

Chapter 1

Spectral Design of Signal-Adapted Tight Frames on Graphs

Hamid Behjat and Dimitri Van De Ville

Abstract Analysis of signals defined on complex topologies modeled by graphs is a topic of increasing interest. Signal decomposition plays a crucial role in the representation and processing of such information, in particular, to process graph signals based on notions of scale (e.g., coarse to fine). The graph spectrum is more irregular than for conventional domains; i.e., it is influenced by graph topology, and, therefore, assumptions about spectral representations of graph signals are not easy to make. Here, we propose a tight frame design that is adapted to the graph Laplacian spectral content of a given class of graph signals. The design exploits the ensemble energy spectral density, a notion of spectral content of the given signal set that we determine either directly using the graph Fourier transform or indirectly through approximation using a decomposition scheme. The approximation scheme has the benefit that (i) it does not require diagonalization of the Laplacian matrix, and (ii) it leads to a smooth estimate of the spectral content. A prototype system of spectral kernels each capturing an equal amount of energy is defined. The prototype design is then warped using the signal set's ensemble energy spectral density such that the resulting subbands each capture an equal amount of ensemble energy. This approach accounts at the same time for graph topology and signal features, and it provides a meaningful interpretation of subbands in terms of coarse-to-fine representations.

1.1 Introduction

Many fields of science rely on network analysis to study complex systems. Networks are modeled mathematically as weighted graphs that have a set of nodes

Hamid Behjat
Lund University, Sweden, e-mail: hamid.behjat@bme.lth.se

Dimitri Van De Ville
École Polytechnique Fédérale de Lausanne, and the Department of Radiology and Medical Informatics, University of Geneva, Switzerland, e-mail: dimitri.vandeville@epfl.ch

(vertices) with interactions between them represented by connections (links) and associated strengths. A rich repertoire of methods have been developed to pursue original queries and integrate the complexity of network structure into the analysis, subsequently providing new interpretations of datasets in diverse scientific disciplines ranging from social sciences to physics and biology. One of the successes in network analysis is the ability to identify sets of nodes based on their connectivity. Traditional graph partitioning goes back to optimizing graph cuts [14], while more recent community detection identifies sets of nodes that are more densely connected inside the set than outside [17]. Community detection has been widely applied and many variants of the corresponding optimization criterion have been proposed [31].

Another significant trend in the field is the emergence of methods to process signals on graphs [32, 41, 38, 7]. Measurements on the nodes of a given network can be considered as graph signals for which classical signal processing operations can be generalized; e.g., how to properly denoise, filter, or transform graph signals by taking into account the underlying connectivity. Many generalization schemes have been proposed to extend classical multi-resolution transforms, filter bank designs and dictionary constructions to the graph setting. These studies fall essentially within two families: spatial (vertex) and spectral (frequency) designs. Schemes that fall within the former family include methods in designing wavelets for hierarchical trees [34, 35, 16] and methods based on lifting schemes [22, 29, 36]. The latter family is based on spectral graph theory [8], which is a powerful approach based on the eigendecomposition of matrices associated with graphs such as the adjacency matrix or graph Laplacian. Its strength originates from the global nature of the eigenvectors that summarize key graph properties and can be used to solve convex relaxed versions of graph cut minimization [50], or to define signal-processing operations by a graph equivalent of the Fourier transform [41, 38]. In its application to graph signal processing, operations are performed in the spectral domain using graph spectral filters. One of the first proposals is the spectral graph wavelet transform (SGWT) frame [20] that is constructed based on a system of scaled cubic spline spectral kernels together with a lowpass spectral polynomial kernel. Moreover, various constructions of systems of spectral graph kernels leading to tight frames were proposed in [26, 18, 13]. Tight frames are particularly interesting because of their property of energy conservation between the original and transformed domain [5]. Other approaches to spectral domain design include diffusion wavelets [10], vertex-frequency frames [42, 44] and approaches to graph filter-bank design using bipartite graph decompositions [30, 45, 46, 37], connected sub-graph decomposition [48], graph coloring [40] and Slepian functions that provide a tradeoff between temporal and spectral energy concentration [49].

One of the difficulties of the graph spectrum is that its construction depends on the graph itself. Consequently, the graph spectral representation of a graph signal is determined by both the domain and the signal. However, the aforementioned spectral designs typically define spectral windows in a way that is independent of the graph and graph signal. One example of adaptation to the spectral properties of the graph domain was recently proposed in [43] for the construction of spectrum-adapted tight graph wavelet and vertex-frequency frames. The spectrum-adapted

kernels account for the non-uniform distribution of Laplacian eigenvalues, and are designed such that a similar number of eigenvalues falls within the support of each spectral kernel. Moreover, in [47, 51], numerical dictionary learning approaches have been proposed in which dictionaries are learnt based on a set of training signals. In these design, the learned kernels are indirectly adapted to the graph Laplacian spectrum as well as to the training data since the graph structure is incorporated into the learning process. In an application specific approach, in [3, 1, 2], the Meyer-like frame design [26] has been tailored to the spectral content of functional MRI signals to obtain a set of narrow-support kernels covering the lower end of the spectrum.

In this chapter, we propose an approach for constructing tight graph frames that account not only for the intrinsic topological structure of the underlying graph as proposed in [43], but also for the characteristics of a given set of signals. This is accomplished by considering a graph-based energy spectral density notion that includes signal and topology properties and encodes the energy-wise significance of the graph eigenvalues. A system of spectral kernels tailored to the energy spectral density is constructed by starting from the design of a prototype Meyer-type tight frame with uniform spectral coverage, followed by a warping step which incorporates the energy spectral density information to the prototype design, resulting in a tight frame with equi-energy subbands.

1.2 Preliminaries

1.2.1 Graphs and Spectral Graph Theory

A graph can be denoted as $G = (V, E)$ with N_g vertices in set V , a set of edges as tuples (i, j) in E where $i, j \in V$. The size of the graph is the number of vertices. In this chapter we only consider undirected graphs without self-loops. Algebraically, G can be represented with the node-to-node adjacency matrix A , with elements $a_{i,j}$ denoting the weight of the edge (i, j) if $(i, j) \in E$; $a_{i,j} = 0$ if $(i, j) \notin E$. The degree matrix D of G is diagonal with elements $d_{i,i} = \sum_j a_{i,j}$. The Laplacian matrices of G in combinatorial form L and normalized form \mathcal{L} are defined as

$$L = D - A, \quad (1.1)$$

$$\mathcal{L} = D^{-1/2} L D^{-1/2}, \quad (1.2)$$

respectively. Both L and \mathcal{L} are symmetric and positive semi-definite, and thus, their diagonalizations lead to a set of N_g real, non-negative eigenvalues that define the graph Laplacian spectrum

$$\Lambda(G) = \{0 = \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_{N_g} = \lambda_{\max}\}. \quad (1.3)$$

The corresponding set of eigenvectors $\{\chi_l\}_{l=1}^{N_g}$ forms a complete set of orthonormal vectors that span the graph spectral domain [8]. When necessary, we use the notations $\Lambda_L(G)$ and $\Lambda_{\mathcal{L}}(G)$ to distinguish between the two definitions of the graph Laplacian. As the eigenvalues may be repetitive, for each λ_l , we denote its algebraic multiplicity by m_{λ_l} and the index of its first occurrence by i_{λ_l} . That is, if λ_l is singular, i.e. $m_{\lambda_l} = 1$, then $i_{\lambda_l} = l$, and if λ_l is repetitive, then $i_{\lambda_l} \leq l$. The multiplicity of eigenvalues equal to zero reflects the number of connected components in the graph. In this paper, only connected graphs are considered, and thus, $m_{\lambda_1} = 1$.

1.2.2 Graph Signals: Vertex versus Spectral Representations

Let $\ell_2(G)$ denote the Hilbert space of all square-summable real-valued vectors $\mathbf{f} \in \mathbb{R}^{N_g}$, with the inner product defined as

$$\langle \mathbf{f}_1, \mathbf{f}_2 \rangle = \sum_{n=1}^{N_g} f_1[n]f_2[n], \quad \forall \mathbf{f}_1, \mathbf{f}_2 \in \ell_2(G) \quad (1.4)$$

and the norm as

$$\|\mathbf{f}\|_2^2 = \langle \mathbf{f}, \mathbf{f} \rangle = \sum_{n=1}^{N_g} |f[n]|^2, \quad \forall \mathbf{f} \in \ell_2(G). \quad (1.5)$$

A real signal defined on the vertices of a graph, $\mathbf{f}: V \rightarrow \mathbb{R}$, can be seen as a vector in $\ell_2(G)$, where the n -th element represents the value of the signal on the n -th vertex.

For any $\mathbf{f} \in \ell_2(G)$, its spectral representation $\widehat{\mathbf{f}} \in \ell_2(G)$, known as the graph Fourier transform of \mathbf{f} , can be used to express \mathbf{f} in terms of the graph Laplacian eigenvectors

$$f[n] = \sum_{l=1}^{N_g} \underbrace{\langle \mathbf{f}, \chi_l \rangle}_{=\widehat{f}[l]} \chi_l[n]. \quad (1.6)$$

With this definition of the Fourier transform, it can be shown that the Parseval relation holds [42]

$$\forall \mathbf{f}_1, \mathbf{f}_2 \in \ell_2(G), \quad \langle \mathbf{f}_1, \mathbf{f}_2 \rangle = \langle \widehat{\mathbf{f}}_1, \widehat{\mathbf{f}}_2 \rangle. \quad (1.7)$$

1.2.3 Filtering of Graph Signals

In the graph setting, the generalized convolution product is defined as

$$(\mathbf{f}_1 * \mathbf{f}_2)[n] = \sum_{l=1}^{N_g} \widehat{f}_1[l] \widehat{f}_2[l] \chi_l[n], \quad \forall \mathbf{f}_1, \mathbf{f}_2 \in \ell_2(G). \quad (1.8)$$

In analogy with conventional signal processing, filtering of graph signals can be viewed as an operation in the spectral domain. For a given graph signal $\mathbf{f} \in \ell_2(G)$ and graph filter $\mathbf{g} \in \ell_2(G)$, defined through its Fourier transform $\widehat{\mathbf{g}}$, the filtered signal, denoted by $(F_{\mathbf{g}}\mathbf{f})$, can be obtained as

$$(F_{\mathbf{g}}\mathbf{f})[n] = (\mathbf{g} * \mathbf{f})[n] \quad (1.9)$$

$$\stackrel{(1.8)}{=} \sum_{l=1}^{N_g} \widehat{g}[l] \widehat{f}[l] \chi_l[n]. \quad (1.10)$$

For the graph filter \mathbf{g} , the filter response of an impulse at vertex m

$$\mathbf{f} = \delta_m \leftrightarrow \widehat{\delta}_m[l] = \langle \delta_m, \chi_l \rangle = \chi_l[m], \quad (1.11)$$

can be obtained as

$$(F_{\mathbf{g}}\delta_m)[n] = \sum_{l=1}^{N_g} \widehat{g}[l] \chi_l[m] \chi_l[n]. \quad (1.12)$$

The impulse response of a graph filter is, in general, shift-variant; i.e., the impulse response at one vertex is not simply a shifted version of the impulse response at any other node. This is due to the absence of a well-defined shift operator in the graph setting as that defined in the Euclidean setting. Therefore, a graph filter is conventionally defined by its spectral kernel $\widehat{\mathbf{g}}$ rather than by its impulse response.

