

Diffusion-adapted spatial filtering of fMRI data for improved activation mapping in white matter

David Abramian and Martin Larsson



LUND
UNIVERSITY

MASTER'S THESIS IN
Biomedical Engineering
Lund, September 2017

Faculty of Engineering LTH
Department of Biomedical Engineering

Supervisor: Hamid Behjat

Abstract

Brain activation mapping using fMRI data has been mostly focused on finding detections in gray matter. Activations in white matter are harder to detect due to anatomical differences between both tissue types, which are rarely acknowledged in experimental design. However, recent publications have started to show evidence for the possibility of detecting meaningful activations in white matter. The shape of the activations arising from the BOLD signal is fundamentally different between white matter and gray matter, a fact which is not taken into account when applying isotropic Gaussian filtering in the preprocessing of fMRI data. We explore a graph-based description of the white matter developed from diffusion MRI data, which is capable of encoding the anisotropic domain. Based on this representation, two approaches to white matter filtering are tested, and their performance is evaluated on both semi-synthetic phantoms and real fMRI data. The first approach relies on heat kernel filtering in the graph spectral domain, and produced a clear increase in both sensitivity and specificity over isotropic Gaussian filtering. The second approach is based on spectral decomposition for the denoising of the signal, and showed increased specificity at the cost of a lower sensitivity.

Acknowledgments

We would like to extend our gratitude to our supervisor Hamid Behjat, who proposed this subject for our thesis and managed to put up with us until it was completed.

David Abramian
Martin Larsson

Acronyms

AD axial diffusivity.

BOLD blood oxygenation level dependent.

DFT discrete Fourier transform.

dMRI diffusion MRI.

DSI diffusion spectrum imaging.

DTI diffusion tensor imaging.

FA fractional anisotropy.

fMRI functional MRI.

FWHM full width half maximum.

GLM general linear model.

GQI generalized Q-sampling imaging.

GRFT Gaussian random field theory.

gSPM graph-based SPM.

HCP Human Connectome Project.

HRF hemodynamic response function.

MD mean diffusivity.

MNI Monteval Neurological Institute.

MRI magnetic resonance imaging.

ODF orientation distribution function.

QBI Q-ball imaging.

RD radial diffusivity.

ROC receiver operating characteristic.

rsfMRI resting state fMRI.

sgWSPM spectral graph WSPM.

SNR signal-to-noise ratio.

SPM statistical parametric mapping.

UMT uniform Meyer-type.

WSPM wavelet SPM.

Contents

Abstract	iii
Acknowledgments	v
Acronyms	viii
1 Introduction	1
1.1 Goal	2
1.2 Contributions	2
1.3 Previous work	3
1.4 Disposition	4
2 Background	5
2.1 Graphs and their spectra	5
2.2 The brain	11
2.3 Functional MRI	12
2.4 Diffusion MRI	17
3 Methodology	23
3.1 White matter graph design	23
3.2 The heat kernel	24
3.3 Signal-adapted filter design	26
3.4 B-spline system of spectral kernels	28
3.5 Improving the performance of sgWSPM	30

4	Datasets	35
4.1	HCP Data	35
4.2	Preprocessing steps	36
4.3	Semi-synthetic data	36
5	Results	39
5.1	Diffusion adapted atoms	39
5.2	Signal-adapted filter design	44
5.3	Performance improvements to sgWSPM	46
5.4	Experimental results	47
6	Discussion	57
6.1	Performance improvements to sgWSPM	57
6.2	Semi-synthetic results	58
6.3	Real data results	58
6.4	Signal-adapted kernels	58
6.5	Conclusions	59
6.6	Future work	60
A	Appendix	63
A.1	The heat kernel and its Gaussianity	63
	References	68

Chapter 1

Introduction

The capacity to study the activity of the brain in a non-invasive way has brought great advances in the understanding of its inner workings. Magnetic resonance imaging (MRI) constitutes a set of techniques with versatile application that are capable of revealing different facets of the brain picture. In particular, functional MRI (fMRI) is used to probe the activity of the brain both during the performance of specific tasks and in a resting state, uncovering functional specialization in brain regions and their collaborative activity in functional networks. Since its development in the early 1990s [1], fMRI has been instrumental as a tool of neuroscience, and has helped in providing answers to numerous questions related to the “wheres” and “hows” of neural activity.

Up to this day, fMRI studies have overwhelmingly had a focus on determining activations in the gray matter of the brain, and disregarded white matter activations as noise. It has also been suggested that the vascularization density in white matter may be so low as to render fMRI activity undetectable [2]. However, the number of studies reporting white matter activations, although relatively small, has been steadily increasing, and a case for measurable and meaningful white matter activations has started to form, thoroughly presented in [3].

Clearly, the relative ease with which activations can be detected in gray or white matter is directly related to their anatomical and physiological differences, such as the significantly higher blood vessel density in gray matter [4], a differently shaped hemodynamic response function (HRF) [5] and different optimal acquisition parameters [6]. Nevertheless, these differences alone do not represent direct evidence against the possibility of detecting functional activity in white matter. On the other hand, a failure to recognize these differences in the design of experiments may be at least partially to blame for the scarcity of reports of white matter activations.

In this work we aim to present and address a further difference between white and gray matter that could serve to increase the amount and quality of the activations

detected in white matter, namely the shape of the activations themselves.

It is generally accepted that activations in the gray matter are isotropic in nature, that is, they extend equally in all directions. This assertion is supported by the nature of the dendritic connections of neural bodies found in gray matter, and corroborated by the isotropic diffusion of water molecules in these regions, as shown by diffusion MRI (dMRI) studies [7]. The isotropic Gaussian smoothing applied on functional data, which is a staple of fMRI analysis, is only justified under this assumption, and is required for increasing the statistical significance of the results.

However, it does not seem reasonable that such an assumption would naturally extend to white matter regions of the brain, where individual fibers and fiber bundles have clear directionality, and are not equally related in all three dimensions. Under such circumstances, two fibers can be adjacent with only one of them being active, and applying an isotropic filtering will tend to combine large amounts of inactive regions with every potential activation. We hypothesize that a different form of filtering which respects the directionality of fibers and their spatial boundaries could increase the accuracy of detections in white matter. We develop such a method and test its effectiveness on both synthetic and real data.

White matter comprises about 50% of the brain, and its functional significance has been long established [8]. By improving the capacity of fMRI to produce results in these regions a whole new realm of possible findings is enabled.

1.1 Goal

The goals of this work are to:

- Develop an approach to white matter filtering for fMRI preprocessing that takes into account the directionality and spatial boundaries of axonal fibers in the human brain.
- Validate the proposed approach on semi-synthetic data created from real structural and diffusion MRI with synthetic functional activations.
- Test the proposed approach on real fMRI data and compare the results with those from the statistical parametric mapping (SPM) toolkit.

1.2 Contributions

The main contributions of this thesis are:

- A method for constructing spatial filters that adapt to the underlying structure of white matter.

-
- Incorporation of the proposed method into a spectral graph decomposition framework for denoising and significance testing of fMRI data.
 - Performance improvements in spectral graph WSPM (sgWSPM) making it applicable for large graphs (graphs with more than 10^4 – 10^5 vertices).
 - A method for constructing signal-adapted tight frames that is applicable for large graphs.
 - A tight frame construction derived from B-splines that provides kernels with local support yet is easy to approximate with polynomials.
 - Theoretical results showing that low-pass filtering with the heat kernel approximates Gaussian smoothing for some regular graphs.

1.3 Previous work

Previous use of graphs in the context of neuroimaging has mostly centered around the construction and analysis of brain networks [9, 10], and the relation between anatomical and functional connectivity in the brain [11, 12]. However, our thesis ties in with a different series of works based on the recently developed field of signal processing on graphs [13, 14, 15]. Of particular relevance are [16, 17, 18], where the wavelet SPM framework was extended to employ graph wavelets, and then used to perform anatomically-adapted smoothing of gray matter regions for fMRI preprocessing. In these works graphs were used to define the irregular domain of the gray matter.

Within the area of graph-based tractography, there have been several attempts at describing the structure of the white matter using graphs weighted by diffusion data [19, 20, 21, 22, 23, 24, 25]. In particular, to improve the angular resolution of outgoing edges, larger neighborhood sizes have been considered (74-connectivity [19], 124-connectivity [22] and 98-connectivity [23, 24]). To avoid excessive branching in crossing fiber populations [23] implemented a multigraph approach in which multiple fiber bundles traverse each voxel, and the weights of the outgoing edges are considered on an individual bundle basis.

Adaptive filtering approaches have also been developed outside of the graph domain, pointing out limitations in ordinary Gaussian smoothing [26, 27]. [26] used bilateral spatial smoothing on fMRI data to preserve edges between voxels of different intensity. This resulted in less mixing of gray matter, white matter and other anatomical regions. In [28] the term *fixel* was coined, denoting a specific fiber population within a voxel. A connectivity measure was introduced where fixels belonging to the same fiber bundle had a stronger connection than others. A method for smoothing

metrics defined on the fixels was also introduced. This is a form of diffusion-adapted filtering, but it was not applied to MRI data.

On a different note, our work also joins a number of others in utilizing one MRI modality to enhance the results of another [29]. One example is [30], where functional connectivity from fMRI was combined with anatomical connectivity from DTI to derive a single, more informative functional connectivity metric. Using a related approach, [31, 32] introduced the concept of track-weighted functional connectivity, in which DTI and fMRI data are used together to generate new functional connectivity maps.

1.4 Disposition

This thesis begins with a background in Chapter 2, followed by a description of the methodology employed in Chapter 3. Afterwards, Chapter 4 introduces the dataset used to generate the results in Chapter 5. Chapter 6 then concludes the main body of the work with a discussion of the results. Finally, a set of appendices is added in A, which serve to clarify or expand on concepts related to the work.

Chapter 2

Background

This chapter introduces the preliminaries required to understand the rest of the work. The first topic explored is spectral graph theory and signal processing on graphs. Following is a brief treatment of the relevant aspects of brain anatomy and physiology. The next section describes functional MRI and the processing required to derive activation maps from it. Finally, a section regarding diffusion MRI is presented.

2.1 Graphs and their spectra

The main mathematical tools used throughout this work are graphs and spectral graph theory. In recent years, several important successes have been achieved in bringing some common signal processing methods (e.g., Fourier transform, wavelets) to the graph domain [13, 15]. In this section we will provide an overview on graphs, their spectra, and the processing of signals defined on graphs.

We define an undirected, weighted graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{A})$ as a set \mathcal{V} of N_v vertices, a set \mathcal{E} of edges connecting pairs (i, j) of vertices, and an adjacency matrix \mathbf{A} , whose nonzero elements $a_{i,j}$ represent the weights of edges $(i, j) \in \mathcal{E}$. Given that the graph is undirected, $a_{i,j} = a_{j,i}$, i.e., \mathbf{A} is symmetric.

Additionally, we define a diagonal degree matrix \mathbf{D} , with elements $d_{i,i} = \sum_j a_{i,j}$, that is, the i -th diagonal element of \mathbf{D} is the degree (sum of the edge weights) of the i -th vertex of \mathcal{G} .

These matrices can be used to obtain the Laplacian matrix of \mathcal{G} , which can be defined both in combinatorial form \mathbf{L} and normalized form \mathcal{L} [33]:

$$\mathbf{L} = \mathbf{D} - \mathbf{A} \tag{2.1}$$

$$\mathcal{L} = \mathbf{D}^{1/2} \mathbf{L} \mathbf{D}^{1/2}. \tag{2.2}$$

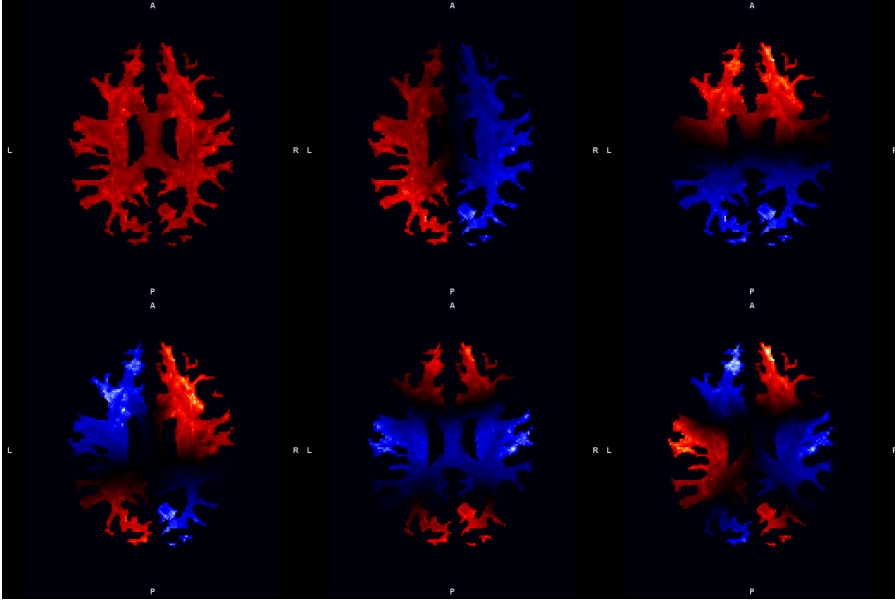


Figure 2.1: Axial slices of the six first eigenvectors (LRTB) of the white matter graph presented in Section 3.1. This clearly illustrates the relationship between eigenvalues and frequency in the graph.

Given that both definitions of the Laplacian matrix are symmetric and positive semi-definite, their eigendecomposition leads to a set of real non-negative eigenvalues:

$$\mathbf{\Lambda}(\mathcal{G}) = \{0 = \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_{N_v} \stackrel{\text{def}}{=} \lambda_{\max}\}. \quad (2.3)$$

The associated eigenvectors $\{\chi_l\}_{l=1}^{N_v}$ can be complex, but a real set can always be found, given that the Laplacian matrices are real and symmetric. The eigenvectors form an orthonormal basis, i.e., $\langle \chi_i, \chi_j \rangle = \delta_{i,j}$, and span the $\ell^2(\mathcal{G})$ space, defined in Section 2.1.1.

The set of eigenvalues of the Laplacian $\mathbf{\Lambda}(\mathcal{G})$ represents the spectrum of the graph. The smallest eigenvalue is always 0, and the multiplicity of zeros matches the number of connected components of the graph. The maximum eigenvalue is unbounded for \mathcal{L} , whereas for \mathcal{L} it is always ≤ 2 . The spectrum of a graph shows similarities to frequency in the classical domain, with the eigenvectors associated with higher eigenvalues being less smooth than those of lower eigenvalues [15] (see Figure 2.1). However, as opposed to the complex exponential in the classical domain, the Laplacian eigenvectors are not necessarily delocalized.

2.1.1 Graph signals

For a given graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{A})$, let $\ell^2(\mathcal{G})$ denote a Hilbert space of all square-summable $\mathbf{f} \in \mathbb{C}^{N_v}$ vectors, with inner product

$$\langle \mathbf{f}_1, \mathbf{f}_2 \rangle = \sum_{m=1}^{N_v} \mathbf{f}_1^*[m] \mathbf{f}_2[m], \quad \forall \mathbf{f}_1, \mathbf{f}_2 \in \ell^2(\mathcal{G}) \quad (2.4)$$

and norm

$$\|\mathbf{f}\|_2^2 = \langle \mathbf{f}, \mathbf{f} \rangle = \sum_{m=1}^{N_v} |\mathbf{f}[m]|^2 < \infty, \quad \forall \mathbf{f} \in \ell^2(\mathcal{G}). \quad (2.5)$$

A real graph signal $\mathbf{f} : \mathcal{V} \rightarrow \mathbb{R}$ defined on the vertex set \mathcal{V} can be seen as a vector $\mathbf{f} \in \ell^2(\mathcal{G})$ whose n -th component represents the value of the signal at the n -th vertex of the graph.

2.1.2 Graph Fourier transform

The classical continuous Fourier transform is defined in terms of the complex exponential $e^{i\omega x}$, which in turn represent the eigenfunctions of the one-dimensional Laplacian operator:

$$\frac{d^2}{dx^2} e^{i\omega x} = -\omega^2 e^{i\omega x}. \quad (2.6)$$

The inverse Fourier transform can therefore be seen as the expansion of a continuous signal f in terms of the eigenfunctions of the Laplacian operator:

$$f(x) = \mathcal{F}^{-1}\{\hat{f}\} = \frac{1}{2\pi} \int \hat{f}(\omega) e^{i\omega x} d\omega. \quad (2.7)$$

In analogy to this, Hammond et al. defined in [13] the graph Fourier transform as the expansion of a graph signal in terms of the eigenvectors of the graph Laplacian. For any graph signal $\mathbf{f} \in \ell^2(\mathcal{G})$, its graph Fourier transformed representation, denoted $\hat{\mathbf{f}} \in \ell^2(\mathcal{G})$, is defined as

$$\hat{\mathbf{f}}[l] = \mathcal{F}\{\mathbf{f}\} = \langle \boldsymbol{\chi}_l, \mathbf{f} \rangle = \sum_{n=1}^{N_v} \mathbf{f}[n] \boldsymbol{\chi}_l^*[n]. \quad (2.8)$$

The corresponding inverse transform is given as

$$\mathbf{f}[n] = \mathcal{F}^{-1}\{\hat{\mathbf{f}}\} = \sum_{l=1}^{N_v} \hat{\mathbf{f}}[l] \boldsymbol{\chi}_l[n]. \quad (2.9)$$

Using this definition for the graph Fourier transform, it can be shown that the Parseval relation holds [13]:

$$\langle \mathbf{f}_1, \mathbf{f}_2 \rangle = \langle \hat{\mathbf{f}}_1, \hat{\mathbf{f}}_2 \rangle, \quad \forall \mathbf{f}_1, \mathbf{f}_2 \in \ell^2(\mathcal{G}). \quad (2.10)$$

2.1.3 Continuous spectral kernels

Having established the graph spectral domain and the graph Fourier transform, it is possible to extend some general signal processing procedures from the classical Euclidean setting to the graph setting. In particular, we are interested in filtering graph signals with a filter defined by its spectral graph profile. A convenient way of defining such a filter is by sampling a continuous function $\mathcal{K} : [0, \lambda_{\max}] \rightarrow \mathbb{R}^+$, which we denote as a *spectral graph kernel*:

$$\hat{\mathbf{k}}[l] = \mathcal{K}(\lambda_l), \quad l = 1, \dots, N_v. \quad (2.11)$$

For a given graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{A})$, let $L^2(\mathcal{G})$ denote a Hilbert space of all square-integrable $\mathcal{K} : [0, \lambda_{\max}] \rightarrow \mathbb{R}^+$, with inner product

$$\langle \mathcal{K}_1, \mathcal{K}_2 \rangle = \int_0^{\lambda_{\max}} \mathcal{K}_1(\lambda) \mathcal{K}_2(\lambda) d\lambda, \quad \forall \mathcal{K}_1, \mathcal{K}_2 \in L^2(\mathcal{G}) \quad (2.12)$$

and norm

$$\|\mathcal{K}\|_2^2 = \langle \mathcal{K}, \mathcal{K} \rangle = \int_0^{\lambda_{\max}} |\mathcal{K}(\lambda)|^2 d\lambda < \infty, \quad \forall \mathcal{K} \in L^2(\mathcal{G}). \quad (2.13)$$

2.1.4 Signal processing on graphs

For any two graph signals $\mathbf{f}_1, \mathbf{f}_2 \in \mathbb{R}^{N_v}$, their convolution product is defined as

$$\begin{aligned} (\mathbf{f}_1 * \mathbf{f}_2)[n] &\stackrel{\text{def}}{=} \sum_{l=1}^{N_v} \hat{\mathbf{f}}_1[l] \hat{\mathbf{f}}_2[l] \chi_l[n] \\ &= \mathcal{F}^{-1}\{\hat{\mathbf{f}}_1[l] \hat{\mathbf{f}}_2[l]\}. \end{aligned} \quad (2.14)$$

As can be seen, the convolution operation in the vertex domain is equivalent to multiplication in the spectral domain, just like in conventional signal processing. The filtering of a graph signal is then defined based on this operation. Given a graph signal

$\mathbf{f} \in \ell^2(\mathcal{G})$ and a spectral graph kernel $\hat{\mathbf{k}}$ specified by its continuous representation $\mathcal{K} \in L^2(\mathcal{G})$, the filtered signal $(F_{\mathbf{k}}\mathbf{f})$ is obtained as

$$\begin{aligned} (F_{\mathbf{k}}\mathbf{f})[n] &\stackrel{\text{def}}{=} (\mathbf{k} * \mathbf{f})[n] \\ &= \sum_{l=1}^{N_v} \hat{\mathbf{k}}[l] \hat{\mathbf{f}}[l] \chi_l[n]. \end{aligned} \quad (2.15)$$

The impulse response associated with a given spectral kernel $\hat{\mathbf{k}}$ can be obtained, as in conventional signal processing, by filtering an impulse signal $\boldsymbol{\delta}$. However, this impulse response is not shift-invariant, and varies depending on the vertex m on which the impulse is localized:

$$\hat{\boldsymbol{\delta}}_m[l] = \langle \chi_l, \boldsymbol{\delta}_m \rangle = \chi_l^*[m], \quad (2.16)$$

$$\begin{aligned} \psi_{\mathcal{K},m}[n] &\stackrel{\text{def}}{=} (F_{\mathbf{k}}\boldsymbol{\delta}_m)[n] \\ &= \sum_{l=1}^{N_v} \hat{\mathbf{k}}[l] \hat{\boldsymbol{\delta}}_m[l] \chi_l[n] \\ &= \sum_{l=1}^{N_v} \hat{\mathbf{k}}[l] \underbrace{\chi_l^*[m]}_{=\hat{\psi}_{\mathcal{K},m}} \chi_l[n]. \end{aligned} \quad (2.17)$$

where $\psi_{\mathcal{K},m}$ denotes the impulse response, referred to as *atom*, associated to spectral kernel $\hat{\mathbf{k}}$ and localized at node m . Therefore, for a given spectral kernel there are N_v possible atoms, produced by filtering an impulse localized on each possible vertex of the graph.

