CRITICAL SAMPLING FOR WAVELET FILTERBANKS ON ARBITRARY GRAPHS

Aamir Anis and Antonio Ortega

Department of Electrical Engineering
University of Southern California, Los Angeles
e-mail: aanis@usc.edu, ortega@sipi.usc.edu

ABSTRACT

Current formulations of critically-sampled graph wavelet filterbanks work only for bipartite graphs where downsampling signals on either partition leads to a spectrum folding phenomenon. The lack of such a natural downsampling scheme for arbitrary graphs poses difficulties in designing filterbanks. In this paper, we propose a critical sampling scheme on an arbitrary graph that chooses a sampling set for each channel, given a set of analysis/synthesis filters, by seeking to minimize a bound on the overall reconstruction error associated with the filterbank. Our algorithm is efficient since it requires a few simple graph filtering operations in each iteration. Our initial experiments show that its output is consistent with the sampling scheme for bipartite graphs and results in superior performance over other methods.

***Index Terms*—** Graph signal processing, graph wavelet filterbanks, critical sampling, graph filtering.

1. INTRODUCTION

Graph wavelet transforms have recently been used for a variety of applications, ranging from multiresolution analysis [1, 2], compression [3, 4, 5], denoising [6], and classification [7]. These transforms allow one to analyze and process signals defined over graphs while taking into consideration the underlying relationships between signal values. The designs of these transforms are generally inspired by traditional wavelet construction schemes and leverage principles from the emerging field of *Graph Signal Processing* (see [8] for a comprehensive overview).

One of the recent techniques for constructing wavelet transforms on graphs is based on filterbanks. This approach is quite appealing because it makes use of graph spectral filters [8] that have a low complexity and enable a good trade-off between vertex-domain and frequency-domain localization. Some of the desirable characteristics for these transforms are critical sampling, compact support, orthogonality and perfect reconstruction. However, state-of-the-art wavelet filterbanks that satisfy most of the mentioned properties require imposing certain structural constraints on the underlying graph. For example, the recently proposed two-channel filterbanks in [9, 10] are designed specifically for bipartite graphs. The special structure leads to a natural downsampling-upsampling scheme (on one of the two partitions) in each channel, accompanied by a spectral folding phenomenon that is exploited while designing the filters. In order to extend the design to arbitrary graphs, these works suggest using a multidimensional framework where the input graph is decomposed into multiple bipartite subgraphs over which filterbanks are designed and implemented independently. Various approaches have been proposed to optimize the bipartite subgraph decomposition [11, 12] for designing these multidimensional filterbanks. However, the limitation of this framework is that one is forced to work with simplified graphs that do not contain all the connectivity information. Additionally, there are also works that suggest expanding the input graph to create a bipartite graph thereby leading to an oversampled filterbank [13], which may not be desirable for some applications such as compression.

The focus of this work is centered on the analysis and design of filterbanks on arbitrary graphs (without altering their topology) that are critically-sampled, have compact support, and satisfy the perfect reconstruction and orthogonality conditions as closely as possible. To this end, we first present a formulation of a generic two-channel filterbank on arbitrary graphs built using polynomial graph filters and accompanied by a simple downsampling-upsampling scheme. We state and analyze the conditions on the filter responses and the sampling scheme that are required to satisfy the aforementioned properties. Unlike previous works, the lack of a special structure in the graph in general makes the problem of jointly designing low-degree filters and the sampling scheme impossible. Therefore, in this work, we decouple the two by focusing only on designing a critical sampling scheme while assuming that the analysis and synthesis filters are predefined for given frequency localization constraints. Our main contributions are: (i) a criterion based on the reconstruction error to evaluate any sampling scheme for given filters, and (ii) a greedy but computationally efficient algorithm to optimize the error criterion in order obtain an approximately optimal solution, along with some theoretical guarantees. When the filters are fixed, the problem of choosing the sampling scheme becomes akin to dictionary or subset selection problems in [14, 15]. However, these works target a specific class of signals and use an objective that is an average of the reconstruction errors for each. The metric considered in our paper is fundamentally different since it is meant for any signal over the graph and captures a bound on the reconstruction error. Further, our greedy algorithm has low complexity making it suitable for large-scale problems. The validity of our ideas is tested on various simple examples. We show that the outcome of our algorithm for a bipartite graph samples signals in each channel on either partition as expected. Moreover, the sampling sets chosen for a general graph result in superior performance in terms of reconstruction error.

