Extending Classical Multirate Signal Processing Theory to Graphs – Part I: Fundamentals

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

***Abstract*—Signal processing on graphs finds applications in many areas. In recent years renewed interest on this topic was kindled by two groups of researchers. Narang and Ortega constructed two-channel filter banks on bipartitie graphs described by Laplacians. Sandryhaila and Moura developed the theory of linear systems, filtering, and frequency responses for the case of graphs with arbitrary adjacency matrices, and showed applications in signal compression, prediction, etc. Inspired by these contributions, this paper extends classical multirate signal processing ideas to graphs. The graphs are assumed to be general with a possibly non-symmetric and complex adjacency matrix. The paper revisits ideas such as noble identities, aliasing, and polyphase decompositions in graph multirate systems. Drawing such a parallel to classical systems allows one to design filter banks with polynomial filters, with lower complexity than arbitrary graph filters. It is shown that the extension of classical multirate theory to graphs is nontrivial, and requires certain mathematical restrictions on the graph. Thus, classical noble identities cannot be taken for granted. Similarly, one cannot claim that the so-called delay chain system is a perfect reconstruction system (as in classical filter banks). It will also be shown that M-partite extensions of the bipartite filter bank results will not work for M-channel filter banks, but a more restrictive condition called M-block cyclic property should be imposed. Such graphs are studied in detail. A detailed theory for M-channel filter banks is developed in a companion paper.**

***Index Terms*—Multirate processing, graph signals, aliasing on graphs, bandlimited graph signals, block-cyclic graphs.**

I. Introduction

The processing of signals defined on graphs has been of interest for many years, and finds applications in a diverse set of fields such as sensor networks [1], social and economic networks [2], biological networks [3] and others [4]. A detailed introduction can be found in [5], and in the tutorial articles [6], [7]. In graph signal processing applications, signals are not defined as functions on a uniform time-domain grid but they are defined as vectors indexed by the vertices of a graph – possibly directed. In recent years renewed interest on this topic was kindled by two groups of researchers. The first set of papers, pioneered by Narang and Ortega [7]–[12] showed how two-channel filter banks can be constructed on graphs, and went on to develop elegant techniques for the design of down-sampled, two-channel perfect reconstruction filter banks on bipartitie graphs. These results were developed for graphs that have a real, symmetric adjacency matrix, and all results were based on the graph Laplacian.

An independent development of graph signal processing was advanced by Sandryhaila and Moura [5], [6], [13], wherein the graph adjacency matrix $\boldsymbol{A}$ was allowed to be arbitrary – possibly non-symmetric (and complex) that allows for directed graphs in the development. By proposing that the adjacency matrix can be regarded as a graph-shift operator, a beautiful extension of the basic concepts of linear shift invariant systems on graphs was developed in [5], giving rise to insightful notions such as filtering and frequency responses on graphs. Many applications of such elegant theory were delineated, such as linear prediction, data compression, and classification. Further extensions of these results were also developed in [14]–[17].

Multirate analysis for graph signals has been of interest in recent years. Studies in [18]–[20] mainly focus on circulant graphs and analyze two-channel decomposition of graph signals. Multirate decomposition can be achieved by iterative application of 2-channel systems. The study in [21] proposes to combine decimators and filters for construction of a filter bank on a graph. Motivated by Haar filter banks in the classical theory, a graph filter bank is developed using the partitions of the graph.

Inspired by the pioneering contributions of [5] and [8], this paper and the companion work [22] extend many of the basic concepts of classical multirate signal processing and filter bank theory to graphs. In this paper (Part I) we develop the equivalent of fundamental ideas such as noble identities, aliasing, and polyphase decompositions in graph multirate systems. A detailed general theory for M-channel filter banks is then developed in the companion paper [22]. The graphs are assumed to be very general as in [5], with a possibly non-symmetric and complex adjacency matrix.

In the context of graph signal processing a linear filter is just a square matrix. By a cascade of such matrices one can trivially construct a graph filter bank. Problems with this approach and reasons why we focus on polynomial filters are detailed in in Sec. III. We will see in these papers that the extension of classical multirate signal processing theory to graphs is nontrivial, and requires certain mathematical restrictions on the graph adjacency matrix $\boldsymbol{A}$. While some of the results of classical filter bank theory extend easily, some of the deeper results unfold a lot of surprises – some extend and some do not extend to the case of graphs. For example, the classical noble identities [23] cannot be taken for granted, and require some restrictions on the graph matrix $\boldsymbol{A}$. Similarly,

The authors are with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA, 91125, USA. Email: oteke@caltech.edu, ppvnath@systems.caltech.edu.

This work was supported in parts by the ONR grants N00014-11-1-0676 and N00014-15-1-2118, and the California Institute of Technology.

one cannot take it for granted that the delay chain system [23] is a perfect reconstruction filter bank (an easily proved result in the case of classical filter banks). It will also be shown that M-partite extensions of the bipartite results in [8] will not in general work for M channel filter banks, but a more restrictive condition called M-block cyclic property should be imposed on the graph. While a number of results in this and the companion paper require this property, there are ways to relax it as explained in Sec. VII of the companion paper [22], and also in specific theorem statements. A detailed outline of this paper is given below, and an outline of the companion paper can be found in Sec. I of [22].

While dealing with graphs, we often make comparisons with conventional multirate systems and filter banks defined in the time domain [23]–[28]. On rare occasions we also make comparisons with systems defined in the cyclic (periodic) time domain that is equivalent to a graph with a specific cyclic adjacency matrix (Eq. (12) in [13]). These systems defined in the time-domain (or cyclical time domain on occasions) will be referred to as "classical" systems, "classical" filter banks, and so forth.

A. Scope and outline

After introducing the canonical downsampling and upsampling operators on graphs, we begin with a study of noble identities. These identities are known to be important in theoretical developments and practical implementations of classical multirate systems [23], [27]. For the case of graphs we will show in Sec. II-B that the noble identities make sense only for graphs with a certain specific structure on the adjacency matrix (Theorems 1 and 2). We then show in Sec. II-C that the delay chain filter bank (or the lazy filter bank) does not in general have perfect reconstruction property for arbitrary graphs. We introduce Type-1 and Type-2 polyphase representation of polynomial filters in Sec. II-D. Section III discusses how one can trivially construct a graph filter bank, and motivates the use of polynomial filters. In order to extend the results for bipartite graphs on 2-channel systems to M-channels, one may propose to use M-partite graphs rather than bipartite graphs. In Sec. IV we briefly discuss that such a generalization is not useful. Section V introduces M-block cyclic graphs that are important for many of the later developments in this and the companion paper [22]. The eigenstructure of M-block cyclic graphs, which forms the foundation for many of these results, is developed in Sec. VI (Theorem 5). Many of the results developed in this and the companion paper [22] are therefore valid only for graphs that satisfy either the M-block cyclic property or the eigenvector structure of M-block cyclic graphs. In Sec. VII of [22] we also discuss how this restriction can be removed, and what the price paid is.

The concepts of spectrum folding and aliasing are developed in Sec. VII for graphs that have an eigenvector structure similar to those of M-block cyclic graphs. These will be used later to develop perfect reconstruction filter banks in [22].

Section VIII embarks on a study of three related properties of linear systems on graphs: namely the polynomial property, the shift invariance property, and the so-called alias-free property. While these properties are identical in classical signal processing theory, such is not the case on graphs. Some of these interrelations were developed in [5], but Sec. VIII goes deeper and establishes the complete picture. This will be useful for obtaining a deeper understanding of alias-free maximally decimated M-channel graph filter banks in [22]. Preliminary conference versions have appeared in [29], [30].

B. Notation

The set of column vectors of size N with complex valued elements is denoted by $\mathcal{C}^N$. The set of $N \times M$ matrices with complex valued elements is denoted by $\mathcal{C}^{N\times M}$. The set of square matrices with size N is denoted by $\mathcal{M}^N$. For a matrix $\boldsymbol{A}$ the conjugate transpose is given by $\boldsymbol{A}^*$, and the transpose is given by $\boldsymbol{A}^T$. The column vector of size N with all 1 entries is denoted by $\mathbb{1}_N$. For the standard basis, k^{th} vector is denoted by $\boldsymbol{e}_k$, that is, $\boldsymbol{e}_k$ has zero elements except for the k^{th} index where it has 1. Identity and zero matrix of size $N \times N$ are denoted by $\boldsymbol{I}_N$ and $\boldsymbol{0}_N$, respectively.

We will use $\otimes$ to denote the Kronecker product with the following definition

$$\boldsymbol{A}\otimes\boldsymbol{B} = \begin{bmatrix} a_{1,1}\boldsymbol{B} & \cdots & a_{1,M}\boldsymbol{B} \\ \vdots & \ddots & \vdots \\ a_{N,1}\boldsymbol{B} & \cdots & a_{N,M}\boldsymbol{B} \end{bmatrix} \in \mathcal{C}^{(NP)\times(MQ)}, \quad (1)$$

where $\boldsymbol{A} \in \mathcal{C}^{N\times M}$ and $\boldsymbol{B} \in \mathcal{C}^{P\times Q}$.

Given a graph, $\boldsymbol{A}$ represents the adjacency matrix of the graph. We often refer to a graph with adjacency matrix $\boldsymbol{A}$ as "graph $\boldsymbol{A}$" for convenience. Throughout the paper, N denotes the size of the graph and length of the signal and M denotes the decimation ratio or the number of filters in a graph filter bank, according to context. Hence, $\boldsymbol{A} \in \mathcal{M}^N$. The $(i,j)^{th}$ block of the adjacency matrix $\boldsymbol{A}$ is denoted by $(\boldsymbol{A})_{i,j}$ and $(\boldsymbol{v})_i$ denotes the $(i)^{th}$ block of the vector $\boldsymbol{v}$. Throughout the paper, when it is not indicated, it should be clear that $(\boldsymbol{A})_{i,j} \in \mathcal{M}^{N/M}$ and $(\boldsymbol{v})_i \in \mathcal{C}^{N/M}$. Otherwise, they are clearly indicated to have the specified sizes. For a vector $\boldsymbol{x} \in \mathcal{C}^N$, $diag(\boldsymbol{x}) \in \mathcal{M}^N$ is a diagonal matrix with elements of $\boldsymbol{x}$ being on the diagonal. The cyclic permutation matrix of size N is denoted by $\boldsymbol{C}_N$, and it is defined as:

$$\boldsymbol{C}_N = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & 0 \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix} \in \mathcal{M}^N. \quad (2)$$

C. Review of DSP on Graphs

We will follow the construction presented in [5], [13]. Let $\boldsymbol{x} \in \mathcal{C}^N$ be a signal on a graph with adjacency matrix $\boldsymbol{A}$. We will assume that the graph is known a priori. The i^{th} vertex in this graph is supposed to produce the i^{th} element of the signal $\boldsymbol{x}$. The $(i,j)^{th}$ element of the adjacency matrix, $a_{i,j}$, denotes the weight of the edge from the j^{th} vertex to the i^{th} vertex. When $a_{i,j} = 0$, it means that there is no edge. We consider the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2016.2617833, IEEE Transactions on Signal Processing

general case and allow $a_{i,j}$ to be different from $a_{j,i}$. When this happens, the graph is called *directed*, or *digraph*. We do *not* assume anything on $a_{i,j}$, and they are allowed to be complex.

In DSP on graphs, the adjacency matrix of the graph of interest is considered to be *the unit shift* operator for a signal on the graph [5]. Namely, let $\boldsymbol{x}$ be a signal on a graph with the adjacency matrix $\boldsymbol{A}$. Then the signal $\boldsymbol{y}$ computed as

$$\boldsymbol{y} = \boldsymbol{A}\,\boldsymbol{x}, \tag{3}$$

is called as the unit shifted version of $\boldsymbol{x}$. We also would like to indicate that the adjacency matrix is not the only choice for the shift operator in general. The study in [31] proposes alternative definitions for the shift operator. Nevertheless, for simplicity, we will stick with the adjacency matrix as done in [5], [13].

In general, any square matrix of size N, $\boldsymbol{H} \in \mathcal{M}^N$, is considered as a *linear graph filter* on the graph. When we have $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x}$, we call $\boldsymbol{y}$ as the filtered version of the signal $\boldsymbol{x}$. In this study, we are interested in a special type of linear filters, namely polynomial filters, which are defined as follows.

