Signal-Adapted Tight Frames on Graphs

Hamid Behjat, Ulrike Richter, Dimitri Van De Ville, *Senior Member, IEEE* and

Leif Sörnmo, *Senior Member, IEEE*

Abstract

Analysis of signals on complex topologies modeled by graphs is a topic of increasing importance. Decompositions play a crucial role in representation and processing of such information. Here we propose a new tight frame design that is adapted to a class of signals on a graph. The construction starts from a prototype Meyer-type system of kernels with uniform subbands. The ensemble energy spectral density is then defined for a given set of signals defined on the graph. The prototype design is then warped such that the resulting subbands capture the same amount of energy for the signal class. This approach accounts at the same time for graph topology and signal features. The proposed frames are constructed for three different graph signal sets and are compared to non signal-adapted frames. Vertex localization of a set of resulting atoms is studied. The frames are then used to decompose a set of real graph signals, and are also used in a setting of signal denoising. The results illustrate the superiority of the designed signal-adapted frames, over frames blind to signal characteristics, in representing data and in denoising.

Index Terms

Spectral graph theory, filter design, tight frames, signal processing on graphs.

I. INTRODUCTION

Graphs provide a flexible framework for representing data that lie on topologically complex domains. Thus, much attention has been given to generalizing fundamental signal processing operations to the graph setting [1]–[3]. In particular, many proposals relate to extending multi-resolution transforms, filter bank designs and dictionary constructions for signals on graphs [4]–[27]. These studies fall essentially within two regimes: spatial (vertex) and spectral (frequency) designs. Constructions that fall within the former regime include methods based on lifting schemes [4]–[6] and methods in designing wavelets for hierarchical trees [7]–[9]. The latter regime can be viewed as defining dictionaries comprised of atoms that are constructed by translating smooth graph spectral filters to different

H. Behjat and L. Sörnmo are with the Department of Biomedical Engineering, Lund University, Sweden (e-mail: hamid.behjat@bme.lth.se; leif.sornmo@bme.lth.se).

U. Richter is with the Department of Experimental Medical Science, Lund University, Sweden (email: ulrike.richter@med.lu.se).

D. Van De Ville is with the Department of Radiology and Medical Informatics, University of Geneva, and the École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland (email: dimitri.vandeville@epfl.ch).

vertices of the graph. The graph spectrum is defined as the eigenspace of a graph Laplacian matrix. One of the first proposals of such a dictionary is the spectral graph wavelet transform (SGWT) frame [10] that is constructed based on a system of spline-based spectral kernels, including a lowpass kernel and a sequence of dilated bandpass kernels. Constructions of systems of spectral graph kernels leading to tight frames were proposed in [11]–[13]. Tight frames are particularly interesting because of their property of energy conservation between the original and transformed domain [28]. Other approaches to spectral domain design include diffusion wavelets [14], vertex-frequency frames [23] and approaches to graph filter-bank design using bipartite graph decompositions [15]–[20], graph coloring [25] and connected sub-graph decomposition [21].

The aforementioned spectral designs are blind to the fact that the eigenvalues of the Laplacian matrix that define the graph spectrum are typically irregularly spaced, unlike the definition of discrete frequencies for regular signals. As a step towards adaptation to the spectral properties of the graph domain, the construction of spectrum-adapted tight graph wavelet and vertex-frequency frames was recently proposed in [24]. The spectrum-adapted spectral kernels are adapted to the distribution of eigenvalues of the graph Laplacian matrix such that a similar number of eigenvalues lies in the support of each spectral kernel.

Besides considering the structural characteristics of the graph, a major improvement to the frame design would be to also consider the properties of the signals realized on the graph. To this aim, Thanou et al. [26], [27] have pursued a structured, numerical dictionary learning approach in which wavelet dictionaries are learnt based on a set of training signals. Since the graph structure is incorporated into the learning process, the learned kernels are indirectly adapted to the graph Laplacian spectrum as well as to the training data. This approach is effective in providing a sparse representation of graph signals that can be described as combinations of overlapping local patterns. In a more empirical approach [29]–[31], the Meyer-like frame design [11] has been tailored to fMRI signals by defining a spectral partitioning such that a small number of narrow-support filters, covering the lower end of the spectrum, are constructed.

In this paper, 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 [24], 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.

The paper is organized as follows. In Section II, definitions related to spectral graph theory, tight frames and graph signal decomposition are briefly reviewed. In Section III, the procedure for constructing signal-adapted tight frames is introduced. In Section IV, insight on vertex localization of the resulting atoms is provided. In Section V, constructions of the proposed frame for three different graph signal sets are presented, vertex localization of a set of resulting atoms is quantified, the coefficients resulting from decomposing a real graph dataset are studied, and finally, the proposed frames are used in the setting of signal denoising.

II. PRELIMINARIES

A. Graphs and Spectral Graph Theory

An undirected, weighted graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$ consists of a set $\mathcal{V}$ of N_g vertices, a set $\mathcal{E}$ of edges (i.e., pairs (i, j) where $i, j \in \mathcal{V}$) and a weighted adjacency matrix $A = [a_{i,j}]$, where $a_{i,j}$ denotes the weight of the edge between vertices i and j. The degree matrix D is diagonal with elements $d_{i,i} = \sum_j a_{i,j}$. The Laplacian matrices of $\mathcal{G}$ in combinatorial form L and normalized form $\mathcal{L}$ are defined as

$$L = D - A, \tag{1}$$

$$\mathcal{L} = D^{-1/2} L D^{-1/2}, \tag{2}$$

respectively. Since both L and $\mathcal{L}$ are symmetric and positive semi-definite, their eigendecompositions lead to a set of N_g real, non-negative eigenvalues that define the graph spectrum

$$\Lambda(\mathcal{G}) = \{0 = \lambda_1 \leq \lambda_2 \cdots \leq \lambda_{N_g} = \lambda_{\max}\}. \tag{3}$$

The corresponding set of eigenvectors $\{\chi_l\}_{l=1}^{N_g}$ forms a complete set of orthonormal vectors that span the graph spectral domain [32]. When necessary, we use the notations $\Lambda_L(\mathcal{G})$ and $\Lambda_\mathcal{L}(\mathcal{G})$ 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, $m_{\lambda_1} = 1$.

B. Graph Signals

Let $\mathbb{R}(\mathcal{V})$ denote a Hilbert space of real-valued signals defined on the vertices of the graph, with the inner product defined as $\forall \boldsymbol{f}_1, \boldsymbol{f}_2 \in \mathbb{R}(\mathcal{V}), \langle \boldsymbol{f}_1, \boldsymbol{f}_2 \rangle = \sum_{n=1}^{N_g} \boldsymbol{f}_1[n] \boldsymbol{f}_2[n]$, and the norm as $\forall \boldsymbol{f} \in \mathbb{R}(\mathcal{V}), \|\boldsymbol{f}\|_2^2 = \langle \boldsymbol{f}, \boldsymbol{f} \rangle = \sum_{n=1}^{N_g} |\boldsymbol{f}[n]|^2$.

Graph Fourier Transform: For any $\boldsymbol{f} \in \mathbb{R}(\mathcal{V})$, its spectral representation $\widehat{\boldsymbol{f}} \in \mathbb{R}(\mathcal{V})$, known as the graph Fourier transform of $\boldsymbol{f}$, can be used to express $\boldsymbol{f}$ in terms of the graph Laplacian eigenvectors

$$\boldsymbol{f}[n] = \sum_{l=1}^{N_g} \underbrace{\langle \boldsymbol{f}, \chi_l \rangle}_{= \widehat{\boldsymbol{f}}[l]} \chi_l[n]. \tag{4}$$

Parseval Relation: With this definition of the Fourier transform, it can be shown that the Parseval relation holds [23]

$$\forall \boldsymbol{f}_1, \boldsymbol{f}_2 \in \mathbb{R}(\mathcal{V}), \quad \langle \boldsymbol{f}_1, \boldsymbol{f}_2 \rangle = \langle \widehat{\boldsymbol{f}}_1, \widehat{\boldsymbol{f}}_2 \rangle. \tag{5}$$

Generalized Convolution Product: For any two graph signals $\boldsymbol{f}_1, \boldsymbol{f}_2 \in \mathbb{R}(\mathcal{V})$, the generalized convolution product is defined as

$$(\boldsymbol{f}_1 * \boldsymbol{f}_2)[n] = \sum_{l=1}^{N_g} \widehat{\boldsymbol{f}}_1[l] \widehat{\boldsymbol{f}}_2[l] \chi_l[n]. \tag{6}$$

Graph Signal Filtering: 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 $\boldsymbol{f} \in \mathbb{R}(\mathcal{V})$ and graph filter $\boldsymbol{g} \in \mathbb{R}(\mathcal{V})$, defined through its Fourier transform $\widehat{\boldsymbol{g}}$, the filtered signal, denoted by $(F_{\boldsymbol{g}}\boldsymbol{f})$, can be obtained as

$$(F_{\boldsymbol{g}}\boldsymbol{f})[n] = (\boldsymbol{g} * \boldsymbol{f})[n] \tag{7}$$

$$\stackrel{(6)}{=} \sum_{l=1}^{N_g} \widehat{\boldsymbol{g}}[l]\widehat{\boldsymbol{f}}[l]\boldsymbol{\chi}_l[n]. \tag{8}$$

The filter response of an impulse at vertex m

$$\boldsymbol{f} = \boldsymbol{\delta}_m \leftrightarrow \widehat{\boldsymbol{\delta}}_m[l] = \langle \boldsymbol{\delta}_m, \boldsymbol{\chi}_l \rangle = \boldsymbol{\chi}_l[m], \tag{9}$$

can then be obtained as

$$(F_{\boldsymbol{g}}\boldsymbol{\delta}_m)[n] = \sum_{l=1}^{N_g} \widehat{\boldsymbol{g}}[l]\boldsymbol{\chi}_l[m]\boldsymbol{\chi}_l[n], \tag{10}$$

which, in general, is shift-variant; i.e, in the vertex domain, it is not a shifted version of the same graph signal. Therefore, for convenience, a graph filter can be defined by its spectral kernel $\widehat{\boldsymbol{g}}$.

To design spectral kernels, it is often more elegant to define an underlying smooth continuous kernel. For instance, we consider $\mathcal{K}(\lambda) : [0, \lambda_{\max}] \to \mathbb{R}^+$, from which we derive a discrete version through sampling as

$$\boldsymbol{k}[l] = \mathcal{K}(\lambda_l), \quad l = 1, \ldots, N_g. \tag{11}$$

Note that although spectral kernels (i.e., $\mathcal{K}(\lambda)$ and $\boldsymbol{k}$) are defined in the spectral domain, they are not linked to any explicit vertex representation, and thus, the Fourier symbol $\widehat{}$ is not used for their denotation. This notation convention will be used throughout the paper.