2. WAVELET FILTERBANKS ON GRAPHS

2.1. Background and notation

In this paper, we work with simple, connected, undirected, and weighted graphs $G = (\mathcal{V}, \mathcal{E})$ consisting of a set of nodes $\mathcal{V} = \{1, 2, \ldots, n\}$ and edges $\mathcal{E} = \{w_{ij}\}, i, j \in \mathcal{V}$, with $w_{ii} = 0$. We

This work was supported in part by NSF grants NSF-CCF-1527874 and NSF-CCF-1410009.

denote the adjacency, degree and Laplacian matrices by $\mathbf{W}$, $\mathbf{D}$ and $\mathbf{L}$ respectively. In order to be able to design graph spectral filters, we shall work with the symmetric normalized form of the Laplacian defined as $\mathbf{L} = \mathbf{I} - \mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2}$. $\mathbf{L}$ is a symmetric positive semi-definite matrix and has a set of real eigenvalues $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n \leq 2$ and a corresponding orthogonal set of eigenvectors denoted as $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n]$. A graph signal is defined as a scalar valued discrete mapping $f : \mathcal{V} \rightarrow \mathbb{R}$ (such that $f(i)$ is the value of the signal on node i) and can also be represented as a vector $\mathbf{f}$ in $\mathbb{R}^n$, with indices corresponding to the nodes in the graph. The downsampling operation on a graph signal $\mathbf{f}$ is defined as the restriction of the signal $\mathbf{f}$ to a certain subset of nodes $\mathcal{S} \subset \mathcal{V}$ (known as the downsampling set), and the downsampled signal is a vector of reduced length $|\mathcal{S}|$. The downsampling operator for $\mathcal{S}$ is obtained by sampling the corresponding rows of the identity matrix $\mathbf{I}$, i.e., $\mathbf{S} = \mathbf{I}_{\mathcal{S},\mathcal{V}} \in \{0,1\}^{|\mathcal{S}|\times n}$. Similarly, the upsampling operation for signals downsampled on $\mathcal{S}$ inserts zeros in place of the missing signal values at appropriate locations and is given by $\mathbf{S}^T$.

The main premise behind graph signal processing is that the eigenvalues and eigenvectors of $\mathbf{L}$ provide a notion of frequency for graph signals, similar to the Fourier transform in traditional signal processing [8]. In this context, the *graph Fourier transform* (GFT) of a signal $\mathbf{f}$ is given by $\tilde{\mathbf{f}} = \mathbf{U}^T\mathbf{f}$. Further, one can design polynomial graph filters $\mathbf{H} = h(\mathbf{L}) = \sum_{i=0}^{k} h_i \mathbf{L}^i$ of different degrees, whose response in the spectral domain is given by the polynomial $h(\lambda) = \sum_{i=0}^{k} h_i \lambda^i$. A k-degree polynomial filter can be implemented in $O(k|\mathcal{E}|)$ complexity. Further, note that for undirected graphs, $\mathbf{L}$ is symmetric, and hence $\mathbf{H}$ is symmetric.

2.2. Two-channel filterbanks

We now describe the general formulation for two-channel wavelet filterbanks on general graphs. A more detailed description can be found in [9, 10] along with the analysis for bipartite graphs. We make certain changes to notation for compactness.

A generic two-channel wavelet filterbank on a graph decomposes any graph signal $\mathbf{x} \in \mathbb{R}^n$ into a lowpass (smooth) and highpass (detail) component (Figure 1). It consists of an analysis filterbank with $\mathbf{H}_0$ and $\mathbf{H}_1$ as lowpass and highpass filters, and a synthesis filterbank with $\mathbf{G}_0$ and $\mathbf{G}_1$ as the lowpass and highpass filters. $\mathbf{S}_0 \in \{0,1\}^{|\mathcal{S}_0|\times n}$ and $\mathbf{S}_1 \in \{0,1\}^{|\mathcal{S}_1|\times n}$ are the downsampling operators for the lowpass and highpass branch respectively, while $\mathbf{S}_0^T$ and $\mathbf{S}_1^T$ are the corresponding upsampling operators. The outputs of the two branches after the analysis filterbank are $\mathbf{y}_0 \in \mathbb{R}^{|\mathcal{S}_0|}$ and $\mathbf{y}_1 \in \mathbb{R}^{|\mathcal{S}_1|}$. These are given as

$$\begin{bmatrix}\mathbf{y}_0\\ \mathbf{y}_1\end{bmatrix} = \begin{bmatrix}\mathbf{S}_0\mathbf{H}_0\\ \mathbf{S}_1\mathbf{H}_1\end{bmatrix}\mathbf{x} = \mathbf{T}_a\mathbf{x}. \tag{1}$$