Definition 1 (Polynomial filters [5], [32])**.** *A linear system $\boldsymbol{H}$ on a graph $\boldsymbol{A}$ is said to be a* polynomial *filter if*

$$\boldsymbol{H} = H(\boldsymbol{A}) = \sum_{k=0}^{L} h_k\, \boldsymbol{A}^k, \tag{4}$$

for a set of $h_k \in \mathcal{C}$. Here L is called the order *of the filter.* $\diamondsuit$

We can assume without loss of generality that $L < N$. This is because, according to Cayley-Hamilton theorem, powers A^k for $k \geqslant N$ can be expressed as linear combinations of smaller powers [33].[1]

For a graph with the adjacency matrix $\boldsymbol{A}$, let the following denote the Jordan decomposition [5], [33] of the adjacency matrix

$$\boldsymbol{A} = \boldsymbol{V}\,\boldsymbol{J}\,\boldsymbol{V}^{\text{-}1}, \tag{5}$$

where $\boldsymbol{V}$ is composed of the (generalized) eigenvectors of the adjacency matrix and $\boldsymbol{J}$ is the Jordan normal form of $\boldsymbol{A}$. When $\boldsymbol{A}$ is diagonalizable, (5) reduces to the following form

$$\boldsymbol{A} = \boldsymbol{V}\,\boldsymbol{\Lambda}\,\boldsymbol{V}^{\text{-}1}, \tag{6}$$

for some diagonal $\boldsymbol{\Lambda}$ consisting of the eigenvalues and some invertible square matrix $\boldsymbol{V}$ consisting of the eigenvectors of the adjacency matrix. When $\boldsymbol{A}$ has distinct eigenvalues, it is necessarily diagonalizable, but not vice versa.

Using the Jordan decomposition in (5), we then have the following definitions.

Definition 2 (Graph Fourier transform [5], [13])**.** *Let $\boldsymbol{x}$ be a signal on a graph with the adjacency matrix $\boldsymbol{A}$. Then the graph Fourier transform of $\boldsymbol{x}$ on the graph $\boldsymbol{A}$ is given by*

$$\widehat{\boldsymbol{x}} = \boldsymbol{V}^{\text{-}1}\,\boldsymbol{x}, \tag{7}$$

where $\boldsymbol{V}$ has the (generalized) eigenvectors of $\boldsymbol{A}$ as in (5). $\diamondsuit$

[1] Also see Theorem 3 of [5].

Definition 3 (Frequency domain operation)**.** *Let $\boldsymbol{H}$ be a linear filter on a graph with the adjacency matrix $\boldsymbol{A}$. Then the* frequency domain operator *corresponding to $\boldsymbol{H}$ is defined by*

$$\widehat{\boldsymbol{H}} = \boldsymbol{V}^{\text{-}1}\,\boldsymbol{H}\,\boldsymbol{V}, \tag{8}$$

where $\boldsymbol{V}$ has the (generalized) eigenvectors of $\boldsymbol{A}$ as in (5). $\diamondsuit$

Definiton 3 does not imply that $\boldsymbol{V}$ diagonalizes the filter $\boldsymbol{H}$, that is, $\widehat{\boldsymbol{H}}$ is not diagonal in general.

Notice that Definitions 2 and 3 are consistent with each other, that is, for a graph signal $\boldsymbol{x}$ and a linear filter $\boldsymbol{H}$, we have $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x}$ if and only if $\widehat{\boldsymbol{y}} = \widehat{\boldsymbol{H}}\widehat{\boldsymbol{x}}$. As explained in Sec. VII (and Definition 7) later, $\widehat{\boldsymbol{H}}$ will be referred to as the *frequency response* of $\boldsymbol{H}$ only when $\widehat{\boldsymbol{H}}$ is a diagonal matrix.

II. Building Blocks for Multirate Processing on Graphs

A. Downsampling and Upsampling Operations

One of the most essential building blocks for multirate signal processing is the *decimation operation* [23]. In the graph signal processing, we will assume that this operator retains N/M samples of the original graph signal $\boldsymbol{x}$ that has N samples. It will be assumed that M is a divisor of N. Since the numbering of the graph vertices is flexible [5], [34], we will assume, without loss of generality, that the first N/M samples of $\boldsymbol{x}$ are retained by the decimator. Thus the *graph decimation* operator is defined as:

Definition 4 (Canonical Decimator)**.** *The M-fold graph decimation operator is defined by the matrix*

$$\boldsymbol{D} = \begin{bmatrix} \boldsymbol{I}_{N/M} & \boldsymbol{0}_{N/M} & \cdots & \boldsymbol{0}_{N/M} \end{bmatrix} \in \mathcal{C}^{(N/M)\times N}. \tag{9}$$

Given a graph signal $\boldsymbol{x}$, decimated graph signal *is then denoted as $\boldsymbol{D}\boldsymbol{x}$.* $\diamondsuit$

We refer to $\boldsymbol{D}$ as *canonical decimator* with decimation ratio M. This is a mapping from N dimensional complex space to N/M dimensional complex space. Similar definitions for the decimator operator have been introduced in the literature [8], [9], [16], [35].

Next, the *upsampling operation* $\boldsymbol{U} \in \mathcal{C}^{N\times(N/M)}$ is a mapping from N/M dimensional complex space to N dimensional complex space. Once we define the downsampling, we cannot arbitrarily select the upsampling operator, they should be consistent with each other. In general, downsample-then-upsample is a lossy operation. Contrary to that, upsample-then-downsample operator is expected to be equal to identity. That is to say

$$\boldsymbol{D}\,\boldsymbol{U} = \boldsymbol{I}_{N/M}. \tag{10}$$

For a given $\boldsymbol{D}$, the right inverse $\boldsymbol{U}$ is not unique. When we look for the minimum norm solution, we get

$$\boldsymbol{U} = \boldsymbol{D}^{*}\big(\boldsymbol{D}\boldsymbol{D}^{*}\big)^{\text{-}1}, \tag{11}$$

assuming that $\boldsymbol{D}$ has full row rank. This result reduces to

$$\boldsymbol{U} = \boldsymbol{D}^{T} = \begin{bmatrix} \boldsymbol{I}_{N/M} \\ \boldsymbol{0}_{N/M} \\ \vdots \\ \boldsymbol{0}_{N/M} \end{bmatrix} \in \mathcal{C}^{N\times(N/M)}. \tag{12}$$

for the decimator operator defined in (9). Hence, the corresponding uniform upsampler with expansion ratio M is defined by the matrix $\boldsymbol{D}^T$. This operation merely inserts blocks of zeros, analogous to conventional expanders [23], [24].

With this selection of the upsampler, we have the following equalities for upsample-then-downsample and downsample-then-upsample operations:

$$\boldsymbol{D}\boldsymbol{D}^T = \boldsymbol{I}_{N/M}, \qquad \boldsymbol{D}^T\boldsymbol{D} = \begin{bmatrix} \boldsymbol{I}_{N/M} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \in \mathcal{M}^N, \quad (13)$$

respectively, where zero blocks have appropriate sizes.

In the following, our result will be based on the simple canonical $\boldsymbol{D}$ defined in (9). More generally, decimator can be selected as an arbitrary $(N/M) \times N$ matrix with full row-rank. Such a definition provides an extension to the results presented in the following sections and allows us to remove some of the restrictions on the adjacency matrix. These details are elaborated in Sec. VII of [22].

B. The Noble Identities

In classical signal processing, we have the first noble identity described in Fig. 1(a), where $H(z)$ denotes the transfer function of an LTI filter [23]. For graph signals, it is possible to obtain a similar result under some conditions on the graph. The result is given in Fig. 1(b) and requires some explanation.

In the classical case, the unit delay z^{-1} has the same meaning for both the original and the decimated signals. But for graph signals, the elementary shift operator should match size of the signal. Since the decimated signal has length N/M, we need to define a different shift operator for the decimated signal. The matrix $\bar{\boldsymbol{A}}$ in the figure denotes this *adjusted shift operator*.

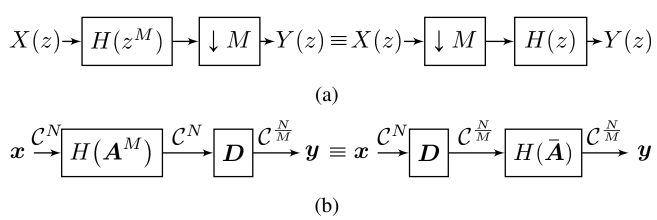


Fig. 1. The first noble identity (a) for classical multirate signal processing where $\downarrow M$ denotes decimator operation, (b) for graph signals on the adjacency matrix $\boldsymbol{A}$.

With the adjusted shift operator for the decimated signal, we have the following form of the first noble identity for graph signals:

$$\boldsymbol{D}\; H(\boldsymbol{A}^M) = H(\bar{\boldsymbol{A}})\; \boldsymbol{D}. \qquad (14)$$

This is shown schematically in Fig. 1(b). It is important to notice that the required adjusted shift operator $\bar{\boldsymbol{A}}$ that satisfies the noble identity in (14) may not exist in general. In the following we will provide the sufficient and necessary condition on $\boldsymbol{A}$ so that an adjusted shift operator exists and satisfies (14).

Theorem 1 (The first noble identity)**.** *Let the decimator* $\boldsymbol{D}$ *be as in* (9)*. If the noble identity* (14) *is satisfied by a graph* $\boldsymbol{A}$ *for all polynomial filters* $H(\cdot)$ *for some* $\bar{\boldsymbol{A}}$*, then* $\boldsymbol{A}^M$ *has to have the form*

$$\boldsymbol{A}^M = \begin{bmatrix} (\boldsymbol{A}^M)_{1,1} & \boldsymbol{0} \\ (\boldsymbol{A}^M)_{2,1} & (\boldsymbol{A}^M)_{2,2} \end{bmatrix}, \qquad (15)$$

where $(\boldsymbol{A}^M)_{1,1} \in \mathcal{M}^{N/M}$*, and furthermore*

$$\bar{\boldsymbol{A}} = \boldsymbol{D}\boldsymbol{A}^M\boldsymbol{D}^T, \qquad (16)$$

i.e., $\bar{\boldsymbol{A}} = (\boldsymbol{A}^M)_{1,1}$*. Conversely if* $\boldsymbol{A}^M$ *and* $\bar{\boldsymbol{A}}$ *have the above form, then noble identity* (14) *holds for all polynomial filters. In short,* (14) *holds for all polynomial filters if and only if both* (15) *and* (16) *are true.* ◊

Proof: First assume (14) holds for all polynomials $H(\cdot)$, i.e., $\boldsymbol{D}\sum_k h_k\, \boldsymbol{A}^{Mk} = \sum_k h_k\, \bar{\boldsymbol{A}}^k\, \boldsymbol{D}$ for all $\{h_k\}$. Then

$$\boldsymbol{D}\; \boldsymbol{A}^{Mk} = \bar{\boldsymbol{A}}^k\; \boldsymbol{D} \qquad (17)$$

for all $k \geqslant 0$. Now express $\boldsymbol{A}^M$ in partitioned form

$$\boldsymbol{A}^M = \begin{bmatrix} (\boldsymbol{A}^M)_{1,1} & (\boldsymbol{A}^M)_{1,2} \\ (\boldsymbol{A}^M)_{2,1} & (\boldsymbol{A}^M)_{2,2} \end{bmatrix}, \qquad (18)$$

where $(\boldsymbol{A}^M)_{1,1} \in \mathcal{M}^{N/M}$. For $k = 1$, (17) yields $\boldsymbol{D}\boldsymbol{A}^M = \bar{\boldsymbol{A}}\boldsymbol{D}$. Using (9) this becomes

$$\begin{bmatrix} (\boldsymbol{A}^M)_{1,1} & (\boldsymbol{A}^M)_{1,2} \end{bmatrix} = \begin{bmatrix} \bar{\boldsymbol{A}} & \boldsymbol{0}_{M,(N\text{-}N/M)} \end{bmatrix}, \qquad (19)$$

which proves $\bar{\boldsymbol{A}} = (\boldsymbol{A}^M)_{1,1}$ and $(\boldsymbol{A}^M)_{1,2} = \boldsymbol{0}$ indeed. Thus (14) implies (15) and (16).

Conversely assume the form (15) and the relation (16) are true. First observe that when (15) holds, we have $(\boldsymbol{A}^{Mk})_{1,1} = ((\boldsymbol{A}^M)_{1,1})^k$. Since (16) also holds, it follows that

$$(\boldsymbol{A}^{Mk})_{1,1} = \bar{\boldsymbol{A}}^k \qquad (20)$$

for all $k \geqslant 0$. This is equivalent to (17), as seen by substituting from (15) and (9). Thus (15) and the relation (16) imply the noble identity (14) indeed. ■

The second noble identity in classical signal processing [23] is described schematically in Fig. 2(a). For graph signals, the analogous identity would be as in Fig. 2(b), where input and output are called as lower and higher rate signal, respectively. Let $\widetilde{\boldsymbol{A}}$ denote the adjusted shift operator for the lower rate signal in the second noble identity. We have the following form of the second noble identity for graph signals.

$$H(\boldsymbol{A}^M)\; \boldsymbol{D}^T = \boldsymbol{D}^T\; H(\widetilde{\boldsymbol{A}}). \qquad (21)$$

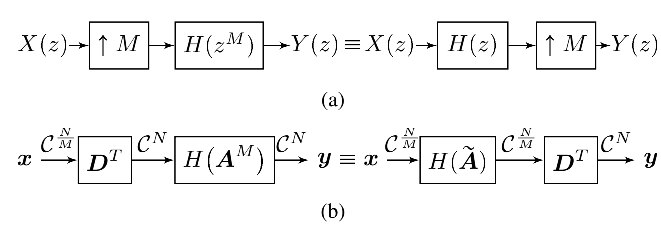


Fig. 2. The second noble identity (a) for classical multirate signal processing where $\uparrow M$ denotes expander operation, (b) for graph signals on the adjacency matrix $\boldsymbol{A}$.