C. Dictionary of Graph Atoms

As noted before, for a given kernel $\boldsymbol{k}$ associated with $\mathcal{K}$, the vertex-domain impulse responses are obtained as

$$\boldsymbol{\psi}_{\mathcal{K},m} = (F_{\boldsymbol{k}}\boldsymbol{\delta}_m) \leftrightarrow \widehat{\boldsymbol{\psi}}_{\mathcal{K},m}[l] = \boldsymbol{k}[l]\boldsymbol{\chi}_l[m]. \tag{12}$$

The collection of impulse responses $\{\boldsymbol{\psi}_{\mathcal{K},m}\}_{m=1}^{N_g}$ are considered graph atoms associated with the spectral kernel $\mathcal{K}(\lambda)$. Given a set of J spectral kernels $\{\boldsymbol{k}_j\}_{j=1}^{J} \in \mathbb{R}(\mathcal{V})$, a dictionary $\mathcal{D}_{\mathcal{G}}$ containing JN_g atoms is obtained as

$$\mathcal{D}_{\mathcal{G}} = \left\{ \{\boldsymbol{\psi}_{\mathcal{K}_j,m}\}_{j=1}^{J} \right\}_{m=1}^{N_g}. \tag{13}$$

The atoms of $\mathcal{D}_{\mathcal{G}}$ form a frame in $\ell_2(\mathcal{V})$ if there exist bounds $B_2 \geq B_1 > 0$ such that [28]

$$\forall \boldsymbol{f} \in \mathbb{R}(\mathcal{V}), \quad B_1||\boldsymbol{f}||_2^2 \leq \sum_{j,m} |\langle \boldsymbol{f}, \boldsymbol{\psi}_{\mathcal{K}_j,m} \rangle|^2 \leq B_2||\boldsymbol{f}||_2^2, \tag{14}$$

where the frame bounds are given by

$$B_1 = \min_{\lambda \in [0,\lambda_{\max}]} G(\lambda), \quad B_2 = \max_{\lambda \in [0,\lambda_{\max}]} G(\lambda), \tag{15}$$

and $G(\lambda)$ is defined as

$$G(\lambda) = \sum_{j=1}^{J} |\mathcal{K}_j(\lambda)|^2. \tag{16}$$

In particular, $\mathcal{D}_\mathcal{G}$ forms a tight frame if

$$\forall \lambda \in [0, \lambda_{\max}], \quad G(\lambda) = C, \tag{17}$$

and a Parseval frame if $C = 1$.

D. Decomposition of Graph Signals

Direct Decomposition: To decompose a graph signal $\boldsymbol{f}$ onto a set of the atoms in $\mathcal{D}_\mathcal{G}$, the coefficients can be obtained as

$$c_{\mathcal{K}_j, m} = \langle \boldsymbol{f}, \boldsymbol{\psi}_{\mathcal{K}_j, m} \rangle \tag{18}$$

$$\stackrel{(5)}{=} \sum_{l=1}^{N_g} \widehat{\boldsymbol{\psi}}_{\mathcal{K}_j, m}[l] \widehat{\boldsymbol{f}}[l], \tag{19}$$

$$\stackrel{(12)}{=} \sum_{l=1}^{N_g} \boldsymbol{k}_j[l] \widehat{\boldsymbol{f}}[l] \boldsymbol{\chi}_l[m]. \tag{20}$$

If $\mathcal{D}_\mathcal{G}$ forms a Parseval frame, the original signal can be recovered as

$$\begin{aligned}
\boldsymbol{f}[n] &= \sum_j \sum_m c_{\mathcal{K}_j, m} \boldsymbol{\psi}_{\mathcal{K}_j, m} \\
&= \sum_j \sum_m \sum_l \boldsymbol{k}_j[l] \widehat{\boldsymbol{f}}[l] \boldsymbol{\chi}_l[m] \sum_{l'} \boldsymbol{k}_j[l'] \boldsymbol{\chi}_{l'}[m] \boldsymbol{\chi}_{l'}[n] \\
&= \sum_l \sum_{l'} \sum_j \boldsymbol{k}_j[l] \boldsymbol{k}_j[l'] \widehat{\boldsymbol{f}}[l] \boldsymbol{\chi}_{l'}[n] \underbrace{\sum_m \boldsymbol{\chi}_l[m] \boldsymbol{\chi}_{l'}[m]}_{\delta_{l-l'}} \\
&= \sum_l \underbrace{\sum_j \boldsymbol{k}_j^2[l]}_{=1} \widehat{\boldsymbol{f}}[l] \boldsymbol{\chi}_l[n]. \tag{21}
\end{aligned}$$

Decomposition Through Polynomial Approximation: The decomposition of $\boldsymbol{f}$ on $\mathcal{D}_\mathcal{G}$ leads to the coefficients $\{\boldsymbol{c}_{\mathcal{K}_j} = [c_{\mathcal{K}_j,1}, c_{\mathcal{K}_j,2}, \ldots, c_{\mathcal{K}_j, N_g}]^T\}_{j=1}^J$ that can be interpreted as filtered versions of $\boldsymbol{f}$ with different spectral kernels $\{\boldsymbol{k}_j\}_{j=1}^J$. Due to the redundancy of such a transform, it is beneficial to implement the transform using a fast algorithm, rather than using the explicit compution of the coefficients through (20). This becomes even more crucial when large graphs are considered. One such algorithm is the Chebyshev polynomial approximation method [10], which is based on considering the expansion of the continuous spectral window functions $\{\mathcal{K}_j(\lambda)\}_{j=1}^J$ with the Chebyshev polynomials $\mathcal{C}_p(x) = \cos(p \arccos(x))$ as

$$\mathcal{K}_j(x) = \frac{1}{2} d_{\mathcal{K}_j, 0} + \sum_{p=1}^{\infty} d_{\mathcal{K}_j, p} \bar{\mathcal{C}}_p(x), \tag{22}$$

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

$$d_{\mathcal{K}_j, p} = \frac{2}{\pi} \int_0^\pi \cos(p\theta) \mathcal{K}_j(b(\cos(\theta) + 1)). \tag{23}$$

By truncating (22) to M terms, $\boldsymbol{c}_{\mathcal{K}_j}$ can then be approximated as

$$\tilde{\boldsymbol{c}}_{\mathcal{K}_j} = \frac{1}{2} d_{\mathcal{K}_j, 0} \boldsymbol{f} + \sum_{p=1}^M d_{\mathcal{K}_j, p} \bar{\mathcal{C}}_p(L) \boldsymbol{f}. \tag{24}$$

We refer to [10] for further details.

III. SIGNAL-ADAPTED FRAME CONSTRUCTION

The objective is to construct a signal-adapted tight frame where the adaptivity is introduced by exploiting the spectral energy content of a given graph signal set. This approach 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 [24]. 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. To formulate the problem, we first introduce a notion of energy for a given graph and a given signal set.

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

$$\bar{\boldsymbol{e}}_{\mathcal{F}}[l] = \frac{1}{N_s} \sum_{s=1}^{N_s} \left| \left\langle \frac{\boldsymbol{f}_s}{||\boldsymbol{f}_s||_2}, \boldsymbol{\chi}_l \right\rangle \right|^2, \quad l = 1, \ldots, N_g, \tag{25}$$

where the normalization term ensures that each signal contributes equally to the ensemble energy and $\sum_l \bar{\boldsymbol{e}}_{\mathcal{F}}[l] = 1$.

Using $\bar{\boldsymbol{e}}_{\mathcal{F}}$, the desired system of spectral kernels $\{\boldsymbol{k}_j\}_{j=1}^J$ needs to be constructed such that each kernel captures an equal amount of ensemble energy, i.e.,

$$\sum_{l=2}^{N_g} \boldsymbol{k}_j[l] \bar{\boldsymbol{e}}_{\mathcal{F}}[l] = \frac{1 - \bar{\boldsymbol{e}}_{\mathcal{F}}[1]}{J}, \quad j = 1, \ldots, J \tag{26}$$

subject to the Parseval frame constraint, i.e.,

$$\sum_{j=1}^J |\boldsymbol{k}_j[l]|^2 = 1, \quad l = 1, \cdots, N_g. \tag{27}$$

The energy contribution from λ_1 is excluded from the design as it merely reflects the ensemble mean of the signal set.

We propose a two-step procedure to obtain such a design. First, a prototype system of continuous spectral kernels, $\{\mathcal{K}_j'(\lambda)\}_{j=1}^J$, is obtained such that they satisfy the following uniformity constraint

$$\exists\, C \in \mathbb{R}^+, \quad \int_0^{\lambda_{\max}} \mathcal{K}_j'(\lambda) \, d\lambda = C, \quad j = 1, \ldots, J, \tag{28}$$

subject to the tight Parseval frame constraint,

$$\sum_{j=1}^{J} |\mathcal{K}'_j(\lambda)|^2 = 1, \quad \forall \lambda \in [0, \lambda_{\max}]. \tag{29}$$

Second, an energy-equalizing transformation $T_\mathcal{F}(\lambda)$ is constructed and incorporated in $\{\mathcal{K}'_j(\lambda)\}_{j=1}^{J}$, leading to a warped version of the prototype design

$$\mathcal{K}_j(\lambda) = \mathcal{K}'_j(T_\mathcal{F}(\lambda)), \quad j = 1, \ldots, J. \tag{30}$$

Provided that $T_\mathcal{F}(\lambda)$ is designed such that it is monotonically increasing and satisfies

$$\sum_{l=a_j}^{b_j} \bar{e}_\mathcal{F}[l] = \frac{1 - \bar{e}_\mathcal{F}[1]}{J}, \quad j = 1, \ldots, J, \tag{31}$$

where

$$a_j = \underset{l \in \{1, \ldots, N_g\}}{\arg\min} \left\{ \left| T_\mathcal{F}(\lambda_l) - \frac{j-1}{J} \lambda_{\max} \right| \right\} \tag{32}$$

$$b_j = \underset{l \in \{1, \ldots, N_g\}}{\arg\min} \left\{ \left| T_\mathcal{F}(\lambda_l) - \frac{j}{J} \lambda_{\max} \right| \right\}, \tag{33}$$

the resulting $\{\mathcal{K}_j(\lambda)\}_{j=1}^{J}$ corresponds to the desired signal-adapted system of spectral kernels in the continuous domain. If a discrete representation is needed for direct decomposition as in (20), $\{\boldsymbol{k}_j\}_{j=1}^{J}$ can be obtained through sampling. In the following, these two steps are explained in detail.

A. Step 1: Prototype System of Spectral Kernels Construction

While there is no unique solution that satisfies (28), prototype systems of spectral kernels satisfying this constraint subject to (29) can be designed. In this paper, we aim at designing spectral kernels similar to those of the Meyer-like graph wavelet frame [11] since they have (i) a finite support of the bandpass type, and (ii) smooth transition bands that have an advantageous effect on localization in the vertex domain [10], [23]. The following proposition introduces the design of a uniform Meyer-type system of spectral kernels.

Proposition 1. *(uniform Meyer-type (UMT) system of spectral kernels) Using the auxiliary function of the Meyer wavelet, given by [33]*

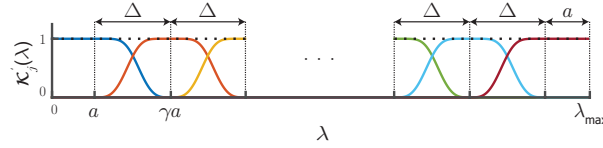$$\nu(x) = x^4(35 - 84x + 70x^2 - 20x^3), \tag{34}$$

Fig. 1. Construction of UMT system of spectral kernels.

