Sparsity-Promoting Spatiotemporal Regularization for Data Mining in Functional Magnetic Resonance Imaging

Fikret Işık Karahanoğlu

Thèse N° — (July 2013)

Thèse présentée à la faculté des sciences et techniques de l'ingénieur pour l'obtention du grade de docteur ès sciences et acceptée sur proposition du jury

Prof. Michael Unser, *président*
Prof. Dimitri Van De Ville, *directeur de thèse*
Dr. Philippe Ciuciu, *rapporteur*
Prof. Daniele Marinazzo, *rapporteur*
Dr. Jean-Marc Vesin, *rapporteur*

École polytechnique fédérale de Lausanne—2013

Cover design by Dimitri Van De Ville and Fikret Işık Karahanoğlu
Printing and binding by Repro-EPFL
Typeset with LaTeX

Available at **http://miplab.epfl.ch/**

Abstract

Magnetic resonance imaging (MRI) has opened unprecedented avenues to observe the human brain non-invasively. In particular, for about two decades, *functional* MRI (fMRI) has enabled to monitor brain function using the blood-oxygen-level-dependent (BOLD) contrast as a proxy for neuronal activity. The impact of fMRI on neurosciences, medicine, and psychology is ever increasing and has been mainly focussing on (1) understanding brain organization in terms of segregation (i.e., localized processing) and integration (i.e., distributed processing), specifically, related to sensory processing and cognition; (2) exploring temporal characteristics of brain processes.

FMRI provides large spatiotemporal datasets, typically one whole-brain volume with spatial resolution of 1–3 mm in each dimension, every 1–4 seconds during several minutes. The structure of neurophysiological contributions in these data is complex and therefore requires advanced data processing. Conventional fMRI analysis is exploiting timing properties of a stimulation or task paradigm designed by the experimenter; i.e., evidence is sought for the presence of a hypothetical BOLD response. More recently, the community has shown increasing interest in spontaneous brain activity acquired during resting-state fMRI (RS-fMRI). In the absence of any task, data-driven or exploratory methods have found great use. In particular, blind source separation such as independent component analysis (ICA) has been widely applied to RS-fMRI data.

One limitation of current data-driven methods is the lack of incorporating knowledge about the hemodynamic system, which governs any activity-related signal component in the fMRI measurements. In this dissertation, we build upon the latest advances in convex optimization and propose a novel framework that can reveal activity-inducing signals at the fMRI timescale. In particular, our regularization strategy, termed "total activation" (TA), allows deconvolving the fMRI signal to remove hemodynamic blur and to improve spatial contrast of activation patterns by incorporating knowledge about meaningful brain regions. The contribution of our method lies in adapting and tailoring state-of-the-art signal processing techniques with specific domain knowledge from fMRI and neurosciences.

First, we extend "total variation" (TV), which is a well-recognized method in image processing for edge-preserving regularization. TV favors signals that are piecewise constant, and, therefore, whose derivatives are sparse. We generalize this concept for signals of which the derivative of an additional linear differential operator is sparse, and build a variational formulation for the denoising problem. The recovered signal can be also studied after applying the regularizing operators; e.g., applying the differential operator will lead to the piecewise constant driving signal, while applying an additional derivative reveals a sparse "innovation" signal. Fast and efficient schemes from convex optimization are deployed to solve the variational problem at hand.

Second, we apply TA for fMRI data analysis to explore the underlying "deblurred" activity-inducing signals. The temporal regularization is based on generalization of TV where the differential operator is chosen to invert the (linearized) hemodynamic system. Consequently, this will favor block-type activity-inducing signals without restrictions on timing nor duration. The spatial regularization uses a mixed-norm regularization to favor coherent activity-inducing signals in brain regions chosen from an anatomical brain atlas. After demonstrating the feasibility of the proposed method using simulated data, we show results on experimental fMRI data consisting of long resting-state periods and a few unanticipated visual stimuli. The method is able to readily recover plausible activation patterns for the visual stimuli without any prior knowledge about their timing. More interestingly, we also recovered complex spatiotemporal patterns of spontaneous activity that were organized in "resting-state networks", as such providing a new approach to study non-stationary dynamics of RS-fMRI, a research direction that will be of major interest in the coming years. Finally, we include convincing results for localizing epileptogenic brain regions in patients using simultaneous EEG-fMRI recordings.

Keywords: Linear system theory, regularization, sparsity, total variation, mixed-norms, shrinkage algorithms, convex optimization, functional magnetic resonance imaging, blood-oxygenation-level-dependent signal, hemodynamic model, deconvolution, spontaneous brain activity, resting state, paradigm-free brain mapping, neurological disease, epilepsy

Résumé

L'imagerie par résonance magnétique (IRM) a ouvert des perspectives sans précédent pour l'observation non-invasive du cerveau humain. En particulier, depuis environ deux décennies, l'IRM fonctionnelle (IRMf) a permis d'observer le fonctionnement du cerveau en utilisant le signal dépendant du niveau d'oxygène (BOLD) comme un indicateur de l'activité neuronale. L'impact de l'IRMf sur les neurosciences, la médecine et la psychologie ne cesse d'augmenter et s'est concentré principalement sur (1) la compréhension de l'organisation du cerveau en termes de ségrégation (c.-à-d. traitement localisé) et d'intégration (c.-à-d. traitement distribué), en particulier, liée aux processus sensoriels et cognitifs, (2) l'exploration de la dynamique des processus du cerveau.

L'IRMf donne accès à un grand quantité de données spatio-temporelles, avec typiquement une résolution spatiale de 1-3 mm selon chaque dimension, et une résolution temporelle de l'ordre de la seconde. L'interaction des différentes contributions neurophysiologiques dans ces données est complexe et nécessite donc un traitement avancé. L'analyse de l'IRMf classique exploite les propriétés de synchronisation d'une stimulation ou d'un paradigme conçu par l'expérimentateur, c.-à-d., consiste à rechercher la présence d'une réponse BOLD hypothétique, concommitante au stimuli. Plus récemment, la communauté a montré un intérêt grandissant à l'étude de l'activité cérébrale spontanée acquise au cours de l'IRMf à l'état de repos, ou ≪resting-state≫ (RS-IRMf). En l'absence de tâche, les méthodes exploratoires ont trouvé une grande utilité. Par exemple, la mise en évidence de réseaux fonctionnels par analyse en composantes indépendantes (ICA) a été largement appliquée aux données RS-IRMf.

Une limitation des méthodes exploratoires actuelles se situe dans l'absence de prise en compte d'information *a priori* sur le système hémodynamique, qui régit toute composante de signal activité liée aux mesures IRMf. Dans cette thèse, nous proposons un nouveau cadre méthodologique qui permet de révéler la dynamique des signaux d'activité par l'IRMf, en s'appuyant sur les dernières avancées en optimisation convexe. Notre stratégie de régularisation, appelée ≪activation totale ≫ (TA), permet de déconvoluer le

signal IRMf pour retrouver la dynamique de l'activité cérébrale sous-jacente. En outre, le contraste spatial des foyers d'activation peut être amélioré en intégrant des connaissances *a priori* sur les régions pertinentes du cerveau. La contribution de notre méthode réside dans l'adaptation et l'application des dernières techniques de l'état de l'art de traitement des signaux dans le domaine spécifique de l'IRMf et des neurosciences.

Tout d'abord, nous étendons la notion de ≪variation totale≫ (TV), méthode de régularisation bien établie et efficace l'en traitement d'images pour préserver les discontinuités. TV favorise les signaux constants par morceaux, et, par conséquent, dont les dérivées sont parcimonieuses. Nous généralisons ce concept aux signaux dont la dérivée d'un opérateur différentiel linéaire supplémentaire est parcimonieux, et construisons une formulation variationnelle pour le problème de débruitage. Le signal ainsi récupéré peut également être analysé après l'application des opérateurs de régularisation, par exemple, en appliquant l'opérateur différentiel qui conduira au signal original constant par morceaux, tout en appliquant une dérivée supplémentaire pour révéler un signal d'≪innovation≫ clairsemé. Une solution, rapide et efficace, d'optimisation convexe est proposée pour résoudre le problème variationnel clé en main.

Deuxièmement, nous appliquons la méthode TA à l'analyse des données IRMf pour étudier les signaux d'activité cérébrale sous-jacents (c.-à-d. en l'absence du filtrage hémodynamique). La régularisation temporelle est basée sur TA oú l'opérateur différentiel est choisi pour inverser le système hémodynamique linéarisé. Cela a pour conséquence de favoriser l'activité des signaux BOLD de type bloc sans restrictions sur la dynamique temporelle ni la durée de l'activation. La régularisation spatiale utilise une régularisation de type ≪norme mixte≫ pour favoriser l'activation cohérente dans les régions cérébrales choisies à l'aide d'un atlas anatomique du cerveau. Après avoir démontré la faisabilité de la méthode proposée sur les données simulées, nous présentons des résultats sur des données expérimentales d'IRMf constituées de longues périodes de repos altérées par quelques stimuli visuels inattendus. La méthode est capable de recouvrer facilement les modèles d'activation plausibles pour les stimuli visuels sans aucune connaissance préalable sur leur temps d'arrivée ni leur durée. Plus intéressant encore, nous avons aussi pu mettre en évidence des *patterns* spatio-temporels complexes de l'activation spontanée, qui sont organisés en ≪réseaux d'état de repos≫. Ainsi, une nouvelle approche pour étudier la dynamique non stationnaire de RS-fMRI est proposée, ouvrant ainsi une nouvelle direction de recherche qui sera d'un grande intérêt dans les années à venir. Enfin, en utilisant des enregistrements EEG-IRMf simultanés, nous avons également obtenu des résultats probants pour localiser les régions cérébrales épileptogènes chez des patients souffrant d'épilepsie pharmaco-résistantes.

Mots-clés : théorie des systèmes linéaires, régularisation, parcimonie, variation totale, mixte-normes, algorithmes shrinkâge, optimisation convexe, imagerie par résonance magnétique fonctionnelle, signal BOLD, modèle hémodynamique, déconvolution, activité cérébrale spontanée, état de repos , cartographie du cerveau sans paradigme, maladie neurologique, épilepsie

Acknowledgements

It is my duty to record my thankfulness to many people who had supported me in my professional and personal life during my doctoral years.

First and foremost, I am deeply indebted to my advisor Prof. Dimitri Van De Ville, whose genuine enthusiasm and optimism I always appreciate, for mentoring and sharing his experience with me. I am very fortunate to witness his early years in professorship that I learned and benefited immensely from his exceptional vision, motivation and expertise. I am also grateful for his guidance and dedication in reading and correcting all the manuscripts with patience to improve my scientific skills.

Next, I would like to thank to my thesis committee, Dr. Philippe Ciuciu, Prof. Daniele Marinazzo, Dr. Jean-Marc Vesin, and the jury president Prof. Michael Unser, for their insightful comments and suggestions to improve my thesis. Moreover, I am grateful to François, Emrah, César, Frédéric, especially to Jeff, Giulia and Zafer for sparing their valuable time for reading, correcting and even printing this dissertation.

I am thankful to all former and current MIPLab members for providing this very friendly and productive research environment. Special thanks go to Dr. Jonas Richiardi, with whom I proudly share the privilege of being the first members of MIPLab, for his eternal energy, integrity, support and collaboration. I owe big thanks to my colleagues Nora, Jeff and Giulia for their companionship; I enjoyed every moment of our stimulating discussions and had a lot of good memories during the conferences. I should also express my sincere gratitude to Zafer who has sometimes challenged and always tried to encourage me to make the most of myself in every aspect of my life. I thank my students for giving me the opportunity to test myself as supervisor.

I am very happy to be a part of CIBM Geneva and thank all my colleagues for creating such an amicable environment. Special thanks to Dr. François Lazeyras, who had always been very supportive, for his guidance during MRI scanning. I am grateful to Prof. Christoph Michel and Dr. Juliane Britz for their collaboration. I am indebted to Dr. César Caballero-Gaudes and Dr. Frédéric Grouiller who shared their experience and ideas with me. I thank Ruth Fiaux and human resources department in University of Geneva for their continuous help on my administrative problems.

I would like to thank all my friends in Lausanne for the great moments we had during all these years. Finally, I owe my deepest gratitude to my family, my parents Aytaç and Yener for their endless love, patience and sacrifices; my brother Burak, who has a great influence on me, and my sister-in-law Akmer for their trust. This thesis is dedicated to the memory of Birol Atakan and to those who have been struggling for the betterment of my dearest country.

Contents

List of Notations

Acronyms	
AAL	Anatomical automatic labelling
AWGN	Additive white Gaussian noise
BOLD	Blood-oxygen-level-dependent
CBF	Cerebral blood flow
CBV	Cerebral blood volume
DCT	Discrete cosine transform
DMN	Default mode network
(d)Hb	(De)Oxygentaed hemoglobin
HRF	Hemodynamic response function
(ic)EEG	(Intracranial) electroencephalography
(F)GP	(Fast) Gradient projection
(F)ISTA	(Fast) Iterative shrinkage thresholding lgorithm
(F)MRI	(Functional) Magnetic resonance imaging
FWHM	Full width half maximum
GLM	General linear model
IED	Intertictal epileptic discharges
ICA	Independent component analysis
ISI	Inter-stimulus interval
MAP	Maximum a posteriori
ML	Maximum likelihood
NMR	Nuclear magentic resonance
RSN	Resting-state networks
RS-FMRI	Resting-state fMRI
SNR	Signal to noise ratio
SPM	Statistical parametric mapping
TA	Total activation
TE	Echo time
TV	Total variation
TR	Repetition time

Mathematical Formulations

$x(t)$	1D continuous signal, $t \in \mathbb{R}$
$\mathbf{x}[n]$	Vector of length N, $n \in \mathbb{Z}$
$\mathbf{X}$	Matrix of size $V \times N$
$X^{T(*)}$	Transpose (Conjugate) of $\mathbf{X}$
l	Iteration number
$\mathbf{x}^l$	Estimate of $\mathbf{x}$ at iteration l
D	Continuous first order derivative operator
L	Continuous general differential operator
$\Delta_{D(L)}\{\cdot\}$	Discrete differential operator (filter) related to continuous $D(L)$
$\boldsymbol{\Delta}_{D(L)}$	Toeplitz matrix form of $\Delta_{D(L)}\{\cdot\}$
$\det(\mathbf{X})$	Determinant of $\mathbf{X}$
$\lambda_{1,(2)}$	Temporal(Spatial) regularization parameter
$\lVert\mathbf{x}\rVert_p$, $p \geq 1$	ℓ_p-norm of a vector $\lVert\mathbf{x}\rVert_p = (\sum \lvert x[i]\rvert^p)^{1/p}$
$\lVert\mathbf{x}\rVert_\infty$,	ℓ_∞-norm of a vector $\lVert\mathbf{x}\rVert_\infty = \max(\mathbf{x}[1], \ldots, \mathbf{x}[N])$
$\lVert\mathbf{x}\rVert_0$	ℓ_0-norm of a vector $\lVert\mathbf{x}\rVert_0 = \sum \operatorname{sign}(\mathbf{x}[i])$
$\lVert\mathbf{X}\rVert_F$	Frobenious norm of a matrix $\lVert\mathbf{X}\rVert_F = \left(\sum_{i,j} \mathbf{X}[i,j]^2\right)^{1/2}$
$\lVert\mathbf{X}\rVert$	Spectral norm of a matrix $\lVert\mathbf{X}\rVert = \max_{\mathbf{x}\neq\mathbf{0}} \frac{\lVert\mathbf{A}\mathbf{x}\rVert_2}{\lVert\mathbf{x}\rVert_2} = \sqrt{\lambda_{\max}(\mathbf{X}^*\mathbf{X})}$
λ_{max}	Maximum eigenvalue of a matrix
$\widehat{x}(\omega)$	Continuous Fourier transform $\int_t x(t)e^{-j\omega t}dt$
$\widehat{X}(z)$	Z-transform $\sum_n \mathbf{x}[n]z^{-n}$
$\widehat{X}(e^{j\omega})$	Discrete time Fourier transform $\sum_n \mathbf{x}e^{-j\omega n}$
SNR	Signal to noise ratio of a noisy signal $\mathbf{y}$ based on the ground truth signal $\mathbf{x}$, $10\log_{10}\left(\frac{\lVert\mathbf{x}\rVert^2}{\lVert\mathbf{y}-\mathbf{x}\rVert^2}\right)$
sign	Signum function $\operatorname{sign}(x) = \frac{x}{\lvert x\rvert}$
$\nabla\mathcal{C}(\mathbf{x})$	Gradient $\nabla\mathcal{C}(\mathbf{x}) = (d\mathcal{C}(\mathbf{x})/d\mathbf{x}[1], \ldots, d\mathcal{C}(\mathbf{x})/d\mathbf{x}[n])^T$

Chapter 1

Introduction

New directions in science are launched by new tools much more often than by new concepts. The effect of a concept-driven revolution is to explain old things in new ways. The effect of a tool-driven revolution is to discover new things that have to be explained.

Imagined Worlds, Freeman Dyson

The brain has not always been conceived as the most vital organ in the center of perception and cognition. In ancient civilizations, the heart was believed to be the seed of the mind. Yet, this prevailing belief survives ironically, such as in the famous idiom *"knowing by heart"*, meaning *"memorizing"* something perfectly. Hippocrates was the first physician to suggest that the brain is the origin of thoughts instead of the heart and further described *epilepsy* as the *"disturbance of the brain"*. The first anatomical studies date from sixteenth century and the first scientific evidence for cerebral localization of psychological functions appeared in nineteenth century. Indeed, it was not until the surgery of Henry Molaison (H.M.), a long-term epilepsy sufferer, in mid-twentieth century that a major breakthrough was accidentally made to understand the organization of *memory*. H.M. suffered from anteriograde amnesia, inability to form new memory, after the resection of particular brain regions in order to cure epilepsy.

Today, the brain's vital role for communication and regulation inside the body and with the external world is widely acknowledged. However, the mystery of *how* the brain accomplishes these tasks has still not been fully uncovered. Scientists are challenged to understand this intriguing and complex mechanism constituting a large network of billions of neurons.

Significant improvements have been made in science for exploring the brain from different aspects; i.e., structural or functional organization at either microscale (i.e., neurons and synapses), mesoscale (i.e., cortical columns and projections), or macroscale (i.e., brain regions and connections). Functional magnetic resonance imaging (FMRI), which enables observation of functioning brain at a macroscopic level, is a non-invasive brain imaging technique introduced during the last decade of the twentieth century. It relies on the nuclear magnetic resonance principle and produces images (volumes) of the brain over time by measuring the difference of oxygen concentration in the blood following neuronal activation. This relationship is known as the *hemodynamic* effect and measured signal is referred to as blood-oxygen-level-dependent (BOLD) signal. FMRI has been one of the vastly exploited modalities not only for visualizing the functioning brain over time (in seconds), but also with good spatial resolution (in millimeters). It is widely used today for medical studies including epilepsy, schizophrenia, depression, stroke, Alzheimer's disease and Parkinson's disease. The most remarkable clinical impact of fMRI is perhaps *"detecting awareness in vegetative state"* [1]. Despite its many advantages, fMRI data have some limitations that require careful consideration:

1. FMRI does not measure direct neuronal activity. The precise mechanism between neuronal activity and fMRI measurements is not yet fully uncovered, however, repeated studies in the literature suggest an indirect relationship between neuronal activations and fMRI by the BOLD response, which is much slower than the actual neuronal events; e.g., the peak activation is observed after around 5-6 secs [2].

2. The data is complex and large-scale; it contains spatial and temporal correlations. Furthermore, even though advanced acquisition schemes are emerging, there is a trade-off between the temporal and spatial resolution.

3. The measurements are hampered by various sources of noise and nuisance contributions, including reconstruction error, subject movements, inter- and intra-subject variability, physiological effects, scanner effects, magnetic field effects and so on.

4. The interpretation of the fMRI signal is challenging; i.e., how to harness most information out of the fMRI signal is one of the biggest questions in neuroimaging. As more sophisticated questions arise about the brain function, deciphering the vast amount of data becomes more demanding.

In order to address (some of) these limitations, there is a dire need for interdisciplinary methodological frameworks that are able to aggregate physics, signal processing and neuroscience.

A lot of research efforts have been devoted in localizing different cognitive functions in the cerebral cortex and understanding its organization. For example, scientists have identified specific brain regions associated with particular tasks; e.g., solving puzzles, thinking of specific persons/objects, watching movie. Conventional fMRI analyses include model-based methods that pinpoint the cortical regions of which the recorded signal is well-matched with a predefined temporal model. These methods are only applicable in the presence of an explicit paradigm. Besides task-based experiments, another concept in fMRI has emerged that allows for the investigation of intrinsic brain activity, referred to as the *resting-state*. Specifically, fMRI measurements are acquired while subjects are not engaged in a prescribed attention-demanding task but rather let their minds wander freely. Resting-state fMRI (RS-fMRI) requires subsequent data-driven methods that could foster new insights into brain function and organization. Blind source separation methods, such as ICA, are able to provide the functional network organization. However, they are invariant to spatial and temporal permutations of the data; i.e., hemodynamic system is not taken into account. Finally, fMRI deconvolution methods have been proposed to analyze the fMRI data when the task is "*implicit*". Specifically, these methods bring out the "underlying activations" by attempting to invert the hemodynamic system. With the reinforcement of signal processing methods and elimination of computational difficulties, these methods are becoming an important part of the fMRI analysis.

While the BOLD signal relates to the underlying neuronal activity, it suffers from significant temporal blurring due to the "slowness" of the hemodynamic system. Deconvolution methods eliminate this blurring effect without the need of any timing information of the events. The following features make these methods stand out among state-of-the-art:

1. Contrary to conventional task-based methods, they are able to reveal the activation patterns during rest or task-implicit experiments.
2. They incorporate the hemodynamic system into their formulation whereas the blind source separation methods are insensitive to the hemodynamic system.

Revealing the patterns of neuronal activities that underlie the BOLD signal enables the fMRI data analysis not only to detect unknown activity, such as spontaneous activity in RS-fMRI and interictal discharges in epilepsy, but also to further elaborate the temporal dynamics, such as mental chronometry and habituation effects.

The idea of deconvolving the hemodynamic blur was first proposed by Glover [3] in order to study the dynamics in sensorimotor and auditory cortices. The solution strategy consisted of applying Wiener deconvolution filtering, which assumes an underlying Gaussian process and in turn recovers smooth activity-inducing signals.

In this dissertation, we aim at developing a novel fMRI data deconvolution scheme, which we named *Total Activation*. We intend to improve the state-of-the-art by considering both the hemodynamic effect of fMRI and anatomical organization of the brain. To that end, we formulate a spatiotemporal regularization problem with appropriate spatial and temporal regularizers.

All the research activities presented in this thesis have been carried out in collaboration between Medical Image Processing Laboratory, École Polytechnique Fédérale de Lausanne and Center for Biomedical Imaging, University Hospital Geneva.

1.1 Contributions

Generalization of Total Variation Regularization

The first contribution of this work is to propose a regularization term that takes into account the presence of a linear system whose inverse is formulated by a series of linear differential equations. The work is inspired by the notable "total variation" regularization in which there is an implicit generating system that favors piecewise constant signals. Total variation combines a sparsity-pursing norm with the first-order derivative operator so that derivative of the signal is enforced to be sparse to minimize the cost. The sparsity constraint uses ℓ_1-norm, which is the relaxed convex counterpart of the ideal sparsity-promoting ℓ_0, so that convex optimization schemes can still be applied.

We extend the notion of total variation in the sense that the underlying system does no longer prefer piecewise constant signals, which can be seen as a combination of weighted and shifted Green's functions of first-order derivative operator, but rather signals composed of the Green's function of a general differential operator. Therefore, we construct a regularization term that combines the sparsity constraint and the discrete differential filter associated to the inverse of the system. We solve a denoising problem whose regularization term acts directly on the driving input signal; i.e., *analysis* formulation. Further tailoring the operator enables to handle different driving signals; i.e., n-th order polynomials.

Total Activation

The second contribution of this work is to develop a novel deconvolution method for fMRI data analysis. The aim is to recover the underlying activity-inducing signals, which are more closely related to neuronal activity, without any restrictions on the timing or duration of the activations. That allows for detection of spontaneous brain activity and observation of the non-stationary dynamics. The variational formulation is convex and contains a data-fitting term and two regularizers for the different dimensions of the data. First, temporal regularization identifies the "innovation" signal (which is spike-type) as the sparse driver of the hemodynamic system. However, the activity-inducing signals are more flexible block-type signals. From the physiological hemodynamic model in state-space representation, a differential operator can be deriven and plugged into the generalized total variation scheme. Spatial regularization is incorporated using mixed-norm based on anatomical priors of brain regions; i.e., activities in the same brain regions are favored to be coherent. We employ the efficient generalized forward-backward splitting algorithm, which is a fast iterative shrinkage algorithm that alternates between temporal and spatial domain solutions until convergence to the final estimate of the underlying activity-inducing signal is reached.