Theorem 2 (The second noble identity)**.** *If the noble identity* (21) *is satisfied by a graph* $\boldsymbol{A}$ *for* all *polynomial filters* $H(\cdot)$ *for some* $\widetilde{\boldsymbol{A}}$*, then* $\boldsymbol{A}^M$ *has to have the form*

$$\boldsymbol{A}^M = \begin{bmatrix} (\boldsymbol{A}^M)_{1,1} & (\boldsymbol{A}^M)_{1,2} \\ \boldsymbol{0} & (\boldsymbol{A}^M)_{2,2} \end{bmatrix}, \tag{22}$$

where $(\boldsymbol{A}^M)_{1,1} \in \mathcal{M}^{N/M}$*, and furthermore*

$$\widetilde{\boldsymbol{A}} = \boldsymbol{D}\boldsymbol{A}^M\boldsymbol{D}^T, \tag{23}$$

i.e., $\widetilde{\boldsymbol{A}} = (\boldsymbol{A}^M)_{1,1}$*. Conversely if* $\boldsymbol{A}^M$ *and* $\widetilde{\boldsymbol{A}}$ *have the above form, then noble identity* (21) *holds for all polynomial filters. In short,* (21) *holds for all polynomial filters if and only if both* (22) *and* (23) *are true.* $\diamondsuit$

Proof: First assume there exists an $\widetilde{\boldsymbol{A}}$ such that (21) is true for *all* polynomial filters $H(\cdot)$. This implies in particular

$$\boldsymbol{A}^{Mk}\boldsymbol{D}^T = \boldsymbol{D}^T\widetilde{\boldsymbol{A}}^k \tag{24}$$

for all $k \geqslant 0$. Now consider the partitioned form in (18). Setting $k = 1$ in (24) and using the form of $\boldsymbol{D}^T$ in (12), we get

$$\begin{bmatrix} \widetilde{\boldsymbol{A}} \\ \boldsymbol{0}_{(N\text{-}N/M),M} \end{bmatrix} = \begin{bmatrix} (\boldsymbol{A}^M)_{1,1} \\ (\boldsymbol{A}^M)_{2,1} \end{bmatrix}, \tag{25}$$

which shows that if (21) has to be true for all polynomial filters, then $\widetilde{\boldsymbol{A}} = (\boldsymbol{A}^M)_{1,1}$ and $(\boldsymbol{A}^M)_{2,1} = \boldsymbol{0}$.

Conversely, suppose the form (22)) and the relation (23) are true. Then $(\boldsymbol{A}^{Mk})_{1,1} = ((\boldsymbol{A}^M)_{1,1})^k = \widetilde{\boldsymbol{A}}^k$. But this is equivalent to (24) as seen by substituting from (22) and (9). So (21) holds for all polynomials $H(\cdot)$ indeed. $\blacksquare$

Combining the preceding two theorems we get

Theorem 3 (The noble identities)**.** *For a graph* $\boldsymbol{A}$*, the two noble identities*

$$\boldsymbol{D}\,H(\boldsymbol{A}^M) = H(\bar{\boldsymbol{A}})\,\boldsymbol{D}, \tag{26}$$

$$H(\boldsymbol{A}^M)\,\boldsymbol{D}^T = \boldsymbol{D}^T\,H(\bar{\boldsymbol{A}}), \tag{27}$$

are simultaneously satisfied for all polynomial filters $H(\cdot)$ *if and only if the following two equations are satisfied:* $\boldsymbol{A}^M$ *has the form*

$$\boldsymbol{A}^M = \begin{bmatrix} (\boldsymbol{A}^M)_{1,1} & \boldsymbol{0} \\ \boldsymbol{0} & (\boldsymbol{A}^M)_{2,2} \end{bmatrix} \tag{28}$$

and

$$\bar{\boldsymbol{A}} = \boldsymbol{D}\,\boldsymbol{A}^M\,\boldsymbol{D}^T \in \mathcal{M}^{N/M}, \tag{29}$$

where $(\boldsymbol{A}^M)_{1,1} \in \mathcal{M}^{N/M}$. $\diamondsuit$

It is clear that an arbitrary graph may not satisfy the condition in (28). Specific examples of graphs that meet, or do not meet, the condition of Theorem 3 will be presented in Sections IV and V.

C. Lazy Graph Filter Banks

An important theoretical example of a maximally decimated filter bank in classical signal processing is the M-channel delay-chain filter bank, also known as the lazy filter bank, shown in Fig. 3(a). This is a perfect reconstruction system [23], and serves as a starting point for developing more useful filter bank systems. Such a development is typically based on the use of polyphase representations and noble identities [23]. We have already developed noble identities for graph signals above. In the following subsection we will develop polyphase representations for graph filters and in Sec. V of [22] we shall develop graph filter banks. In the present subsection we consider the graph equivalent of the lazy filter bank shown in Fig. 3(b). In this system the graph signal $\boldsymbol{x} \in \mathcal{C}^N$ is passed through a chain of graph shift operators $\boldsymbol{A}^k$, $0 \leqslant k \leqslant M\text{-}1$ and each shifted version is passed through the downsampler $\boldsymbol{D}$ and upsampler $\boldsymbol{D}^T$. The resulting M signals are then graph-shifted again and added. It is clear that the system is linear with the input-output relation $\boldsymbol{y} = T(\boldsymbol{A})\,\boldsymbol{x}$ where

$$T(\boldsymbol{A}) = \sum_{k=0}^{M\text{-}1} \boldsymbol{A}^{M\text{-}1\text{-}k}\,\boldsymbol{D}^T\,\boldsymbol{D}\,\boldsymbol{A}^k. \tag{30}$$

For the classical lazy filter bank we have $Y(z) = z^{\text{-}(M\text{-}1)}X(z)$, and it is a perfect reconstruction system. Similarly, we say that the lazy graph filter bank has perfect reconstruction (PR) if $T(\boldsymbol{A}) = \boldsymbol{A}^{M\text{-}1}$, that is,

$$\sum_{k=0}^{M\text{-}1} \boldsymbol{A}^{M\text{-}1\text{-}k}\,\boldsymbol{D}^T\,\boldsymbol{D}\,\boldsymbol{A}^k = \boldsymbol{A}^{M\text{-}1}. \tag{31}$$

This will be referred to as the *lazy FB PR condition*. We will return to more general filter banks on graphs, along with the theory of perfect reconstruction and alias cancellation in Sections II-V of [22].

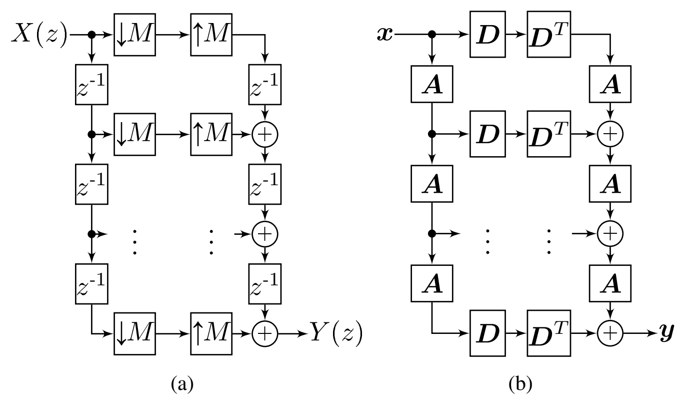


Fig. 3. (a) M-channel lazy filter bank in classical multirate signal processing, (b) M-channel lazy filter bank on a graph with adjacency matrix $\boldsymbol{A}$. The decimation matrix $\boldsymbol{D}$ is as in (9) with decimation ratio M.

D. Polyphase Implementation of Decimation and Interpolation Filters

A useful tool in multirate signal processing is the polyphase representation of linear time-invariant filters [23], [24], [27]. Similar to the classical case, for a given polynomial graph filter, we can write Type-1 polyphase decomposition of the filter as follows:

$$H(\boldsymbol{A}) = \sum_{k=0}^{M\text{-}1} \boldsymbol{A}^k\,E_k(\boldsymbol{A}^M), \tag{32}$$

and Type-2 polyphase decomposition as follows:

$$H(\boldsymbol{A}) = \sum_{k=0}^{M\text{-}1} \boldsymbol{A}^{M\text{-}1\text{-}k}\,R_k(\boldsymbol{A}^M). \tag{33}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2016.2617833, IEEE Transactions on Signal Processing

Notice that there is no assumption on the structure of the adjacency matrix, hence any polynomial filter on any graph has a polyphase representation. As in the classical theory [23], Type-1 and Type-2 polyphase components are related as $R_k(\boldsymbol{A}) = E_{M\text{-}1\text{-}k}(\boldsymbol{A})$.

Fig. 4(a) shows a graph filter followed by a graph decimator on $\boldsymbol{A}$. This is called a decimation filter, in analogy with classical theory. Similarly the system in Fig. 5(a) is called an interpolation filter. For graphs that satisfy the conditions of the noble identities (28), these filters can be implemented in simplified form using the polyphase representation as shown next.

Let $T(\boldsymbol{A})$ denote the overall response of the system in Fig. 4(a). Then we can write it as:

$$\begin{aligned} T(\boldsymbol{A}) &= \boldsymbol{D}\,H(\boldsymbol{A}) = \boldsymbol{D}\sum_{k=0}^{M\text{-}1}\boldsymbol{A}^k\,E_k(\boldsymbol{A}^M) \qquad (34)\\ &= \sum_{k=0}^{M\text{-}1}\boldsymbol{D}\,E_k(\boldsymbol{A}^M)\,\boldsymbol{A}^k = \sum_{k=0}^{M\text{-}1}E_k(\bar{\boldsymbol{A}})\,\boldsymbol{D}\,\boldsymbol{A}^k, \end{aligned}$$

where we use the fact that $E_k(\boldsymbol{A}^M)$ and $\boldsymbol{A}$ commute since E_k is a polynomial in $\boldsymbol{A}$, hence it is shift invariant (see Sec. VIII). We then utilize the noble identity in (26) to get the final result. The adjusted shift operator given in (29) is denoted by $\bar{\boldsymbol{A}}$. Fig. 4(b) and Fig. 4(c) schematically show the steps in (34).

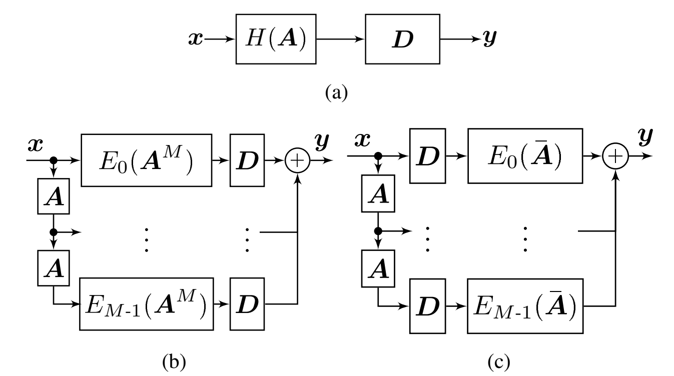


Fig. 4. (a) Polynomial filtering then decimation operation on a graph with the adjacency matrix $\boldsymbol{A}$. (b) Polyphase implementation of (a). (c) Simplification of (b) using the first noble identity (26). The decimation matrix $\boldsymbol{D}$ is as in (9) with decimation ratio M. Implementation in (b) exists without any restriction on the adjacency matrix. However, $\boldsymbol{A}$ should satisfy (15) in order to utilize the first noble identity for the implementation in (c).

Complementary to (34), upsampling followed by a filtering operation can be implemented via Type-2 polyphase decomposition of the filter. Namely, let $T(\boldsymbol{A})$ denote the overall response of the system in Fig. 5(a). Then we can write it as:

$$\begin{aligned} T(\boldsymbol{A}) &= H(\boldsymbol{A})\,\boldsymbol{D}^T = \sum_{k=0}^{M\text{-}1}\boldsymbol{A}^{M\text{-}1\text{-}k}\,R_k(\boldsymbol{A}^M)\,\boldsymbol{D}^T\\ &= \sum_{k=0}^{M\text{-}1}\boldsymbol{A}^{M\text{-}1\text{-}k}\,\boldsymbol{D}^T\,R_k(\bar{\boldsymbol{A}}). \qquad (35) \end{aligned}$$

where we use the fact that $R_k(\boldsymbol{A}^M)$ and $\boldsymbol{A}$ commute since R_k is a polynomial in $\boldsymbol{A}$, hence it is shift invariant (see Sec. VIII). We then utilize the noble identity in (27) to get the final result. Fig. 5(b) and Fig. 5(c) schematically show the steps in (35).