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

$$\mathcal{K}'_1(\lambda) = \begin{cases} 1 & \forall \lambda \in [0, a] \\ \cos(\frac{\pi}{2}\nu(\frac{1}{\gamma-1}(\frac{\lambda}{a} - 1))) & \forall \lambda \in]a, \gamma a] \\ 0 & \text{elsewhere} \end{cases} \tag{35a}$$

$$\mathcal{K}'_j(\lambda) = \begin{cases} \sin(\frac{\pi}{2}\nu(\frac{1}{\gamma-1}(\frac{\lambda-(j-2)\Delta}{a} - 1))) & \forall \lambda \in]\lambda_I, \lambda_{II}] \\ \cos(\frac{\pi}{2}\nu(\frac{1}{\gamma-1}(\frac{\lambda-(j-1)\Delta}{a} - 1))) & \forall \lambda \in]\lambda_{II}, \lambda_{II} + \Delta] \\ 0 & \text{elsewhere} \end{cases} \tag{35b}$$

$$\mathcal{K}'_J(\lambda) = \begin{cases} \sin(\frac{\pi}{2}\nu(\frac{1}{\gamma-1}(\frac{\lambda-(J-2)\Delta}{a} - 1))) & \forall \lambda \in]\lambda_I, \lambda_{II}] \\ 1 & \forall \lambda \in]\lambda_{II}, \lambda_{II} + a] \\ 0 & \text{elsewhere} \end{cases} \tag{35c}$$

can be constructed, where

$$\Delta = \gamma a - a, \tag{36a}$$

$$\lambda_I = a + (j - 2)\Delta, \tag{36b}$$

$$\lambda_{II} = \gamma a + (j - 2)\Delta, \tag{36c}$$

$$a = \frac{\lambda_{max}}{J\gamma - J - \gamma + 3}. \tag{36d}$$

Fig. 1 illustrates the notations used. By setting $\gamma = 2.73$, the set of kernels defined in (35) satisfies the uniformity constraint given in (28). The atoms of a dictionary constructed using this set of spectral kernels form a Parseval frame on $\mathbb{R}(\mathcal{V})$.

Proof. see Appendix A.   □

Figs. 2(a) and (b) show realizations of the resulting UMT system of spectral kernels for a fixed λ_{max} and two different J. For comparison, the half-cosine uniform translate (HCUT) system of spectral kernels [24] are shown in Fig. 2(c) and (d). There are three main differences when comparing these two designs. First, the atoms of a dictionary constructed using the UMT system of spectral kernels not only form a tight frame but also a Parseval frame. Second, the UMT system of spectral kernels has better passband characteristics compared to HCUT system

of spectral kernels as the support of each kernel is a more strict subset of the spectrum, with less overlap of adjacent kernels. Third, the UMT system of spectral kernels satisfies (28). The latter two differences are in favour of the proposed UMT design for our purpose.

B. Step 2: Energy-Equalizing Spectral Warping

Although the UMT system of filters satisfy the uniformity constraint (28), they satisfy the equi-energy constraint (26) only if 1) the graph's eigenvalues are uniformly distributed across the spectrum and 2) the ensemble energy contribution at each eigenvalue is equal. In fact, for a given graph and signal set, these two conditions do not hold in general. Using $\bar{e}_{\mathcal{F}}$, the desired continuous energy-equalizing transformation function $T_{\mathcal{F}}(\lambda) : [0, \lambda_{\max}] \to [0, \lambda_{\max}]$, can then be obtained through monotonic cubic interpolation [34] of the pair of points

$$\left\{ \left(\lambda_l \, , \, \frac{\lambda_{\max}}{m_{\lambda_l}(1 - \bar{e}_{\mathcal{F}}[1])} \sum_{r=i_{\lambda_l}}^{i_{\lambda_l}+m_{\lambda_l}} \sum_{k=2}^{r} \bar{e}_{\mathcal{F}}[k] \right) \right\}_{l=2}^{N_g} , \tag{37}$$

together with $(\lambda_1, 0)$. The resulting $T_{\mathcal{F}}(\lambda)$ satisfies (31). Thus, the set of spectral kernels obtained by warping the UMT system of kernels using $T_{\mathcal{F}}(\lambda)$, cf. (30), correspond to the desired signal-adapted system of spectral kernels.

C. Generalization

The proposed procedure for the design of signal-adapted frames can be generalised in two aspects. First, prototype designs other than the UMT can be designed and used as the base for the proposed frames, with a general case formulation given in Appendix B. Second, rather than the proposed ensemble energy spectral density measure given in (25), other forms of stationary signal information can be exploited to construct a suitable signal-adapted spectral transformation, with example measures recently extended to the graph setting given in [35]–[37]. Finally, we also
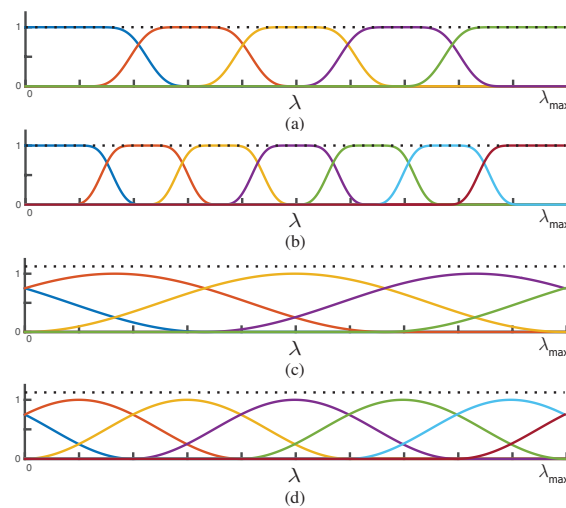

Fig. 2. UMT system of spectral kernels as given in Proposition 1 displayed for (a) $J = 5$ and (b) $J = 7$ spectral scales. The HCUT system of spectral kernels [24] displayed for (c) $J = 5$ and (d) $J = 7$ spectral scales.

note that, depending on the dataset, it can be beneficial to slightly smooth $T_{\mathcal{F}}(\lambda)$, in particular important at spectral regions where adjacent eigenvalues are minutely different in value and the ensemble energy significantly changes across them. A smooth $T_{\mathcal{F}}(\lambda)$, in turn, comes in favour of obtaining smoother spectral kernels.

IV. VERTEX LOCALIZATION OF SPECTRAL KERNELS

As in classical signal processing where the uncertainty principle determines the trade-off between fine localization in time (spatial) and frequency domain, a similar notion naturally extends to the graph setting, with example proposals given in [23], [38]–[41]. As the present proposal of constructing frames is defined in the spectral domain, it is interesting to consider the theoretical constraints in the vertex localization of the resulting atoms. In particular, we first consider the bound on the vertex localization of atoms provided through polynomial approximation [23]. For a polynomial spectral kernel of degree D, defined as

$$\mathcal{P}_D(\lambda_l) = \sum_{k=0}^{D} \alpha_k \lambda_l^k \tag{38}$$

for some coefficients $\{\alpha_k\}_{k=0}^{D}$, it can be shown that its associated atoms satisfy the localization constraint [23]

$$\forall n \in \mathcal{V}, \quad \boldsymbol{\psi}_{\mathcal{P}_D,n}[m] = 0 \quad \text{if} \quad d_{\mathcal{G}}(n,m) > D, \tag{39}$$

where $d_{\mathcal{G}}(\cdot,\cdot)$ is a distance metric. As suggested in [38], the geodesic distance is one metric that can be used, in which case, $d_{\mathcal{G}}(i,j)$ is the length of the shortest path connecting vertices i and j. For a generic spectral kernel $\mathcal{K}(\lambda)$, the localization of its associated atoms can be quantified as [23]

$$|\boldsymbol{\psi}_{\mathcal{K},n}[m]| \leq \sqrt{N_g} \inf_{\mathcal{P}_{D'}} \{||\mathcal{K} - \mathcal{P}_{D'}||_\infty\}, \tag{40}$$

where the infimum is taken over all polynomial kernels of degree D', as defined in (38), with $D' = d_{\mathcal{G}}(n,m) - 1$. An upper bound can be determined for the minimax polynomial approximation error of the infimum term. We refer to [23] for further details and a proof of (39) and (40).

Based on (39) and (40), if a given spectral kernel can be perfectly represented as a polynomial of degree M, its realization at a given vertex n will be localized in a sphere of radius M edges around vertex n; i.e., the lower the degree, the finer the localization in the vertex domain. If the approximation is not perfect, the vertices that fall outside the sphere will not necessarily be zero, but constrained by the bound in (40). This can be seen as the trade-off of vertex-spectral localization: the smoother the given spectral kernel $\mathcal{K}(\lambda)$ and the wider its transitions bands, the lower is its approximation error to a polynomial, and thus, the finer is the localization of its associated atoms in the vertex domain. Compared to the HCUT spectral kernels, the proposed UMT spectral kernels provide narrower transition bands (cf. Fig. 2), and therefore, the resulting atoms are expected to be less localized. As signal-adapted kernels result from incorporating the energy-equalizing transformation into the UMT design, cf. (30), the specific bound depends on the signal set used.

For an alternative description, the localization of realizations of graph atoms can be quantified using the graph spread measure [38]; the graph spread of a signal $\boldsymbol{f} \in \mathbb{R}(\mathcal{V})$ around vertex $k \in \mathcal{V}$ is defined as

$$\Delta_{\mathcal{G},k}^2(\boldsymbol{f}) = \frac{1}{||\boldsymbol{f}||_2^2} \sum_{v \in \mathcal{V}} d_{\mathcal{G}}^2(k,v) \boldsymbol{f}^2[v], \tag{41}$$

where $d_{\mathcal{G}}(\cdot,\cdot)$ is the same distance metric as that used in (39) and (40). The smaller is $\Delta_{\mathcal{G},k}^2(\boldsymbol{f})$, the more localized is $\boldsymbol{f}$ around vertex k, and vice versa. A global spread measure can also be obtained as

$$\Delta_{\mathcal{G}}^2(\boldsymbol{f}) = \min_{k \in \mathcal{V}} \{\Delta_{\mathcal{G},k}^2(\boldsymbol{f})\}. \tag{42}$$

By substituting an atom $\boldsymbol{\psi}_{\mathcal{K},n}$ for $\boldsymbol{f}$ in (41) and (42), its vertex localization can be quantified. In Section V-B, these two measures are used to quantify vertex localization of the different frames.

V. RESULTS

A. Signal-Adapted Tight Frame Constructions

We present constructions of the proposed frames for signal sets realized on the Minnesota road graph [42] and the Alameda graph [26]. The Minnesota road graph is chosen as it is considered as a benchmark in many recent studies, e.g., [15]–[25]. The choice of the Alameda graph is to highlight the fact that, although the proposed method is developed based on spectral energy characteristics of a signal set, it is also implicitly adapted to the graph's spectrum.