Although the graph spectrum is discrete, to design spectral kernels, it is often more elegant to define an underlying smooth continuous kernel. Let $L_2(G)$ denote the Hilbert space of all square-integrable spectral functions $K(\lambda) : [0, \lambda_{\max}] \rightarrow \mathbb{R}^+$, with the inner product defined as

$$\langle K_1, K_2 \rangle_{L_2} = \int_{-\infty}^{+\infty} K_1(\lambda) K_2(\lambda) d\lambda, \quad \forall K_1, K_2 \in L_2(G), \quad (1.13)$$

and the L_2 -norm defined as

$$\|K\|_{L_2}^2 = \langle K, K \rangle_{L_2}, \quad \forall K \in L_2(G). \quad (1.14)$$

A discrete version of $K(\lambda) \in L_2(G)$ can then be determined as

$$k[l] = K(\lambda_l), \quad l = 1, \dots, N_g. \quad (1.15)$$

Note that although \mathbf{k} is defined in the spectral domain, it is not linked to any explicit vertex representation, and thus, the Fourier symbol $\widehat{\cdot}$ is not used for their denotation. This notation convention will be used throughout the chapter.

1.2.4 Dictionary of Graph Atoms

For a given spectral kernel \mathbf{k} associated with $K(\lambda)$, the vertex-domain impulse responses are obtained as

$$\Psi_{K,m} = (F_{\mathbf{k}} \delta_m) \leftrightarrow \widehat{\Psi}_{K,m}[l] = k[l] \chi_l[m]. \quad (1.16)$$

The collection of impulse responses $\{\Psi_{K,m}\}_{m=1}^{N_g}$ are considered as graph *atoms* associated with spectral kernel $K(\lambda)$. Given a set of J spectral kernels $\{\mathbf{k}_j \in \ell_2(G)\}_{j=1}^J$, a dictionary of graph atoms D_G with JN_g elements can be obtained

$$D_G = \left\{ \left\{ \Psi_{K_j,m} \right\}_{m=1}^{N_g} \right\}_{j=1}^J. \quad (1.17)$$

The atoms of D_G form a frame in $\ell_2(G)$ if there exist bounds $B_2 \geq B_1 > 0$ such that [5]

$$\forall \mathbf{f} \in \ell_2(G), \quad B_1 \|\mathbf{f}\|_2^2 \leq \sum_{j,m} |\langle \mathbf{f}, \Psi_{K_j,m} \rangle|^2 \leq B_2 \|\mathbf{f}\|_2^2, \quad (1.18)$$

where the frame bounds are given by

$$B_1 = \min_{\lambda \in [0, \lambda_{\max}]} G(\lambda), \quad B_2 = \max_{\lambda \in [0, \lambda_{\max}]} G(\lambda), \quad (1.19)$$

and $G(\lambda) \in L_2(G)$ is defined as

$$G(\lambda) = \sum_{j=1}^J |K_j(\lambda)|^2. \quad (1.20)$$

In particular, D_G forms a tight frame if

$$\forall \lambda \in [0, \lambda_{\max}], \quad G(\lambda) = C, \quad (1.21)$$

and a Parseval frame if $C = 1$.

1.2.5 Decomposition of Graph Signals

Direct Decomposition

To decompose a graph signal \mathbf{f} onto a set of the atoms in D_G , the coefficients can be obtained as

$$c_{K_j,m} = \langle \mathbf{f}, \Psi_{K_j,m} \rangle \quad (1.22)$$

$$\stackrel{(1.7)}{=} \sum_{l=1}^{N_g} \widehat{\Psi}_{K_j,m}[l] \widehat{f}[l], \quad (1.23)$$

$$\stackrel{(1.16)}{=} \sum_{l=1}^{N_g} k_j[l] \widehat{f}[l] \chi_l[m]. \quad (1.24)$$

Relation (1.24) shows that the direct decomposition requires a full eigendecomposition of the L since it requires i) the Laplacian eigenvectors $\{\chi_l\}_{l=1}^{N_g}$ and ii) the graph Fourier transform of the signal $\widehat{\mathbf{f}}$.

If D_G forms a Parseval frame, the coefficients can be used to recover the original signal as

$$\begin{aligned} f[n] &= \sum_j \sum_m c_{K_j,m} \Psi_{K_j,m} \\ &= \sum_j \sum_m \sum_l k_j[l] \widehat{f}[l] \chi_l[m] \sum_{l'} k_j[l'] \chi_{l'}[m] \chi_{l'}[n] \\ &= \sum_l \sum_{l'} \sum_j k_j[l] k_j[l'] \widehat{f}[l] \chi_{l'}[n] \underbrace{\sum_m \chi_l[m] \chi_{l'}[m]}_{\delta_{l-l'}} \\ &= \sum_l \underbrace{\sum_j k_j^2[l]}_{=1} \widehat{f}[l] \chi_l[n]. \end{aligned} \quad (1.25)$$

Decomposition Through Polynomial Approximation

The decomposition of \mathbf{f} on D_G leads to a coefficient vector associated to each \mathbf{k}_j given as

$$\mathbf{c}_{K_j} = [c_{K_j,1}, c_{K_j,2}, \dots, c_{K_j,N_g}]^T \quad (1.26)$$

$$\stackrel{(1.24)}{=} \sum_{l=1}^{N_g} k_j[l] \widehat{f}[l] \chi_l, \quad (1.27)$$

that can be interpreted as filtered versions of \mathbf{f} with different spectral kernels $\{\mathbf{k}_j\}_{j=1}^J$. Due to the redundancy of such a transform, it is beneficial to implement the transform using a fast algorithm, rather than using the explicit computation of the coefficients through (1.24). Moreover, for large graphs, it can be cumbersome to compute the full eigendecomposition of L , and in extensively large graphs this can in fact be impossible. One solution to overcome this computational burden is to use a polynomial approximation scheme.

One such algorithm is the truncated Chebyshev polynomial approximation method [20], which is based on considering the expansion of the continuous spectral window

functions $\{K_j(\lambda)\}_{j=1}^J$ with the Chebyshev polynomials $C_p(x) = \cos(p \arccos(x))$ as

$$K_j(\lambda) = \frac{1}{2}d_{K_j,0} + \sum_{p=1}^{\infty} d_{K_j,p} \bar{C}_p(\lambda), \quad (1.28)$$

where $\bar{C}_p(x) = C_p(\frac{x-b}{b})$, $b = \lambda_{\max}/2$ and $d_{K_j,p}$ denote the Chebyshev coefficients obtained as

$$d_{K_j,p} = \frac{2}{\pi} \int_0^{\pi} \cos(p\theta) K_j(b(\cos(\theta) + 1)) d\theta. \quad (1.29)$$

By truncating (1.28) to M terms, $K_j(\lambda)$ can be approximated as an M -th order polynomial $P_j(\lambda) \in L_2(G)$. Consequently, \mathbf{c}_{K_j} can be approximated as

$$\mathbf{c}_{K_j} \stackrel{(1.27)}{=} \sum_{l=1}^{N_g} \underbrace{k_j[l]}_{K_j(\lambda_l)} \widehat{f}[l] \chi_l \quad (1.30)$$

$$\approx \sum_{l=1}^{N_g} P_j(\lambda_l) \widehat{f}[l] \chi_l \quad (1.31)$$

$$= P_j(L) \sum_{l=1}^{N_g} \widehat{f}[l] \chi_l \quad (1.32)$$

$$\stackrel{(1.6)}{=} P_j(L) \mathbf{f} \quad (1.33)$$

where in (1.32) we exploit the property $L\chi_l = \lambda_l\chi_l \Rightarrow P_j(L)\chi_l = P_j(\lambda_l)\chi_l$.

1.3 Ensemble Energy Spectral Density

The ensemble energy spectral density can be either computed using the graph Fourier transform or approximated through decomposition of the signals using polynomial approximation. In the former approach the ensemble energy is determined at the resolution of eigenvalues whereas in the latter approach it is determined at the resolution of a given number of subbands. The direct computation approach has two shortcomings. Firstly, it requires explicit computation of the graph spectrum and the associated eigenvectors; i.e., a full eigendecomposition of the graph Laplacian matrix, which is computationally cumbersome for large graphs and infeasible for extensively large graphs. Secondly, it typically results in a non-smooth description of the ensemble energy. These shortcomings are resolved by using the polynomial approximation scheme.

1.3.1 Direct Computation: Using the Graph Fourier Transform

Definition (ensemble energy spectral density)

For a given graph G , with spectrum $\Lambda(G)$, and graph signal set $F = \{\mathbf{f}_s\}_{s=1}^{N_s}$, the ensemble energy spectral density of F is obtained as

$$e_F[l] = \frac{1}{N_s} \sum_{s=1}^{N_s} \left| \widehat{f}_s[l] \right|^2, \quad l = 1, \dots, N_g, \quad (1.34)$$

where \widehat{f}_s denotes the de-meaned and normalized version of \mathbf{f}_s obtained as

$$\widehat{f}_s = \frac{\mathbf{f}_s - \sum_{r=1}^{1+m\lambda_1} \langle \mathbf{f}_s, \chi_r \rangle \chi_r}{\|\mathbf{f}_s - \sum_{r=1}^{1+m\lambda_1} \langle \mathbf{f}_s, \chi_r \rangle \chi_r\|_2}, \quad s = 1, \dots, N_s. \quad (1.35)$$

The ensemble energy spectral density has the following properties: (i) $\{e_F[r] = 0\}_{r=1}^{1+m\lambda_1}$, and (ii) $\sum_l e_F[l] = 1$.

1.3.2 Approximation: Using Decomposition through Polynomial Approximation

The ensemble energy spectral density can be approximated through a multi subband decomposition scheme. In the sequel, we first design a B-spline based system of spectral kernels. The benefit in using a B-spline basis is in the smoothness characteristic of such kernels. Smooth overlapping kernels are advantageous in that i) they enable obtaining a smooth estimation of the ensemble energy spectral density and ii) they can be approximated as low order polynomials. We then decompose the graph signals using the designed system of kernels with a large number of subbands by exploiting the polynomial approximation scheme in decomposition. With such a decomposition, we approximate the ensemble spectral content of the signal set at the resolution of subbands.

