Tight Wavelet Frames on Multislice Graphs

Leonardi, Nora; Van De Ville, Dimitri

How to cite

LEONARDI, Nora, VAN DE VILLE, Dimitri. Tight Wavelet Frames on Multislice Graphs. In: IEEE transactions on signal processing, 2013, vol. 61, n° 13, p. 3357–3367. doi: 10.1109/TSP.2013.2259825

This publication URL: https://archive-ouverte.unige.ch/unige:39809
Publication DOI: 10.1109/TSP.2013.2259825

Tight Wavelet Frames on Multislice Graphs

Nora Leonardi, *Student Member, IEEE*, and Dimitri Van De Ville, *Senior Member, IEEE*

***Abstract*—We present a framework for the design of wavelet transforms tailored to data defined on multislice graphs (i.e., multiplex or dynamic graphs). Graphs with multiple types of interactions are ubiquitous in real life, motivating the extension of wavelets to these complex domains. Our framework generalizes the recently proposed spectral graph wavelet transform (SGWT) [D. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on Graphs via Spectral Graph Theory," *Appl. Comput. Harmon. Anal.*, vol. 30, pp. 129–150, Mar. 2011], which is designed in the spectral (frequency) domain of an arbitrary finite weighted graph. We extend the SGWT to form a tight frame, which conserves energy in the wavelet domain, and define the relationship between conventional and spectral graph wavelets. We then propose a design for multislice graphs that is based on the higher-order singular value decomposition (HOSVD), a powerful tool from multilinear algebra. In particular, the multiple adjacency matrices are stacked to form a tensor and the HOSVD decomposition provides information about its third-order structure, analogous to that provided by matrix factorizations. We obtain a set of "eigennetworks" and from these graph wavelets, which exploit the variability across the graphs. We demonstrate the feasibility of our method 1) by capturing different orientations of a gray-scale image and 2) by decomposing brain signals from functional magnetic resonance imaging. We show its effectiveness to identify variability across graph edges and provide meaningful decompositions.**

***Index Terms*—Higher-order singular value decomposition (HOSVD), multislice graph, spectral graph theory, tensor decompositions, wavelet transform.**

I. Introduction

BRAIN networks, human society or the internet are only some examples of systems that can be modelled as networks. Even though many networks have multiple types of interactions or interactions that change over time, emphasis has traditionally been put on the study of simple or static networks. Recent work has however illustrated the advantages of exploiting the rich information of multislice networks: dynamic changes in modular organization of the human brain during learning are linked to the mastery of a novel skill [1], [2]; the multiple types of relations in human society help to understand the different roles humans play in different networks [3]; broad degree distributions enhance the vulnerability to failure of interdependent networks, in contrast to simple networks [4]; and community detection in datasets with multiple types of similarities can be improved [5].

Often, we want to study a signal *on* the network, not only the network itself. Wavelets are powerful tools in signal processing since they can localize signal contents in both space and frequency, and provide compact and multiscale representations of piecewise smooth signals. However, they are designed for signals defined on regular Euclidean spaces. Since real signals are often defined on irregular domains multiple extensions to signals defined on graphs and manifolds have recently been proposed [6]–[10]. (We will use the terms network and graph interchangeably in this article.) These proposals can be split into spatial and spectral (frequency) designs. The former group includes Jansen *et al.* [6] who based their method on lifting schemes, and Ram *et al.* [7] who constructed wavelets for hierarchical trees. Hammond *et al.* [8], Coifman and Maggioni [9] and Narang and Ortega [10] used spectral graph theory to construct spectral graph wavelets, "diffusion wavelets", and two-channel filter-banks for bipartite graphs, respectively. Importantly, graph wavelet transforms are adapted to the topology of the graph. Our proposed framework generalizes the spectral graph wavelet transform (SGWT) introduced in [8]: a redundant wavelet transform for arbitrary weighted graphs that is based on the eigenspace of the graph Laplacian matrix and that can be efficiently applied to graphs with a large number of nodes by expressing the wavelets as iterates of the graph Laplacian operator.

Motivated by the promising results from multislice networks and the power of wavelets to analyze non-stationary signals, we sought to extend wavelets to multislice graphs. We expect multislice graph wavelets, which can adapt to the variable underlying graph topology, to be better suited to analyze signals defined on multislice graphs than "single slice" graph wavelets, constructed either separately for each slice or after simply averaging across them as is commonly done.

Our contributions in this paper are (1) to introduce a tight frame design for the SGWT, which is desirable because it preserves energy in the transformed domain and is self-reversible, which leads to a more efficient reconstruction scheme than in [8], (2) to make the connection between conventional and spectral graph wavelets explicit, and (3) to generalize the SGWT to multislice graphs. A multislice network, which consists of a set of networks, can be naturally represented with tensors (multi-way arrays). To reveal "hidden" structure in the variable graph topology, we propose to use tensor decompositions. Tensor decompositions are higher-order analogues of matrix factorizations, which are powerful tools, for, e.g., feature selection, but which are limited to two-dimensional data. To study data with three or more dimensions, tensor decompositions have

Manuscript received October 16, 2012; revised February 26, 2013; accepted April 11, 2013. Date of publication April 24, 2013; date of current version June 07, 2013. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Xiang-Gen Xia. This work was supported by the Swiss National Science Foundation (PP00P2-123438 and PP00P2-146318), the Société Académique de Genève, the FOREMANE foundation, the Center for Biomedical Imaging (CIBM) of the Geneva and Lausanne Universities and EPFL, as well as the Leenaards and Louis-Jeantet foundations.

The authors are with the Department of Radiology and Medical Informatics, University of Geneva, CH-1211 Geneva 14, Switzerland (e-mail: nora.leonardi@epfl.ch; dimitri.vandeville@epfl.ch).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2013.2259825

been used in a wide variety of applications, such as chemometrics [11], psychoactivemetrics [12], face recognition [13] and neuroscience (for a review see [14]). The two most commonly used decompositions are the higher order singular value decomposition (HOSVD) and CANDECOMP/PARAFAC (CP), both of which are often considered as higher-order generalizations of the matrix singular value decomposition (SVD) and principal component analysis (PCA) [15]–[18]. Each decomposition extends a different property of the SVD: The HOSVD decomposes a third-order tensor into three orthonormal subspaces associated with the different dimensions of a tensor, and CP decomposes it into a linear combination of rank-1 terms. We propose to use the HOSVD to learn "eigennetworks" from which we obtain novel networks by combining them in different ways. These networks and the resulting graph wavelets capitalize on the variability across a collection of networks.

In related work, Dong *et al.* [5] introduced spectral clustering methods for multislice networks. The authors suggested that their methods may also be used to generalize other methods that are based on the eigenspace of the Laplacian matrix, but no applications other than clustering were shown. Their method differs from ours in that they derive one joint eigenspace for the multislice graph, whereas we derive multiple eigenspaces that capture the variability across the slices. We are not aware of any other proposals to generalize wavelets to multislice graphs.

We apply the proposed method to different types of regular grid graphs and functional magnetic resonance imaging (fMRI) data to demonstrate that it can be a useful tool for analyzing signals defined on arbitrary weighted graphs with multiple types of interactions.

The paper is organized as follows. In Section II, we briefly review the SGWT on which our framework is based. In Section III we introduce the construction of tight graph wavelet frames, rational dilation factors and the relationship between conventional and spectral graph wavelets. In Section IV, we propose a framework to design wavelets on multislice graphs and we present experimental results in Section V.

Notations

Scalars are denoted by lowercase letters (e.g., a), vectors by boldface lowercase letters (e.g., $\mathbf{a}$), matrices by boldface capital letters (e.g., $\mathbf{A}$), and tensors by Euler script letters (e.g., $\mathcal{A}$).

The order of a tensor is its number of dimensions, also known as modes. For third-order tensors, slices are two-dimensional subarrays obtained by fixing one index: $\mathbf{A}_{i::}$, $\mathbf{A}_{:j:}$ and $\mathbf{A}_{::k}$ for a horizontal, vertical and frontal slice, respectively.

The multiplication of a tensor with a matrix in mode n is denoted by $\mathcal{A} \times_n \mathbf{B}$. For a third-order tensor $\mathcal{A} \in \mathbb{R}^{I \times J \times K}$, its mode-1 product with the matrix $\mathbf{B} \in \mathbb{R}^{M \times I}$ is a tensor of size $(M \times J \times K)$ with entries

$$(\mathcal{A} \times_1 \mathbf{B})_{mjk} = \sum_{i=1}^{I} a_{ijk} b_{mi}. \tag{1}$$

Using the n-mode product notation, the matrix product $\mathbf{G} = \mathbf{U} \cdot \mathbf{F} \cdot \mathbf{V}^T$ can be rewritten as $\mathbf{G} = \mathbf{F} \times_1 \mathbf{U} \times_2 \mathbf{V}$.

A tensor can be converted to a matrix by unfolding it along any mode. The mode-1 matricization or unfolding of a third-order tensor $\mathcal{A}$ is the $I \times JK$ matrix $\mathbf{A}_{(1)}$ with entries $(\mathbf{A}_{(1)})_{im} = l_{ijk}$ where $m = (j-1)K + k$, i.e., $\mathbf{A}_{(1)}$ is a reordered concatenation of the frontal slices of the tensor.

II. Review of Spectral Graph Wavelets

We start with a short overview of the construction of the SGWT and refer to [8] for details and proofs.