1.2 Thesis Outline

This thesis is organized as follows: In Chapter 2, we provide a theoretical background on sparsity-promoting regularization methods, especially Total Variation (TV) and its variants. We provide a detailed description of the methods employed in this work, and briefly explain related state-of-the-art. In Chapter 3, we extend the concept of TV regularization for a differential operator L. Specifically, we introduce the generalized L-TV framework that takes into account the presence of a linear system represented by a series of linear differential equations. We devise a variational formulation that can be solved by fast iterative methods. We validate the strength of our method by comparing it with other conventional methods on both synthetic and real data. In Chapter 4, we then explain the basics of fMRI principle and data analysis methods. We start with the working principle of MRI, specifically, the contrast mechanism in fMRI. Then, we discuss state-of-the-art fMRI analysis methods in three subcategories; confirmatory, exploratory and deconvolution methods. In Chapter 5, we introduce *Total Activation*, spatiotemporal regularization tailored for exploring spontaneous brain activity in fMRI. We further extend the generalized L-TV by including an anatomical prior, so that both spatial and temporal characteristics are taken into account. Finally, in Chapter 6, we present the results for two real fMRI datasets. The first experiment is based on a few short and unpredicted visual stimulation during resting-state for three healthy subjects. Total activation is able to reconstruct the underlying experimental paradigm requiring neither timing nor duration of the stimuli. The second experiment aims at detecting and mapping of sources of epileptic activity on pharmaco-resistant epilepsy patients.

Chapter 2

Reconstruction with Sparsity-Promoting Regularization

Inverse problems are common in many engineering applications, in particular, in signal and image processing. The aim is to retrieve the original signal from noisy and degraded measurements. However, this problem is typically *ill-posed*, by that means, unique and stable solution might not exist due to the nature of the observation model. Therefore, additional assumptions are required to achieve a feasible solution. In this chapter, we describe how a linear inverse problem can be cast as a convex optimization problem. Additional constraints are incorporated into the formulation as *regularizers* to favor some desirable properties of the recovered signal, such as smoothness and sparsity. We specifically elaborate sparsity-promoting regularization and discuss state-of-the-art optimization tools that lead to efficient iterative schemes.

2.1 Observation Model

We consider a typical problem setting widely encountered in signal and image processing applications: an unknown signal goes through a physical process whose output is measured through some acquisition scheme. The observed data is possibly degraded and corrupted by the system and/or several sources of noise during the process. The measured signal is composed of discrete samples and can be formulated through a linear mapping

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \boldsymbol{\epsilon}, \tag{2.1}$$

where $\mathbf{A} \in \mathbb{R}^{M\times N}$ is the observation model, $\mathbf{y} \in \mathbb{R}^{\mathbb{M}}$ is the measured signal, $\mathbf{x} \in \mathbb{R}^{\mathbb{N}}$ is the input signal and $\boldsymbol{\epsilon}$ is the additive noise. The matrix $\mathbf{A}$ can possibly represent some dictionary (wavelets, DCT, ...), blurring kernel, subsampling kernel, a binary map or a combination of these depending on the application at hand. The system is referred to as underdetermined when $M < N$ and overdetermined when $M > N$. In its simplest version (i.e., $\mathbf{A} = \mathbf{I}$), the signal model reduces to noise-only case which will be covered mostly throughout this thesis.

Often, the primary goal is to solve the inverse problem and access the source signal; i.e., recovering the original signal $\mathbf{x}$ from noisy measurements $\mathbf{y}$.

2.2 Variational Formulation

The problem of estimating the signal of interest $\mathbf{x}$ can be investigated via different schemes. The fundamental solution strategy aims at minimizing the squared-error, also known as least squares, leading to the following variational formulation:

$$\hat{\mathbf{x}} \in \arg\min_{\mathbf{x}\in\mathbb{R}^N} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2. \tag{2.2}$$

The solution of the above formulation is not necessarily unique. The set of all $\mathbf{x}$, which satisfy the above criteria, highly depends on the null space of $\mathbf{A}$, as the solution set of $\mathbf{A}\mathbf{x}_{null} = 0$. Note that, when $\mathbf{A}$ represents an orthogonal basis, the solution is unique; i.e., $\mathbf{x}_{null} = \mathbf{0}$. Otherwise, any solution that involves the null space, $\mathbf{x} + c\mathbf{x}_{null}$, where c is a constant, automatically satisfies (2.2). Therefore, the problem is considered as ill-posed. The minimum-norm least squares solution, which maps (2.2) into a unique solution of

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_2 \quad \text{subject to} \quad \mathbf{A}\mathbf{x} = \mathbf{y}, \tag{2.3}$$

becomes $\hat{\mathbf{x}} = \mathbf{A}^{\dagger}\mathbf{y}$, where $\mathbf{A}^{\dagger}$ is the *pseudo-inverse* of $\mathbf{A}$. Specifically, when $\mathbf{A}$ is a full rank overcomplete ($M < N$) or undercomplete matrix ($M > N$), $\mathbf{A}^{\dagger} = \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}$ is the right inverse or $\mathbf{A}^{\dagger} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$ left inverse of $\mathbf{A}$, respectively. The *condition-number* of $\mathbf{A}$ plays an important role on the robustness to small perturbations. Indeed, for the denoising case ($\mathbf{A} =\mathbf{I}$) the solution becomes the measured noisy signal itself, $\hat{\mathbf{x}} = \mathbf{y}$.

Generally, additional constraints on the reconstructed signal could be incorporated into (2.2) to obtain a unique solution as

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathrm{R}^N} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \mathcal{R}(\mathbf{x}), \tag{2.4}$$

where the first term is the *data-(fitting) term* and $\mathcal{R}(\mathbf{x})$ is the *regularization term*. This regularization problem is also known as penalized least squares problem.

If we prefer smooth representations, the variational formulation could use Tikhonov regularization as

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \|\boldsymbol{F}\mathbf{x}\|_2^2, \tag{2.5}$$

where λ is the regularization term and $\mathbf{F}$ is a suitable linear mapping. The unique solution is then achieved as $\hat{\mathbf{x}} = (\mathbf{A}^T\mathbf{A} + \lambda\mathbf{F}^T\mathbf{F})^{-1}\mathbf{A}^T\mathbf{y}$.

2.2.1 Bayesian Interpretation

The same aforementioned concepts can be elaborated from a statistical point of view. The simplest estimation exploits the maximum likelihood (ML) principle, which maximizes the probability of the measured signal given the estimated signal; i.e., $\hat{\mathbf{x}} = \arg\max_{\mathbf{x}} \mathbf{p}(\mathbf{y}|\mathbf{x})$. Clearly, the ML formulation depends on the adapted noise model for the specific problem at hand, and requires different solution schemes for different noise models. The widely accepted noise model is the additive white Gaussian noise (AWGN); i.e., $\mathbf{p}(\boldsymbol{\epsilon}) \sim \mathcal{N}(0, \mathbf{I})$, maximizing the ML follows

$$\begin{aligned}\hat{\mathbf{x}} &= \arg\max_{\mathbf{x}} \mathbf{p}(\mathbf{A}\mathbf{x} + \boldsymbol{\epsilon}|\mathbf{x}) = \arg\max_{\mathbf{x}} \mathbf{p}(\boldsymbol{\epsilon}),\\ &= \arg\max_{\mathbf{x}} \left(\frac{1}{2\pi\sigma^2}\right)^{1/2} e^{-\frac{\|\mathbf{y}-\mathbf{A}\mathbf{x}\|^2}{2\sigma^2}}.\end{aligned}$$

The log-likelihood leads to the same variational formulation in (2.2) with a quadratic cost function through

$$\begin{aligned}\hat{\mathbf{x}} &= \arg\max_{\mathbf{x}} \log\left(\left(\frac{1}{2\pi\sigma^2}\right)^{1/2} e^{-\frac{\|\mathbf{y}-\mathbf{A}\mathbf{x}\|^2}{2\sigma^2}}\right),\\ &= \arg\max_{\mathbf{x}} \left(-\frac{\|\mathbf{y}-\mathbf{A}\mathbf{x}\|_2^2}{2\sigma^2}\right) + K,\\ &= \arg\min_{\mathbf{x}} \|\mathbf{y}-\mathbf{A}\mathbf{x}\|_2^2,\end{aligned} \tag{2.6}$$

where K is a constant independent of $\mathbf{x}$. When the correlated noise is adapted (i.e., $\mathbf{p}(\boldsymbol{\epsilon}) \sim \mathcal{N}(0, \sigma^2\Sigma)$ with a covariance matrix Σ), ML leads to weighted least squares as

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{y}-\mathbf{A}\mathbf{x}\|_{\Sigma}^2 = \arg\min_{\mathbf{x}} (\mathbf{y}-\mathbf{A}\mathbf{x})^T\Sigma^{-1}(\mathbf{y}-\mathbf{A}\mathbf{x}).$$

If the prior probability density of the unknown signal $\mathbf{x}$ can be anticipated the Bayesian inference solves for the *maximum a posteriori* probability (MAP); i.e., $\hat{\mathbf{x}} = \arg\max_{\mathbf{x}} \mathbf{p}(\mathbf{x}|\mathbf{y}) = \arg\max_{\mathbf{x}} \mathbf{p}(\mathbf{y}|\mathbf{x})\mathbf{p}(\mathbf{x})$. Then, the regularization term can be driven from the prior as $\mathcal{R}(\mathbf{x}) \approxeq -\log(\mathbf{p}(\mathbf{x}))$. When the prior follows a Gaussian distribution with $\mathbf{p}(\mathbf{x}) \sim \mathcal{N}(0, \sigma_x^2\boldsymbol{\Sigma}_{\boldsymbol{x}})$ and the noise is AWGN, the variational formulation becomes penalized least squares as in (2.5) regularization term $\lambda = \frac{\sigma^2}{\sigma_x^2}$ and whitening operator $\mathbf{F} = \mathbf{\Sigma}^{-1/2}$. The same solution is also known as the minimum mean square error (MMSE) estimator. A more general solution of the same problem for a general correlation structure in signal and image processing applications is known as the Wiener filter solution.

The link between the Bayesian inference and the variational formulation in (2.4) is sometimes qualitative as the prior needs to correspond to a proper probability density distribution. In this work, we deal with variational viewpoint and do not elaborate on statistical interpretation of the regularizers.

2.3 Analysis and Synthesis Formulations

The incorporation of the regularizer leads to a variational formulation such that the optimal solution depends on the compromise between the data- and the regularization terms. Regularization can impose constraints either on the coefficients $\mathbf{c}$ of the signal in a proper basis or frame; i.e., $\mathbf{x} = \mathbf{Tc}$, known as *synthesis prior*

$$\tilde{\mathbf{x}} = \mathbf{T}\tilde{\mathbf{c}} \text{ with } \tilde{\mathbf{c}} = \arg\min_{\mathbf{c}} \|\mathbf{y} - \mathbf{ATc}\|_2^2 + \lambda\|\mathbf{c}\|_p^p, \tag{2.7}$$

or on signal as a result of a specific transformation, $\mathbf{Fx}$, known as *analysis prior*

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{Ax}\|_2^2 + \lambda\|\mathbf{Fx}\|_p^p, \tag{2.8}$$

where $\mathbf{T}$ is the synthesis operator of a decomposition of $\mathbf{x}$ on a dictionary (e.g., wavelets), $\mathbf{F}$ is the operator that extracts meaningful information from $\mathbf{x}$.
The analysis and synthesis formulations do not always lead to the same solution. Two solutions are *analytically* equivalent only for some specific cases:

(i) when $\mathbf{F}$ is a square and invertible matrix, $\mathbf{T} = \mathbf{F}^{-1}$,
(ii) when $\mathbf{A} = \mathbf{I}$ and $\mathbf{T}$ represents an undercomplete dictionary, $\mathbf{T} = \mathbf{F}^T(\mathbf{FF}^T)^{-1}$,
(iii) when $\mathbf{T}$ is an overcomplete dictionary and $p = 2$, $\mathbf{T} = (\mathbf{F}^T\mathbf{F})^{-1}\mathbf{F}^T$.

However, it is difficult to anticipate which prior (analysis or synthesis) provides a better estimation of the underlying solution [4].

2.4 Sparsity-Pursing Norms

Many signal and image processing studies have concentrated on finding the best possible priors by looking for analysis/synthesis operators that lead to desirable properties of the solution; i.e., by questioning the signal's representation in these domains and considering the computational cost of the optimization. One of the popular choices is Tikhonov regularization which corresponds to energy minimization of the signal via ℓ_2-norm and hence yields smooth solutions. The quadratic norm is particularly interesting since it leads to a closed form and analytical solution. However, *sparsity-pursuing* norms can be a better substitute since it allows for representing the signal with least amount of elements.

ℓ_0-(quasi)norm Constrained Regularization

The sparse solutions, which provide representations with only a few large coefficients, are promoted by the ℓ_0-(quasi)norm. The problem is now cast as

$$\hat{\mathbf{x}} \in \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 + \lambda\|\mathbf{x}\|_0. \tag{2.9}$$

The regularization is not convex and computationally demanding; the global solution is only obtained with combinatorial complexity, therefore, it quickly becomes intractable even for small sized problems. Search based methods, matching pursuit (MP) and its variants, are heuristic solution proposed to handle with these problems by tracking every coefficient and updating the estimate iteratively [5, 6].

Convex Relaxation

We consider the convex relaxation of ℓ_0-(quasi)norm by ℓ_p-norms $p \geq 1$, especially the ℓ_1-norm which provides closest sparse solution through convex (however, not strictly) formulations. ℓ_1-based regularization still favor few large elements where the minimum length is achieved. The relaxation of ℓ_0 to ℓ_1-norm allows the use of efficient algorithms from convex optimization to find a minimizer to the cost function $\mathcal{C}$ in the form of

$$\tilde{\mathbf{x}} \in \arg\min_{\mathbf{x}} \mathcal{C}(\mathbf{x}) = \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_1. \tag{2.10}$$

This specific type of problem is referred to as basis pursuit (BP) or least absolute shrinkage and selection operator (LASSO) problem in the literature [7, 8]. Different versions of LASSO exist in literature; e.g., group LASSO imposes sparsity constraint on a group of variables [9], overlapping group LASSO and graph LASSO exploit specific group structure [10], fused LASSO exploits both sparse and smooth regularization terms [11], etc.

ℓ_p-(quasi)norm Constrained Regularization

Another family of regularization problems incorporate the 'smoothed' ℓ_0-(quasi)norm, generally referred to as ℓ_p-(quasi)norm with $0 < p < 1$, postulating that the solution approaches to ℓ_0 solution as $p \to 0$, and hence provides better estimation than ℓ_1 regularized solutions. Even though these problems did not attract a lot of attention in the early stage of sparsity-inducing regularization–probably due to its non-convex nature–, with recent improvements in the optimization algorithms and regularization, ℓ_p-(quasi)norm constrained regularization is emerging, especially in compressive sensing community. The variational problem is casted as

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_p, \quad 0 < p < 1. \tag{2.11}$$

The problem might be considered as a weighted ℓ_1-norm regularization, weighted least squares, iterative thresholding, smoothed ℓ_0 or specific type of gradient descent, etc [12, 13]. The application on real data problems; an example of MRI reconstruction problem from few Fourier samples is particularly interesting [14]. A recent review of ℓ_p-(quasi)norm constrained regularization methods can be found in [15].

Total Variation Regularization

One of the notable analysis priors in signal and image processing is the total variation (TV) norm, introduced by Rudin *et al.* as a regularization that measures the first degree information in the signal [16]. The regularization is casted as a denoising problem that favors piecewise constant signal representation preserving sharp edges.

TV regularization has found wide real-world data applications in signal and image processing. So far, we have discussed only the discrete setting, however understanding the role of TV in continuous domain is more intuitive than its discrete counterpart. Hence, we now start with the formal definition of total variation in the continuous setting.

Definition 1. *[Total Variation] The TV of a continuous-domain function $x(t)$ on interval $[a, b]$, is defined as the supremum of absolute differences for any partition $P = \{\ldots, t_n, \ldots\}_n$ on its support:*

$$\mathrm{TV}\{x\} = \sup_P \sum_n |x(t_n) - x(t_{n-1})| \,. \tag{2.12}$$

If the first derivative of $x(t)$ is well defined, then TV can be shown to be equivalent to

$$\mathrm{TV}\{x\} = \int_a^b |D\{x\}(t)| \, dt, \tag{2.13}$$

where D is the regular (continuous-domain) derivative [17].

In the discrete setting, TV can be represented as the ℓ_1-norm of the finite-difference operator:

$$\mathrm{TV}\{\mathbf{x}\} = \sum_{n\in\mathbb{Z}} |\Delta_D \{\mathbf{x}\} [n]| = \|\Delta_D\{\mathbf{x}\}\|_1,$$

where Δ_D is the finite difference operator and can be seen as the discrete counterpart of derivative operator. Specifically, we adopt the shortest length filter representation; i.e., $\Delta_D\{\mathbf{x}\}[n] = x(t_n) - x(t_{n-1})$ [1].

Combining the quadratic data-term with TV-norm leads to TV regularization:

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}\in \mathrm{R}^N} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda\|\Delta_D\mathbf{x}\|_1. \tag{2.14}$$

The regularization parameter λ lets the result vary between the ML solution ($\lambda = 0$) and the null space of TV ($\lambda \to \infty$).

Here, we elaborate the TV definition to give an insight into the concept that is introduced in the following chapter. The ideal solution promoted by TV regularization and the Green's function of the derivative operator have a connection. TV regularization in (2.14) promotes sparsity on the derivative of the signal; i.e., piecewise constant signals whose derivatives only have a few spikes are the optimal candidates. Moreover, the Green's function $u(t)$ of the derivative operator D is the Heaviside step function satisfying $D\{u\}(t) = \delta(t)$. Since piecewise constant signals are constructed by shifted Heaviside step functions, TV regularization, indeed, promotes signals that look like the Green's function of its derivative operator.

It is important to note that basis-pursuit denoising [7] can offer a "synthesis" counterpart of TV through a dictionary with (shifted) Heaviside step functions that is the inverse of the TV analysis dictionary. Therefore, the solutions of analysis and synthesis regularizations are equivalent [18]. However, the synthesis dictionary becomes unstable as the size of the problem increases. In some cases, a specific dictionary can be built for the synthesis operator [19, 20], including generalizations of wavelet design [21, 22]. The selection of the solution scheme depends on the application at hand [23].

Structured Sparsity Constrained Regularization

Group or structured sparsity is first introduced as a synthesis formulation known as group LASSO [9]. Structured sparsity assumes that the signal is composed of M partitions, and promotes the sparsity of the partitions of variables; i.e., instead of penalizing each coefficient separately, it penalizes

[1] The operator Δ_D is defined in terms of its transfer function and represented in time domain as $\Delta_D\{x\}[n]$. However, the matrix equivalent formulation is straight-forward (Toeplitz matrix) and it is represented as $\Delta_D\mathbf{x}$.

all coefficients in a partition. In the general form, the group constraint is achieved through a $\ell_{2,1}$ mixed-norm,

Definition 2 (Mixed Norm)**.** *For any non-overlapping partition* $R_k, k = 1, \ldots, M$ *of signal* $\mathbf{x} \in \mathrm{R}^N$*, the* $\ell_{(p,q)}$ *mixed-norm is defined as*

$$\|\mathbf{x}\|_{(p,q)} = \left(\sum_{k=1}^{M} (\|\mathbf{x}[R_k]\|_p)^q \right)^{1/q} = \left(\sum_{k=1}^{M} \left(\sum_{i \in R_k} \mathbf{x}[i]^p \right)^{q/p} \right)^{1/q} . \tag{2.15}$$

The structured sparsity is induced by the $\ell_{(2,1)}$-norm, which ensures the recovery of a few partitions in which smoothness is enforced. The analysis counterpart asserts the $\ell_{(2,1)}$ mixed norm over a linear transform, $\mathbf{F}$, and the regularization becomes [24]

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \|\mathbf{F}\mathbf{x}\|_{(2,1)}. \tag{2.16}$$

2.5 Optimization Algorithms

We have introduced the variational formulations exploiting sparsity-pursing norms that are relevant to many problems in signal processing. Now, we briefly review the state-of-the-art convex optimization techniques.

2.5.1 Quadratic Priors

The minimizer of the cost function with an ℓ_2 regularizer (i.e., Tikhonov regularization)

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \mathcal{C}(\mathbf{x}) = \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \|\mathbf{F}\mathbf{x}\|_2^2, \tag{2.17}$$

can be computed analytically by deriving $\nabla \mathcal{C}(\tilde{\mathbf{x}}) = \mathbf{0}$. Since both the data-term and the regularization term are quadratic (differentiable) the minimization directly yields $\tilde{\mathbf{x}} = \left(\mathbf{A}^T\mathbf{A} + \lambda \mathbf{F}^T\mathbf{F}\right)^{-1} \mathbf{A}^T\mathbf{y}$.

2.5.2 Non-Quadratic Priors

The sparsity-pursuing regularization introduces non-quadratic functions, hence differentiability is not always guaranteed. A vast amount of optimization methods have emerged to address the problem in (2.10), such as orthogonal matching pursuit OMP, least angle regression (LARS), iterative reweighed least squares (IRLS), and forward-backward splitting [6, 25–27]. For a complete overview we refer to Chapter 3-5 of [18]. Now, we explain the alternative popular *iterative-shrinkage* schemes that allow for dealing with non-quadratic priors.

2.5.3 Proximal Maps for Sparse Priors

The denoising problem with a non-quadratic regularization term $\mathcal{R}(\cdot)$ can be associated with a general framework referred to as *proximal map.*

Definition 3 (Proximal Map [28])**.** *For a lower semicontinuous convex function,* $\mathcal{R}(\mathbf{x})$*; i.e.,* $\lim_{\mathbf{x}\to\mathbf{x_0}} \inf \mathcal{R}(\mathbf{x}) \geq \mathcal{R}(\mathbf{x_0})$ *the proximal map,* $\mathrm{prox}_{\mathcal{R}(\mathbf{x})}(\cdot)$*, solves for the following minimization problem*

$$\mathrm{prox}_{\mathcal{R}(\mathbf{x})}(\mathbf{y}) = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y}-\mathbf{x}\|_2^2 + \mathcal{R}(\mathbf{x}) \quad \textit{with} \tag{2.18}$$

$$\mathbf{0} \in \mathbf{x} - \mathbf{y} + \partial\mathcal{R}(\mathbf{x}) \Rightarrow \mathrm{prox}_{\mathcal{R}(\mathbf{x})}(\mathbf{y}) \in (\mathbf{I} + \partial\mathcal{R}(\mathbf{x}))^{-1}\mathbf{y} \tag{2.19}$$

When $f(\mathbf{x}) = i_C(\mathbf{x})$*, indicator function of a convex set,* C*, the proximal map becomes the projection,* $P_C(\mathbf{y})$*, onto the set. A list of interesting properties and proximal maps of some notable functions are found in [29].*

Soft Thresholding

Consider the following denoising problem

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y}-\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_1. \tag{2.20}$$

The solution can be computed by explicitly taking the derivative of the scalar cost function separately over each coefficient $\mathbf{x}[i]$; i.e., $\frac{1}{2}(\mathbf{y}[i]-\mathbf{x}[i])^2+\lambda|\mathbf{x}[i]|$. Then the proximal map; i.e., the minimizer of the above function is obtained via soft thresholding operator [30]

$$\tilde{\mathbf{x}}[i] = \mathrm{prox}_{\lambda|\mathbf{x}[i]|}(\mathbf{y}[i]) = (|\mathbf{y}[i]| - \min(\lambda, |\mathbf{y}[i]|))\,\mathrm{sign}(\mathbf{y}[i]). \tag{2.21}$$

TV Denoising

We now show how to compute the proximal map minimizing the TV denoising problem

$$\tilde{\mathbf{x}} = \mathrm{prox}_{TV}(\mathbf{y}) = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y}-\mathbf{x}\|_2^2 + \lambda\|\Delta_D\{\mathbf{x}\}\|_1. \tag{2.22}$$

Unfortunately, the non-smoothness and non-separability of the TV term do not let coefficient-wise derivation of the functional. Instead of a closed form, an iterative scheme will lead to the minimizer. The *primal* problem in (2.22) has no direct analytical solution. As proposed by Chambolle, the dual definition allows for constructing a gradient based optimization to TV denoising problem [31].

Definition 4. *(Dual norm of ℓ_1) $\|\mathbf{F}\mathbf{x}\|_1$ can be written in its dual form as*

$$\|\mathbf{F}\mathbf{x}\|_1 = \max_{|\mathbf{p}[i]|\leq 1} \langle \mathbf{F}\mathbf{x}, \mathbf{p} \rangle. \tag{2.23}$$

The dual of ℓ_1-norm is the ℓ_∞-norm (i.e., $|\mathbf{p}[i]| \leq 1$). More generally, for any ℓ_p-norm its dual ℓ_q satisfies $\frac{1}{p} + \frac{1}{q} = 1$.