Finally, it can be shown that filtering a signal \mathbf{f} with spectral kernel $\hat{\mathbf{k}}$ is equivalent to calculating the inner product between the signal and the atoms of the kernel:

$$\begin{aligned} (F_{\mathbf{k}}\mathbf{f})[m] &= \sum_{l=1}^{N_v} \hat{\mathbf{k}}[l] \hat{\mathbf{f}}[l] \chi_l[m] \\ &= \sum_{l=1}^{N_v} \hat{\psi}_{\mathcal{K},m}^*[l] \hat{\mathbf{f}}[l] \\ &= \langle \hat{\psi}_{\mathcal{K},m}, \hat{\mathbf{f}} \rangle \\ &\stackrel{(2.10)}{=} \langle \psi_{\mathcal{K},m}, \mathbf{f} \rangle. \end{aligned} \quad (2.18)$$

2.1.5 Spectral graph signal decomposition

A set of spectral kernels $\{\mathcal{K}_j\}_{j=1}^{N_s}$ forms a *tight frame* if

$$\sum_{j=1}^{N_s} |\mathcal{K}_j(\lambda)|^2 = C \quad \forall \lambda \in [0, \lambda_{\max}]. \quad (2.19)$$

that is, if the squared sum of the values of each kernel is equal to a constant for every λ . In the case that $C = 1$, the spectral kernels form a *Parseval frame*.

Given a set of spectral kernels $\{\mathcal{K}_j\}_{j=1}^{N_s}$, (2.18) can be used to decompose a graph signal \mathbf{f} onto them, yielding the coefficients

$$c_{\mathcal{K}_j, m} = \langle \psi_{\mathcal{K}_j, m}, \mathbf{f} \rangle. \quad (2.20)$$

Provided the kernels form a Parseval frame, the original signal can be reconstructed using

$$\mathbf{f}[n] = \sum_j \sum_m c_{\mathcal{K}_j, m} \psi_{\mathcal{K}_j, m}[n]. \quad (2.21)$$

2.1.6 Polynomial kernel approximation

Both the filtering of graph signals and their decomposition onto sets of spectral kernels are operations realized through (2.18). However, this equation depends on the availability of the full set of atoms for every kernel, which in turn requires the calculation of all the eigenvectors of the Laplacian matrix. Such an approach becomes infeasible for larger graphs. Instead, a fast approximation algorithm can be used.

Let $\mathcal{P} \in L^2(\mathcal{G})$ be a polynomial approximation of kernel $\mathcal{K} \in L^2(\mathcal{G})$. For a graph signal \mathbf{f} , its filtering with kernel \mathcal{K} , or equivalently, its decomposition coefficients when projected on the atoms of \mathcal{K} , can be found using \mathcal{P} as

$$\tilde{\mathbf{c}}_{\mathcal{K}} = \sum_{l=1}^{N_v} \mathcal{P}(\lambda_l) \hat{\mathbf{f}}[l] \chi_l \quad (2.22)$$

$$= \mathcal{P}(\mathcal{L}) \sum_{l=1}^{N_v} \hat{\mathbf{f}}[l] \chi_l \quad (2.23)$$

$$= \mathcal{P}(\mathcal{L}) \mathbf{f} \quad (2.24)$$

where $\tilde{\mathbf{c}}_{\mathcal{K}} \in \ell^2(\mathcal{G})$ with $\tilde{\mathbf{c}}_{\mathcal{K}}[m] = c_{\mathcal{K}, m}$. In (2.23) we use the fact that for any polynomial \mathcal{P}

$$\mathcal{L} \chi_l = \lambda_l \chi_l \Rightarrow \mathcal{P}(\mathcal{L}) \chi_l = \mathcal{P}(\lambda_l) \chi_l. \quad (2.25)$$

This approximation has the benefit that it does not require the explicit calculation of the eigenvectors. Instead, a polynomial of the Laplacian matrix is applied to the signal, which can be efficiently implemented with matrix-vector multiplication. In [13], a truncated Chebyshev expansion was used as it has the benefit of approximating a minimax polynomial. This minimizes an upper bound on the approximation error in the coefficients and is the method we used for this work.

2.2 The brain

MRI methods are both heavily reliant on and helpful in characterizing certain essential aspects of brain anatomy and function. The tissues in the brain can be broadly divided by their functions into those responsible for processing and those responsible for the transmission of signals [34].

The processing burden of the brain is carried out in *gray matter*, and is distributed among functionally specialized interdependent regions. This type of tissue is found on the outermost layer of the brain (cerebral and cerebellar cortex) as well as in a series of deep nuclei and the brainstem. Its constitution is mostly neuronal cell bodies.

The interdependent nature of gray matter regions relies on the existence of a channel for information transfer among them. This function is performed by the *white matter*, which comprises most of the subcortical regions of the brain. It is composed of neuronal *axons*, which are covered in a myelin sheath that insulates the electrical signals traveling through them. The left and right cerebral hemispheres are connected through the *corpus callosum*, a large white matter structure.

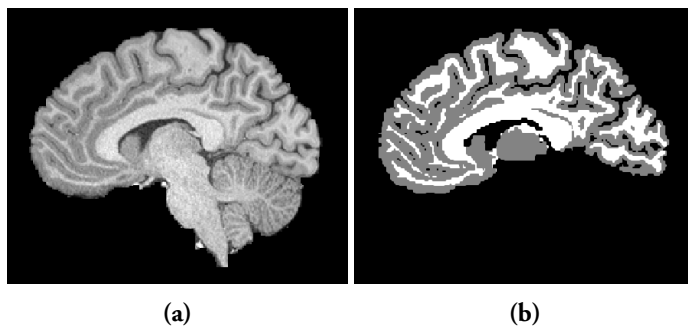


Figure 2.2: (a) A T1 image of a sagittal slice of the brain. (b) Segmentation showing cerebral gray matter in gray and white matter in white.

2.3 Functional MRI

Functional MRI (fMRI) is a magnetic resonance technique that facilitates the non-invasive study of functional activation in the brain. It most commonly works by utilizing the so called blood oxygenation level dependent (BOLD) effect: when a brain region is activated, it experiences an increase in the flow of oxygenated blood, which, due to the magnetic properties of hemoglobin, is measurable through the application of an electromagnetic gradient [35]. The acquisition process generates a time series of 3D brain volumes, each of them composed of multiple individual volume elements, called *voxels*.

Two main modalities of fMRI can be distinguished:

- In *task-based fMRI* participants are subjected to a series of stimuli arranged in time, referred to as a *paradigm*, with the purpose of eliciting in them a measurable BOLD response. The analysis of the resulting data can reveal which brain regions are engaged in the performance of specific task or are responsible for the actions of a given body part.
- In *resting-state fMRI* subjects are scanned in the absence of a paradigm, that is, during rest. Because the subjects are not engaged in a specific task, spontaneous brain activity can be detected, which can be used to find regions of correlated behavior, revealing the functional organization of the brain [36].

2.3.1 fMRI activation mapping

The analysis of fMRI data is performed in the time domain, and it has the goal of establishing temporal correlations in the BOLD signal of brain regions. Task-based fMRI attempts to establish correlations between regions of the brain and the paradigm used, while in resting-state fMRI the correlations are drawn between pairs of brain regions [37]. This section will provide a description of the analysis techniques used in the former, since that is the modality employed throughout this work.

SPM

SPM is a statistical framework for the analysis of functional imaging brain data. A software toolkit with the same name is implementing the framework. It was developed in the 1990s [38] and has become one of the most established tools for fMRI analysis.

The SPM framework is summarized in Figure 2.3. The procedure can be divided into three main steps [39]:

Preprocessing of the raw data. A number of corrections are applied to the data, both for quality control and in order to allow for proper statistical modeling.

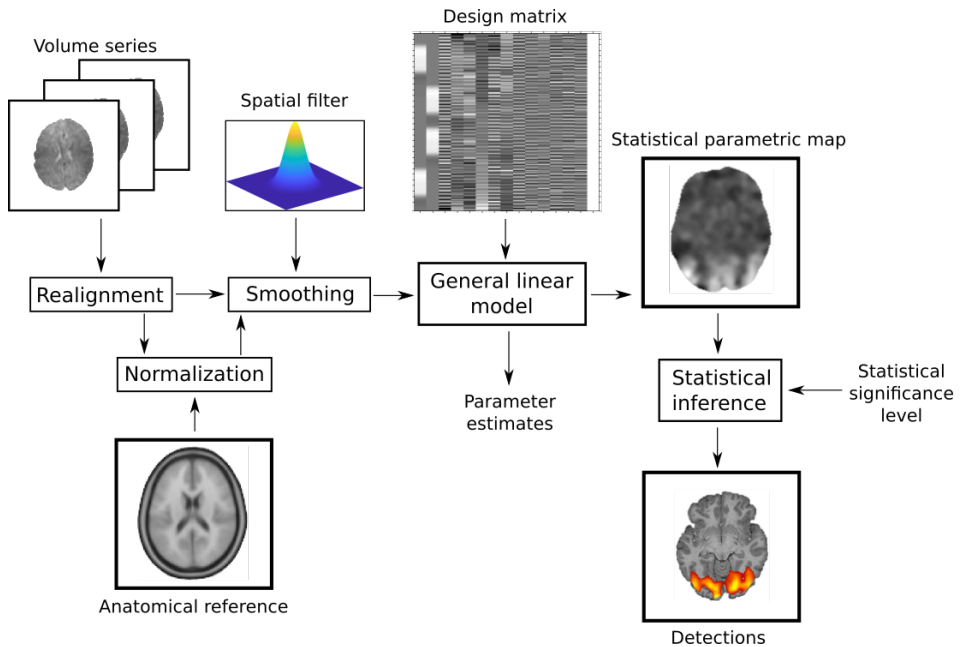


Figure 2.3: Overview of the SPM framework

These can include inspecting the raw data, distortion correction, motion correction (coregistration) and slice timing correction. Spatial alignment of the volumes ensures that each voxel corresponds to the same region of the brain across time. The final and essential step in preprocessing is the spatial smoothing of the data, typically with a Gaussian kernel. The importance of this is twofold: it increases the signal-to-noise ratio (SNR) of the data and it enables the use of methods for the correction for multiple comparisons that are less strict than Bonferroni correction, such as correction based on Gaussian random field theory (GRFT).

Modeling of data. The BOLD value of each voxel throughout the volume series constitutes a signal in time. Using a GLM, the temporal variation in the signal can be described as a linear combination of experimental effects (i.e., the actions that the subject is asked to perform), confounding effects (e.g., slight uncorrected movement, oscillations in the signal level) and residual variability.

For a series of N_t volumes consisting of N_v voxels, and with a set of K regressors, the GLM of the timecourse of all voxels can be written in matrix form

as

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \quad (2.26)$$

where \mathbf{Y} is a $N_t \times N_v$ matrix, each column representing the timecourse of a voxel, \mathbf{X} is a $N_t \times K$ matrix where each column contains one of the covariates used to explain the variation in the data, $\boldsymbol{\beta}$ is a $K \times N_v$ matrix containing the model parameters for every voxel, and $\boldsymbol{\epsilon}$ is a $N_t \times N_v$ residuals matrix.

The known parameters are \mathbf{Y} and \mathbf{X} , and the goal of the GLM fitting is to find a set of values for the parameters $\boldsymbol{\beta}$ that minimizes the residuals $\boldsymbol{\epsilon}$. In the case of $\boldsymbol{\epsilon}_j \in \mathcal{N}(0, \sigma_j^2 \mathbf{I})$, the solution in the least-squares sense is given by

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y} \quad (2.27)$$

where, for voxel j , $E(\hat{\boldsymbol{\beta}}_j) = \boldsymbol{\beta}_j$ and $\text{Var}(\hat{\boldsymbol{\beta}}_j) = \sigma_j^2 (\mathbf{X}^T \mathbf{X})^{-1}$.

Statistical inference on the modeled data. The significance of the estimated effect size associated to each voxel is statistically tested, generating a volume in which the voxel values represent some statistic, in our case a t-value. This volume is a statistical parametric map.

The t-values are given by

$$t_j = \frac{\hat{\mu}_j}{\hat{\sigma}_j} \quad (2.28)$$

where

$$\hat{\mu}_j = \mathbf{c}^T \hat{\boldsymbol{\beta}} \quad (2.29)$$

is the effect size for voxel j ,

$$\hat{\sigma}_j = \sqrt{\frac{\boldsymbol{\epsilon}_j^T \boldsymbol{\epsilon}_j}{N_t - \text{rank}(\mathbf{X})} \mathbf{c}^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{c}} \quad (2.30)$$

is an estimate of its standard error. \mathbf{c} denotes a $1 \times K$ contrast vector used to select the effects of interest from the set of covariates by linearly combining a subset of the estimated parameters $\hat{\boldsymbol{\beta}}_j$. In the basic case when the effect size of the i -th covariate is estimated, the contrast vector becomes $\boldsymbol{\delta}_i$.

The final activation map is obtained by thresholding the statistical parametric map according to some statistical significance level α . The resulting map contains all voxels which have been found to have at least a $(1 - \alpha)$ probability of being significantly different from the null hypothesis ($\mu_j = 0$).

Multiple comparisons

SPM is a mass-univariate approach, in that every voxel forms a time series that has to be individually tested for significance, resulting in N_v statistical tests. The statistical significance level α of a test represents the probability of rejecting the null hypothesis given that it is true. This is equivalent to the false-positive rate. However, for N_v individual tests, the probability of at least one of them producing a false positive becomes $(1 - (1 - \alpha)^{N_v})$, which quickly approaches one as N_v increases. This has been dubbed the *multiple comparisons* problem.

A conservative solution to this problem is the application of *Bonferroni correction*, by which a corrected significance level is calculated as $\tilde{\alpha} = \alpha/N_v$. Using this significance level sets the probability of having at least one false-positive detection at α . However, fMRI volumes can commonly have upwards of $10^4 - 10^5$ voxels, which poses a very high requirement on the significance of the effects that can be detected.

A different approach, employed in SPM, relies on the use of GRFT in order to exploit the spatial correlation in the brain volume [40], which is disregarded by Bonferroni correction. The significance level is then corrected based on the number of unique resolution elements (*resels*), which is related to the number or independent values that the data takes, and can be far lower than the total number of voxels N_v . However, this method places an assumption on the smoothness of the data, and therefore requires the use of Gaussian smoothing in preprocessing. In effect, this supposes a trade-off between the increased sensitivity of the test and the spatial specificity of the shape and size of the detected activations. Furthermore, a number of studies have shown that this approach yields higher error rates than expected from the significance level [41, 42].

Spectral graph wavelet SPM

Advanced approaches to SPM have been developed in which the brain volumes are subjected to wavelet decomposition and a GLM is used to model the time variance of the resulting coefficients. This has the advantages of increasing the SNR and producing a sparse representation of the activation maps. One such approach, termed *wavelet SPM (WSPM)* [43, 44], integrates wavelet-based processing with statistical testing in the spatial domain. This approach is described adapted to a graph setting and employing a spectral graph decomposition in place of the discrete wavelet transform. This modified approach is referred to as *spectral graph WSPM (sgWSPM)* [16, 17].

For a given graph \mathcal{G} representing the brain volume, a set of spectral graph kernels $\{\mathcal{K}_j\}_{j=1}^{N_s}$ with associated atoms $\{\{\psi_{\mathcal{K}_j,m}\}_{j=1}^{N_s}\}_{m=1}^{N_v}$ is defined. Each brain volume in the fMRI time series, taken as a graph signal \mathbf{f} , can be decomposed onto the atoms using (2.20), yielding the set of coefficients $c_{\mathcal{K}_j,m}$. The time series for all coefficients are arranged into a matrix \mathbf{Y} , which is used in a GLM fitting according to (2.26).

Optimal parameters are then estimated with (2.27).

Effect sizes $\hat{\mu}_{\mathcal{K}_j,m}$, their standard error $\hat{\sigma}_{\mathcal{K}_j,m}$, and the corresponding t-values for each coefficient $t_{\mathcal{K}_j,m}$ are calculated with (2.29), (2.30) and (2.28) respectively. Two parameter maps are then defined: one with all effect sizes, and another with effect sizes that are above a statistical significance threshold τ_T . Both of these contrasts are reconstructed using (2.21):

$$R_i = \sum_{j=1}^{N_s} \sum_{m=1}^{N_v} \hat{\mu}_{\mathcal{K}_j,m} \psi_{\mathcal{K}_j,m}[i] \quad (2.31)$$

$$\tilde{R}_i = \sum_{j=1}^{N_s} \sum_{m=1}^{N_v} H(|t_{\mathcal{K}_j,m}| - \tau_T) \hat{\mu}_{\mathcal{K}_j,m} \psi_{\mathcal{K}_j,m}[i] \quad (2.32)$$

where $H(\cdot)$ represents the Heavyside step function. The first of these reconstructed maps corresponds to the parameter map that would have been obtained by running ordinary SPM, while the second reconstructs only the coefficients that are deemed significant in the decomposed domain. In order to avoid spatial bias in the result, the final spatial parameter map is obtained as [44]:

$$\tilde{u}_i = \min(R_i, \tilde{R}_i). \quad (2.33)$$

Finally, to find the activations, the estimated standard errors are reconstructed with a modified transform, and are used in the thresholding of the spatial parameter map:

$$u_i = H\left(\frac{\tilde{u}_i}{\sum_{j=1}^{N_s} \sum_{m=1}^{N_v} \hat{\sigma}_{\mathcal{K}_j,m} |\psi_{\mathcal{K}_j,m}[i]|} - \tau_S\right) \tilde{u}_i. \quad (2.34)$$

This is a time consuming step, since it requires the calculation of every atom in every subband. This work contributes with significant improvements in its computation time, which are described in Section 3.5.