1.3.2.1 B-spline based Parseval Frames on Graphs

The central B-spline of degree n , denoted $\beta^{(n)}(x)$, is a compactly-supported function in the interval $[-\Delta^{(n)}, \Delta^{(n)}]$, i.e., $\beta^{(n)}(x) = 0$ for all $|x| \geq \Delta^{(n)}$ where $\Delta^{(n)} = (n+1)/2$, and is obtained through the $(n+1)$ -fold convolution as

$$\beta^{(n)}(x) = \underbrace{\beta^{(0)}(x) * \beta^{(0)}(x) * \dots * \beta^{(0)}(x)}_{(n+1)\text{times}}, \quad (1.36)$$

where

$$\beta^{(0)}(x) = \begin{cases} 1, & -\frac{1}{2} < x < \frac{1}{2} \\ \frac{1}{2}, & |x| = \frac{1}{2} \\ 0, & \text{otherwise.} \end{cases} \quad (1.37)$$

Proposition 1 (*B-spline based Parseval Frame on Graphs*) For a given graph G and B-spline generating function $\beta^{(n)}(x)$, $n \geq 2$, a set of B-spline based spectral kernels $\{B_j(\lambda) \in L_2(G)\}_{j=1}^J$ can be defined as

$$B_j(\lambda) = \begin{cases} \tilde{B}_j(\lambda) + \sum_{i=-\Delta}^0 \tilde{B}_i(\lambda), & j = 1 \\ \tilde{B}_j(\lambda), & j = 2, \dots, J-1 \\ \tilde{B}_j(\lambda) + \sum_{i=J+1}^{J+\Delta+1} \tilde{B}_i(\lambda), & j = J \end{cases} \quad (1.38)$$

where $\Delta = \lfloor n/2 \rfloor - 1$ and $\tilde{B}_l(\lambda) \in L_2(G)$ is defined as

$$\tilde{B}_l(\lambda) = \sqrt{\beta^{(n)}\left(\frac{\lambda_{\max}}{J-1}(\lambda - l + 1)\right)}, \quad l = -\Delta, \dots, J + \Delta + 1. \quad (1.39)$$

The system of kernels $\{B_j(\lambda)\}_{j=1}^J$ satisfy

$$\sum_{j=1}^J |B_j(\lambda)|^2 = 1, \quad \forall \lambda \in [0, \lambda_{\max}], \quad (1.40)$$

and, thus, their associated dictionary of atoms forms a Parseval frame.

Proof. See Appendix A.

Fig. 1.1 shows two realizations of spline-type systems of spectral kernels. The spline-type system of spectral kernels have wide, overlapping passbands. Moreover, the kernels are smooth and can thus be approximated as low order polynomials.

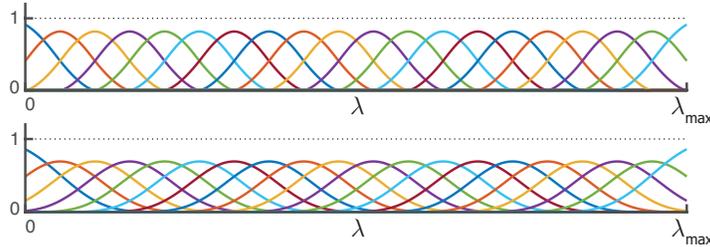


Fig. 1.1 Spline-type system of spectral kernels with 20 spectral bands constructed based on B-splines of order 3 (top) and 7 (bottom).

Using a system of N_a B-spline based spectral kernels, $\{B_i(\lambda)\}_{i=1}^{N_a}$, the ensemble spectral energy of F can be approximated at N_a overlapping bands across the

spectrum as

$$a_F[i] = \frac{1}{N_s} \sum_{s=1}^{N_s} \sum_{n=1}^{N_g} \left| \langle \tilde{\mathbf{f}}_s, \boldsymbol{\psi}_{B_i, n} \rangle \right|^2, \quad i = 1, \dots, N_a, \quad (1.41)$$

where $\tilde{\mathbf{f}}_s$ is as given in (1.35). Let $\mathbf{b}_j \in \ell_2(G)$ denote the discrete version of $B_j(\lambda)$, i.e.,

$$b_j[l] = B_j(\lambda_l), \quad l = 1, \dots, N_g. \quad (1.42)$$

We have $\sum_i a_F[i] = 1$ since

$$\sum_i a_F[i] \stackrel{(1.24)}{=} \frac{1}{N_s} \sum_{i=1}^{N_a} \sum_{s=1}^{N_s} \sum_{n=1}^{N_g} \left| \sum_{l=1}^{N_g} b_i[l] \widehat{f}_s[l] \chi_l[n] \right|^2 \quad (1.43)$$

$$= \frac{1}{N_s} \sum_{s=1}^{N_s} \sum_{n=1}^{N_g} \left| \underbrace{\sum_{l=1}^{N_g} \sum_{i=1}^{N_a} b_i^2[l] \widehat{\mathbf{f}}_s[l] \chi_l[n]}_{\stackrel{(1.40)}{=} 1} \right|^2 \quad (1.44)$$

$$= \frac{1}{N_s} \sum_{s=1}^{N_s} \sum_{n=1}^{N_g} \left| \sum_{l=1}^{N_g} \widehat{f}_s[l] \chi_l[n] \right|^2 \quad (1.45)$$

$$\stackrel{(1.6)}{=} \frac{1}{N_s} \sum_{s=1}^{N_s} \sum_{n=1}^{N_g} |\tilde{f}_s[n]|^2 \quad (1.46)$$

$$= \frac{1}{N_s} \sum_{s=1}^{N_s} \|\tilde{\mathbf{f}}_s\|_2^2 \quad (1.47)$$

$$\stackrel{(1.35)}{=} 1. \quad (1.48)$$

If desired, an explicit approximation of the ensemble energy spectral density of F , denoted $e_F^{(a)}[l]$, can also be determined. First, a continuous ensemble spectral energy representation, denoted $E_F^{(a)}(\lambda)$, is obtained through interpolating the set of points

$$\left\{ (0, 0) \cup \left\{ \left(\frac{\lambda_{\max}}{C} \sum_{k=1}^i \|B_k(\lambda)\|_2^2, a_F[k] \right) \right\}_{i=1}^{N_a} \right\}. \quad (1.49)$$

where $C = \sum_{k=1}^{N_a} \|B_k(\lambda)\|_2^2$. Then, $e_F^{(a)}[l]$ is obtained through sampling $E_F^{(a)}(\lambda)$ at $\Lambda(G)$ as

$$e_F^{(a)}[l] = E_F^{(a)}(\lambda_l), \quad l = 1, \dots, N_g. \quad (1.50)$$

1.4 Signal-Adapted System of Spectral Kernels

The construction of a signal-adapted system of spectral kernels is motivated by two observations: (i) the eigenvalues of the graph Laplacian that define the graph's spectrum are irregularly spaced, and depend in a complex way on the graph topology; (ii) the distribution of graph signals' energy is generally non-uniform across the spectrum. Based on these observations, the idea is to construct an 'adapted' frame, such that the energy-wise significance of the eigenvalues is taken into account, rather than only adapting based on the distribution of the eigenvalues as proposed in [43]. In this way, also the topological information of the graph is implicitly incorporated in the design, since the energy content is given in the graph spectral domain that is in turn defined by the eigenvalues.

For the design of a signal-adapted system of spectral kernels with J subbands, denoted $\{S_j(\lambda)\}_{j=1}^J$, we start off from a prototype system of spectral kernels $\{U_j(\lambda)\}_{j=1}^J$ that satisfies the following two properties:

- (Uniformity constraint)

$$\exists C \in \mathbb{R}^+, \quad \int_0^{\lambda_{\max}} U_j(\lambda) d\lambda = C, \quad j = 1, \dots, J. \quad (1.51)$$

- (Tight Parseval frame constraint)

$$\sum_{j=1}^J |U_j(\lambda)|^2 = 1, \quad \forall \lambda \in [0, \lambda_{\max}]. \quad (1.52)$$

We then exploit the ensemble energy spectral density \mathbf{e}_F or the approximated ensemble spectral energy \mathbf{a}_F to introduce the desired signal adaptivity. The adaptivity is introduced by first transforming the ensemble spectral energy measures to an energy-equalizing transformation $T_F(\lambda) : [0, \lambda_{\max}] \rightarrow [0, \lambda_{\max}]$, which is then in turn incorporated into the prototype design.

1.4.1 Prototype Uniform System of Spectral Kernels

There is no unique system of kernels that satisfies (1.51) and (1.52). We present a design in which the kernels have a finite support of the bandpass type.

Proposition 2 (*uniform Meyer-type (UMT) system of spectral kernels*) Using the auxiliary function of the Meyer wavelet, given by [28]

$$v(x) = x^4(35 - 84x + 70x^2 - 20x^3), \quad (1.53)$$

a set of $J \geq 2$ spectral kernels defined as

$$U_1(\lambda) = \begin{cases} 1 & \forall \lambda \in [0, a] \\ \cos(\frac{\pi}{2} \nu(\frac{1}{\gamma-1}(\frac{\lambda}{a} - 1))) & \forall \lambda \in]a, \gamma a] \\ 0 & \text{elsewhere} \end{cases} \quad (1.54a)$$

$$U_j(\lambda) = \begin{cases} \sin(\frac{\pi}{2} \nu(\frac{1}{\gamma-1}(\frac{\lambda-(j-2)\Delta}{a} - 1))) & \forall \lambda \in]\lambda_I, \lambda_{II}] \\ \cos(\frac{\pi}{2} \nu(\frac{1}{\gamma-1}(\frac{\lambda-(j-1)\Delta}{a} - 1))) & \forall \lambda \in]\lambda_{II}, \lambda_{II} + \Delta] \\ 0 & \text{elsewhere} \end{cases} \quad (1.54b)$$

$$U_J(\lambda) = \begin{cases} \sin(\frac{\pi}{2} \nu(\frac{1}{\gamma-1}(\frac{\lambda-(J-2)\Delta}{a} - 1))) & \forall \lambda \in]\lambda_I, \lambda_{II}] \\ 1 & \forall \lambda \in]\lambda_{II}, \lambda_{II} + a] \\ 0 & \text{elsewhere} \end{cases} \quad (1.54c)$$

can be constructed, where

$$\Delta = \gamma a - a, \quad (1.55a)$$

$$\lambda_I = a + (j-2)\Delta, \quad (1.55b)$$

$$\lambda_{II} = \gamma a + (j-2)\Delta, \quad (1.55c)$$

$$a = \frac{\lambda_{\max}}{J\gamma - J - \gamma + 3}. \quad (1.55d)$$

Fig. 1.2 illustrates the notations used. By setting $\gamma = 2.73$, the set of kernels defined in (1.54) satisfies the uniformity constraint given in (1.51). The atoms of a dictionary constructed using this set of spectral kernels form a Parseval frame on $\ell_2(G)$.