The dual formulation of TV denoising then becomes

$$\begin{aligned}\tilde{\mathbf{x}} &= \arg\min_{\mathbf{x}} \max_{|\mathbf{p}[i]|\leq 1} \frac{1}{2}\|\mathbf{y} - \mathbf{x}\|_2^2 + \lambda\langle \Delta_D\{\mathbf{x}\}, \mathbf{p}\rangle, \qquad (2.24)\\ &= \arg\min_{\mathbf{x}} \max_{\mathbf{p}} \mathcal{C}(\mathbf{x}, \mathbf{p}).\end{aligned}$$

We observe that the above function is concave (but not strictly so) in $\mathbf{p}$ and convex in $\mathbf{x}$, which satisfies the saddle point criteria [32] as

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \left\{ \max_{\mathbf{p}} \mathcal{C}(\mathbf{x}, \mathbf{p}) \right\}, \quad \tilde{\mathbf{p}} = \arg\max_{\mathbf{p}} \left\{ \min_{\mathbf{x}} \mathcal{C}(\mathbf{x}, \mathbf{p}) \right\} \quad \text{and,} \tag{2.25}$$

$$\mathcal{C}(\tilde{\mathbf{x}}, \mathbf{p}) \leq \mathcal{C}(\mathbf{x}, \mathbf{p}) \leq \mathcal{C}(\mathbf{x}, \tilde{\mathbf{p}}). \tag{2.26}$$

The existance of a saddle point allows for exchanging the minimum and maximum in the functional and lets us work with the dual formulation as

$$\min_{\mathbf{x}} \max_{|\mathbf{p}[i]|\leq 1} \frac{1}{2}\|\mathbf{y} - \mathbf{x}\|_2^2 + \lambda\langle \Delta_D\{\mathbf{x}\}, \mathbf{p}\rangle = \max_{|\mathbf{p}[i]|\leq 1} \min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y} - \mathbf{x}\|_2^2 + \lambda\langle \Delta_D\{\mathbf{x}\}, \mathbf{p}\rangle \tag{2.27}$$

The optimal solution for the inner minimization problem is obtained via computing the derivative[2]

$$\tilde{\mathbf{x}} = \mathbf{y} - \lambda\Delta_D^T\{\mathbf{p}\}. \tag{2.28}$$

The inner solution is fed into the original formulation and solved for the maximization

$$\begin{aligned}\tilde{\mathbf{p}} &= \arg\max_{|\mathbf{p}[i]|\leq 1} \frac{1}{2}\|\mathbf{y} - (\mathbf{y} - \lambda\Delta_D^T\{\mathbf{p}\})\|_2^2 + \lambda\langle \Delta_D\{\mathbf{y} - \lambda\Delta_D^T\{\mathbf{p}\}\}, \mathbf{p}\rangle\\ &= \arg\min_{|\mathbf{p}[i]|\leq 1} \frac{\lambda^2}{2}\|\Delta_D^T\{\mathbf{p}\})\|_2^2 - \lambda\langle \Delta_D\{\mathbf{y}\}, \mathbf{p}\rangle\\ &= \arg\min_{|\mathbf{p}[i]|\leq 1} -\mathcal{C}(\tilde{\mathbf{x}}, \mathbf{p}). \qquad (2.29)\end{aligned}$$

The gradient descent allows for computing the minimizer in an iterative way

$$\begin{aligned}\tilde{\mathbf{p}}^{k+1} &= P_B\left(\tilde{\mathbf{p}}^k - \frac{1}{L}\nabla\mathcal{C}(\tilde{\mathbf{x}}, \mathbf{p})\right)\\ &= P_B\left(\tilde{\mathbf{p}}^k - \frac{1}{L}(\lambda^2\Delta_D\Delta_D^T\{p\} - \lambda\Delta_D\{\mathbf{y}\})\right), \qquad (2.30)\end{aligned}$$

[2]The transpose of the derivative operator (in matrix form) indeed converts to filtering with the time reversed transfer function; i.e., $\Delta_D^T[n] = \Delta_D^T[-n]$.

where $P_B(\cdot) = \text{sign}(\cdot)\,\min(|\cdot|, 1)$ is the element-wise projection to satisfy the constraint and $1/L$ is the step-size with $L = \lambda^2\|\Delta_D \Delta_D^T\| = 4\lambda^2$ selected as the smallest Lipschitz constant of $\mathcal{C}(\tilde{\mathbf{x}}, \mathbf{p})$; i.e., $\|\nabla\mathcal{C}(\tilde{\mathbf{x}}, \mathbf{p}_1) - \nabla\mathcal{C}(\tilde{\mathbf{x}}, \mathbf{p}_2)\|_2 \leq L\|\mathbf{p}_1 - \mathbf{p}_2\|_2$. The algorithm is summarized in Algorithm 1.

Algorithm 1 TV Denoising $\text{prox}_{TV}(\mathbf{y}) = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y} - \mathbf{x}\|_2^2 + \lambda\,||\Delta_D\,\{\mathbf{x}\}||_1$

input: Noisy signal $\mathbf{y}$

1: $l \leftarrow 1$
2: Initialize: $\mathbf{p}^0 = \mathbf{0}$
3: **repeat**
4: Update $\mathbf{p}^l = P_B\left(\Delta_D\,\{\mathbf{y}\}/(4\lambda) + (I - \Delta_D\Delta_D^T/4)\,\left\{\mathbf{p}^{l-1}\right\}\right)$
where Δ_D^T is the adjoint and $P_B = \text{sign}(\cdot)\,\min(|\cdot|, 1)$ denotes the elementwise clipping function,
5: $l \leftarrow l + 1$
6: **until** convergence or the number of maximum iterations are reached.
7: Set $\tilde{\mathbf{x}} = \mathbf{y} - \lambda\,\Delta_D^T\,\left\{\mathbf{p}^{l-1}\right\}$.

A variation to the Chambolle's TV denoising algorithm in [31] is proposed by Beck and Teboulle in [33]. The algorithm, so called (fast) Gradient Projection (FGP) algorithm, includes a constraint set for the minimizer $\mathbf{x}$. The fast implementation is described in Algorithm 2 (steps 5-6).

Algorithm 2 Fast Gradient Projection Algorithm for TV Denoising
$\text{prox}_{TV}(\mathbf{y}) = \arg\min_{\mathbf{x}\in X} \frac{1}{2}\|\mathbf{y} - \mathbf{x}\|_2^2 + \lambda\,||\Delta_D\,\{\mathbf{x}\}||_1$

input: Noisy signal $\mathbf{y}$

1: $l \leftarrow 1$
2: Initialize: $k^1 = 1, \mathbf{p}^0 = \mathbf{0}, \mathbf{v}^0 = \mathbf{0}$
3: **repeat**
4: Update $\mathbf{p}^l = P_B\left(\mathbf{v}^l + \Delta_D\left\{P_X\left(\mathbf{y} - \lambda\Delta_D^T\right)\left\{\mathbf{v}^l\right\}\right\}/(4\lambda)\right)$
where $P_B = \text{sign}(\cdot)\,\min(|\cdot|, 1)$ denotes the elementwise clipping function and P_X is the orthogonal projection onto set X,
5: Update $k^{l+1} = \frac{1+\sqrt{1+4(k^l)^2}}{2}$
6: Update $\mathbf{v}^{l+1} = \mathbf{p}^l + \frac{k^l-1}{k^{l+1}}(\mathbf{p}^l - \mathbf{p}^{l-1})$
7: $l \leftarrow l + 1$
8: **until** convergence or the number of maximum iterations are reached.
9: Set $\tilde{\mathbf{x}} = P_X\left(\mathbf{y} - \lambda\,\Delta_D^T\,\left\{\mathbf{p}^{l-1}\right\}\right)$.

Denoising with Structured Sparse Prior

Instead of promoting piecewise constant solutions as in TV regularization, one can solve for group sparsity prior. Then the denosing problem becomes similar to (2.16) with $\mathbf{A} = \mathbf{I}$ as

$$\tilde{\mathbf{x}} = \text{prox}_{\lambda\|\mathbf{F}\mathbf{x}\|_{(2,1)}}(\mathbf{y}) = \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{x}\|_2^2 + \lambda\|\mathbf{F}\mathbf{x}\|_{(2,1)}. \tag{2.31}$$

The proximal map is computed in a similar fashion as TV denoising by starting with the dual formulation of the $\ell_{(2,1)}$-norm as

$$\|\mathbf{Fx}\|_{(2,1)} = \max_{\left(\sum_{i\in R_k} \mathbf{x}[i]^2\right)^{1/2} \leq 1} \langle \mathbf{Fx}, \mathbf{p} \rangle, \tag{2.32}$$

for each partition $R_k, k = 1, \ldots, M$. Then, the algorithm differs from TV only in the clipping function P_B in (2.30) which is applied over each partition instead of each coefficient of $\mathbf{x}$. In [24], Baritaux *et al.* formulate the proximal map in (2.31) for various linear mappings $\mathbf{F}$.

2.5.4 Forward-Backward Splitting

Until now, we have elaborated the optimization algorithms to solve the denoising problems, and computed the proximal maps. However, in a general framework we would like to find the minimizer of the cost function, $\mathcal{C}$, in the form

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \mathcal{C}(\mathbf{x}) = \arg\min_{\mathbf{x}} \mathcal{F}(\mathbf{x}) + \mathcal{R}(\mathbf{x}), \tag{2.33}$$

where $\mathcal{F}(\mathbf{x})$ is the quadratic data-term and $\mathcal{R}(\mathbf{x})$ is possibly non-smooth regularization term which is the case for sparsity-pursuing problems considered so far. Then, the minimizer, $\tilde{\mathbf{x}}$, can be computed via *forward-backward* splitting [34] as

$$0 \in \partial\mathcal{C}_T(\tilde{\mathbf{x}}) \in \nabla\mathcal{F}(\tilde{\mathbf{x}}) + \partial\mathcal{R}(\tilde{\mathbf{x}}), \quad \tilde{\mathbf{x}} - \mu\nabla\mathcal{F}(\tilde{\mathbf{x}}) \in (\mathbf{I} + \mu\partial\mathcal{R})(\tilde{\mathbf{x}}), \tag{2.34}$$

$$\Rightarrow \tilde{\mathbf{x}} = \underbrace{(\mathbf{I} + \mu\partial\mathcal{R})^{-1}}_{\text{backward step}} \underbrace{(\tilde{\mathbf{x}} - \mu\nabla\mathcal{F}(\tilde{\mathbf{x}}))}_{\text{forward step}} = \operatorname{prox}_{\mu\mathcal{R}}(\tilde{\mathbf{x}} - \mu\nabla\mathcal{F}(\tilde{\mathbf{x}})). \tag{2.35}$$

where μ is a proper step-size of the gradient descent. The forward step involves an explicit gradient step on $\mathcal{F}$ and the backward step employs an implicit projection on $\mathcal{R}$. The denoising model diminishes the forward model; i.e., $\mathcal{F}(\mathbf{x}) = \frac{1}{2}\|\mathbf{y} - \mathbf{x}\|_2$ leads to $\tilde{\mathbf{x}} - \mu\nabla\mathcal{F}(\tilde{\mathbf{x}}) = \mathbf{y}$ in (2.35) with $\mu = 1$. The forward-backward splitting method is described in Algorithm 3.

In the iterative scheme of forward step, the gradient descent at iteration l initiates from the linear approximation of $\mathcal{F}(\tilde{\mathbf{x}}^{l+1})$ around $\tilde{\mathbf{x}}^l$. Then, the optimization leads to the forward-backward splitting as

$$
\begin{aligned}
\tilde{\mathbf{x}}^{l+1} &= \arg\min_{\mathbf{x}} \mathcal{F}(\tilde{\mathbf{x}}^l) + (\mathbf{x}-\tilde{\mathbf{x}}^l)^T \nabla \mathcal{F}(\tilde{\mathbf{x}}^l) + \frac{1}{2}(\mathbf{x}-\tilde{\mathbf{x}}^l)^T \nabla^2 \mathcal{F}(\tilde{\mathbf{x}}^l)(\mathbf{x}-\tilde{\mathbf{x}}^l) \\
&= \arg\min_{\mathbf{x}} \mathcal{F}(\tilde{\mathbf{x}}^l) + (\mathbf{x}-\tilde{\mathbf{x}}^l)^T \nabla \mathcal{F}(\tilde{\mathbf{x}}^l) + \frac{1}{2\mu}(\mathbf{x}-\tilde{\mathbf{x}}^l)^T (\mathbf{x}-\tilde{\mathbf{x}}^l) \qquad (2.36) \\
&= \arg\min_{\mathbf{x}} \frac{1}{2\mu}\|\mathbf{x} - (\tilde{\mathbf{x}}^l - \mu \nabla \mathcal{F}(\tilde{\mathbf{x}}^l))\|_2^2 + \mathcal{R}(\mathbf{x}), \\
&= \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{x} - \underbrace{(\tilde{\mathbf{x}}^l - \mu \nabla \mathcal{F}(\tilde{\mathbf{x}}^l))}_{\tilde{\mathbf{x}}^{l+1/2}}\|_2^2 + \mu \mathcal{R}(\mathbf{x}), \\
&= \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{x} - \tilde{\mathbf{x}}^{l+1/2}\|_2^2 + \mu \mathcal{R}(\mathbf{x}) = \operatorname{prox}_{\mu\mathcal{R}}(\tilde{\mathbf{x}}^{l+1/2}), \qquad (2.37)
\end{aligned}
$$

where step-size should satisfy $\mu \in (0, 1/L)$ to guarantee the convergence with L being the smallest Lipshitz constant of $\mathcal{F}(\mathbf{x})$. The upper bound on the step-size μ in (2.36) is attained through the linear approximation of $\nabla\mathcal{F}$.

The forward-backward method is a specific class of splitting methods, for a general overview of proximal splitting methods see [29] and references therein. The (fast) Iterative Shrinkage Thresholding Algorithm (FISTA) [35, 36] is a special case of forward-backward splitting that solves ℓ_1-norm regularizations

$$
\begin{aligned}
\tilde{\mathbf{x}} &= \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_1, \quad \text{with steps} \\
\tilde{\mathbf{x}}^{l+1} &= \operatorname{prox}_{\frac{\lambda\|\mathbf{x}\|_1}{\|\mathbf{A}^T\mathbf{A}\|}} \left(\tilde{\mathbf{x}}^l - \frac{1}{\|\mathbf{A}^T\mathbf{A}\|}\mathbf{A}^T(\mathbf{A}\tilde{\mathbf{x}}^l - \mathbf{y}) \right).
\end{aligned}
$$

FISTA includes an intermediate variable to achieve quadratic convergence (optimally) similar to FGP presented in steps 5-6 of Algorithm 2. The fast step can be incorporated into forward-backward splitting (instead of step 6 of Algorithm 3).

Algorithm 3 Forward-Backward Splitting Algorithm $\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \mathcal{F}(\mathbf{x}) + \mathcal{R}(\mathbf{x})$

input: Noisy measurements $\mathbf{y}, \mu = 1/L, L$ is the smallest Lipschitz constant of $\mathcal{F}, \kappa_l \in]0, 3/2]$.

1: $l \leftarrow 1$
2: Initialize: $\tilde{\mathbf{x}}^0 = \mathbf{0}$
3: **repeat**
4: Update forward step $\tilde{\mathbf{x}}^{l+1/2} = \mathbf{x}^l - \mu\nabla\mathcal{F}(\tilde{\mathbf{x}}^l)$
5: Update backward step $\tilde{\mathbf{x}}^{l+1} = \operatorname{prox}_{\mu\mathcal{R}}(\tilde{\mathbf{x}}^{l+1/2})$
6: Update $\tilde{\mathbf{x}}^{l+1} = \tilde{\mathbf{x}}^l + \kappa_l(\tilde{\mathbf{x}}^{l+1} - \tilde{\mathbf{x}}^l)$
7: $l \leftarrow l + 1$
8: **until** convergence or the number of maximum iterations are reached.

Remarks

The sparsity-pursuing regularization problems can be solved with many algorithms other than proximal splitting methods:

1. The iterative reweighted norm (IRN) method represents the ℓ_1-norm in terms of weighted ℓ_2-norm and allows for inverting a quadratic function [37, 38].

2. Majorization-minimization methods compute a quadratic upper bound converging to the desired solution iteratively [39, 40].

3. Alternating direction method of multipliers (ADMM) forms an unconstrained optimization and minimizes the augmented Lagrangian function [41, 42].

Generalized Forward-Backward Splitting

A more sophisticated optimization problem with multiple regularization functions can be formed in order to impose different constraints on the solution as

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \mathcal{F}(\mathbf{x}) + \sum_{k=1}^{N} \mathcal{R}_k(\mathbf{x}), \tag{2.38}$$

where $\mathcal{F}$ is a smooth function and $\mathcal{R}_k$'s are non-smooth lower semicontinuous functions. In [43] a *generalized* forward-backward algorithm is proposed which ideally leads to a joint solution of multiple regularizations. The generalized forward-backward method is described in Algorithm 4. Other generalizations of the splitting method exist; especially when $\mathcal{F}$ is not smooth, Parallel Proximal Algorithm (PPXA) solves for the minimizer in (2.38). Another specific case, denoising, with $\mathcal{F}(\mathbf{x}) = \frac{1}{2}\|\mathbf{x} - \mathbf{y}\|_2^2$, reduces to the parallel Dykstra-like Proximal algorithm with appropriate weighting of each regularization term [44].

Algorithm 4 Generalized Forward-Backward Splitting Algorithm $\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \mathcal{F}(\mathbf{x}) + \sum_{k=1}^{N} \mathcal{R}_k(\mathbf{x})$

input: Noisy measurements $\mathbf{y}$, $\tilde{\mathbf{x}} = \mathbf{0}$, $\mathbf{p}_k = \mathbf{0}$, $\omega_k \in [0,1]$, $\sum_{k=1}^{N} \omega_k = 1$, $\lambda \in]0,1]$ and $\mu = 1/L$, L is the smallest Lipschitz constant of $\mathcal{F}$.

1: **repeat**
2:    **for** $k = 1$ to N **do**
3:       Solve for each regularization
$$\mathbf{p}_k = \mathbf{p}_k + \lambda \left(\mathrm{prox}_{\frac{\mu}{\omega_k}\mathcal{R}_k} \left(2\tilde{\mathbf{x}} - \mathbf{p}_k - \mu\nabla\mathcal{F}(\tilde{\mathbf{x}})\right) - \tilde{\mathbf{x}} \right)$$
4:    **end for**
5:    Update: $\tilde{\mathbf{x}} = \sum_k \omega_k \mathbf{p}_k$
6: **until** convergence or the number of maximum iterations are reached.

Chapter 3

Generalization of Total Variation Regularization

In this chapter[1], we propose a generalization of TV regularization for 1-D signals. As discussed in Chapter 2.4, TV regularization imposes sparsity on the derivative of the signal, which favors reconstruction with piecewise constant signals. We describe generalized L-TV for 1D signals by extending the conventional TV concept for any linear differential operator L. Basically, choice of the differential operator is chosen in accordance with the underlying linear system and the type of the driving signal we expect to recover. We start with the mathematical definition of the generalized L-TV and discuss the corresponding linear system and driving signal properties. We build the regularization using a filter representation of the general differential operator and employ a TV-like regularization adapted for the general operator. We validate our method on both synthetic examples and real audio signal.

3.1 Related Work and Contributions

TV regularization promotes piecewise constant signals by imposing sparsity on the signal's derivative. TV regularization is commonly applied where solutions with sharp edges are preferred, but when the underlying signal is smooth (e.g., piecewise polynomial signals), TV might lead to blocking artifacts. Such result is known as *staircase* effect. In order to deal with this problem, higher-order derivatives are adopted instead of first-order derivative inside the TV norm [46–50]. For example, TV with the second-order

[1]This chapter is based on the publication:
F. I. Karahanoglu, I. Bayram and D. Van De Ville. "A Signal Processing Approach to Generalized 1-D Total Variation", IEEE Transactions on Signal Processing, vol. 59, no. 11, pp. 5265-5274, Nov 2011 [45].

derivative is optimal for piecewise linear signals [51]. This extension has been used to retain smooth transitions while keeping sharp edges [50], for texture extraction [52], and has recently been reintroduced for MRI reconstruction [53].

The contributions of this work are the following:

1. Tailoring the derivative operator of TV to any linear differential operator. We generalize the TV denoising to be able to *deconvolve* any linear system that admits state-space representation. The problem is formulated as a denoising setting. As the regularizer is based on sparsity of the "innovation" signal, we also have access to the deconvolved versions of the solution.

2. Efficient convex optimization. The problem is cast as a convex optimization problem, and solved using fast and efficient solution schemes.

3.2 Generalized L-TV

In what follows, we explain all necessary steps and ingredients for developing generalized L-TV approach.

3.2.1 Mathematical Formulation

Instead of the first-order derivative in TV norm, generalized total variation admits a linear differential operator L. In the case of uniformly sampled discrete data, akin to TV regularization, the concept can easily be extended for discrete filters that are associated with general linear differential operators.

Definition 5 (Generalized Total Variation)**.** *For a discrete signal* $\mathbf{x}$*, we define the generalized* L*-TV regularizer as*

$$\mathrm{TV}_L\{\mathbf{x}\} = \sum_{n\in\mathbb{Z}} \left|\Delta_L \left\{\mathbf{x}\right\}[n]\right|, \tag{3.1}$$

where Δ_L *is the discrete version of the differential operator*

$$L = \prod_{i=1}^{N}(D-\alpha_i I)\left(\prod_{i=1}^{M}(D-\gamma_i I)\right)^{-1}, \tag{3.2}$$

with I *the identity operator,* $\alpha_i \in \mathbb{C}$*,* $i = 1,\dots,N$*, and* $\gamma_i \in \mathbb{C}$*,* $i = 1,\dots,M$*, the zeros and poles of the operator, respectively. We conveniently characterize the operator by* $\boldsymbol{\alpha} = (\alpha_1,\dots,\alpha_N)$ *and* $\boldsymbol{\gamma} = (\gamma_1,\dots,\gamma_M)$*.*

Clearly, the definition in (3.1) includes conventional TV; i.e., the case $N = 1$ with $\alpha_1 = 0$ and $M = 0$ reverts to that case when continuous operator D is associated with finite difference operator Δ_D. Moreover, we remark that for $M = 0$ the discrete operator Δ_L can be implemented as a filter with minimal finite support $N + 1$. From digital signal processing view, we build the simplest discrete approximation of the differential operator L; e.g., $\Delta_D[n] = [1\ -1]$ is a 2-tap filter. More sophisticated discrete filter designs ideally lead to better approximation of the continuous-domain operator in frequency domain. For $M > 0$, the support of the discrete operator Δ_L becomes infinite in general. Then, the operator can be constructed by a proper combination of causal and anti-causal filtering depending on the poles γ_i of the system. The following proposition summarizes how to obtain the discrete counterpart Δ_L of L. The proof can be found in Appendix A.1.

Proposition 1 (Discrete Implementation of L)**.** *Consider the continuous-domain linear differential operator $L = L_n L_d^{-1}$, where*

$$\begin{aligned} L_n &= \prod_{i=1}^{N}(D - \alpha_i I), \\ L_d &= \prod_{i=1}^{M}(D - \gamma_i I). \end{aligned}$$

We separate L_d into its causal (characterized by $\boldsymbol{\gamma}' = (\gamma_1', \dots, \gamma_{m_1}')$, $\mathrm{Re}(\gamma_k') < 0$) and anticausal ($\boldsymbol{\gamma}'' = (\gamma_1'', \dots, \gamma_{m_2}'')$, $\mathrm{Re}(\gamma_k'') > 0$) parts to ensure stability. Then, the discrete operator Δ_L associated with L can be obtained by a cascade of filtering operations corresponding to

$$\begin{aligned} \mathbf{y}[n] &= \sum_{k\in\mathbb{Z}} \mathbf{y}''[n-k]\Delta_{L_n}[k], \\ \mathbf{y}''[n] &= \mathbf{y}'[n+m_2]\Delta_{L_d''}[-m_2] - \sum_{k=-m_2}^{-1} \mathbf{y}''[n-k]\Delta_{L_d''}[k], \\ \mathbf{y}'[n] &= \mathbf{x}[n] - \sum_{k=1}^{m_1} \mathbf{y}'[n-k]\Delta_{L_d'}[k], \end{aligned} \tag{3.3}$$

The constituting filters are given in the time domain by

$$\begin{aligned} \Delta_{L_n}[n] &= (-1)^n \sum_{|\mathbf{m}|=n} (e^{\boldsymbol{\alpha}})^{\mathbf{m}}, \mathbf{m} \in [0,1]^N, 0 \le n \le N, \\ \Delta_{L_d'}[n] &= (-1)^n \sum_{|\mathbf{m}|=n} (e^{\boldsymbol{\gamma}'})^{\mathbf{m}}, \mathbf{m} \in [0,1]^{m_1}, 0 \le n \le m_1, \\ \Delta_{L_d''}[n] &= (-1)^n \sum_{|\mathbf{m}|=-n} (e^{-\boldsymbol{\gamma}''})^{\mathbf{m}}, \mathbf{m} \in [0,1]^{m_2}, -m_2 \le n \le 0, \end{aligned}$$

where we use the multi-index notation $\mathbf{m} = (m_1, \ldots, m_N)$*, with* $|\mathbf{m}| = \sum_{k=1}^{N} m_k$ *and the conventions* $c^{\mathbf{m}} = (c^{m_1}, \ldots, c^{m_N})$ *and* $\mathbf{c}^{\mathbf{m}} = \prod_{k=1}^{N} c_k^{m_k}$*.*

3.2.2 Problem Definition

We define an operator L_h associated with a linear system $h(t)$ as illustrated in Fig. 1. To focus the attention, we consider a linear system

$$x(t) = h(t) * u(t), \tag{3.4}$$

where $u(t)$ and $x(t)$ are the driving signal and the system response, respectively, and $*$ denotes the convolution operator.

We can define a pseudoinverse $L_h\{x\} = g * x$, where g is defined by

$$\hat{g}(\omega) = \begin{cases} 1/\hat{h}(\omega), & \text{if } \hat{h}(\omega) \neq 0, \\ \hat{h}(\omega) = 0, & \text{if } \hat{h}(\omega) = 0. \end{cases}$$

Then L_h will cancel the effect of the linear system (Fig. 1 second row):

$$L_h\{x\} = L_h\{h * u\} = g * h * u = u + u_{\text{null}},$$

where u_{null} is a null-space component, if it exists, of the system h, in other words, $h * u_{\text{null}} = 0$. Notice that a null-space component of L_h cannot be recovered.