1) Data Realized on the Minnesota Road Graph: The Minnesota road graph is considered as a benchmark in many recent studies, e.g., [15]–[25]. The edges represent major roads and the vertices their intersection points, which often correspond to towns or cities. We consider a general model for realizing signals on this graph as

$$\boldsymbol{y}_{\eta,n} = \boldsymbol{x}_{\eta,n} + \boldsymbol{e}, \tag{43}$$

where $\boldsymbol{x}_{\eta,n} \in \mathbb{R}(\mathcal{V})$ denotes the graph signal of interest with density $\eta \in]0,1]$ and smoothness $n \in \mathbb{Z}^+$, and $\boldsymbol{e} \in \mathbb{R}(\mathcal{V})$ denotes additive white Gaussian noise of variance σ_e^2. In particular, $\boldsymbol{x}_{\eta,n}$ is constructed as

$$\boldsymbol{x}_{\eta,n} = A^n \boldsymbol{p}_\eta, \tag{44}$$

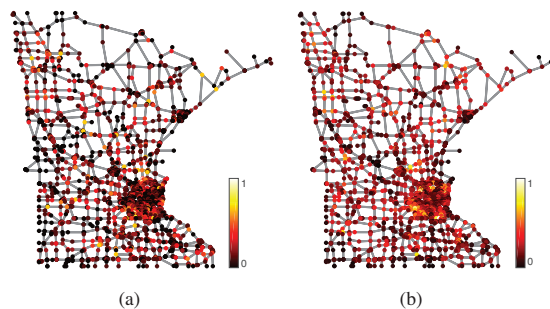(a)   (b)

Fig. 3. Sample signal realizations on the Minnesota road graph, (a) $\boldsymbol{x}_{0.2,2}$ and (b) $\boldsymbol{x}_{0.5,4}$. The plots are normalized as $\boldsymbol{x}_{\eta,n}/||\boldsymbol{x}_{\eta,n}||_\infty$.
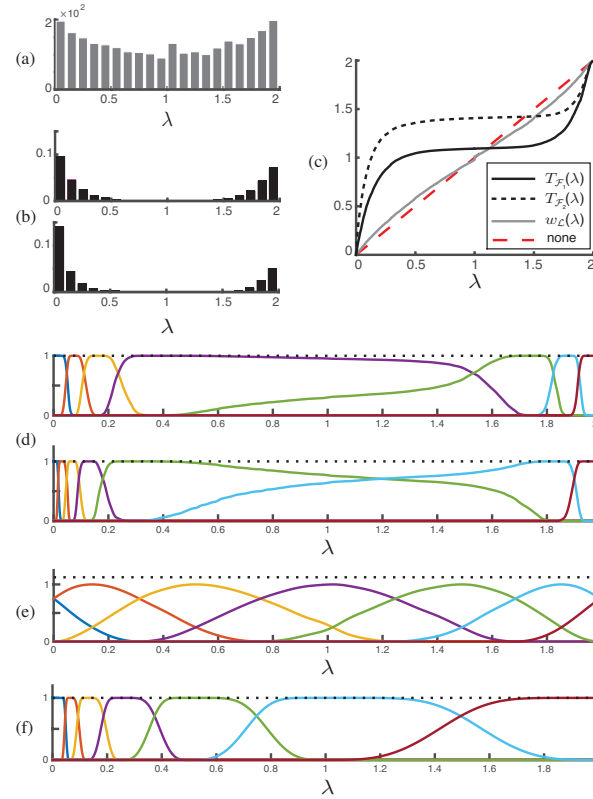
Fig. 4. (a) Histograms of the eigenvalues $\Lambda_{\mathcal{L}}(\mathcal{G})$ of the Minnesota road graph. Each bar indicates the number of eigenvalues that lie in the corresponding spectral range. (b) Distribution of the ensemble energy spectral density $\bar{e}_{\mathcal{F}_1}$ (top) and $\bar{e}_{\mathcal{F}_2}$ (bottom). Each bar indicates the sum of ensemble energies of the eigenvalues lying in the corresponding spectral range. (c) Constructed energy-equalizing transformation functions, $T_{\mathcal{F}_1}(\lambda)$ and $T_{\mathcal{F}_2}(\lambda)$, and the spectrum-adapting warping function $w_{\mathcal{L}}(\lambda)$ [24]. (d) Signal-adapted system of spectral kernels constructed by warping the UMT system of spectral kernels ($J = 7$) using $T_{\mathcal{F}_1}(\lambda)$ (top) and $T_{\mathcal{F}_2}(\lambda)$ (bottom). (e) Spectrum-adapted system of spectral kernels [24] constructed by warping the HCUT system of spectral kernels ($J = 7$) using $w_{\mathcal{L}}(\lambda)$. (f) Meyer-like wavelet frame system of spectral kernels [11]. The dashed line in (d)–(f) corresponds to the function $G(\lambda)$ in (15).

where $\boldsymbol{p}_\eta \in \mathbb{R}(\mathcal{V})$ denotes a random realization of a spike signal as $\{\boldsymbol{p}_\eta[i] \in \{0,1\}\}_{i=1,\dots,N_g}$ such that $\sum_i \boldsymbol{p}_\eta[i] = \eta N_g$, A^n incorporates the intrinsic structure of the graph into the signal, and the power n controls the extent of signal smoothness. Fig. 3 shows two signals realized on the Minnesota road graph using this scheme. By setting $\sigma_e^2 = 0$, two sets of clean graph signals with different smoothness were constructed as $\mathcal{F}_1 = \left\{ \boldsymbol{y}_{\eta,2}^{[i]} \right\}_{\eta=0.2,0.5}$ and $\mathcal{F}_2 = \left\{ \boldsymbol{y}_{\eta,4}^{[i]} \right\}_{\eta=0.2,0.5}$, where $i = 1,\dots,10$ denotes random realizations of $\boldsymbol{p}_\eta$, leading to 20 signals in each set.

While the distribution of the eigenvalues of this graph is rather uniform, see Fig. 4(a), the ensemble energy is concentrated towards the two ends of the spectrum for both $\mathcal{F}_1$ and $\mathcal{F}_2$, see Fig. 4(b). Fig. 4(c) shows the resulting energy-equalizing transformation functions. By incorporating $T_{\mathcal{F}_1}(\lambda)$ and $T_{\mathcal{F}_2}(\lambda)$ into the UMT system of spectral kernels, the desired signal-adapted systems of spectral kernels are obtained, see Fig. 4(d). A comparison of Figs. 4(d) and (b) 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 $\mathcal{F}_2$ frame than for the $\mathcal{F}_1$ frame, and vice versa at the upper end of the spectrum.

For comparison, Fig. 4(e) shows the corresponding spectrum-adapted system of spectral kernels for the Minnesota road graph, where the HCUT system of spectral kernels [24], shown in Fig. 2(d), are warped using a spectrum-adapting warping function $w_{\mathcal{L}}(\lambda)$, shown in Fig. 4(c). $w_{\mathcal{L}}(\lambda)$ is constructed such that the distribution of eigenvalues is equalized [24]. As the distribution of the eigenvalues of this graph is almost uniform, the spectrum-adapted filters almost resemble the non-warped HCUT filters. Comparing Figs. 4(d) and (e), it can be observed that the proposed method optimizes the construction of the filters 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. 4(f) shows the Meyer-like wavelet frame system of spectral kernels [11], in which the kernels have the classical dyadic dilation of the wavelet scheme. The frame is neither adapted to the distribution of the eigenvalues nor to the distribution of the ensemble energy. By comparing Figs. 4(d) and (f), it is observed that the partitioning of the kernels at the lower end of the spectrum are similar, whereas a major difference is seen at the remainder of the spectrum. The similarity is more pronounced for kernels of the signal-adapted frame associated with $\mathcal{F}_1$, due to the particular initial decay pattern observed in the ensemble energy of $\mathcal{F}_1$, see Fig. 4(b), top. The benefit of the signal-adapted frame is observed when comparing the spread of the kernels in the spectral interval $[0.2, 2]$: the signal-adapted design approach allocates more subbands to the spectral interval where signal energy is expected, whereas the wavelet frame design approach, cannot allocate more subbands to spectral intervals other than the lower end.

2) Traffic Data on the Alameda Graph: The data are part of the Caltrans Performance Measurement System dataset[1]. The monthly bottlenecks occurring across 17 freeways in Alameda county between January 2010 and December 2015 are considered. A bottleneck could be any location where there is a persistent drop in speed, such as merges, large on-ramps and incidents. The vertices of the graph represent detector stations where bottlenecks

---

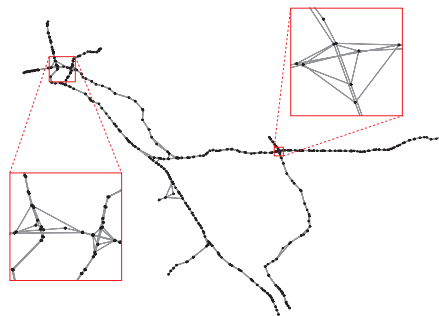[1]The data are publicly available at http://pems.dot.ca.gov.

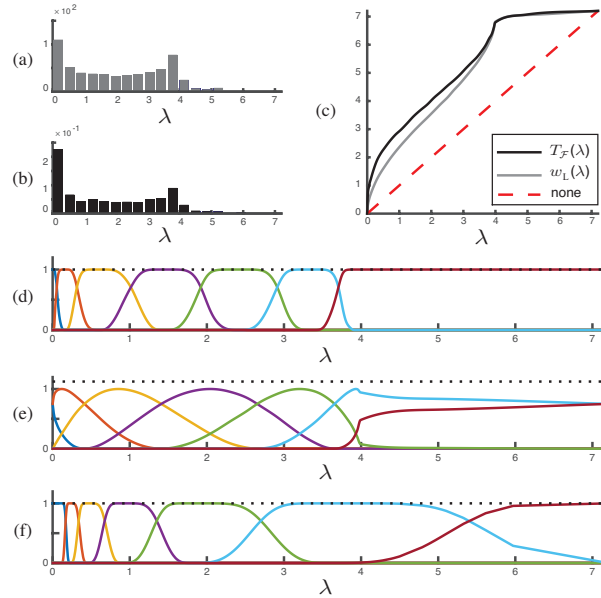Fig. 5. Illustration of the Alameda graph.

Fig. 6. (a) Histogram of the eigenvalues $\Lambda_L(\mathcal{G})$ of the Alameda graph. (b) Distribution of the ensemble energy spectral density $\bar{e}_{\mathcal{F}}$ of the traffic dataset. (c) Constructed energy-equalizingtransformation functions $T_{\mathcal{F}}(\lambda)$ and the spectrum-adapting warping function $w_L(\lambda)$ [24]. (d) Signal-adapted system of spectral kernels. (e) Spectrum-adapted system of spectral kernels [24]. (f) Meyer-like wavelet frame system of spectral kernels [11].

were identified over the mentioned period, see Fig. 5. Two stations i and j are considered as connected through an edge if either 1) they are adjacent across a freeway, or 2) there is a connection at a crossing between freeways near the two stations. The latter type of edges were defined based on satellite maps of the county available on Google Maps [43]. The signal on the graph is the average duration of bottlenecks for each specific month during three shifts (AM shift: 5am-10am, noon shift: 10am-3pm, and PM shift: 3pm-8pm), resulting in 180 signals in total.

The spectral characteristics of this signal set deviate considerably from that of the Minnesota graph. The total ensemble energy is almost uniformly spread across the eigenvalues, as observed by comparing the histogram of the eigenvalues $\Lambda_L(\mathcal{G})$ in Fig. 6(a) and the distribution of the ensemble energy spectral density in Fig. 6(b). As a result, the warping function defined for equalizing the distribution of energy across the spectrum, i.e., $T_{\mathcal{F}}(\lambda)$, closely resembles that defined for equalizing the distribution of eigenvalues; i.e., $w_L(\lambda)$, see Fig. 6(c). Consequently, almost an equal number of kernels span each part of the spectrum, with more kernels allocated to the lower half of the spectrum and vice versa, see Figs. 6(d) and (e). On the other hand, the non-adapted, Meyer-like wavelet frame has a kernel at the far end of the spectrum where there is very few eigenvalues and almost no energy, see Fig. 6(f).