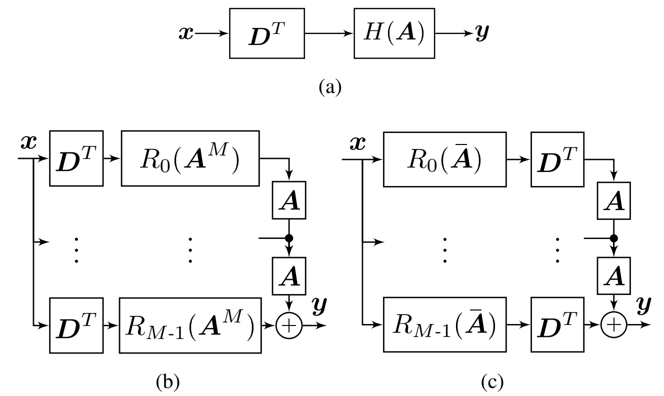


Fig. 5. (a) Expansion then polynomial filtering on a graph with the adjacency matrix $\boldsymbol{A}$. (b) Polyphase implementation of (a). (c) Simplification of (b) using the second noble identity (27). The expansion matrix $\boldsymbol{D}^T$ with expansion ratio M is transpose of $\boldsymbol{D}$, which is in (9). Implementation in (b) exists without any restriction on the adjacency matrix. However, $\boldsymbol{A}$ should satisfy (22) in order to utilize the second noble identity for the implementation in (c).

We will use polyphase implementation of decimation and interpolation filters when we develop polyphase implementation of filter banks in Sec. V of [22].

III. Graph Filter Banks and Polynomial Filters

The ultimate goal in this paper and the companion paper [22] is to develop a theory of analysis/synthesis filter banks for graphs with properties such as perfect reconstruction, alias cancellation, and so forth. Fig. 6 shows such a filter bank for a signal $\boldsymbol{x} \in \mathcal{C}^N$ defined on the graph $\boldsymbol{A}$. Here each analysis filter $\boldsymbol{H}_k$ is an $N \times N$ matrix, and the decimator $\boldsymbol{D}$ is as defined in Sec. II. Since there are M analysis filters and each decimator retains N/M samples, this constitutes a *maximally decimated* analysis bank. The expanders $\boldsymbol{D}^T$ (defined as in Sec. II-A) are followed by synthesis filters $\boldsymbol{F}_k$, which are also $N \times N$ matrices.

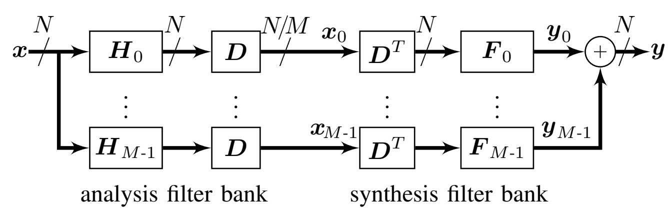


Fig. 6. A maximally decimated graph filter bank where the filters $\boldsymbol{H}_k$ and $\boldsymbol{F}_k$ are arbitrary matrices (i.e., not necessarily polynomials in $\boldsymbol{A}$).

When the filters and decimators are cascaded, the maximally decimated analysis bank can clearly be defined by the M matrices $\{\boldsymbol{DH}_0, \boldsymbol{DH}_1, \ldots, \boldsymbol{DH}_{M\text{-}1}\}$ where $\boldsymbol{DH}_k \in \mathcal{C}^{(N/M)\times N}$. Similarly, the expander and the synthesis filters can be lumped into one matrix $\boldsymbol{F}_k\,\boldsymbol{D}^T \in \mathcal{C}^{N\times(N/M)}$. Therefore, the entire analysis bank, $\boldsymbol{H}_{\text{anl}}$, and the synthesis bank, $\boldsymbol{F}_{\text{syn}}$, are just $N \times N$ matrices as follows:

$$\boldsymbol{F}_{\text{syn}} = \begin{bmatrix}\boldsymbol{F}_0\boldsymbol{D}^T & \cdots & \boldsymbol{F}_{M\text{-}1}\boldsymbol{D}^T\end{bmatrix}, \; \boldsymbol{H}_{\text{anl}} = \begin{bmatrix}\boldsymbol{DH}_0\\ \vdots\\ \boldsymbol{DH}_{M\text{-}1}\end{bmatrix} \qquad (36)$$

Thus, perfect reconstruction property is equivalent to having $\boldsymbol{F}_{\text{syn}}\boldsymbol{H}_{\text{anl}} = \boldsymbol{I}$, so that as long as $\boldsymbol{H}_{\text{anl}}$ has full rank, we can

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2016.2617833, IEEE Transactions on Signal Processing

find synthesis filters for perfect reconstruction. But there are practical difficulties in taking this "brute force" approach with "unconstrained" filter matrices. The complexity of the analysis bank (including decimators) is N^2 **multiplications**, and so is the complexity of the synthesis bank. For large graphs (large N), this complexity can be impractical.

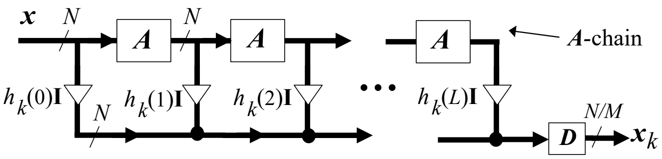


Fig. 7. Implementation of the polynomial graph filter $H_k(\mathbf{A}) = \sum_{n=0}^{L} h_k(n)\mathbf{A}^n$. When the graph is sparse and has simple edge weights this system requires only LN multiplications for its implementation, compared to N^2 in the brute force method of Fig. 6.

In parallel to the classical case, we can force the filters, $\{\boldsymbol{H}_k, \boldsymbol{F}_k\}$, to be polynomial in $\boldsymbol{A}$ as given in (4). This construction is also used in [5], [8], [32]. The advantage is that the graph filters can now be implemented as in Fig. 7, where L is the polynomial degree, and the scalars $h_k(n)$ are the coefficients of the k^{th} filter. This implementation is especially attractive when $\boldsymbol{A}$ is sparse and has simple entries like $0, 1, \text{-}1$, etc., as in many practical graphs (e.g., the Minnesotta traffic graph in [8], [36]). In this case the implementation of the matrix multipliers $\boldsymbol{A}$ has negligible complexity since they only require additions. Therefore an order L polynomial requires only LN multiplications (the coefficients $h_k(i)\boldsymbol{I}$) compared to the N^2 multipliers in the case of unconstrained filters. This is a significant saving when $L \ll N$. Another important point is that, as discussed in [32], [37], an order L polynomial is L-hop localized on the graph, hence, it can be implemented at each node locally via L message passings between neighbors.

IV. Relations to M-Partite Graphs

From Theorem 1 and 2, it is clear that some structure on the adjacency matrix is required in order to generalize the basic concepts in the classical multirate signal processing theory to graph signals. In order to investigate the required structure, we start with bipartite graphs. In [8], bipartite graphs are shown to be useful for 2-channel filter banks where the development was based on the graph Laplacian. We observe the same when we focus on the adjacency matrix. Let $\boldsymbol{A}$ be the adjacency matrix of a directed or undirected bipartite graph in the following form

$$\boldsymbol{A} = \begin{bmatrix} \mathbf{0} & \boldsymbol{A}_2 \\ \boldsymbol{A}_1 & \mathbf{0} \end{bmatrix} \tag{37}$$

where $\boldsymbol{A}_1, \boldsymbol{A}_2 \in \mathcal{M}^{N/2}$. Then, it is straightforward to verify that noble identity condition in (28) is satisfied for $M = 2$. Furthermore one can show that $T(\boldsymbol{A}) = \boldsymbol{A}$, that is, 2-channel lazy filter bank provides perfect reconstruction on a bi-partite graph. Even though (37) considers the case of bipartite graphs with equal sized partitions, it extends to arbitrary bipartite graphs with a proper update on the size of the decimation operator. More importantly, one can also show that bipartiteness is *necessary* for 2-channel lazy FB to provide perfect reconstruction [38].

Even though bipartite graphs are in conformity with the 2-channel systems as discussed above, this relation cannot be generalized to M-channel systems on M-partite graphs. An M-partite graph is one whose vertex set can be partitioned into M subsets so that no edge has both ends in any one subset [39]. Under suitable labelling of the vertices, the adjacency matrix of an M-partite graph can be written as follows:

$$\boldsymbol{A} = \begin{bmatrix} \mathbf{0} & (\boldsymbol{A})_{1,2} & \cdots & (\boldsymbol{A})_{1,M} \\ (\boldsymbol{A})_{2,1} & \mathbf{0} & \ddots & \vdots \\ \vdots & \ddots & \ddots & (\boldsymbol{A})_{M\text{-}1,M} \\ (\boldsymbol{A})_{M,1} & \cdots & (\boldsymbol{A})_{M,M\text{-}1} & \mathbf{0} \end{bmatrix} \tag{38}$$

where $(\boldsymbol{A})_{i,j}$'s have arbitrary but appropriate sizes. In particular, the diagonal blocks of the adjacency matrix of an arbitrary M-partite graph are zero. We have the following *negative* observations for M-partite graphs with $M > 2$:

1) The M-partite property is neither necessary nor sufficient for validity of the noble identity condition (28).
2) The M-partite property is in general not sufficient to ensure perfect reconstruction property of the lazy filter bank of Fig. 3(b).

All proofs, counter-examples and further details of these results can be found in [38].

V. M-Block Cyclic Graphs

Contrary to intuition, the two channel filter bank results on bipartite graphs do not extend to M-channel filter banks on M-partite graphs, as discussed in Section IV. In the following, we will show that, with more restrictive conditions on the graph, it is possible to generalize the classical multirate theory to graph signals for arbitrary M. For this purpose we define the following graph.

Definition 5 (M-block cyclic graphs)**.** *A graph is said to be M-block cyclic if the adjacency matrix of the graph has the following form:*

$$\boldsymbol{A} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \boldsymbol{A}_M \\ \boldsymbol{A}_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{A}_2 & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{A}_3 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{A}_{M\text{-}1} & \mathbf{0} \end{bmatrix} \in \mathcal{M}^N, \tag{39}$$

where each $\boldsymbol{A}_j$ has arbitrary but appropriate sizes. Furthermore, such a graph is said to be balanced *M-block cyclic, when $\boldsymbol{A}_j$'s have the same size, that is $\boldsymbol{A}_j \in \mathcal{M}^{N/M}$. In this case, we can write the adjacency matrix as:*

$$(\boldsymbol{A})_{i,j} = \boldsymbol{A}_j \; \delta(j\text{-}i\text{+}1), \tag{40}$$

where $(\cdot)_{i,j}$ denotes the $(i,j)^{th}$ block of the adjacency matrix and $\delta(\cdot)$ is the M-periodic discrete Dirac function, that is $\delta(Mj) = 1$ for all integer j and zero otherwise. ◇

In the rest of the paper, when we refer to M-block cyclic graphs, *we always mean balanced M-block cyclic graphs.*

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2016.2617833, IEEE Transactions on Signal Processing

Some of the results presented in this study can be generalized to unbalanced M-block cyclic graphs also. However, the adjacency matrix of an unbalanced M-block cyclic graph can be shown to be non-invertible and non-diagonalizable. Such a case requires a careful treatment that falls outside of the scope of this study and will be elaborated elsewhere.

For the visual representation of M-block cyclic graphs, see Fig. 9(a) for a balanced 5-block cyclic graph of size 20. Also consider Fig. 8 to see the relation between the cyclic shift matrix in (2) and M-block cyclic matrices. We now state some properties of M-block cyclic graphs that can be readily verified:

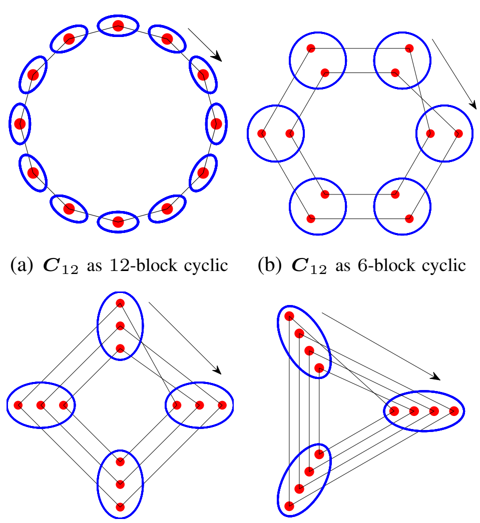
Fact 1. *If a graph is M-block cyclic, then it is M-partite, but not vice-versa.*

Fact 2. *A graph is 2-block cyclic if and only if it is bi-partite.*

Fact 3. *An M-block cyclic graph is necessarily a directed graph for $M > 2$, hence its adjacency matrix does not have any symmetry property in terms of edge weights.*

Fact 4. *A cyclic graph of size N, $\boldsymbol{C}_N$, is an M-block cyclic graph for all M that divides N. See Fig. 8.*

Some other properties of the adjacency matrix of an M-block cyclic graph are presented in Sec. II of the supplementary document [38].