We now illustrate these concepts by considering a linear system where the Fourier transform of its impulse response has the form:

$$\widehat{h}(\omega) = \frac{(j\omega - \tilde{\gamma}_1)}{(j\omega - \tilde{\alpha}_1)(j\omega - \tilde{\alpha}_2)(j\omega - \tilde{\alpha}_3)} \tag{3.5}$$

with three poles and one zero. Consequently, the differential operator L_h represents a third order differential equation and can be characterized in its turn in the Fourier domain as the inverse of h:

$$\widehat{L}_h(\omega) = \frac{(j\omega - \tilde{\alpha}_1)(j\omega - \tilde{\alpha}_2)(j\omega - \tilde{\alpha}_3)}{(j\omega - \tilde{\gamma}_1)},$$

where the system's poles take the role of the operator's zeros (and vice versa). The time-domain operator then corresponds to

$$L_h = (D - \tilde{\alpha}_1 I)(D - \tilde{\alpha}_2 I)(D - \tilde{\alpha}_3 I)(D - \tilde{\gamma}_1 I)^{-1}.$$

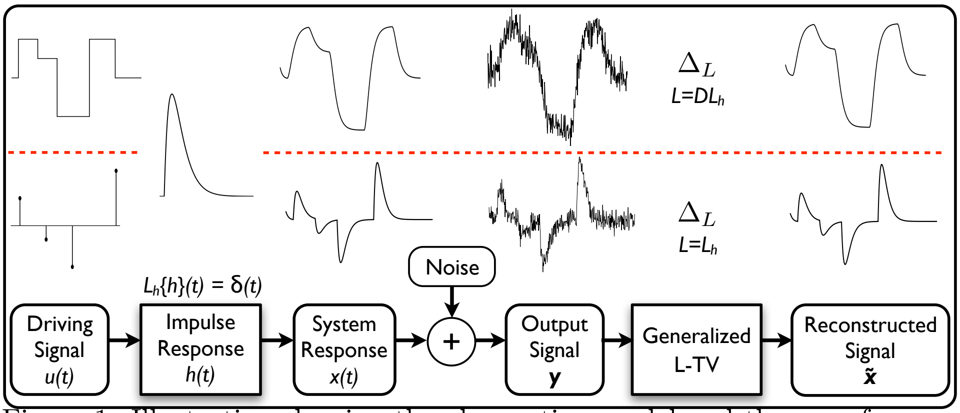


Figure 1: Illustration showing the observation model and the use of generalized L-TV. First, the driving signal serves as input to the linear system. Next, the (ideal) system response gets corrupted by noise. Generalized L-TV regularization aims at minimizing the ℓ_1-norm of the differential operator, which is tuned to the linear system and driving signal, applied to the reconstruction.

Variational Formulation

The system response $x(t)$ is corrupted by AWGN and sampled uniformly to obtain the output signal $\mathbf{y}$. Now, we cast our problem of finding the approximation $\tilde{\mathbf{x}}$ from the noisy measurements $\mathbf{y}$ into a variational formulation where we use the L-TV regularizer from (3.1). Then, the minimization problem becomes

$$\tilde{\mathbf{x}} = \text{prox}_{L-TV}(\mathbf{y}) = \arg\min_{\mathbf{x}} \mathcal{C}(\mathbf{x}) = \arg\min_{\mathbf{x}\in\mathbb{R}^N} \frac{1}{2} ||\mathbf{y}-\mathbf{x}||_2^2 + \lambda \, ||\Delta_L \{\mathbf{x}\}||_1 \, , \tag{3.6}$$

where λ is the regularization parameter and Δ_L is the discretized form of a differential operator of the form (3.2) that depends on L_h and on the type of the driving signal.

3.2.3 The Influence of The Differential Operator L

Linear System and Driving Signal

When dealing with a linear system with impulse response $h(t)$, as in Fig. 1, the differential operator L_h can be tuned to the inverse of the system response, such that $L_h\{h\}(t) = \delta(t)$. Therefore, $h(t)$ can be seen as the Green's function of the differential operator L_h. If the operator L is matched to the system in (3.6), $L = L_h$, the differential operator will "undo" the effect of the linear system, and regularization will be guided by the driving signal.

Table 3.1: The differential operator L needs to be chosen according to the linear system and the type of driving signal. The linear system is associated with a general operator L_h such that $L_h\{h\}(t) = \delta(t)$, $h_u(t)$ is Heaviside step function and I is the identity operator.

driving signal	linear system	optimal operator L
spikes	$\delta(t)$	I
piecewise constant	$\delta(t)$	D (conventional TV)
spikes	$h_u(t)$	D (conventional TV)
spikes	$h(t)$	L_h
piecewise constant	$h(t)$	DL_h
piecewise linear	$h(t)$	D^2L_h

The ℓ_1-norm leads to the optimal performance when the signal $L\{x\}(t)$ is spike-like or $\Delta_L\{\mathbf{x}\}$ is sparse. Therefore, tuning the operator $L = L_h$ will promote spike-like driving signals (Fig. 1 second row). More complex driving signals can be dealt with by further refining the operator; e.g., for a step-like driving signal a regular derivative can be added to the regularizing operator $L = DL_h$. Depending on the assumptions on the driving signal $u(t)$, the optimal differential operator includes higher order derivatives that sparsify the signal and make the ℓ_1-norm effective. Specifically, in Table 3.1, we give an overview of how the operator L should be chosen for various types of driving signals; e.g., spikes, piecewise constant, and piecewise linear.

Null Space Considerations

The optimal solution to (3.6) satisfies a compromise between the data- and the regularization terms. The signals with sparse representation after applying Δ_L have a low regularizer cost for the ℓ_1-norm and will be favored. In the continuous domain, any homogeneous solution $x_h(t)$ of the differential operator L has no cost for the regularizer since $L\{x_h\}(t) = 0$. This property also holds for the sampled version $\mathbf{x}_h$ and the discrete filter Δ_L; i.e., $\Delta_L\{\mathbf{x}_h\} = 0$. Therefore, null space components of the differential operator can be used at no regularization cost to minimize the residual data-fitting error. The regularization parameter determines the amount of adaption of the signal towards the prior. Setting $\lambda = 0$ diminishes the regularization whereas increasing λ strengthens the effect of regularization till only the null-space components fitting on the data are recovered.

3.2.4 Optimization Algorithm

The early application of TV regularization has been hampered by computationally expensive algorithms, but recent advances in convex optimization have led to fast algorithms for the ℓ_1 regularization problem. In what follows, we briefly explain our preferred algorithm to obtain a minimizer of the cost functional in (3.6). This minimization problem can be regarded as a generalized form of the analysis-prior denoising problem (see [4, 20] and the references therein). The Algorithm 5 is essentially an adaptation of the one provided by Chambolle for TV denoising [33, 54] in Algorithms 1-2 (see also [31] for a slightly different algorithm) for the discrete operator Δ_L.

Algorithm 5 Generalized L-TV Algorithm for Denoising
$\mathrm{prox}_{L-TV}(\mathbf{y}) = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y}-\mathbf{x}\|_2^2 + \lambda\, ||\Delta_L\{\mathbf{x}\}||_1$

input: Noisy signal $\mathbf{y}$, the regularization operator Δ_L and Lipschitz constant c subject to

$$c > \sup_{\omega} \left|\widehat{\Delta}_L(e^{j\omega})\right|^2 = \sup_{\omega} \frac{\prod_{i=1}^{N} |1 - e^{\alpha_i} e^{-j\omega}|^2}{\prod_{i=1}^{M} |1 - e^{\gamma_i} e^{-j\omega}|^2}.$$

1: $l \leftarrow 1$
2: Initialize: $k^1 = 1, \mathbf{p}^0 = \mathbf{0}, \mathbf{v}^0 = \mathbf{0}$
3: **repeat**
4: Update $\mathbf{p}^l = P_B\left(\Delta_L\{\mathbf{y}\}/(\lambda c) + (I - \Delta_L \Delta_L^T/c)\left\{\mathbf{v}^l\right\}\right)$
where the adjoint of Δ_L is $\Delta_L^T[n] = \Delta_L^*[-n]$ and
$P_B = \mathrm{sign}(\cdot)\,\min(|\cdot|, 1)$ denotes the elementwise clipping function,
5: Update $k^{l+1} = \frac{1+\sqrt{1+4(k^l)^2}}{2}$
6: Update $\mathbf{v}^{l+1} = \mathbf{p}^l + \frac{k^l-1}{k^{l+1}}(\mathbf{p}^l - \mathbf{p}^{l-1})$
7: $l \leftarrow l+1$
8: **until** convergence or number of maximum iterations are reached.
9: Set $\tilde{\mathbf{x}} = \mathbf{y} - \lambda\,\Delta_L^T\left\{\mathbf{p}^l\right\}$.

It is important to note that generalized L-TV can be combined with a general inverse problem as

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda\, ||\Delta_L\{\mathbf{x}\}||_1$$

Then, the solution can be obtained through a two-step optimization (forward-backward splitting Algorithm 3). The first (outer) loop tackles the deblurring problem while the second (inner) loop solves the denoising problem. The forward-backward scheme is discussed in Section 2.5.4.

3.3 Results

In this section, we present several examples to illustrate the performance of generalized L-TV. First, we demonstrate the signal reconstruction of a simulated linear system response with different driving signals and show how it outperforms conventional TV. Finally, we perform waveform analysis of audio signals by tuning the zeros of the operator to the central tone.

3.3.1 Proof of Concept

We considered a third-order system defined in (3.5) by its impulse response $h(t)$ with three poles and one zero. We assigned the poles $\tilde{\alpha} = \tilde{\alpha}_1 = \tilde{\alpha}_2 = \tilde{\alpha}_3 = -2$ and zero $\gamma = -0.1$. When the operator was tuned to the system directly $L = L_h$, the spike-like driving signal was reconstructed. The discrete version Δ_{L_n} of the forward operator corresponded to a finite impulse response (FIR) filter with four taps: $[1, -3e^{\tilde{\alpha}}, 3e^{2\tilde{\alpha}}, -e^{3\tilde{\alpha}}]$ and the discrete version of the inverse operator corresponded to a causal infinite impulse response (IIR) filtering with $\Delta_{L'_d} = [1, -e^{\tilde{\gamma}}]$. In Fig. 3.2(a), we show the original signal with the spike-like driving signal in the inset. We also added a random null-space component of the operator L (i.e., $c_1 e^{\tilde{\alpha}t} + c_2 t e^{\tilde{\alpha}t} + c_3 t^2 e^{\tilde{\alpha}t}$) as a "background". Next, the signal was corrupted by AWGN 15 dB (see Fig. 3.2(b)). Three different denosing methods are exploited: generalized L-TV, conventional TV and oracle Wiener filter, as shown in Figs. 3.2(c), 3.2(d) and 3.2(e), respectively. Moreover, we computed the filtered version Δ_L of the regularized solution in order to explore how well we reconstruct the underlying driving signal, see the insets in Figs. 3.2(c)-(e).

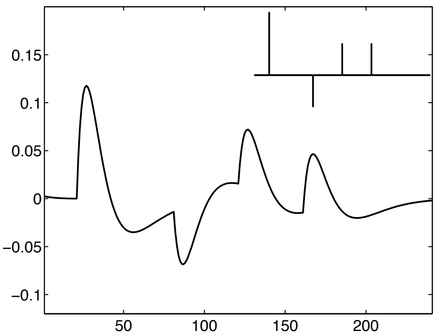
As an additional experiment, we considered piecewise constant driving signal, as shown in Fig. 3.3. Accordingly, we adapted the regularizing operator into $L = DL_h$, which only leads to an alteration of the FIR filter, $\Delta_{L_n} = [1, -1 - 3e^{\tilde{\alpha}}, 3e^{2\tilde{\alpha}} + 3e^{\tilde{\alpha}}, 3e^{\tilde{\alpha}} - e^{3\tilde{\alpha}}, e^{3\tilde{\alpha}}]$. The results are shown in Figs. 3.3c-e.

Regularization Parameter

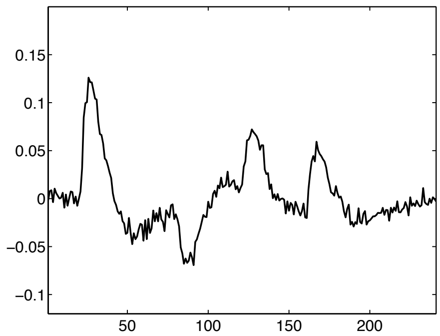
For each method, the oracle selected the optimal regularization parameter λ since it had access to the ground truth. When the oracle is not available at hand, the choice of the regularization parameter λ is important to calibrate the solution towards the desired constraints. Different strategies can be adapted for selecting λ; see [55] for classical references to methods such as generalized cross-validation and the L-curve, or [56] for a recent Monte-Carlo adaptation of Stein's Unbiased Risk Estimate that works well for TV, and [31, 57–60].

Signal to Noise Ratio Measurements

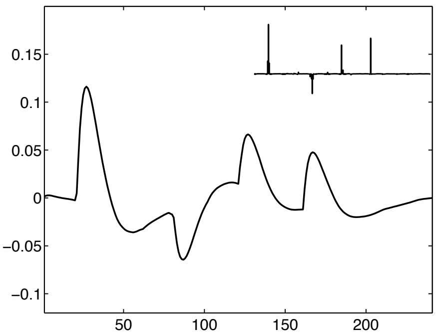
We deployed three different noise settings: AWGN corresponding to SNR level of 5, 10, and 20 dB. We report average SNR levels with standard deviation (over 100 realizations), with the optimal regularization parameter and maximum SNR, in Table 3.2. We compared the reconstruction quality obtained by oracle Wiener filtering (optimal for a Gaussian-process driving signal corrupted with AWGN), conventional TV, and generalized L-TV. As



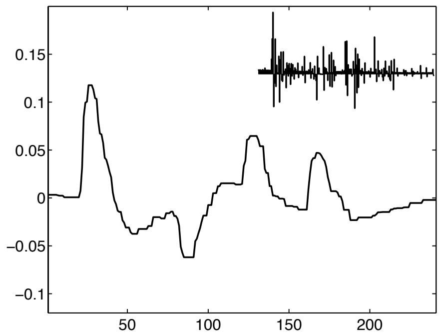
(a) Original signal
(inset: driving signal)



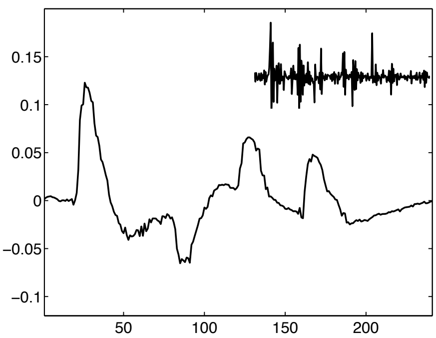
(b) Noisy signal



(c) Reconstruction using generalized L-TV regularization



(d) Reconstruction using conventional TV regularization



(e) Reconstruction using oracle Wiener filter

Figure 3.2: generalized L-TV, conventional TV and oracle Wiener filter solutions for a third-order linear system (with three poles and one zero) driven by spike-like signal.

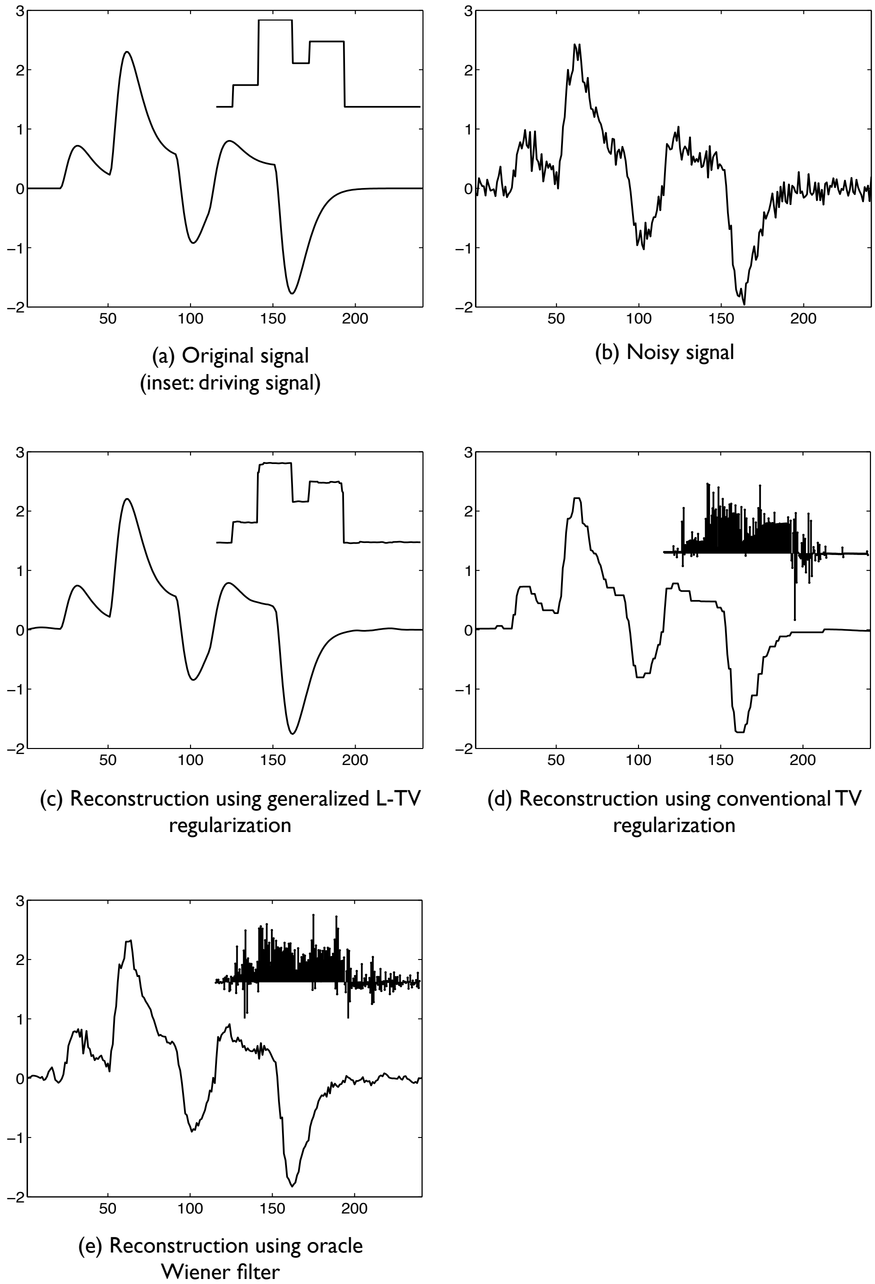


Figure 3.3: generalized L-TV, conventional TV and oracle Wiener filter solutions for a third-order linear system (with three poles and one zero) driven by piecewise constant signal.

expected, the results reveal that generalized L-TV is superior to conventional TV and the Wiener filter when tuning the operator L different from D is

Table 3.2: Overview of the performance measured as SNR (dB). The optimal tuning parameter λ was determined using an oracle. Average SNR and its standard deviation are reported for 100 realizations of the noise.

	SNR (dB)	λ	SNR (dB)	λ	SNR (dB)	λ
noisy signal 1	5	—	10	—	20	—
oracle Wiener	14.09 ± 0.87	—	17.45 ± 0.65	—	24.31 ± 0.41	—
conventional TV	12.32 ± 0.68	0.32 ± 0.04	15.89 ± 0.63	0.16 ± 0.05	22.15 ± 0.46	0.10 ± 0.01
generalized L-TV	14.86 ± 1.96	0.82 ± 0.38	19.58 ± 2.22	0.52 ± 0.25	28.71 ± 1.92	0.17 ± 0.09
noisy signal 2	5	—	10	—	20	—
oracle Wiener	14.17 ± 0.88	—	18.87 ± 0.84	—	26.90 ± 0.59	—
conventional TV	13.08 ± 0.72	1.96 ± 0.36	16.70 ± 0.64	0.98 ± 0.17	23.99 ± 0.41	0.22 ± 0.05
generalized L-TV	14.83 ± 1.68	10.02 ± 1.86	19.78 ± 1.86	4.90 ± 1.49	29.35 ± 1.60	1.70 ± 0.63

appropriate. In addition, as it can be appreciated from the corresponding Figures, the reconstruction of the underlying driving signal has high quality and can be useful for further processing in applications.

Tuning the Parameters of Differential Operator

An important concern is the robustness of the choice of the regularization operator L with respect to the underlying "true" linear system. To that aim, we generated signals for a spike-like input of a third-order linear system $L_h = (D - \tilde{\alpha}I)^3(D - \gamma I)^{-1}$ with $\tilde{\alpha}$ in the range $[-2, 2]$. We repeated the regularization for 10 surrogate signals corrupted by AWGN resulting into 10 dB SNR. Next, we applied several regularization strategies:

1. Generalized L-TV with α tuned exactly to the system $L = L_h$,
2. Generalized L-TV with $\alpha = -1$,
3. Generalized L-TV with $\alpha = 0$ to illustrate second-order TV,
4. Conventional TV.

We kept $\gamma = 4$ constant to eliminate pole-zero cancellation. In Fig. 3.4, we plot the average SNR (10 regularizations) for different $\tilde{\alpha}$ values of the linear system. As expected, we observe that second-order TV and matched generalized L-TV have equal performance at $\tilde{\alpha} = 0$. Similarly, generalized L-TV with fixed $\alpha = -1$ meets matched generalized L-TV at $\tilde{\alpha} = -1$. Moreover, we notice that SNR levels for matched generalized L-TV tend to increase further for larger values of $\tilde{\alpha}$. Conventional TV underperforms as the derivative operator is not well tuned to the linear system.

3.3.2 Audio Signal Example

We show that it is possible to tune the operator of generalized L-TV to include information about modulation, which can be useful for audio signals;

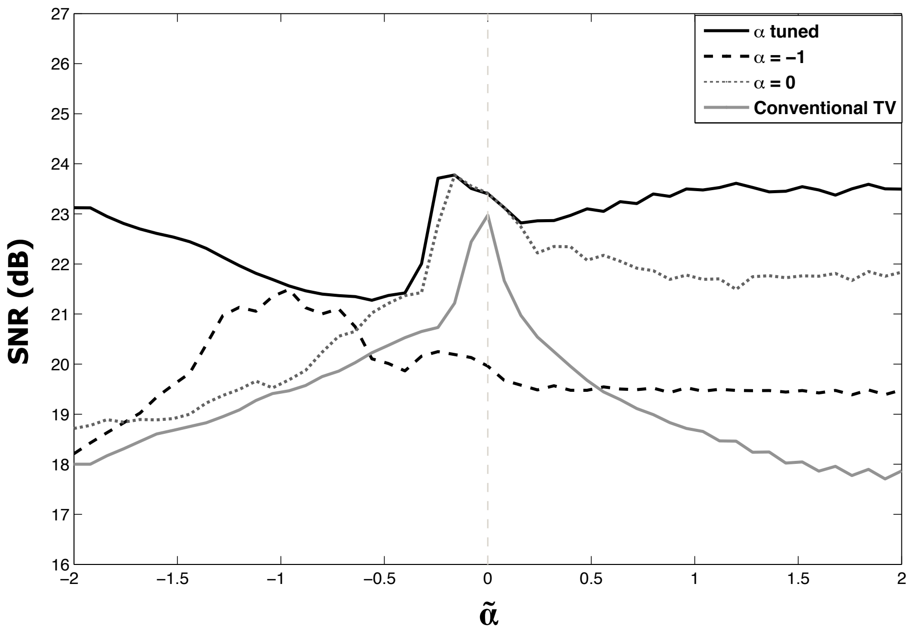


Figure 3.4: Performance measured as SNR (dB) for generalized L-TV and conventional TV regularization of a third-order linear system with equivalent differential operator $L_h = (D - \tilde{\alpha} I)^3 (D - \gamma I)^{-1}$, with $\gamma = 4$, for varying $\tilde{\alpha}$ values. The reported SNR measures are averaged over 10 realizations of AWGN (SNR 10 dB).

e.g., processing of tonal and transients layers [61].

Specifically, let us assume the simplified signal model as a sum of shifted decaying exponentials, each one modulated by a high-frequency sinusoidal function:

$$y(t) = \sum_i A_i \sin(\omega_{0,i}(t - t_i)) e^{\alpha(t - t_i)} u(t - t_i), \tag{3.7}$$

where $\alpha < 0$ is the decay rate. Eq. (3.7) can be considered as the sum of responses of linear systems with impulse responses $\sin(\omega_{0,i} t) e^{\alpha t} u(t)$ for spikes $A_i \delta(t - t_i)$, respectively. The corresponding transfer function is

$$\widehat{h}(\omega, \omega_{0,i}) = \frac{\omega_{0,i}}{(j\omega + (j\omega_{0,i} + \alpha))(j\omega + (-j\omega_{0,i} + \alpha))}. \tag{3.8}$$

Here, we exploit the differential operator

$$L = (D + (j\omega_0 + \alpha)I)(D + (-j\omega_0 + \alpha)I),$$

where ω_0 is the average frequency.

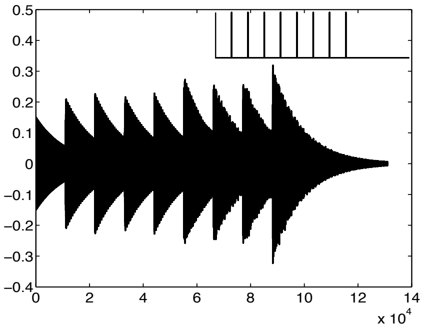
We generated a synthetic signal (Fig. 3.5(a) according to (3.7) with the first 9 notes of "Für Elise", whose frequencies range from $329 - 1318$ Hz, at sampling frequency $f_s = 44100$ Hz. The decay rate was $\alpha = -4\text{s}^{-1}$. We used the average frequency of the notes and $\alpha = -4\text{s}^{-1}$. The effect of tuning α is negligible on the output since the decay is very slow compared to the sampling frequency. We created a noisy realization of the audio signal by corrupting it with AWGN at SNR=15 dB, see Fig. 3.5(b). We show the output of generalized (25.3 dB), conventional TV (18.3 dB) and oracle Wiener (18.54 dB) in Fig. 3.5(c)-(e), respectively. The regularization parameter λ was chosen using an oracle.