The values for the thresholds τ_T and τ_S are defined based on the desired statistical significance level α , and are given by

$$\tau_T = \sqrt{-W_{-1}(-2\pi\alpha^2)}, \quad \tau_S = 1/\tau_T \quad (2.35)$$

where W_{-1} is the -1 -branch of the Lambert W -function.

2.3.2 Activations in white matter

Previous studies using fMRI have mainly focused on finding detections in gray matter, and as such, experimental designs have been tailored for this tissue type as opposed to the white matter. Optimal acquisition parameters, for example, are not identical for gray and white matter, and the HRF of each tissue may have a different shape. This, in combination with a lower cerebral blood flow and cerebral blood volume, results in weaker BOLD signals, making any potential white matter activations harder to detect [3].

However, there is an increasing body of literature both presenting detected activations and examining the properties of the BOLD signal in white matter. Of particular interest for this work is [45], where the authors performed a rsfMRI experiment showing that, for the same separation distance, the BOLD signal of pairs of white matter voxels belonging to the same fiber bundle is significantly more correlated than that of random pairs of white matter voxels. These results were expanded in [46], where it was shown that the magnitude of the correlations among voxels from the same bundle is about twice as strong in a task-based study than in rsfMRI. Additionally, in [47] the authors performed a clustering analysis on rsfMRI data, and produced functional white matter networks that showed correspondence with streamlines produced by diffusion tensor imaging (DTI).

Based on this evidence, we formulate the fundamental assumption and hypothesis of our work:

Assumption: The BOLD signal in white matter shows a stronger spatial correlation along the anatomical white matter tracts than in other directions.

Hypothesis: The use of an isotropic filter confounds genuine white matter activations with the surrounding inactive regions. Therefore, adopting a filtering scheme which adapts to the direction and spatial extent of the underlying white matter tracts would increase the specificity and sensitivity in the detection of white matter activations.

2.4 Diffusion MRI

Particles suspended in a fluid will naturally exhibit a Brownian motion, moving randomly without the introduction of bulk motion (motion caused by pressure or temperature differences). For example, if a drop of dye is introduced to a glass of water it will diffuse and after some time the color of the water will be homogeneous.

Diffusion MRI (dMRI) is an MRI technique that measures the amount of diffusion experienced by water molecules in the brain along a particular gradient. The movement of molecules will be hindered by anatomical features and thus water will

diffuse more or less in certain directions. Within the white matter, the diffusion is assumed to be greater in the direction parallel to the axonal fibers than in the direction perpendicular to them. Based on this fact, one can infer the fiber structure of the brain by sampling the diffusion along multiple gradients [48].

2.4.1 Diffusion tensors

The most common way of modeling the diffusion within a voxel is to use a *diffusion tensor*, which is a real symmetric matrix D with eigenvectors χ_1, χ_2, χ_3 and eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3$. The assumption being made in the model is that the displacement distribution of water molecules in a given time is Gaussian with covariance matrix D [49]. The eigenvalues are used to define various measures such as mean diffusivity (MD), fractional anisotropy (FA), radial diffusivity (RD) and axial diffusivity (AD). These find usage when studying diseases such as Alzheimer's disease, autism, schizophrenia and multiple sclerosis [50].

A common way of visualizing a tensor is by using an ellipsoid (see Figure 2.4) where the principal axes are $\sqrt{\lambda_i}\chi_i, i = 1, 2, 3$. The largest eigenvalue corresponds to the direction of greatest diffusion and thus a probable direction for axonal fibers [51]. However, a limitation of the tensor is that it can only represent a single fiber direction, and to assume that a voxel only contains a single fiber direction is a strong assumption to make, as it has been estimated that 33%–90% of voxels contain multiple fiber populations [50].

2.4.2 ODFs

More advanced approaches have been developed for reconstructing the dMRI data, such as diffusion spectrum imaging (DSI) [52], Q-ball imaging (QBI) [53] and generalized Q-sampling imaging (GQI) [54], which result in the generation of *orientation distribution functions (ODFs)* (see Figure 2.5). An ODF is a probability distribution defined on the sphere that indicates the probability of a water molecule at a given point in space diffusing in a certain direction. Thus, the ODF effectively captures the local fiber structure. One of the main benefits of the ODF over other common representations (e.g., diffusion tensors [49]) is that it is capable of representing crossing fibers [55].

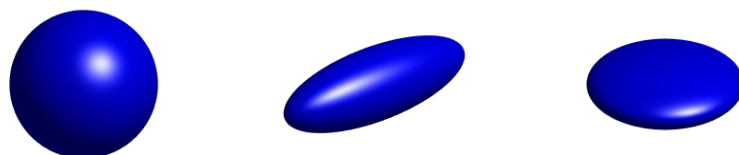


Figure 2.4: (a) An isotropic tensor. (b) An (anisotropic) prolate tensor. (c) An oblate tensor arising from two crossing fibers. Note that it is not possible to discern the individual directions of the crossing fibers.

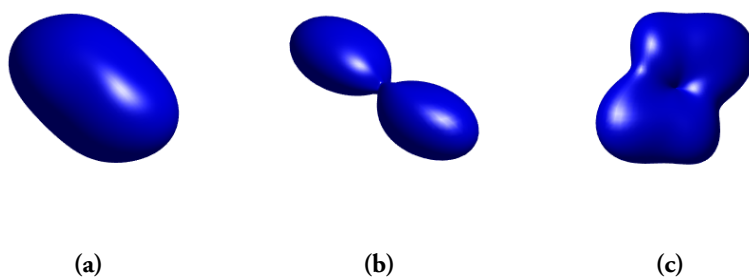


Figure 2.5: (a) A very anisotropic ODF. (b) The same ODF normalized by subtracting its minimum value. (c) A normalized ODF containing two fiber directions

2.4.3 Tractography

One of the main uses given to dMRI data is to reconstruct sets of plausible pathways within the white matter, based on the assumption that water diffusion is strongest in directions parallel to the underlying axonal fibers. This process is known as tractography. Although many different tractography methods have been developed, our discussion will focus only on deterministic tractography [56], which is the approach used in this work.

In deterministic tractography, the calculated fibers are seen as 3D curves, called streamlines, whose direction is always tangential to an underlying vector field. This vector field represents the directions of maximum diffusion at each point, and can be defined by, for example, finding peak directions in the ODFs.

The process of producing each streamline starts by selecting a seeding point, from which the streamline will start. Then, small steps are taken in the direction of maximum propagation of the voxel in which the seeding point is located. When the streamline transitions into a new voxel, its direction of maximum propagation is followed. This process is repeated until some termination criteria are met, which usually involve a minimum anisotropy in the traversed voxel or a maximum turning angle for the streamline [57]. An example streamline is shown in Figure 2.6.

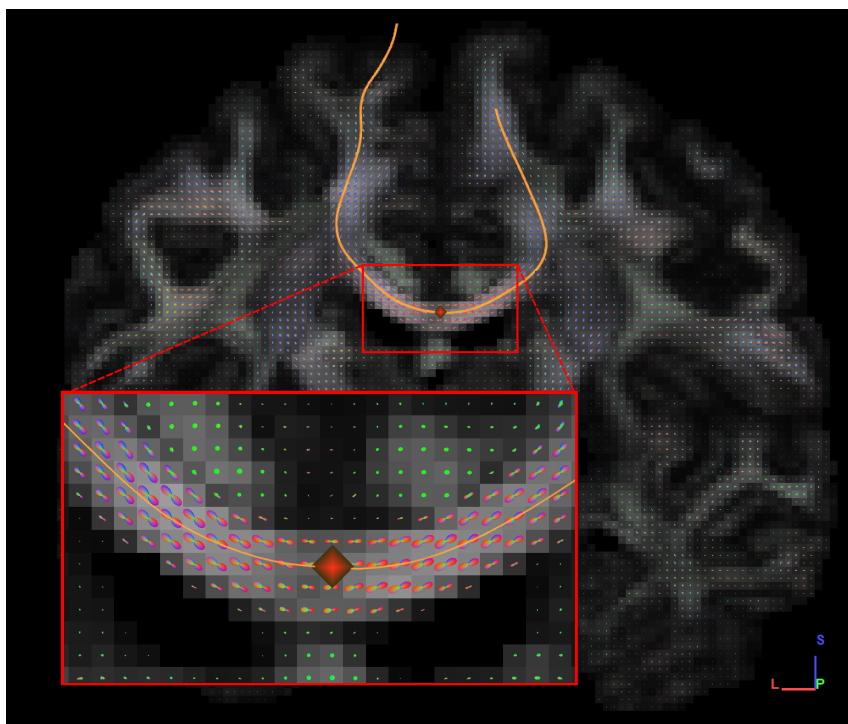


Figure 2.6: A coronal slice illustrating a single streamline passing through the corpus callosum overlaid on the constructed ODFs for that slice. The seeding point is indicated with a red diamond.

Chapter 3

Methodology

In this work we explore two different approaches for the filtering of white matter in fMRI data. Both of them rely on the use of a graph for the description of the white matter, which enables the encoding of the underlying directionality of the axonal fibers. The design of a suitable graph is described in Section 3.1.

The first filtering approach consists in replacing the isotropic Gaussian filtering commonly used in fMRI data preprocessing with a heat kernel filtering in the graph spectral domain. Filters designed in such a way retain approximately Gaussian isotropic behavior in isotropic white matter regions, but are also capable of adapting their shape to the anisotropic regions. Section 3.2 describes the use of the heat kernel and its properties.

The second approach is based on spectral decomposition of the data, using the WSPM framework adapted to the graph domain. Within this framework, a signal-adapted procedure for the design of spectral kernels was adopted, which is presented in Section 3.3. An efficient implementation of this procedure requires the fast estimation of the ensemble energy spectral density of the fMRI volume series. This was achieved by employing a B-spline kernel design, described in Section 3.4. Finally, a series of improvements in the efficiency of the graph domain WSPM framework are shown in Section 3.5.

3.1 White matter graph design

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{A})$ denote a graph where the vertex set \mathcal{V} is taken to be all voxels that are classified as white matter. Edges are added from a voxel to all 124 voxels within a $5 \times 5 \times 5$ neighborhood. A smaller neighborhood, such as $3 \times 3 \times 3$, has worse angular resolution in the edges and would not be able to model the directions of the axonal fibers with the same accuracy. On the other hand, a larger neighborhood, in addition

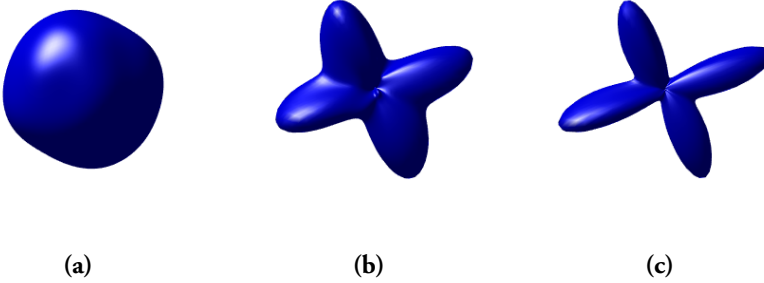


Figure 3.1: (a) An ODF sharpened by taking it to a power of $p_{ODF} = 1$, (b) $p_{ODF} = 20$ and (c) $p_{ODF} = 40$.

to increased complexity, could result distant connections that are not justifiable as a series of local ones. The weighting of the edges is similar to that in [20] and [23]:

$$a_{ij} = a_{ji} = P_{mat}(i)P_{mat}(j) [P_{diff}(i, \vec{r}_{ij}) + P_{diff}(j, \vec{r}_{ji})] \quad (3.1)$$

where the P_{mat} factor is derived from probability maps of the white and gray matter. In our construction we are assuming a binary classification of the white matter and thus the factor reduces to $P_{mat}(i) = 1 \forall i \in \mathcal{V}$. P_{diff} is determined by integrating the ODF over a solid angle $\omega = 4\pi/98$ around \vec{r}_{ij} , where \vec{r}_{ij} is a vector pointing from vertex i to vertex j . Since the ODF is sampled at a discrete set of points, we approximate the integral with a sum:

$$P_{diff}(i, \vec{r}_{ij}) = \frac{1}{Z_i} \int_{\omega} ODF(i, \vec{r}) dS \approx \frac{1}{Z_i} \sum_{q=1}^Q ODF(i, \vec{r}_q) \cdot \Delta S_q. \quad (3.2)$$

where Z_i is set such that $\max_{j|(i,j) \in \mathcal{E}} P_{diff}(i, \vec{r}_{ij}) = 0.5$. A slight modification to the above formula is that, prior to integrating, we sharpen the ODFs by taking them to some power p_{ODF} (see Figure 3.1).

3.2 The heat kernel

In order to improve the results of the filtering of fMRI data over those obtained from the typical isotropic Gaussian [58], it is necessary to design a filter kernel that would mimic its behavior in particularly isotropic regions of the white matter, while at the same time adapting its shape to follow the underlying fibers in anisotropic regions.

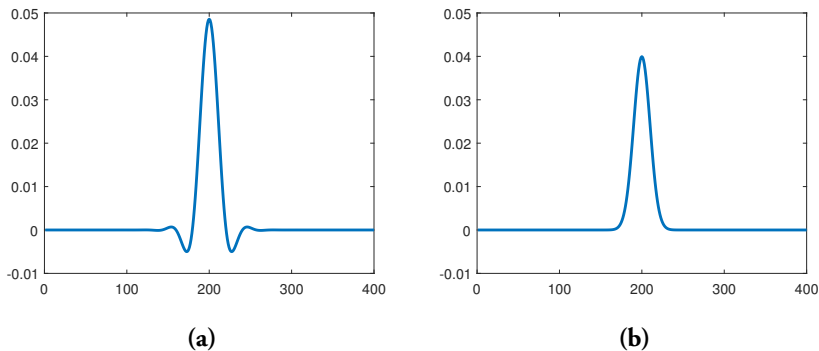


Figure 3.2: (a) an atom associated with a Gaussian kernel and (b) an atom associated with a heat kernel at vertex $m = 200$ in a cycle graph with $N_v = 400$ vertices.

The graph setting allows for the design of filters with such level of adaptation, but the problem of eliciting Gaussian behavior in the atoms becomes nontrivial.

In Euclidean continuous and discrete domains, the Fourier transform of a Gaussian signal is another Gaussian, with the pair following simple and well known relations [59, Table 1.14.1]. Thus, it is easy to obtain Gaussian signals by specifying their spectral representation. However, this property does not extend to atoms and spectral kernels in a graph setting. This can be shown empirically to be the case by examining atoms produced by a Gaussian kernel on a cycle graph. The results, presented in Figure 3.2a, show that the atoms oscillate, have negative values, and do not resemble the shape of a Gaussian.

The cause of this disparity between the classical and graph domains is that their Fourier transforms, while produced in an analogous way as expansions into the eigenfunctions of the Laplacian operator, are not entirely similar in their behavior. For example, as opposed to Fourier elements in the Euclidean setting, the spectral elements (eigenvalues) in the graph setting are non-uniformly spread across the spectrum. Moreover, the basis (eigenvectors) are not guaranteed to be delocalized as their Euclidean counterparts (complex exponentials) are [15].

Instead of a Gaussian kernel, a heat kernel

$$\mathcal{K}(\lambda) = e^{-\tau\lambda} \quad (3.3)$$

where τ is a free parameter, produces results that more closely resemble the Gaussian filters of the Euclidean domain. Atoms produced by such a kernel will always be non-negative [33, Lemma 10.4] and are approximately Gaussian for some regular graphs (see Figure 3.2b). In Appendix A.1 we show this for unweighted grid graphs

of any dimension and derive a relation between the parameter τ and the standard deviation of the atoms. In particular, a 3D grid graph with 6-connectivity produces approximately Gaussian atoms with standard deviation $\sigma = \sqrt{\tau/3}$.

Unfortunately, the Gaussianity of the atoms is dependent on the topology and weighting of the graph, as well as on the τ parameter. In particular, the shape of the atoms in lattice graphs closely resembles a Gaussian, but with an additional component of variable prominence at the center voxel. This peak is nonexistent in a 3D 6-connectivity configuration, but gets emphasized by incorporating additional neighborhood directions and increasing the weights for these additional directions. The peak becomes significant for the 124-connectivity used in this work. This poses a compromise between angular resolution in the atoms and their Gaussianity.

3.3 Signal-adapted filter design

The description of sgWSPM in Section 2.3.1 leaves out the design of spectral kernels required for the decomposition step. For the decomposition to be meaningful we do not want all the energy of the signal to be confined to a single or just a few of the subbands. Rather, a good system of spectral kernels would divide up the energy evenly between the subbands. For this to work for an arbitrary signal the design has to be able to adapt to the spectral content of the signal.

In [60], a method for creating a signal-adapted system of spectral kernels was introduced. First, a prototype uniform system of spectral kernels $\{\mathcal{K}'_j\}_{j=1}^{N_s}$ forming a Parseval frame is designed such that

$$\int_0^{\lambda_{max}} \mathcal{K}'_j(\lambda) d\lambda = C, \quad j = 1, \dots, N_s \quad (3.4)$$

for some $C \in \mathbb{R}^+$.

The method makes use of the so called *ensemble energy spectral density*, defined as

$$\bar{e}_{\mathcal{F}}[l] = \frac{1}{N_s} \sum_{s=1}^{N_s} \left| \left\langle \chi_l, \frac{\mathbf{f}_s}{\|\mathbf{f}_s\|} \right\rangle \right|^2, \quad l = 1, \dots, N_v. \quad (3.5)$$

A continuous warping function $T_{\mathcal{F}}(\lambda) : [0, \lambda_{max}] \rightarrow [0, \lambda_{max}]$ is obtained through monotonic cubic interpolation of the set of points

$$\left\{ \left(\lambda_l, \frac{\lambda_{max}}{m_{\lambda_l}(1 - \bar{e}_{\mathcal{F}}[1])} \sum_{r=i_{\lambda_l}}^{i_{\lambda_l} + m_{\lambda_l}} \sum_{k=2}^r \bar{e}_{\mathcal{F}}[k] \right) \right\}_{l=2}^{N_v} \quad (3.6)$$

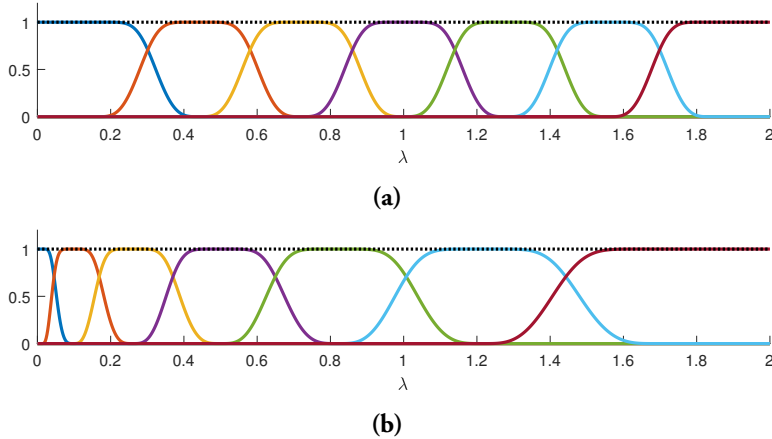


Figure 3.3: (a) A UMT filter design with 7 subbands. (b) The same design warped using $T_{\mathcal{F}}(\lambda) = \sqrt{\lambda_{max}\lambda}$. This kind of warping could arise from a signal with more energy in the lower frequencies.

together with $(\lambda_1, 0)$, where m_{λ_l} is the multiplicity of the eigenvalue λ_l and i_{λ_l} is the index of its first occurrence. The signal-adapted system of kernels is then given by the warped kernels

$$\mathcal{K}_j(\lambda) = \mathcal{K}'_j(T_{\mathcal{F}}(\lambda)), \quad j = 1, \dots, N_s. \quad (3.7)$$

See Figure 3.3 for an example using the uniform Meyer-type (UMT) filter design presented in [60].