Similarly, the output of the synthesis filterbank (i.e., the reconstructed signal) is denoted as $\hat{\mathbf{x}} \in \mathbb{R}^n$ and is given by

$$\hat{\mathbf{x}} = \begin{bmatrix}\mathbf{G}_0\mathbf{S}_0^T & \mathbf{G}_1\mathbf{S}_1^T\end{bmatrix}\begin{bmatrix}\mathbf{y}_0\\ \mathbf{y}_1\end{bmatrix} = \mathbf{T}_s\begin{bmatrix}\mathbf{y}_0\\ \mathbf{y}_1\end{bmatrix}, \tag{2}$$

with the complete transfer equation for the system given by

$$\hat{\mathbf{x}} = \left(\mathbf{G}_0\mathbf{S}_0^T\mathbf{S}_0\mathbf{H}_0 + \mathbf{G}_1\mathbf{S}_1^T\mathbf{S}_1\mathbf{H}_1\right)\mathbf{x}. \tag{3}$$

We now state some desirable characteristics of the filterbank along with the conditions needed to satisfy each.

Compact support requires that the filters $\{\mathbf{H}_i, \mathbf{G}_i\}_{i=0,1}$ be expressible as finite polynomials of the graph Laplacian, a notion analogous to FIR filters in classical DSP. A k-degree polynomial filter requires collecting information from a k-degree neighborhood for each node.

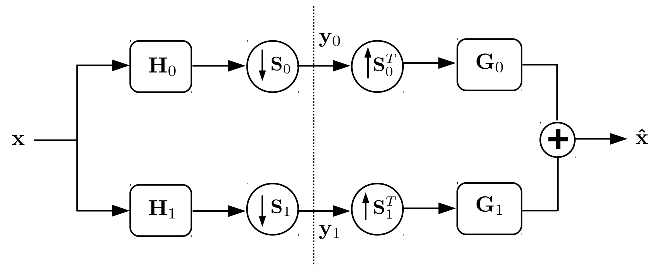

Fig. 1: A generic two-channel filterbank on graphs

Critical sampling requires that the total number of samples after downsampling in both branches should be equal to the dimension of the signal, i.e., $|\mathcal{S}_0| + |\mathcal{S}_1| = n$.

Perfect reconstruction requires that the transfer function of the entire system be identity, i.e.,

$$\mathbf{G}_0\mathbf{S}_0^T\mathbf{S}_0\mathbf{H}_0 + \mathbf{G}_1\mathbf{S}_1^T\mathbf{S}_1\mathbf{H}_1 = \mathbf{I}. \tag{4}$$

Orthogonality requires the filterbanks to satisfy $\mathbf{T}_s = \mathbf{T}_a^T$ and $\mathbf{T}_a^T\mathbf{T}_a = \mathbf{I}$, which translates to substituting $\mathbf{G}_0 = \mathbf{H}_0$ and $\mathbf{G}_1 = \mathbf{H}_1$ in (4).

Note that the perfect reconstruction condition in (4) can also be interpreted using the eigendecomposition of $\mathbf{L} = \mathbf{U}\Lambda\mathbf{U}^T$ as

$$g_0(\Lambda)\mathbf{U}^T\mathbf{S}_0^T\mathbf{S}_0\mathbf{U}h_0(\Lambda) + g_1(\Lambda)\mathbf{U}^T\mathbf{S}_1^T\mathbf{S}_1\mathbf{U}h_1(\Lambda) = \mathbf{I}, \tag{5}$$

For an arbitrary $\mathbf{U}$, it is impossible to satisfy (5) using low-degree polynomial filters, since the number of constraints ($= n^2$) is much larger than the available degrees of freedom. Therefore, one would like to design the system such that $\mathbf{G}_0\mathbf{S}_0^T\mathbf{S}_0\mathbf{H}_0 + \mathbf{G}_1\mathbf{S}_1^T\mathbf{S}_1\mathbf{H}_1$ is as close as possible to identity. Special structure in the graph results in a structured $\mathbf{U}$ and therefore simplification (5) by elimination of several constraints. For example, in bipartite graphs with $\mathcal{S}_0, \mathcal{S}_1$ denoting opposite partitions, it can be shown that $\mathbf{U}^T\mathbf{S}_0^T\mathbf{S}_0\mathbf{U} = \frac{1}{2}(\mathbf{I}+\mathbf{I}^*)$ and $\mathbf{U}^T\mathbf{S}_1^T\mathbf{S}_1\mathbf{U} = \frac{1}{2}(\mathbf{I}-\mathbf{I}^*)$, where $\mathbf{I}^*$ is the identity matrix mirrored about one of the vertical sides (i.e., with ones running from top-right to bottom-left). This observation leads exactly to the perfect reconstruction conditions stated in [9, 10] for bipartite graphs. $\mathbf{I}^*$ causes the spectral folding phenomenon for bipartite graphs and thus generates n additional aliasing constraints besides the n diagonal constraints, resulting in a total of $2n$ constraints that are easier to satisfy with low-degree filters. Note that if we are not restricted to using polynomial filters in the synthesis flterbank, least-squares inversion can be used for inverting the analysis transfer function $\mathbf{T}_a$, provided it is non-singular.