Performance of the real data was tested through the analysis of "Glockenspiel" audio waveform [61]. We increased the multiplicity of zeros to make the frequency response of L increasingly flat around ω_0, and thus cancel also tones with nearby frequencies.[2] Finally, in Fig. 3.6(a), we show the "Glockenspiel", which was directly adopted from [61]. We degraded the signal with AWGN at SNR=15 dB, see Fig. 3.6(b). The output for generalized L-TV (18.3 dB), conventional TV (16.67 dB) and oracle Wiener (17.01 dB) is shown in 3.6(c)-(e). Note that the SNR values were computed against the real "Glockenspiel" signal, which contains some noise itself. Moreover, the signal contains different frequency components from perturbed harmonics in the tonal layer and sharp transitions in the transient layer, neither of which were modeled by the operator in generalized L-TV—nevertheless, the result is still better than the other methods we considered here. More advanced applications of generalized L-TV for sound wave processing can be devised in the future, such as the inclusion of multiple regularization terms with different operators each (e.g., for different frequencies and harmonics) and an additional model to deal with the transient layer.

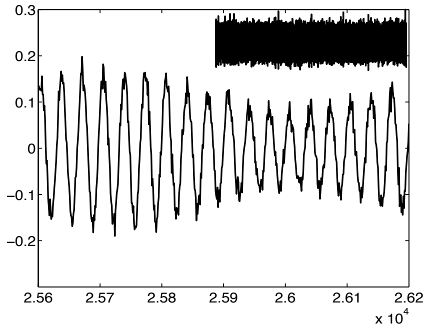
3.4 Discussion and Summary

Extensions to TV regularization have mainly focused on the use of higher-order derivatives [46–50, 52] and recently also on non-local generalizations [62]. We extend the basic TV concept further by introducing a general differential operator L instead of the derivative D, and motivate this choice by a linear system in the observation model. This allows a great deal of flexibility since we can take into account the presence of a linear system and different types of driving signals. The formulation is constructed as a denoising problem, however, it permits access to both the denoised signal and underlying driving signal (deconvolved). This work can be considered as the analysis prior counterpart of exponential spline wavelets [21] or generalized Daubechies wavelets [22], which are generalizations of Nth order derivatives. Indeed, these wavelets can be tuned to a given differential operator and their use in regularized reconstruction concurs with a synthesis prior.

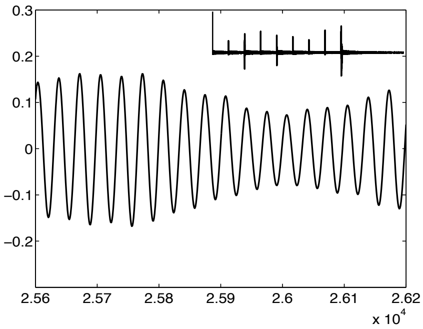
[2] Increasing the multiplicity of the zeros increases the size of the nullspace.



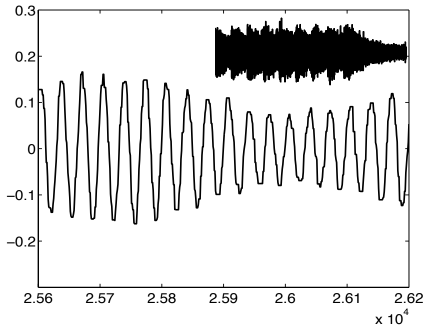
(a) Original signal (Für Elise)



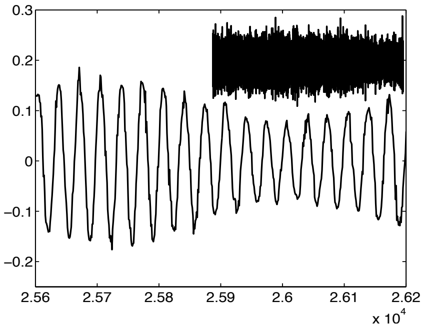
(b) Noisy signal



(c) Reconstruction using generalized L-TV regularization



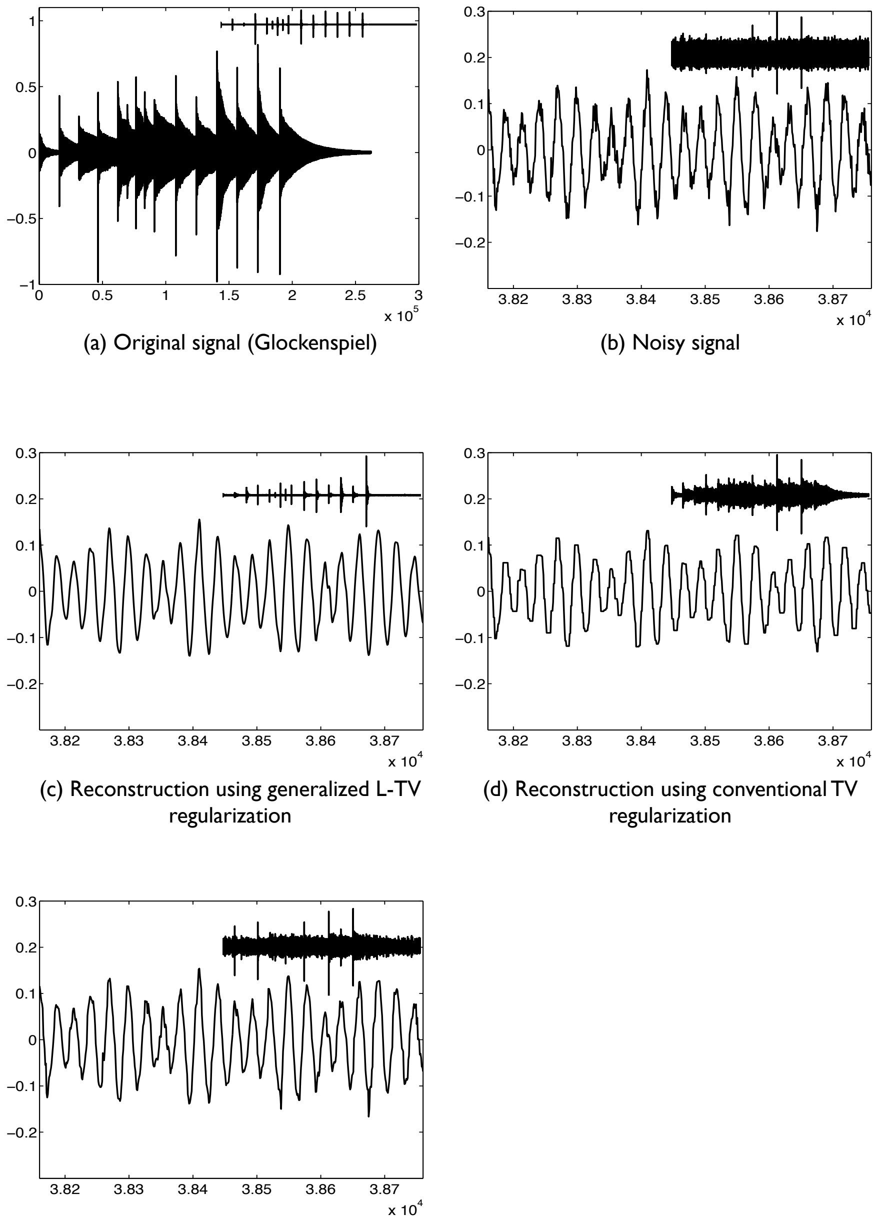
(d) Reconstruction using conventional TV regularization



(e) Reconstruction using oracle Wiener filter

Figure 3.5: Zoom of the denoised audio signal, Für Elise corrupted with AWGN 15 dB with generalized L-TV, conventional TV and oracle Wiener filter. We employ differential operator $L = (D+(j\omega_0+\alpha)I)(D+(-j\omega_0+\alpha)I)$ for Für Elise where ω_0 is the average frequency and $\alpha = -4\text{s}^{-1}$.

The generalized L-TV scheme will be further elaborated in Chapter 5 for analyzing the functional magnetic resonance imaging (fMRI) data, where the transfer function is considered as the fMRI's slowly varying hemodynamic system.



(a) Original signal (Glockenspiel)

(b) Noisy signal

(c) Reconstruction using generalized L-TV regularization

(d) Reconstruction using conventional TV regularization

(e) Reconstruction using oracle Wiener filter

Figure 3.6: Zoom of the denoised audio signal Glockenspiel signal corrupted with AWGN 15 dB with generalized L-TV, conventional TV and oracle Wiener filter. We employ differential operator $L = ((D + (j\omega_0 + \alpha)I)(D + (-j\omega_0 + \alpha)I))^3$ for Glockenspiel signal, where ω_0 is the average frequency and $\alpha = -4\mathrm{s}^{-1}$.

Chapter 4

Functional Magnetic Resonance Imaging

Functional magnetic resonance imaging (fMRI) is a non-invasive modality for visualizing brain function. Since its introduction in the 1990s, fMRI has been studied and applied extensively [63]. In this chapter, we explain how neuronal activity leads to changes in the fMRI signal, identify the underlying physiological processes, and describe fMRI data analysis by state-of-the-art methods. We begin by introducing some of the fundamental principles behind basic MRI acquisition. Then, we describe what the term *functional* implies. A complete overview of the concepts introduced here can be found in [64–66] and references therein.

4.1 Introduction

We discuss the underlying principle of MR image formation for different contrasts, which are related to the different intensities between tissues in the acquired images.

4.1.1 Basic MRI Principle

Quantum theory ascribes to each atomic nucleus an intrinsic property known as spin, which when non-zero gives rise to a magnetic moment. Normally, these moments are randomly aligned, producing zero net magnetization per unit volume. However, when placed within an external magnetic field, the spins start to align parallel (low-energy state) or anti-parallel (high-energy state) to the field direction. At equilibrium, there is a slight abundance of spins in the lower energy state, thereby producing a net magnetization, which precesses around the external field direction, considered henceforth as the longitudinal axis. The frequency of this precession, known as the Larmor frequency, depends linearly on both the external field strength, and a constant known as the gyromagnetic ratio, which varies with each nucleus.

In order to measure the magnetization, the equilibrium condition must be perturbed. Early experiments showed that this is possible by applying a circularly polarized magnetic field (i.e., the RF pulse) to the system. If the frequency of the RF pulse is matched to the Larmor frequency, known as the *resonance* condition, then spins in the low-energy state begin to absorb energy and pass to the high-energy state, effectively tipping the net magnetization into the transverse (measurable) plane. The magnetization can then be measured through simple circuitry. This principle is defined as nuclear magnetic resonance (NMR). Early experiments were performed by Rabi for the lithium nucleus [67], then by Purcell for solid substances [68], and by Bloch for water [69]. Indeed, NMR was first applied in chemistry in order to better understand the chemical composition of substances, known as NMR spectroscopy. Later, the first medical application was suggested by Damadian, who hypothesized that cancerous tissues might be delineated by exciting the water molecules with NMR [70]. Yet, it was unclear how NMR could be used to generate spatial images. The solution to MR image formation was found through the use of additional linear gradient fields, as proposed by Lauterbur and Mansfield [71, 72]. By allowing the resonance frequency to vary as a function of position, spatial information could be ascertained through the use of the inverse Fourier transform. Hence, the concept of MRI was born.

One of the primary reasons MRI has been so successful clinically rests in the diversity of contrasts it can provide. At a basic level, MRI contrast is defined by two intrinsic sample tissue properties, known as the T1 and T2 relaxation times, which are described further below.

1. Longitudinal relaxation (T1 recovery): Following an RF pulse (or else a general perturbation of the spin system from equilibrium), the net magnetization will tend to grow back to its equilibrium value along the longitudinal axis as spins interact with the surrounding lattice. The rate at which this occurs is known as the spin-lattice relaxation time, or T1.

2. Transverse relaxation (T2 - T2* decay): Once the net magnetization has been flipped onto the transverse plane, each local spin system (commonly referred to as an "isochromat") experiences a slightly different local magnetic field. The resulting variations in resonance frequency lead to a loss of coherence of the system as a whole, thus diminishing the net transverse magnetization–contributing to what is known as the spin-spin relaxation time, or T2. Additionally, inho-

mogeneities in the primary magnetic field compound this dephasing effect, further reducing the transverse relaxation time to T2*.

Biological tissues have different T1 and T2 relaxation times (T2 $\ll$ T1). Table 4.1 shows typical T1 and T2 relaxation times for brain tissues at 1.5 Tesla. The MR image is formed by choosing appropriate parameters and pulse sequences such that the images provide the maximum intensity difference (contrast) between tissues of interest. The two important parameters are: (1) the repetition time (TR), which is the elapsed time between successive applications of the selected pulse sequence, and (2) the echo time (TE), which is the time passed between the rf excitation and acquisition of the MR signal. Fig. 4.1 depicts T1 and T2 relaxation curves for gray matter, white matter and cerebrospinal fluid. Maximum T1 contrast between white matter and gray matter is usually obtained with a very short TE; i.e., minimum T2 effect, and relatively short TR. T2 weighted images are typically obtained with long TR; minimum T1 effect, and relatively long TE; maximizing the T2 effect. Another contrast that minimizes both T1 and T2 effects, with longer TR and very short TE, is called proton density-weighted imaging. Fig. 4.2 shows T1-weighted, T2-weighted and proton density MR images of the brain. For a complete overview of basic MR theory, pulse sequences, and image reconstruction, we refer the reader to [73, 74].

Table 4.1: T1 and T2 relaxation times for brain tissues at 1.5 Tesla

	Gray Matter	White Matter	Cerebrospinal Fluid
T1	900ms	600ms	4200ms
T2	100ms	80ms	2000ms

4.1.2 From MRI to FMRI: A BOLD Connection

FMRI measures physiological changes that are related to neuronal activity. Specifically, it relies on an endogenous contrast agent, hemoglobin, whose density can be monitored through the blood-oxygen-level-dependent (BOLD) response. Instead of revealing direct neuronal activation, fMRI indicates hemodynamic changes; e.g., alterations in the veins, cerebral blood volume (CBV) and blood flow (CBF). The relationship between neuronal activity and the BOLD signal is described through *neurovascular coupling*. Fig. 4.3(a) provides a step-by-step illustration of the vascular effect following neuronal activity. Basically, neuronal activity–the synaptic and spiking activity during information transmission–causes local energy and oxygen consumption. Therefore, a demand for nutrients is signaled to the veins, which triggers a vascular response. As a result, the CBV and CBF are increased. Specifically, the BOLD signal measures the ratio of deoxygenated

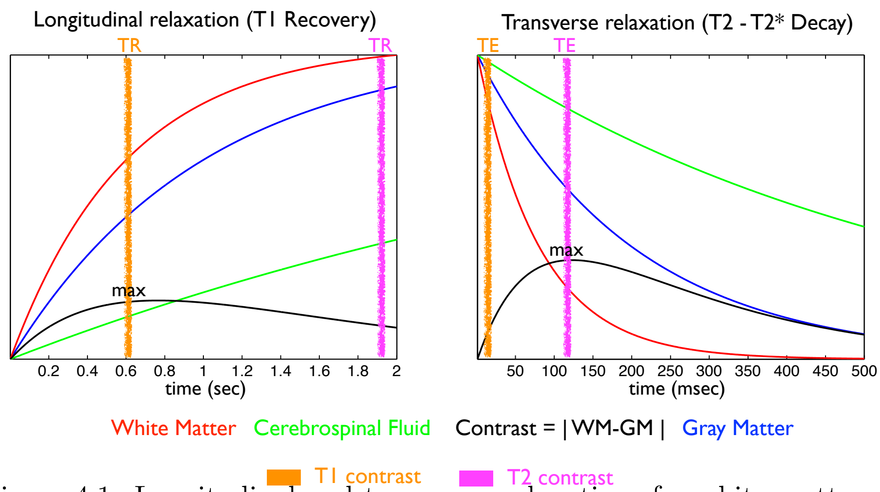


Figure 4.1: Longitudinal and transverse relaxations for white matter, gray matter and cerebrospinal fluid in the brain. Longitudinal relaxation time (left) shows the recovery of the net magnetization following excitation, the transverse relaxation (right) shows the decay in the transverse plane due to spin-spin interactions and magnetic field inhomogeneities. The contrast highlights the difference between the gray matter and white matter. Maximum T1 contrast (orange), in which white matter is represented by the highest intensities in the image, is achieved with short TR and short TE. Maximum T2 contrast is obtained with longer TR and longer TE, and cerebrospinal fluid is represented with the highest intensity.

hemoglobin (dHb) to oxygenated hemoglobin (Hb) during this chain of metabolic events. The dHb is paramagnetic, that is to say, has a high magnetic susceptibility, whereas the Hb is diamagnetic. The decrease in dHb/Hb ratio causes an increase in the fMRI signal, which can be detected with (T2-T2*)-weighted imaging. One expects that the dHb ratio would increase and hence BOLD would decrease due to oxygen consumption following neuronal activity, however, the vascular effect causes overcompensation of oxygen, which leads to an increase in the BOLD signal [77]. Fig. 4.4(b) shows the changes in BOLD, CBF and CBV. The BOLD signal increases a couple of seconds after the stimulus and slowly returns to the baseline value.

It is known that neuronal activity is correlated with the BOLD signal in fMRI, however, the exact relation between the neuronal activity and hemodynamic effect is not yet completely understood, and is still a topic of active research [64, 78, 79]. Simultaneous fMRI and direct single-neuron intracortical recordings suggest that the BOLD signal indirectly reflects integrative synaptic activity [2].

FMRI data is acquired using (T2-T2*)-weighted imaging, which maximizes

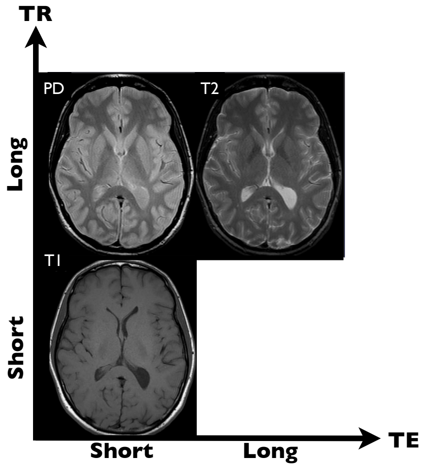


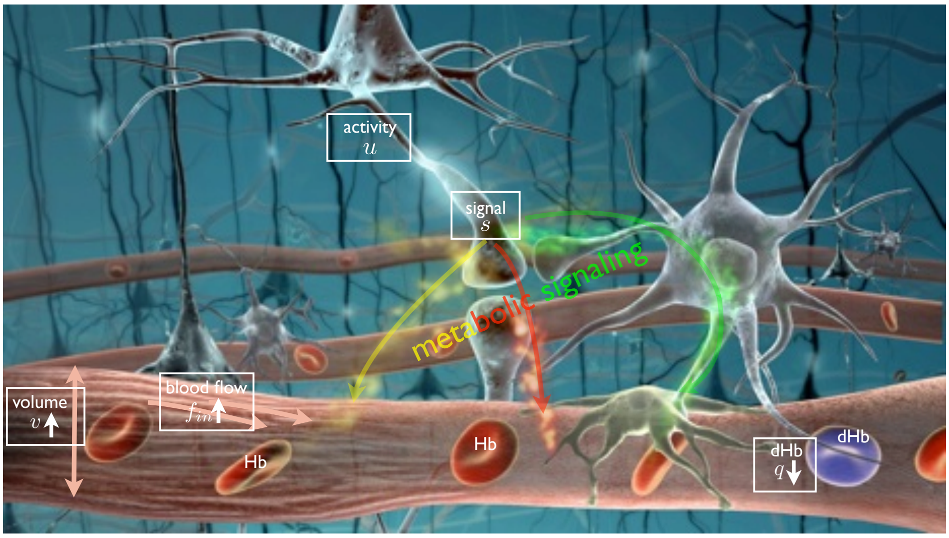
Figure 4.2: T1, T2 and proton density contrast MR images of the brain. Proton density image shows minimum T1 and T2 effects with long TR and very short TE. White matter is represented with highest intensity in T1 contrast, and T2 effect is minimized. In T2 contrasted images, the highest and lowest intensity values belong the cerebrospinal fluid and white matter, respectively.

the dHb/Hb concentration. The dataset is 4D, i.e., it is composed of the of time series of individual voxels in the brain. Temporal resolution is defined by TR (typically around $1-3$ sec.). Recently, faster fMRI acquisition schemes have been proposed [80]. The spatial resolution depends on the volume of interest. For a whole brain scan, the voxel size typically ranges between $1-4$ mm. per spatial dimension. Each volume consists of around (10'000–100'000) voxels. The duration of the fMRI experiment varies with the type of experiment respecting subjects' comfort in the scanner (typically around $2-20$ min.). Ultimately, a compromise must be reached between temporal and spatial resolution.

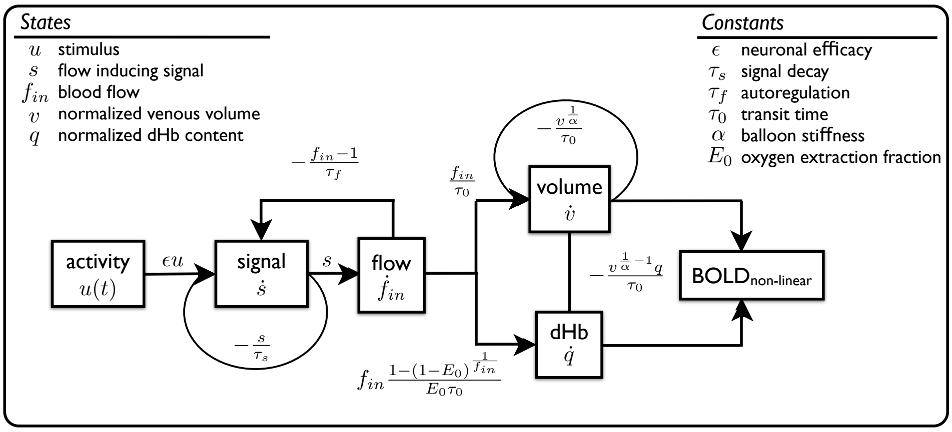
4.1.3 Noise in FMRI

The variability of the BOLD signal is not solely due to brain activity. The fMRI signal contains various sources of noise including:

1. Thermal noise due to the motion of electrons caused by both the subject and the scanner (e.g., eddy currents, heating), which tends to be spatially and temporally independent [81, 82],
2. (Aliased) physiological noise due to the subject's cardiac and respiratory fluctuations, as well as the interference of intravascular fluctuations (e.g., CBV and CBF) [83],
3. Subject head motion during scanning,



(a) Illustration of neurovascular coupling
(Courtesy, Arthur W. Toga, Laboratory of Neuro Imaging at UCLA)



(b) State-space representation of BOLD response including blood flow, blood volume and oxygentation concentration

Figure 4.3: The graphical representation of neurovascular coupling (Courtesy of Arthur W. Toga, Laboratory of Neuro Imaging at UCLA) and schematic representation of balloon model (Courtesy of [75, 76]). The neuronal activity signal $u(t)$ is the input of the system which is then converted to flow inducing signal s triggering the vascular effects. The blood flow f_{in}, blood volume v and dHb concentration are the other intermediate dynamic states in the system. The output is a nonlinear function of all states.

4. Other subject/experiment specific effects such as variability of the BOLD response, anatomy, experiment related artifacts (e.g., anticipation, habituation,variability across sessions etc.), [84, 85],
5. Gradual low frequency drifts and varying baselines in the voxel time series due to the magnetic field inhomogeneities, physiological effects

[86,87], which has been characterized by 1/f power spectrum [88],

Noise in fMRI can be typically characterized by autoregressive models [89]. Furthermore, if head motion and specific noise factors constituting low frequency drifts are carefully dealt with, then the residual errors might be assumed as white noise [83].