A. Classical One-Dimensional Wavelets

To solve the problem of how to define shifting and scaling of wavelets on irregular domains, Hammond *et al.* appealed to the Fourier domain. The one-dimensional continuous wavelet $\psi_{s,a}(x) = \frac{1}{s}\psi(\frac{x-a}{s})$, at scale s and location a, can be defined in the Fourier domain as

$$\psi_{s,a}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{\psi}(s\omega) e^{-j\omega a} e^{j\omega x} d\omega. \tag{2}$$

For ease of understanding of the SGWT design that will follow, we highlight several properties of (2):

- the wavelet is represented on complex exponentials, which are eigenfunctions of the one-dimensional Laplacian operator; i.e., $\frac{d^2}{dx^2} e^{j\omega x} = -\omega^2 e^{j\omega x}$;
- shifting the wavelet to location a corresponds to a multiplication by $e^{-j\omega a}$;
- the wavelet can be interpreted as scaled bandpass filter and scaling ψ by $1/s$ corresponds to scaling $\hat{\psi}$ with s.

Shifting and scaling of the classical wavelet can thus be defined in the Fourier domain. To generalize wavelets to graph, we will need the analogue of the Fourier domain for graphs.

B. Weighted Graphs

Let $\mathcal{G} = (V, E, W)$ be an undirected graph consisting of $|V| = N$ nodes that are connected by the set of edges E with nonnegative weights W. The $N \times N$ adjacency matrix $\mathbf{A}$ for a graph without loops is then given by the non-diagonal entries

$$a_{ij} = \begin{cases} w(i,j), & \text{if } (i,j) \in E, \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

The diagonal degree matrix $\mathbf{D}$ is given by $d_{ii} = \sum_j a_{ij}$ and the Laplacian matrix $\mathbf{L}$ by $\mathbf{L} = \mathbf{D} - \mathbf{A}$. The normalized version of the Laplacian matrix $\mathbf{L}$ is defined as $\widetilde{\mathbf{L}} = \mathbf{D}^{-\frac{1}{2}} \mathbf{L} \mathbf{D}^{-\frac{1}{2}}$.

The eigenspaces of the Laplacian matrices form the graph analogue of the Fourier domain. $\mathbf{L}$ and $\widetilde{\mathbf{L}}$ are symmetric, positive semi-definite matrices and can thus be decomposed into non-negative eigenvalues and eigenvectors. For $\mathbf{L}$ the decomposition is given by

$$\mathbf{L} = \mathbf{X} \Lambda \mathbf{X}^T, \tag{4}$$

where $\mathbf{X} = [\boldsymbol{\chi}_1 | \chi_2 | \dots | \chi_N]$ is an orthonormal matrix containing N eigenvectors and $\boldsymbol{\Lambda}$ is a diagonal matrix whose entries equal the associated eigenvalues $\lambda_1 = 0 \leq \lambda_1 \leq \dots \leq \lambda_N$[19]. For $\widetilde{\mathbf{L}}$, $\lambda_\ell \in [0, 2]$ with $\ell = 1, 2, \dots N$, and the eigenvectors are different from those of $\mathbf{L}$ (except if $\widetilde{\mathbf{L}} = \mathbf{L}$, of course).

C. Spectral Graph Wavelets

In analogy to the classical wavelet of II-A, we can now define the spectral graph wavelet $\psi_{t,n}(m)$, at scale t and node n, in the eigenspace of the graph Laplacian matrix $\mathbf{L}$ or $\widetilde{\mathbf{L}}$:

$$\psi_{t,n}(m) = \sum_{\ell=0}^{N-1} g(t\lambda_\ell)\chi_\ell^*(n)\chi_\ell(m), \tag{5}$$

where g is the wavelet generating function defined in the spectral graph domain. We now compare (2) and (5):

- the wavelet is represented on the eigenvectors $\chi_\ell(m)$ of the Laplacian matrix, replacing $e^{i\omega x}$;
- shifting the wavelet to node n corresponds to a multiplication by $\chi_\ell^*(n)$, replacing $e^{-i\omega a}$;
- g behaves as a scaled bandpass filter, replacing $\hat{\psi}$;
- the role of the frequency ω is played by the eigenvalue λ_ℓ.

Note that in practice the continuous scaling parameter t will be sampled to a finite number of scales; i.e., we have $\{t_j\}_{j=1}^J$ for J scales.

We also need to define the scaling function $\phi_n(m)$ to capture the residual "low-pass" components:

$$\phi_n(m) = \sum_{l=0}^{N-1} h(t_J\lambda_\ell)\chi_\ell^*(n)\chi_\ell(m), \tag{6}$$

where h is the scaling function generator. An example of exact specification of $g(\cdot)$ and $h(\cdot)$ was given in .

As in the one-dimensional setting, the wavelet and scaling coefficients are given by the inner product of the graph signal f with the wavelets $\psi_{t,n}$ and the scaling function ϕ_n, respectively:

$$W_f(t,n) = \langle \psi_{t,n}, f\rangle = \sum_{m=1}^{N} \psi_{t,n}^*(m)f(m), \tag{7}$$

$$S_f(n) = \langle \phi_n, f\rangle = \sum_{m=1}^{N} \phi_n^*(m)f(m). \tag{8}$$

Because it is infeasible to explicitly compute the eigenspace of the Laplacian matrix for large graphs, the authors proposed a fast method to compute the forward transform. The method is based on approximating the kernels g and h with a low-dimensional Chebyshev polynomial:

$$g(t_j x) = \frac{1}{2}c_{j,0} + \sum_{k=1}^{M_j} c_{j,k}T_a^k(x), \tag{9}$$

where $c_{j,k}$ denotes the Chebyshev coefficients, M_j the order of the polynomial approximation, and $T_a^k(x)$ the shifted Chebyshev polynomials $T_a^k(x) = T^k(\frac{x-a}{a})$ of order k. The fast approximate spectral graph wavelet forward transform is given by

$$\widetilde{W}_f(t_j,n) = \frac{1}{2}c_{j,0}f + \sum_{k=1}^{M_j} c_{j,k}T_a^k(\mathbf{L})f, \tag{10}$$

$$\widetilde{S}_f(t_j,n) = \frac{1}{2}c_{0,0}f + \sum_{k=1}^{M_0} c_{0,k}T_a^k(\mathbf{L})f. \tag{11}$$

To recover a graph signal f from a given set of wavelet coefficients $c = \widetilde{W}f$, where c includes the wavelet and scaling coefficients $\widetilde{W}_f$ and $\widetilde{S}_f$, and $\widetilde{W} : \mathbb{R}^N \to \mathbb{R}^{N(J+1)}$ represents the approximate SGWT, we use a left-inverse M s.t. $MWf = f$. Since the SGWT is a redundant transform, there is an infinite number of possible inverses. The authors in chose to use the pseudoinverse that solves $(\widetilde{W}^*\widetilde{W})f = \widetilde{W}^*c$, where $\widetilde{W}^*$ is the adjoint operator. For large graphs, the pseudoinverse can be calculated by an iterative conjugate gradient method.

III. Tight Frame Graph Wavelets

A. Tight Frame Design

The authors in highlighted the need to optimize the design of the wavelet generating function $g(\cdot)$ and scaling function $h(\cdot)$ for specific applications. Here, we are interested in constructions that lead to tight (or Parseval) wavelet frames because of their important property of energy conservation between the original and transformed domain and the easy signal reconstruction [20].

Definition 1: A family of wavelets $\{\psi_{t,n}\}_{t\in J, n\in N}$ forms a frame of $\ell_2(V)$, if there exist frame bounds $A, B > 0$ such that

$$\forall f \in \ell_2(V), \quad A\|f\|^2 \le \sum_t \sum_n |\langle f, \psi_{t,n}\rangle|^2 \le B\|f\|^2. \tag{12}$$

If $A = B = 1$, the family $\{\psi_{t,n}\}$ forms a Parseval frame and preserves energy:

$$\forall f \in \ell_2(V), \quad \|f\|^2 = \sum_t \sum_n |\langle f, \psi_{t,n}\rangle|^2. \tag{13}$$

A Parseval frame is the normalized version of a tight frame ($A = B$). Notably, Parseval frames have a simple reconstruction formula where the analysis functions are used at the synthesis side:

$$\forall f \in \ell_2(V), \quad f = \sum_t \sum_n \langle f, \psi_{t,n}\rangle \psi_{t,n}. \tag{14}$$

Thus, the fast approximate SGWT can be used for both the analysis and synthesis steps, avoiding the iterative computation of the pseudoinverse. Note that the proposed analysis/synthesis scheme uses an approximation for both steps, whereas the pseudoinverse proposed in [8] compensates for approximation errors made in the analysis step. Therefore, the polynomial order M_j needs to be chosen high enough in the former case to ensure perfect reconstruction.

We propose to generate Parseval graph wavelet frames by designing the wavelet generating kernel g and scaling function h in the spectral graph domain in analogy to classical definitions in the Fourier frequency domain (Table I, Fig. 1(b)). The J dyadic spatial scales are defined as $t_j = \frac{a}{\lambda_{\max}}2^j$ for $j = 1, \ldots, J$ and the factor $\frac{a}{\lambda_{\max}}$ is chosen to ensure full spectral coverage of the kernels.