Due to the explicit use of eigenvalues and eigenvectors for computing $\bar{e}_{\mathcal{F}}$, this approach becomes computationally heavy for large graphs, and infeasible for extensively large graphs whose corresponding Laplacian cannot be diagonalized. Here, we instead approximate the ensemble energy spectral density by decomposing the signals using a uniform system of spectral kernels with a large number of subbands. We start by removing the DC component of the signal, i.e., the contribution of the first eigenvector:

$$\mathbf{g}_s = \mathbf{f}_s - \langle \boldsymbol{\chi}_1, \mathbf{f}_s \rangle \boldsymbol{\chi}_1, \quad (3.8)$$

where $\boldsymbol{\chi}_1$ can be found without doing a full eigenvalue decomposition of \mathcal{L} [61]. If there are multiple eigenvalues equal to zero the contribution of their respective eigenvectors would also have to be removed in the same way.

With $\{\mathcal{K}''_j\}_{j=1}^{N''_s}$ denoting a system of N''_s spectral kernels (e.g. the B-spline design

presented in the next section) we approximate ensemble spectral energy density as

$$\tilde{e}_{\mathcal{F}}[j] = \frac{1}{C} \sum_{s=1}^{N_s} \sum_{m=1}^{N_v} \left| \left\langle \psi_{\mathcal{K}_j'',m}, \frac{\mathbf{g}_s}{\|\mathbf{g}_s\|} \right\rangle \right|^2, \quad j = 1, \dots, N_s'' \quad (3.9)$$

where C is a normalization constant ensuring that $\sum_j \tilde{e}_{\mathcal{F}}[j] = 1$.

Similar to before we define the points

$$\left\{ \left(\frac{\lambda_{max}}{K} \sum_{k=1}^j \|\mathcal{K}_k(\lambda)\|^2, \lambda_{max} \sum_{k=1}^j \tilde{e}_{\mathcal{F}}[k] \right) \right\}_{j=1}^{N_s''} \quad (3.10)$$

together with $(0, 0)$, where $K = \sum_{k=1}^{N_s''} \|\mathcal{K}_k(\lambda)\|^2$. Interpolation is performed resulting in the warping function $T_{\mathcal{F}}(\lambda)$, which is used to create the final signal-adapted system of spectral kernels $\mathcal{K}_j(\lambda) = \mathcal{K}_j'(T_{\mathcal{F}}(\lambda))$, $j = 1, \dots, N_s$.

3.4 B-spline system of spectral kernels

For the purpose of estimating the ensemble energy spectral density described in the previous section we need a suitable system of spectral kernels. Such a system should have smooth kernels of the bandpass type. The smoothness makes it possible to estimate the kernels with polynomials of lower degree which speeds up the calculations. It also results in more localized atoms [60], but that is of less importance for our usage. In this section we propose such a design, derived from the B-spline basis functions.

Splines are piecewise polynomial functions defined with certain continuity requirements. B-splines, or basis splines, are a particular kind of splines defined by the recursion formula [62]

$$N_i^0(u) = \begin{cases} 1 & \text{if } a_i \leq u < a_{i+1} \\ 0 & \text{otherwise} \end{cases} \quad (3.11)$$

and

$$N_i^n(u) = \frac{u - a_i}{a_{i+n} - a_i} N_i^{n-1}(u) + \frac{a_{i+n+1} - u}{a_{i+n+1} - a_{i+1}} N_{i+1}^{n-1}(u), \quad (3.12)$$

where n is the degree of the polynomials and a_0, \dots, a_{m+n+1} is a sequence of non-decreasing numbers called knots. The splines are \mathcal{C}^{n-r} ($n - r$ times differentiable) at

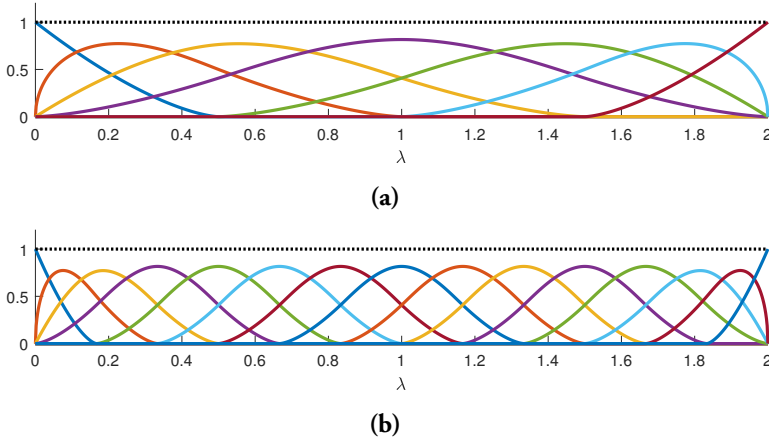


Figure 3.4: (a) A Parseval frame consisting of 7 subbands constructed from B-spline basis functions of degree 3 with uniform knot sequence $[0, 0, 0, 0, 0.5, 1.0, 1.5, 2, 2, 2, 2]$. (b) A similar construction with 15 subbands illustrating the local support of the kernels.

a knot of multiplicity r . Additionally, N_i^n are non-negative, have compact support $[a_i, a_{i+n+1}]$ and most importantly form a partition of unity:

$$\sum_{i=0}^m N_i^n(u) = 1, \quad \text{for } u \in [a_n, a_{m+1}]. \quad (3.13)$$

Using this property we define the filter kernels $\mathcal{K}_j(\lambda) = \sqrt{N_{j-1}^n(\lambda)}$ for $j = 1, \dots, N_s = m + 1$ which form a Parseval frame. In particular, with $n = 3$ and the uniform knot sequence

$$a_0 = \dots = a_3 = 0 \quad (3.14)$$

$$a_i = \frac{i-3}{N_s-3} \lambda_{max}, \quad i = 4, \dots, N_s - 1 \quad (3.15)$$

$$a_{N_s} = \dots = a_{N_s+3} = \lambda_{max}, \quad (3.16)$$

the kernels $\mathcal{K}_j(\lambda)$ are \mathcal{C}^2 for $\lambda \in (0, \lambda_{max})$ and are uniformly spaced, apart from the ends of the spectra. Figure 3.4 shows two such constructions where $\lambda_{max} = 2$.

3.5 Improving the performance of sgWSPM

In this section we present two methods for improving the performance of sgWSPM (cf. Section 2.3.1). In particular we are concerned with the spatial thresholding performed in equation (2.34).

Using (2.17) we can derive a symmetric relationship between the atoms:

$$\psi_{\mathcal{K},m}[n] = \sum_{l=1}^{N_v} \hat{\mathbf{k}}[l] \chi_l^*[m] \chi_l[n] = \psi_{\mathcal{K},n}^*[m]. \quad (3.17)$$

Applying this to (2.34) results in

$$u_i = H \left(\frac{\tilde{u}_i}{\sum_{j=1}^{N_s} \sum_{m=1}^{N_v} \hat{\sigma}_{\mathcal{K}_j,m} |\psi_{\mathcal{K}_j,i}^*[m]|} - \tau_S \right) \tilde{u}_i \quad (3.18)$$

which has the benefit that only N_s atoms need to be constructed for a particular u_i and these atoms are not needed for any other u_i . We can also note that in the absence of the absolute value, the denominator becomes a graph signal reconstruction (cf. Eq. (2.21)).

Due to this absolute value the denominator requires explicit construction of all atoms, which has been pointed out as a limiting factor and the most time consuming stage in this kind of scheme [18]. This is unpractical for large graphs even when using the approximation technique presented in Section 2.1.5. Since the atoms are intrinsic to the graph and independent of any signal, they could be precomputed and stored. However, this still becomes a problem for large graphs and systems of spectral kernels with many subbands.

Below follow two approaches for improving the performance. The first one works by reducing the number atoms that needs to be calculated and the second one by increasing the speed of the calculation.

3.5.1 Bounding the denominator

Since $\hat{\sigma}_{\mathcal{K}_j,m} \geq 0$ the denominator in (3.18) is always non-negative. A trivial way of reducing the number of atoms created is thus to directly discard all negative \tilde{u}_i as they will never pass the positive threshold τ_S . While this is useful, we can do better by imposing upper and lower bounds on the denominator, which in turn put bounds on the quotient. Quotients with an upper bound lower than τ_S will for sure not pass the thresholding and similarly, values with a lower bound above τ_S will indeed survive the thresholding. Explicit construction is then only required for the remaining u_i . The bounds of the denominator are presented in the following proposition:

Proposition 3.5.1. *Given a Parseval frame and $\hat{\sigma}_{\mathcal{K}_j, m} \geq 0$, the denominator in (3.18) is bounded by*

$$\sum_{j=1}^{N_s} |\langle \psi_{\mathcal{K}_j, i}, \hat{\sigma}_{\mathcal{K}_j} \rangle| \leq \sum_{j=1}^{N_s} \sum_{m=1}^{N_v} \hat{\sigma}_{\mathcal{K}_j, m} |\psi_{\mathcal{K}_j, i}^*[m]| \leq \sqrt{\sum_{j=1}^{N_s} \langle \hat{\sigma}_{\mathcal{K}_j}, \hat{\sigma}_{\mathcal{K}_j} \rangle} \quad (3.19)$$

where $\hat{\sigma}_{\mathcal{K}_j} \in \ell^2(\mathcal{G})$ with $\hat{\sigma}_{\mathcal{K}_j}[m] = \hat{\sigma}_{\mathcal{K}_j, m}$. The lower bound can be efficiently estimated using N_s filter operations as described in Section 2.1.6.

Proof. The lower bound follows from

$$\begin{aligned} \sum_{j=1}^{N_s} \sum_{m=1}^{N_v} \hat{\sigma}_{\mathcal{K}_j, m} |\psi_{\mathcal{K}_j, m}^*[i]| &\geq \sum_{j=1}^{N_s} \left| \sum_{m=1}^{N_v} \hat{\sigma}_{\mathcal{K}_j, m} \psi_{\mathcal{K}_j, i}^*[m] \right| = \\ &= \sum_{j=1}^{N_s} |\langle \psi_{\mathcal{K}_j, i}, \hat{\sigma}_{\mathcal{K}_j} \rangle|. \end{aligned} \quad (3.20)$$

The upper bound is given by the Cauchy–Schwarz inequality:

$$\begin{aligned} \left(\sum_{j=1}^{N_s} \sum_{m=1}^{N_v} \hat{\sigma}_{\mathcal{K}_j, m} |\psi_{\mathcal{K}_j, i}^*[m]| \right)^2 &\leq \left(\sum_{j=1}^{N_s} \sum_{m=1}^{N_v} \hat{\sigma}_{\mathcal{K}_j, m}^2 \right) \left(\sum_{j=1}^{N_s} \sum_{m=1}^{N_v} |\psi_{\mathcal{K}_j, i}[m]|^2 \right) \\ &= \sum_{j=1}^{N_s} \langle \hat{\sigma}_{\mathcal{K}_j}, \hat{\sigma}_{\mathcal{K}_j} \rangle \end{aligned} \quad (3.21)$$

where the last equality follows from

$$\begin{aligned} \sum_{j=1}^{N_s} \langle \psi_{\mathcal{K}_j, i}, \psi_{\mathcal{K}_j, i} \rangle &\stackrel{(2.10)}{=} \sum_{j=1}^{N_s} \langle \hat{\psi}_{\mathcal{K}_j, i}, \hat{\psi}_{\mathcal{K}_j, i} \rangle \\ &\stackrel{(2.17)}{=} \sum_{j=1}^{N_s} \sum_{l=1}^{N_v} \hat{\mathbf{k}}_j^2[l] \chi_l^*[i] \chi_l[i] \\ &= \sum_{l=1}^{N_v} \chi_l^*[i] \chi_l[i] \underbrace{\sum_{j=1}^{N_s} \hat{\mathbf{k}}_j^2[l]}_{\stackrel{(2.19)}{=} 1} = 1. \end{aligned} \quad (3.22)$$

In the last step we utilized the orthonormality property of the eigenvectors. \square

3.5.2 Cropping the Laplacian

If the approximation technique presented in Section 2.1.5 is used to calculate the atoms, the main operation will be sparse matrix-vector multiplications [13]. These have linear time complexity in the number of non-zero elements in \mathcal{L} [63], and consequently, a reduction in the number of non-zero elements will provide a proportional speed improvement.

Let $d_G(p, m)$ represent the geodesic distance between vertices p and m , i.e., the length of the shortest path connecting p and m . With this notation we present the following lemma:

Lemma 3.5.2 ([15, Lemma 2]). *Given a polynomial kernel $\mathcal{P} \in L^2(\mathcal{G})$ of degree M , $d_G(p, m) > M \Rightarrow \psi_{\mathcal{P},p}[m] = 0$.*

In other words, atom $\psi_{\mathcal{P},p}$ is non-zero only on a local neighborhood surrounding p . It is therefore reasonable to expect that only that part of the graph plays a role in defining the atom. Because of this we propose an improvement in speed by reducing the size of the Laplacian to only include rows and columns corresponding to vertices in the set $\mathcal{R}_{p,M} = \{m \in \mathcal{V} \mid d_G(p, m) \leq M\}$. In practice this is done by extracting the relevant rows and columns creating a smaller matrix, but here we instead set all other rows and columns to zero, which is equivalent and simplifies the theory.

Proposition 3.5.3. *For a polynomial kernel P of degree M we have*

$$\psi_{\mathcal{P},p} = P(\mathcal{L})\delta_p = P(\tilde{\mathcal{L}}_p)\delta_p, \quad p = 1, \dots, N_v \quad (3.23)$$

where

$$(\tilde{\mathcal{L}}_p)_{m,n} = \begin{cases} \mathcal{L}_{m,n} & \text{if } m, n \in \mathcal{R}_{p,M} \\ 0 & \text{otherwise.} \end{cases} \quad (3.24)$$

Proof. To prove (3.23), it is equivalent to prove that

$$\mathcal{L}^s \delta_p = \tilde{\mathcal{L}}_p^s \delta_p, \quad s = 0, \dots, M, \quad (3.25)$$

i.e., $(\mathcal{L}^s)_{m,p} = (\tilde{\mathcal{L}}_p^s)_{m,p} \forall m \in \mathcal{V}$. For this, we use a proof similar to that in [13, Lemma 5.2].

Matrix multiplication can be expressed as a sum of products:

$$(\mathcal{L}^s)_{m,p} = \sum \mathcal{L}_{m,k_1} \mathcal{L}_{k_1,k_2} \cdots \mathcal{L}_{k_{s-1},p} \quad (3.26)$$

where the sum is taken over all $s-1$ length sequences k_1, k_2, \dots, k_{s-1} with $1 \leq k_r \leq N_v$. Assume for contradiction that (3.25) is false, i.e., $\exists m \in \mathcal{V} (\mathcal{L}^s)_{m,p} \neq (\tilde{\mathcal{L}}_p^s)_{m,p}$.

Then there has to be a non-zero term in the sum above containing elements of \mathcal{L} that are zero in $\tilde{\mathcal{L}}_p$. This is not possible as it would imply the existence of a path of length less than or equal to s from p to a vertex not in $\mathcal{R}_{p,M}$, i.e. a vertex m where $d_G(p, m) > M \geq s$. \square

Chapter 4

Datasets

In the proposed scheme we make use of structural MRI data for identifying the white matter and dMRI data for constructing the white matter graph. Additionally, task-based fMRI data constitutes what is to be analyzed using the proposed scheme. In this work, the MRI data were provided by the Human Connectome Project (HCP)¹, but other datasets can be used as long as the previously mentioned MRI modalities are available.

4.1 HCP Data

The MRI data used in the project was provided by the HCP, a collaborative endeavor between multiple institutions, with the goal of studying and mapping the neural connectivity of the human brain [64, 65].

The HCP is divided into two research consortia, with different but complementary goals. The WU-Minn-Oxford consortium is lead by Washington University in St. Louis, the University of Minnesota and Oxford University, and has collected MRI data from 1 200 subjects, consisting of twin pairs and their siblings from 300 families. The MGH/Harvard-UCLA consortium is lead by the Massachusetts General Hospital, Harvard University and the University of California in Los Angeles, and focused its efforts in further developing the acquisition and processing methods employed on MRI data.

The HCP makes available multiple datasets for a single subject, including structural, task-based fMRI, resting-state fMRI, and diffusion data. The data is provided

¹Data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

in a preprocessed form that facilitates its use in research. A description of the extensive preprocessing pipeline can be found in [66].

This project makes use of the structural, task-based fMRI and dMRI data for subject 100307. The specific task-based fMRI dataset used was generated from subjects completing a motor task.

4.2 Preprocessing steps

The HCP datasets are provided in a mixture of three spatial resolutions, and two different spaces (ACPC and MNI): 0.7 mm^3 ACPC for the structural data, 1.25 mm^3 ACPC for the diffusion data and 2 mm^3 MNI for the fMRI data. ACPC is the native space of the subject and MNI is a standard space provided by the Montreal Neurological Institute (MNI). For the purposes of this project it is necessary to reconcile the three datasets into a single set of working parameters. However, the resampling process and the nonlinear conversion between ACPC and MNI spaces have the potential of negatively affecting the data. The number of voxels is also a parameter to be taken into account, as it determines to a great extent the memory use and computation time of the various processing steps.

From these considerations, it was deemed that the diffusion data was the most sensitive to alteration, and that in general upsampling was less damaging to the data than downsampling. Therefore, the chosen parameters were 1.25 mm^3 ACPC, which additionally provide a suitable balance between resolution and computational tractability.

The relevant structural data to be converted was a parcellated volume with labeled white and gray matter regions (`aparc+aseg.nii`). After downsampling, the cerebral white matter and corpus callosum regions were extracted to create the binary white matter mask.

The fMRI data is provided as a temporal series of volumes. In order to convert it into the working space it was necessary to apply a nonlinear transformation from MNI space into ACPC, which requires the use of a displacement map, included with the dataset.

4.3 Semi-synthetic data

To test the performance of the diffusion-adapted filtering we construct a set of semi-synthetic phantoms for simulating streamline-shaped activations. We denote the phantoms as semi-synthetic since the activation patterns were derived from real diffusion data. The phantoms consist of random activations diffused along streamlines gathered from deterministic tractography. A clean version is used as ground truth

while Gaussian noise is added to actual phantoms. Below follows a precise description of the process.

Let $\mathcal{S} = \{S_i(x) \in \mathbb{R}^3, i = 1 \dots N_S\}$ denote the set of N_S streamlines extracted using deterministic tractography, where $S_i(x)$ represents the i -th streamline, composed of L_{S_i} points, and $x = 1 \dots L_{S_i}$ is the index of those points.

A random subset $\mathcal{T} \subseteq \mathcal{S}$ of streamlines is selected to produce a phantom. Let $\mathcal{I}_{\mathcal{T}}$ be the index of the streamlines that form \mathcal{T} , i.e., $\mathcal{T} = \{S_i(x) \in \mathbb{R}^3, i \in \mathcal{I}_{\mathcal{T}}\}$. The streamlines in this set are voxelized, forming a new set $\mathcal{U} = \{\mathbf{u}_i \in \mathbb{N}^{L_{\mathbf{u}_i}}, i \in \mathcal{I}_{\mathcal{T}}\}$, where the \mathbf{u}_i are vectors containing indices of the voxels through which the original streamline passes.

To model the activation centers of the voxelized streamlines, indicator vectors \mathbf{a}_i , of the same length as the corresponding \mathbf{u}_i , are constructed. Each indicator vector has one random element set to one, corresponding to the voxel where the activation center in the associated streamline is found, while the rest of the elements are set to zero.

After defining the activation center for each streamline, the activation is diffused along the length of the streamline. In order to do this, a stochastic matrix \mathbf{C}_i is defined, describing the connections between the voxels in \mathbf{u}_i . Voxels are connected to themselves and to the voxels in their 26-neighborhood, all with equal weight. For example, if voxel $\mathbf{u}_i(k)$ is adjacent to voxels $\mathbf{u}_i(k-1)$ and $\mathbf{u}_i(k+1)$, \mathbf{C}_i will on the k^{th} row have three elements with the value $1/3$ at columns $k-1$, k and $k+1$.

