of any feasible solution compared to our ideal design, and proposed a fast greedy algorithm to minimize this dispersion. The greedy algorithm offers theoretical approximation guarantees by virtue of matroid theory and results for submodular function optimization.

With regards to future directions, the complexity of finding the maximum frequency ($f_{\max}$) on a digraph is an interesting open question. If NP-hard, it will be interesting to find the best achievable approximation factor (here we gave a $1/2$ -approximation). Furthermore, our performance guarantee measures the quality of the proposed basis with respect to the optimal set one can get by using the eigenvectors of the Laplacian matrix. A significant improvement would be to generalize this guarantee to any orthonormal basis, or otherwise establish that the global minimizer of dispersion (among all orthonormal bases) lies within the considered feasible set $\mathcal{B}$.

6. REFERENCES

- [1] C. J. Honey, R. Kötter, M. Breakspear, and O. Sporns, "Network structure of cerebral cortex shapes functional connectivity on multiple time scales," *Proceedings of the National Academy of Sciences*, vol. 104, no. 24, pp. 10 240–10 245, 2007.
- [2] E. D. Kolaczyk and G. Csárdi, *Statistical analysis of network data with R*. Springer, 2014, vol. 65.
- [3] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE Signal Processing Magazine*, vol. 30, no. 3, pp. 83–98, 2013.
- [4] A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs," *IEEE Transactions on Signal Processing*, vol. 61, no. 7, pp. 1644–1656, 2013.
- [5] O. Teke and P. P. Vaidyanathan, "Extending classical multirate signal processing theory to graphs-Part I: Fundamentals," *IEEE Transactions on Signal Processing*, vol. 65, no. 2, pp. 409–422, 2017.
- [6] W. Huang, L. Goldsberry, N. F. Wymbs, S. T. Grafton, D. S. Bassett, and A. Ribeiro, "Graph frequency analysis of brain signals," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 7, pp. 1189–1203, 2016.
- [7] O. Teke and P. P. Vaidyanathan, "Extending classical multirate signal processing theory to graphs-Part II: M-channel filter banks," *IEEE Transactions on Signal Processing*, vol. 65, no. 2, pp. 423–437, 2017.
- [8] A. G. Marques, S. Segarra, G. Leus, and A. Ribeiro, "Sampling of graph signals with successive local aggregations," *IEEE Transactions on Signal Processing*, vol. 64, no. 7, pp. 1832–1843, 2016.
- [9] S. Chen, R. Varma, A. Sandryhaila, and J. Kovačević, "Discrete signal processing on graphs: Sampling theory," *IEEE Transactions on Signal Processing*, vol. 63, no. 24, pp. 6510–6523, 2015.
- [10] A. G. Marques, S. Segarra, G. Leus, and A. Ribeiro, "Stationary graph processes and spectral estimation," *arXiv preprint arXiv:1603.04667*, 2016.
- [11] R. Shafipour, S. Segarra, A. G. Marques, and G. Mateos, "Network topology inference from non-stationary graph signals," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2017.
- [12] S. Segarra, G. Mateos, A. G. Marques, and A. Ribeiro, "Blind identification of graph filters," *IEEE Transactions on Signal Processing*, vol. 65, no. 5, pp. 1146–1159, 2017.
- [13] D. Thanou, D. I. Shuman, and P. Frossard, "Learning parametric dictionaries for signals on graphs," *IEEE Transactions on Signal Processing*, vol. 62, no. 15, pp. 3849–3862, 2014.
- [14] X. Zhu and M. Rabbat, "Approximating signals supported on graphs," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2012, pp. 3921–3924.
- [15] A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs: Frequency analysis," *IEEE Transactions on Signal Processing*, vol. 62, no. 12, pp. 3042–3054, June 2014.
- [16] S. Sardellitti, S. Barbarossa, and P. Di Lorenzo, "Graph fourier transform for directed graphs based on lovász extension of min-cut," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2017.
- [17] J. A. Deri and J. M. F. Moura, "Spectral projector-based graph fourier transforms," *arXiv preprint arXiv:1701.02690*, 2017.
- [18] R. Lai and S. Osher, "A splitting method for orthogonality constrained problems," *Journal of Scientific Computing*, vol. 58, no. 2, pp. 431–449, 2014.
- [19] G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher, "An analysis of approximations for maximizing submodular set functions—I," *Mathematical Programming*, vol. 14, no. 1, pp. 265–294, 1978.
- [20] A. Schrijver, *Combinatorial optimization: polyhedra and efficiency*. Springer Science & Business Media, 2002, vol. 24.
- [21] J. A. Kelner and E. Nikolova, "On the hardness and smoothed complexity of quasi-concave minimization," in *48th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, 2007, pp. 472–482.
- [22] S. Mittal and A. S. Schulz, "An FPTAS for optimizing a class of low-rank functions over a polytope," *Mathematical Programming*, pp. 1–18, 2013.
- [23] J. Lee, V. S. Mirrokni, V. Nagarajan, and M. Sviridenko, "Non-monotone submodular maximization under matroid and knapsack constraints," in *Proceedings of the 41st Annual ACM Symposium on Theory of Computing*, 2009, pp. 323–332.
- [24] M. L. Fisher, G. L. Nemhauser, and L. A. Wolsey, "An analysis of approximations for maximizing submodular set functions—II," *Polyhedral combinatorics*, pp. 73–87, 1978.
- [25] G. Calinescu, C. Chekuri, M. Pál, and J. Vondrák, "Maximizing a monotone submodular function subject to a matroid constraint," *SIAM Journal on Computing*, vol. 40, no. 6, pp. 1740–1766, 2011.
- [26] M. Rubinov and O. Sporns, "Complex network measures of brain connectivity: Uses and interpretations," *Neuroimage*, vol. 52, no. 3, pp. 1059–1069, 2010.