All of these constructions are bandlimited functions, and the union of the functions $|g(t_j\lambda)|^2$ and $|h(t_j\lambda)|^2$ forms a partition of unity: $\sum_{j=1}^J |g(t_j\lambda)|^2 + |h(t_J\lambda)|^2 = 1$.

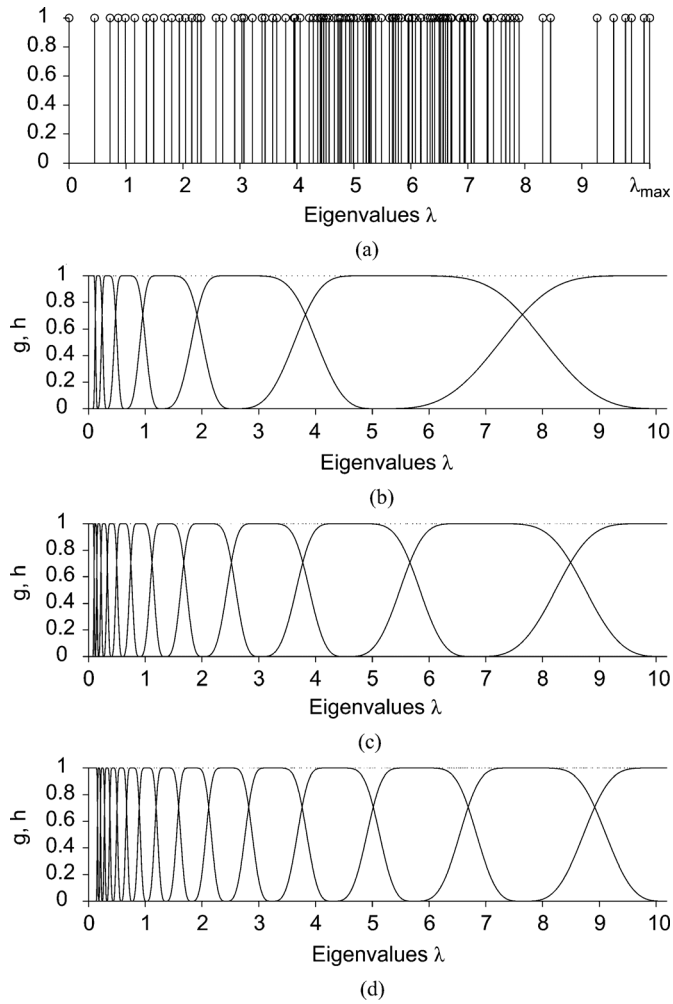

Fig. 1. (a) Eigenvalues of the Swiss Roll graph from [8] with $N = 50$ nodes, and $\sigma = 0.3$. Scaling function h, wavelet generating kernels g, and sum of squares (dotted line) of a Meyer-like rational wavelet frame with $M = 2$ (b), $\frac{3}{2}$ (c) and $\frac{4}{3}$ (d) and $\epsilon = (1+M)^{-1}$.

B. Rational Dilation Factors

The eigenvalues λ_ℓ can be unevenly distributed depending on the nature of the graph (see Fig. 1(a) for an example). For increased flexibility in defining the wavelet subbands, we also provide rational frame designs. Extensions to the rational case of Meyer and Shannon constructions for classical wavelets were originally proposed by Auscher [25] and have further been generalized by [26].

The rational wavelet generating kernel $g(\cdot)$ and scaling function $h(\cdot)$ with rational dilation factor $M = \frac{q+1}{q}$, $q \in \mathbb{Z}$ are defined as follows for a Meyer-like wavelet frame:

$$g(\lambda) = \begin{cases} \sin\left(\frac{\pi}{2}\nu\left(q\left(\frac{\lambda}{a}-1\right)\right)\right) & \forall\lambda \in]a, Ma] \\ \cos\left(\frac{\pi}{2}\nu\left(q\left(\frac{\lambda}{Ma}-1\right)\right)\right) & \forall\lambda \in]Ma, M^2a] \\ 0 & \text{elsewhere} \end{cases} \quad (15)$$

$$h(\lambda) = \begin{cases} 1 & \forall\lambda \in [0, a] \\ \cos\left(\frac{\pi}{2}\nu\left(q\left(\frac{\lambda}{a}-1\right)\right)\right) & \forall\lambda \in]a, Ma] \\ 0 & \text{elsewhere} \end{cases} \quad (16)$$

where $a = (q-\epsilon)b, \epsilon =]0, (1+M)^{-1}]$ and $a, b \in \mathbb{R}^+$. Clearly, the classical dyadic dilation corresponds to $M = 2$. The J wavelet scales are defined as $t_j = \frac{a}{\lambda_{\max}} M^j$ for $j = 1, \ldots, J$. Fig. 1(c) and (d) show the spectra of two Meyer-like rational wavelets.

TABLE I
SPECIFICATION OF WAVELET AND SCALING GENERATING FUNCTIONS THAT LEAD TO PARSEVAL WAVELET FRAMES ($a \in \mathbb{R}^+$)

Wavelet type	Spectral frequency response
Meyer [21]	$g(\lambda) = \begin{cases} \sin\left(\frac{\pi}{2}\nu\left(\frac{\lambda}{a}-1\right)\right) & \forall\lambda \in]a, 2a] \\ \cos\left(\frac{\pi}{2}\nu\left(\frac{\lambda}{2a}-1\right)\right) & \forall\lambda \in]2a, 4a] \\ 0 & \text{elsewhere} \end{cases}$
	$h(\lambda) = \begin{cases} 1 & \forall\lambda \in [0, a] \\ \cos\left(\frac{\pi}{2}\nu\left(\frac{\lambda}{a}-1\right)\right) & \forall\lambda \in]a, 2a] \\ 0 & \text{elsewhere} \end{cases}$
	$\nu(x) = x^4(35 - 84x + 70x^2 - 20x^3)$
Simoncelli [22]	$g(\lambda) = \begin{cases} \cos\left(\frac{\pi}{2}\log_2\left(\frac{\lambda}{2a}\right)\right) & \forall\lambda \in]a, 4a] \\ 0 & \text{elsewhere} \end{cases}$
	$h(\lambda) = \begin{cases} 1 & \forall\lambda \in [0, a] \\ \cos\left(\frac{\pi}{2}\log_2\left(\frac{\lambda}{a}\right)\right) & \forall\lambda \in]a, 2a] \\ 0 & \text{elsewhere} \end{cases}$
Shannon	$g(\lambda) = \begin{cases} 1 & \forall\lambda \in]a, 2a] \\ 0 & \text{elsewhere} \end{cases}$
	$h(\lambda) = \begin{cases} 1 & \forall\lambda \in [0, a] \\ 0 & \text{elsewhere} \end{cases}$
Papadakis [23]	$g(\lambda) = \begin{cases} \sqrt{\frac{1+\sin\left(\frac{3\pi\lambda}{2a}\right)}{2}} & \forall\lambda \in]a, \frac{5}{3}a] \\ 1 & \forall\lambda \in]\frac{5}{3}a, 2a] \\ \sqrt{\frac{1-\sin\left(\frac{3\pi\lambda}{4a}\right)}{2}} & \forall\lambda \in]2a, \frac{10}{3}a] \\ 0 & \text{elsewhere} \end{cases}$
	$h(\lambda) = \begin{cases} 1 & \forall\lambda \in [0, a] \\ \sqrt{\frac{1-\sin\left(\frac{3\pi\lambda}{2a}\right)}{2}} & \forall\lambda \in]a, \frac{5}{3}a] \\ 0 & \text{elsewhere} \end{cases}$
Held [24]	$g(\lambda) = \begin{cases} \cos\left(2\pi\mu\left(\frac{\lambda}{8a}\right)\right) & \forall\lambda \in]a, 2a] \\ \sin\left(2\pi\mu\left(\frac{\lambda}{16a}\right)\right) & \forall\lambda \in]2a, 4a] \\ 0 & \text{elsewhere} \end{cases}$
	$h(\lambda) = \begin{cases} 1 & \forall\lambda \in [0, a] \\ \sin\left(2\pi\mu\left(\frac{\lambda}{8a}\right)\right) & \forall\lambda \in]a, 2a] \\ 0 & \text{elsewhere} \end{cases}$
	$\mu(x) = -1 + 24x - 144x^2 + 256x^3$

C. Conventional and Spectral Graph Wavelets

As stated in Section II-C, the eigenvalues λ_ℓ of the graph Laplacian play the role of frequencies ω. However, the exact relation between spectral and conventional Fourier frequencies was not further deepened in [8]. Here, we want to find this relation. As we have introduced wavelet generating functions from previously proposed wavelet bases, we want to define the mapping $w : \lambda_\ell \to \omega_\ell$ that leads to the classical Meyer wavelet when using the Meyer generating function and a Cartesian grid. We use a cycle graph because it corresponds to an equispaced sampling grid of a one-dimensional signal with periodic boundary conditions, and the eigenvectors of its Laplacian matrix are the basis vectors of the discrete Fourier transform

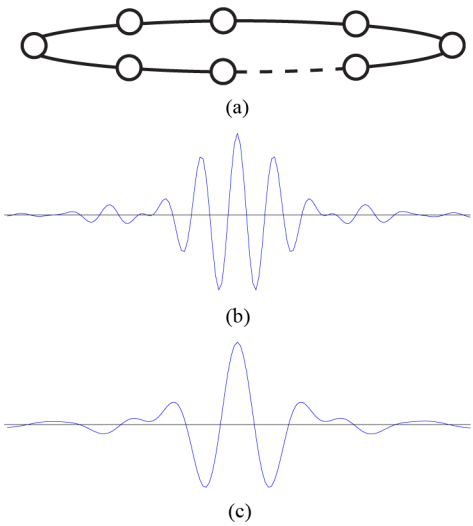

Fig. 2. (a) Schematic of a cycle graph that corresponds to an equispaced sampling grid of a one-dimensional signal; (b) the spectral graph wavelet for the Meyer generating window $g(\lambda_\ell)$ does not equal the Meyer wavelet, (c) but it does for the generating window $g(\omega(\lambda_\ell))$ with $\omega = \arccos(\frac{1-\lambda_\ell}{2})$.