Modeling The BOLD Response

Apart from examining neural correlates of the BOLD response to a single stimulus, many studies have investigated explicit modeling of the BOLD response following stimuli. One important issue related to neurovascular coupling is the linearity of the BOLD response, which has been an ongoing research topic since the early days of fMRI. For example, nonlinear effects caused by successive applications of the stimuli have been observed [90,91]. The hemodynamic system, which reflects the neuronal activity underlying the BOLD signal, can be expressed via intermediate variables; blood flow, volume and oxygen concentration play an important role, and should be included in mathematical modeling [76, 92–95]. One such model, the so-called balloon model, was proposed by Buxton and colleagues [76], and later extended by Friston using Volterra kernel series [75] to represent the hemodynamic system by partial differential equations. In the state-space representation, the single-input single-output system is built with four state variables as:

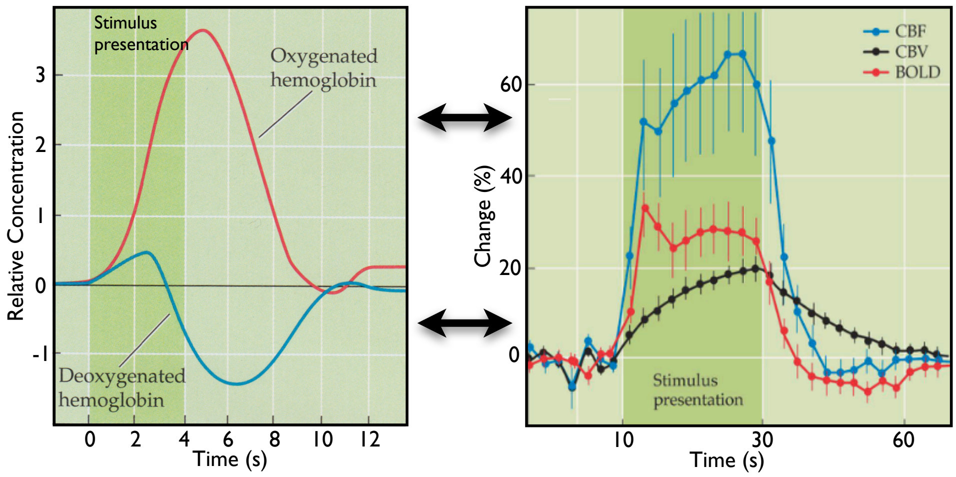
$$
\begin{aligned}
\dot{s} &= \epsilon u - \frac{s}{\tau_s} - \frac{f_{in} - 1}{\tau_f} \\
\dot{f}_{in} &= s \\
\dot{v} &= \frac{1}{\tau_0}\left(f_{in} - v^{\frac{1}{\alpha}}\right) \\
\dot{q} &= \frac{1}{\tau_0}\left(f_{in}\frac{1-(1-E_0)^{\frac{1}{f_{in}}}}{E_0} - v^{\frac{1}{\alpha}-1}q\right).
\end{aligned}
$$

The schematic representation and list of variables of the balloon model are depicted in Fig. 4.3(b). Then, the nonlinear BOLD signal is represented as BOLD

$$
\mathrm{BOLD}_{\mathrm{nonlinear}} = V_0\left(k_1(1-q) + k_2(1-\frac{q}{v}) + k_3(1-v)\right),
$$

where V_0 is the resting blood volume fraction and $k_1 = 7E_0, k_2 = 2, k_3 = 2E_0 - 2$ are the BOLD constants [82].

Nonlinear models, however, limit the fMRI data analysis. Besides the associated computational challenges, it leads to difficulties when quantitative



(a) Changes in oxygenated and deoxygenated hemoglobin due to the neuronal activity

(b) Changes in cerebral blood volume, cerebral blood flow and BOLD due to the neuronal activity

Figure 4.4: Hemodynamic response following neuronal activation (courtesy of [64, 77, 92]). The concentration of deoxygenated hemoglobin (dHb) increases towards the beginning of the stimulation due to oxygen consumption (a). The cerebral blood volume (CBV) and cerebral blood flow (CBF) increase with neuronal activity (b), which elevates the oxygen concentration (overcompensation) and reduces the dHB (a). The ratio of dHb to Hb forms the BOLD signal change (b).

comparisons of fMRI responses are needed; e.g., group analyses, comparisons of responses over different tasks or in different brain regions. Alternatively, linear models have been extensively exploited in fMRI studies due to their simplicity and further ease of data interpretation; e.g., group and longitudinal studies. There is indeed experimental evidence suggesting linear behavior of BOLD in the literature [96–100]. The BOLD signal is thus represented through a linear shift-invariant system as a convolution of the input (e.g., stimulus) with the hemodynamic response function (HRF), which is the impulse response.

The most acknowledged HRF approximation is the canonical HRF, which uses a combination of gamma functions [82]. Another model is the linear approximation of the balloon model proposed by Khalidov *et al.* with first-order Volterra kernel [101]. Here, we describe the model introduced in [101] in detail. The state-space representation for the new four states $\{x_1, x_2, x_3, x_4\} = \{s, 1 - f_{in}, 1 - v, 1 - q\}$ is represented as

$$\begin{aligned}
\dot{x_1} &= \epsilon u - \frac{x_1}{\tau_s} + \frac{x_2}{\tau_f} \\
\dot{x_2} &= -x_1 \\
\dot{x_3} &= \frac{1}{\tau_0}\left(x_2 - \frac{x_3}{\alpha}\right) \\
\dot{x_4} &= c x_2 - \frac{1-\alpha}{\alpha\tau_0} x_3 - \frac{1}{\tau_0} x_4,
\end{aligned}$$

leading to

$$\mathrm{BOLD}_{\mathrm{linear}} = V_0\left(\left(k_1 + k_2\right) x_4 + \left(k_3 - k_2\right) x_3\right).$$

Then for such a system, $\mathrm{BOLD}_{\mathrm{linear}}$ for a single input $u(t)$ is derived as

$$\begin{aligned}
&\dot{\mathbf{x}}(t) = A\mathbf{x}(t) + Bu(t), \quad \mathrm{BOLD}_{\mathrm{linear}}(t) = C\mathbf{x}(t), \\
&\mathrm{BOLD}_{\mathrm{linear}} = (C(D - A)^{-1}B)\{u\}
\end{aligned}$$

where $\mathbf{x}$ is the state vector, $\dot{\mathbf{x}}(t) = D\{\mathbf{x}\}(t)$, A and B are the system and input matrices, respectively, C is the output matrix,

$$A = \begin{bmatrix} \frac{-1}{\tau_s} & \frac{1}{\tau_f} & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & \frac{1}{\tau_0} & -\frac{1}{\alpha\tau_0} & 0 \\ 0 & c & -\frac{(1-\alpha)}{\alpha\tau_0} & -\frac{1}{\tau_0} \end{bmatrix}, \quad B = \begin{bmatrix} \epsilon \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 0 & V_0(k_3 - k_2) & V_0(k_1 + k_2) \end{bmatrix}.$$

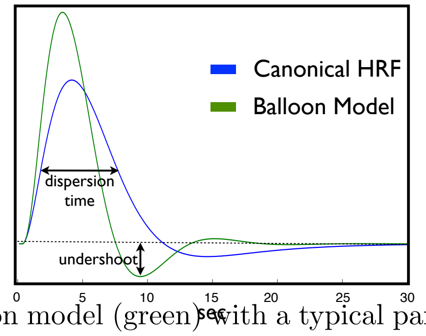


Figure 4.5: The balloon model (green) with a typical parameter set [75] and the canonical HRF (blue). The canonical HRF has a longer dispersion and smaller undershoot than balloon model.

The corresponding $BOLD_{\text{linear}}$ is defined explicitly as

$$\begin{aligned} \text{BOLD}_{\text{linear}} &= \frac{\frac{V_0\epsilon}{\tau_0}\left[\left(-\left(k_1+k_2\right)c\tau_0 - k_3 + k_2\right)D + \left((k_1+k_2)(\frac{1-\alpha}{\alpha\tau_0} - \frac{c}{\alpha}) - (k_3-k_2)\frac{1}{\tau_0}\right)\right]}{(D+\frac{1}{\tau_0}I)(D+\frac{1}{\alpha\tau_0})(D^2+\frac{1}{\tau_s}D+\frac{1}{\tau_f}I)}\{u\} \\ &= G\frac{D-\gamma I}{\prod_{i=1}^{i=4}(D-\alpha_i I)^{-1}}\{u\}, \end{aligned}$$

where $G = \frac{V_0\epsilon}{\tau_0}(-(k_1+k_2)c\tau_0 - k_3 + k_2)$.

A differential operator L_h that represents the inverse of the impulse response of the hemodynamic system is derived by assigning $u = \delta$. The differential operator L_h and its zeros and pole is represented explicitly [101] as

$$L_h = \prod_{i=1}^{i=4}(D-\alpha_i I)(D-\gamma I)^{-1}, \tag{4.1}$$

$$\begin{aligned} \vec{\boldsymbol{\alpha}} &= \left\{-\frac{1}{\tau_0}, -\frac{1}{\alpha\tau_0}, -\frac{1}{2\tau_s}\left(1 \pm j\sqrt{\frac{4\tau_s^2}{\tau_f} - 1}\right)\right\}, \\ \vec{\boldsymbol{\gamma}} &= \left\{\frac{(k_1+k_2)(\frac{1-\alpha}{\alpha\tau_0} - \frac{c}{\alpha}) - (k_3-k_2)\frac{1}{\tau_0}}{-(k_1+k_2)c\tau_0 - k_3 + k_2}\right\}. \end{aligned}$$

Fig. 4.5 depicts the canonical HRF and the linear balloon model. The canonical HRF is slowly decaying, whereas the balloon model has a shorter dispersion time with a higher undershoot.

The measured BOLD signal variability is not solely due to the neuronal response. Indeed, the measured fMRI signal is corrupted by different types of artifacts including, but not limited to subject motion in the scanner, inference from cardiac and respiratory signals (aliased), random thermally-generated noise, intersubject anatomical variability, subject specific effects, magnetic field inhomogeneity, and scanner drifts [81–83, 86]. Some of the noise effects can be reduced by using appropriate techniques, which are briefly discussed in Section 4.3.1.

4.2 FMRI Experiments

FMRI data is recorded sequentially, typically while subjects engage in a task (for example, watching a video), and is then analyzed in order to understand how the particular task is processed by the brain, e.g., which regions are interrelated. The experimenter derives a hypothesis and designs an experiment to address his/her question. The experiment often consists of different conditions such as presenting images of human faces vs. images of objects. The acquired data is later analyzed to verify the research hypothesis.

4.2.1 Task Based Stimuli

In task based studies (also known as task-evoked or task-induced in the literature), subjects are asked to perform a task, referred to as the experimental paradigm/stimuli, during the fMRI scan. There are generally three types of experimental designs for fMRI studies. Event-related designs consist of several short stimuli (events) presented at particular intervals of time. The events could be different conditions related to the cognitive study, such as presenting images of objects with different semantic meaning, a motor task, etc. The time elapsed between two consecutive events is referred to as the interstimulus interval (ISI), during which subjects do not perform any activity; i.e., the baseline condition. The aim is for the subject to demonstrate differences between specific conditions. Another typical experimental design strategy consists of longer stimuli, called a block based design, such as auditory stimuli. Every block can represent a different kind of condition to be examined. Block based designs are known to provide higher and persistent BOLD contrast, as longer stimulation causes an accumulation of activation. However, the prolonged stimuli might introduce anticipation and habituation effects (i.e., subject interest is lost due to repetition of the condition) or lead to saturation in the measured BOLD response. Finally, a mixed design constitutes a combination of block based and event-related designs, where a series of short stimuli are included in blocks. Fig. 4.6 illustrates the types of design schemes employed in fMRI analysis.

4.2.2 Resting-State fMRI

Task-related studies are intended for comparing different experimental conditions that are successively manipulated during scanning. On the other

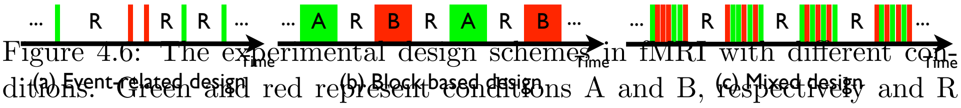


Figure 4.6: The experimental design schemes in fMRI with different conditions. Green and red represent conditions A and B, respectively and R represents resting (baseline). Event-related designs consist of short stimuli (a) whereas in block based designs each condition is presented for a prolonged period of time (b). Mixed designs are composed of several different stimuli in blocks (c).

hand, resting-state fMRI (RS-fMRI) examines brain activity while subjects are not engaged in any particular task, but rather relax and do not think about specific events. The concept of RS-fMRI was introduced by Biswal *et al.* [102] in a study showing that slow fluctuations in the time courses of regions that are co-activated during a (motor) task are also correlated during rest. These co-active brain regions during rest suggest the presence of an intrinsic brain mechanism–when the task is implicit. Even though there is no consensus on the definition and the role of the resting baseline state, it is clear that the brain consumes a high amount of energy during this paradigm-free period [103]. Many studies validate *task-negative* brain regions, which are suspended during the task. These universal regions are designated as the default-mode network [104–106]. Scientists have been further able to uncover other characteristic patterns of activation during RS-fMRI, which are referred to as resting-state networks (RSN) [107].

4.3 FMRI Data Analysis

Selecting a suitable method depends highly upon the aim of the study. Two classes of techniques can be distinguished: *confirmatory* approaches put forward a hypothesis to be verified with the data from the experimental paradigm. *Exploratory* approaches, or data-driven methods, aim to discover the underlying structures of the data that are not anticipated a priori. We discuss the state-of-the-art of confirmatory and exploratory analyses employed for fMRI studies.

4.3.1 Preprocessing

The literature contains a broad range of preprocessing methods proposed for removing noise factors in fMRI data to enable further processing. Many prominent preprocessing steps have been included in fMRI software toolboxes; e.g., SPM, FSL, AFNI, BrainVoyager, FreeSurfer etc. Fig. 4.7 illustrates a prototypical preprocessing scheme. The order of the methods may vary or some methods might be refrained according to the purpose of the study. Quantitative analysis of different preprocessing pipelines can be found in [108,109]. Basically, the acquired fMRI data are corrected for head motion, then functional images are registered with structural images, or vice versa, and projected into a common image space (e.g., MNI, Talairach). Projecting from subjects' individual space to a common space is practical for referential purposes and allows group analysis. The temporal filtering might include the detrending for low frequency components, high pass, low pass filtering, and physiological noise correction [110]. Spatial smoothing reduces the inter voxel variability and introduces spatial correlation that is required by the Gaussian random field theory to assess the statistical power of the analysis.

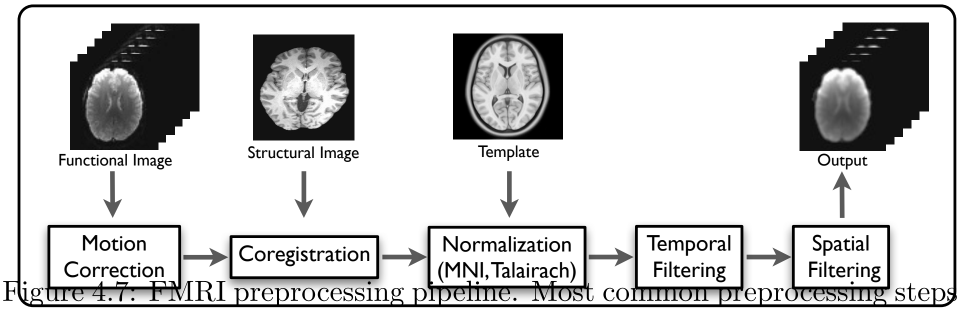


Figure 4.7: FMRI preprocessing pipeline. Most common preprocessing steps include: (1) the realignment of functional images for subject head motion correction, (2) projection of functional images into the structural space, or vice versa, (3) normalization of functional and structural images into a common space (e.g., MNI space), (4) detrending for low frequency drifts, high pass filtering, or physiological noise correction, etc., (5) spatial filtering, generally smoothing, in order to reduce spatial variability and make use of the Gaussian random field theory for thresholding.

4.3.2 Confirmatory Methods

Based on the research question, the experimenter first designs an experimental paradigm, and then exploits a confirmatory analysis to affirm or reject the hypothesis. General linear model (GLM) is one of the most acknowledged method in task-related fMRI analysis.

General Linear Model

Here, we give the key ingredients for GLM analysis at a single subject level; for a review we refer to [65, 111] and references therein. GLM is a linear regression model where the experimental paradigm is used to construct temporal regressors and then fitted into every voxel's time series to recover the weights of each regressor. Finally, statistical hypothesis testing reveals task-related activation maps. Specifically, consider

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \tag{4.2}$$

where $\mathbf{y} \in \mathrm{R}^N$ vector representing the measured and sampled time course of a single voxel, N being the number of time points, $\mathbf{X} \in \mathrm{R}^{N\times M}$ is the deterministic design matrix composed of M regressors, $\boldsymbol{\beta} \in \mathrm{R}^M$ is the parameter weight vector, and $\boldsymbol{\epsilon}$ is the additive noise. We assume a correlated Gaussian noise with $\mathbf{p}_\epsilon \sim \mathcal{N}(0, \sigma^2\boldsymbol{\Sigma})$ with $\boldsymbol{\Sigma}$ is the covariance matrix. The aim of the analysis is to recover the parameter weights $\boldsymbol{\beta}$ given the design matrix $\mathbf{X}$, then one can assess the voxels best represent an experimental condition up to a certain statistic.

Fig. 4.8 depicts a complete picture of GLM (colored). The GLM is a mass-univariate method; i.e., it is applied for each voxel time series separately. The task conditions are first convolved with canonical HRF $h(t)$ to retain the temporal characteristic of fMRI and then fed into the design matrix $\mathbf{X}$. Generally, the design matrix embodies additional factors, nuisance regressors $n(t)$ that are related to the non-experimental sources of variability, such as temporal derivatives of HRF, subject's motion parameters, basis sets representing scanner drift effects or subject cardiac and respiratory effects, etc [81–83, 86].

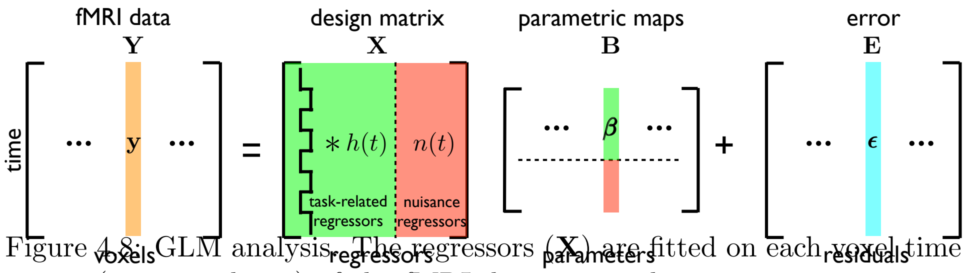


Figure 4.8: GLM analysis. The regressors ($\mathbf{X}$) are fitted on each voxel time course (orange column) of the fMRI data $\mathbf{y}$ in order to recover parameter weights $\boldsymbol{\beta}$. The design matrix can include both experiment related regressors (green) and nuisance parameters (red). The aim is to find the parameter weights $\boldsymbol{\beta}$ that minimize the empirical error $\boldsymbol{\epsilon}$ (cyan).

Then, exploiting generalized least squares or the Bayesian inference (ML), the weights are estimated as

$$\hat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}\in R^M} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \tag{4.3}$$

$$= (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{y} = \mathbf{P}\mathbf{y}, \tag{4.4}$$

which requires an estimate of the covariance matrix $\boldsymbol{\Sigma}$. The covariance structure can be estimated by a two step regularization iteratively; i.e., given covariance matrix we recover the weights, and then we estimate the covariance matrix from the residuals [112]. Note that, assuming an i.i.d. Gaussian noise $\boldsymbol{\Sigma} = \mathbf{I}$ yields $\hat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$.

Once the weights are estimated, t-test may be used to verify the hypothesis. If the hypothesis is "the effect of performing task 1 or task 2 is the same"; i.e., $\mathcal{H}_0 = \mathbb{E}\left[\mathbf{c}^T\hat{\boldsymbol{\beta}}\right] = 0$, with the contrast vector $\mathbf{c} = [1, -1, \mathbf{0}]^T$, it asserts that the difference between weights can be explained by chance. Finally, each voxel's statistic is compared to the null hypothesis with certain probability (p-value). The null hypothesis is rejected for the voxels whose test statistic is above a threshold that depends on the α-value, reflecting the desired p-value.

The t statistic yields

$$t = \frac{\mathbf{c}^T\hat{\boldsymbol{\beta}}}{\sqrt{\mathrm{var}(\mathbf{c}^T\hat{\boldsymbol{\beta}})}}, \tag{4.5}$$

where

$$\mathrm{var}(\mathbf{c}^T\hat{\boldsymbol{\beta}}) = \sigma^2\mathbf{c}^T(\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}^T)\mathbf{c} = \sigma^2\mathbf{c}^T(\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{c}. \tag{4.6}$$

The error variance is estimated from the residual error $\hat{\boldsymbol{\epsilon}} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{y} - \mathbf{P}\mathbf{y} = \mathbf{R}\mathbf{y}$ as

$$\hat{\sigma}^2 = \frac{\hat{\boldsymbol{\epsilon}}^{\boldsymbol{T}}\hat{\boldsymbol{\epsilon}}}{\mathrm{tr}(\mathbf{R})}. \tag{4.7}$$

The effective degrees of freedom of the Student's t-distribution can be computed [113] as

$$\nu = \frac{\mathrm{tr}(\mathbf{R})^2}{\mathrm{tr}(\mathbf{R}^2)}, \tag{4.8}$$

in order to be able to compare the t-values which are represented explicitly as

$$t = \frac{\mathbf{c}^T\hat{\boldsymbol{\beta}}}{\hat{\sigma}\sqrt{(\mathbf{c}^T(\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{c})}}. \tag{4.9}$$

The F-test compares the residual errors of the complete model to a reduced model. Basically the regressors are partitioned as $\mathbf{X} = [\mathbf{X}_1|\mathbf{X}_0]$; regressors of interest in $\mathbf{X}_1$ and those of no interest in $\mathbf{X}_0$. The projection matrices are computed to get the weights for the complete and reduced ($\mathbf{y} = \mathbf{X}_0\boldsymbol{\beta}_0 + \boldsymbol{\epsilon}_0$) models as $\hat{\boldsymbol{\beta}} = \mathbf{P}\mathbf{y}$ and $\hat{\boldsymbol{\beta}}_0 = \mathbf{P}_0\mathbf{y}$, respectively. Finally, F-statistic is computed as

$$F = \frac{\mathbf{y}^T\mathbf{M}\mathbf{y}/\mathrm{tr}\mathbf{M}\boldsymbol{\Sigma})}{\mathbf{y}^T\mathbf{R}\mathbf{y}/\mathrm{tr}(\mathbf{R}\boldsymbol{\Sigma})}, \tag{4.10}$$

where $\mathbf{R} = \mathbf{I} - \mathbf{P}$, $\mathbf{R}_0 = \mathbf{I} - \mathbf{P}_0$ are the corresponding residual-forming matrices and $\mathbf{M} = \mathbf{R}_0 - \mathbf{R}$.

Since the statistical testing is performed over massive amount of voxels, it will result (under $\mathcal{H}_0$) in many false positives, which are incorrectly rejected under the null hypothesis. The reliability and validity of the tests can be improved by correcting for multiple hypothesis testing by exploiting Benferonni correction, Gaussian Random Fields and False Discovery Rate [114–116].

4.3.3 Exploratory Methods

Not all brain activity can be modeled beforehand using stimulus functions. For example, spontaneous activity cannot be introduced as regressors in traditional GLM analysis [117]. Therefore, there is an increasing need for methodologies that enable the exploration of brain activity without predefined responses [66]. Exploratory methods are essentially employed when the experimenter is interested in extracting *useful* information from the data. These methods perform a data-driven analysis instead of imposing a prior model (e.g., design matrix), therefore they constitute a perfect candidate for analysing the spontaneous brain activity. Data-driven methods have been proposed for that purpose such as fuzzy clustering [118], temporal clustering analysis (TCA) [119,120], seed correlation analysis [102], or subspace decomposition methods such as independent component analysis (ICA) [121, 122], partial least squares [123, 124], canonical correlation analysis (CCA) [125] and agnostic canonical variates analysis (agnostic-CVA) [126]. We explain seed based analysis which have been the primer RS-fMRI analysis method and ICA which is probably the most commonly used data-driven method.

Seed-Based Analysis

This method has been used in the first RS-fMRI data analysis to reveal the coherent structures during rest [102]. Specifically, it is based on selection of an a priori seed voxel/region, e.g., posterior cingulate cortex (PCC), and computing the correlation between the seed's time course with the rest of the voxel time courses. Subsequent methods propose to feed the time course of the selected voxel/region into further analyses, such as GLM analysis and partial least squares (PLS) [127–129]. It is, however, computationally exhaustive to perform this analysis and assess the connectivity patterns for all voxels in the brain, and choice of the seed might not be optimal. Moreover, intrinsic activity in the seed voxel/region potentially introduces biased networks [130].

Independent Component Analysis

Independent component analysis (ICA) is a blind source separation method that aims at segregating the data into different compartments based on their statistical independence [131]. Basically, the original data is represented through a bilinear model; i.e., composition of the *independent* sources with associated weights (mixing matrix), so that the original data can be perfectly reconstructed. In fMRI, ICA can be performed either spatially or temporally; then the sources become either spatial maps or time courses, respectively. Spatial ICA is generally preferred since the spatial dimension is higher than the temporal dimension. ICA was first employed in fMRI for analyzing task-related data [132]. Later, it has been applied for RS-fMRI studies and for separating the noise components from the data [133]. Consider the following data model as in Fig. 4.9

$$\mathbf{Y} = \mathbf{A}\mathbf{S}, \tag{4.11}$$

where $\mathbf{Y} \in \mathrm{R}^{N \times V}$ is a matrix representing the measured and sampled fMRI data, N and V being the number of time points and number of voxels, respectively, $\mathbf{A} \in \mathrm{R}^{N \times M}$ is the mixing matrix with M components, and $\mathbf{S} \in \mathrm{R}^{M \times V}$ is source matrix where each component is a spatial map (in rows). Note the similarity of ICA to GLM in terms of graphical interpretation, that is, the mixing matrix $\mathbf{A}$ acts as the design matrix $\mathbf{X}$ that should be estimated.