Having defined the indicator vector and the stochastic matrix for each streamline, a diffuse activation pattern \mathbf{p}_i can be calculated by the following formula:

$$\mathbf{p}_i = \frac{1}{c_i} \mathbf{C}_i^{m_i} \mathbf{a}_i \quad (4.1)$$

where c_i is a normalization constant such that $\max_k \mathbf{p}_i(k) = 1$ and $m_i \in \mathbb{N}$ is a parameter controlling the extent of the diffusion, which can be set independently for the different streamlines. For this work we used a value of 250.

Finally, to create the complete phantom embedded in a 3D volume \mathbf{P} , the diffuse activation patterns \mathbf{p}_i of each streamline \mathbf{u}_i are combined:

$$\mathbf{P}_j = 1 - \prod_{i \in \mathcal{I}_{\mathcal{T}}} \sum_{k \in \mathcal{K}_i} [1 - \mathbf{p}_i(k)] \quad \text{where } \mathcal{K}_i = \{k | \mathbf{u}_i(k) = j\} \quad (4.2)$$

where j is a linear indexing of the voxels in \mathbf{P} . We treat \mathbf{P} as our ground truth activation patterns. Noise added versions of \mathbf{P} are obtained by convolving it with additive zero-mean Gaussian noise of standard deviation σ_e :

$$\mathbf{P}_G = \mathbf{P} + \mathbf{e}, \quad \mathbf{e} \in \mathcal{N}(\mathbf{0}, \sigma_e^2 \mathbf{I}). \quad (4.3)$$

A time-series phantom can then be created by combining a single volume phantom P with different realizations of random noise.

Chapter 5

Results

In this chapter, we present results obtained using the proposed methods. A single subject (subject ID: 100307) from the HCP is used throughout. MRI data from the subject is used to define the white matter graph, and also serves as real data when performing activation mapping.

The white matter graph was constructed as explained in Section 3.1. The ODFs were reconstructed with DSI Studio (<http://dsi-studio.labsolver.org>), using the GQI [54] algorithm with default parameters. The ODFs were sharpened by setting p_{ODF} to 40, as this value showed good results in some initial tests.

5.1 Diffusion adapted atoms

In this section, we show some examples of how the atoms in the white matter graph are adapted to both the topology of the white matter, i.e., tissue boundaries, and to the varying microstructure within white matter. We start by inspecting how the atoms behave in an unweighted graph defined on the whole brain volume. In such a graph the atoms are not restricted to diffuse within the white matter alone, but instead diffuse isotropically. Such an approach mimics the results obtained by using a conventional isotropic Gaussian filter. A contour plot of such an atom is shown in Figure 5.1b.

In Figure 5.1c, an unweighted version of the white matter graph (cf. Section 3.1) is used. It can be seen that the atom is confined to the boundaries of the white matter, but is otherwise unrestricted and free to diffuse in any direction. Defining the graph on the white matter only makes the resulting atoms adapt to the white matter topology, but they still diffuse isotropically instead of reflecting the tissue microstructure.

In Figure 5.1d the proposed white matter graph is used. It can be seen that the atom adapts not only to the topology of the white matter, but also to the microstruc-

ture of white matter tissue. This microstructure is presented in Figure 5.1e) as the directions of the axonal fibers found within white matter. Figures 5.2 and 5.3 show additional atoms in the same coronal slice.

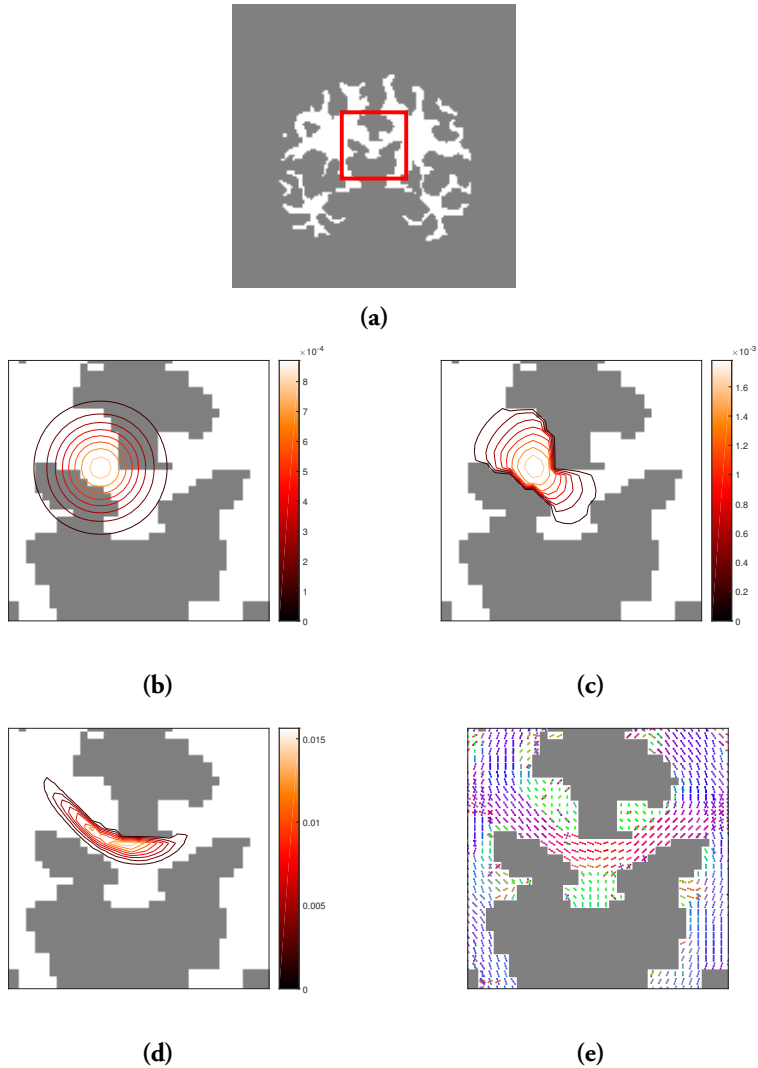


Figure 5.1: Coronal slices of atoms resulting from the heat kernel in different graphs. (a) marks the region being shown in (b)-(d). In (b) the graph is unweighted and is not confined to the white matter, causing the atom to diffuse isotropically in all directions. In (c) the graph is still unweighted but is only defined on the white matter. This causes the atom to adapt to the topology of the graph. In (d) the white matter graph presented in Section 3.1 is used and, as can be seen, the atom adapts to both the topology of the graph and the diffusion data. (e) shows the white matter fiber directions. Red, green and blue lines indicates fiber directions normal to the sagittal, coronal and axial planes, respectively.

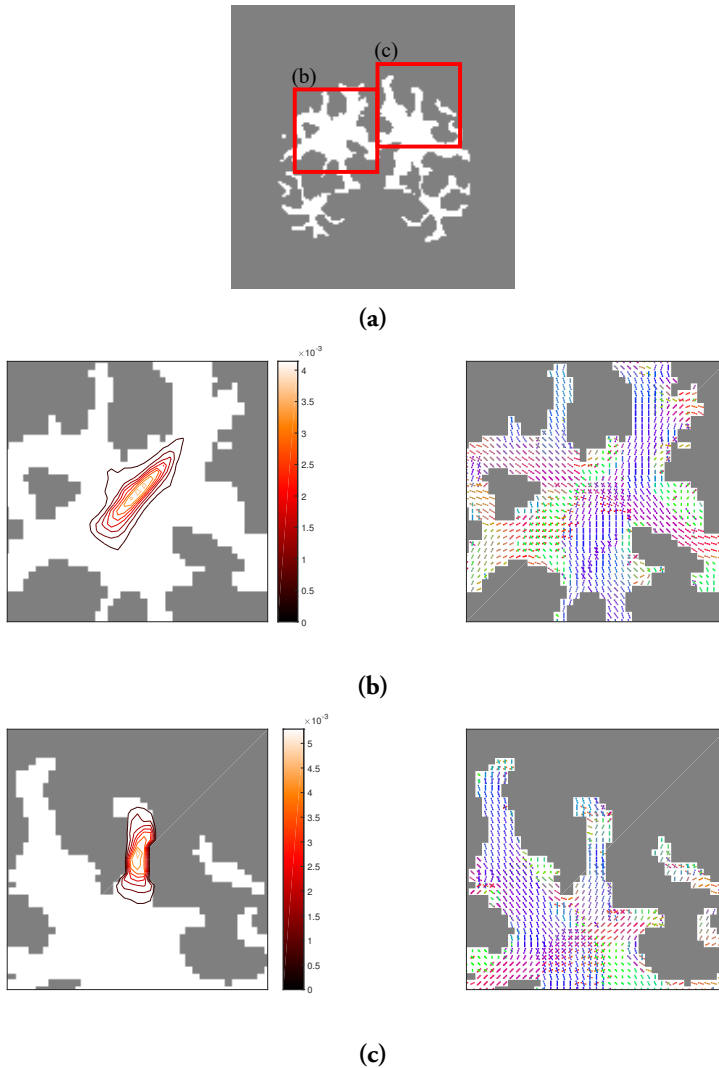


Figure 5.2: (a) indicates what regions are shown in (b) and (c). (b) and (c) shows two additional atoms from the coronal slice shown in Figure 5.1a together with the fiber directions.

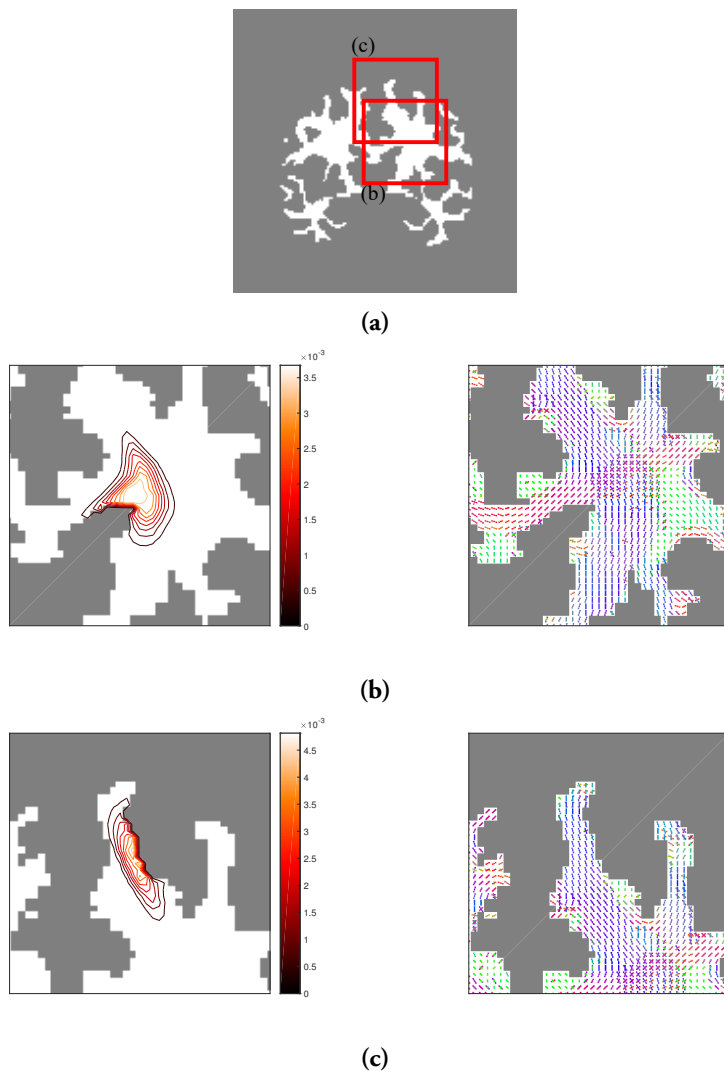


Figure 5.3: (a) indicates what regions are shown in (b) and (c). (b) and (c) shows two additional atoms from the coronal slice shown in Figure 5.1a together with the fiber directions.

5.2 Signal-adapted filter design

To test the signal-adapted approach described in Section 3.3, we adapted a seven subband UMT system of kernels to the spectral characteristics of a set of task-based functional data. Specifically, we first selected ten random volumes from the motor task acquired with right-to-left phase encoding, and then extracted our set of signals $\{\mathbf{f}_s\}_{s=1}^{10}$ as the signals defined on the white matter voxels of these volumes. The ensemble energy spectral density was then estimated with (3.9), using a B-spline frame consisting of $N_s'' = 50$ subbands. As can be seen in Figure 5.4, most of the energy is associated to eigenvalues in the lower range of the spectrum, leading to the UMT system of kernels being warped accordingly (see Figure 5.5).

Next, we want to verify whether the kernels of the warped filter design capture and approximately equal amount of energy from the signals. We start with a single signal \mathbf{f}_s and remove the DC component as in (3.8) to produce the signal \mathbf{g}_s . The relative energy captured by each subband is then calculated as

$$e_s[j] = \sum_{m=1}^{N_v} \left| \left\langle \psi_{\mathcal{K}_j, m}, \frac{\mathbf{g}_s}{\|\mathbf{g}_s\|} \right\rangle \right|^2, \quad j = 1, \dots, 7. \quad (5.1)$$

Figure 5.6 shows e_s for the ten sampled signals before and after the warping. The energy is much more evenly distributed across the subbands in the signal-adapted system of kernels design than in the original UMT system of kernels.

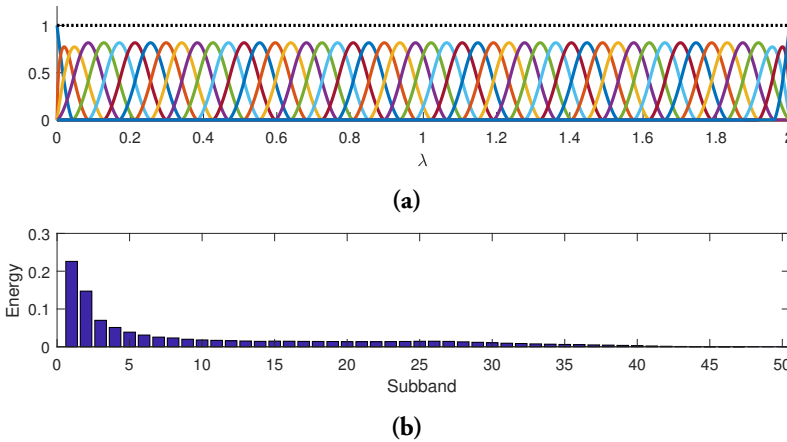


Figure 5.4: (a) A B-spline frame with 50 subbands. (b) The estimated ensemble energy spectral density.

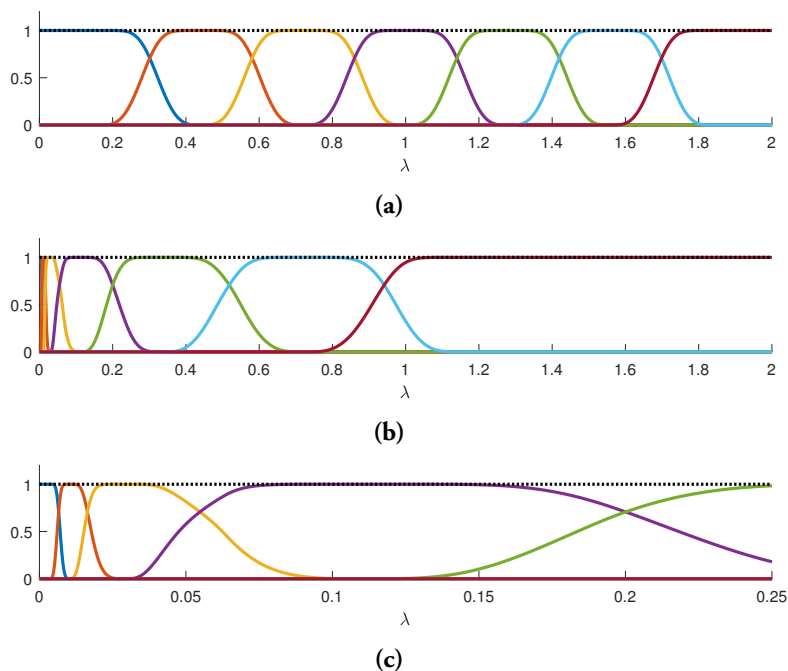


Figure 5.5: (a) A UMT system of spectral kernels with 7 subbands. (b) Signal-adapted version of the system of kernels in (a). (c) The design in (b) within the spectral range 0–0.25.

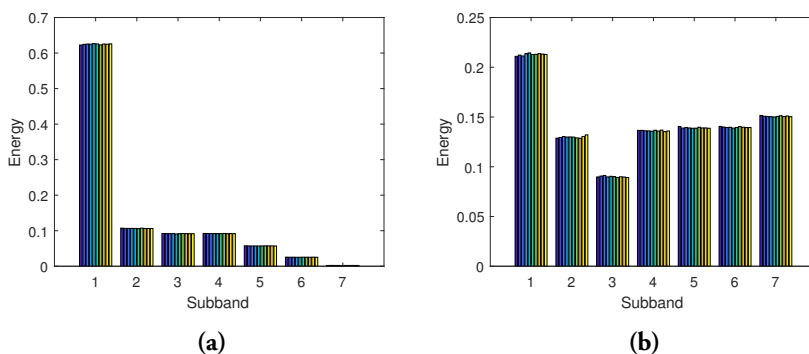


Figure 5.6: Relative energy distribution for the ten selected volumes of fMRI data before (a) and after (b) adapting the system of kernels to the spectral content of the signals.

5.3 Performance improvements to sgWSPM

The improvements presented in Section 3.5 had a significant effect on the resulting execution time of wavelet SPM (WSPM). Table 5.1 shows the result from using the bounds in Section 3.5.1. The explicit calculation of the vast majority of the atoms could be avoided by recognizing that all voxels with negative \tilde{u}_i are non-activations. The upper bound provided an additional improvement while the lower bound provided no improvement. In the end, only 2.9 % and 2.1 % of the atoms were required to be calculated explicitly for the semi-synthetic and real datasets, respectively. These datasets are described in the next section.

To measure the improvement from cropping the Laplacian as presented in Section 3.5.2, we constructed atoms associated with a B-spline spectral kernel design with seven subbands. The kernels were approximated using polynomials (cf. Section 2.1.6) of degrees in the range 5–50. A random subset of 100 white matter voxels were then chosen and the associated atoms were calculated, resulting in a total of 700 atoms. The measured execution times are shown in Figure 5.7. The proposed approach proved most useful for lower polynomial degrees. For example, for a polynomial degree of 5, the calculated speed was an order of magnitude faster. For polynomial degrees above 25 the proposed approach had an opposite effect and lead to an increase in the execution time.

	Semi-synthetic dataset	Real dataset
Total number of atoms	207829	207829
Non-activations due to $\tilde{u}_i \leq 0$	153377	154601
Non-activations due to upper bound $< \tau_S$	201766	203530
Activations due to lower bound $> \tau_S$	0	0
Atoms required to be calculated	6063	4299

Table 5.1: The result of implementing the bounds presented in Section 3.5.1 when running sgWSPM using the signal-adapted system of kernels shown in Figure 5.5b.

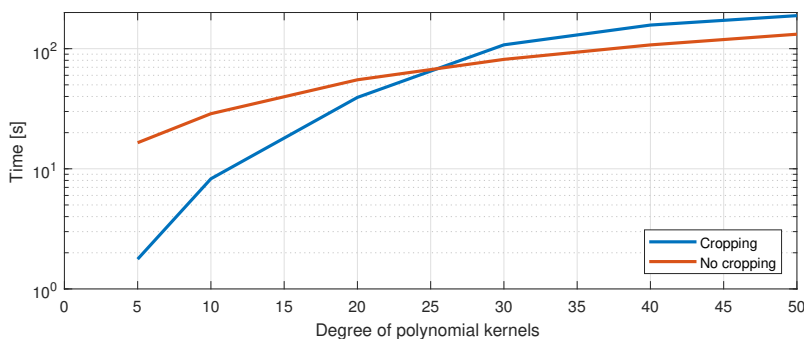


Figure 5.7: Execution time when calculating 700 atoms using polynomial kernels of different degrees. The calculations were performed on an ordinary desktop computer.