[27] (Fig. 2(a)). Without any mapping w, the spectral graph wavelet does not yield the classical Meyer wavelet (Fig. 2(b)). Intuitively, one would expect a mapping $\lambda_\ell = \omega_\ell^2$ since the Laplacian corresponds to a second-order derivative operator ($\frac{d^2}{dx^2} e^{j\omega x} = -\omega^2 e^{j\omega x}$), but the relation is slightly more subtle. To understand it, we consider the eigenvalues of the cycle graph [19]:

$$\lambda_\ell = 2(1 - \cos(2\pi\ell/N)) \quad \in [0, 4] \tag{17}$$

with $\ell = 0, 1, \ldots, N-1$; N the number of nodes, 2 the degree of the nodes and $2\pi\ell/N = \omega_\ell$ the conventional Fourier frequency. Inversion of (17) leads to the mapping $w : \lambda_\ell \rightarrow \omega_\ell$

$$\omega_\ell = \arccos\left(1 - \frac{\lambda_\ell}{2}\right) \quad \in [0, \pi], \tag{18}$$

which we incorporate in the wavelet generating function as $g(w(\cdot))$. This mapping linearizes the eigenvalues and the spectral graph wavelet yields the classical Meyer wavelet (Fig. 2(c)). Note that the behavior of $\arccos$ around the origin is quadratic, as expected, but that the eigenvalues at the upper part of the Laplacian spectrum are saturated.

Because the spectrum of any graph Laplacian matrix $\mathbf{L}$ or $\widetilde{\mathbf{L}}$ is contained in $[0, 2d_{\max}]$, we will incorporate the mapping

$$\omega_\ell = \arccos(1 - \lambda_\ell / d_{\max}) \tag{19}$$

in the wavelet generating function g and scaling function generator h.

Interestingly, the same relation between spectral and conventional Fourier frequencies has been reported when considering the wave equation on graphs. For the interested reader, we further discuss this result in the appendix.

IV. Wavelets on Multislice Graphs

We propose to construct wavelets on multislice graphs by first estimating eigennetworks of a multislice graph, which capture the variability across slices (Section IV-A), and then, within this eigennetwork construction, apply the SGWT (Section IV-B).

A. Eigenspace of Multislice Graphs

Dynamic networks or networks connected by different types of edges can be represented by a multislice network, where each slice corresponds to one time point or type of interaction (Fig. 3). A multislice graph with N nodes and K different types of edges is described by a third-order adjacency tensor $\mathcal{A} \in \mathbb{R}^{N\times N\times K}$ with the individual adjacency matrices $\mathbf{A}_k \in \mathbb{R}^{N\times N}$ as its frontal slices: $\mathbf{A}_{::k} = \mathbf{A}_k$, for $k = 1, \ldots, K$.

Here, we use tools from multilinear algebra, the HOSVD in particular, to extract orthonormal subspaces from the adjacency tensor, from which "novel" graphs can be constructed that capture the variability across the edge measures.

Proposition 1: The HOSVD of a tensor $\mathcal{A} \in \mathbb{R}^{N\times N\times K}$ with symmetric frontal slices ($\mathbf{A}_k = \mathbf{A}_k^T$ for all $k = 1, \ldots, K$) yields the decomposition

$$\mathcal{A} = \mathcal{S} \times_1 \mathbf{U} \times_2 \mathbf{U} \times_3 \mathbf{V} \tag{20}$$

where $\mathcal{S} \in \mathbb{R}^{N\times N\times K}$ is called the core tensor and $\mathbf{U} \in \mathbb{R}^{N\times N}$ and $\mathbf{V} \in \mathbb{R}^{K\times K}$ are orthonormal mode matrices (Fig. 3). The 1-mode singular values are defined by $\sigma_i^{(1)} = \sqrt{\langle \mathbf{S}_{1::}, \mathbf{S}_{1::} \rangle}$ for $i = 1, .., N$, and are ordered $\sigma_1^{(1)} \geq \sigma_2^{(1)} \geq \cdots \geq \sigma_N^{(1)}$. The singular values for the other two modes are analogous.

The core tensor $\mathcal{S}$ governs the interactions between the factors represented in the mode matrices; i.e., between the nodes and various edge measures, and the mode matrix $\mathbf{U}$ spans a joint subspace of all adjacency matrices.

Definition 2: The core tensor $\mathcal{S}$ can be transformed into an "eigennetwork" tensor $\mathcal{S}'$ by the joint subspace $\mathbf{U}$:

$$\mathcal{S}' = \mathcal{S} \times_1 \mathbf{U} \times_2 \mathbf{U} \tag{21}$$

with $\mathcal{S}' \in \mathbb{R}^{N\times N\times K}$. The individual eigennetworks $\mathbf{S}'_{::k}$ are given by

$$\mathbf{S}'_{::k} = \mathbf{S}_{::k} \times_1 \mathbf{U} \times_2 \mathbf{U} = \mathbf{U}\mathbf{S}_{::k}\mathbf{U}^T \tag{22}$$

with $k = 1, \ldots, K$ and using the slice-wise representation of [28]. The eigennetworks are orthogonal in the sense of the scalar product $\langle \mathbf{S}'_{::i}, \mathbf{S}'_{::j} \rangle = 0$, for $i \neq j$, and symmetric since $\mathcal{S}$ shares the symmetry of $\mathcal{A}$ [16]: ${\mathbf{S}'_{::k}}^T = \left(\mathbf{U}\mathbf{S}_{::k}\mathbf{U}^T\right)^T = \mathbf{U}{\mathbf{S}_{::k}}^T\mathbf{U}^T = \mathbf{U}\mathbf{S}_{::k}\mathbf{U}^T = \mathbf{S}'_{::k}$.

The first eigennetwork $\mathbf{S}'_{::1}$ is the average over all networks and the additional $K-1$ eigennetworks $\mathbf{S}'_{::k}$ capture the variability across them. The idea of extracting eigennetworks from a tensor using HOSVD is similar to work from Vasilescu and Terzopoulos [13] who proposed "TensorFaces", a facial image recognition technique.

The adjacency tensor $\mathcal{A} = \mathcal{S}' \times_3 \mathbf{V}$ can now also be interpreted as a weighted sum of eigennetworks, as illustrated in Fig. 3 [29]:

$$\mathbf{A}_{::k} = \sum_t^T v_{kt} \mathbf{S}'_{::t}. \tag{23}$$

For a graph with a large number of nodes N, it is infeasible to compute the eigennetworks $\mathbf{S}'_{::k}$. However, noting that $\mathcal{S}' = \mathcal{A} \times_3 \mathbf{V}^T$, since $(\mathcal{S}' \times_3 \mathbf{V}) \times_3 \mathbf{V}^T = \mathcal{S}' \times_3 (\mathbf{V}\mathbf{V}^T) = \mathcal{S}'$ [16], we can efficiently compute them in an indirect way if $K \ll N$:

$$\mathbf{S}'_{::k} = \sum_t^T v_{tk} \mathbf{A}_{::t}, \tag{24}$$

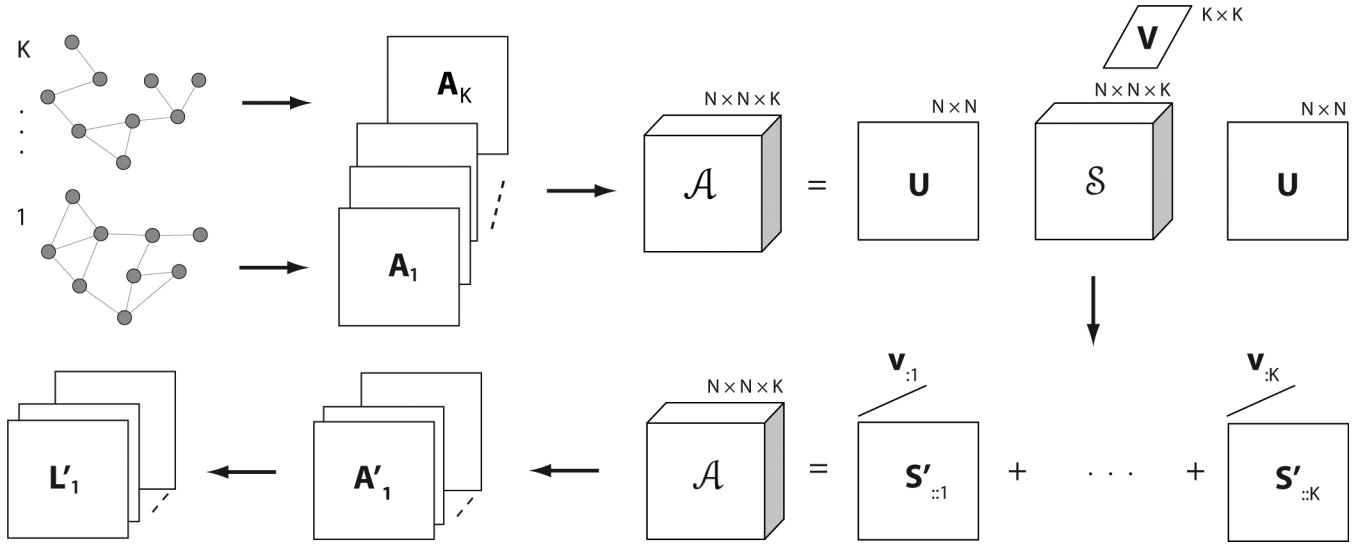

Fig. 3. An adjacency tensor $\mathcal{A}$ is constructed from a stack of K adjacency matrices $\mathbf{A}_k$. The HOSVD decomposes $\mathcal{A}$ into a core tensor $\mathcal{S}$ with symmetric frontal slices and orthonormal matrices $\mathbf{U}$ and $\mathbf{V}$. The decomposition can also be interpreted as a weighted sum of eigennetworks $\mathbf{S}'_{::k} = \mathbf{U}\mathbf{S}_{::k}\mathbf{U}^T$. New adjacency matrices $\mathbf{A}'_m$ are built from linear combinations of the eigennetworks, which each reflect different properties of the underlying network. The adjacency matrices are then converted to Laplacian matrices $\mathbf{L}'_m$ to apply the SGWT.

which is a weighted sum of the frontal slices of $\mathcal{A}$. $\mathbf{V}$ can be directly computed as the left singular matrix of the mode-3 matrix unfolding of $\mathcal{A}$: $\mathbf{A}_{(3)} = \mathbf{V}\Sigma\mathbf{W}^T$. For large graphs it is more efficient to compute $\mathbf{V}$ by decomposing $\mathbf{A}_{(3)}\mathbf{A}_{(3)}^T = \mathbf{V}\Sigma^2\mathbf{V}^T$.

B. Graph Wavelet Design

In many real-world cases, most of the variability in the edge measures can be reasonably assumed to be captured by the first few eigennetworks with the largest singular values. We can thus learn a multilinear subspace of our original adjacency tensor by keeping only the eigennetworks with the largest singular values. Note that the tensor $\hat{\mathcal{A}}$ obtained by discarding the smallest 3-mode singular values is in general not the optimal rank-N, N, κ ($\kappa < K$) approximation of a tensor $\mathcal{A}$, because in HOSVD the mode matrices are optimized for each mode separately, neglecting interactions among them. However, the truncated HOSVD is a good approximation with the error bounded by $\|\mathcal{A} - \hat{\mathcal{A}}\|^2 \leq \sum_{i=\kappa+1}^{K}(\sigma_i^{(3)})^2$ and is much faster than the otherwise required iterative algorithm [30].

We propose to design adjacency matrices $\mathbf{A}'_m$, $m = 1, \ldots, M$, that capture different properties of the underlying networks by combining the eigennetworks:

$$\mathbf{A}'_m = \sum_{k=1}^{\kappa} \alpha_k^{(m)} \mathbf{S}'_{::k} \quad \text{s.t. } \mathbf{A}'_m \geq \mathbf{0} \tag{25}$$

where $\alpha_k^{(m)} \in \mathbb{R}$. The SGWT is only defined for graphs with nonnegative weights and the set of $\alpha_k^{(m)}$ must therefore be chosen such that $\mathbf{A}'_m \geq \mathbf{0}$. Since $\mathbf{S}'_{::1}$ is the average graph, or "backbone" of the multislice graph, and only all-postiive or all-negative eigennetwork, $\alpha_1^{(m)}$ must be $\neq 0$.

The weights $\alpha_k^{(m)}$ can also be obtained by a constrained least-squares fitting problem, which exploits the weights in $\mathbf{V}$ associated with each eigennetwork k:

$$\min_{\alpha_k^{(m)}} \left\| \sum_{k=1}^{\kappa} \alpha_k^{(m)} \mathbf{v}_{::k} - \mathbf{g} \right\|^2 \quad \text{s.t. } \mathbf{A}'_m \geq \mathbf{0} \tag{26}$$

where $\mathbf{g}$ is a pre-defined variation between the slices of the network that we like $\mathbf{A}'_m$ to capture. This strategy will not be further explored here.

After calculating $\mathbf{L}'_m$, we can then use the SGWT or its approximation for large graphs to construct wavelets.

V. Experimental Results

A. Graph Wavelets on Multiplex Graphs

An image can be interpreted as a grid graph where each pixel is a node that can be connected to neighboring pixels/nodes according to chosen properties. To illustrate the potential of our design to extract meaningful information from multislice graphs, we constructed a tensor from 4 different grid graphs. Each node was connected to at most 6 neighbors along different directions, and two neighboring nodes were disconnected if their intensity values in the image differed by more than 40. The image was pre-smoothed prior to estimating the adjacency matrix to better approximate the gradient operator (Gaussian kernel, $\sigma = 1$) [31]. The slices of the adjacency tensor $\mathcal{A}$ were ordered as shown in Fig. 4.

We obtained 4 eigennetworks and associated weights $\mathbf{V}$ from the HOSVD of $\mathcal{A}$. The columns of $\mathbf{V}$ capture the variability across mode 3 of $\mathcal{A}$ (i.e., the direction of the edges):

$$\begin{bmatrix} -0.47 & 0.73 & 0 & 0.50 \\ -0.49 & -0.69 & 0 & 0.54 \\ -0.52 & -0.01 & 0.71 & -0.47 \\ -0.52 & -0.01 & -0.71 & -0.49 \end{bmatrix}$$

with associated singular values $\sigma_1^{(3)} = 97, \sigma_2^{(3)} = 48, \sigma_3^{(3)} = 31, \sigma_4^{(3)} = 18$. We observe that the first column is the average of all slices, the second one contrasts the horizontal and vertical edges (large values in row 1 and 2), the third one the diagonal ones, and the last one rectangular and diagonal ones.

In Fig. 5 we show 5 "directional" graph wavelets that were implemented by combining the eigennetworks in different

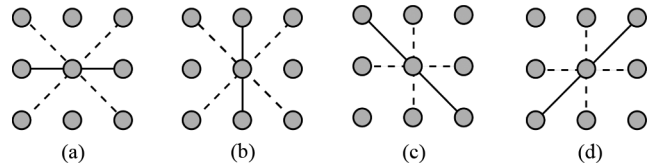

Fig. 4. Neighborhoods of 4 different grid graphs. Dashed lines indicate edge weights that were penalized by a factor $1/2$ so as to enhance one direction: (a) horizontal, (b) vertical, (c) diagonal top left to bottom right, (d) diagonal bottom left to top right.

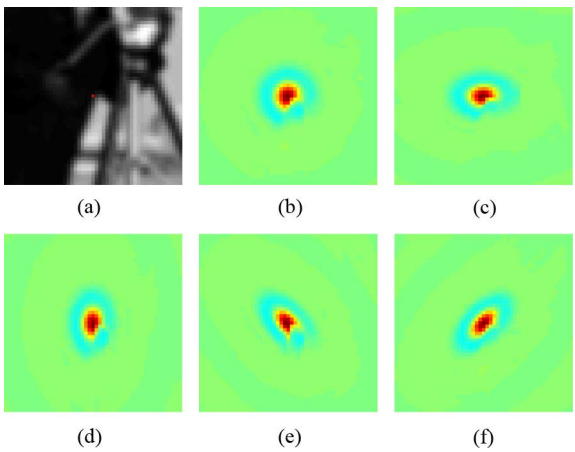

Fig. 5. Graph wavelets constructed from various combinations of eigennetworks diffuse along different directions: (a) location of the wavelet on the smoothed cameraman image, (b) $\mathbf{A}'_1 = -\mathbf{S}'_{::1}$: isotropic for the average graph, (c) $\mathbf{A}'_2 = -0.5\mathbf{S}'_{::1} + 0.7\mathbf{S}'_{::2}$: horizontal, (d) $\mathbf{A}'_3 = -0.5\mathbf{S}'_{::1} - 0.7\mathbf{S}'_{::2}$: vertical, (e) $\mathbf{A}'_4 = -0.5\mathbf{S}'_{::1} + 0.8\mathbf{S}'_{::3}$: diagonal, (f) $\mathbf{A}'_5 = -0.5\mathbf{S}'_{::1} - 0.8\mathbf{S}'_{::3}$: diagonal. The first frontal slice $\mathbf{S}'_{::1}$ was multiplied by $-0.5 = 1/\sum_i \mathbf{v}_{i1}$, which turns it into the average graph. Wavelets are shown at a coarse scale.

ways. Whereas wavelets constructed using the average graph diffuse homogeneously (i.e., there is no directional information, Fig. 5(b)), they diffuse preferentially along a certain direction when the average graph is combined with one of the other 3 eigennetworks, making them sensitive to different types of orientations in an image (Figs. 5(c)–(f)). Since neighboring nodes were only connected if their intensity values were sufficiently close, the wavelets diffuse along underlying boundaries, such as the coat.