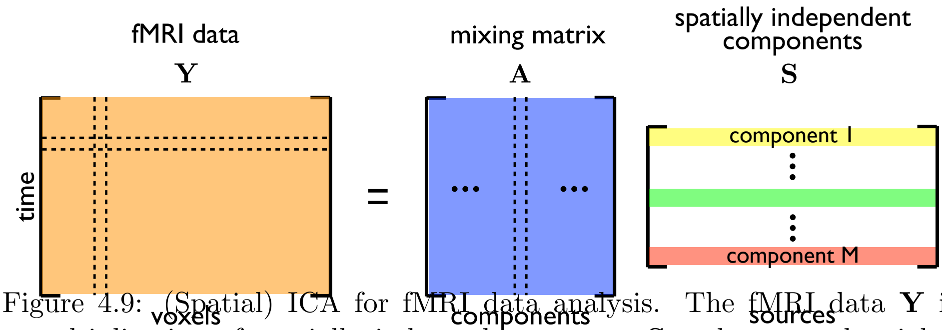


Figure 4.9: (Spatial) ICA for fMRI data analysis. The fMRI data $\mathbf{Y}$ is a multiplication of spatially independent sources $\mathbf{S}$ and temporal weights (mixing matrix $\mathbf{A}$)

The main idea is to find the unmixing matrix $\mathbf{W}$ that recovers the sources $\hat{\mathbf{S}}$ from the mixture data $\mathbf{Y}$; i.e., $\hat{\mathbf{S}} = \mathbf{W}\mathbf{Y}$. Indeed, ICA has some inherent assumptions that; (1) the sources are generated from independent processes, (2) the sources follow non Gaussian distribution, and (3) the observed data is a linear mixture of sources and follow a Gaussian distribution (according to the central limit theorem) [131]. Another multivariate decomposition similar to ICA is principal component analysis (PCA) which estimates the uncorrelated (orthogonal) sources best explain the variance in the data. Note that PCA can be interpreted as the Gaussian version of ICA. In the literature there are several methods to estimate the unmixing matrix $\mathbf{W}$ such as entropy maximization (infomax) [134], ML-ICA [135], maximization of non-gaussianity [136] etc. Even though ICA has been one of the most popular methods in fMRI analysis, it has some disadvantages:

1. ICA is not directly designed to identify "activation like" components since no knowledge is taken into account about the hemodynamics or about the type of activity-driven signal (e.g., spikes versus sustained activity).
2. Due to random initialization and complex criterion of the method its reproducibility is not always guaranteed [137].
3. The selection of the model order (how many components to estimate) remains as an issue [138, 139].

4.3.4 FMRI Deconvolution Methods

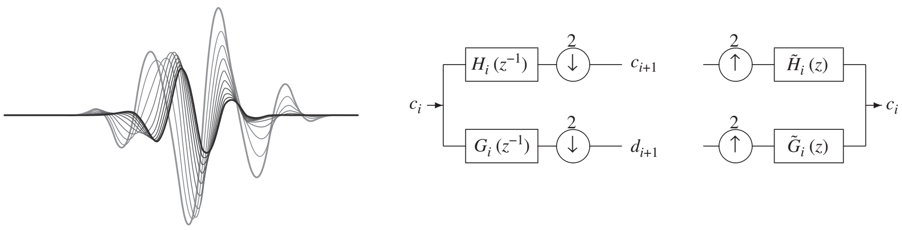
The GLM and ICA constitute the mainstream state-of-the-art methods in fMRI data analysis. Alternatively, temporal analysis of fMRI is of interest to the neuroscience community, especially the elaboration of unpredicted activations. Since GLM is only applicable when the task is explicit, and ICA does not incorporate any hemodynamic effect, new tools should be developed. FMRI deconvolution methods have been proposed to uncover the underlying activation signals in BOLD signal. Initially, Glover introduced Wiener deconvolution filtering that is optimal for Gaussian sources and thus results in very smooth activity-inducing signals [3]. This work was generalized by Gitelman *et al.* to study the psychophysiologic interactions at the neuronal level [140]. We shortly explain the prominent temporal deconvolution schemes that are of particular interest in this work;

1. *Activelets*; wavelets mimicking the hemodynamic system, are designed and used to decompose BOLD signals that should be ideally represented by sparse activelet coefficients. Extended from traditional wavelets, which behave like Nth-order derivatives, activelets are a family of exponential spline wavelets that annihilate the null-space of a general differential operator L; i.e., exponential vanishing moments. Fig. 4.10(a) illustrates the evolution of B-spline wavelets to activelets. Specifically, the activelets are constructed from the shifted replicates of Green's function of operator L_h, operator related to the inverse hemodynamic system in (4.1), in a multiscale formalism [21, 101, 141].

   Exploiting the activelet dictionary as the building blocks of BOLD signal, a sparsity-inducing variational formulation is constructed as a *synthesis* problem

$$\tilde{\mathbf{c}} = \arg\min_{\mathbf{c}} \|\mathbf{y} - \mathbf{\Phi}\mathbf{c}\|_2^2 + \lambda\|\mathbf{c}\|_1, \tag{4.12}$$

where $\mathbf{y}$ is the measured BOLD signal, $\mathbf{c}$ is the activelet coefficient, $\mathbf{\Phi}$ is the activelet dictionary, which satisfies the BOLD signal estimate $\tilde{\mathbf{x}} = \mathbf{\Phi}\tilde{\mathbf{c}}$. The activelet coefficient c are implicitly computed using the filter bank formulation; i.e., $\mathbf{\Phi}$ is not explicitly computed, see Fig. 4.10 (b).



(a) From B-spline wavelets (gray), related to Nth order derivative, to activelets (black), related to general differential operator

(b) Filter-bank implementation of activelets dictionary for one level

Figure 4.10: Activelets and filter-bank implementation. The B-spline wavelets are designed to annihilate high order polynomials, instead, activelets annihilate the operator L_h of the hemodynamic system (a). The implementation is performed exploiting the filter bank representation, at each level the coefficients are filtered with the low-pass (scaling $g(t)$) and high-pass (wavelet, $h(t)$) filters. The synthesis dictionary $\mathbf{\Phi}^T$ is implicitly computed using the dual filters in Fourier domain. (Courtesy of Khalidov *et al.* [101]).

2. *(Sparse) Paradigm-Free Mapping* is another deconvolution method aiming at recovering the underlying activity without any timing information. The method consists of two stages; (1) temporal deconvolution based on a *synthesis* dictionary, here canonical HRF, and solved through Tikhonov regularization, and (2) statistical analysis of the time courses [142]. Sparse paradigm-free mapping is later proposed imposing sparsity constraint on the dictionary coefficients $\mathbf{c}$ [143] as

$$\tilde{\mathbf{c}} = \arg\min_{\mathbf{c}} \|\mathbf{y} - \mathbf{H}\mathbf{c}\|_2^2 + \lambda\|\mathbf{c}\|_1, \tag{4.13}$$

   where dictionary $\mathbf{H}$ is composed of shifted replicates of canonical HRF, and finally the estimated BOLD activity $\tilde{\mathbf{x}}$ is recovered from the sparse coefficients $\tilde{\mathbf{c}}$ convolved with HRF; i.e., $\tilde{\mathbf{x}} = \mathbf{H}\tilde{\mathbf{c}}$.

3. *Event Detection by Iterative Estimation*; Garcia *et al.* [144] exploits the synthesis dictionary, shifted canonical HRFs, and formulates a regularization problem consisting multiple regularization terms; both ℓ_1-norm and TV as

$$\tilde{\mathbf{c}} = \arg\min_{\mathbf{c}\geq\mathbf{0}} \|\mathbf{y} - \mathbf{H}\mathbf{c}\|_2^2 + \lambda_1\|\mathbf{c}\|_1 + \lambda_2\|\Delta_D\mathbf{c}\|_1. \tag{4.14}$$

> The regularization presumes a combination of non-negative spike-like and block-like activity-inducing signals.

Interestingly, Wu *et al.* proposed an insightful method [145]. They first pinpoint the local maxima points in the BOLD signal and construct indicators function of its time shifted replicates. Then, regressors are built by the convolution of the indicator functions with the canonical HRF and its second order variants. GLM analysis yields the HRF estimates for all indicator functions where minimum error criteria determines the best representative of the neuronal activity and HRF. Finally, BOLD deconvolution is performed using a Wiener filter with the corresponding HRF estimate.

The aforementioned methods presume the presence of a linear hemodynamic model. Nonlinear models have also attracted a lot of attention for blind deconvolution. They basically solve the continuous state-space representation constructed by stochastic differential equations. Note that, the inversion of the nonlinear hemodynamic system to uncover the underlying state transitions has been studied for a while; for example, dynamical causal modelling for effective connectivity analysis between brain regions [146]. Instead, Riera *et al.* proposed the deconvolution of the underlying neuronal signal and estimation hidden states through Bayesian filtering [147]. Later, Friston *et al.* developed dynamic expectation maximization (DEM), variational filtering and generalised filtering with extended Kalman filtering [148–150]. An extended method is developed by Havlicek *et al.* using nonlinear cubature-Kalman filtering [151]. In a recent work, Bush *et al.* presume a parametric model for underlying activations and exploit nonlinear regression to deconvolve the hemodynamic system [152]. These methods, however, have high computational cost compared with linear models and are generally preferred rather for region of interest (ROI) or network analysis.

Besides the deconvolution methods with a fixed HRF model, many studies in the literature investigate temporal dynamics via HRF identification. In standard GLM approaches [82], the HRF is predefined using gamma functions which does not take the subject variability into account. The temporal and dispersion derivatives are often incorporated into GLM to account for intra-subject variability. More flexible techniques are suggested to estimate HRF components in a subject- or time-dependent way in order to deal with inter- and intra-subject variability [84], such as parametric model, non-parametric models using voxel-wise or region-wise priors [153–156]. More notably, parcel-based HRF estimation methods through joint detection estimation (JDE) framework are studied based on Bayesian approaches [157–159]. Recently, an adaptive parcel identification driven by the hemodynamics is proposed using JDE [160, 161]. These HRF identification methods are mainly combined with GLM analysis to explore the parcel/subject/group/task specific hemodynamic models.

Chapter 5

Total Activation

In this chapter [1], we introduce a novel spatiotemporal deconvolution method, for which we coin the term *total activation* (TA), for fMRI data analysis. Our framework allows exploring the underlying activations related to the BOLD signal without any timing information. We begin by defining the fMRI signal model upon which we build TA framework. We then formulate a regularization problem with carefully chosen temporal and spatial priors that take into account the specific characteristics of fMRI data. First, the temporal regularization—based on generalized L-TV framework (introduced in Chapter 3)—is adapted to the hemodynamic system in order to recover the system's driving (activity-inducing) signals. Second, the spatial regularization term promotes coherent activation patterns in anatomically-defined brain regions. The utility of TA is corroborated by 3D phantom experiments where TA significantly reduced the deviations in the activity-related signals.

5.1 Related Work and Contributions

We mentioned three particular state-of-the-art deconvolution methods in Chapter 4.3.4 that have deployed similar ideas to recover activity-inducing signals [101, 142–144]. All of these methods were cast as variational formulations that aim at temporal deconvolution. However, they operate solely in the temporal domain in a voxel-wise fashion and do not use spatial information. Moreover, the variational formulation admits a *synthesis* prior, which means that a dictionary of atoms is built. Khalidov *et al.* designed activelets,

[1] This chapter is based on the publication:
F. I. Karahanoglu, C. Caballero-Gaudes, F. Lazeyras, and D. Van De Ville, "Total activation: fMRI deconvolution through spatio-temporal regularization", NeuroImage, vol. 73, pp. 121-134, June 2013 [162].

as new flexible wavelets tailored to the hemodynamic system, to decompose BOLD signals such that the BOLD signal is ideally represented with a few large activelet coefficients if the underlying activity is spike-like [101]. A similar idea was developed by Caballero-Gaudes *et al.* where an explicit dictionary with all possible shifts of the canonical HRF was constructed to recover spike-like activity [142, 143]. Akin to the dictionary based methods, Garcia *et al.* exploited two regularization terms, which in turn provide the spike-like and block-like atoms. This, however, raises the issue of how to choose the regularization parameters to adapt to the ratio of spikes and blocks for each voxel [144].

We have proposed TA to deconvolve fMRI data based on hemodynamic and anatomical properties of the brain. The unique features of TA can be summarized as:

1. Reveal temporal properties of the activity-inducing signal. The deconvolution identifies the "innovation" signal (which is spike-type) as the sparse driver of the BOLD signal. The innovation signal then defines the activity-inducing signal (by integration), which is a flexible block-type signal where the timing and duration is driven from the data.

2. Combine temporal and spatial regularization. Spatial regularization is incorporated using structured sparsity as expressed by mixed-norms based on a priori knowledge of spatial structure of the data [24]; i.e., time courses of voxels in the same brain regions are favored to be coherent.

3. Take advantage of efficient optimization schemes. We employ the efficient generalized forward-backward splitting algorithm [43], which is a fast iterative shrinkage algorithm that alternates between temporal and spatial domain solutions until convergence to the final estimate of the underlying activity-inducing signal. Moreover, we utilize an adaptive scheme to systematically calibrate the regularization parameter.

5.2 Total Activation

5.2.1 Generative BOLD Signal Model

In the sequel, we describe the temporal properties of the fMRI signal model adopted for our work. We exploit a linear shift-invariant system, which specify the activity-related signals $x(i,t)$ as convolution of the hemodynamic response function (HRF) $h(t)$ with the activity-inducing signals $u(i,t)$:

$$x(i,t) = u(i,t) * h(t), \tag{5.1}$$

where $i \in \mathbb{Z}$ is the voxel index.

Essentially, we model the activity-inducing signal as the block-like driving signal of the hemodynamic system. Thus, we represent $u(i,t)$ as a weighted sum of shifted and dilated box functions $b(t)$,

$$u(i,t) = \sum_k c_k(i) b(t/a_k - t_k), \tag{5.2}$$

where $b(t) = \mathrm{rect}(t - 1/2)$, $c_k(i)$ is the amplitude of the k-th block, a_k is the block length, and $a_k t_k$ is the onset timing of activity. We define the innovation signal $u_s(i,t)$ as the derivative of the activity-inducing signal:

$$\begin{aligned} D\{u(i,\cdot)\}(t) &= \sum_k c'_k(i) \left(\delta(t - a_k t_k) - \delta(t - a_k(t_k + 1))\right), \\ &= \sum_{k'} c_{k'}(i) \delta(t - t_{k'}) = u_s(i,t), \end{aligned} \tag{5.3}$$

where D is the derivative operator and $\delta(t)$ is the Dirac-delta function. Ideally, for block-type signals, the innovation signal consists of spikes indicating the onsets and offsets of the blocks and zeros elsewhere. Note that, we have reparameterized the innovation signal to clearly reflect its sparse nature. Hence, $u_s(i,t)$ indicates the timing when the activity inducing signal's $u(i,t)$ amplitude changes. In Fig. 1, we illustrate the fMRI signal

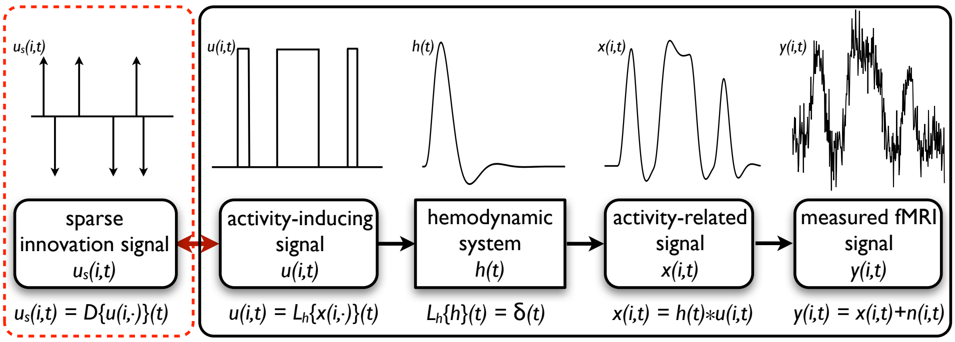


Figure 1: FMRI signal model. Assuming that the activity-inducing signal is block-type, its derivative is the innovation signal, which will be sparse. The activity-related signal can be obtained by convolving the activity-inducing signal with the impulse response of the hemodynamic system. The activity-related signal is then further corrupted with noise and signal artifacts, and finally sampled at the fMRI temporal resolution (TR).

model and its underlying sparse structure.

The HRF is adopted from the linear approximation of the balloon model, which was introduced by Khalidov et al. [101], and formulated explicitly in Chapter 4.1.2. The linear differential operator L_h in (4.1) inverting the hemodynamic system is represented by four zeros α_i $(i = 1, \ldots, 4)$ and one pole γ_1

$$L_h = \prod_{i=1}^{4} (D - \alpha_i I) \left(D - \gamma_1 I\right)^{-1},$$

and satisfies

$$L_h\{h\}(t) = \delta(t). \tag{5.4}$$

Then, we recover the activity-inducing signal by exploiting the differential operator that inverts the system as

$$L_h\{x(i, \cdot)\}(t) = u(i, t). \tag{5.5}$$

Given the link between innovation and activity-inducing signal, we also have $L\{x(i, \cdot)\}(t) = D\{u(i, \cdot)\}(t) = u_s(i, t)$, where the sparsifying operator $L = DL_h$ combines L_h with the regular derivative; i.e., adding one more zero $\alpha_5 = 0$ into its expression.

In practice, the activity-related signal $x(i, t)$ is corrupted by different sources of noise, such as non-neurophysiological contributions (e.g., aliased cardiac and respiratory fluctuations), movement, scanner drifts and thermal noise [83]. The fMRI signal $y(i, t)$ then becomes

$$y(i, t) = u(i, t) * h(t) + \sum_k \beta_k(i) n_k(t) + \epsilon(i, t), \tag{5.6}$$

where $n_k(t)$ represent known nuisance regressors (e.g., movement, low frequency drifts), β_k are associated weights, and $\epsilon(i, t)$ is AWGN with variance σ_i^2.

FMRI Data Representation

Before we move on with the problem formulation, it is necessary to explain the fMRI data representation. We represent the sampled and discretized full dataset as a matrix $\mathbf{Y}[i, n] = [y(i, t)]_{i, t_n}, n \in \mathbb{Z}$ of size $V \times N$, where V is the total number of voxels and N is the number of scans. Following the same convention in previous chapters, the discretised operators are denoted by Δ; e.g., Δ_D indicates the finite-difference for the derivative D, and so on. We also consider a predefined structural atlas that contains an anatomical parcellation of the brain; i.e., we assume M different parcels where R_k, $k = 1, \ldots, M$, are the sets of voxels for each region. Fig. 1 illustrates the fMRI dataset structure that is exploited further on.

5.2.2 Variational Formulation

We now aim at reconstructing activity-related signals from noisy fMRI measurements by casting a variational formulation [163,164]. Within the context

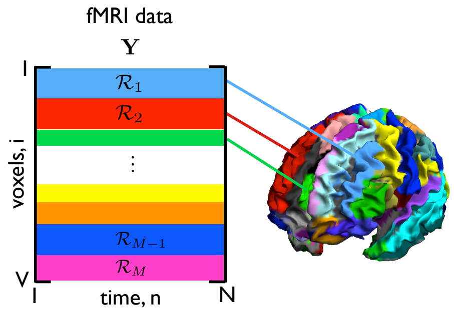


Figure 1: FMRI data representation. The dataset is represented as a matrix $\mathbf{Y}$ of size $V \times N$, where V is the total number of voxels and N is the number of scans. An anatomical atlas enables to define region of interests in the data and is exploited in our method.

of fMRI data processing, we introduce a novel spatiotemporal formulation based on the minimization of a cost function that is composed of a data-term, and two regularization terms. Specifically, our cost function reads

$$\tilde{\mathbf{X}} = \arg \min_{\mathbf{X} \in R^{V \times N}} \frac{1}{2} \|\mathbf{Y} - \mathbf{X}\|_F^2 + \mathcal{R}_T(\mathbf{X}) + \mathcal{R}_S(\mathbf{X}), \tag{5.7}$$

where $\mathbf{Y} \in R^{V \times N}$ is the noisy fMRI measurements, $\mathcal{R}_T$ and $\mathcal{R}_S$ are the temporal and spatial regularization terms, respectively. The optimal solution $\tilde{\mathbf{X}}$ is a compromise between the data fitness and the regularization penalties, which will now be further elaborated.

Temporal regularization

We have discussed in Chapter 3 how the generalized L-TV framework allows recovering sparse driving signals by inverting a linear system. Following the same idea, but this time on the hemodynamic system, we define the regularization term $\mathcal{R}_T(\mathbf{X})$ such that the sparsity of the innovation signal is emphasized when a differential operator $\Delta_L = \Delta_D \Delta_{L_h}$ is applied to the recovered activity-inducing signals $\mathbf{X}$:

$$\mathcal{R}_T(\mathbf{X}) = \sum_{i=1}^{V} \lambda_1(i) \, ||\Delta_L \{\mathbf{X}\} [i, \cdot]||_1 \, , \tag{5.8}$$

where

$$||\Delta_L \{\mathbf{X}[i, \cdot]\}||_1 = \sum_{n=1}^{N} |\Delta_L \{\mathbf{X}\} [i, n]| \, , \tag{5.9}$$

and $\lambda_1(i)$ is the regularization parameter for voxel i.

Spatial regularization

Since fMRI data has large amount of spatial correlations, we design a spatial regularization term $\mathcal{R}_S(\mathbf{X})$ that promotes coherent activity within the same region. To that end, we use a mixed $\ell_{(2,1)}$-norm to express spatially coherent (i.e., smooth) activity *inside* a region and possibly crisp changes in activity at regional borders [24, 165]:

$$\mathcal{R}_S(\mathbf{X}) = \sum_{n=1}^{N} \lambda_2(n) \left|\left|\Delta_{\text{Lap}} \left\{\mathbf{X}\right\} [\cdot, n]\right|\right|_{(2,1)}, \tag{5.10}$$

where

$$\left|\left|\Delta_{\text{Lap}} \left\{\mathbf{X}[\cdot, n]\right\}\right|\right|_{(2,1)} = \sum_{k=1}^{M} \sqrt{\sum_{i \in \mathrm{R}_k} \Delta_{\text{Lap}} \left\{\mathbf{X}\right\} [i, n]^2}, \tag{5.11}$$

and Δ_{Lap} is the second-order finite difference (Laplacian) operator and $\lambda_2(n)$ is the regularization parameter for each time point.

5.2.3 Optimization Algorithm

We now focus on the minimization problem given in (5.7). Since our cost function consists of a quadratic data fitting term and multiple sparsity-promoting regularization terms, we employ the generalized forward-backward algorithm [43]. The solution is obtained by incorporating the proximal maps of each sparse prior defined as

$$\tilde{\mathbf{X}}_T = \arg\min_{\mathbf{X}} \frac{1}{2} \left|\left|\mathbf{Y} - \mathbf{X}\right|\right|_F^2 + \sum_{i=1}^{V} \lambda_1(i) \left|\left|\Delta_L \left\{\mathbf{X}\right\} [i, \cdot]\right|\right|_1 = \text{prox}_{\mathcal{R}_\mathbf{T}}(\mathbf{Y}), \tag{5.12}$$

$$\tilde{\mathbf{X}}_S = \arg\min_{\mathbf{x}} \frac{1}{2} \left|\left|\mathbf{Y} - \mathbf{X}\right|\right|_F^2 + \sum_{n=1}^{N} \lambda_2(n) \left|\left|\Delta_{Lap} \left\{\mathbf{X}\right\} [\cdot, n]\right|\right|_{(2,1)} = \text{prox}_{\mathcal{R}_\mathbf{S}}(\mathbf{Y}). \tag{5.13}$$

Algorithm 6 summarizes the generalized forward-backward scheme adapted to the fMRI denoising problem. Similar to the generalized L-TV algorithm (Algorithm 5 in Section 3.2.4), each proximal map is solved exploiting the dual norm.

The good news is that we can calibrate the regularization parameter $\lambda_1(i)$ such that the residual noise converges to the pre-estimated noise level of the data fit. Then, at iteration n, we update the temporal regularization parameter $\lambda_1(i)$ as in [31]:

$$\lambda_1(i)^{[n+1]} = \frac{N\tilde{\sigma}(i)}{\|\mathbf{X}[i, \cdot] - \mathbf{Y}[i, \cdot]^{[n]}\|_2} \lambda_1(i)^{[n]},$$

where $\tilde{\sigma}(i)$ is the pre-estimated noise level of voxel i, $\lambda_1(i)^n$ and $\mathbf{X}[i,\cdot]^n$ are the regularization parameter and recovered activity-related signal of the ith voxel at nth iteration of the algorithm, respectively. We pre-estimate the noise level $\tilde{\sigma}$ using the median absolute deviation of fine-scale wavelet coefficients (Daubechies, order 3) [166].

Algorithm 6 Spatiotemporal Regularization for fMRI
$\tilde{\mathbf{X}} = \arg\min_{\mathbf{X}} \frac{1}{2}\|\mathbf{Y} - \mathbf{X}\|_F^2 + \mathcal{R}_{\mathbf{T}}(\mathbf{X}) + \mathcal{R}_{\mathbf{S}}(\mathbf{X})$

INPUTS: Noisy fMRI data $\mathbf{y}$,
1: Initialize: $l \leftarrow 1, \hat{\mathbf{X}}_T^0 = \mathbf{0}, \hat{\mathbf{X}}_S^0 = \mathbf{0}, \tilde{\mathbf{X}}^0 = \mathbf{0}$,
2: **repeat**
3: Solve for temporal prior: $\hat{\mathbf{X}}_T^{l+1} = \hat{\mathbf{X}}_T^l + \mathrm{prox}_{\mathcal{R}_T}(\tilde{\mathbf{X}}^l - \hat{\mathbf{X}}_T^l + \mathbf{Y}) - \tilde{\mathbf{X}}^l$
4: Solve for spatial prior: $\hat{\mathbf{X}}_S^{l+1} = \hat{\mathbf{X}}_S^l + \mathrm{prox}_{\mathcal{R}_S}(\tilde{\mathbf{X}}^l - \hat{\mathbf{X}}_S^l + \mathbf{Y}) - \tilde{\mathbf{X}}^l$
5: Update $\tilde{\mathbf{X}}^{l+1} = \hat{\mathbf{X}}_T^{l+1}/2 + \hat{\mathbf{X}}_S^{l+1}/2$
6: $l \leftarrow l + 1$
7: **until** convergence or number of maximum iterations are reached.