Although the kernels of the proposed frame have narrower passband characteristics compared to those of the spectrum-adapted frame due to the different base set of kernels used (compare Figs. 6(d) and (e)), this example demonstrates where the proposed frame design approach and the spectrum-adapted approach coincide in terms of their respective approach to adaptivity: if the spectral energy is equally spread across the eigenvalues, the
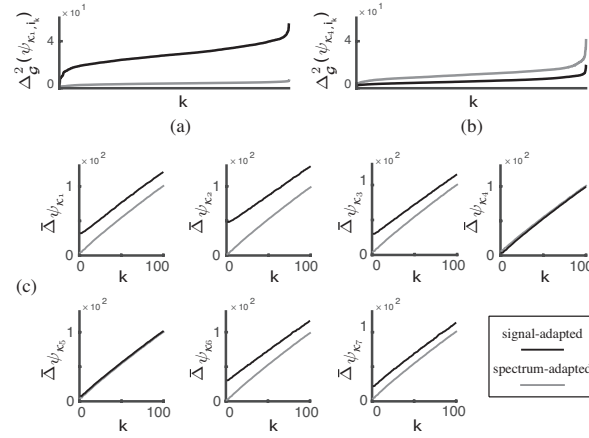
Fig. 7. Global graph spread of the atoms associated with (a) $\mathcal{K}_1(\lambda)$ and (b) $\mathcal{K}_4(\lambda)$ of the signal-adapted frame (black curves) and spectrum-adapted frame (gray curves). (c) Ensemble graph spread measure for the seven subbands of the signal-adapted and spectrum-adapted frames.

energy-equalizing transformation function $T_\mathcal{F}(\lambda)$ and the spectrum-adapting warping function $w_L(\lambda)$ become almost identical. Thus, although the proposed method is developed based on spectral energy characteristics of a signal set, it is optimal in the sense that it will indirectly adapt based on the graph's spectrum in the event of uniformly spread energy across the eigenvalues.

B. Vertex Localization

Vertex localization of atoms realized using the frames constructed on the Minnesota road graph are quantified by computing the graph spread around all graph vertices, cf. (41), for $\{\{\psi_{\mathcal{K}_j,l}\}_{j=1,\ldots,J}\}_{l=1,\ldots,N_g}$. The atoms were then sorted based on their global graph spread measure, cf. (42), leading to a set of sorted indices $\{i_k\}_{k=1}^{N_g}$, such that $\Delta_\mathcal{G}^2(\psi_{\mathcal{K}_j,i_1}) \le \Delta_\mathcal{G}^2(\psi_{\mathcal{K}_j,i_2}) \le \cdots \le \Delta_\mathcal{G}^2(\psi_{\mathcal{K}_j,i_{N_g}})$. Figs. 7(a) and (b) show the global graph spread of the atoms associated to $\mathcal{K}_1$ and $\mathcal{K}_4$, respectively, for both the signal-adapted and spectrum-adapted frames of the Minnesota road graph. For $\mathcal{K}_1$, the global graph spread is greater for the atoms of the signal-adapted frame. This can be associated with the narrower spectral spread of $\mathcal{K}_1$ of the signal adapted frame compared to that of the spectrum-adapted frame, see blue kernel in Fig. 4(d), top, and that in Fig. 4(e). On the other hand, for $\mathcal{K}_4$, the global graph spread of the atoms of the two frames become almost identical, which can be linked to the fact that $\mathcal{K}_4$ is approximately equally spread for both the signal-adapted and the spectrum-adapted frames, see purple kernel in Fig. 4(d), top, and that in Fig. 4(e).

To express the behaviour of the atoms in the vicinity of their central vertex, an ensemble measure of the change in graph spread as a function of geodesic distance from their central vertex can be obtained as

$$\bar{\Delta}_{\psi_{\mathcal{K}_j}}[k] = \frac{1}{N_g} \sum_{l=1}^{N_g} \Delta_{\mathcal{G},v_k^{(l)}}^2(\psi_{\mathcal{K}_j,l}), \quad k = 1, \ldots, N_g. \tag{45}$$

where $v_k^{(l)} \in \{1, \ldots, N_g\}$ denotes the index of the kth closest vertex to the vertex where atom $\psi_{\mathcal{K}_j,l}$ is centered at, based on the the geodesic distance measure; i.e., $d_\mathcal{G}(l,v_1^{(l)}=l)=0 < d_\mathcal{G}(l,v_2^{(l)}) \le \cdots \le d_\mathcal{G}(l,v_{N_g}^{(l)})$. Fig. 7(c)

shows $\{\bar{\Delta}_{\psi_j}\}_{j=1}^{7}$ associated with the spectral kernels of the signal-adapted and spectrum-adapted frames of the Minnesota road graph, for geodesic distances of up to 100 vertices. For all subbands, and for both frames, the ensemble graph spread monotonically increases as the geodesic distance increases. As expected, for subbands $j = \{1, 2, 3, 6, 7\}$, where the kernels of the spectrum-adapted frame are more spectrally spread compared to that of the signal-adapted frame (see Fig. 4(d), top, and Fig. 4(e)), $\bar{\Delta}_{\psi_j}[k]$ is lower, whereas for subbands $j = \{4, 5\}$ the values are almost identical since the corresponding kernels of both frames are widely spread across the spectrum. These results are in line with the theoretical insight provided by (39) and (40).

C. Efficient Decomposition of Brain fMRI Data

Functional magnetic resonance imaging (fMRI) is a key bioimaging modality used for performing non-invasive studies of the brain. The principle behind fMRI is the detection of a signal known as the blood-oxygen-level-dependent (BOLD) signal, which arises as a result of increased blood flow to activated regions of the brain. As high temporal resolution data is required to correlate brain activity with the experimental paradigm, the resulting fMRI data are generally corrupted with an extensive amount of noise, requiring denoising.

Filters and wavelets in the most classical form used in image processing share several basic properties: they are (i) defined within Euclidean spaces (a square in 2-D or a cube in 3-D), (ii) isotropic in structure and (iii) stationary and quasi shift-invariant, meaning that their structure does not vary when applied to different regions within a volume. However, the BOLD signal is expected only within the thin convoluted layer of gray matter of the brain, but not within the white matter or cerebrospinal fluid [44]. 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 this aim, many approaches have been proposed (see for example, [45]–[48]), in particular, the construction of anatomically-adapted graph wavelets [31]. 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 [29]–[31]. In fact, these findings motivated us to pursue the idea of designing the proposed signal-adapted frames.

Here, we consider the cerebellum region of the brain. The graph, as designed in [31], encodes the 3-D geometry of the cerebellar gray matter, and its construction is based on an atlas template of the cerebellum [49]. The vertices of the graph represent voxels within the cerebellar structure that correspond to gray matter, and the edges are assigned by computing connections between adjacent voxels in 3-D neighbourhood, see Fig. 8. The fMRI data
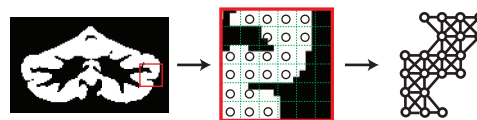

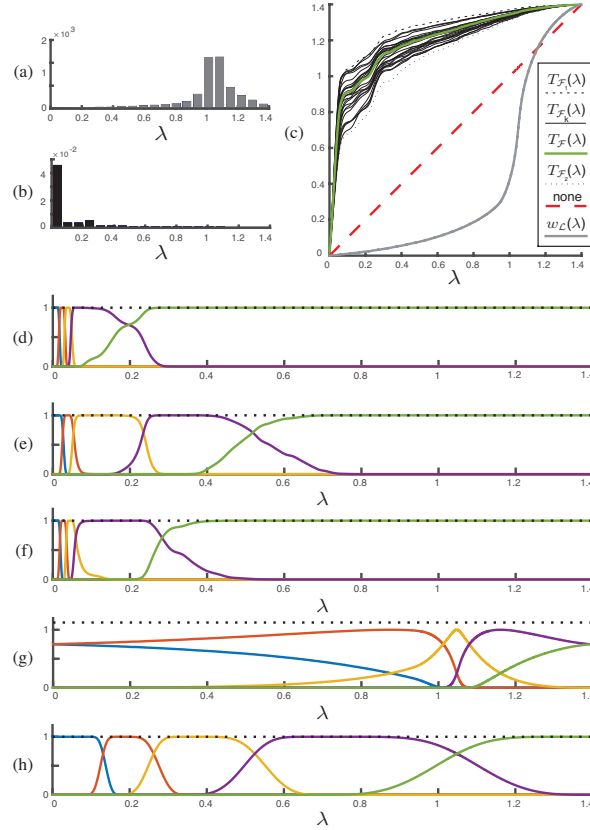
Fig. 8. Illustration of the cerebellum graph.

Fig. 9. (a) Histogram of the eigenvalues $\Lambda_{\mathcal{L}}(\mathcal{G})$ of the cerebellum graph. (b) Distribution of the ensemble energy spectral density $\bar{e}_{\mathcal{F}}$ of $\mathcal{F}$. (c) Constructed energy-equalizing transformation functions $\{T_{\mathcal{F}_k}(\lambda)\}_{k=1}^{26}$, $T_{\mathcal{F}}(\lambda)$ and the spectrum-adapting warping function $w_{\mathcal{L}}(\lambda)$ [24]. Note that the black curves correspond to the transformation for each subjects signal set, with the transformation at the two extremes presented as dashed curves. (d)-(f) Signal-adapted system of spectral kernels based on $\mathcal{F}_1$, $\mathcal{F}_2$ and $\mathcal{F}$, respectively. (g) Spectrum-adapted system of spectral kernels [24]. (h) Meyer-like wavelet frame system of spectral kernels [11].

were acquired from 26 healthy subjects while performing a slow event-related Eriksen flanker task [50].[2] Whole-brain functional volumes were acquired throughout the task, one every two seconds, with a total of 292 volumes per subject. An anatomical scan of each subject's brain was also collected. By registering each subject's anatomical scan to the template cerebellum, and accordingly mapping the functional volumes, cerebellar graph signals were constructed by extracting functional voxels matching the defined graph vertices. This resulted in one signal set for each subject, $\{\mathcal{F}_k\}_{k=1}^{26}$, each including 292 signals. A set including the signals from all subjects was also constructed as $\mathcal{F} = \mathcal{F}_1 \cup \mathcal{F}_2 \cup \cdots \cup \mathcal{F}_{26}$.

The histogram of the eigenvalues $\Lambda_{\mathcal{L}}(\mathcal{G})$ of the cerebellum gray matter graph and the distribution of the ensemble energy spectral density of $\mathcal{F}$ are shown in Figs. 9(a) and (b), respectively. Unlike the Alameda traffic graph and data, a major difference is observed between the two distributions: most eigenvalues are located at the upper end of

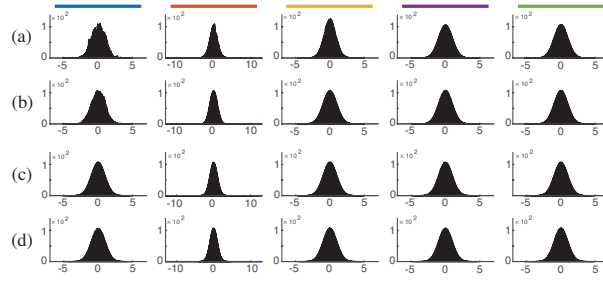[2]The data are publicly available at https://openfmri.org/dataset/ds000102.