Fig. 6 shows two graph wavelet decompositions of the cameraman image of size 128×128, where the SGWT was built from "horizontal" $\mathbf{A}'_2$ and "vertical" $\mathbf{A}'_3$ (wavelets shown in Figs. 5(c) and 5(d)). The energy of the wavelet coefficients is high for different types of features of the image: At the finest scale, which detects edges not expected by the underlying graph topology, the energy for the "horizontal" transform is high around vertical edges, such as the tripod, and for the "vertical" transform around horizontal edges, such as the camera.

For a graph corresponding to a Cartesian grid and when using the mapping introduced in III-C, we know that the spectral graph wavelets are equivalent to classical wavelets. The advantage of using graph wavelets, however, is that they are not restricted to Cartesian grids. Weighted graphs provide a flexible model to describe irregular data domains and the wavelets can thus be adapted to many different applications. Here, we have used this

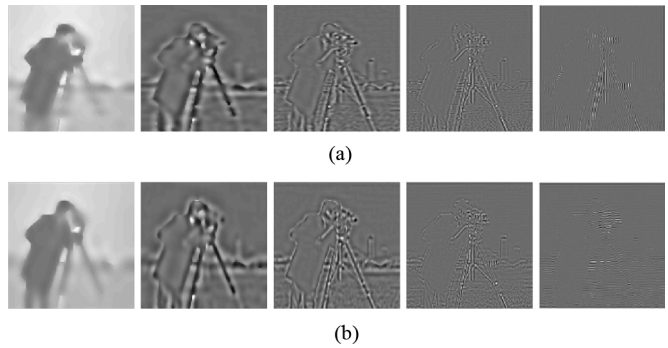

Fig. 6. Wavelet coefficients of the cameraman image (128×128 pixels), where the graph wavelet transforms were built from (a) $\mathbf{A}'_2$, (b) $\mathbf{A}'_3$. $J = 4$ scales, Simoncelli-like design, and scaling and coarse to fine scale wavelet coefficients from left to right.

flexibility to incorporate knowledge about directionality and underlying boundaries. With the proposed approach for multislice graphs, we also not only have access to an average graph ($\mathbf{S}'_{::1}$), which corresponds to the common way of handling multislice graphs, but several eigennetworks that capture the variability across the direction of the edges.

B. Graph Wavelets on Dynamic Graphs

Cognitive function and dynamic adaptation emerge through the integrated activity of many brain regions. Using functional magnetic resonance imaging (fMRI), functionally connected brain regions are identified as those exhibiting temporal dependence in their activity patterns. Studies representing the brain as a static, functionally connected network have already led to unique insights into the organization of the brain, but understanding how the architecture of the brain varies with cognitive function or across subjects is a crucial frontier [32], [33]. Chang and Glover [34] were among the first to highlight that functional connectivity is not stationary throughout an fMRI scanning session (with a focus on one brain region), and Bassett *et al.* [2] recently studied dynamic reconfigurations in the modular organization of brain networks during learning.

We explored the potential of our proposed framework to capture dynamic changes of the human brain network in a previously studied dataset [35], [36]. It was shown that average functional connectivity differs between a "resting" and "movie watching" cognitive state and enough so that a subject's cognitive state could be predicted using a classifier algorithm [36]

The fMRI data was acquired from 15 healthy subjects during a simple block design, which consisted of alternating cycles of watching a short movie excerpt (40–60 sec), followed by a period where subjects were instructed to rest with eyes closed and let their mind freely wander (90 sec) and ending with a beep sound that instructed the subject to answer a short question about the content of their thoughts [35].[1] The fMRI data was realigned using SPM8 and parcellated into $N = 90$ regions corresponding to the Automated Anatomic Labeling atlas [37]

[1]Data was acquired on a Siemens 3T Trio TIM in two scanning runs, TR/TE/FA $= 1.1\ \mathrm{s}/27\ \mathrm{ms}/90^\circ$, matrix $= 64 \times 64$, voxel size $= 3.75 \times 3.75 \times 4.2\ \mathrm{mm}^3$, 21 contiguous transverse slices, 1.05 mm gap, 2598 volumes.

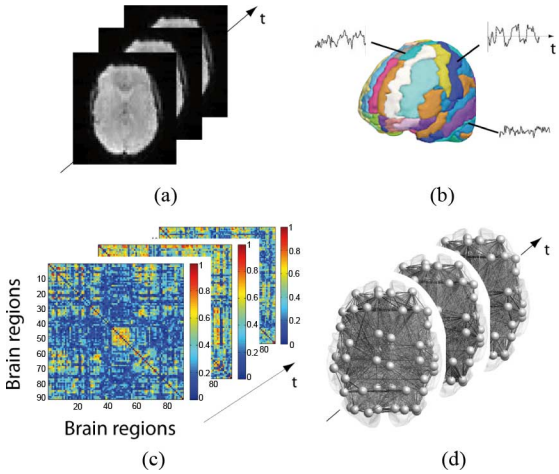

Fig. 7. fMRI analysis pipeline: (a) raw fMRI data; (b) parcellation of the brain into regions (network nodes), and estimation of their regional spatially-averaged activity; (c) to estimate changes in network architecture over time, we computed the temporal correlation between the regional activity, averaged across subjects, of all pairs of regions using a sliding window approach (weighted network edges); (d) and obtained $K = 125$ functional brain networks.

(Fig. 7). Two regions were excluded because of signal dropout, leaving 88 regions. We estimated regional mean activity by averaging voxel time series in each brain region, and averaged across all subjects. We constructed dynamic whole-brain functional networks by estimating temporal correlations between all pairs of brain regions using a sliding window approach with a window length of 30 scans. The absolute values of the correlation measures for each pair of regions were used as edge weights to construct $K = 125$ weighted graphs.

In this example, mode 3 captures the variability of the edges across time and the eigennetworks hence contain information on network dynamics. Fig. 8 shows eigennetworks $\mathbf{S}'_{::k}$ for $k = 1, .., 4, 100$ and their associated weights $\mathbf{v}_{:k}$. We observe well-known functional clusters in the average brain network $\mathbf{S}'_{::1}$ (Fig. 8(a)), such as the interconnected visual and auditory processing systems (nodes $43 - 56$ and $79 - 86$), the frontal attention network (nodes $7, 8, 29-34$), somato-motor cortex network (nodes $1, 2, 19, 20, 57, 58$) and language-related areas (nodes $11 - 14$).

The other eigennetworks represent directions of variation from the average network across time. The second and third eigennetworks capture variability associated with the experimental paradigm, as can be seen form comparing $\mathbf{v}_{:2}$ and $\mathbf{v}_{:3}$ to the paradigm (Figs. 8(b) and 8(c)). Since the sign of $\mathbf{v}_{:2}$ is positive during the resting condition, edges with red color have larger weights during the resting condition, and those with blue color have larger weights during the transitions of the movies condition. For example, connectivity in the visual cortex is stronger during the resting condition, as previously observed [36], and in frontal regions during the transitions of the movies condition.

As is the case when applying PCA or independent component analysis to noisy data, some of these components will represent noise (e.g., Fig. 8(e) depicting component 100). Discarding some of the eigennetworks could thus also help to reduce noise in a multislice graph.

To explore the multislice SGWT, we added and subtracted $\mathbf{S}'_{::2}$ from $\mathbf{S}'_{::1}$ to obtain the adjacency matrices $\mathbf{A}'_1$ and $\mathbf{A}'_2$, respectively (Figs. 9(a), 9(b), 9(f), 9(g)). From $\mathbf{v}_{:2}$ we know that $\mathbf{A}'_1$ enhances edges typical for the resting condition, and $\mathbf{A}'_2$ for the transitions of the movies condition. $\mathbf{A}'_1$ has strongly connected visual and auditory cortices and a wavelet localized in the primary visual cortex diffuses to visual and auditory regions at coarser scales, as shown in Figs. 9(c)–9(e). In contrast, as expected from the reduced connectivity within the visual cortex in $\mathbf{A}'_2$, the wavelet diffuses only within the primary visual cortex at coarser scales (nodes $43 - 48$), but not to extrastriate areas (nodes $49 - 56$), as shown in Figs. 9(h)–9(j).

Instead of adding and subtracting $\mathbf{S}'_{::2}$ from $\mathbf{S}'_{::1}$, we could also have explored the strategy proposed in (26) with $\mathbf{g}$ equal to the stimulus paradigm to find the combination of eigennetwork weights that best approximates the stimulus paradigm.

We then implemented two SGWTs built from $\mathbf{A}'_1$ (resting) and $\mathbf{A}'_2$ (movies), respectively, and applied them to the normalized regional brain activity. We calculated the total energy of the wavelet coefficients at each scale and for each decomposition by summing over all brain regions. Due to energy conservation of the tight frame design, the ℓ^2-norm of the wavelet coefficients across all scales sums up to 1.

Fig. 10 shows the difference of the energy of the scaling (red line) and wavelet coefficients (blue dashed line), where a positive value indicates that the energy is higher for the resting than for the movies transform. We observe that at the finest scale, the energy is increased during the condition not well represented by the graph, i.e., the energy of the resting transform is larger than the movies transform during the movies conditions (the blue line goes up). In previous work we estimated two separate brain networks for each condition (though not in a data-driven manner) and also observed an increase of energy when the SGWT was not adapted to the data [38]. For the scaling coefficients the picture is reversed, i.e., the energy of the resting transform is smaller than the movies transform during the movies condition (red line goes down).

Our results suggest that the variability identified by the HOSVD and illustrated with the eigennetworks, corresponds to relevant changes in network architecture across cognitive states. The column vectors of $\mathbf{V}$ give additional information regarding the importance of these networks across time, and the proposed multislice SGWT provides a meaningful decomposition of fMRI data.