(a) $\boldsymbol{C}_{12}$ as 12-block cyclic (b) $\boldsymbol{C}_{12}$ as 6-block cyclic

(c) $\boldsymbol{C}_{12}$ as 4-block cyclic (d) $\boldsymbol{C}_{12}$ as 3-block cyclic

Fig. 8. Under suitable permutation of the vertices, cyclic graph of size N can be represented as an M-block cyclic graph of size N where M divides N. Notice that cyclic graph of size N is equivalent to N-block cyclic graph of size N. All the edges are directed clock-wise as indicated by the arrow.

Even though arbitrary M-partite graphs are not suitable for M-channel systems as discussed in Section IV, imposing more restrictions and having M-block cyclic structure in (39) provides much more freedom in terms of multirate processing on graphs, which is formally stated in the following theorem.

Theorem 4 (M-block cyclic graphs, noble identities, and lazy filter banks)**.** *Let $\boldsymbol{A}$ be the adjacency matrix of a balanced M-block cyclic graph. Then, noble identity condition in Theorem 3 and lazy FB PR condition in* (31) *are satisfied.* $\diamondsuit$

Proof: According to Corollary 2 of [38], $\boldsymbol{A}^M$ is a block diagonal matrix with blocks of size $\mathcal{M}^{N/M}$, which satisfies the condition in Theorem 3. Therefore, noble identities hold true with the adjusted shift operator

$$\bar{\boldsymbol{A}} = \boldsymbol{D}\,\boldsymbol{A}^M\,\boldsymbol{D}^T = \boldsymbol{A}_M \cdots \boldsymbol{A}_1. \tag{41}$$

For the lazy filter bank condition, consider Corollary 4 and 5 of [38]. Since $\boldsymbol{A}^{\text{-}k}\boldsymbol{D}^T$ is a block-column vector and $\boldsymbol{D}\boldsymbol{A}^k$ is a block-row vector, we have

$$\begin{aligned}\left(\boldsymbol{A}^{\text{-}k}\boldsymbol{D}^T\boldsymbol{D}\boldsymbol{A}^k\right)_{i,j} &= (\boldsymbol{A}^{\text{-}k}\boldsymbol{D}^T)_i\,(\boldsymbol{D}\boldsymbol{A}^k)_j \\ &= \boldsymbol{I}_{N/M}\;\;\delta(i\text{-}1{+}k)\;\delta(j\text{-}1{+}k)\end{aligned} \tag{42}$$

Therefore,

$$\left(\sum_{k=0}^{M\text{-}1}\boldsymbol{A}^{\text{-}k}\boldsymbol{D}^T\boldsymbol{D}\boldsymbol{A}^k\right)_{i,j} = \boldsymbol{I}_{N/M}\;\;\delta(i\text{-}j), \tag{43}$$

that is $\sum_{k=0}^{M\text{-}1}\boldsymbol{A}^{\text{-}k}\boldsymbol{D}^T\boldsymbol{D}\boldsymbol{A}^k = \boldsymbol{I}_N$. Hence, $T(\boldsymbol{A}) = \boldsymbol{A}^{M\text{-}1}$, that is, M-channel lazy filter bank provides perfect reconstruction due to condition in (31). Notice that this proof implicitly assumes that $\boldsymbol{A}$ is invertible. However, the result still holds true even if the adjacency matrix is not invertible as long as it is M-block cyclic. We omit these details for brevity. ∎

VI. Eigen-properties of M-block cyclic graphs

M-block cyclic graphs have an important eigenvalue-eigenvector structure that will play a key role in the development of graph filter banks. This property is as follows.

Theorem 5 (Eigen-families of M-block cyclic graphs)**.** *Eigenvalues and eigenvectors of the adjacency matrix of an M-block cyclic graph come as families of size M. That is, if $(\lambda, \boldsymbol{v})$ is an eigenpair of M-block cyclic graph, then $\{(\lambda,\boldsymbol{v}),\ (w\lambda, \boldsymbol{\Omega}\boldsymbol{v}),\ (w^2\lambda, \boldsymbol{\Omega}^2\boldsymbol{v}),\ \cdots\ (w^{M\text{-}1}\lambda, \boldsymbol{\Omega}^{M\text{-}1}\boldsymbol{v})\}$ are all eigenpairs of the same graph, where*

$$w = e^{-j2\pi/M}, \tag{44}$$

$$\boldsymbol{\Omega} = diag\Big(\big[1\ w^{\text{-}1}\ w^{\text{-}2}\ \cdots\ w^{\text{-}(M\text{-}1)}\big]\Big) \otimes \boldsymbol{I}_{N/M}. \tag{45}$$

$\diamondsuit$

Proof: Let $(\lambda, \boldsymbol{v})$ be an eigenpair of a balanced M-block cyclic graph. Assume that we have the following partitions for the eigenvector

$$\boldsymbol{v} = \Big[(\boldsymbol{v})_1^*\ (\boldsymbol{v})_2^*\ \cdots\ (\boldsymbol{v})_M^*\Big]^* \tag{46}$$

where $(\boldsymbol{v})_i \in \mathcal{C}^{N/M}$ for all $1 \leqslant i \leqslant M$. Then,

$$\boldsymbol{A}\boldsymbol{v} = \begin{bmatrix}\boldsymbol{A}_M\ (\boldsymbol{v})_M \\ \boldsymbol{A}_1\ (\boldsymbol{v})_1 \\ \vdots \\ \boldsymbol{A}_{M\text{-}1}\ (\boldsymbol{v})_{M\text{-}1}\end{bmatrix} = \lambda\boldsymbol{v} = \begin{bmatrix}\lambda\ (\boldsymbol{v})_1 \\ \lambda\ (\boldsymbol{v})_2 \\ \vdots \\ \lambda\ (\boldsymbol{v})_M\end{bmatrix}, \tag{47}$$

that is,

$$\boldsymbol{A}_i\ (\boldsymbol{v})_i = \lambda\ (\boldsymbol{v})_{i+1}. \tag{48}$$

When both sides of (48) are multiplied by $w^{1\text{-}i}$, we get

$$\boldsymbol{A}_i\ \left(w^{1\text{-}i}(\boldsymbol{v})_i\right) = (w\lambda)\ \left(w^{\text{-}i}(\boldsymbol{v})_{i+1}\right). \tag{49}$$

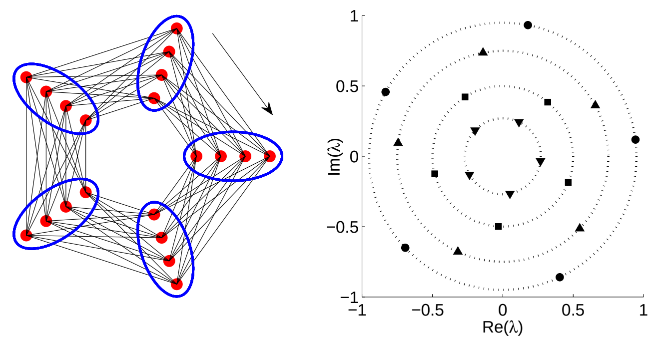
Therefore $w\lambda$ is also an eigenvalue with the corresponding eigenvector $\boldsymbol{v}' = \left[w^0\,(\boldsymbol{v})_1^T\;\; w^{-1}\,(\boldsymbol{v})_2^T\;\;\cdots\;\; w^{-(M-1)}\,(\boldsymbol{v})_M^T\right]^T$. Due to definition of $\boldsymbol{\Omega}$ in (45), we have the following:

$$\left[w^0\,(\boldsymbol{v})_1^T\; w^{-1}\,(\boldsymbol{v})_2^T\;\cdots\; w^{-(M-1)}\,(\boldsymbol{v})_M^T\right]^T = \boldsymbol{\Omega}\,\boldsymbol{v}, \qquad (50)$$

hence, $(w\lambda, \boldsymbol{\Omega v})$ is also an eigen-pair.

Iterating this argument k times, we get $(w^k\lambda, \boldsymbol{\Omega}^k\boldsymbol{v})$ as an eigenpair. However notice that $w^{M+k} = w^k$ and $\boldsymbol{\Omega}^{M+k} = \boldsymbol{\Omega}^k$. Therefore, starting from an eigenpair and iteratively using (49), we can produce at most M-1 distinct eigenpairs. As a result, if $(\lambda, \boldsymbol{v})$ is an eigenpair, $(w^k\lambda, \boldsymbol{\Omega}^k\boldsymbol{v})$ is also an eigenpair for $0 \leqslant k \leqslant M$-1. ■

This eigenvalue relation of block cyclic matrices has also been observed in earlier studies [34], [40]–[42].



(a) (b)

Fig. 9. (a) 5-block cyclic graph of size 20, (b) eigenvalues of the graph. Notice that all the edges are directed along with the clock-wise direction and they have complex valued weights. As given in Theorem 5, eigenvalues of a balanced M-block cyclic graph come as families of size M. Eigenvalues belonging to the same family are equally spaced on a circle in the complex plane. Actual values of the eigenvalues depend on the weight of the edges.

Fig. 9(b) visualizes the relation between the eigenvalues of an M-block cyclic graph. There are N/M concentric circles centered at the origin. Each circle has M eigenvalues equispaced in angle. The circles need not have distinct radii.

One immediate consequence of this eigenfamily structure of the M-block cyclic graph is that eigenvalues can be real only for $M = 2$. We formally state this property as follows.

Corollary 1 (Complex eigenvalues of M-block cyclic)**.** *For $M > 2$, if an M-block cyclic graph has a non-zero eigenvalue, then it has at least one complex valued eigenvalue.*

Proof: Let λ be a non-zero eigenvalue of an M-block cyclic graph. Then $w^k\lambda$ is also eigenvalue for $0 \leqslant k \leqslant M$-1 due to Theorem 5. Therefore for $M > 2$, there exists a k such that $w^k\lambda$ is complex valued. ■

It should be clear that Theorem 5 gives information about only one eigen-family and does *not* imply diagonalizability of the adjacency matrix in general. When $\boldsymbol{A}$ is not diagonalizable, Theorem 5 still applies to its proper eigenvectors, whereas we cannot say too much for the generalized eigenvectors coming from the Jordan chain. However, we note that a randomly generated balanced M-block cyclic matrix is diagonalizable with probability 1.

Assuming that the adjacency matrix is diagonalizable, we will use double indexing to represent the eigenvalues and the eigenvectors of M-block cyclic graphs, since they come as families of size M. That is, the eigenpair $(\lambda_{i,j}, \boldsymbol{v}_{i,j})$ will denote the j^{th} eigenpair of the i^{th} family, where $1 \leqslant i \leqslant N/M$ and $1 \leqslant j \leqslant M$. Using this indexing scheme, with the use of Theorem 5, we have the following form:

$$\lambda_{i,j+k} = w^k\,\lambda_{i,j}, \qquad (51)$$

$$\boldsymbol{v}_{i,j+k} = \boldsymbol{\Omega}^k\,\boldsymbol{v}_{i,j}. \qquad (52)$$

It is important to state that this indexing scheme has a circular structure. Even though we do not explicitly indicate this fact in the notation, it should be clear that $\lambda_{i,j+M} = \lambda_{i,j}$ and $\boldsymbol{v}_{i,j+M} = \boldsymbol{v}_{i,j}$ for all i and j. This property comes from the fact that $w^M = 1$ and $\boldsymbol{\Omega}^M = \boldsymbol{I}$.

With this specific family structure of the eigenvalues of an M-block cyclic graph, when we talk about the eigenvalue decomposition of the adjacency matrix,

$$\boldsymbol{A} = \boldsymbol{V}\,\boldsymbol{\Lambda}\,\boldsymbol{V}^{-1}, \qquad (53)$$

we will assume that eigenvalues and the eigenvectors are ordered as follows:

$$\boldsymbol{\Lambda} = diag\Big(\left[\lambda_{1,1}\cdots\lambda_{1,M}\;\cdots\;\lambda_{N/M,1}\cdots\lambda_{N/M,M}\right]\Big), \qquad (54)$$

$$\boldsymbol{V} = \Big[\boldsymbol{v}_{1,1}\cdots\boldsymbol{v}_{1,M}\;\cdots\;\boldsymbol{v}_{N/M,1}\cdots\boldsymbol{v}_{N/M,M}\Big]. \qquad (55)$$

It is also important to notice that the eigenfamily structure described in (51) and (52) is unique to M-block cyclic graphs. This fact is stated in the following theorem whose proof is given in Sec. I-A of the supplementary document [38].