Fig. 10. (a)-(d) The mean distribution of coefficients resulting from decomposing the set of signal in $\mathcal{F}_{\text{noise}}$ using the system of spectral kernels shown in Figs. 9(e)-(h), respectively. Each plot shows the mean distribution of coefficients in one subband, with the order from left to right corresponding to the blue, red, yellow, purple and green spectral kernels, respectively. Note that the distributions are shown for the same range of coefficient values for each subband across the frames, whereas the counts (y axis) are adjusted for each distribution.
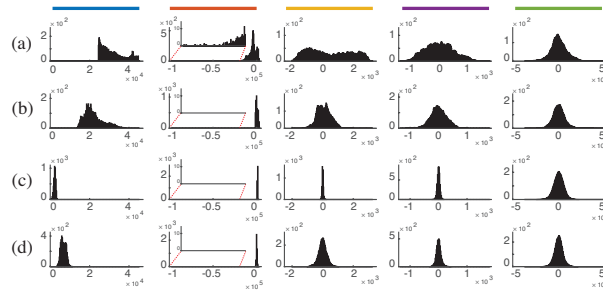

Fig. 11. (a)-(e) The same as in Fig. 10 but for $\mathcal{F}_2$. For clarity, the lower end of the distributions in the second subbands are vertically magnified.

the spectrum, whereas the ensemble energy is significantly concentrated at the lower end of the spectrum. This leads to a major discrepancy between $T_{\mathcal{F}}(\lambda)$ and $w_{\mathcal{L}}(\lambda)$, see Fig. 9(c), and consequently, the resulting spectral kernels, see Figs. 9(d)-(g): the kernels of the signal-adapted frames are localized at the lower end of the spectrum, whereas those of the spectrum-adapted frame are localized at the higher end of the spectrum. As a result, the signal-adapting scheme leads to an optimal configuration of filters in the sense that more filters are allocated to spectral regions where higher ensemble signal energy is present rather than being allocated to part of the spectrum where more eigenvalues are located.

The concentration of the ensemble energy at the lower end of the spectrum makes the non-adapted tight Meyer-like wavelet design a relatively suitable design, see Fig. 9(h). However, comparing Figs. 9(f) and (h) demonstrates the efficiency of the proposed construction: the spectral range lying in the support of the first kernel of the Meyer-like wavelet frame (blue kernel) is spanned by approximately three kernels in the signal-adapted frame (blue, red and orange kernels). This suggests the superiority of the signal-adapted frame in providing a more efficient multi-scale representation of the data compared to the dyadically scaled wavelet design. Interestingly, the narrowband configuration of the proposed signal-adapted frame closely resembles that found to be optimal for

July 5, 2016 DRAFT

analyzing cerebellar data by tunning the Meyer-like wavelet frame in [30], [31].[3]

To provide further insight than that provided by visual comparison of the frames, the set of cerebellar signals in $\mathcal{F}_2$ as well as a set of 292 white Gaussian noise signals ($\sigma^2 = 1$), denoted $\mathcal{F}_{\text{noise}}$, were decomposed using the frames shown in Figs. 9(e)-(h). The resulting coefficients $c_{\mathcal{K}_j,m}$ (cf. (18)) were normalized as $c_{\mathcal{K}_j,m}/\|\psi_{\mathcal{K}_j,m}\|_2$ to account for the difference in the norm of their associated atoms. Figs. 10 and 11 show the mean distribution of the normalized decomposition coefficients of $\mathcal{F}_{\text{noise}}$ and $\mathcal{F}$, respectively. For $\mathcal{F}_{\text{noise}}$, the distributions of the coefficients in all five subbands of all four frames are Gaussian-like as expected, see Fig. 10. For $\mathcal{F}_2$, the distributions of the coefficients of the last subband of all four frames closely resembles a Gaussian (see the 4 plots at the right end of Figs. 11(a)-(d)), similar to the distributions obtained on $\mathcal{F}_{\text{noise}}$, cf. Figs. 11. This suggests that the corresponding atoms of this subband have captured noise. A similar observation is made for the distributions associated with the third and fourth subbands of the spectrum-adapted and Meyer-like wavelet frame (Figs. 11(c)-(d)) as well as the fourth subband of the signal-adapted frame constructed based on $\mathcal{F}$ (Fig. 11(b)). On the other hand, the coefficients of the first four subbands of the signal-adapted frame constructed based on $\mathcal{F}_2$ significantly deviate from zero, cf. first four plots in Fig. 11(a), and have distributions unlike that expected to result from decomposing noise (cf. Fig. 10). This suggests the efficiency of the signal-adapted frame in resulting in atoms that capture signal components.

Although the coefficients of the first three subbands of the signal-adapted frame constructed based on $\mathcal{F}$ (Fig. 11(b)) also provide more information than those of the spectrum-adapted and Meyer-like wavelet frame (Figs. 11(c)-(d)), the coefficients of the signal-adapted frame constructed based on $\mathcal{F}_2$ (Fig. 11(a)) show yet greater significance. This suggestes the superiority of the subject-based frame, shown in Fig. 9(e), over the frame constructed based on the energy content of the signals from the group of subjects, shown in Fig. 9(f), for decomposing $\mathcal{F}_2$. Despite this fact, if a comparison between decomposition of the signals from different subjects is required, as for instance performed in [31], the group-based frame constructed based on $\mathcal{F}$ is the suitable choice.

D. Signal Denoising

To illustrate one application of the proposed frames, we denoise a set of signals realized on the Minnesota road graph. For different σ_e^2, noise corrupted signal sets were created using (43) as $\mathcal{F}_{\sigma_e} = \left\{ \boldsymbol{y}_{\eta,2}^{[i]} \right\}_{\eta=0.2,0.5}$, where $i = 1, \ldots, 10$ denotes different random realizations of $\boldsymbol{p}_\eta$ and $\boldsymbol{e}$. Denoising was performed through soft thresholding of the decomposition coefficients. In particular, the denoised signal $\tilde{\boldsymbol{x}}^{(\vartheta)}$, where ϑ denotes the global threshold, were obtained from the coefficients $c_{\mathcal{K},m}$ (cf. (18)) as

$$\tilde{\boldsymbol{x}}^{(\vartheta)} = \sum_{j,m} H(|c_{\mathcal{K}_j,m}| - \tau_{\mathcal{K}_j,m}^{(\vartheta)}) c_{\mathcal{K}_j,m} \psi_{\mathcal{K}_j,m}, \tag{46}$$

---

[3]The similarity can be observed by comparing the spectral coverage of the first two kernels in Fig. 9(f) with that in figure 3(d) in [31] and figure 1(b) in [30].
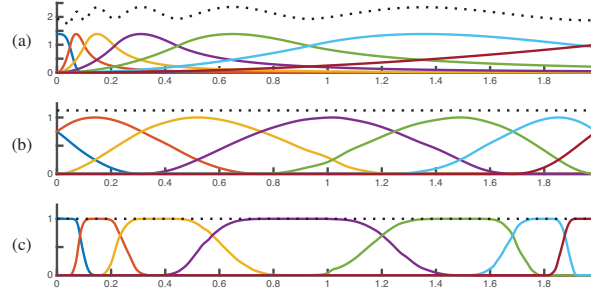
Fig. 12. Systems of spectral kernels of the (a) SGWT frame [10], (b) spectrum-adapted frame and (c) the signal-adapted frame, used for denoising.

where $H(\cdot)$ denotes the Heaviside step function and $\tau^{(\vartheta)}_{\mathcal{K}_j,m}$ denote element-adapted thresholds. In order to account for that the frame elements generally have different norms, element-adapted thresholds of the form [12]

$$\tau^{(\vartheta)}_{\mathcal{K}_j,m} = \sigma_e ||\boldsymbol{\psi}_{\mathcal{K}_j,m}||_2 \vartheta, \tag{47}$$

were used that are also adjusted with respect to the noise level.

To evaluate performance, the average mean square error reduction (ΔMSE) was computed for each ϑ as

$$\Delta\text{MSE}^{(\vartheta)} = \frac{1}{|\mathcal{F}|} \sum_{k=1}^{|\mathcal{F}|} 10 \log_{10}\left(\frac{||\boldsymbol{x}_k - \tilde{\boldsymbol{x}}_k^{(\vartheta)}||_2^2}{||\boldsymbol{x}_k - \boldsymbol{y}_k||_2^2}\right), \tag{48}$$

where k runs over the signals in the signal set, and $|\mathcal{F}|$ denotes the cardinality of the set. To compare the denoising performance across different σ_e^2 and J, the minimum MSE reduction is considered

$$\Delta\text{MSE}^{(\text{min})} = \min_\vartheta \Delta\text{MSE}^{(\vartheta)}, \tag{49}$$

where ϑ is varied within the interval $[1, 5]$ with a step size of 0.25.

The signals in the sets $\mathcal{F}_{\sigma_e}$ were denoised using: 1) the SGWT frame [10], neither adapted to the spectrum nor to the graph signals, 2) the spectrum-adapted frame [24] and 3) the proposed frame, using up to 10 spectral scales. Fig. 12 illustrates an example of the used frames for $J = 7$ scales and $\sigma_e^2 = \sigma_x^2$. By comparing the signal-adapted frame constructed based on the noisy signals, shown in Fig. 12(c), with its respective counterpart constructed based on the clean signals, shown in the top of Fig. 4(d), it is observed that although the addition of noise has affected the support of the filters, the overall distribution of the support of the kernels across the spectrum is still consistent between the two cases. As white Gaussian noise exhibits a uniform energy spectral density, its influence on the overall distribution of ensemble energy is negligible.

Fig. 13 presents a comparison of denoising performances of the different types of frames. Fig. 13(a) presents ΔMSE measures when using $J = 7$ scales. Compared to the spectrum-adapted frame, the proposed frames show better performance across ϑ. Compared to the SGWT frame, the difference in performance is generally better for ϑ values close to the minimum ΔMSE. Fig. 13(b) presents the $\Delta\text{MSE}^{(\text{min})}$ measures as a function of the SNR $= \sigma_x^2/\sigma_e^2$, when using $J = 7$. The relative performance of the signal-adapted frame and SGWT frame is
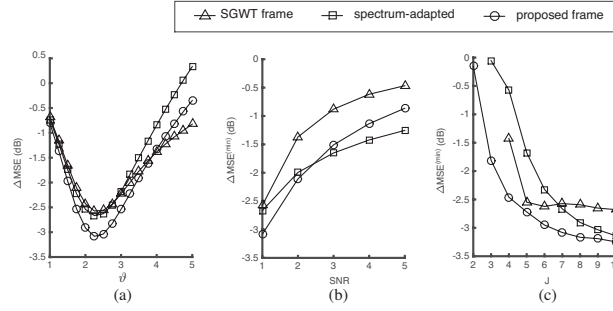
Fig. 13. Comparison of denoising performance using different frames on signal set $\mathcal{F}_{\sigma_e}$. (a) ΔMSE as a function of threshold level ϑ using $J = 7$ spectral scales, where $\sigma_e^2 = \sigma_x^2$ in signal realization for all cases. (b) ΔMSE$^{(\min)}$ as a function of noise level, SNR $= \sigma_x^2/\sigma_e^2$, using frames with (b) $J = 7$ spectral scales. (c) ΔMSE$^{(\min)}$ as a function of the number of scales J, where $\sigma_e^2 = \sigma_x^2$ in signal realization for all cases. Note that in (c) some sample points are not displayed for the spectrum-adated and the SGWT frame; this is due to that the former is only defined for $J \geq 3$, and that the latter is not optimally designed for small J and leads to ΔMSE$^{(\min)} > 0$.

consistent across different SNRs, with the proposed frame offering, in general, better noise reduction. However, the spectrum-adapted frame shows better noise reduction at higher SNRs.