Proof. See Appendix B.

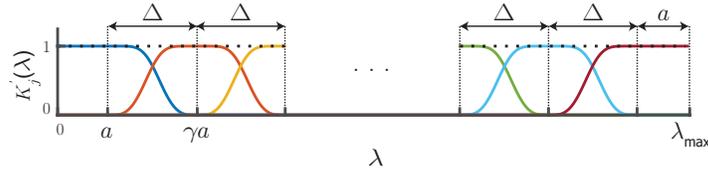


Fig. 1.2 Construction of UMT system of spectral kernels.

Figs. 1.3(a) and (b) show realizations of the resulting UMT system of spectral kernels for a fixed λ_{\max} and two different J . The UMT system of spectral kernels have a narrow passband characteristic with the support of each kernel being a rather strict subset of the spectrum, with minimal overlap of adjacent kernels.

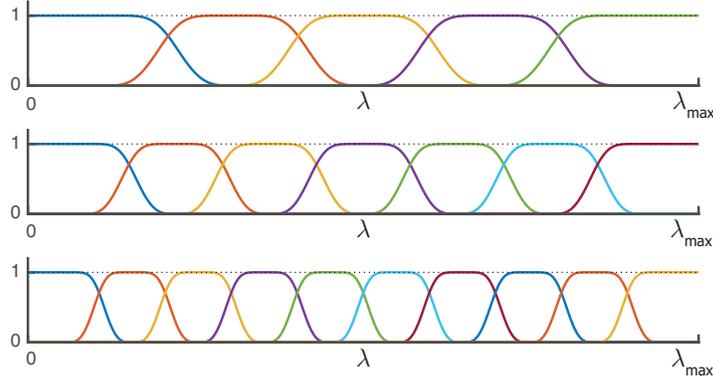


Fig. 1.3 UMT system of spectral kernels with $J = 5$ (top), $J = 7$ (middle) and $J = 10$ (bottom) spectral scales.

1.4.2 Energy-Equalizing Transformation

If the ensemble spectral density function is available, $T_F(\lambda)$ is obtained through monotonic cubic interpolation [15] of the pair of points

$$\left\{ \left(\lambda_l, \frac{\lambda_{\max}}{m\lambda_l} \sum_{r=i\lambda_l}^{i\lambda_l+m\lambda_l} \sum_{k=1}^r e_F[k] \right) \right\}_{l=1}^{N_g}. \quad (1.56)$$

If the ensemble energy spectral density is approximated using a system of N_a B-spline based spectral kernels (cf. 1.3.2), $T_F(\lambda)$ can instead be obtained through monotonic cubic interpolation of the set of points

$$\left\{ (0, 0) \cup \left\{ \left(\frac{\lambda_{\max}}{C} \sum_{k=1}^i \|B_k(\lambda)\|_2^2, \lambda_{\max} \sum_{k=1}^i a_F[k] \right) \right\}_{i=1}^{N_a} \right\}, \quad (1.57)$$

where $C = \sum_{j=1}^J \|B_j(\lambda)\|_2^2$.

1.4.3 Warping the Prototype Design

By incorporating $T_F(\lambda)$ in $\{U_j(\lambda)\}_{j=1}^J$, a warped version of the prototype design is obtained as

$$S_j(\lambda) = U_j(T_F(\lambda)), \quad j = 1, \dots, J. \quad (1.58)$$

We refer to $\{S_j(\lambda)\}_{j=1}^J$ as a signal-adapted system of spectral kernels. The atoms of a dictionary constructed using $\{S_j(\lambda)\}_{j=1}^J$ form a Parseval frame on $\ell_2(G)$ since

$$\begin{aligned} \sum_{j=1}^J |S_j(\lambda)|^2 &\stackrel{(1.58)}{=} \sum_{j=1}^J |U_j(\underbrace{T_F(\lambda)}_{:=\lambda'})|^2, \quad \forall \lambda \in [0, \lambda_{\max}] \\ &= \sum_{j=1}^J |U_j(\lambda')|^2, \quad \forall \lambda' \in [0, \lambda_{\max}] \\ &= 1 \end{aligned}$$

where the last equality follows from Proposition 2.

If a discrete representation is needed for direct decomposition as in (1.24), $\{s_j\}_{j=1}^J$ can be obtained through sampling $S_j(\lambda)$ at $\Lambda(G)$.

With this design, each of the J spectral kernel $\{s_j\}_{j=1}^J$ capture an equal amount of *ensemble* energy. That is, if the ensemble energy spectral density is used we have

$$\sum_{l=1}^{N_g} s_j[l] e_F[l] = \frac{1}{J}, \quad j = 1, \dots, J, \quad (1.59)$$

and if the approximation scheme is used we have

$$\sum_{l=1}^{N_g} s_j[l] e_F^{(a)}[l] = \frac{1}{J}, \quad j = 1, \dots, J. \quad (1.60)$$

Moreover, the resulting system of spectral kernels form a partition of unity, i.e.,

$$\sum_{j=1}^J |s_j[l]|^2 = 1, \quad l = 1, \dots, N_g, \quad (1.61)$$

and thus, their associated dictionary of atoms, i.e., $\left\{ \left\{ \psi_{S_j, m} \right\}_{j=1}^J \right\}_{m=1}^{N_g}$, forms a Parseval frame.

1.5 Example Spectral Designs of Signal-Adapted Tight Frame

We present constructions of signal-adapted systems of spectral kernels for signal sets realized on the Minnesota road graph, the Alameda graph [47] and the cerebellum gray matter graph [2, 1]. Before proceeding to the constructions, let us consider a model for simulating random graph signals of varying smoothness. The model will be used to realize signals on the Minnesota and Alameda graphs, although there exists also real data for the latter graph. For a given graph with adjacency matrix A , we consider a general model for realizing graph signals of density $\eta \in [0, 1]$ and

smoothness $n \in \mathbb{Z}^+$ as

$$\mathbf{x}_{\eta,n} = A^n \mathbf{p}_\eta, \quad (1.62)$$

where $\mathbf{p}_\eta \in \ell_2(G)$ denotes a random realization of a spike signal as $\{\mathbf{p}_\eta[i] \in \{0, 1\}\}_{i=1, \dots, N_g}$ such that $\sum_i \mathbf{p}_\eta[i] = \eta N_g$. Application of the n -th power of A to \mathbf{p}_η leads to a signal that i) respects the intrinsic structure of the graph and ii) has a desired smoothness determined by n , a higher n leading to a smoother graph signal.

1.5.1 The Minnesota Road Graph

The edges of the Minnesota Road Graph represent major roads and its vertices their intersection points, which often correspond to towns or cities, see Fig. 1.4(a). Fig. 1.4(b) shows the graph's normalized Laplacian spectrum presented as the distribution of the eigenvalues.

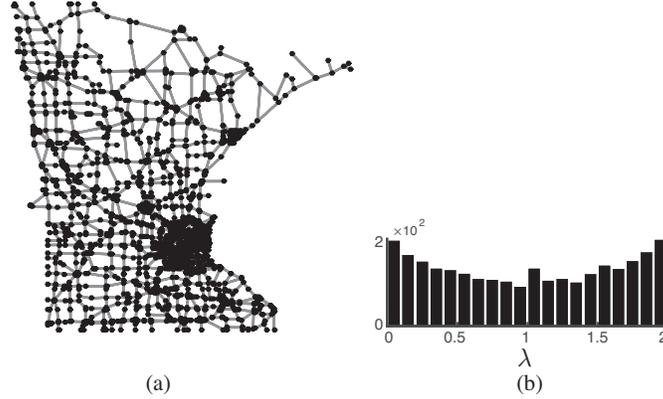


Fig. 1.4 (a) Minnesota road graph. (b) Histograms of the eigenvalues $\Lambda_{\mathcal{L}}(G)$ of the Minnesota road graph. Each bar indicates the number of eigenvalues that lie in the corresponding spectral range.

Two sets of graph signals were constructed as

$$F_1 = \left\{ \left\{ \mathbf{x}_{\eta,2}^{[i]} \right\}_{\eta=0.2,0.5} \right\}_{i=1, \dots, 10},$$

$$F_2 = \left\{ \left\{ \mathbf{x}_{\eta,4}^{[i]} \right\}_{\eta=0.2,0.5} \right\}_{i=1, \dots, 10},$$

where index i denotes random realizations of p_η in (1.62), resulting in 20 signals in each set. Figs. 1.5(a) and (b) show a realizations of a signal from F_1 and F_2 , respectively.

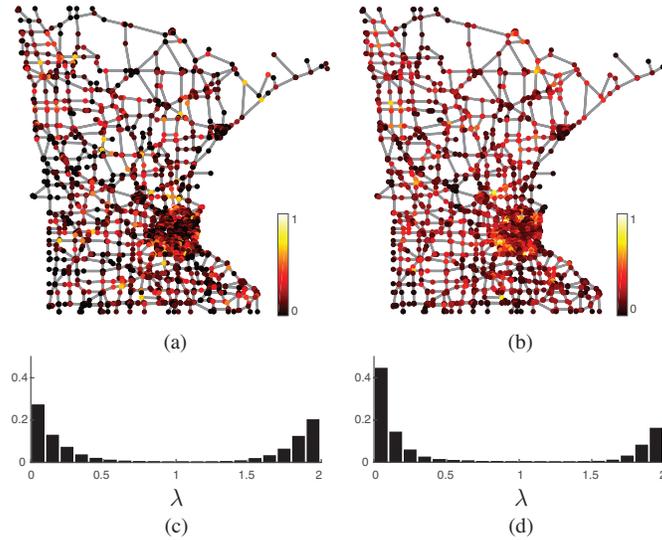


Fig. 1.5 Sample signal realizations on the Minnesota road graph, (a) $\mathbf{x}_{0.2,2}$ and (b) $\mathbf{x}_{0.5,4}$. The plots are normalized as $\mathbf{x}_{\eta,n}/\|\mathbf{x}_{\eta,n}\|_{\infty}$ (c)-(d) Distribution of the ensemble energy spectral density e_{F_1} and e_{F_2} , respectively. Each bar indicates the sum of ensemble energies of the eigenvalues lying in the corresponding spectral range.

Fig. 1.6(a) shows the energy-equalizing transformation functions associated to F_1 and F_2 . The transformations constructed based on a_F , cf. (1.57) closely matches that constructed based on e_F , cf. (1.56). The former transformation has the benefit of being smooth, and indeed, that it was computed without the explicit need to diagonalize L . By incorporating the transformations in the UMT system of spectral kernels, signal-adapted systems of spectral kernels are obtained, see Figs. 1.6(b)-(c).