3. CRITICAL SAMPLING FOR FILTERBANKS

3.1. Approximately optimal sampling scheme

For a critically sampled design, we must choose $\mathcal{S}_0$ and $\mathcal{S}_1$ such that $|\mathcal{S}_0|+|\mathcal{S}_1| = n$ and the filterbank is as close to perfect reconstruction as possible. One way to achieve this is by minimizing the deviation of the overall transfer function of the system from identity in terms of Frobenius form, i.e., $\|\mathbf{G}_0\mathbf{S}_0^T\mathbf{S}_0\mathbf{H}_0 + \mathbf{G}_1\mathbf{S}_1^T\mathbf{S}_1\mathbf{H}_1 - \mathbf{I}\|_F^2$, which is in fact an upper bound on the squared relative error for all signals on the graph. Further, in our design, we assume that we have already designed filters $\mathbf{H}_0, \mathbf{H}_1, \mathbf{G}_0, \mathbf{G}_1$ to satisfy $\mathbf{G}_0\mathbf{H}_0 + \mathbf{G}_1\mathbf{H}_1 = 2\mathbf{I}$

Algorithm 1 Basic greedy minimization

Initialize: $\mathcal{S} = \{\emptyset\}$.
while $|\mathcal{S}| < n$ **do**
  $\mathcal{S} \leftarrow \mathcal{S} \cup \{u\}$, where $u = \text{argmin}_{v \in \mathcal{S}^c} \ \phi(\mathcal{S} \cup \{v\})$.
end while

(for example, using the methods of [9, 10]). In order to minimize the reconstruction error over the choice of sampling sets $\mathcal{S}_0$ and $\mathcal{S}_1$, we first introduce a concatenated setting (of $2n$ dimensions) by defining

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_0 \\ \mathbf{H}_1 \end{bmatrix} \in \mathbb{R}^{2n \times n}, \ \mathbf{G} = \begin{bmatrix} \mathbf{G}_0 \\ \mathbf{G}_1 \end{bmatrix} \in \mathbb{R}^{2n \times n},$$
$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \end{bmatrix} \in \mathbb{R}^{|\mathcal{S}_0| + |\mathcal{S}_1|}, \ \mathbf{S} = \begin{bmatrix} \mathbf{S}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_1 \end{bmatrix} \in \{0,1\}^{(|\mathcal{S}_0| + |\mathcal{S}_1|) \times 2n}.$$

Note that the concatenated downsampling operator $\mathbf{S}$ can be obtained by sampling rows of the $2n$-dimensional identity corresponding to indices in a concatenated sampling set $\mathcal{S} \subset \{1, \ldots, 2n\}$ that contains sampled nodes for both the channels such that $|\mathcal{S}| = |\mathcal{S}_0| + |\mathcal{S}_1|$. Further, $\mathcal{S}^c = \{1, \ldots, 2n\} \setminus \mathcal{S}$ and $\mathcal{S}_0$ and $\mathcal{S}_1$ can be recovered from $\mathcal{S}$ as $\mathcal{S}_0 = \{v | v \in \mathcal{S}, 1 \le v \le n\}$ and $\mathcal{S}_1 = \{v - n | v \in \mathcal{S}, n+1 \le v \le 2n\}$. With these definitions, the transfer function of the system can be written as $\mathbf{G}^T \mathbf{S}^T \mathbf{S} \mathbf{H}$ and finding a critical sampling scheme requires solving

$$\min_{\mathcal{S}: |\mathcal{S}| = n} \left\| \mathbf{G}^T \mathbf{S}^T \mathbf{S} \mathbf{H} - \mathbf{I} \right\|_F^2. \tag{6}$$