It is advantageous to also compare the denoising performance for frames with different number of subbands. Fig. 13(c) shows plots of the ΔMSE$^{(\min)}$ in denoising the signals using the different frames as a function of J. Using the signal-adapted frames leads to a lower ΔMSE$^{(\min)}$ than for the SGWT and spectrum-adapted frames. It is also observed that the difference in performance between the proposed frame and the spectrum-adapted frame is more pronounced when using up to eight scales. In recent studies that use the SGWT [31], [51]–[55], typically, up to six scales are used, which can be explained by observing that the performance for this frame saturates at $J = 5$. In contrast, the proposed frames show a monotonic increase in performance also for larger number of scales. This suggests that the proposed signal-adapted frames have the potential to provide a more meaningful multi-scale representation of graph signals.

VI. CONCLUSION

We have presented a construction of signal-adapted tight frames for graph signals. The adaptivity of the approach is introduced by exploiting the ensemble energy spectral density, which describes the second-order statistics of the signal class at hand. It should be noted that wavelet-type decompositions are optimal for non-stationary signals, such as the ones considered in our experimental results. From this point-of-view, the design only uses the stationary information, and its flexibility to represent non-stationary features comes with the width and smoothness of the bandpass characteristics. Conceptually, this approach is similar to optimal pyramid designs that were proposed for image processing in the 90's [56]. A MATLAB implementation of the proposed approach and the datasets used in this paper are available online at miplab.epfl.ch/software/.

APPENDIX A

PROOF OF PROPOSITION 1

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

$$\lambda_{\max} \overset{(35c)}{=} \lambda_{\mathrm{II}} + a \overset{(j=J)}{=} \gamma a + (J-2)\Delta + a,$$

which using (36a) leads to $a = \frac{\lambda_{\max}}{J\gamma - J - \gamma + 3}$.

To prove that the UMT system of spectral kernels form a tight frame, (17) needs to be fulfilled. Since, for all j, the supports of $\mathcal{K}'_{j-1}(\lambda)$ and $\mathcal{K}'_{j+1}(\lambda)$ are disjoint, $G(\lambda)$ can be determined as

$$G(\lambda) = \sum_{j=1}^{J} |\mathcal{K}'_j(\lambda)|^2$$

$$\overset{(35)}{=} \begin{cases} |\mathcal{K}'_1(\lambda)|^2 \overset{(35a)}{=} 1 & \forall \lambda \in [0, a] \\ |\mathcal{K}'_1(\lambda)|^2 + |\mathcal{K}'_2(\lambda)|^2 & \forall \lambda \in]a, \gamma a] \\ |\mathcal{K}'_2(\lambda)|^2 + |\mathcal{K}'_3(\lambda)|^2 & \forall \lambda \in]\gamma a, \gamma a + \Delta] \\ \vdots & \vdots \\ |\mathcal{K}'_J(\lambda)|^2 \overset{(35c)}{=} 1 & \forall \lambda \in]\lambda_{\max} - a, \lambda_{\max}] \end{cases}$$

$$\overset{(35b)}{=} \begin{cases} 1 & \forall \lambda \in [0, a] \\ \cos^2(x_{\mathrm{I}}) + \sin^2(x_{\mathrm{I}}) & \forall \lambda \in]a, \gamma a] \\ \cos^2(x_{\mathrm{II}}) + \sin^2(x_{\mathrm{II}}) & \forall \lambda \in]\gamma a, \gamma a + \Delta] \\ \vdots & \vdots \\ 1 & \forall \lambda \in]\lambda_{\max} - a, \lambda_{\max}] \end{cases}$$

$$= 1 \quad \forall \lambda \in [0, \lambda_{\max}] \tag{50}$$

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

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 (28), the appropriate γ needs to be determined. From (35b), we have $\forall j \in \{2, \ldots, J-2\}$

$$\mathcal{K}'_j(\lambda) = \mathcal{K}'_{j+1}(\lambda + \Delta) \quad \forall \lambda \in]\lambda_{\mathrm{I}}, \lambda_{\mathrm{II}} + \Delta]. \tag{51}$$

By considering an inverse linear mapping of the spectral support where $\mathcal{K}'_1(\lambda) \neq 0$, i.e. $[0, \gamma a]$, to the spectral support where $\mathcal{K}'_J(\lambda) \neq 0$, i.e. $[\lambda_{\max} - \gamma a, \lambda_{\max}]$, we have

$$\mathcal{K}'_1(\lambda) = \mathcal{K}'_J(-\lambda + 2a + J\Delta) \quad \forall \lambda \in [0, \gamma a]. \tag{52}$$
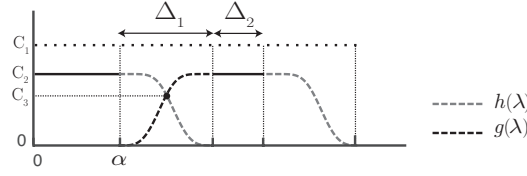
Fig. 14. Design of a uniform system of spectral kernels satisfying (28) and (29). The displayed function $h(\lambda)$ and $g(\lambda)$ correspond to the ones used in the UMT design. However, any two functions passing through $(\alpha + \Delta_1/2, C_3)$ and satisfying (56) and (57) can be used to design a desired prototype system of uniform spectral kernels.

Thus, from (51) and (52) we have

$$\int_0^{\lambda_{\max}} \mathcal{K}'_j(\lambda)\,d\lambda = C_2, \quad j = 2, \ldots, J-1 \tag{53a}$$

$$\int_0^{\lambda_{\max}} \mathcal{K}'_1(\lambda)\,d\lambda = \int_0^{\lambda_{\max}} \mathcal{K}'_J(\lambda)\,d\lambda = C_1, \tag{53b}$$

respectively, where $C_1, C_2 \in \mathbb{R}^+$. Thus, in order to satisfy (28), γ should be chosen such that

$$C_1 = C_2$$

$$\int_0^{\lambda_{\max}} \mathcal{K}'_1(\lambda)\,d\lambda = \int_0^{\lambda_{\max}} \mathcal{K}'_2(\lambda)\,d\lambda$$

$$a + \int_a^{\gamma a} \mathcal{K}'_1(\lambda)\,d\lambda = \int_a^{\gamma a} \sin\left(\frac{\pi}{2}\nu\left(\frac{1}{\gamma-1}\left(\frac{\lambda}{a}-1\right)\right)\right)\,d\lambda$$

$$+ \int_{\gamma a}^{\gamma a + \Delta} \mathcal{K}'_2(\lambda)\,d\lambda$$

$$a \overset{(51)}{=} \int_a^{\gamma a} \sin\left(\frac{\pi}{2}\nu\left(\frac{1}{\gamma-1}\left(\frac{\lambda}{a}-1\right)\right)\right)\,d\lambda. \tag{54}$$

The optimal γ that satisfies (54) was obtained numerically by defining

$$Q(\gamma) = \int_a^{\gamma a} \sin\left(\frac{\pi}{2}\nu\left(\frac{1}{\gamma-1}\left(\frac{\lambda}{a}-1\right)\right)\right)\,d\lambda - a, \tag{55}$$

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 $\lambda_{\max}$ and J, was found to be $\gamma = 2.73$.

APPENDIX B
UNIFORM SYSTEM OF KERNELS – GENERAL CASE

The UMT system of spectral kernels (cf. Proposition 1) is one prototype design that can be used as the basis for the proposed frame design approach. For the general case, other prototype system of kernels can be designed, with their design essentially breaking down to (see Fig. 14 for notations):

1) the proper design of the pair of functions $h(\lambda)$ and $g(\lambda)$ such that they lead to smooth quadrature mirror style filters that satisfy

$$|h(\lambda)|^2 + |g(\lambda)|^2 = C_1 \in \mathbb{R}^+, \quad \forall \lambda \in [\alpha, \alpha + \Delta_1], \tag{56}$$

where $\alpha, \Delta_1 \in\,]0, \lambda_{\max}[$, $C_2 = \sqrt{C_1}$, $C_3 = \sqrt{C_1/2}$,

2) determination of α, Δ_1 and Δ_2 such that

$$\alpha C_2 = \Delta_2 C_2 + \int_\alpha^{\alpha+\Delta_1} h(\lambda)\, \mathrm{d}\lambda. \tag{57}$$

Observe that α, Δ_1, Δ_2, $h(\lambda)$ and $g(\lambda)$ all depend on J and $\lambda_{\max}$. Fig. 14 shows an illustration of the non-zero segments of the resulting $\mathcal{K}_1'(\lambda)$ and $\mathcal{K}_2'(\lambda)$. $\{\mathcal{K}_j'(\lambda)\}_{j=3}^{J-1}$ can be obtained by translating the non-zero segments of $\mathcal{K}_2'(\lambda)$ by $(j-2)\Delta_3$, and $\mathcal{K}_J'(\lambda)$ by mirroring the non-zero segment of $\mathcal{K}_1'(\lambda)$ and translating it such that it spans the upper end of the spectrum.

REFERENCES

[1] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. 83–98, 2013.

[2] A. Sandryhaila and J. Moura, "Discrete signal processing on graphs." *IEEE Trans. Signal Process.*, vol. 61, pp. 1644–1656, 2013.

[3] S. Chen, R. Varma, A. Sandryhaila, and J. Kovačević, "Discrete signal processing on graphs: sampling theory," *IEEE Trans. Signal Process.*, vol. 63, no. 24, pp. 6510–6523, 2015.

[4] M. Jansen, G. P. Nason, and B. W. Silverman, "Multiscale methods for data on graphs and irregular multidimensional situations," *J. Roy. Statistical Soc.: Series B (Statistical Methodology)*, vol. 71, no. 1, pp. 97–125, 2009.

[5] S. K. Narang and A. Ortega, "Lifting based wavelet transforms on graphs," in *Proc. APSIPA ASC*, 2009, pp. pp. 441–444.

[6] R. Rustamov and L. Guibas, "Wavelets on graphs via deep learning," in *Proc. Adv. Neural Info. Proces. Syst.*, 2013, pp. 998–1006.

[7] I. Ram, M. Elad, and I. Cohen, "Generalized tree-based wavelet transform," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4199–4209, 2011.

[8] ——, "Redundant wavelets on graphs and high dimensional data clouds," *IEEE Signal Process. Lett.*, vol. 19, no. 5, pp. 291–294, 2012.

[9] M. Gavish, B. Nadler, and R. R. Coifman, "Multiscale wavelets on trees, graphs and high dimensional data: Theory and applications to semi supervised learning." in *Proc. Int. Conf. Mach. Learn.*, 2010, pp. 367–374.

[10] D. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Appl. Comput. Harmon. Anal.*, vol. 30, no. 2, pp. 129–150, 2011.

[11] N. Leonardi and D. Van De Ville, "Tight wavelet frames on multislice graphs," *IEEE Trans. Signal Process.*, vol. 61, no. 13, pp. 3357–3367, 2013.

[12] F. Göbel, G. Blanchard, and U. von Luxburg, "Construction of tight frames on graphs and application to denoising," Aug. 2014 [online]. Available: http://arXiv.org/abs/1408.4012.

[13] B. Dong, "Sparse representation on graphs by tight wavelet frames and applications," *Appl. Comput. Harmon. Anal.*, 2015, doi:10.1016/j.acha.2015.09.005.

[14] R. R. Coifman and M. Maggioni, "Diffusion wavelets," *Appl. Comput. Harmon. Anal.*, vol. 21, no. 1, pp. 53–94, 2006.

[15] S. K. 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, 2012.

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

[17] Y. Tanaka and A. Sakiyama, "M-channel oversampled graph filter banks," *IEEE Trans. Signal Process.*, vol. 62, no. 14, pp. 3578–3590, 2014.