A comparison of Figs. 1.6(b) and (c) and Figs. 1.5(c) and (d) highlights the energy-wise optimality of the proposed signal-adapted frame construction; i.e., more filters are allocated to spectral ranges that have higher ensemble energy. The support of the filters in the two sets vary relative to the difference in the distribution of the ensemble energy of the two signal sets, with more filters allocated to the lower end of the spectrum for the F_2 frame than for the F_1 frame, and vice versa at the upper end of the spectrum. For comparison, a spectrum-adapted system of kernels is shown in Figs. 1.6(d). The spectrum-adapted system of kernels is obtained by warping the UMT prototype system of kernels with a spectrum-equalizing transformation function $T_{\mathcal{L}}(\lambda)$ which equalizes the distribution of the eigenvalues [43]. As the distribution of the eigenvalues of the Minnesota Road graph minimally deviate from a uniform distribution, so does the spectrum-adapted system of kernels relative to the UMT prototype, compare Figs. 1.3 and 1.6(d). On the contrary, the signal-adapted design optimizes the construction of the kernels such that the energy-wise significance of the eigenvalues is taken into account, rather than only considering the distribution of the eigenvalues as in the spectrum-adapted frame.

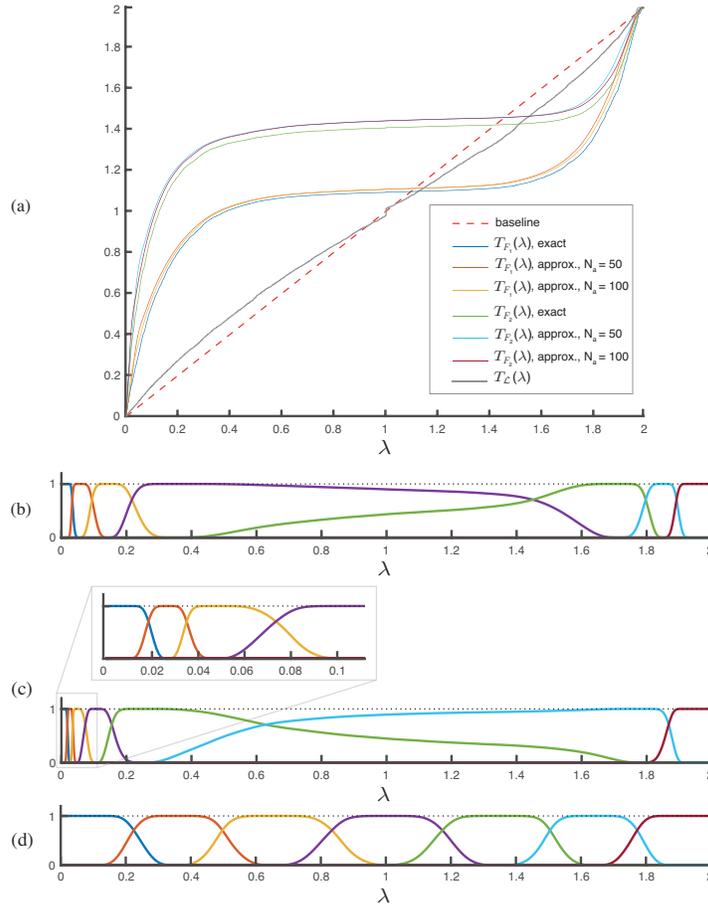


Fig. 1.6 (a) Constructed energy-equalizing transformation functions, $T_{F_1}(\lambda)$ and $T_{F_2}(\lambda)$ using the exact and approximation schemes. N_a denotes the number of spectral kernels used for the approximation, cf. (1.41). (b)-(c) Signal-adapted system of spectral kernels constructed by warping the UMT system of spectral kernels ($J = 7$) using $T_{F_1}(\lambda)$ (approx., $N_a = 100$) and $T_{F_2}(\lambda)$ (approx., $N_a = 100$), respectively. (d) Spectrum-adapted system of spectral kernels constructed by warping the UMT system of spectral kernels ($J = 6$) using $T_{\mathcal{L}}(\lambda)$. In (b)-(d), the dashed lines corresponds to the function $G(\lambda)$ in (1.19).

Such adaptation results in a system of spectral kernels that largely deviate from the UMT prototype.

1.5.1.1 Robustness to Noise

It is interesting to study the robustness of the design to possible additive noise. Let F_{1,σ_e} denote the noise added version of signal set F_1 computed as

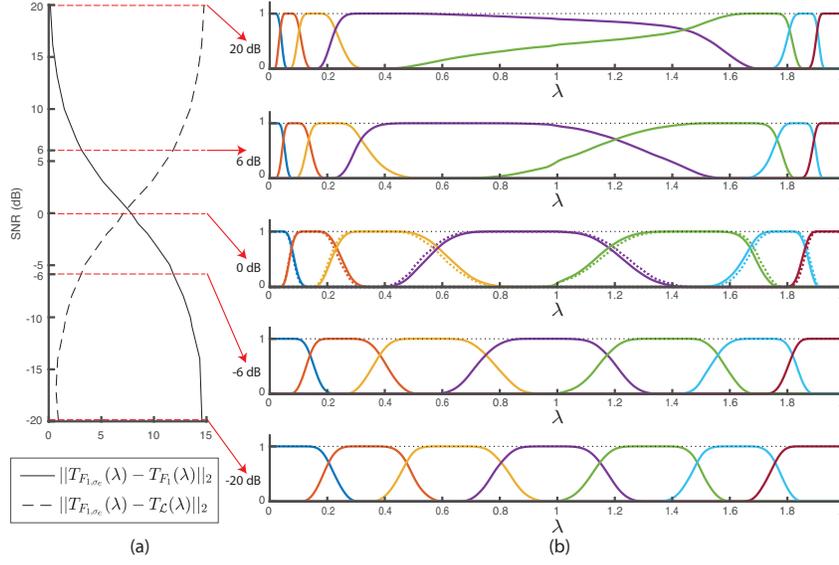


Fig. 1.7 (a) Deviation of energy-equalizing transformation functions of noise added signal sets $T_{F_1, \sigma_e}(\lambda)$ relative to $T_{\mathcal{L}}(\lambda)$ and $T_{F_1}(\lambda)$ (cf. Fig. 1.6(a)) as a function of the signal sets' SNRs. (b) Signal-adapted system of spectral kernels constructed by warping the UMT system of spectral kernels ($J = 7$) using $T_{F_1, \sigma_e}(\lambda)$ of noise-added signal sets at five different SNRs. At 0 dB, the resulting system of kernels are overlaid on the system of kernels obtained by warping the UMT system of spectral kernels using the transformation function $(T_{F_1}(\lambda) + T_{\mathcal{L}}(\lambda))/2$, shown in dashed lines.

$$F_{1, \sigma_e} = \{\mathbf{y}_i = \mathbf{x}_i + \mathbf{e}_i \mid \mathbf{x}_i \in F_1\}_{i=1, \dots, 20}, \quad (1.63)$$

where $\{\mathbf{e}_i\}_{i=1}^{20}$ denote random realizations of additive white Gaussian noise of standard deviation σ_e . We construct signal sets F_{1, σ_e} of varying $\text{SNR} = \sigma_x^2 / \sigma_e^2$, where σ_x denotes the standard deviation of each signal $\mathbf{x}_i \in F_1$. Let $T_{F_1, \sigma_e}(\lambda)$ denote the energy-equalizing transformation function associated to F_{1, σ_e} . Fig. 1.7(a) shows mean-square error metrics $\|T_{F_1, \sigma_e}(\lambda) - T_{F_1}(\lambda)\|_2$ and $\|T_{F_1, \sigma_e}(\lambda) - T_{\mathcal{L}}(\lambda)\|_2$ across signal sets F_{1, σ_e} of varying SNR, where $T_{\mathcal{L}}(\lambda)$ and $T_{F_1}(\lambda)$ are the transformation functions shown in Fig. 1.6(a), $T_{F_1}(\lambda)$ being the approximated version using $N_a = 100$. The estimated energy-equalizing transformation functions $T_{F_1, \sigma_e}(\lambda)$ become more similar to $T_{F_1}(\lambda)$ as the SNR increases. At low SNRs, $T_{F_1, \sigma_e}(\lambda)$ become more similar to $T_{\mathcal{L}}(\lambda)$. The signal-adapted system of spectral kernels using noise-added signal sets of five different SNRs are shown in Fig. 1.7(b). At the two extremes, i.e., +20 dB and -20dB, the system of kernels become almost identical to the system of kernels shown in Figs. 1.6(b) and (d), respectively. At 0dB, the signal-adapted system of kernels at each subband can be seen as the average of the corresponding kernels in the associated subbands in Figs. 1.6(b) and (d). Equivalently, this can be seen as constructing a system of kernels through warping the

the UMT prototype system of kernels with a warping function defined as the average of the spectrum-equalizing and energy-equalizing transformation functions, i.e., $(T_{F_1}(\lambda) + T_{\mathcal{L}}(\lambda))/2$, see Fig. 1.7(b) at 0 dB.

1.5.2 The Alameda Graph

The Alameda Graph is constructed based on Caltrans Performance Measurement System database¹, see Fig. 1.8(a). The vertices of the graph represent detector stations where bottlenecks were identified over the period January 2011 and December 2015. A bottleneck is a location where there is a persistent drop in speed, such as merges, large on-ramps and incidents. Two stations are considered as connected through an edge if either 1) they are adjacent along a freeway, or 2) there is a connection near the two stations at crossings between freeways. The latter type of edges were defined based on Google Maps' satellite images of Alameda county.

We use (1.62) to simulate a synthetic graph signal set as

$$F_s = \left\{ \mathbf{x}_{0.8,3}^{[i]} \right\}_{i=1,\dots,20},$$

where index i denotes random realizations of p_η in (1.62), resulting in a set of 20 signals. As real data, we treat the average duration of bottlenecks for each specific month and shift (AM shift: 5am-10am, noon shift: 10am-3pm, and PM shift: 3pm-8pm) as a graph signal, resulting in 180 signals in total. We denote this dataset as F_r .