Since we choose the filters such that $\mathbf{G}^T \mathbf{H} = 2\mathbf{I}$, we can rewrite the objective as

$$\begin{aligned} \phi(\mathcal{S}) &= \left\| \mathbf{G}^T \mathbf{S}^T \mathbf{S} \mathbf{H} - \frac{1}{2} \mathbf{G}^T \mathbf{H} \right\|_F^2 \\ &= \left\| \frac{1}{2} \sum_{i \in \mathcal{S}} \mathbf{g}_i \mathbf{h}_i^T - \frac{1}{2} \sum_{j \in \mathcal{S}^c} \mathbf{g}_j \mathbf{h}_j^T \right\|_F^2, \end{aligned} \tag{7}$$

where $\mathbf{g}_i$ and $\mathbf{h}_i$ denote the i^{th} columns of $\mathbf{G}^T$ and $\mathbf{H}^T$ respectively. In order to minimize $\phi(\mathcal{S})$, we propose to use a simple greedy procedure (Algorithm 1) that begins with an empty $\mathcal{S}$ and keeps adding nodes one-by-one that minimize $\phi(\mathcal{S})$ at each step. This algorithm requires $O(n^2)$ evaluations of the objective $\phi(\mathcal{S})$ which can be quite expensive. Explicitly storing the matrices $\mathbf{G}$ and $\mathbf{H}$ requires $O(n^2)$ space. We now show how one can efficiently implement the algorithm in $O(n)$ graph filtering operations and $O(n)$ space. Using (7), the change in the objective $\phi(\mathcal{S})$ when a node $v \subset \{1, \ldots, 2n\}$ is added to $\mathcal{S}$ is given by:

$$\begin{aligned} \phi(\mathcal{S} \cup \{v\}) &= \left\| \left(\frac{1}{2} \sum_{i \in \mathcal{S}} \mathbf{g}_i \mathbf{h}_i^T - \frac{1}{2} \sum_{j \in \mathcal{S}^c} \mathbf{g}_j \mathbf{h}_j^T \right) + \mathbf{g}_v \mathbf{h}_v^T \right\|_F^2 \\ &= \phi(\mathcal{S}) + p_v(\mathcal{S}) + q_v, \end{aligned} \tag{8}$$

where we defined

$$p_v(\mathcal{S}) = \text{Tr} \left[\mathbf{h}_v \mathbf{g}_v^T \left(\sum_{i \in \mathcal{S}} \mathbf{g}_i \mathbf{h}_i^T - \sum_{j \in \mathcal{S}^c} \mathbf{g}_j \mathbf{h}_j^T \right) \right], \tag{9}$$

$$q_v = \|\mathbf{g}_v\|^2 \|\mathbf{h}_v\|^2. \tag{10}$$

Thus, we have

$$\underset{v \in \mathcal{S}^c}{\text{argmin}} \ \phi(\mathcal{S} \cup \{v\}) = \underset{v \in \mathcal{S}^c}{\text{argmin}} \ (p_v(\mathcal{S}) + q_v). \tag{11}$$

Algorithm 2 Efficient algorithm for critical sampling

Input: Graph $G = \{\mathcal{V}, E\}$, concatenated filters $\mathbf{H}$, $\mathbf{G}$.
Initialize: $\mathcal{S} = \{\emptyset\}$, $\mathbf{p}, \mathbf{q} \in \mathbb{R}^{2n}$ such that $p_v = -2 \langle \mathbf{g}_v, \mathbf{h}_v \rangle$, $q_v = \|\mathbf{g}_v\|^2 \|\mathbf{h}_v\|^2$.
while $|\mathcal{S}| < n$ **do**
  $\mathcal{S} \leftarrow \mathcal{S} \cup \{u\}$, where $u = \text{argmin}_{v \in \mathcal{S}^c} \ (p_v + q_v)$.
  $\mathbf{p} \leftarrow \mathbf{p} + 2(\mathbf{G}\mathbf{g}_u) \circ (\mathbf{H}\mathbf{h}_u)$.
end while

In order to compute $p_v(\mathcal{S})$ for each $\mathcal{S}$, we first note that

$$p_v(\emptyset) = \text{Tr} \left[\mathbf{h}_v \mathbf{g}_v^T (-\mathbf{G}^T \mathbf{H}) \right] = -2 \langle \mathbf{g}_v, \mathbf{h}_v \rangle. \tag{12}$$

Further, for a node u, $p_v(\mathcal{S} \cup \{u\})$ can be computed as

$$\begin{aligned} p_v(\mathcal{S} \cup \{u\}) &= \text{Tr} \left[\mathbf{h}_v \mathbf{g}_v^T \left(\sum_{i \in \mathcal{S}} \mathbf{g}_i \mathbf{h}_i^T - \sum_{j \in \mathcal{S}^c} \mathbf{g}_j \mathbf{h}_j^T + 2 \mathbf{g}_u \mathbf{h}_u^T \right) \right] \\ &= p_v(\mathcal{S}) + 2 \langle \mathbf{g}_v, \mathbf{g}_u \rangle \langle \mathbf{h}_v, \mathbf{h}_u \rangle. \end{aligned} \tag{13}$$