Theorem 6 (Eigen-structure of M-block cyclic graphs)**.** *Let $\boldsymbol{V}$ be an invertible matrix indexed as in* (55) *with columns that have the property in* (52)*. Let $\boldsymbol{\Lambda}$ be a diagonal matrix indexed as in* (54) *with diagonal entries that have the property in* (51)*. Then $\boldsymbol{A} = \boldsymbol{V\Lambda V}^{-1}$ is diagonalizable M-block cyclic graph. Conversely the adjacency matrix of a diagonalizable M-block cyclic graph always has the form $\boldsymbol{A} = \boldsymbol{V\Lambda V}^{-1}$ where $\boldsymbol{V}$ and $\boldsymbol{\Lambda}$ are as described above.* ◇

In order to enhance our motivation for M-block cyclic graphs, we would like to consider a specific case where $M = 2$. Due to Fact 2 in Sec. V, this is equivalent to bipartite graphs.

For bipartite graphs, we now present Theorem 7 given below. In [8], 2-channel filter banks on bipartite graphs are developed using this result from the spectral graph theory. Note here that the Laplacian of a graph is given as $\boldsymbol{L} = \boldsymbol{D} - \boldsymbol{A}$, where $\boldsymbol{D}$ is the diagonal degree matrix and the normalized Laplacian is given as $\mathcal{L} = \boldsymbol{D}^{-1/2}\boldsymbol{L}\boldsymbol{D}^{-1/2}$.

Theorem 7 (Lemma 1.8 in [43] or Lemma 1 in [8])**.** *The following statements are equivalent for an undirected graph with real non-negative edge weights:*

1) *$\boldsymbol{A}$ is bipartite.*
2) *The spectrum of $\mathcal{L}$ is symmetric about 1 and the minimum and maximum eigenvalues of $\mathcal{L}$ are 0 and 2, respectively.*
3) *If $\boldsymbol{v} = [(\boldsymbol{v})_1^*\;(\boldsymbol{v})_2^*]^*$ is an eigenvector of $\mathcal{L}$ with eigenvalue λ, then the deformed vector $\hat{\boldsymbol{v}} = [(\boldsymbol{v})_1^*\;\text{-}(\boldsymbol{v})_2^*]^*$ is also an eigenvector of $\mathcal{L}$ with eigenvalue 2-λ.* ◇

Notice that Theorem 7 is valid for the normalized Laplacian of the graph. Since we work directly on the adjacency matrix rather than the Laplacian, we will not utilize this result in our development. Interestingly, Theorem 5 provides a very similar statement for the adjacency matrix of the graph when $M = 2$. To see this, observe the following corollary.

Corollary 2 (Bipartite as 2-block cyclic)**.** *If λ is an eigenvalue of the adjacency matrix of an arbitrary balanced bipartite graph with the eigenvector $\boldsymbol{v} = [(\boldsymbol{v})_1^* \ (\boldsymbol{v})_2^*]^*$, then -$\lambda$ will be an eigenvalue of the same graph with the eigenvector $\boldsymbol{v}' = [(\boldsymbol{v})_1^* \ \text{-}(\boldsymbol{v})_2^*]^*$.* ◇

Proof: Set $M = 2$ in Theorem 5. Then $w =$ -1. ■

When the graph is bipartite, $\mathcal{L}$ and $\boldsymbol{A}$ have the same eigenvector structure even though they may have different eigenvectors. However, due to normalization by the degree matrix $\left(\mathcal{L} = \boldsymbol{I} - \boldsymbol{D}^{-1/2}\boldsymbol{A}\boldsymbol{D}^{-1/2}\right)$, symmetric eigenvalues of $\mathcal{L}$ add up to 2, whereas symmetric eigenvalues of $\boldsymbol{A}$ add up to 0, which agrees with the fact that trace of $\boldsymbol{A}$ is zero when it is M-block cyclic. Notice that Corollary 2 is valid for arbitrary bipartite graphs with complex edge values whereas Theorem 7 is constrained to undirected graphs with non-negative edge weights. From this comparison we can conclude that use of $\boldsymbol{A}$ as the unit shift operator rather than $\mathcal{L}$ allows more general class of bipartite graphs. Furthermore, Theorem 5 generalizes this property of 2-channel systems on arbitrary bipartite graphs to M-channel systems on M-block cyclic graphs.

Due to Theorem 4 and 5, we conclude that M-block cyclic graphs defined in (39) have all the necessary properties to generalize the classical multirate theory to the graph signals.

At this point it is interesting to notice the connection to circulant graphs discussed in [18]–[20]. Circulant graphs *do* satisfy the eigenvector condition in (52). This result follows from the fact that DFT matrix diagonalizes *any* circulant matrix. Further, with proper permutations (relabelling of the nodes), DFT matrix satisfies the condition in (52). An example of such a permutation will be demonstrated on the directed cyclic graph, which is a circulant graph, in the following paragraph. *This is very interesting because some of our theorems (Theorem 8 of this paper, Theorems 3 and 4 of [22]) that* only *require the eigenvector condition are now applicable to circulant graphs*. However, the eigenvalue condition of an M-block cyclic graph, (51), is not satisfied by the circulant graphs in general.

The connection to the classical cyclic graph (Sec. II-C of [13]) is also important to understand. For the classical cyclic shift matrix, the classical time domain decimator retains every M^{th} sample (rather than the first N/M samples). But our convention for graphs is that the first N/M samples are retained. To match with our convention, we permute the vertices (i.e., change the numbering convention). This converts the classical cyclic shift matrix into an M-block cyclic matrix. For example, suppose $N = 4$ and $M = 2$. The classical cyclic shift matrix, $\boldsymbol{C}_4$, is

$$\begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \tag{56}$$

where rows and columns are numbered as $0, 1, 2,$ and 3. The classical decimator retains samples 0 and 2 whereas our canonical decimator, by convention, retains 0 and 1. So we simply exchange columns 1 and 2 and also exchange rows 1 and 2. The resulting matrix is:

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \tag{57}$$

which satisfies the requirements of Theorem 1, 2, and 3 (for $M = 2$). In fact the above matrix is a 2-block cyclic matrix. As stated in Fact 4, this permutation is possible for any (N, M) pair where M divides N. For a visual example with $N = 12$, please see Fig. 8.

VII. Concept of Spectrum Folding and Aliasing

In order to talk about alias-free and perfect reconstruction graph filter banks, we need to first define what aliasing is in graph signals. For this purpose we now revisit the downsample-then-upsample (DU) operation. According to our canonical definition of decimator in (9), DU operator is given in (13). Since DU replaces samples with zeros, it is a lossy operation and the erased samples cannot be reconstructed back from the remaining data in general. We now analyze the effect of the DU operator from the frequency domain viewpoint, and explain the *spectrum folding* or *aliasing* effect. A similar approach is presented for two-channel systems in [8], where graph signal processing is based on the graph Laplacian. In our development the graph $\boldsymbol{A}$ is allowed to have complex edge weights and can be directed.

Using the canonical definition of the decimator in (9) and eigenvector-shift operator $\boldsymbol{\Omega}$ in (45), the DU operator can be written as a sum of powers of $\boldsymbol{\Omega}$. That is,

$$\boldsymbol{D}^T\boldsymbol{D} = \frac{1}{M}\sum_{k=0}^{M\text{-}1}\boldsymbol{\Omega}^k. \tag{58}$$

Now consider the DU version of a graph signal $\boldsymbol{x}$, namely

$$\boldsymbol{y} = \boldsymbol{D}^T\,\boldsymbol{D}\,\boldsymbol{x}. \tag{59}$$

Remember that graph Fourier transform of a graph signal $\boldsymbol{x}$ is given in Definition 2. Let $\widehat{\boldsymbol{x}}$ and $\widehat{\boldsymbol{y}}$ denote the graph Fourier transform of the input and output signal of the DU system. Let $\boldsymbol{G}$ denote the frequency domain operation of the DU operator. That is, $\widehat{\boldsymbol{y}} = \boldsymbol{G}\,\widehat{\boldsymbol{x}}$. Due to Definition 3 we have

$$\boldsymbol{G} = \boldsymbol{V}^{\text{-}1}\boldsymbol{D}^T\,\boldsymbol{D}\,\boldsymbol{V}. \tag{60}$$

Using (58), we can write $\boldsymbol{G}$ as follows:

$$\boldsymbol{G} = \frac{1}{M}\sum_{k=0}^{M\text{-}1}\boldsymbol{V}^{\text{-}1}\boldsymbol{\Omega}^k\,\boldsymbol{V}. \tag{61}$$

In the following, we will not constrain ourselves to M-block cyclic graphs and assume that $\boldsymbol{A}$ is diagonalizable and only the eigenvectors of $\boldsymbol{A}$ satisfy (52) and let eigenvalues be arbitrary. In Sec. VII of [22] we will discuss how this assumption on the eigenvectors can be removed by appropriately generalizing the definition of the decimator $\boldsymbol{D}$.

Now notice that $\boldsymbol{\Omega}$ is the eigenvector-shift operator for the eigenvectors satisfying (52). Therefore, $\boldsymbol{\Omega}^k \boldsymbol{V}$ will be the column permuted version of $\boldsymbol{V}$. Due to (52), $\boldsymbol{\Omega}^k$ will circularly shift each vector of an eigen-family to the left by k times. Due to our ordering convention on the eigenvectors in (55), we have the following

$$\boldsymbol{\Omega}^k \boldsymbol{V} = \left[\boldsymbol{v}_{1,1+k} \cdots \boldsymbol{v}_{1,M+k} \cdots \boldsymbol{v}_{N/M,1+k} \cdots \boldsymbol{v}_{N/M,M+k} \right]. \quad (62)$$

Notice that this permutation of the columns of $\boldsymbol{V}$ can also be written with a column permutation matrix. Therefore we have

$$\boldsymbol{\Omega}^k \, \boldsymbol{V} = \boldsymbol{V} \, \boldsymbol{\Pi}_k, \quad (63)$$

where

$$\boldsymbol{\Pi}_k = \boldsymbol{I}_{N/M} \otimes \boldsymbol{C}_M^k = \begin{bmatrix} \boldsymbol{C}_M^k & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \ddots & \boldsymbol{0} \\ \boldsymbol{0} & \cdots & \boldsymbol{C}_M^k \end{bmatrix}, \quad (64)$$

where $\boldsymbol{C}_M$ is the size M cyclic matrix defined in (2). Using (63) and (64), the frequency domain operation $\boldsymbol{G}$ in (60) can be written as:

$$\boldsymbol{G} = \boldsymbol{I}_{N/M} \otimes \frac{1}{M} \sum_{k=0}^{M\text{-}1} \boldsymbol{C}_M^k. \quad (65)$$

Since the first M powers of cyclic matrix of size M add up to matrix with all 1 entries, this response further simplifies to

$$\boldsymbol{G} = \boldsymbol{I}_{N/M} \otimes \frac{1}{M} \mathbb{1}_M \mathbb{1}_M^T, \quad (66)$$

where $\mathbb{1}_M$ denotes the column vector of size M with all 1 entries.

To be consistent with double indexing of the eigenvectors, we will stick to that scheme for the frequency components of a graph signal. That is to say,

$$\widehat{\boldsymbol{x}} = [\widehat{x}_{1,1} \cdots \widehat{x}_{1,M} \; \cdots \; \widehat{x}_{N/M,1} \cdots \widehat{x}_{N/M,M}]^T. \quad (67)$$

Due to (66), we have the following relation between the graph Fourier transform of the original signal and the graph Fourier transform of the downsampled-then-upsampled signal

$$\widehat{y}_{i,1} = \widehat{y}_{i,2} = \cdots = \widehat{y}_{i,M} = \frac{1}{M} \sum_{j=1}^{M} \widehat{x}_{i,j}, \quad (68)$$

for all $1 \leqslant i \leqslant N/M$. We state this result in the following theorem.