The spectral characteristics of both F_s and F_r deviate considerably from that of the Minnesota Road graph. The distribution of the ensemble energy spectral density

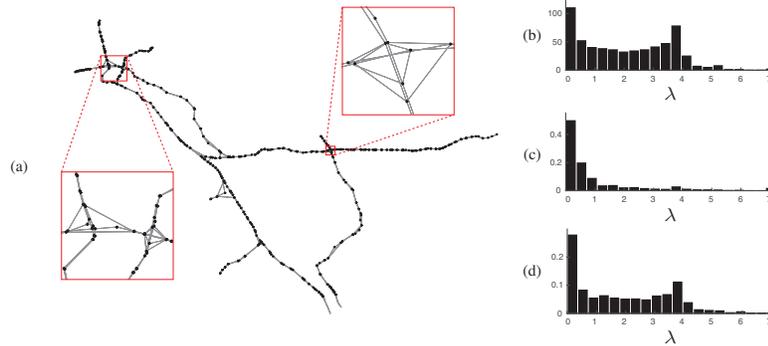


Fig. 1.8 (a) The Alameda graph. (b) Histogram of the eigenvalues $\Lambda_L(G)$ of the Alameda graph. (c)-(d) Distribution of the ensemble energy spectral density e_F of the simulated dataset F_s and the real traffic dataset F_r , respectively.

¹ The data are publicly available at <http://pems.dot.ca.gov>.

of F_s emulates an exponential distribution. Comparing the histogram of the eigenvalues $\Lambda_L(G)$ in Fig. 1.8(b) and the distribution of the ensemble energy spectral density of F_r in Fig. 1.8(c) shows that the ensemble energy is almost uniformly spread across the spectrum.

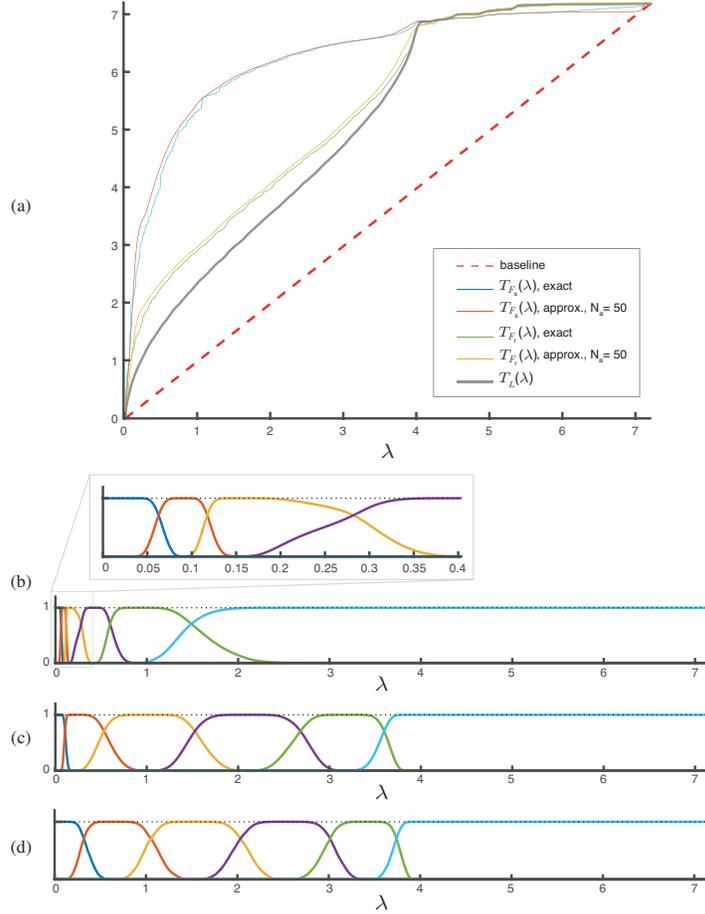


Fig. 1.9 (a) Energy-equalizing and spectrum-equalizing transformation functions. (b)-(c) Signal-adapted system of spectral kernels constructed by warping the UMT system of spectral kernels ($J = 6$) using $T_{F_s}(\lambda)$ (approx., $N_a = 50$) and $T_{F_r}(\lambda)$ (approx., $N_a = 50$), respectively. (d) Spectrum-adapted system of spectral kernels constructed by warping the UMT system of spectral kernels ($J = 6$) using $T_L(\lambda)$. In (b)-(d), the dashed lines corresponds to the function $G(\lambda)$ in (1.19).

Fig. 1.9(a) shows the energy-equalizing transformation functions associated to F_s and F_r . Also, a spectrum-equalizing transformation $T_L(\lambda)$ function is displayed. $T_L(\lambda)$ is constructed such that the distribution of eigenvalues is equalized [43]. Due to the similarity of the distributions of ensemble energy of Fig. 1.8(b) (see) and

the distribution of eigenvalues (see Fig. 1.8(a)), $T_{F_r}(\lambda)$ closely resembles $T_L(\lambda)$. Fig. 1.9(b) shows the signal-adapted system of spectral kernels associated to F_s . The majority of the spectral kernels are realized in the lower end of the spectrum where the majority of the ensemble energy is present. The zoomed-in inset in Fig. 1.9(b) show the benefit of the signal-adapted scheme in allocating a large number of spectral kernels to a narrow band of the spectrum, and yet result in smooth kernels. Figs. 1.9(c) and (d) show the signal-adapted system of spectral kernels associated to F_r and the spectrum-adapted system of spectral kernels, respectively. The similarity between $T_{F_r}(\lambda)$ and $T_L(\lambda)$, leads to the resulting signal-adapted and spectrum-adapted systems of kernels having a similar distribution of kernels across the spectrum, with more kernels allocated to the lower half of the spectrum and vice versa. This example demonstrates where the signal-adapted frame design coincides with the spectrum-adapted frame design [43] coincide in terms of their respective approach to adaptivity: if the ensemble spectral energy is equally spread across the eigenvalues, the energy-equalizing and spectrum-equalizing transformation functions become almost identical. Thus, although the signal-adapted design approach is developed based on spectral energy characteristics of a signal set, it is inherently also adapted to the graph's spectrum.

1.5.3 The Cerebellum Gray Matter Graph

Functional magnetic resonance imaging (fMRI) is a conventional neuroimaging technique used in the study of brain functionality. Its principle is in detecting a contrast that arises as a result of increased blood flow to activated regions of the brain, the so called blood-oxygen-level-dependent (BOLD) signal. Acquired fMRI data are generally corrupted with an extensive amount of noise mainly due to the fast acquisition rate; high temporal resolution is necessary to enable correlate brain activity with the experimental paradigm. The BOLD signal is not detectable across the entire brain tissue. Rather, the signal is only expected within the brain's gray matter [27]. The gray matter is convoluted layer interleaved with the brain's white matter tissue as well as the cerebrospinal fluid. As such, the BOLD signal exhibits spatial patterns that are not well suited to be characterized within a Euclidean setting. In the classical Euclidean setting, filters and wavelets used in image processing are isotropic in structure and quasi shift-invariant. The latter property infers that their structure does not vary when applied to different regions within an image/volume.

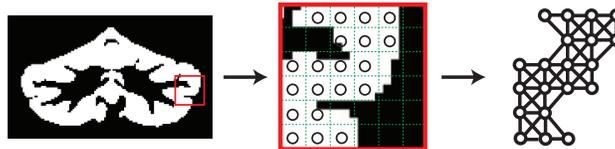


Fig. 1.10 Illustration of the cerebellum graph.

Such filters are thus not well suited for detecting the BOLD signal, with its aforementioned spatial characteristics. At the spatial resolution of fMRI, isotropically shaped basis functions will cross boundaries of gray matter, even at the finest scale. Thus, it is advantageous to construct filters that adapt to this intricately convoluted domain rather than to assume that the spatial characteristics of the underlying signal is independent of its location. To date, various approaches have been proposed to address this concern (see for example, [25, 9, 19, 33]). In particular, the construction of anatomically-adapted graph wavelets was recently proposed [2]. Yet, the deficiency of a fixed graph frame design and the lack of a systematic approach in determining the spectral coverage of spectral bands for analyzing fMRI data have been pointed out in [2, 1, 3]. These findings motivated the need for a frame design that adapts to the spectral characteristics of fMRI graph signals.

We consider a graph that encodes the 3-D topology of the cerebellar gray matter [2], which is constructed based on an atlas template of the cerebellum [12]. The graph vertices represent gray matter voxels within the cerebellum. The graph edges are defined by determining the adjacency of the gray matter voxels within their $3 \times 3 \times 3$ voxel neighbourhood, see Fig. 1.10. The fMRI data were acquired from 26 healthy subjects performing an event-related visual stimulation task [24].² For each subject, a structural MRI scan of the brain anatomy and a series of functional volumes were acquired. The structural and functional volumes were registered together and mapped to the same spatial resolution, leading to a one-to-one correspondence between functional and structural voxels. Functional voxels associated to cerebellar gray matter were then extracted and treated as cerebellar graph signals. A signal set was constructed for each subject, $\{F_k\}_{k=1}^{26}$, by randomly selecting 20 signals from each subject's functional signal set. A signal set including the signals from all subjects was also constructed as $F = F_1 \cup F_2 \cup \dots \cup F_{26}$.

Fig. 1.11(a) shows the distribution of the eigenvalues $\Lambda_{\mathcal{L}}(G)$ of the cerebellum gray matter graph. The distribution of the ensemble energy spectral density of signals sets F_1 , F_2 and F are shown in Figs. 1.11(b), (c) and (d), respectively. The distribution of eigenvalues is significantly different from that of the ensemble energy spectral densities; most eigenvalues are located at the upper end of the spectrum, whereas the ensemble energy is significantly concentrated at the lower end of the spectrum. The ensemble energy spectral densities also vary across the signal sets. Signal set F_1 has more low energy spectral content than F_2 (compare the height of the first bins of the histograms in Figs. 1.11(b) and (c)), whereas F_2 show greater spectral content at higher harmonics. F_1 and F_2 represent the two extremes in spectral content distribution among the 26 subjects. The distribution of the ensemble energy content of F falls in between that of F_1 and F_2 , see Figs. 1.11(d). This is better observed by comparing the energy-equalizing transformation functions, see Fig. 1.11(e). The transformations associated to $\{F_k\}_{k=3}^{26}$ span the space in between $T_{F_1}(\lambda)$ and $T_{F_2}(\lambda)$, and $T_F(\lambda)$ falls almost in the mid range. Moreover, the significant difference between the distribution of the eigenvalues and that of the ensemble signal energies is reflected as a major discrepancy between $T_{\mathcal{L}}(\lambda)$ and the

² The data are publicly available at <https://openfmri.org/dataset/ds000102>.