To make the notation compact, we introduce the vectors $\mathbf{p}(\mathcal{S}), \mathbf{q} \in \mathbb{R}^{2n}$, whose v^{th} elements are $p_v(\mathcal{S})$ and q_v. Therefore, using "$\circ$" to denote element-wise vector product (Hadamard product), we have

$$\mathbf{p}(\mathcal{S} \cup \{u\}) = \mathbf{p}(\mathcal{S}) + 2(\mathbf{G}\mathbf{g}_u) \circ (\mathbf{H}\mathbf{h}_u). \tag{14}$$

We summarize the efficient method for choosing $\mathcal{S}$ in Algorithm 2. Note that the vectors $\mathbf{h}_v$ and $\mathbf{g}_v$ can be computed using two filtering operations each as $\mathbf{H}^T \delta_v$ and $\mathbf{G}^T \delta_v$ respectively, where δ_v is the graph delta signal on node v. Therefore, in terms of time complexity, computing $\mathbf{p}(\emptyset)$ and $\mathbf{q}$ require $4n$ one-time graph filtering operations in total. Further, each greedy iteration requires performing 8 filtering operations. Therefore, Algorithm 2 requires $O(n)$ graph filtering operations. The space complexity of the algorithm is $O(n)$ since it is matrix-free, i.e., $\mathbf{L}$ is the only matrix that needs to be stored.

3.2. Theoretical guarantees

We now show that it is possible to obtain some theoretical insight into the performance of (a randomized variant of) our greedy algorithm when $\mathbf{G} = \mathbf{H}$. Note that for $\mathcal{S}_1 \subseteq \mathcal{S}_2$ and a $v \notin \mathcal{S}_1, \mathcal{S}_2$,

$$\begin{aligned} p_v(\mathcal{S}_1) &= \sum_{i \in \mathcal{S}_1} \langle \mathbf{h}_v, \mathbf{h}_i \rangle^2 - \sum_{j \in \mathcal{S}_1^c} \langle \mathbf{h}_v, \mathbf{h}_j \rangle^2 \\ &\le \sum_{i \in \mathcal{S}_2} \langle \mathbf{h}_v, \mathbf{h}_i \rangle^2 - \sum_{j \in \mathcal{S}_2^c} \langle \mathbf{h}_v, \mathbf{h}_j \rangle^2 = p_v(\mathcal{S}_2) \end{aligned} \tag{15}$$

Using this in (8), we obtain

$$\phi(\mathcal{S}_1 \cup \{v\}) - \phi(\mathcal{S}_1) \le \phi(\mathcal{S}_2 \cup \{v\}) - \phi(\mathcal{S}_2), \tag{16}$$

which implies $\phi(\mathcal{S})$ is supermodular in $\mathcal{S}$. Therefore, the function $\psi(\mathcal{S}) = \phi(\emptyset) - \phi(\mathcal{S})$ is submodular, non-monotone and normalized ($\psi(\emptyset) = 0$). As a result, the set $\mathcal{S}^*$ obtained by the greedy maximization of $\psi(\mathcal{S})$ (or equivalently greedy minimization of $\phi(\mathcal{S})$) with a randomized version of Algorithm 1, that selects one node uniformly at random from the best n nodes at each iteration, is at least a 0.3-approximation of the optimal set $\mathcal{S}_{\text{OPT}}$ [16]. To be precise, we have the following guarantees for $\mathcal{S}^*$ obtained from the randomized greedy algorithm

$$\psi(\mathcal{S}_{\text{OPT}}) \ge \psi(\mathcal{S}^*) \ge 0.3 \psi(\mathcal{S}_{\text{OPT}}) \tag{17}$$
$$\Rightarrow \phi(\mathcal{S}_{\text{OPT}}) \le \phi(\mathcal{S}^*) \le 0.3 \phi(\mathcal{S}_{\text{OPT}}) + 0.7n \tag{18}$$

Although, guarantees for the deterministic version of the greedy algorithm are part of ongoing research, we observe empirically that its performance is competitive. Note that for the biorthogonal design when $\mathbf{G} \neq \mathbf{H}$, $\phi(\mathcal{S})$ is no longer supermodular, hence we cannot state guarantees on the performance of the greedy algorithm in this case. However, experiments show that the algorithm performs well in this case as well.