Theorem 8 (Spectrum folding in graph signals)**.** *Let $\boldsymbol{A}$ be the adjacency matrix of a graph. Assume that $\boldsymbol{A}$ is diagonalizable and has the eigenvector structure in* (52) *as indexed in* (55) *with arbitrary eigenvalues. Let $\boldsymbol{x}$ be a signal on the graph and $\boldsymbol{y} = \boldsymbol{D}^T \boldsymbol{D} \boldsymbol{x}$ where $\boldsymbol{D}$ is as in* (9). *Then, the graph Fourier transforms of $\boldsymbol{x}$ and $\boldsymbol{y}$ are related as:*

$$\widehat{\boldsymbol{y}} = \frac{1}{M} \left(\boldsymbol{I}_{N/M} \otimes \mathbb{1}_M \mathbb{1}_M^T \right) \widehat{\boldsymbol{x}}. \quad (69)$$

◇

Thus the DU operation results in the phenomenon described by (68) in the frequency domain. This is similar to *aliasing* or *spectral folding* because multiple frequency components of the input overlap into the same frequency component of the output. This is similar to the effect of decimation in classical signal processing [23]. From the folded spectrum (68) we cannot in general recover the original signal, which is consistent with the fact that decimation is in general a information-lossy operation. It should be remembered however that the expression (68) has been derived only for graphs $\boldsymbol{A}$ for which the eigenvectors have the restricted structure (52).

VIII. Linear Systems on Graphs: Interconnection Between Shift Invariance, Alias-Free Property, and Polynomial Property

The above notion of aliasing or spectrum folding due to the DU operator on a graph can be generalized. Thus consider *any* system $\mathcal{S}$ defined on a diagonalizable graph $\boldsymbol{A}$, producing output $\boldsymbol{y} = \mathcal{S}(\boldsymbol{x})$ in response to an input $\boldsymbol{x}$. Let $\widehat{\boldsymbol{x}}$ and $\widehat{\boldsymbol{y}}$ denote the graph Fourier transforms of $\boldsymbol{x}$ and $\boldsymbol{y}$. We say that the *system $\mathcal{S}$ is alias-free* if each component of $\widehat{\boldsymbol{y}}$ is determined by the corresponding component of $\widehat{\boldsymbol{x}}$, i.e., $\widehat{y}_i = g_i(\widehat{x}_i)$. In other words, there is no interference between Fourier components. For the special case of *linear systems* on the graph $\boldsymbol{A}$, this reduces to $\widehat{y}_i = \alpha_i \, \widehat{x}_i$, where α_i is analogous to frequency response.

In classical signal processing, it is well known that linear shift invariant systems are automatically alias-free. For the case of graph signals this equivalence is not always true as we shall elaborate. It was proved in [5] that shift invariance of a linear system on a graph $\boldsymbol{A}$ is equivalent to the statement that the system $\boldsymbol{H}$ be a polynomial (under some conditions, see Theorem 9 below). In this section we will see that the shift invariance, alias-free property, and polynomial property do not imply each other in general. Their inter relationship depends on whether the graph $\boldsymbol{A}$ has distinct eigenvalues or not. These results are elaborated in Theorems 10 and 11, which we shall prove in this section. For clarity we begin with the following formal definitions.

Definition 6 (Shift-invariant filters [5])**.** *Let $\boldsymbol{A}$ be the adjacency matrix of a graph. Let $\boldsymbol{H}$ be a linear system on the graph. It is said that $\boldsymbol{H}$ is* shift-invariant *if it commutes with $\boldsymbol{A}$, that is, $\boldsymbol{A}\boldsymbol{H} = \boldsymbol{H}\boldsymbol{A}$.* ◇

Definition 7 (Alias-free filters)**.** *Let the graph be such that $\boldsymbol{A}$ is diagonalizable, i.e., $\boldsymbol{A} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{\text{-}1}$ for some diagonal $\boldsymbol{\Lambda}$ and invertible $\boldsymbol{V}$. Let $\boldsymbol{H}$ be a linear system on $\boldsymbol{A}$ with frequency domain operation $\widehat{\boldsymbol{H}} = \boldsymbol{V}^{\text{-}1}\boldsymbol{H}\boldsymbol{V}$. We say $\boldsymbol{H}$ is a alias-free filter on graph $\boldsymbol{A}$ if $\widehat{\boldsymbol{H}}$ is a diagonal matrix. In this case $\widehat{\boldsymbol{H}}$ is called the* frequency response *of the filter $\boldsymbol{H}$.* ◇

A polynomial filter is always shift invariant because

$$\boldsymbol{H}\boldsymbol{A} = \left(\sum_{k=0}^{N\text{-}1} \alpha_k \, \boldsymbol{A}^k \right) \boldsymbol{A} = \boldsymbol{A} \left(\sum_{k=0}^{N\text{-}1} \alpha_k \, \boldsymbol{A}^k \right) = \boldsymbol{A}\boldsymbol{H}. \quad (70)$$

But in general the converse is not true. The following result was proved in [5]:

Theorem 9 (Polynomial and shift-invariant graph filters, Theorem 1 in [5])**.** *Let $\boldsymbol{A}$ be the graph adjacency matrix and assume that its characteristic and minimal polynomials are* *equal. Then a graph filter* $\boldsymbol{H}$ *is linear and shift-invariant if and only if* $\boldsymbol{H}$ *is a polynomial on the graph shift* $\boldsymbol{A}$. ♢

We now state and prove the following results.

Theorem 10 (Linear systems on diagonalizable graphs)**.** *Let* $\boldsymbol{H}$ *be a linear system on the graph* $\boldsymbol{A}$*. Assume* $\boldsymbol{A}$ *is diagonalizable. Then the following are true:*

1) *If* $\boldsymbol{H}$ *is a polynomial in* $\boldsymbol{A}$ *then it is alias free.*
2) *If* $\boldsymbol{H}$ *is alias-free then it is shift invariant.* ♢

Proof: 1) Let $\boldsymbol{A} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{-1}$ be the eigenvalue decomposition of $\boldsymbol{A}$. Since $\boldsymbol{H}$ is polynomial in $\boldsymbol{A}$ we have $\boldsymbol{H} = H(\boldsymbol{A})$ where $H(\cdot)$ is a polynomial. Then $\boldsymbol{H} = \boldsymbol{V}H(\boldsymbol{\Lambda})\boldsymbol{V}^{-1}$, where $\boldsymbol{\Lambda}$ is the diagonal matrix consisting of the eigenvalues of $\boldsymbol{A}$. Notice that $H(\boldsymbol{\Lambda}) = \boldsymbol{V}^{-1}\boldsymbol{H}\boldsymbol{V}$ is the frequency domain operation of the system, which is a polynomial of a diagonal matrix. Therefore the overall frequency domain operation is a diagonal matrix, hence it is alias-free.

2) Let $\boldsymbol{A} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{-1}$ be the eigenvalue decomposition of $\boldsymbol{A}$. Assume that $\boldsymbol{H}$ is alias-free. Then it can be written as $\boldsymbol{H} = \boldsymbol{V}\boldsymbol{Z}\boldsymbol{V}^{-1}$ for a diagonal $\boldsymbol{Z}$ due to Definition 7. Then, we have $\boldsymbol{H}\boldsymbol{A} = \boldsymbol{V}\boldsymbol{Z}\boldsymbol{\Lambda}\boldsymbol{V}^{-1} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{Z}\boldsymbol{V}^{-1} = \boldsymbol{A}\boldsymbol{H}$ since diagonal matrices commute. Hence, $\boldsymbol{H}$ is shift-invariant. ■

Theorem 11 (Linear systems on graphs with distinct eigenvalues)**.** *Let* $\boldsymbol{H}$ *be a linear system on the graph* $\boldsymbol{A}$*. Assume* $\boldsymbol{A}$ *has distinct eigenvalues (so that it is, in particular, diagonalizable). Then the following statements are equivalent:*

1) $\boldsymbol{H}$ *is a polynomial in* $\boldsymbol{A}$*.*
2) $\boldsymbol{H}$ *is alias-free.*
3) $\boldsymbol{H}$ *is shift invariant.* ♢

Proof: Since $\boldsymbol{A}$ is diagonalizable, it follows from Theorem 10 that (1) implies (2) and (2) implies (3).

We now prove that (3) implies (2): Assume $\boldsymbol{H}$ is shift invariant, that is, $\boldsymbol{A}\boldsymbol{H} = \boldsymbol{H}\boldsymbol{A}$. Since $\boldsymbol{A}$ has distinct eigenvalues, this implies the following: $\boldsymbol{H}$ is also diagonalizable; $\boldsymbol{A}$ and $\boldsymbol{H}$ are *simultaneously diagonalizable*. (These two claims follows from Problem 13 on page 56 of [33]). But since $\boldsymbol{A}$ has distinct eigenvalues, $\boldsymbol{V}$ is its only diagonalizing matrix (up to a permutation and scaling of columns). So, $\boldsymbol{V}$ in particular, diagonalizes $\boldsymbol{H}$, which (by Definition 7) shows that $\boldsymbol{H}$ is alias-free.

We finally prove that (2) implies (1): Assume $\boldsymbol{H}$ is alias-free, that is, $\boldsymbol{V}^{-1}\boldsymbol{H}\boldsymbol{V} = \boldsymbol{Z}$ is a diagonal matrix with N diagonal elements z_i. Since the eigenvalues λ_i, $1 \leqslant i \leqslant N$ of $\boldsymbol{A}$ are distinct, we can always find a set of N numbers h_i such that the following holds:

$$\begin{bmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{N-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{N-1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \lambda_N & \cdots & \lambda_N^{N-1} \end{bmatrix} \begin{bmatrix} h_0 \\ h_1 \\ \vdots \\ h_{N-1} \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_N \end{bmatrix}. \quad (71)$$

This is because the matrix on the left, being Vandermonde with distinct λ_i, is invertible. Thus, there exists a polynomial $H(\lambda) = \sum_{k=0}^{N-1} h_k\, \lambda^k$ such that $H(\lambda_i) = z_i$. In matrix notation we can rewrite this as $H(\boldsymbol{\Lambda}) = \boldsymbol{Z}$, or

$$\sum_{k=0}^{N-1} h_k\, \boldsymbol{\Lambda}^k = \boldsymbol{Z}, \quad \text{i.e.,} \quad \boldsymbol{V} \sum_{k=0}^{N-1} h_k\, \boldsymbol{\Lambda}^k\, \boldsymbol{V}^{-1} = \boldsymbol{H}, \quad (72)$$

or equivalently $\sum_{k=0}^{N-1} h_k\, \boldsymbol{A}^k = \boldsymbol{H}$, which proves that $\boldsymbol{H}$ is a polynomial in $\boldsymbol{A}$. ■

According to Definition 1 and 6, we can talk about polynomial and shift-invariant filters on a graph with an arbitrary adjacency matrix. However, *the definition of alias-free filters is exclusive to graphs with diagonalizable adjacency matrices*. We intentionally exclude the graphs with non-diagonalizable adjacency matrices due to following reasons. In [13], authors use total variation to quantify the notion of frequency in the graph signals. When the adjacency matrix is not diagonalizable, total variation of a generalized eigenvector inherently depends on the next generalized eigenvector in the Jordan chain that makes it difficult to interpret. Furthermore, when the adjacency matrix is not diagonalizable, even the unit shift element, $\boldsymbol{A}$, has a non-diagonal frequency domain operation. Hence, relation between the polynomial filtering and aliasing in the case of non-diagonalizable adjacency matrices is out of the scope of this work and deserves an independent study.

In the following we will provide three examples to demonstrate the necessity of distinct eigenvalues for the equivalence of the above mentioned three properties. Let $\boldsymbol{A} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{-1}$ be the eigenvalue decomposition of the graph as in (6).

1) Let $\boldsymbol{H}$ be an alias-free filter: $\boldsymbol{H} = \boldsymbol{V}\boldsymbol{Z}\boldsymbol{V}^{-1}$ where $\boldsymbol{Z}$ is a diagonal matrix with *distinct* diagonal entries z_i such that $z_i \neq z_j$ for $i \neq j$. Let λ be a repeated eigenvalue of $\boldsymbol{A}$ with algebraic multiplicity 2. In order to represent $\boldsymbol{H}$ as a polynomial, we need to find a polynomial $H(\cdot)$ such that $H(\lambda) = z_i$ and $H(\lambda) = z_j$ for some $i \neq j$. Since z_i's are distinct, such a function does not exist. Hence, $\boldsymbol{H}$ is alias-free but not polynomial in $\boldsymbol{A}$.
2) Let λ be an eigenvalue of $\boldsymbol{A}$ with algebraic multiplicity m [33]. Assume $m > 1$. Hence, $\boldsymbol{A}$ has repeated eigenvalues. Then we can write $\boldsymbol{\Lambda}$ as follows (by ordering the eigenvectors):

$$\boldsymbol{\Lambda} = \begin{bmatrix} \lambda\boldsymbol{I}_m & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Lambda}' \end{bmatrix}, \qquad \boldsymbol{Z} = \begin{bmatrix} \boldsymbol{Z}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Z}_2 \end{bmatrix} \quad (73)$$

   Let $\boldsymbol{H}$ be such that $\boldsymbol{H} = \boldsymbol{V}\boldsymbol{Z}\boldsymbol{V}^{-1}$ with $\boldsymbol{Z}$ is as in (73) where $\boldsymbol{Z}_2$ is a diagonal but $\boldsymbol{Z}_1$ is a non-diagonalizable square matrix of size m. Notice that $\boldsymbol{\Lambda}$ and $\boldsymbol{Z}$ commute. Hence, $\boldsymbol{A}$ and $\boldsymbol{H}$ commute, that is, $\boldsymbol{H}$ is shift invariant on the graph. But $\boldsymbol{Z}$, which is the frequency domain operation of $\boldsymbol{H}$, is not diagonal since $m > 1$ and $\boldsymbol{Z}_1$ is non-diagonalizable. As a result, $\boldsymbol{H}$ is shift-invariant but not alias-free.
3) Consider the construction in the previous example (73). Since $\boldsymbol{\Lambda}$ is a diagonal matrix, any polynomial of $\boldsymbol{\Lambda}$ will be diagonal. That is, no polynomial of $\boldsymbol{\Lambda}$ is equal to non-diagonal $\boldsymbol{Z}$. Hence, $\boldsymbol{H}$ is shift-invariant but not polynomial.