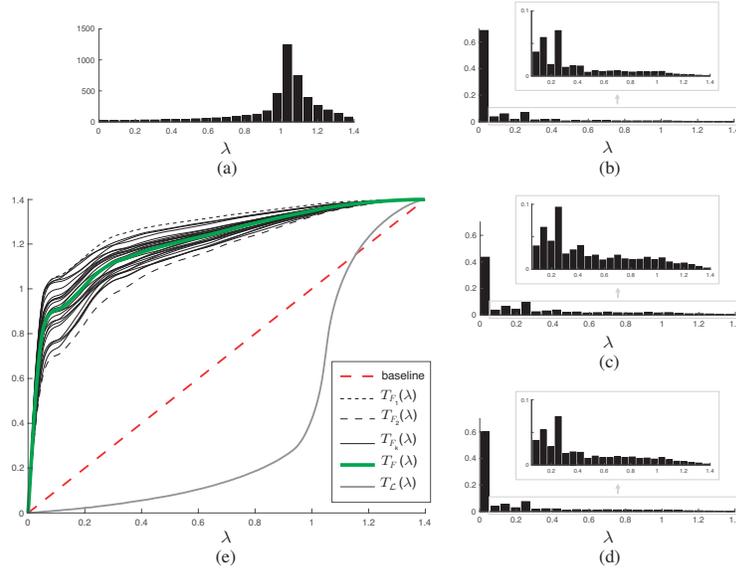


Fig. 1.11 (a) Histogram of the eigenvalues $\Lambda_{\mathcal{L}}(G)$ of the cerebellum graph. (b)-(d) Distribution of the ensemble energy spectral density of F_1 , F_2 and F . (c) Energy-equalizing and spectrum-equalizing transformation functions. The black curves correspond to the energy-equalizing transformation for each subject's signal set. The upper and lower extreme transformations represented with dashed curves are associated to signal sets F_1 and F_2 , respectively.

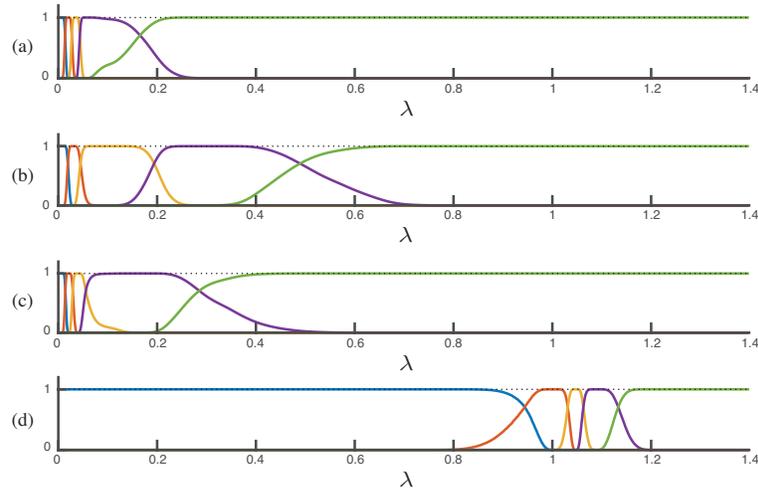


Fig. 1.12 (a)-(c) Signal-adapted system of spectral kernels adapted to the ensemble spectral content of F_1 , F_2 and F , respectively. (d) Spectrum-adapted system of spectral kernels.

energy equalizing transformations. Fig. 1.12 shows the resulting signal-adapted and spectrum-adapted systems of spectral kernels.

The kernels of the spectrum-adapted frame are localized at the higher end of the spectrum where a significant proportion of the eigenvalues fall. In contrast, kernels of the signal-adapted frames are localized at the lower end of the spectrum. This shows that the signal-adapting scheme leads to an optimal configuration of filters relative to the given ensemble energy content. Laplacian eigenmodes corresponding to large eigenvalues tend to become localized and less stable (i.e., influenced by small changes to the structure of the graph). The ensemble energy will capture the consistency of the energy for each mode across signals of the class, and thus these eigenmodes will in practice aggregate in larger subbands. The narrowband configuration of the proposed signal-adapted frame at the lower end of the spectrum closely resembles the design previously adopted for analyzing cerebellar data in [2, 1], which was obtained by empirically tuning the spectral design of the Meyer-like graph wavelet frame [26].

1.6 Conclusion & Outlook

We presented a scheme for the spectral design of signal-adapted frames on graphs. The scheme exploits the ensemble energy spectral density of a given signal class to introduce adaptivity of the spectral kernels to signal content. The design only uses stationary signal information, with a flexibility to represent non-stationary features based on the width and smoothness of the bandpass characteristics. The design has been formulated on the graph Laplacian spectrum but can be readily extended to the spectrum of the graph adjacency matrix to enable signal-adapted decomposition of signals defined on directed graphs. Various potential applications can be envisioned for the proposed developments. For instance, in functional brain imaging, another major research theme where graph signal processing can be advantageous is the study of intrinsic brain activity that fully takes into account the dynamic aspects [21]. In such case, the moment-to-moment functional data can be analyzed on a graph “backbone” [23, 21]. Time-dependent functional data can then be used to constitute the ensemble energy spectral density. As alternative avenues, signal decompositions provided by the proposed signal-adapted system of kernels can also be found beneficial in applications such as graph signal compression [39] and deep neural networks learning schemes over graphs [6, 11].

Acknowledgements This chapter draws in part on material previously published in [4].

Appendix 1 - Proof of Proposition 1

The sum of squared magnitudes of B-spline based spectral kernels $\{B_j(\lambda)\}_{j=1}^J$ forms a partition of unity since

$$\begin{aligned}
 \sum_{j=1}^J |B_j(\lambda)|^2 &\stackrel{(1.38)}{=} \sum_{i=\Delta}^{J+\Delta+1} |\tilde{B}_i(\lambda)|^2 \\
 &\stackrel{(1.39)}{=} \sum_{i=\Delta}^{J+\Delta+1} \beta^{(n)} \left(\frac{\lambda_{\max}}{J-1} (\lambda - i + 1) \right) \\
 &\stackrel{i-1 \rightarrow k}{=} \sum_{k=\Delta-1}^{J+\Delta} \beta^{(n)} \left(\frac{\lambda_{\max}}{J-1} (\lambda - k) \right) \\
 &= 1.
 \end{aligned}$$

where in the last equality we use the property that integer shifted splines form a partition of unity.

Appendix 2 - Proof of Proposition 2

In order to ensure that the spectral kernels cover the full spectrum, a must be chosen such that

$$\lambda_{\max} \stackrel{(1.54c)}{=} \lambda_{\Pi} + a \stackrel{(j=J)}{=} \gamma a + (J-2)\Delta + a,$$

which using (1.55a) leads to $a = \frac{\lambda_{\max}}{J\gamma - J - \gamma + 3}$.

To prove that the UMT system of spectral kernels form a tight frame, (1.21) needs to be fulfilled. Since, for all j , the supports of $U_{j-1}(\lambda)$ and $U_{j+1}(\lambda)$ are disjoint, $G(\lambda)$ can be determined as

$$\begin{aligned}
G(\lambda) &= \sum_{j=1}^J |U_j(\lambda)|^2 \\
&\stackrel{(1.54)}{=} \begin{cases} |U_1(\lambda)|^2 \stackrel{(1.54a)}{=} 1 & \forall \lambda \in [0, a] \\ |U_1(\lambda)|^2 + |U_2(\lambda)|^2 & \forall \lambda \in]a, \gamma a] \\ |U_2(\lambda)|^2 + |U_3(\lambda)|^2 & \forall \lambda \in]\gamma a, \gamma a + \Delta] \\ \vdots & \vdots \\ |U_J(\lambda)|^2 \stackrel{(1.54c)}{=} 1 & \forall \lambda \in]\lambda_{\max} - a, \lambda_{\max}] \end{cases} \\
&\stackrel{(1.54b)}{=} \begin{cases} 1 & \forall \lambda \in [0, a] \\ \cos^2(x_I) + \sin^2(x_I) & \forall \lambda \in]a, \gamma a] \\ \cos^2(x_{II}) + \sin^2(x_{II}) & \forall \lambda \in]\gamma a, \gamma a + \Delta] \\ \vdots & \vdots \\ 1 & \forall \lambda \in]\lambda_{\max} - a, \lambda_{\max}] \end{cases} \\
&= 1 \quad \forall \lambda \in [0, \lambda_{\max}] \tag{1.64}
\end{aligned}$$

where $x_I = \frac{\pi}{2} \nu(\frac{1}{\gamma-1}(\frac{\lambda}{a} - 1))$ and $x_{II} = \frac{\pi}{2} \nu(\frac{1}{\gamma-1}(\frac{\lambda-\Delta}{a} - 1))$.

For any given γ , the constructed set of spectral kernels form a tight frame. However, in order for the frame to satisfy the uniformity constraint given in (1.51), the appropriate γ needs to be determined. From (1.54b), we have $\forall j \in \{2, \dots, J-2\}$

$$U_j(\lambda) = U_{j+1}(\lambda + \Delta) \quad \forall \lambda \in]\lambda_I, \lambda_{II} + \Delta]. \tag{1.65}$$

By considering an inverse linear mapping of the spectral support where $U_1(\lambda) \neq 0$, i.e. $[0, \gamma a]$, to the spectral support where $U_J(\lambda) \neq 0$, i.e. $[\lambda_{\max} - \gamma a, \lambda_{\max}]$, we have

$$U_1(\lambda) = U_J(-\lambda + 2a + J\Delta) \quad \forall \lambda \in [0, \gamma a]. \tag{1.66}$$

Thus, from (1.65) and (1.66) we have

$$\int_0^{\lambda_{\max}} U_j(\lambda) d\lambda = C_2, \quad j = 2, \dots, J-1 \tag{1.67a}$$

$$\int_0^{\lambda_{\max}} U_1(\lambda) d\lambda = \int_0^{\lambda_{\max}} U_J(\lambda) d\lambda = C_1, \tag{1.67b}$$

respectively, where $C_1, C_2 \in \mathbb{R}^+$. Thus, in order to satisfy (1.51), γ should be chosen such that

$$\begin{aligned}
C_1 &= C_2 \\
\int_0^{\lambda_{\max}} U_1(\lambda) d\lambda &= \int_0^{\lambda_{\max}} U_2(\lambda) d\lambda \\
a + \int_a^{\gamma a} U_1(\lambda) d\lambda &= \int_a^{\gamma a} \sin\left(\frac{\pi}{2} v\left(\frac{1}{\gamma-1}\left(\frac{\lambda}{a} - 1\right)\right)\right) d\lambda \\
&\quad + \int_{\gamma a}^{\gamma a + \Delta} U_2(\lambda) d\lambda \\
&\stackrel{(1.65)}{=} \int_a^{\gamma a} \sin\left(\frac{\pi}{2} v\left(\frac{1}{\gamma-1}\left(\frac{\lambda}{a} - 1\right)\right)\right) d\lambda. \tag{1.68}
\end{aligned}$$

The optimal γ that satisfies (1.68) was obtained numerically by defining

$$Q(\gamma) = \int_a^{\gamma a} \sin\left(\frac{\pi}{2} v\left(\frac{1}{\gamma-1}\left(\frac{\lambda}{a} - 1\right)\right)\right) d\lambda - a, \tag{1.69}$$

and discretizing $Q(\gamma)$ within the range $(a, \gamma a]$, with a sampling factor of 1×10^{-4} . Testing for $\gamma \geq 1$, with a step size of 1×10^{-2} , the optimal value, which is independent of λ_{\max} and J , was found to be $\gamma = 2.73$.