3.3. Multi-channel extension

In order to extend our formulation to m-channel filterbanks with analysis/synthesis filter pairs $\{\mathbf{H}_k, \mathbf{G}_k\}_{k=0,\ldots,m-1}$ and sampling sets $\{\mathcal{S}_k\}_{k=0,\ldots,m-1}$ ($\mathcal{S}_k \subset \{1,\ldots,n\}$), one can create the concatenated filters $\mathbf{H}, \mathbf{G} \in \mathbb{R}^{(\sum_{k=0}^{m-1} |\mathcal{S}_i|) \times n}$ and the concatenated sampling set $\mathcal{S} \subset \{1,\ldots,mn\}$ in a manner similar to that of the two-channel case. Note that each $\mathcal{S}_k$ can be then be recovered from $\mathcal{S}$ as $\mathcal{S}_k = \{v - kn | v \in \mathcal{S}, kn+1 \leq v \leq kn+n\}$. Further, in this case, we require predesigned filters such that $\mathbf{G}^T\mathbf{H} = m\mathbf{I}$, resulting in the objective

$$\phi(\mathcal{S}) = \left\| \left(1 - \frac{1}{m}\right) \sum_{i \in \mathcal{S}} \mathbf{g}_i \mathbf{h}_i^T - \frac{1}{m} \sum_{j \in \mathcal{S}^c} \mathbf{g}_j \mathbf{h}_j^T \right\|_F^2, \qquad (19)$$

that can be optimized under the constraint $|\mathcal{S}| = n$ using the same technique as that of the two-channel case. Moreover, it admits the same theoretical guarantees.

4. EXPERIMENTS

In this section, we present simple experiments to demonstrate the effectiveness of our critical sampling scheme for two-channel filterbanks. In our first experiment, we test its performance on two simple bipartite graphs (Figure 2) with filters $\mathbf{H}_0, \mathbf{H}_1$ obtained using the graph-QMF design of [9], that approximates the Meyer kernel with a polynomial filter of chosen length 8. We observe that the output of Algorithm 2 matches exactly with that of the optimal sampling scheme for the bipartite graph, which is to downsample the filtered signal in each channel on either partition.

For our second experiment, we design a critically-sampled two-channel filterbank on the Minnesota road network graph using two configurations of analysis/synthesis filters: (i) Graph-QMF design [9] with 8-degree polynomial approximations of the Meyer kernel, and (ii) *graphBior*(6,6) [10]. The sampling scheme obtained for each of these configurations is plotted in Figures 3a and 3b. We observe that the sampling pattern for each channel colors nodes in a predominantly alternating fashion indicating a propensity towards bipartition. The response of the filterbank after determining the sampling set is plotted in Figures 3c and 3d for unit magnitude delta functions in the spectral domain. We observe that it is close to 1 for all frequencies. Since the transfer function is not diagonalizable in the GFT basis $\mathbf{U}$, there is an associated aliasing effect with the filterbank. We characterize this by plotting the maximum aliasing coefficient in terms of magnitude for each frequency component in Figures 3e and 3f. Finally, we also compare the recontruction performance (in terms ratio of energies of error signal and original) of our proposed method against a random sampling scheme, and a spectral approximation of MaxCut for 1000 random signals. The average squared relative errors along with the standard deviations are listed in Table 1, we observe that our method has superior performance.

Fig. 2: Sampling scheme obtained using Algorithm 2 for bipartite graphs with graph-QMF design filters. Red and blue indicate nodes in low-pass and high-pass channels.

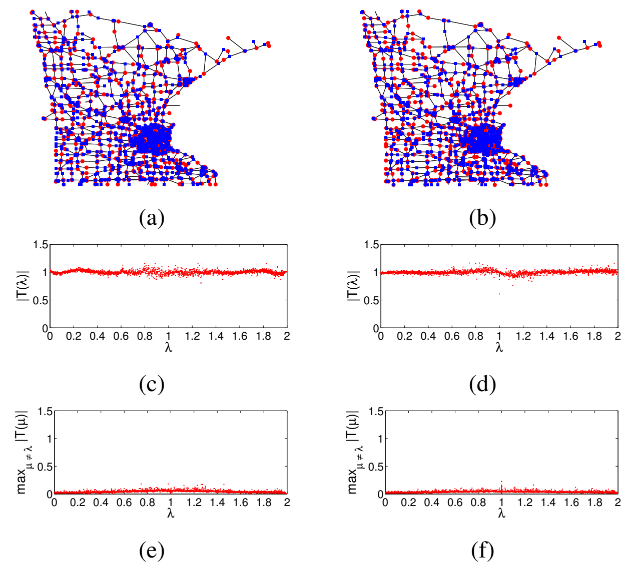

Fig. 3: Performance of critical sampling scheme (Algorithm 2) on Minnesota road network graph. (a), (c) and (e) denote sampling scheme obtained, spectral response, and maximum aliasing component for graph-QMF design. (b), (d) and (f) illustrate corresponding results for graphBior(6,6).