5.4 Experimental results

The capacity for discerning streamline-shaped activations of the two proposed diffusion-adapted filtering approaches was tested against that of an isotropic 4mm full width half maximum (FWHM) Gaussian filter. This was done using a semi-synthetic time-series phantom. The filtering approaches were also tested on real fMRI data from the HCP.

Three different approaches to the filtering of white matter fMRI data were implemented:

- The data was masked with the white matter and subjected to the regular SPM pipeline, using an isotropic 4mm FWHM Gaussian kernel for smoothing. This approach represents a typical processing pipeline applied to fMRI data, and will be referred to as *SPM*.
- A white matter graph was defined in the manner described in 3.1, and used to filter the data with a heat kernel. Afterwards, the filtered data was processed with the regular SPM pipeline. This approach will be referred to as *gSPM*.
- The data was processed with the *sgWSPM* framework described in 2.3.1, using the white matter graph from 3.1. A UMT system of seven spectral kernels was used to perform the signal decomposition, and it was adapted to the energy content of the signal using the procedure described in 3.3. This approach will be referred to as *sgWSPM*.

In order to match the amount of smoothing done by the SPM and *gSPM* approaches. This was done by first filtering the real data volume series with the isotropic

Gaussian filter, and then using the SPM toolkit to estimate the smoothness of the resulting filtered volumes. A value for τ was then chosen that produced a similar amount of smoothness when filtering the volume series with gSPM. For a FWHM of 4 mm, this resulted in a τ of 2.2.

5.4.1 Semi-synthetic data

The filtering approaches were tested on a synthetic time series designed to mimic real fMRI data. This time series had a length of 100 volumes, and consisted of a Gaussian noise ($\sigma = 1$) background with two stretches of 18 volumes where activity was present. The activations were created from a single 100 streamline phantom (see Figure 5.8) produced using the method described in Section 4.3. The regressor corresponding to the periods of activity in the time series is shown in Figure 5.9.

The ensemble energy spectral density of the time-series phantom was estimated in order to adapt the spectral kernels to the signal in the sgWSPM approach. The ensemble energy density and the resulting signal-adapted system of kernels are shown in Figure 5.10. The signal-adapted system of spectral kernels was found to be inadequate (see Section 6), and thus a seven subband uniform B-spline kernel frame (identical to Figure 3.4a) was used for the remainder of the work.

A ground truth volume was created by thresholding and binarizing the 100 streamline phantom at a level of 0.5. The performance of the three approaches was then evaluated by thresholding at multiple levels the t-value parametric produced by each approach and comparing the resulting detections with the ground truth volume. Each threshold level produced a true positive and a false positive rate for each approach, and these were collected in a receiver operating characteristic (ROC) curve, which relates the two metrics for a given threshold value. Figure 5.11 shows the ROC curves for the three approaches.

Figures 5.12, 5.13 and 5.14 show the detected activations at a significance level of $\alpha = 0.05$ for the three approaches.

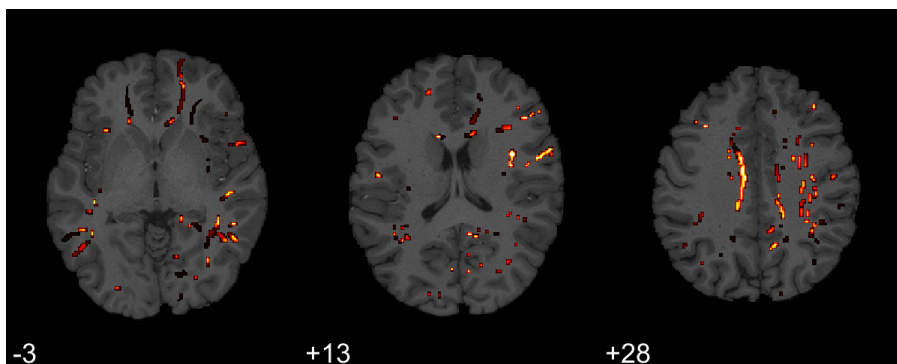


Figure 5.8: Semi-synthetic phantom used in the construction of the synthetic time series.

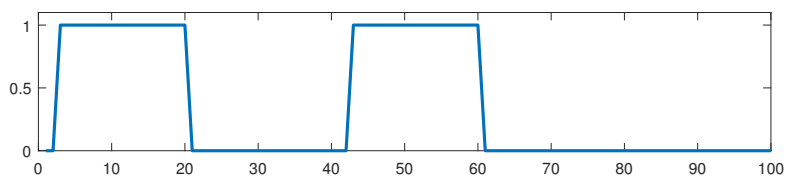
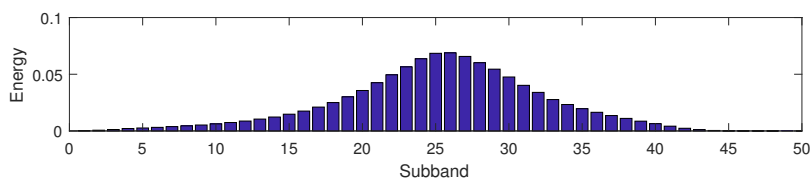
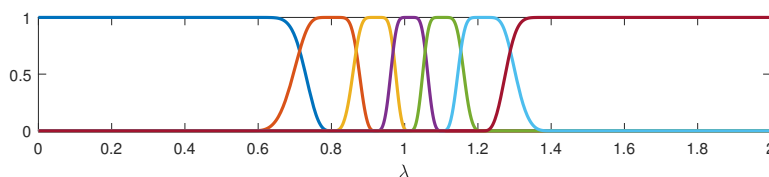


Figure 5.9: Regressor representing the periods of activity in the time series.



(a)



(b)

Figure 5.10: (a) Ensemble energy spectral density of the time series phantom. (b) Signal adapted system of spectral kernels.

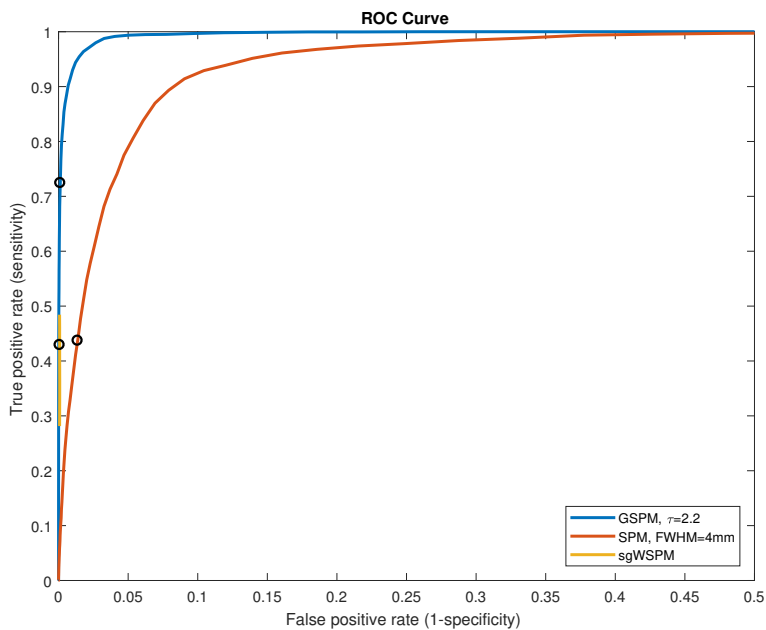


Figure 5.11: ROC curve for all three approaches. The circled positions show results obtained at statistical significance level $\alpha = 0.05$. The dashed line represents an equal true and false positive rate.

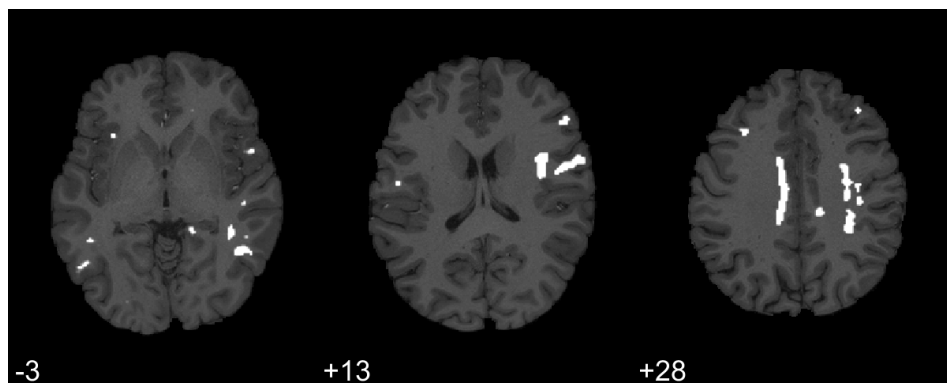


Figure 5.12: Activations from phantom detected with SPM, 4mm FWHM.

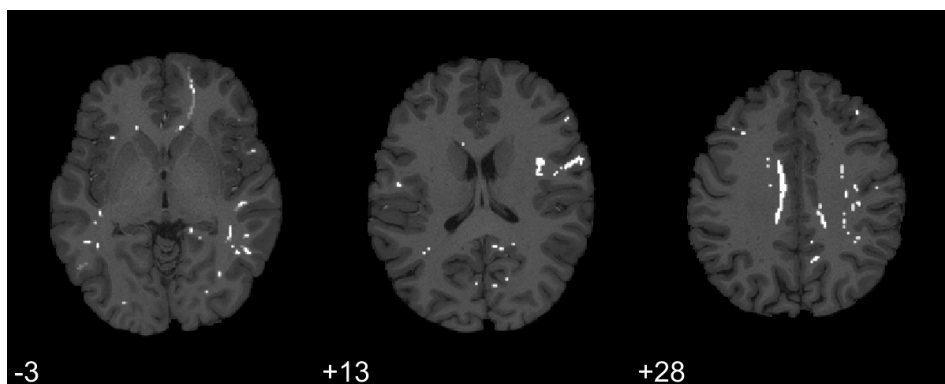


Figure 5.13: Activations from phantom detected with gSPM, $\tau=2.2$.

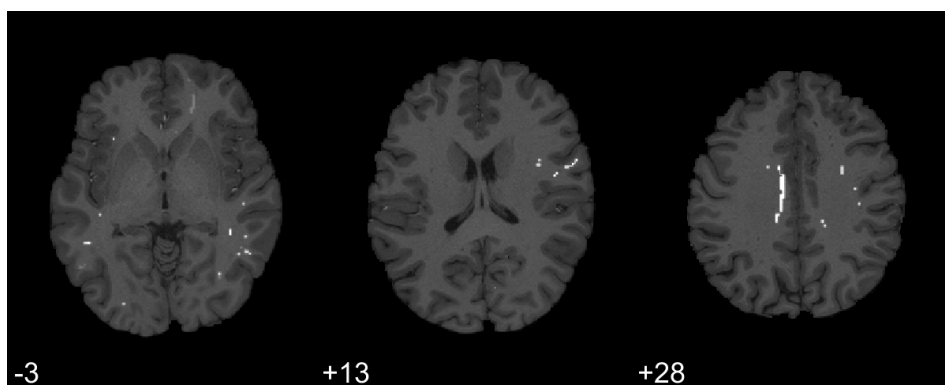


Figure 5.14: Activations from phantom detected with sgWSPM.

5.4.2 Real data

The fMRI data from the HCP was subjected to the three filtering approaches described, with the statistical significance level α set at 0.05. The paradigm involved five different event types, of which two produced significant activations patterns (right hand and tongue movement) [67]. The fMRI data was masked to include only the white matter, so gray matter activations are absent from the results.

The SPM approach used a corrected statistical significance level obtained from GRFT, while Bonferroni correction was used for both gSPM and sgWSPM.

The following set of figures presents the results obtained using the three approaches. The detected activations are overlaid on a T1 weighted structural scan of the subject. The Results for the ordinary SPM approach are presented in Figures 5.15 and 5.16, for graph SPM in Figures 5.17 and 5.18, and for sgWSPM in Figures 5.19 and 5.20.

Some of the detected activations appear very close to the white-gray matter boundary, and can be reasonably assumed to be the result of nearby activations in gray matter. However, some isolated white matter activation were detected, most notably in the corpus callosum, which can be seen in slices +13 to +15 for the right hand event results and slices +4 to +10 for the tongue results.

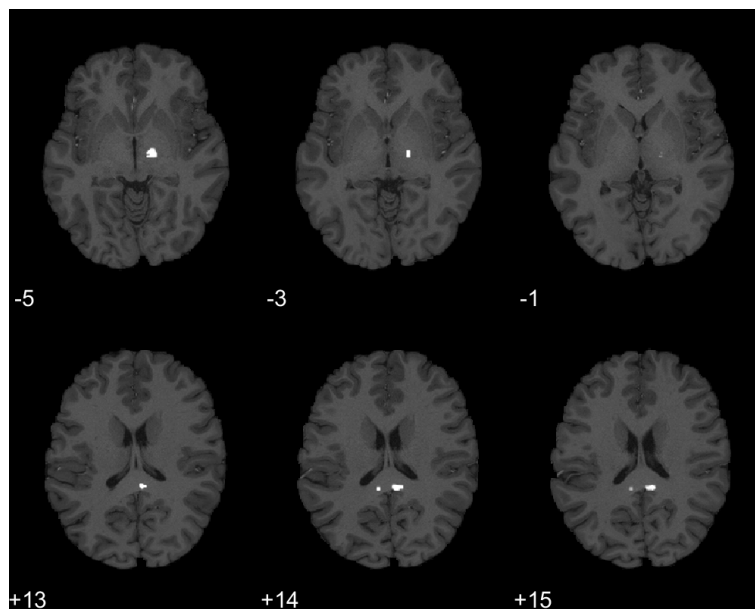


Figure 5.15: Activations from right hand event detected with SPM, 4mm FWHM.

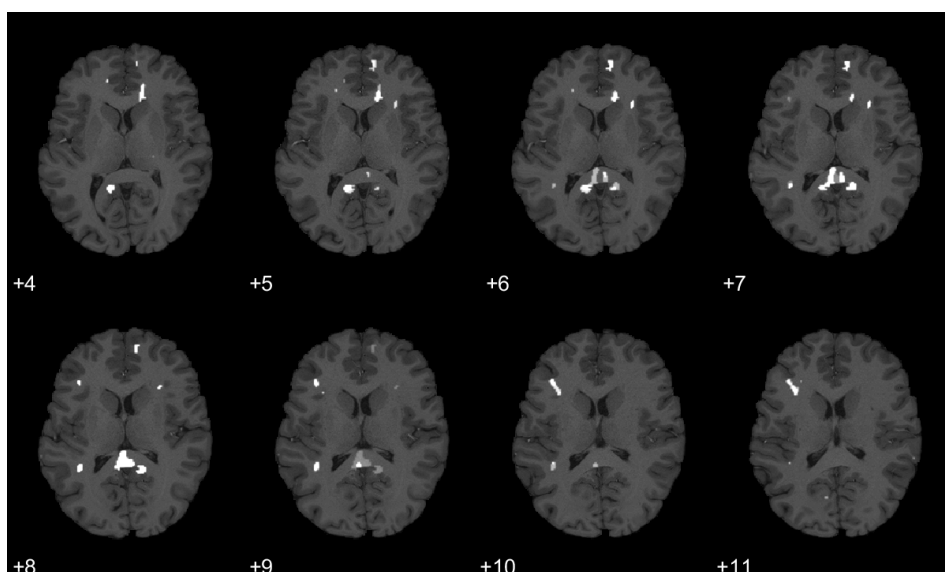


Figure 5.16: Activations from tongue event detected with SPM, 4mm FWHM.

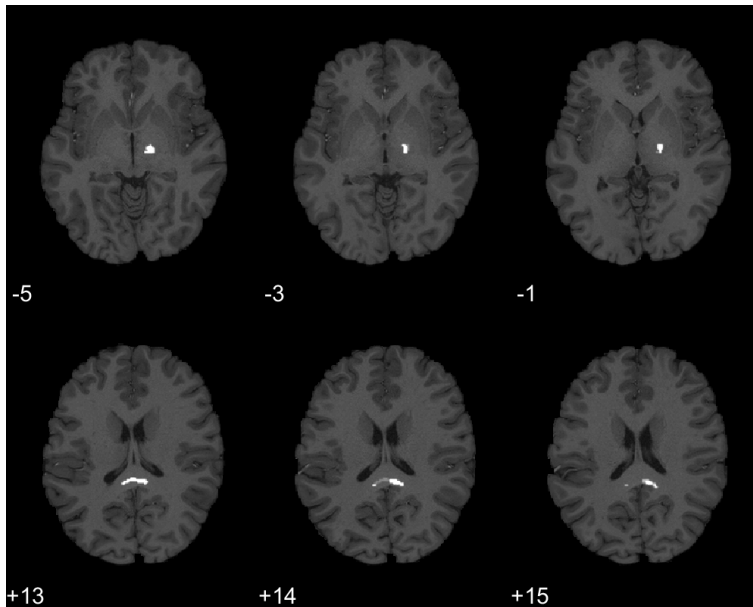


Figure 5.17: Activations from right hand event detected with gSPM, $\tau=2.2$.

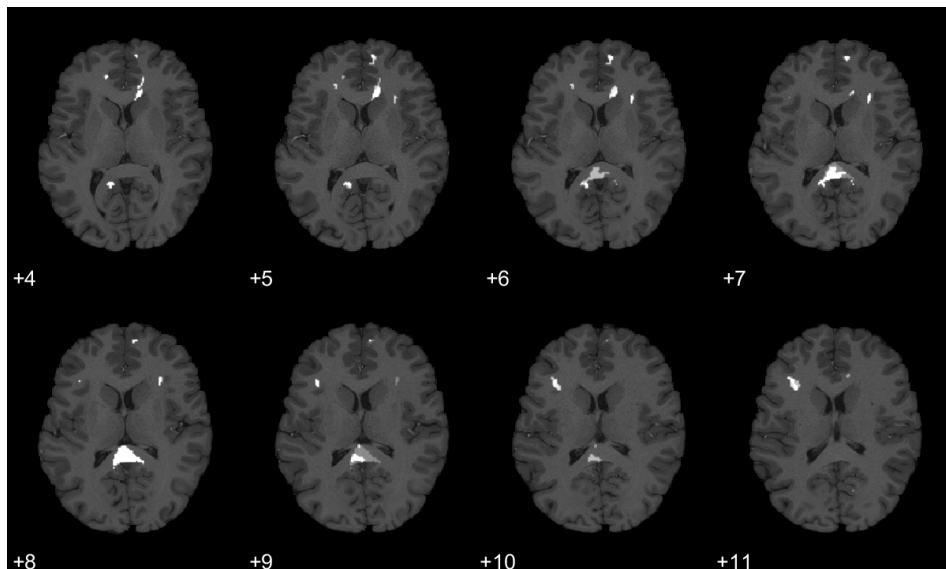


Figure 5.18: Activations from tongue event detected with gSPM, $\tau=2.2$.

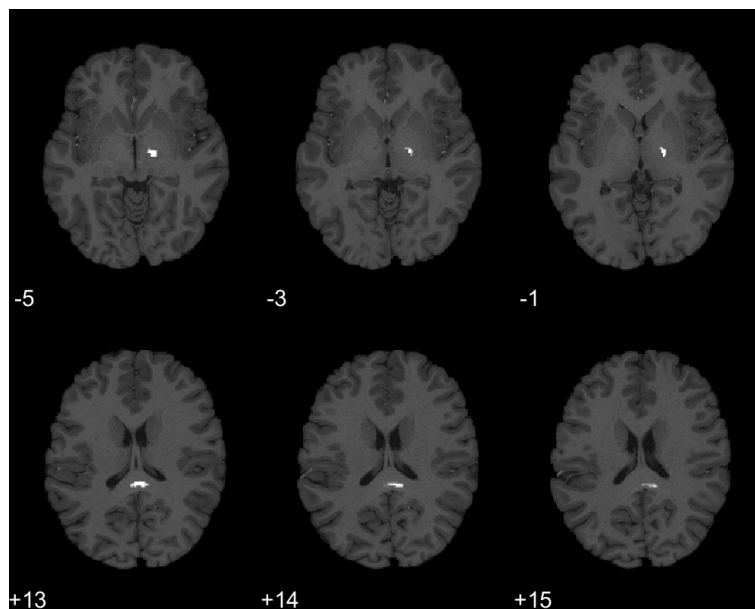


Figure 5.19: Activations from right hand event detected with sgWSPM.