References

1. Behjat, H., Leonardi, N., Sörnmo, L., Van De Ville, D.: Canonical cerebellar graph wavelets and their application to fMRI activation mapping. In: Proc. IEEE Int. Conf. Eng. Med. Biol. Soc., pp. 1039–1042 (2014)
2. Behjat, H., Leonardi, N., Sörnmo, L., Van De Ville, D.: Anatomically-adapted graph wavelets for improved group-level fMRI activation mapping. *Neuroimage* **123**, 185–199 (2015)
3. Behjat, H., Leonardi, N., Van De Ville, D.: Statistical parametric mapping of functional MRI data using wavelets adapted to the cerebral cortex. In: Proc. IEEE Int. Symp. Biomed. Imaging, pp. 1070–1073 (2013)
4. Behjat, H., Richter, U., Van De Ville, D., Sörnmo, L.: Signal-adapted tight frames on graphs. *IEEE Trans. Signal Process.* **64**(22), 6017–6029 (2016)
5. Benedetto, J., Fickus, M.: Finite normalized tight frames. *Adv. Comput. Math.* **18**(2), 357–385 (2003)
6. Bronstein, M.M., Bruna, J., LeCun, Y., Szlam, A., Vandergheynst, P.: Geometric deep learning: Going beyond euclidean data. *IEEE Signal Process. Mag.* **34**(4), 18–42 (2017). DOI 10.1109/MSP.2017.2693418
7. Chen, S., Varma, R., Sandryhaila, A., Kovačević, J.: Discrete signal processing on graphs: sampling theory. *IEEE Trans. Signal Process.* **63**(24), 6510–6523 (2015)
8. Chung, F.: *Spectral graph theory*. AMS, Providence, RI (1997)
9. Chung, M.K., Robbins, S.M., Dalton, K.M., Davidson, R.J., Alexander, A.L., Evans, A.C.: Cortical thickness analysis in autism with heat kernel smoothing. *Neuroimage* **25**(4), 1256–1265 (2005)
10. Coifman, R.R., Maggioni, M.: Diffusion wavelets. *Appl. Comput. Harmon. Anal.* **21**(1), 53–94 (2006)
11. Defferrard, M., Bresson, X., Vandergheynst, P.: Convolutional neural networks on graphs with fast localized spectral filtering. In: *Advances in Neural Information Processing Systems*, pp. 3844–3852 (2016)

12. Diedrichsen, J., Balsters, J.H., Flavell, J., Cussans, E., Ramnani, N.: A probabilistic MR atlas of the human cerebellum. *Neuroimage* **46**(1), 39–46 (2009)
13. Dong, B.: Sparse representation on graphs by tight wavelet frames and applications. *Appl. Comput. Harmon. Anal.* (2015). Doi:10.1016/j.acha.2015.09.005
14. Fiedler, M.: Algebraic connectivity of graphs. *Czechoslov. Math. J.* **23**(2), 298–305 (1973)
15. Fritsch, F.N., Carlson, R.E.: Monotone piecewise cubic interpolation. *SIAM J. Numer. Anal.* **17**(2), 238–246 (1980)
16. Gavish, M., Nadler, B., Coifman, R.R.: Multiscale wavelets on trees, graphs and high dimensional data: Theory and applications to semi supervised learning. In: *Proc. Int. Conf. Mach. Learn.*, pp. 367–374 (2010)
17. Girvan, M., Newman, M.E.: Community structure in social and biological networks. *Proc. Natl Acad. Sci.* **99**(12), 7821–7826 (2002)
18. Göbel, F., Blanchard, G., von Luxburg, U.: Construction of tight frames on graphs and application to denoising. Aug. 2014 [online]. Available: <http://arXiv.org/abs/1408.4012>
19. Hagler Jr, D.J., Saygin, A.P., Sereno, M.I.: Smoothing and cluster thresholding for cortical surface-based group analysis of fMRI data. *Neuroimage* **33**(4), 1093–1103 (2006)
20. Hammond, D., Vandergheynst, P., Gribonval, R.: Wavelets on graphs via spectral graph theory. *Appl. Comput. Harmon. Anal.* **30**(2), 129–150 (2011)
21. Huang, W., Bolton, T.A.W., Medaglia, J.D., Bassett, D.S., Ribeiro, A., Ville, D.V.D.: A graph signal processing perspective on functional brain imaging. *Proceedings of the IEEE* **106**(5), 868–885 (2018). DOI 10.1109/JPROC.2018.2798928
22. Jansen, M., Nason, G.P., Silverman, B.W.: Multiscale methods for data on graphs and irregular multidimensional situations. *J. Roy. Statistical Soc.: Series B (Statistical Methodology)* **71**(1), 97–125 (2009)
23. Karahanoğlu, F.I., Van De Ville, D.: Dynamics of large-scale fMRI networks: Deconstruct brain activity to build better models of brain function. *Curr. Opin. Biomed. Eng.* **3**, 28–36 (2017). DOI 10.1016/j.cobme.2017.09.008
24. Kelly, A.M.C., Uddin, L.Q., Biswal, B.B., Castellanos, F.X., Milham, M.P.: Competition between functional brain networks mediates behavioral variability. *Neuroimage* **39**(1), 527–537 (2008)
25. Kiebel, S.J., Goebel, R., Friston, K.J.: Anatomically informed basis functions. *Neuroimage* **11**(6), 656–667 (2000)
26. Leonardi, N., Van De Ville, D.: Tight wavelet frames on multislice graphs. *IEEE Trans. Signal Process.* **61**(13), 3357–3367 (2013)
27. Logothetis, N., Wandell, B.: Interpreting the BOLD signal. *Annu. Rev. Physiol.* **66**, 735–769 (2004)
28. Meyer, Y.: Principe d’incertitude, bases hilbertiennes et algèbres d’opérateurs. *Seminaire Bourbaki (in French)* **662**, 209–223 (1986)
29. Narang, S.K., Ortega, A.: Lifting based wavelet transforms on graphs. In: *Proc. APSIPA ASC*, pp. pp. 441–444 (2009)
30. Narang, S.K., Ortega, A.: Perfect reconstruction two-channel wavelet filter banks for graph structured data. *IEEE Trans. Signal Process.* **60**(6), 2786–2799 (2012)
31. Newman, M.: *Networks*. OUP Oxford (2010)
32. Ortega, A., Frossard, P., Kovačević, J., Moura, J.M.F., Vandergheynst, P.: On the optimality of ideal filters for pyramid and wavelet signal approximation. *Proc. IEEE* **106**(5), 808–828 (2018)
33. Ozkaya, S., Van De Ville, D.: Anatomically adapted wavelets for integrated statistical analysis of fMRI data. In: *Proc. IEEE Int. Symp. Biomed. Imaging*, pp. 469–472. Chicago, IL (2011)
34. Ram, I., Elad, M., Cohen, I.: Generalized tree-based wavelet transform. *IEEE Trans. Signal Process.* **59**(9), 4199–4209 (2011)
35. Ram, I., Elad, M., Cohen, I.: Redundant wavelets on graphs and high dimensional data clouds. *IEEE Signal Process. Lett.* **19**(5), 291–294 (2012)
36. Rustamov, R., Guibas, L.: Wavelets on graphs via deep learning. In: *Proc. Adv. Neural Info. Proces. Syst.*, pp. 998–1006 (2013)

37. Sakiyama, A., Tanaka, Y.: Oversampled graph Laplacian matrix for graph filter banks. *IEEE Trans. Signal Process.* **62**(24), 6425–6437 (2014)
38. Sandryhaila, A., Moura, J.: Discrete signal processing on graphs. *IEEE Trans. Signal Process.* **61**, 1644–1656 (2013)
39. Sandryhaila, A., Moura, J.M.: Big data analysis with signal processing on graphs: Representation and processing of massive data sets with irregular structure. *IEEE Signal Process. Mag.* **31**(5), 80–90 (2014)
40. Shuman, D.I., Faraji, M.J., Vandergheynst, P.: A multiscale pyramid transform for graph signals. *IEEE Trans. Signal Process.* **64**(8), 2119–2134 (2016)
41. Shuman, D.I., Narang, S.K., Frossard, P., Ortega, A., Vandergheynst, P.: The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains. *IEEE Signal Process. Mag.* **30**(3), 83–98 (2013)
42. Shuman, D.I., Ricaud, B., Vandergheynst, P.: Vertex-frequency analysis on graphs. *Appl. Comput. Harmon. Anal.* (2015). Doi: 10.1016/j.acha.2015.02.005
43. Shuman, D.I., Wiesmeyr, C., Holighaus, N., Vandergheynst, P.: Spectrum-adapted tight graph wavelet and vertex-frequency frames. *IEEE Trans. Signal Process.* **63**(16), 4223–4235 (2015)
44. Stankovic, L., Dakovic, M., Sejdic, E.: Vertex-frequency analysis: A way to localize graph spectral components [lecture notes]. *IEEE Signal Process. Mag.* **34**(4), 176–182 (2017)
45. Tanaka, Y., Sakiyama, A.: M-channel oversampled graph filter banks. *IEEE Trans. Signal Process.* **62**(14), 3578–3590 (2014)
46. Tay, D.B.H., Zhang, J.: Techniques for constructing biorthogonal bipartite graph filter banks. *IEEE Trans. Signal Process.* **63**(21), 5772–5783 (2015). Doi: 10.1109/TSP.2015.2460216
47. Thanou, D., Shuman, D.I., Frossard, P.: Learning parametric dictionaries for signals on graphs. *IEEE Trans. Signal Process.* **62**(15), 3849–3862 (2014)
48. Tremblay, N., Borgnat, P.: Subgraph-based filterbanks for graph signals. *IEEE Trans. Signal Process.* **64**(15), 3827–3840 (2016)
49. Van De Ville, D., Demesmaeker, R., Preti, M.G.: When slepian meets fiedler: Putting a focus on the graph spectrum. *IEEE Signal Process. Lett.* **24**(7), 1001–1004 (2017)
50. Von Luxburg, U.: A tutorial on spectral clustering. *Stat. Comput.* **17**(4), 395–416 (2007)
51. Yankelevsky, Y., Elad, M.: Dual graph regularized dictionary learning. *IEEE Trans. Signal Info. Process. Netw.* **2**(4), 611–624 (2016)