VI. Conclusion

We have extended the previously proposed SGWT to a tight frame, shown the link between classical and spectral graph wavelets, and presented a novel framework for the design of wavelets on multislice graphs based on the HOSVD. The proposed approach proceeds in two steps, by first estimating eigennetworks from an adjacency tensor and then building adjacency matrices that exploit the variability across slices of the network. Only then are these adjacency matrices converted to Laplacian matrices to apply the SGWT. The proposed framework of estimating eigennetworks and their combination is therefore not limited to the SGWT, but can be applied to

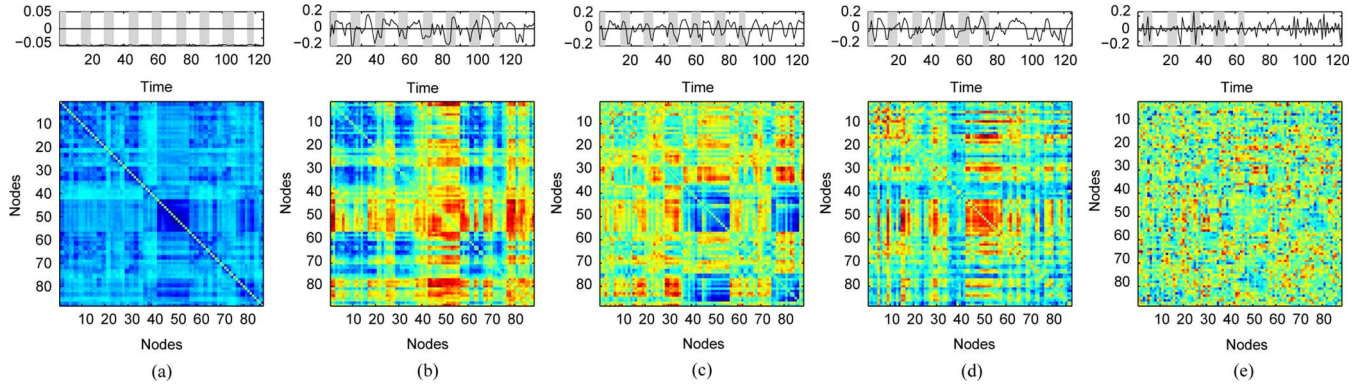

Fig. 8. Eigennetworks $\mathbf{S}'_{::k}$ for $k = 1, \ldots, 4, 100$. The top plots show the corresponding column $\mathbf{v}_{:k}$ overlaid on the stimulus paradigm (gray-shaded areas indicate the movie condition). The eigennetworks show the edge weights between 88 brain regions, ordered by lobe (frontal, limbic, occipital, parietal, subcortical, temporal) and with homologous regions adjacent to each other (e.g., left precentral cortex followed by right precentral cortex). The colorbars are symmetric around zero, i.e., red colors indicate positive edge weights, blue colors negative ones.

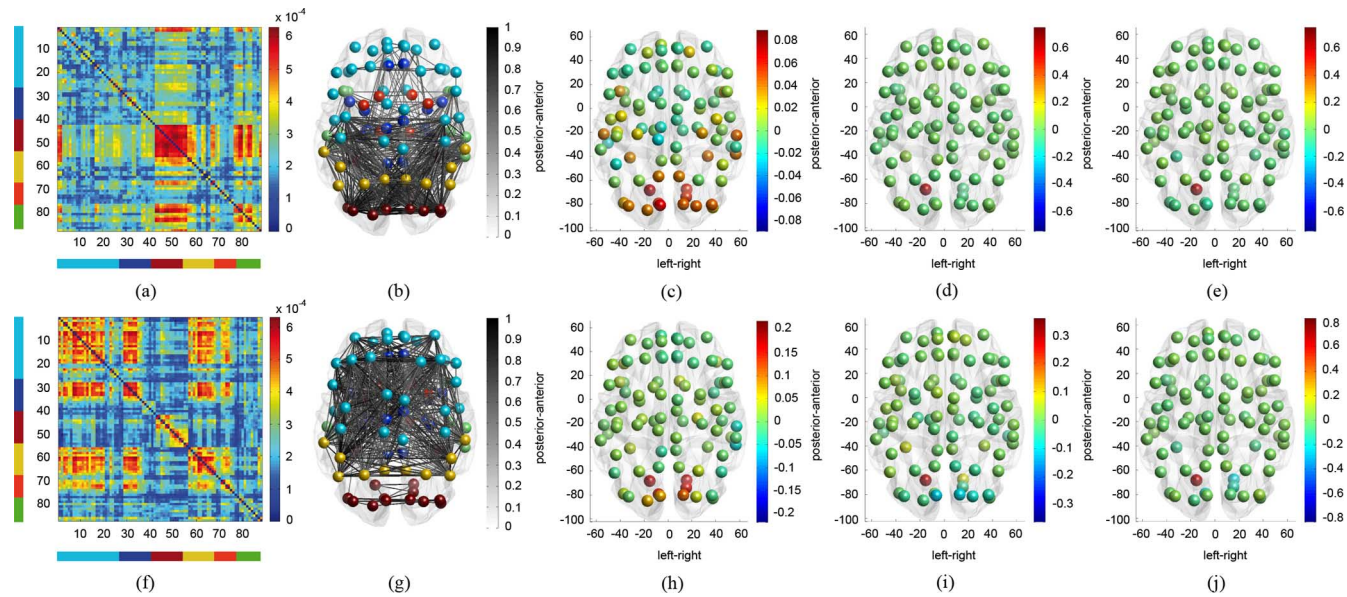

Fig. 9. (a) Adjacency matrix of the 88×88 brain graph built from the first and second eigennetworks, $\mathbf{A}'_1 = -0.1\mathbf{S}'_{::1} + 0.2\mathbf{S}'_{::2}$; (b) visualization of the graph in brain space, with nodes colored according to their anatomical location (light blue = frontal lobe, dark blue = limbic, red = occipital, yellow = parietal, orange = subcortical, green = temporal); (c-e) the energy of scaling and wavelet coefficients from coarse to fine scale, $J = 2$ scales and a Simoncelli-like design; (f) adjacency matrix of the graph built from $\mathbf{A}'_2 = -0.1\mathbf{S}'_{::1} - 0.3\mathbf{S}'_{::2}$; (g-j) as (b-e) and $-0.1 = 1/\sum_i \mathbf{v}_{i1}$.

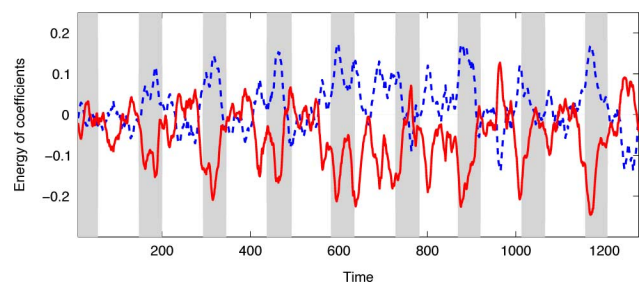

Fig. 10. Difference of the energy of the wavelet coefficients of the fMRI data decomposed with the SGWTs built from $\mathbf{A}'_1$ (resting condition) and $\mathbf{A}'_2$ (transitions of movie condition), $J = 2$ scales and a Simoncelli-like frame design. Difference of the energy for the scaling (red), and finest scale (blue dashed), where a positive value indicates a higher energy in the "resting" SGWT. A moving average with equal window length as for calculating the dynamic connectivity (i.e., 30) was applied, and gray-shaded areas indicate the movie condition.

other graph wavelet transforms [9] or discrete signal processing frameworks [39] that are based on the adjacency matrix. Our results suggest that eigennetworks obtained from the HOSVD of a collection of adjacency matrices capture relevant variability across graph edges and that they provide meaningful graph wavelet transforms. In contrast to the common method of averaging or summing across multiple slices of a graph, the proposed framework allows to exploit the variability across them.

Appendix
Wave Equation on Graphs

The relation $\omega_\ell = \arccos(1 - \lambda_\ell)$ between conventional Fourier frequencies and eigenvalues of the normalized Laplacian $\widetilde{\mathbf{L}}$ (i.e., $d_{\max} = 1$) has also been found when considering the wave equation on complete and binary graphs [40]–[42]:

$$\frac{\partial^2 u}{\partial t^2} = -\Delta u, \tag{27}$$

where Δ is a so-called edge-based Laplacian that acts as an operator on functions on the edges (in contrast to the conventional graph Laplacian that acts on signals defined on the nodes). The edge-based wave equation has a physical interpretation, where the edges are taut strings that are fused together at the nodes. While the node-based wave equation has infinite wave propagation speed, the edge-based wave equation has unit propagation speed [42].

From a purely signal-processing point-of-view, we can consider the wave equation (27) with the normalized graph Laplacian and discretized time. After applying the discrete temporal Fourier transform to (27), we retrieve the left-hand side as $(\cos(\omega) - 1)u$. The equivalence with the eigendecomposition $\Delta u = \lambda u$ of the right-hand side then leads to $1 - \cos(\omega) = \lambda$ with all eigenvalues contained in $[0, 2]$.