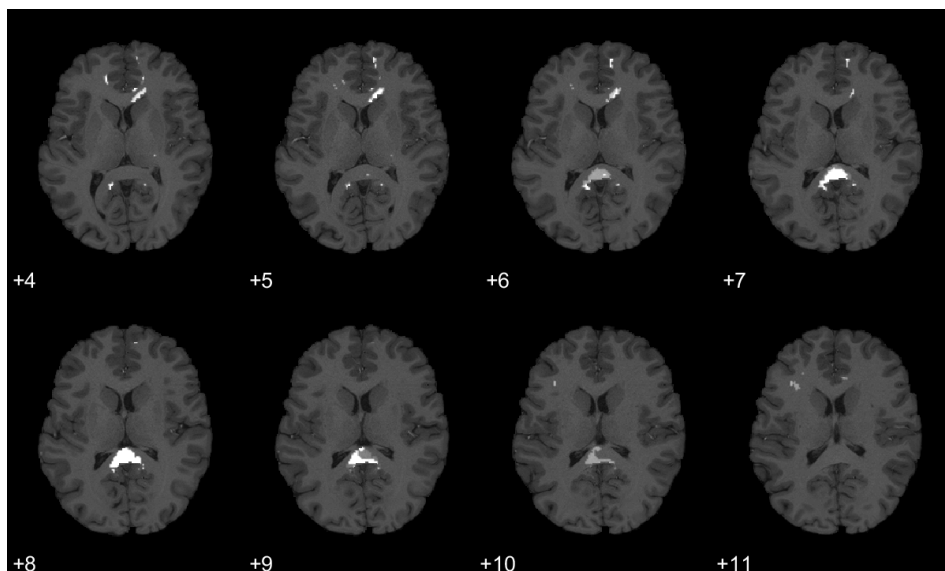


Figure 5.20: Activations from tongue event detected with sgWSPM.

Chapter 6

Discussion

Figures 5.1, 5.2 and 5.3 show that the atoms produced in our white matter graph construction indeed adapt to both the topology and the microstructure of the white matter. This results in smoothing along the axonal fibers which is inline with our assumption that the BOLD signals have a greater spatial correlation in those directions.

6.1 Performance improvements to sgWSPM

The combined improvement of the two proposed approaches will vary depending on the functional data analyzed and the filter design used, but the results presented here indicate that it is not unreasonable to achieve decreases in the execution time of the spatial thresholding (cf. (2.34)) by a factor of 400. The performance improvement is particularly significant when using spectral kernels with low requirements for their polynomial estimation, such as the presented B-spline system of spectral kernels. This contribution was imperative to running sgWSPM on the size of graphs we worked on.

As was shown in Table 5.1 the lower bound provided no reduction in the number of atoms that needed to be calculated explicitly. The reasons for this is that the bound was simply too conservative, being an order of magnitude lower than τ_S for the real dataset. However, even if the bound had been less conservative the performance gain would have been limited as the number of activations found in the white matter is relatively few compared to the total number of voxels.

Cropping the Laplacian offered a great improvement but only for polynomial kernels of degree lower than 25. This is to be expected. When not cropping \mathcal{L} the execution time increases linearly in the degree M . However, when the cropping is performed the relation becomes more complex as the set $\mathcal{R}_{p,M}$ (cf. Section 3.5.2) grows cubically in M . At some point the overhead of exploring the graph to determine

$\mathcal{R}_{p,M}$ cancels any potential gain. This is clearly the case, since for a sufficiently large M $\mathcal{R}_{p,M} = \mathcal{V}$. In our white matter graph the longest geodesic distance between any pair of voxels (i.e. the *diameter* of the graph) is around 50. Thus, for $M > 50$ the cropping will for sure not improve the performance, and in the case of a very central voxel neither will an $M > 25$.

6.2 Semi-synthetic results

The ROC curve in Figure 5.11 shows the increased sensitivity of graph-based SPM (gSPM) compared to an isotropic Gaussian filter. In addition to this, comparing Figures 5.12 and 5.13 reveals a significantly improved spatial specificity in the detections. These results become even more positive when considering that while SPM employed the more lenient corrected statistical significance level obtained from GRFT, gSPM managed to obtain better results using the more strict Bonferroni correction.

The specificity of sgWSPM can be seen to be high both from the ROC curve results and from the sharpness of the detected streamlines in Figure 5.14. However, sgWSPM showed significantly lower sensitivity than the two other approaches. These results can be caused by the use of Bonferroni correction in the determination of significant activations, leading to excessively conservative detections.

6.3 Real data results

The performance of the three filtering approaches matches the conclusions from the semi-synthetic data results. The activations produced by gSPM are generally narrower and have a slightly greater spatial extent than those of SPM. sgWSPM also produces activations with greater spatial specificity, but has a lower sensitivity than the other approaches.

The results in general showed few and small activations, mostly located close to gray matter regions, with the notable exception of the corpus callosum. This is not unexpected, as according to [3] most reports of BOLD fMRI activations in white matter involve the corpus callosum.

6.4 Signal-adapted kernels

The use of a signal-adapted approach was found to produce unsatisfactory results for the semi-synthetic data. The purpose of the method is to find spatial patterns with high energy in the data. However, the determination of what constitutes an activation can only be made in regard to the temporal changes in the data. Therefore, a system

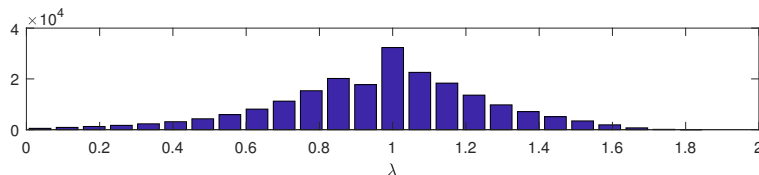


Figure 6.1: Eigenvalue density distribution of the white matter graph.

of kernels that optimally adapts to the spatial distribution of energy is not necessarily optimal at detecting temporal activations. If there is a large amount of noise present, the signal-adapted system of kernels will adapt to the noise rather than the signal, in which case spatial components that are significant in the temporal domain can be removed due to being insignificant in the spatial domain.

In our case, the amount of noise seems to be excessive. The energy distribution of the phantom (Figure 5.10a) mostly reflects the white noise adapting to the distribution of the graph eigenvalues (see Figure 6.1), which are centered around $\lambda = 1$. The relevant signal, however, is known to be concentrated in the low end of the spectrum, which is what justifies the use of a Gaussian filter to enhance detections. Due to this, the signal-adapted kernels failed to produce meaningful detections. Nevertheless, the use of a more realistic noise model based on rsfMRI would have resulted in a time series whose spectral profile is more strongly concentrated in the low spectral range. Consequently, the signal-adapted system of kernels would have been centered around the lower end of the spectrum, in which case we expect the approach would have produced better results.

In the case of real data, the signal-adapted system of kernels was concentrated around the low spectral range, which, we hypothesize, should have made it suitable for discerning the types of activations present in the fMRI data. Nevertheless, the procedure did not detect any significant activations for the right hand event. The tongue event was not tested due to the long computation time required to run signal-adapted sgWSPM with a UMT system of kernels. The cause of such unexpected behavior is not clear to us, but multiple runs of sgWSPM with different systems of kernels could clarify this issue. The fact that a non-signal-adapted B-spline system of kernels managed to produce results may suggest that the problem resides in the UMT system of kernels and not in the signal adaptation itself.

6.5 Conclusions

The activation shape arising from the BOLD signal is fundamentally different between white matter and gray matter. We have developed a graph-based description of the

white matter that is capable of encoding the underlying structure of the axonal fibers. Based on this description, we have examined two approaches to the filtering of white matter fMRI data, achieving better results than those of the commonly used isotropic Gaussian filtering with one of them. In addition, several general developments were made for the use of sgWSPM with graph signals, including an approach to the estimation of the ensemble energy spectral density of a volume series, a B-spline frame with very low requirements for polynomial approximation, and several performance improvements on the framework itself.

6.6 Future work

There are a number of ways in which the work done in this thesis can be expanded upon. An immediate improvement would be achieved by testing the filtering approaches on a larger number of subjects and paradigms, in order to guarantee that the results seen here are not spurious, but an actual reflection of the performance of the methods.

The white Gaussian noise used in the semi-synthetic dataset may not ideally represent the types of noise seen in rel fMRI data. Therefore, it could be interesting to explore different noise models, for example, using resting state fMRI (rsfMRI) [42, 68].

An interesting avenue for development would be to extend the graph description to the whole brain. This would require careful consideration on the weighting scheme of the graph in order to adequately represent both gray and white matter, and could potentially show activations that extend isotropically through the gray matter but become anisotropic after transitioning into white matter.

The signal-adapted filter design could be modified to adapt to the spectral components that show temporal correlation with the regressors of the general linear model (GLM). This could result in a more meaningful decomposition and avoid the problems that arise from the mismatch between the energy distribution of the signal in the spectral domain and the temporal correlation of certain spectral components to a given regressor.

The use of bounds in the spatial thresholding (cf. (2.34)) is something that potentially could be applied in other methods. For example, in [60] the Euclidean norm of the atoms were used to determine a threshold used for denoising in the graph spectral domain. This too requires explicit calculation of the atoms and thus suffers from the same limitations as sgWSPM.

The B-spline frame could possibly be developed further. Although the knots are uniformly space, the kernels do not have the same L^1 -norm nor L^2 -norm. This could be desirable properties and might be achievable by using a non-uniform knot

sequence.

Figure 5.3b illustrates a possible limitation in our approach in that it does not explicitly handle crossing fibers. An atom localized in a region of crossing fibers will spread along all fiber bundles which may or may not be desirable. More research is need to determine the spatial correlation in the fMRI data in such regions.

Chapter A

Appendix

A.1 The heat kernel and its Gaussianity

In this section, we reintroduce the heat kernel as a spectral graph kernel, and show that its associated atoms are approximately Gaussian on lattice graphs. To do this we utilize some nice properties of circulant graphs. Throughout this section zero-based indices are used to simplify the notation.

The heat kernel is given by

$$\mathcal{K}(\lambda) = e^{-\tau\lambda} \quad (\text{A.1})$$

where τ is a free parameter.

A.1.1 Circulant graphs

A circulant graph has a circulant adjacency matrix where each row is a circular shift of the one above. Consequently, the Laplacian is circulant as well:

$$\mathcal{L} = \begin{bmatrix} c_0 & c_1 & c_2 & \cdots & c_{N_v-1} \\ c_{N_v-1} & c_0 & c_1 & \cdots & c_{N_v-2} \\ c_{N_v-2} & c_{N_v-1} & c_0 & \cdots & c_{N_v-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_1 & c_2 & c_3 & \cdots & c_0 \end{bmatrix}. \quad (\text{A.2})$$

In the case of an undirected graph, \mathcal{L} is symmetric and $c_m = c_{N_v-m}$, $m = 1, \dots, N_v - 1$. A useful property of circulant matrices is that the eigenvectors and eigenvalues are given by [69]

$$\chi_l[m] = \frac{1}{\sqrt{N_v}} W_l^m, \quad (\text{A.3})$$

$$\lambda_l = \sum_{m=0}^{N_v-1} c_m W_l^m, \quad (\text{A.4})$$

where $W_l = e^{-2\pi il/N_v}$. Worth noting is that the eigenvector matrix $\chi = [\chi_0, \dots, \chi_{N_v-1}]$ form the transformation matrix for the discrete Fourier transform (DFT) and that the eigenvalues λ_l are the (unnormalized) DFT of the vector $\mathbf{c} = [c_0, \dots, c_{N_v-1}]$. Additionally, only the eigenvalues depend on \mathbf{c} , meaning that the eigenvectors are the same for all circulant matrices.

A.1.2 The Gaussianity of the atoms

The construction of an atom can be seen as the DFT of the kernel $\hat{\mathbf{k}}[l]$ modulated by $\chi_l[m]$:

$$\begin{aligned} \psi_{\mathcal{K},m}[n] &\stackrel{(2.17)}{=} \sum_{l=0}^{N_v-1} \hat{\mathbf{k}}[l] \chi_l^*[m] \chi_l[n] \\ &\stackrel{(A.3)}{=} \frac{1}{\sqrt{N_v}} \sum_{l=0}^{N_v-1} \hat{\mathbf{k}}[l] e^{2\pi ilm/N_v} \chi_l[n]. \end{aligned} \quad (\text{A.5})$$

The modulation results in a circular shift property in the resulting atoms¹:

$$\psi_{\mathcal{K},m}[n] = \psi_{\mathcal{K},0}[(n - m) \bmod N_v]. \quad (\text{A.6})$$

The resulting atoms will have the same vertex spread relative to the vertex they are located at. Thus, it is sufficient to show that $\psi_{\mathcal{K},0}$ is Gaussian. We can also use the property that the Fourier transform of a Gaussian is itself a Gaussian. If $\hat{\mathbf{k}}$ is Gaussian with standard deviation $\sigma_{\hat{\mathbf{k}}}$, $\psi_{\mathcal{K},0}$ is Gaussian as well with standard deviation [59, Table 1.14.1]

$$\sigma_{\psi} = \frac{1}{\sigma_{\hat{\mathbf{k}}}}. \quad (\text{A.7})$$

A.1.3 1D case - the cycle graph

A cycle graph is a circulant graph consisting of a single closed chain of nodes (See Figure A.1). Its Laplacian is defined as in (A.2) with $c_0 = 1$ and $c_1 = c_{N_v-1} =$

¹In the context of the DFT, the kernel $\hat{\mathbf{k}}$ is actually in what is commonly regarded as the time or space domain, while the atoms $\psi_{\mathcal{K},m}$ are in the frequency domain.

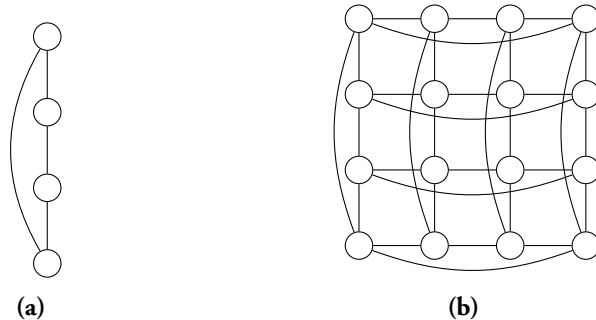


Figure A.1: (a) a cycle graph with 4 vertices and (b) the 2D analog — a torus grid graph.

$-1/2$, and has the eigenvalues

$$\lambda_l \stackrel{(A.4)}{=} 1 - \frac{1}{2}W_l - \frac{1}{2}W_l^{-1} = 1 - \cos\left(\frac{2\pi l}{N_v}\right). \quad (\text{A.8})$$

Using the truncated Taylor expansion of $\cos ax$ at $x = 0$

$$\cos ax \approx 1 - \frac{a^2 x^2}{2} \quad (\text{A.9})$$

the eigenvalues can be approximated as²

$$\lambda_l \approx \frac{1}{2} \left(\frac{2\pi l}{N_v} \right)^2. \quad (\text{A.10})$$

Inserting (A.10) into the heat kernel (A.1) yields

$$\hat{\mathbf{k}}[l] = e^{-\tau\lambda_l} \approx e^{-\tau\frac{1}{2}\left(\frac{2\pi l}{N_v}\right)^2} \quad (\text{A.11})$$

which shows that $\hat{\mathbf{k}}[l]$ is approximately Gaussian with mean zero and standard deviation $\sigma_{\hat{\mathbf{k}}} = 1/\sqrt{\tau}$. Using (A.7), we can calculate the standard deviation of the atoms as

$$\sigma_\psi = \sqrt{\tau}. \quad (\text{A.12})$$

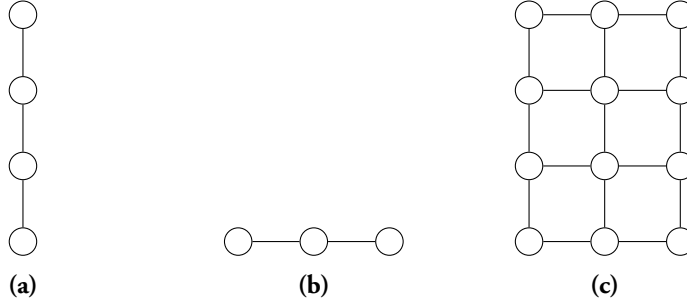


Figure A.2: If G and H are the graphs in (a) and (b) respectively, (c) shows $G \square H$.

A.1.4 2D case - the torus grid graph

The general idea here is to use the nice properties of the cycle graph and extend it to two dimensions using graph products. In the following, we will show that this generates similar results as in the previous section but using the two-dimensional DFT.

Let $u \sim v$ indicate the existence of an edge between vertices u and v , \times denote the Cartesian product of two sets and \otimes denote the Kronecker product. Additionally, let $A(\cdot)$ and $D(\cdot)$ denote the adjacency and degree matrix of a graph respectively.

Definition A.1.1 (Cartesian product of graphs). *The Cartesian product $G \square H$ of two graphs G and H is a graph with vertex set $\mathcal{V}(G \square H) = \mathcal{V}(G) \times \mathcal{V}(H)$ where $(u, v) \sim (u', v')$ if $u = u'$ and $v \sim v'$ or $v = v'$ and $u \sim u'$.*

The adjacency matrix for $G \square H$ is given by $A(G \square H) = A(G) \otimes I + I \otimes A(H)$ [70]. If x and y are eigenvectors of $A(G)$ and $A(H)$ respectively, with corresponding eigenvalues λ and μ , then $(x \otimes y)$ is an eigenvector of $A(G \square H)$ with eigenvalue $\lambda + \mu$. These properties also apply when swapping the adjacency matrix for the unnormalized Laplacian. See Figure A.2 for examples.

Let C_{N_1} and C_{N_2} denote two unweighted cycle graphs with N_1 and N_2 vertices, respectively. The Cartesian product of C_{N_1} and C_{N_2} will become a torus grid graph (see Figure A.1), $T_{N_1, N_2} = C_{N_1} \square C_{N_2}$, with a set of $N_1 \times N_2$ normalized Laplacian eigenvectors and eigenvalues given as

$$\left\{ \chi_{(l_1, l_2)}[(m_1, m_2)] = \frac{1}{\sqrt{N_1 N_2}} e^{-2\pi i \left(\frac{m_1 l_1}{N_1} + \frac{m_2 l_2}{N_2} \right)} \right\}_{l_1=1, \dots, N_1; l_2=1, \dots, N_2} \quad (\text{A.13})$$

²Strictly speaking the approximation only holds for $l \in [-N_v/2, N_v/2 - 1]$ if N_v is even and $l \in [-(N_v - 1)/2, (N_v - 1)/2]$ if N_v is odd. However, λ_l and χ_l are periodic in l so (A.5) will produce the same result for these ranges as for $l \in [0, N_v - 1]$.

$$\left\{ \lambda_{(l_1, l_2)} = 1 - \frac{1}{2} \cos\left(\frac{2\pi l_1}{N_1}\right) - \frac{1}{2} \cos\left(\frac{2\pi l_2}{N_2}\right) \right\}_{l_1=1, \dots, N_1; l_2=1, \dots, N_2}. \quad (\text{A.14})$$

Note that the eigenvectors are still vectors but are for better interpretation indexed using tuples. Constructing the atoms we have

$$\psi_{\mathcal{K}, (m_1, m_2)}[(n_1, n_2)] \stackrel{(2.17)}{=} \sum_{l_1=0}^{N_1-1} \sum_{l_2=0}^{N_2-1} \hat{\mathbf{k}}[(l_1, l_2)] \chi_{(l_1, l_2)}^*[m_1, m_2] \chi_{(l_1, l_2)}[(n_1, n_2)]. \quad (\text{A.15})$$

Similar to as in the 1D case, cf. (A.5), it can be seen that (A.15) represents a DFT; in particular the 2D DFT.