[18] A. Sakiyama and Y. Tanaka, "Oversampled graph laplacian matrix for graph filter banks," *IEEE Trans. Signal Process.*, vol. 62, no. 24, pp. 6425–6437, 2014.

[19] D. B. H. Tay and J. Zhang, "Techniques for constructing biorthogonal bipartite graph filter banks," *IEEE Trans. Signal Process.*, vol. 63, no. 21, pp. 5772–5783, 2015, doi: 10.1109/TSP.2015.2460216.

[20] D. B. Tay and Z. Lin, "Design of near orthogonal graph filter banks," *IEEE Signal Process. Lett.*, vol. 22, no. 6, pp. 701–704, 2015.

[21] N. Tremblay and P. Borgnat, "Subgraph-based filterbanks for graph signals," *IEEE Trans. Signal Process.*, vol. 64, no. 15, pp. 3827–3840, 2016.

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

[23] D. I. Shuman, B. Ricaud, and P. Vandergheynst, "Vertex-frequency analysis on graphs," *Appl. Comput. Harmon. Anal.*, 2015, doi: 10.1016/j.acha.2015.02.005.

[24] D. I. Shuman, C. Wiesmeyr, N. Holighaus, and P. Vandergheynst, "Spectrum-adapted tight graph wavelet and vertex-frequency frames," *IEEE Trans. Signal Process.*, vol. 63, no. 16, pp. 4223–4235, 2015.

[25] D. I. Shuman, M. J. Faraji, and P. Vandergheynst, "A multiscale pyramid transform for graph signals," *IEEE Trans. Signal Process.*, vol. 64, no. 8, pp. 2119–2134, 2016.

[26] D. Thanou, D. I. Shuman, and P. Frossard, "Learning parametric dictionaries for signals on graphs," *IEEE Trans. Signal Process.*, vol. 62, no. 15, pp. 3849–3862, 2014.

[27] D. Thanou and P. Frossard, "Multi-graph learning of spectral graph dictionaries." in *Proc. IEEE Int. Conf. Acoust. Speech and Signal Process.*, 2015.

[28] J. Benedetto and M. Fickus, "Finite normalized tight frames," *Adv. Comput. Math.*, vol. 18, no. 2, pp. 357–385, 2003.

[29] H. Behjat, N. Leonardi, and D. Van De Ville, "Statistical parametric mapping of functional MRI data using wavelets adapted to the cerebral cortex," in *Proc. IEEE Int. Symp. Biomed. Imaging*, 2013, pp. 1070–1073.

[30] H. Behjat, N. Leonardi, L. Sörnmo, and D. Van De Ville, "Canonical cerebellar graph wavelets and their application to fMRI activation mapping," in *Proc. IEEE Int. Conf. Eng. Med. Biol. Soc.*, 2014, pp. 1039–1042.

[31] ——, "Anatomically-adapted graph wavelets for improved group-level fMRI activation mapping," *NeuroImage*, vol. 123, pp. 185–199, 2015.

[32] F. Chung, *Spectral graph theory*. Providence, RI: AMS, 1997.

[33] Y. Meyer, "Principe d'incertitude, bases hilbertiennes et algebres d'operateurs," *Seminaire Bourbaki (in French)*, vol. 662, pp. 209–223, 1986.

[34] F. N. Fritsch and R. E. Carlson, "Monotone piecewise cubic interpolation," *SIAM J. Numerical Anal.*, vol. 17, no. 2, pp. 238–246, 1980.

[35] A. G. Marques, S. Segarra, G. Leus, and A. Ribeiro, "Stationary graph processes and spectral estimation," arXiv preprint arXiv:1603.04667 (2016).

[36] N. Perraudin and P. Vandergheynst, "Stationary signal processing on graphs," arXiv preprint arXiv:1601.02522 (2016).

[37] S. Chepuri and G. Leus, "Subsampling for graph power spectrum estimation," arXiv preprint arXiv:1603.03697 (2016).

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

[39] M. Tsitsvero, S. Barbarossa, and P. Di Lorenzo, "Signals on graphs: uncertainty principle and sampling," *IEEE Trans. Signal Process.*, 2016, DOI 10.1109/TSP.2016.2573748.

[40] P. Koprowski, "Graph theoretic uncertainty and feasibility," arXiv preprint arXiv:1603.02059 (2016).

[41] N. Perraudin, B. Ricaud, D. Shuman, and P. Vandergheynst, "Global and local uncertainty principles for signals on graphs," arXiv preprint arXiv:1603.03030 (2016).

[42] D. Gleich. The MatlabBGL matlab library [online]. Available: http://www.cs.purdue.edu/homes/dgleich/packages/matlab˙bgl/index.html.

[43] Google, [google maps for alameda county.]. Available: https://goo.gl/maps/X4Ye6CHmvXA2. Retrieved March, 2016.

[44] N. Logothetis and B. Wandell, "Interpreting the BOLD signal." *Annu. Rev. Physiol.*, vol. 66, pp. 735–769, 2004.

[45] S. J. Kiebel, R. Goebel, and K. J. Friston, "Anatomically informed basis functions," *NeuroImage*, vol. 11, no. 6, pp. 656–667, 2000.

[46] M. K. Chung, S. M. Robbins, K. M. Dalton, R. J. Davidson, A. L. Alexander, and A. C. Evans, "Cortical thickness analysis in autism with heat kernel smoothing," *NeuroImage*, vol. 25, no. 4, pp. 1256–1265, 2005.

[47] D. J. Hagler Jr, A. P. Saygin, and M. I. Sereno, "Smoothing and cluster thresholding for cortical surface-based group analysis of fMRI data," *NeuroImage*, vol. 33, no. 4, pp. 1093–1103, 2006.

[48] S. Ozkaya and D. Van De Ville, "Anatomically adapted wavelets for integrated statistical analysis of fMRI data," in *Proc. IEEE Int. Symp. Biomed. Imaging*, Chicago, IL, 2011, pp. 469–472.

[49] J. Diedrichsen, J. H. Balsters, J. Flavell, E. Cussans, and N. Ramnani, "A probabilistic MR atlas of the human cerebellum," *NeuroImage*, vol. 46, no. 1, pp. 39–46, 2009.

[50] A. M. C. Kelly, L. Q. Uddin, B. B. Biswal, F. X. Castellanos, and M. P. Milham, "Competition between functional brain networks mediates behavioral variability," *NeuroImage*, vol. 39, no. 1, pp. 527–537, 2008.

[51] W. H. Kim, V. Singh, M. K. Chung, C. Hinrichs, D. Pachauri, O. C. Okonkwo, and S. C. Johnson, "Multi-resolutional shape features via non-Euclidean wavelets: Applications to statistical analysis of cortical thickness," *NeuroImage*, vol. 93, pp. 107–123, 2014.

[52] W. H. Kim, N. Adluru, M. K. Chung, O. C. Okonkwo, S. C. Johnson, B. Bendlin, and V. Singh, "Multi-resolution statistical analysis of brain connectivity graphs in preclinical Alzheimer's disease." *NeuroImage*, vol. 118, pp. 103–117, 2015.

[53] C. Li and A. B. Hamza, "A multiresolution descriptor for deformable 3D shape retrieval," *Vis. Comput.*, vol. 29, no. 6, pp. 513–524, 2013.

[54] M. Zhong and H. Qin, "Sparse approximation of 3D shapes via spectral graph wavelets," *Vis. Comput.*, vol. 30, no. 6, pp. 751–761, 2014.

[55] X. Yan, H. Qin, J. Li, H. Zhou, and J. G. Zong, "Infrared and visible image fusion with spectral graph wavelet transform," *J. Opt. Soc. Am. A*, vol. 32, no. 9, pp. 1643–1652, 2015.

[56] M. Unser, "On the optimality of ideal filters for pyramid and wavelet signal approximation," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3591–3596, 1993.
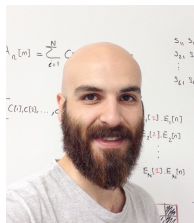
Hamid Behjat received the B.A. degree in biomedical engineering from Amirkabir University of Technology, Tehran, Iran, in 2007, and the M.Sc. degree in biomedical engineering from Linköping University, Linköping, Sweden, in 2012. He is currently working toward the Ph.D. degree at Lund University, Lund, Sweden, and is part of the Signal Processing Group, Department of Biomedical Engineering. His research interests include signal processing and wavelets, with particular focus on developing methodologies for fMRI data analysis.

Ulrike Richter was born in 1981. She received the Diplom-Ingenieur (M.Sc.) degree in electrical engineering from the Leibniz University, Hanover, Germany, in 2005, and the Ph.D. degree in signal processing from Lund University, Lund, Sweden, in 2010. Her Ph.D. research was focused on the development of methods for spatial characterization and estimation of intracardiac propagation patterns during atrial fibrillation. Since 2011, she has had positions as a Postdoctoral Fellow and Researcher in the Integrative Neurophysiology Group at the Department of Experimental Medical Science and the Neuronano Research Center, Lund University. Her research focusses on signal processing methods for the analysis of neurophysiological signals, in particular extracellular spikes and local field potentials. A main interest is the study of changes in functional connectivity between different brain areas in relation to disease states.

Dimitri Van De Ville (M'02–SM'12) received the M.S. degree in engineering and computer sciences and the Ph.D. degree from Ghent University, Belgium, in 1998, and 2002, respectively. After a post-doctoral stay (2002-2005) at the lab of Prof. Michael Unser at the École Polytechnique Fédérale de Lausanne (EPFL), Switzerland, he became responsible for the Signal Processing Unit at the University Hospital of Geneva, Switzerland, as part of the Centre dImagerie Biomédicale (CIBM). In 2009, he has been the recipient of a Swiss National Science Foundation professorship and since 2015 became Professor of Bioengineering at the EPFL and the University of Geneva, Switzerland. His research interests include wavelets, sparsity, pattern recognition, and their applications in computational neuroimaging. He received the Pfizer Research Award 2012 and the NARSAD Independent Investigator award in 2014.

Dr. Van De Ville served as an Associate Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING from 2006 to 2009 and the IEEE SIGNAL PROCESSING LETTERS from 2004 to 2006. He was the Chair of the Bio Imaging and Signal Processing (BISP) TC of the IEEE Signal Processing Society (2012-2013). He is Co-Chair of the biennial Wavelets & Sparsity series conferences, together with V. Goyal and M. Papadakis.

Leif Sörnmo (S'80–M'85–SM'02) received the M.Sc. and Ph.D. degrees in electrical engineering from Lund University, Lund, Sweden, in 1978 and 1984, respectively.

Since 1990, he has been with the Biomedical Signal Processing Group, Department of Biomedical Engineering, Lund University, where he is currently Professor. He is, together with Pablo Laguna, the author of *Bioelectrical Signal Processing in Cardiac and Neurological Applications* (Elsevier, 2005). His research interests include statistical signal processing, modeling of biomedical signals, methods for analysis of atrial fibrillation, multimodal signal processing in hemodialysis, and power-efficient signal processing in implantable devices.

Dr. Sörnmo is a Fellow of International Academy of Medical and Biological Engineering and European Alliance for Medical and Biological Engineering. He is Associate Editor of IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING, *Journal of Electrocardiology*, and *Medical and Biological Engineering & Computing*; he is on the editorial board of *Journal of Biomedical Engineering*. He was an Associate Editor of *Computers in Biomedical Research* (1997–2000). He is founder and director of the undergraduate and graduate program in biomedical engineering at Lund University. He serves on the board of directors of Computing in Cardiology.