ACKNOWLEDGMENT

The authors would like to thank Dr. H. Eryilmaz, Prof. P. Vuilleumier and Prof. S. Schwartz for sharing their fMRI data, Dr. J. Richiardi for help with the preprocessing and the anonymous reviewers for their valuable comments and suggestions. A preliminary version of the tight frame design for the SGWT was presented at [38].

REFERENCES

[1] P. J. Mucha, T. Richardson, K. Macon, M. A. Porter, and J.-P. Onnela, "Community structure in time-dependent, multiscale, and multiplex networks," *Science*, vol. 328, no. 5980, pp. 876–878, May 2010.

[2] D. S. Bassett, N. F. Wymbs, M. A. Porter, P. J. Mucha, J. M. Carlson, and S. T. Grafton, "Dynamic reconfiguration of human brain networks during learning," *Proc. Natl. Acad. Sci. USA*, vol. 108, no. 18, pp. 7641–7646, May 2011.

[3] M. Szell, R. Lambiotte, and S. Thurner, "Multirelational organization of large-scale social networks in an online world," *Proc. Natl. Acad. Sci. USA*, vol. 107, no. 31, pp. 13636–13641, Jul. 2010.

[4] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin, "Catastrophic cascade of failures in interdependent networks," *Nature*, vol. 464, no. 7291, pp. 1025–1028, Apr. 2010.

[5] X. Dong, P. Frossard, P. Vandergheynst, and N. Nefedov, "Clustering with multi-layer graphs: A spectral perspective," *IEEE Trans. Signal Process.*, vol. 60, pp. 5820–5831, Nov. 2012.

[6] M. Jansen, G. P. Nason, and B. W. Silverman, "Multiscale methods for data on graphs and irregular multidimensional situations," *J. R. Stat. Soc. Ser. B Stat. Methodol.*, vol. 71, no. 1, pp. 97–125, Jan. 2009.

[7] I. Ram, M. Elad, and I. Cohen, "Generalized tree-based wavelet transform," *IEEE Trans. Signal Process.*, vol. 59, pp. 4199–4209, Sep. 2011.

[8] D. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Appl. Comput. Harmon. Anal.*, vol. 30, pp. 129–150, Mar. 2011.

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

[10] S. K. Narang and A. Ortega, "Perfect reconstruction two-channel wavelet filter banks for graph structured data," *IEEE Trans. Signal Process.*, vol. 60, pp. 2786–2799, Jun. 2012.

[11] R. Bro, "Parafac tutorial and applications," *Chem. Intell. Lab. Syst.*, vol. 38, pp. 149–171, 1997.

[12] C. J. Appellof and E. R. Davidson, "Strategies for analyzing data from video fluorometric monitoring of liquid chromatographic effluents," *Anal. Chem.*, vol. 53, no. 13, pp. 2053–2056, 1981.

[13] M. A. O. Vasilescu and D. Terzopoulos, "Multilinear analysis of image ensembles: Tensorfaces," in *Proc. Eur. Conf. Comput. Vis.*, Copenhagen, 2002, pp. 447–460.

[14] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, Aug. 2009.

[15] L. Tucker, "Some mathematical notes of three-mode factor analysis," *Psychometrika*, vol. 31, pp. 279–311, 1966.

[16] L. De Lathauwer, B. De Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 21, pp. 1253–1278, Apr. 2000.

[17] J. D. Carroll and J.-J. Chang, "Analysis of individual differences in multidimensional scaling via an n-way generalization of eckart-young decomposition," *Psychometrika*, vol. 35, pp. 283–319, 1970.

[18] R. Harshman, "Foundations of the PARAFAC procedure: Models and conditions for an explanatory multi-modal factor analysis," *UCLA Working Papers in Phonetics*, vol. 16, no. 1, pp. 1–84, 1970.

[19] F. Chung, *Spectral Graph Theory*. Providence, RI, USA: AMS, 1997.

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

[21] Y. Meyer, "Principe d'incertitude, bases Hilbertiennes et algèbres d'opérateurs," in *Seminaire Bourbaki* (in French), 1986, vol. 662, pp. 209–223.

[22] J. Portilla and E. P. Simoncelli, "A parametric texture model based on joint statistics of complex wavelet coefficients," *Int. J. Comput. Vis.*, vol. 40, pp. 49–70, Oct. 2000.

[23] J. Romero, S. Alexander, S. Baid, S. Jain, and M. Papadakis, "The geometry and the analytic properties of isotropic multiresolution analysis," *Adv. Comput. Math.*, vol. 31, pp. 283–328, 2009.

[24] S. Held, M. Storath, P. Massopust, and B. Forster, "Steerable wavelet frames based on the Riesz transform," *IEEE Trans. Image Process.*, vol. 19, pp. 653–667, Mar. 2010.

[25] P. Auscher, "Wavelet bases for $L^2(\mathbb{R})$ with rational dilation factor," in *Wavelets and Their Applications*. Boston, MA, USA: Jones and Bartlett, 1992, pp. 439–451.

[26] A. Baussard, F. Nicolier, and F. Truchetet, "Rational multiresolution analysis and fast wavelet transform: application to wavelet shrinkage denoising," *Signal Process.*, vol. 84, no. 10, pp. 1735–1747, Aug. 2004.

[27] G. Strang, "The discrete cosine transform," *SIAM Rev.*, vol. 41, no. 1, pp. 135–147, Aug. 1999.

[28] R. Harshman and S. Hong, "'Stretch' vs 'slice' methods for representing three-way structure via matrix notation," *J. Chemometrics*, vol. 16, no. 4, pp. 198–205, Apr. 2002.

[29] L. Eldén, *Matrix Methods in Data Mining and Pattern Recognition SIAM*. Philadelphia, PA, USA: SIAM, 2007.

[30] L. De Lathauwer, B. De Moor, and J. Vandewalle, "On the best rank-1 and rank-$R_1, R_2, \ldots, R_N$) approximation of higher-order tensors," *SIAM J. Matrix Anal. Appl.*, vol. 21, no. 4, pp. 1324–1342, Mar. 2000.

[31] S. Di Zenzo, "A note on the gradient of a multi-image," *Comput. Vis. Graph. Image Process.*, vol. 33, no. 1, pp. 116–125, Jan. 1986.

[32] O. Sporns, *Networks of the brain*. Cambridge, MA, USA: MIT Press, 2010.

[33] D. S. Bassett and M. S. Gazzaniga, "Understanding complexity in the human brain," *Trends Cogn. Sci.*, vol. 15, no. 5, pp. 200–209, May 2011.

[34] C. Chang and G. H. Glover, "Time-frequency dynamics of resting-state brain connectivity measured with fMRI," *NeuroImage*, vol. 50, no. 1, pp. 81–98, Mar. 2010.

[35] H. Eryilmaz, D. V. D. Ville, S. Schwartz, and P. Vuilleumier, "Impact of transient emotions on functional connectivity during subsequent resting state: A wavelet correlation approach," *NeuroImage*, vol. 54, no. 3, pp. 2481–2491, Oct. 2011.

[36] J. Richiardi, H. Eryilmaz, S. Schwartz, P. Vuilleumier, and D. V. D. Ville, "Decoding brain states from fMRI connectivity graphs," *NeuroImage*, vol. 56, no. 2, pp. 616–626, Jun. 2011.

[37] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, and M. Joliot, "Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain," *NeuroImage*, vol. 15, no. 1, pp. 273–289, Jan. 2002.

[38] N. Leonardi and D. Van De Ville, "Wavelet frames on graphs defined by fMRI functional connectivity," in *Proc. IEEE Int. Symp. Biomed. Imag.*, Chicago, IL, USA, 2011, pp. 2136–2139.

[39] A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs," *IEEE Trans. Signal Process.*, vol. 61, no. 7, pp. 1644–1656, Apr. 2013.

[40] J. von Below, "A characteristic equation associated to an eigenvalue problem on c^2-networks," *Linear Algebra Appl.*, vol. 71, pp. 309–325, 1985.
[41] C. Cattaneo, "The spectrum of the continuous Laplacian on a graphs," *Monatshefte fuer Mathematik*, vol. 124, pp. 215–235, 1997.
[42] J. Friedman and J.-P. Tillich, "Wave equations for graphs and the edge-based Laplacian," *Pacific J. Math.*, vol. 216, pp. 229–266, 2004.

Nora Leonardi (S'13) received the M.Sc. degree in biomedical engineering from the École Polytechnique Fédérale de Lausanne (EPFL) in 2009. She currently works towards a Ph.D. degree at the Medical Image Processing Lab, which is affiliated with the EPFL and the University of Geneva, Switzerland.

Her research interests include wavelets, graphs, and their applications in functional magnetic resonsance imaging.

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 Biomedical Imaging Group of Prof. Michael Unser at theÉcole 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 dÕImagerie Biomédicale (CIBM). In 2009, he has been the recipient of a Swiss National Science Foundation professorship and currently holds a joint position at the University of Geneva, Switzerland, and the EPFL. His research interests include wavelets, sparsity, pattern recognition, and their applications in biomedical imaging, such as functional magnetic resonance imaging. He received the Pfizer Research Award 2012 in the category Neurosciences for his work on scale-free dynamics of EEG microstates.

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 is the Chair of the Bio Imaging and Signal Processing (BISP) TC of the IEEE Signal Processing Society (2012-2013). Since 2003, he has also been an Editor and Webmaster of The Wavelet Digest. He is Co-Chair of the biennial Wavelets & Sparsity conferences, together with V. Goyal and M. Papadakis.