Using the same approximation approach as in (A.10) we obtain

$$\hat{\mathbf{k}}[(l_1, l_2)] = e^{-\tau \lambda_{(l_1, l_2)}} \approx e^{-\tau \frac{1}{4} \left(\left(\frac{2\pi l_1}{N_1} \right)^2 + \left(\frac{2\pi l_2}{N_2} \right)^2 \right)} \quad (\text{A.16})$$

which is a bivariate Gaussian with covariance matrix $\Sigma = \sigma_{\hat{\mathbf{k}}} I$. Again, using the reciprocal relationship between the standard deviations of Gaussian kernels, we have

$$\sigma_{\psi} = \frac{1}{\sigma_{\hat{\mathbf{k}}}} = \sqrt{\frac{\tau}{2}}. \quad (\text{A.17})$$

It is straight-forward to extend this property to higher dimensions, which leads to the relation

$$\sigma_{\psi} = \frac{1}{\sigma_{\hat{\mathbf{k}}}} = \sqrt{\frac{\tau}{D}} \quad (\text{A.18})$$

where D denotes the dimension of the space.

References

- [1] S. Ogawa, T. M. Lee, A. R. Kay, and D. W. Tank, “Brain magnetic resonance imaging with contrast dependent on blood oxygenation,” *Proceedings of the National Academy of Sciences*, vol. 87, no. 24, pp. 9868–9872, 1990.
- [2] N. K. Logothetis and B. A. Wandell, “Interpreting the BOLD signal,” *Annual Review of Physiology*, vol. 66, no. 1, pp. 735–769, 2004. PMID: 14977420.
- [3] J. R. Gawryluk, E. L. Mazerolle, and R. C. D’Arcy, “Does functional MRI detect activation in white matter? A review of emerging evidence, issues, and future directions,” *Frontiers in neuroscience*, vol. 8, 2014.
- [4] T. H. Jochimsen, D. Ivanov, D. V. Ott, W. Heinke, R. Turner, H. E. Möller, and J. R. Reichenbach, “Whole-brain mapping of venous vessel size in humans using the hypercapnia-induced bold effect,” *NeuroImage*, vol. 51, no. 2, pp. 765 – 774, 2010.
- [5] T. Yarkoni, D. M. Barch, J. R. Gray, T. E. Conturo, and T. S. Braver, “BOLD correlates of trial-by-trial reaction time variability in gray and white matter: A multi-study fMRI analysis,” *PLOS ONE*, vol. 4, pp. 1–15, 01 2009.
- [6] A. Peters, M. Brookes, F. Hoogenraad, P. Gowland, S. Francis, P. Morris, and R. Bowtell, “Comparison of T2* measurements in human brain at 1.5, 3 and 7 T,” *Proceedings of the International Society for Magnetic Resonance in Medicine*, no. 14.
- [7] A. L. Alexander, J. E. Lee, M. Lazar, and A. S. Field, “Diffusion tensor imaging of the brain,” *Neurotherapeutics*, vol. 4, no. 3, pp. 316 – 329, 2007. Advances in Neuroimaging/Neuroethics.

-
- [8] M. Catani and D. H. ffytche, “The rises and falls of disconnection syndromes,” *Brain*, vol. 128, no. 10, pp. 2224–2239, 2005.
- [9] E. Bullmore and O. Sporns, “Complex brain networks: Graph theoretical analysis of structural and functional systems,” *Nature reviews. Neuroscience*, vol. 10, no. 3, p. 186, 2009.
- [10] M. Rubinov and O. Sporns, “Complex network measures of brain connectivity: Uses and interpretations,” *NeuroImage*, vol. 52, no. 3, pp. 1059 – 1069, 2010. Computational Models of the Brain.
- [11] J. S. Damoiseaux and M. D. Greicius, “Greater than the sum of its parts: A review of studies combining structural connectivity and resting-state functional connectivity,” *Brain Structure and Function*, vol. 213, pp. 525–533, Oct 2009.
- [12] P. Hagmann, L. Cammoun, X. Gigandet, R. Meuli, C. J. Honey, V. J. Wedeen, and O. Sporns, “Mapping the structural core of human cerebral cortex,” *PLOS Biology*, vol. 6, pp. 1–15, 07 2008.
- [13] D. K. Hammond, P. Vandergheynst, and R. Gribonval, “Wavelets on graphs via spectral graph theory,” *Applied and Computational Harmonic Analysis*, vol. 30, no. 2, pp. 129 – 150, 2011.
- [14] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, “The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains,” *IEEE Signal Processing Magazine*, vol. 30, pp. 83–98, May 2013.
- [15] D. I. Shuman, B. Ricaud, and P. Vandergheynst, “Vertex-frequency analysis on graphs,” *Applied and Computational Harmonic Analysis*, vol. 40, no. 2, pp. 260 – 291, 2016.
- [16] 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*, pp. 1070–1073, 2013.
- [17] 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.*, pp. 1039–1042, 2014.
- [18] H. Behjat, N. Leonardi, L. Sörnmo, and D. Van De Ville, “Anatomically-adapted graph wavelets for improved group-level fMRI activation mapping,” *NeuroImage*, vol. 123, pp. 185–199, 2015.

-
- [19] D. Merhof, M. Richter, F. Enders, P. Hastreiter, O. Ganslandt, M. Buchfelder, C. Nimsky, and G. Greiner, “Fast and accurate connectivity analysis between functional regions based on DT-MRI,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2006: 9th International Conference, Copenhagen, Denmark, October 1-6, 2006. Proceedings, Part II* (R. Larsen, M. Nielsen, and J. Sporring, eds.), pp. 225–233, Berlin, Heidelberg: Springer Berlin Heidelberg, 2006.
- [20] Y. Iturria-Medina, E. Canales-Rodríguez, L. Melie-García, P. Valdés-Hernández, E. Martínez-Montes, Y. Alemán-Gómez, and J. Sánchez-Bornot, “Characterizing brain anatomical connections using diffusion weighted MRI and graph theory,” *NeuroImage*, vol. 36, no. 3, pp. 645 – 660, 2007.
- [21] S. Lifshits, A. Tamir, and Y. Assaf, “Combinatorial fiber-tracking of the human brain,” *NeuroImage*, vol. 48, no. 3, pp. 532 – 540, 2009.
- [22] S. N. Sotiropoulos, L. Bai, and C. R. Tench, “Fuzzy anatomical connectedness of the brain using single and multiple fibre orientations estimated from diffusion MRI,” *Computerized Medical Imaging and Graphics*, vol. 34, no. 6, pp. 504 – 513, 2010. Biomedical Image Technologies and Methods - BIBE 2008.
- [23] S. N. Sotiropoulos, L. Bai, P. S. Morgan, C. S. Constantinescu, and C. R. Tench, “Brain tractography using Q-ball imaging and graph theory: Improved connectivities through fibre crossings via a model-based approach,” *NeuroImage*, vol. 49, no. 3, pp. 2444 – 2456, 2010.
- [24] R. S. Vorburger, C. Reischauer, and P. Boesiger, “Bootgraph: Probabilistic fiber tractography using bootstrap algorithms and graph theory,” *NeuroImage*, vol. 66, pp. 426 – 435, 2013.
- [25] N. Kasenburg, M. Liptrot, N. L. Reisle, S. N. Ørting, M. Nielsen, E. Garde, and A. Feragen, “Training shortest-path tractography: Automatic learning of spatial priors,” *NeuroImage*, vol. 130, pp. 63 – 76, 2016.
- [26] S. A. Walker, D. Miller, and J. Tanabe, “Bilateral spatial filtering: Refining methods for localizing brain activation in the presence of parenchymal abnormalities,” *NeuroImage*, vol. 33, no. 2, pp. 564 – 569, 2006.
- [27] L. Harrison, W. Penny, J. Ashburner, N. Trujillo-Barreto, and K. Friston, “Diffusion-based spatial priors for imaging,” *NeuroImage*, vol. 38, no. 4, pp. 677 – 695, 2007.

- [28] D. A. Raffelt, R. E. Smith, G. R. Ridgway, J.-D. Tournier, D. N. Vaughan, S. Rose, R. Henderson, and A. Connelly, “Connectivity-based fixel enhancement: Whole-brain statistical analysis of diffusion MRI measures in the presence of crossing fibres,” *NeuroImage*, vol. 117, no. Supplement C, pp. 40 – 55, 2015.
- [29] D. Zhu, T. Zhang, X. Jiang, X. Hu, H. Chen, N. Yang, J. Lv, J. Han, L. Guo, and T. Liu, “Fusing DTI and fMRI data: A survey of methods and applications,” *NeuroImage*, vol. 102, no. Part 1, pp. 184 – 191, 2014. Multimodal Data Fusion.
- [30] F. D. Bowman, L. Zhang, G. Derado, and S. Chen, “Determining functional connectivity using fMRI data with diffusion-based anatomical weighting,” *NeuroImage*, vol. 62, no. 3, pp. 1769 – 1779, 2012.
- [31] F. Calamante, R. A. Masterton, J.-D. Tournier, R. E. Smith, L. Willats, D. Raffelt, and A. Connelly, “Track-weighted functional connectivity (TW-FC): A tool for characterizing the structural–functional connections in the brain,” *NeuroImage*, vol. 70, no. Supplement C, pp. 199 – 210, 2013.
- [32] F. Calamante, R. E. Smith, X. Liang, A. Zalesky, and A. Connelly, “Track-weighted dynamic functional connectivity (TW-dFC): A new method to study time-resolved functional connectivity,” *Brain Structure and Function*, Apr 2017.
- [33] F. R. Chung, *Spectral graph theory*. No. 92, American Mathematical Soc., 1997.
- [34] “Chapter 1 - Studying the nervous system,” in *Neuroscience* (D. Purves, G. J. Augustine, D. Fitzpatrick, W. C. Hall, A.-S. LaMantia, J. O. McNamara, and L. E. White, eds.), pp. 1 – 22, Sinauer Associates, 4th ed., 2008.
- [35] R. B. Buxton, “Chapter 1 - Neural activity and energy metabolism,” in *Introduction to Functional Magnetic Resonance Imaging: Principles and Techniques*, Cambridge University Press, 2nd ed., 2009.
- [36] B. Biswal, F. Zerrin Yetkin, V. M. Haughton, and J. S. Hyde, “Functional connectivity in the motor cortex of resting human brain using echo-planar MRI,” *Magnetic Resonance in Medicine*, vol. 34, no. 4, pp. 537–541, 1995.
- [37] M. P. van den Heuvel and H. E. H. Pol, “Exploring the brain network: A review on resting-state fMRI functional connectivity,” *European Neuropsychopharmacology*, vol. 20, no. 8, pp. 519 – 534, 2010.

-
- [38] K. J. Friston, A. P. Holmes, K. J. Worsley, J.-P. Poline, C. D. Frith, and R. S. J. Frackowiak, "Statistical parametric maps in functional imaging: A general linear approach," *Human Brain Mapping*, vol. 2, no. 4, pp. 189–210, 1994.
- [39] G. Flandin and M. J. Novak, "Chapter 6 - fMRI data analysis using SPM," in *fMRI: Basics and clinical applications* (S. Ulmer and O. Jansen, eds.), pp. 51 – 76, Springer, 2nd ed., 2013.
- [40] J.-B. Poline, K. Worsley, A. Evans, and K. Friston, "Combining spatial extent and peak intensity to test for activations in functional imaging," *NeuroImage*, vol. 5, no. 2, pp. 83 – 96, 1997.
- [41] A. Eklund, M. Andersson, C. Josephson, M. Johannesson, and H. Knutsson, "Does parametric fMRI analysis with SPM yield valid results?—An empirical study of 1484 rest datasets," *NeuroImage*, vol. 61, no. 3, pp. 565 – 578, 2012.
- [42] A. Eklund, T. E. Nichols, and H. Knutsson, "Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates," *Proceedings of the National Academy of Sciences*, vol. 113, no. 28, pp. 7900–7905, 2016.
- [43] D. van De Ville, T. Blu, and M. Unser, "Integrated wavelet processing and spatial statistical testing of fMRI data," *NeuroImage*, vol. 23, no. 4, pp. 1472 – 1485, 2004.
- [44] D. V. D. Ville, M. L. Seghier, F. Lazeyras, T. Blu, and M. Unser, "WSPM: Wavelet-based statistical parametric mapping," *NeuroImage*, vol. 37, no. 4, pp. 1205 – 1217, 2007.
- [45] Z. Ding, A. T. Newton, R. Xu, A. W. Anderson, V. L. Morgan, and J. C. Gore, "Spatio-temporal correlation tensors reveal functional structure in human brain," *PLOS ONE*, vol. 8, pp. 1–10, 12 2013.
- [46] X. Wu, Z. Yang, S. K. Bailey, J. Zhou, L. E. Cutting, J. C. Gore, and Z. Ding, "Functional connectivity and activity of white matter in somatosensory pathways under tactile stimulations," *NeuroImage*, vol. 152, pp. 371 – 380, 2017.
- [47] M. Peer, M. Nitzan, A. S. Bick, N. Levin, and S. Arzy, "Evidence for functional networks within the human brain's white matter," *Journal of Neuroscience*, vol. 37, no. 27, pp. 6394–6407, 2017.
- [48] P. J. Basser and E. Özarıslan, "Chapter 1 - Introduction to diffusion MR," in *Diffusion MRI* (H. Johansen-Berg, , and T. E. Behrens, eds.), pp. 3 – 9, San Diego: Academic Press, 2nd ed., 2014.

- [49] P. Basser, J. Mattiello, and D. LeBihan, “MR diffusion tensor spectroscopy and imaging,” *Biophysical Journal*, vol. 66, no. 1, pp. 259 – 267, 1994.
- [50] D. K. Jones, T. R. Knösche, and R. Turner, “White matter integrity, fiber count, and other fallacies: The do’s and don’ts of diffusion MRI,” *NeuroImage*, vol. 73, pp. 239 – 254, 2013.
- [51] D. K. Jones, “Chapter 3 - Gaussian modeling of the diffusion signal,” in *Diffusion MRI* (H. Johansen-Berg and T. E. Behrens, eds.), pp. 37 – 54, San Diego: Academic Press, 2009.
- [52] V. Wedeen, T. Reese, D. Tuch, M. Weigel, J. Dou, R. Weiskoff, and D. Chessler, “Mapping fiber orientation spectra in cerebral white matter with fourier-transform diffusion MRI,” in *Proceedings of the 8th Annual Meeting of ISMRM, Denver*, p. 82, 2000.
- [53] D. S. Tuch, “Q-ball imaging,” *Magnetic Resonance in Medicine*, vol. 52, no. 6, pp. 1358–1372, 2004.
- [54] F. C. Yeh, V. J. Wedeen, and W. Y. I. Tseng, “Generalized q -sampling imaging,” *IEEE Transactions on Medical Imaging*, vol. 29, pp. 1626–1635, Sept 2010.
- [55] K. K. Seunarine and D. C. Alexander, “Chapter 6 - Multiple fibers: Beyond the diffusion tensor,” in *Diffusion MRI* (H. Johansen-Berg, , and T. E. Behrens, eds.), pp. 105 – 123, San Diego: Academic Press, 2nd ed., 2014.
- [56] S. Mori, B. J. Crain, V. P. Chacko, and P. C. M. Van Zijl, “Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging,” *Annals of Neurology*, vol. 45, no. 2, pp. 265–269, 1999.
- [57] T. E. Behrens, S. N. Sotiropoulos, and S. Jbabdi, “Chapter 19 - MR diffusion tractography,” in *Diffusion MRI* (H. Johansen-Berg, , and T. E. Behrens, eds.), pp. 429 – 451, San Diego: Academic Press, 2nd ed., 2014.
- [58] K. Friston, “Chapter 2 - Statistical parametric mapping,” in *Statistical Parametric Mapping* (K. Friston, J. Ashburner, S. Kiebel, T. Nichols, and W. Penny, eds.), pp. 10 – 31, London: Academic Press, 2007.
- [59] “NIST Digital Library of Mathematical Functions.” <http://dlmf.nist.gov/>, Release 1.0.15 of 2017-06-01. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller and B. V. Saunders, eds.

-
- [60] H. Behjat, U. Richter, D. Van De Ville, and L. Sörnmo, “Signal-adapted tight frames on graphs,” *IEEE Trans. Signal Process.*, vol. 64, no. 22, pp. 6017–6029, 2016.
- [61] R. B. Lehoucq and D. C. Sorensen, “Deflation techniques for an implicitly restarted arnoldi iteration,” *SIAM Journal on Matrix Analysis and Applications*, vol. 17, no. 4, pp. 789–821, 1996.
- [62] H. Prautzsch, W. Boehm, and M. Paluszny, “B-spline representation,” in *Bézier and B-Spline Techniques*, pp. 59–75, Berlin, Heidelberg: Springer Berlin Heidelberg, 2002.
- [63] J. R. Gilbert, C. Moler, and R. Schreiber, “Sparse matrices in MATLAB: Design and implementation,” *SIAM Journal on Matrix Analysis and Applications*, vol. 13, no. 1, pp. 333–356, 1992.
- [64] D. V. Essen, K. Ugurbil, E. Auerbach, D. Barch, T. Behrens, R. Bucholz, A. Chang, L. Chen, M. Corbetta, S. Curtiss, S. D. Penna, D. Feinberg, M. Glasser, N. Harel, A. Heath, L. Larson-Prior, D. Marcus, G. Michalareas, S. Moeller, R. Oostenveld, S. Petersen, F. Prior, B. Schlaggar, S. Smith, A. Snyder, J. Xu, and E. Yacoub, “The human connectome project: A data acquisition perspective,” *NeuroImage*, vol. 62, no. 4, pp. 2222 – 2231, 2012. Connectivity.
- [65] National Institutes of Health, “The human connectome project.” Internet: [www . neuroscienceblueprint . nih . gov / connectome](http://www.neuroscienceblueprint.nih.gov/connectome). [09-Sep-2017].
- [66] M. F. Glasser, S. N. Sotiropoulos, J. A. Wilson, T. S. Coalson, B. Fischl, J. L. Andersson, J. Xu, S. Jbabdi, M. Webster, J. R. Polimeni, D. C. V. Essen, and M. Jenkinson, “The minimal preprocessing pipelines for the human connectome project,” *NeuroImage*, vol. 80, pp. 105 – 124, 2013. Mapping the Connectome.
- [67] D. M. Barch, G. C. Burgess, M. P. Harms, S. E. Petersen, B. L. Schlaggar, M. Corbetta, M. F. Glasser, S. Curtiss, S. Dixit, C. Feldt, D. Nolan, E. Bryant, T. Hartley, O. Footer, J. M. Bjork, R. Poldrack, S. Smith, H. Johansen-Berg, A. Z. Snyder, and D. C. V. Essen, “Function in the human connectome: Task-fMRI and individual differences in behavior,” *NeuroImage*, vol. 80, no. Supplement C, pp. 169 – 189, 2013. Mapping the Connectome.
- [68] T. E. Nichols, A. Eklund, and H. Knutsson, “A defense of using resting-state fMRI as null data for estimating false positive rates,” *Cognitive Neuroscience*, pp. 1–2, 2017.

- [69] R. M. Gray, "Toeplitz and circulant matrices: A review," *Foundations and Trends® in Communications and Information Theory*, vol. 2, no. 3, pp. 155–239, 2006.
- [70] L. W. Beineke and R. J. Wilson, *Topics in algebraic graph theory*, vol. 102. Cambridge University Press, 2004.