Table 1: Recontruction error results for random signals on the Minnesota road network graph.

	graph-QMF (poly 8)	graphBior(6,6)
Random	0.4842 ± 0.0113	0.4629 ± 0.0108
MaxCut	0.1125 ± 0.0069	0.0972 ± 0.0061
Proposed	$\mathbf{0.0779 \pm 0.0049}$	$\mathbf{0.0664 \pm 0.0045}$

5. CONCLUSIONS AND FUTURE WORK

We presented a technique for designing critically-sampled near-perfect reconstruction wavelet filterbanks on arbitrary graphs. In the absence of a special structure in the graph, it is not possible in general to attain perfect reconstruction with low-degree polynomial graph filters with any critical sampling scheme. Therefore, we decouple the two design problems in this work and focus only on choosing the best possible sampling scheme. Specifically, given a predesigned set of analysis/synthesis filters, our algorithm efficiently approximates the best sampling set for each channel in order to minimize a bound on the overall reconstruction error associated with the filterbank. As a future extension, we would like to jointly optimize the filter responses and the critical sampling scheme in order to further reduce the reconstruction error.

6. REFERENCES

[1] S. K. Narang, Y. H. Chao, and A. Ortega, "Graph-wavelet filterbanks for edge-aware image processing," in *2012 IEEE Statistical Signal Processing Workshop (SSP)*, Aug 2012, pp. 141–144.

[2] A. Sakiyama and Y. Tanaka, "Edge-aware image graph expansion methods for oversampled graph laplacian matrix," in *2014 IEEE International Conference on Image Processing (ICIP)*, Oct 2014, pp. 2958–2962.

[3] S. K. Narang, Y. H. Chao, and A. Ortega, "Critically sampled graph-based wavelet transforms for image coding," in *Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2013 Asia-Pacific*, Oct 2013, pp. 1–4.

[4] H. Q. Nguyen, P. A. Chou, and Y. Chen, "Compression of human body sequences using graph wavelet filter banks," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014, pp. 6152–6156.

[5] Y. H. Chao, A. Ortega, and S. Yea, "Graph-based lifting transform for intra-predicted video coding," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2016, pp. 1140–1144.

[6] S. Deutsch, A. Ortega, and G. Medioni, "Manifold denoising based on spectral graph wavelets," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2016, pp. 4673–4677.

[7] V. N. Ekambaram, G. Fanti, B. Ayazifar, and K. Ramchandran, "Wavelet-regularized graph semi-supervised learning," in *Global Conference on Signal and Information Processing (GlobalSIP), 2013 IEEE*, Dec 2013, pp. 423–426.

[8] 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, May 2013.

[9] S. K. Narang and A. Ortega, "Perfect reconstruction two-channel wavelet filter banks for graph structured data," *IEEE Transactions on Signal Processing*, vol. 60, no. 6, pp. 2786–2799, June 2012.

[10] S. K. Narang and A. Ortega, "Compact support biorthogonal wavelet filterbanks for arbitrary undirected graphs," *IEEE Transactions on Signal Processing*, vol. 61, no. 19, pp. 4673–4685, Oct 2013.

[11] H. Q. Nguyen and M. N. Do, "Downsampling of signals on graphs via maximum spanning trees," *IEEE Transactions on Signal Processing*, vol. 63, no. 1, pp. 182–191, Jan 2015.

[12] J. Zeng, G. Cheung, and A. Ortega, "Bipartite subgraph decomposition for critically sampled wavelet filterbanks on arbitrary graphs," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2016, pp. 6210–6214.

[13] A. Sakiyama and Y. Tanaka, "Oversampled graph laplacian matrix for graph filter banks," *IEEE Transactions on Signal Processing*, vol. 62, no. 24, pp. 6425–6437, Dec 2014.

[14] Andreas Krause and Volkan Cevher, "Submodular dictionary selection for sparse representation," in *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, Haifa, Israel, June 2010, pp. 567–574.

[15] Abhimanyu Das and David Kempe, "Submodular meets spectral: Greedy algorithms for subset selection, sparse approximation and dictionary selection," in *Proceedings of the 28th International Conference on Machine Learning (ICML-11)*, New York, NY, USA, June 2011, pp. 1057–1064.

[16] Niv Buchbinder, Moran Feldman, Joseph (Seffi) Naor, and Roy Schwartz, "Submodular maximization with cardinality constraints," in *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms*, 2014, pp. 1433–1452.