Notice that Theorem 9 applies to any graph whether its adjacency matrix is diagonalizable or not. In the case of diagonalizable matrices, the minimal polynomial is equal to characteristic polynomial if and only if the matrix has distinct eigenvalues [33]. Fig. 10 schematically shows the relation between shift-invariance, alias-free property, and polynomial property of a linear system on a graph.

When the adjacency matrix is the directed cycle $\boldsymbol{C}_N$, graph signal processing reduces to the classical theory [5]. Since

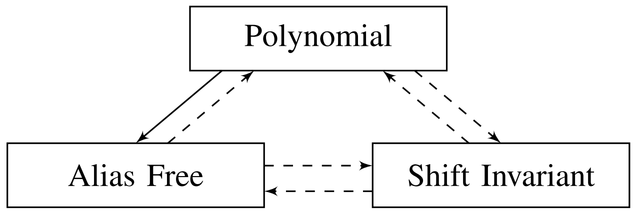


Fig. 10. Relations between the alias-free, shift invariant and polynomial graph filters. Implications shown with solid lines exist for diagonalizable adjacency matrices whereas broken lines further require all eigenvalues to be distinct. In fact, polynomial filters imply shift invariance even if the adjacency matrix is not diagonalizable.

$\boldsymbol{C}_N$ has distinct eigenvalues in the form of $e^{-j2\pi k/N}$ for $0 \leqslant k \leqslant N$-1, polynomial, alias-free and time-invariant filters are equivalent to each other. Therefore, relations in Fig. 10 are consistent with classical signal processing theory but show that our understanding of these properties do not extend to graph case trivially.

IX. Conclusions

In this paper we first developed fundamental blocks for multirate signal processing on graphs by drawing a parallel with classical multirate systems. We started with the canonical definition of the decimator and identified the corresponding expander. We then defined noble identities for graph multirate DSP. Contrary to the classical case, we showed that a certain structure needs to be imposed on the graph to establish these identities. We then studied some graphs that satisfy the conditions and defined M-block cyclic graphs in this context. The unique eigenstructure of such graphs was also shown, and the concept of spectrum folding in such graphs was thereby established. Finally we showed that alias-free systems, polynomial systems, and shift-invariant systems on graphs do not imply each other for arbitrary graphs, and established conditions under which these three concepts are equivalent. In the companion paper [22], we will build upon these results to develop M-channel filter banks on graphs, and study the alias cancellation and perfect reconstruction properties in such graph filter banks.

X. Acknowledgments

The authors wish to thank the anonymous reviewers for the encouragement, criticisms, thoughtful questions, and very useful suggestions.

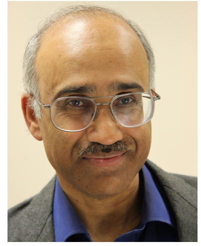
References

[1] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102–114, Aug 2002.
[2] M. O. Jackson, *Social and Economic Networks*. Princeton University Press, 2008.
[3] M. Weber and S. Kube, "Robust perron cluster analysis for various applications in computational life science," in *Computational Life Sciences*. Springer Berlin Heidelberg, 2005, vol. 3695, pp. 57–66.
[4] M. E. J. Newman, *Networks: An Introduction*. Oxford University Press, 2010.
[5] A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs," *IEEE Trans. Signal Process.*, vol. 61, no. 7, pp. 1644–1656, April 2013.
[6] 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.
[7] 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.
[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] S. Narang and A. Ortega, "Downsampling graphs using spectral theory," in *Proc. Int. Conf. Acoust. Speech, Signal Process. (ICASSP)*, May 2011, pp. 4208–4211.
[10] S. Narang, A. Gadde, and A. Ortega, "Signal processing techniques for interpolation in graph structured data," in *Proc. Int. Conf. Acoust. Speech, Signal Process. (ICASSP)*, May 2013, pp. 5445–5449.
[11] 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.
[12] 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.
[13] A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs: Frequency analysis," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3042–3054, June 2014.
[14] Y. Tanaka and A. Sakiyama, "M-channel oversampled graph filter banks," *IEEE Trans. Signal Process.*, vol. 62, no. 14, pp. 3578–3590, July 2014.
[15] A. Sakiyama and Y. Tanaka, "Oversampled graph laplacian matrix for graph filter banks," *IEEE Trans. Signal Process.*, vol. 62, no. 24, pp. 6425–6437, Dec 2014.
[16] S. Chen, R. Varma, A. Sandryhaila, and J. Kovacevic, "Discrete signal processing on graphs: Sampling theory," *IEEE Trans. Signal Process.*, vol. 63, no. 24, pp. 6510–6523, Dec 2015.
[17] S. Chen, A. Sandryhaila, and J. Kovacevic, "Sampling theory for graph signals," in *Proc. Int. Conf. Acoust. Speech, Signal Process. (ICASSP)*, April 2015, pp. 3392–3396.
[18] V. Ekambaram, G. Fanti, B. Ayazifar, and K. Ramchandran, "Critically-sampled perfect-reconstruction spline-wavelet filterbanks for graph signals," in *Global Conf. on Signal and Information Process. (GlobalSIP)*, Dec 2013, pp. 475–478.
[19] V. Ekambaram, G. Fanti, B. Ayazifar, and K. Ramchandran, "Circulant structures and graph signal processing," in *Int. Conf. on Image Process. (ICIP)*, Sept 2013, pp. 834–838.
[20] V. Ekambaram, G. Fanti, B. Ayazifar, and K. Ramchandran, "Spline-like wavelet filterbanks for multiresolution analysis of graph-structured data," *IEEE Trans. Signal and Information Process. over Networks*, vol. 1, no. 4, pp. 268–278, Dec 2015.
[21] N. Tremblay and P. Borgnat, "Subgraph-based filterbanks for graph signals," *IEEE Trans. Signal Process.*, vol. 64, no. 15, pp. 3827–3840, Aug. 2016.
[22] O. Teke and P. P. Vaidyanathan, "Extending classical multirate signal processing theory to graphs – Part II: M-Channel filter banks," *IEEE Trans. Signal Process.*, to appear.
[23] P. P. Vaidyanathan, *Multirate systems and filter banks*. Englewood Cliffs, N.J. Prentice Hall, 1993.
[24] P. P. Vaidyanathan, "Multirate digital filters, filter banks, polyphase networks, and applications: a tutorial," *Proceedings of the IEEE*, vol. 78, no. 1, pp. 56–93, Jan 1990.
[25] M. Vetterli and J. Kovacevic, *Wavelets and Subband Coding*. Prentice-Hall, 1995.
[26] G. Strang and T. Nguyen, *Wavelets and Filter Banks*. Wellesley-Cambridge Press, 1996.
[27] R. E. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing*. Prentice Hall, 1987.
[28] F. J. Harris, *Multirate Signal Processing for Communication Systems*. Prentice Hall, 2004.
[29] O. Teke and P. P. Vaidyanathan, "Fundamentals of multirate graph signal processing," in *Asilomar Conf. on Signals, Systems and Computers*, Nov 2015, pp. 1791–1795.
[30] O. Teke and P. P. Vaidyanathan, "Graph filter banks with m-channels, maximal decimation, and perfect reconstruction," in *Proc. Int. Conf. Acoust. Speech, Signal Process. (ICASSP)*, March 2016, pp. 4089–4093.
[31] A. Gavili and X.-P. Zhang, "On the shift operator and optimal filtering in graph signal processing," *arXiv: 1511.03512v2*, Nov. 2015.

[32] S. Narang and A. Ortega, "Compact support biorthogonal wavelet filterbanks for arbitrary undirected graphs," *IEEE Trans. Signal Process.*, vol. 61, no. 19, pp. 4673–4685, Oct 2013.
[33] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1990.
[34] D. M. Cvetkovic, M. Doob, and H. Sachs, *Spectra of Graphs: Theory and Application (Pure & Applied Mathematics)*. Academic Press, 1980.
[35] H. Nguyen and M. Do, "Downsampling of signals on graphs via maximum spanning trees," *IEEE Trans. Signal Process.*, vol. 63, no. 1, pp. 182–191, Jan 2015.
[36] S. Narang and A. Ortega. (2013) Graph bior wavelet toolbox. [Online]. Available: http://biron.usc.edu/wiki/index.php/Graph_Filterbanks
[37] 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.
[38] O. Teke and P. P. Vaidyanathan. (2016) Supplementary material for extending classical multirate signal processing theory to graphs – Part I. [Online]. Available: http://systems.caltech.edu/dsp/students/oteke/files/supp1.pdf
[39] J. A. Bondy and U. Murty, *Graph Theory With Applications*. Elsevier Science Ltd/North-Holland, 1976.
[40] D. S. Watkins, "Product eigenvalue problems," *SIAM Review*, vol. 47, no. 1, pp. 3–40, 2005.
[41] M. Fiedler, *Special Matrices and Their Applications in Numerical Mathematics: Second Edition*. Dover Publications, 2008.
[42] C. D. Meyer, *Matrix analysis and applied linear algebra*. Society for Industrial and Applied Mathematics, 2000.
[43] F. R. K. Chung, *Spectral Graph Theory*. American Mathematical Society, 1997.

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

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



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

Dr. Vaidyanathan served as Vice-Chairman of the Technical Program committee for the 1983 IEEE International symposium on Circuits and Systems, and as the Technical Program Chairman for the 1992 IEEE International symposium on Circuits and Systems. He was an Associate editor for the IEEE Transactions on Circuits and Systems for the period 1985-1987, and is currently an associate editor for the journal IEEE Signal Processing letters, and a consulting editor for the journal Applied and computational harmonic analysis. He has been a guest editor in 1998 for special issues of the IEEE Trans. on Signal Processing and the IEEE Trans. on Circuits and Systems II, on the topics of filter banks, wavelets and subband coders.

Dr. Vaidyanathan has authored nearly 500 papers in journals and conferences, and is the author/coauthor of the four books Multirate systems and filter banks, Prentice Hall, 1993, Linear Prediction Theory, Morgan and Claypool, 2008, and (with Phoong and Lin) Signal Processing and Optimization for Transceiver Systems, Cambridge University Press, 2010, and Filter Bank Transceivers for OFDM and DMT Systems, Cambridge University Press, 2010. He has written several chapters for various signal processing handbooks. He was a recipient of the Award for excellence in teaching at the California Institute of Technology for the years 1983-1984, 1992-93 and 1993-94. He also received the NSF's Presidential Young Investigator award in 1986. In 1989 he received the IEEE ASSP Senior Award for his paper on multirate perfect-reconstruction filter banks. In 1990 he was recepient of the S. K. Mitra Memorial Award from the Institute of Electronics and Telecommuncations Engineers, India, for his joint paper in the IETE journal. In 2009 he was chosen to receive the IETE students' journal award for his tutorial paper in the IETE Journal of Education. He was also the coauthor of a paper on linear-phase perfect reconstruction filter banks in the IEEE SP Transactions, for which the first author (Truong Nguyen) received the Young outstanding author award in 1993. Dr. Vaidyanathan was elected Fellow of the IEEE in 1991. He received the 1995 F. E. Terman Award of the American Society for Engineering Education, sponsored by Hewlett Packard Co., for his contributions to engineering education. He has given several plenary talks including at the IEEE ISCAS-04, Sampta-01, Eusipco-98, SPCOM-95, and Asilomar-88 conferences on signal processing. He has been chosen a distinguished lecturer for the IEEE Signal Processing Society for the year 1996-97. In 1999 he was chosen to receive the IEEE CAS Society's Golden Jubilee Medal. He is a recepient of the IEEE Signal Processing Society's Technical Achievement Award for the year 2002, and the IEEE Signal Processing Society's Education Award for the year 2012. He is a recipient of the IEEE Gustav Kirchhoff Award (an IEEE Technical Field Award) in 2016, for "Fundamental contributions to digital signal processing." In 2016 he also received the Northrup Grumman Prize for excellence in Teaching at Caltech.