In Fig. 2, we schematically outline the TA algorithm for fMRI data analysis. The extracted time courses are fed into a two step forward-backward splitting algorithm where a joint solution is achieved. Temporal regularization works with each voxel time course since the differential operator Δ_L acts on the temporal domain, whereas spatial regularization term works with each fMRI volume exploting the Laplacian operator Δ_{Lap} in space. The algorithm solves for the denoised activity-related signals $\tilde{\mathbf{X}}$: We can access the activity-inducing signals as $\Delta_{L_h}\{\tilde{\mathbf{X}}\}$, which are driving the system as the neuronal-related activity at fMRI time scale.

5.3 Numerical Simulations

In this section, we present the results of TA applied on a 3D synthetic phantom. We elaborated extensively TA for various settings, including temporal and spatial mismatches. For this purpose, we created a software phantom with $10 \times 10 \times 10$ voxels divided into four regions. The activity-inducing signal was fixed within a region, but different across regions. Two regions had spike-like activity-inducing signals: Region 1 had spike trains with gradually increasing inter stimulus interval (ISI) from 1 to 12 sec.; Region 2 had short events with duration uniformly distributed between $[1, 2]$ sec. The other two regions had longer block-like activity (duration uniformly distributed between $[1, \ldots, 15]$ sec.). The onset timings of the events had uniform distribution such that 12 and 6 events on average were generated in regions with spikes and blocks, respectively. A very short event was included into region 4 to test TA's robustness for short events in the middle of sustained

$$\tilde{\mathbf{X}} = \arg\min_{\mathbf{X}} \frac{1}{2}\|\mathbf{Y}-\mathbf{X}\|_F^2 + \mathcal{R}_T(\mathbf{X}) + \mathcal{R}_S(\mathbf{X})$$

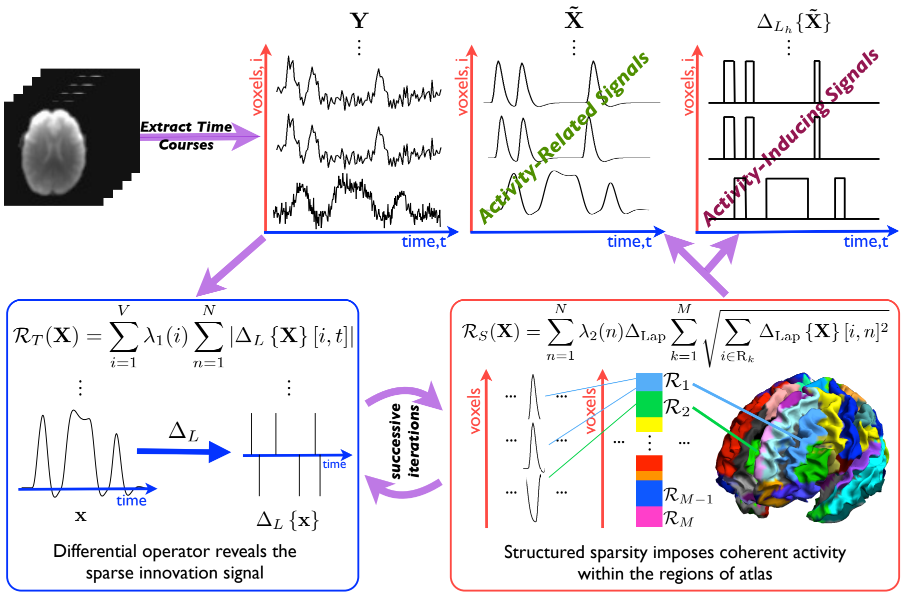


Figure 2: Flowchart of TA. Successive regularization in temporal and spatial domains are applied to the noisy BOLD signals. The algorithm alternates between temporal regularization (blue window) and spatial regularization (red window) until the convergence of activity-related signals. Finally, we derive activity-inducing signals which reveal the neuronal-related activity.

events. All activity-inducing signals were sampled on a grid with temporal resolution (TR) of 1 s and had 200 timepoints. The activity-induced signals were then convolved with the HRF and corrupted with AWGN such that SNR was 1 dB. The phantom is depicted in Fig. 5.4 with the associated time courses for each region.

5.3.1 Matched Model

TA analysis was first performed for the perfect setting; i.e., the temporal differential operator was matched with the generative HRF and the spatial regularization exploited the same regions as the phantom itself. In Fig. 5.5 (a) and (b), we show the activity-related and activity-induced signals for randomly selected voxels in four regions, respectively. The recovered activity-inducing signals match very closely with the ground-truth activity with no prior information on the timing or duration of the simulated events. In Fig. 5.5 (first row), we observe for Region 1 that TA can resolve for events with ISI down to 2 TRs. Our model is able to successfully recover differ-

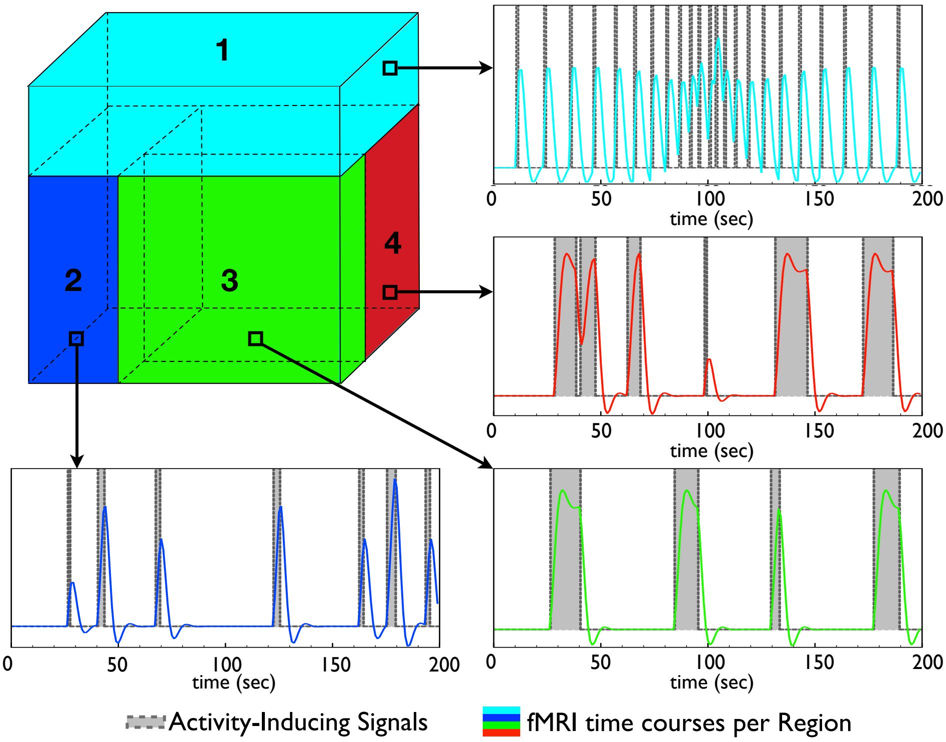


Figure 5.4: The phantom contains 4 regions in a cube of $10 \times 10 \times 10$ voxels. The first region (cyan, 300 voxels) was simulated as spike train with gradually increasing ISI from 1 sec. to 12 sec. and the second region (blue, 210 voxels) was simulated with random events with uniform duration in [1,2] sec. The third region (green, 245 voxels) and the fourth region (red, 245 voxels) were simulated with random events with uniform duration in [1,15] sec. The time resolution was chosen as TR=1 sec. The activity-inducing signals (in grey) were convolved with HRF to obtain the BOLD activity for each region. Each voxel time series was then corrupted with AWGN such that voxel time series had a resulting SNR of 1 dB.

ent types of activity-inducing signals; i.e., short spike-like and long block-like stimuli, especially the short event in Region 4 is well detected with a slightly lower amplitude.

We also analyzed the evolution of the total cost function minimized by the generalized forward-backward algorithm (see Fig. 5.6(a)). At each outer iteration, we computed the proximal map of the temporal and spatial priors as described in Algorithm 6 steps 3-4. The cost functions of these regularizations are plotted in Fig. 5.6 (b)-(c). The total cost decreased as expected and the inner regularizations also converged to a proxy solution at each outer iteration. In particular, we plot the cost of the proximal maps that solve $\mathrm{prox}_{\mathcal{R}_T}(\tilde{\mathbf{X}}^l - \hat{\mathbf{X}}_T^l + \mathbf{Y})$ and $\mathrm{prox}_{\mathcal{R}_S}(\tilde{\mathbf{X}}^l - \hat{\mathbf{X}}_S^l + \mathbf{Y})$ for temporal and spatial regularizations, respectively. We remark that the proximal maps are computed using variants of generalized L-TV regularization algorithms in Algorithm 5, which is globally convergent. By that means, the algorithms converge to a minimizer regardless of the initial condition, that was assigned as zero in our case (the peaks in Figs. 5.6(b)-(c)). We also note that, the algorithm does not have to be very precise in the early iterations, a coarse reconstruction is sufficient; therefore, we gradually increase the maximum number of inner iterations in temporal regularization to reduce the total computational cost. It is known that such an iteration strategy achieves better convergence performance [167].

5.3.2 Effect of Spatial Regularization

Our spatial regularization opts for smooth activitation patterns in the same region. Here, we illustrate the contribution of the spatial prior. In this regard, we compared three methods:

1. Full TA;

$$\tilde{\mathbf{X}} = \arg\min_{\mathbf{X}} \frac{1}{2}\|\mathbf{Y} - \mathbf{X}\|_F^2 + \mathcal{R}_T(\mathbf{X}) + \mathcal{R}_S(\mathbf{X}),$$

2. Regularization with the temporal prior only; this method basically simulates the state-of-the-art fMRI temporal deconvolution models

$$\tilde{\mathbf{X}} = \arg\min_{\mathbf{X}} \frac{1}{2}\|\mathbf{Y} - \mathbf{X}\|_F^2 + \mathcal{R}_T(\mathbf{X}),$$

3. Regularization with the temporal and a spatial smoothing prior lacking the atlas; i.e., Tikhonov regularization;

$$\tilde{\mathbf{X}} = \arg\min_{\mathbf{X}} \frac{1}{2}\|\mathbf{Y} - \mathbf{X}\|_F^2 + \mathcal{R}_T(\mathbf{X}) + \mathcal{R}_{S_2}(\mathbf{X}),$$

   where

$$\mathcal{R}_{S_2}(\mathbf{X}) = \sum_{n=1}^{N} \lambda_2(n) \left|\left|\Delta_{\mathrm{Lap}}\left\{\mathbf{X}\right\}[\cdot, n]\right|\right|_2 .$$

Fig. 5.7 shows all reconstructed activity-related signals per region. In Fig. 5.7, (a) depicts the TA reconstructed signals, (b) shows the reconstructed signals exploiting only the temporal prior, and finally (c) depicts the reconstructed signals with a global smoothing constraint (i.e., Tikhonov regularization) in the spatial domain. Despite the relatively high noise level

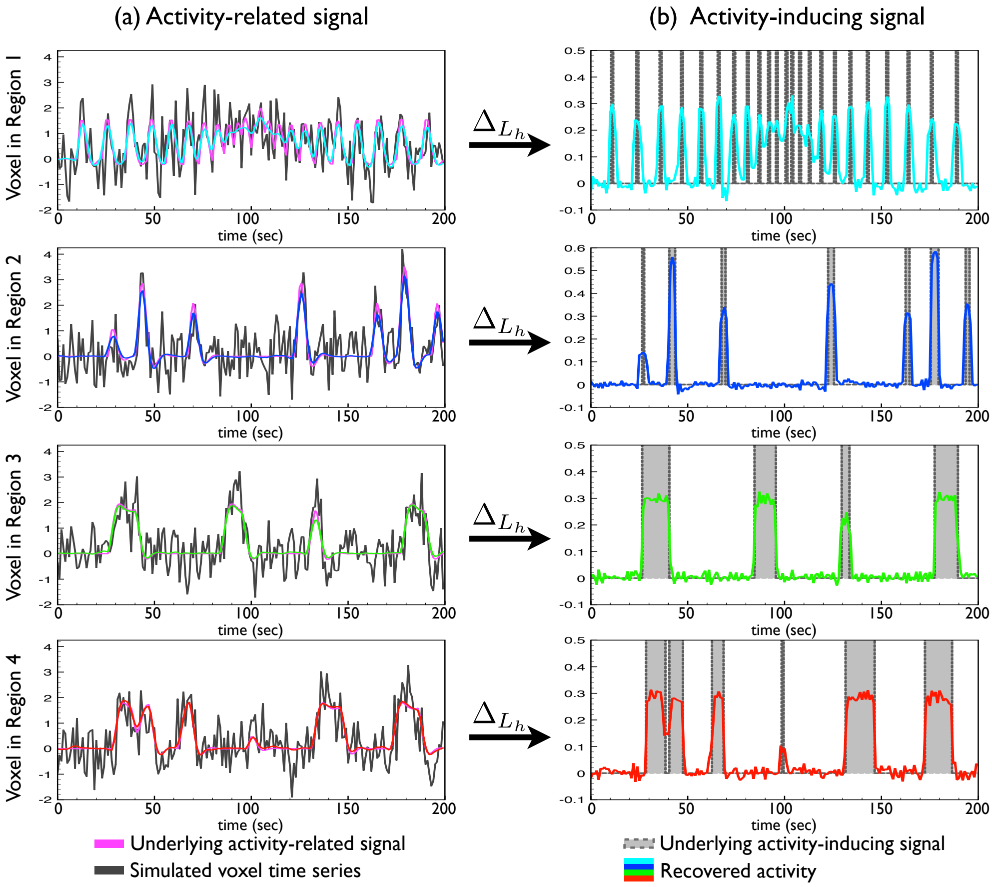


Figure 5.5: Results for the software phantom. The left column (a) shows simulated noisy data (black), underlying BOLD signals (magenta), and recovered activity-related signals of a random voxel in each region (cyan, blue, red, green, respectively). The right column (b) shows the underlying activity-inducing signal (grey) and the associated recovered activity signals.

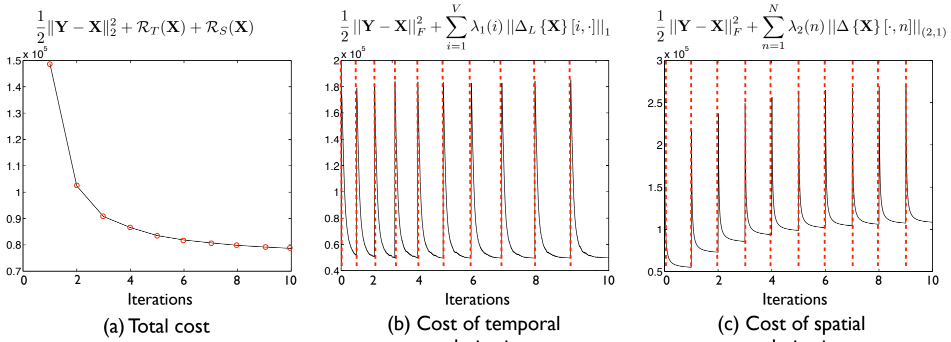


Figure 5.6: The cost function of TA algorithm. The left figure (a) shows the total cost of generalized forward-backward algorithm per iteration. At each iteration two suboptimization problems are again solved iteratively; temporal and spatial regularizations. The corresponding inner costs are plotted in (b) and (c).

in the simulated time courses, TA-regularized signals have smallest variance in a specific region. When there is no spatial constraint each voxel is treated independently, thereby, the variations are substantially higher. We notice that Tikhonov regularization imposes global smoothness, which in turn brings on interferences with other regions; i.e., false activations or loss of the weak ones.

5.3.3 Temporal Model Mismatch

Under the linear time-invariant system assumption, the inter-subject or intra-subject HRF fluctuations should be directly captured by TA. Therefore, we tested TA for HRF variations by adopting different hemodynamic models for signal generation (canonical HRF) and analysis (balloon model).

We generated a synthetic time course (1dB SNR) and analyzed with both temporally matched model and mismatched model. Fig. 5.8 depicts the HRF variations (time-to-peak, amplitude, dispersion, and undershoot). Fig. 5.8(a) shows the recovered activity-inducing signal (blue) with the temporally matched model, and Fig. 5.8(b) illustrates the mismatched case; i.e., ground-truth was generated with canonical HRF and analyzed with balloon model. We observe a time shift of the activity-inducing signal due to the differences in the temporal characteristics. However, both models were able to reveal similar activation patterns.

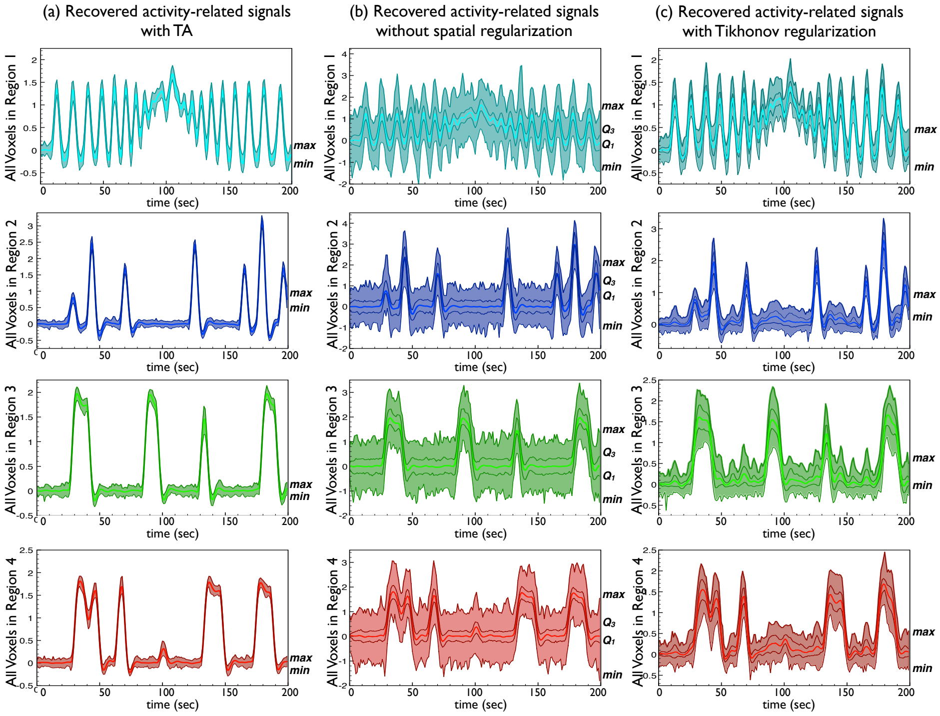


Figure 5.7: The effect of different spatial regularizations. The left column (a) shows all recovered activity-related signals obtained with TA analysis indicated by their mean, maximum, minimum per region. Small deviations within each region are observed. The middle column (b) shows all recovered activity-related signals without the spatial regularization. Maximum, minimum, first and third quartiles are indicated. The variation is considerably higher as each voxel is treated independently. The right column (c) shows activity-related signals produced by exploiting Tikhonov regularization. The interference among the time courses may lead to false activations.

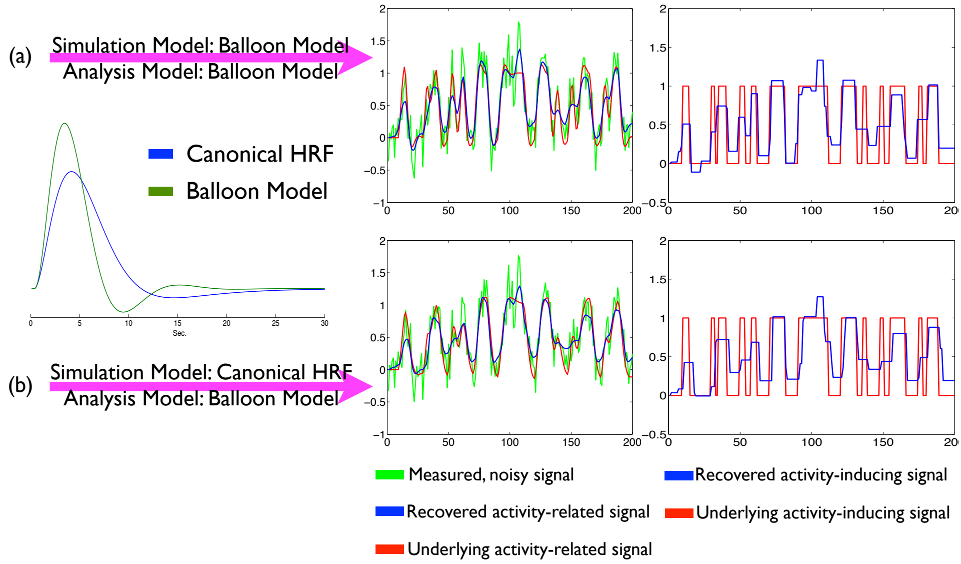


Figure 5.8: HRF variation. We generated a synthetic time course (SNR 1 dB) with two settings; (a) matched model, we simulate and analyze using the same HRF model and (b) mismatched model. Both models resulted in showing similar activation patterns, however, the mismatched model had an expected time shift in the underlying activations due to different HRF characteristics.

5.4 Model Extensions

5.4.1 Correlated Noise

The noise in fMRI data is known to exhibit temporal correlations ("serial correlations"), which are typically handled by using autoregressive models [89]. TA's cost function is optimal for uncorrelated noise, nevertheless, an autoregressive noise model can be easily integrated into the framework. The colored noise should be whitened based on an estimated covariance of the residuals, which then leads to a weighted ℓ_2-norm for the data-term of the cost function. Considering the fMRI generative model with a first order auto-regressive noise AR(1), we have

$$y(i,t) = x(i,t) + \epsilon(t), \quad \epsilon(t) = \rho\epsilon(t-1) + n(t),$$

where $n(t)$ is AWGN with $\mathcal{N}(0, \sigma_n)$ and $\epsilon(i)$ is AR(1) with covariance matrix $\mathbf{\Sigma}$. We can therefore whiten the noise and define the regularization in time as

$$\tilde{\mathbf{X}}_T = \arg\min_{\mathbf{X}} \frac{1}{2}(\mathbf{Y} - \mathbf{X})^T \Sigma^{-1}(\mathbf{Y} - \mathbf{X}) + \mathcal{R}_T(\mathbf{X}), \tag{5.14}$$

Consider $\mathbf{x}$ and $\mathbf{y}$ denote the activity-related and measured BOLD signals at the ith voxel. Then, given an estimate of the correlation (e.g., FSL-like estimation based on the correlation of residuals) we derive the following updates for each iteration l;

$$\mathbf{p}^l = P_B\left(\Delta_L\ \{\mathbf{y}\}/(\lambda_1(i)^l\, c) + (I - \Delta_L \mathbf{\Sigma} \Delta_L^T/c)\ \left\{\mathbf{v}^l\right\}\right), \text{(Algorithm 5, step 4)} \tag{5.15}$$

$$\mathbf{x}^l = \mathbf{y} - \lambda_1(i)^l\, \mathbf{\Sigma}\Delta_L^T\ \{\mathbf{p}\}, \text{(Algorithm 5, step 9).} \tag{5.16}$$

The above algorithm integrates the temporal regularization with the noise model, where only the data term is affected by the noise; the spatial regularization is not modified since the AR noise is in the temporal domain.

5.5 Discussion and Summary

In this chapter, we try to give an insight into the underlying principle of TA through performing experiments on a synthetic 3D phantom. The dataset consisted of four regions with distinct temporal characteristics; e.g., long blocks, short blocks, and fast stimuli (spike train with various ISI). Furthermore, we perturbed the perfectly matched generative and analysis models to study how TA deals with these discrepancies. We specifically highlight the flexibility of the HRF operator and the contribution of the spatial regularization. TA shows promising results to handle the synthetic data, which is encouraging to perform analysis of real experimental fMRI data. In the next chapter, we study two such fMRI datasets.

Chapter 6

Data Mining with Total Activation: Application to FMRI Data

In this chapter [1], we describe how TA proves to be a useful tool for various aspects of fMRI data. The primary goal of TA is to reveal the activity-inducing signals, which provide rich information on task-related as well as spontaneous brain activity, thereby, enabling further exploration that can not be performed with traditional fMRI data analysis methods. Here, we specifically evaluate TA in cognitive and clinical experimental conditions:

1. An event-related fMRI experiment with prolonged resting-state periods that are disturbed by unpredicted visual stimuli. We illustrate how TA captures the paradigm, without any prior knowledge on timing, and further recovers task-related as well as meaningful resting-state networks (RSN) that cannot be inferred from conventional analyses.
2. RS-fMRI data acquired from pharmacoresistant epilepsy patients for presurgical planning. We show that TA is able to locate the epileptogenic regions from simultaneous recording of fMRI and electroencephalography (EEG).

[1] This chapter is based partially on the publications:
F. I. Karahanoglu, C. Caballero-Gaudes, F. Lazeyras, and D. Van De Ville, "Total activation: fMRI deconvolution through spatio-temporal regularization", NeuroImage, vol. 73, pp. 121-134, June 2013 [162];
F. I. Karahanoglu, F. Grouiller, C. Caballero-Gaudes, M. Seeck, S. Vulliemoz, and D. Van De Ville, "Spatial Mapping of Interictal Epileptic Discharges in FMRI with Total Activation", in Proceedings of IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2013, in press [168].

6.1 Potential Applications

We will show that TA can be useful for elucidating brain function and dysfunction. Among fMRI data analysis tools, TA explicitly aims at retrieving the patterns of cortical activity that underlie the BOLD signal, thereby, enable the exploration of many remaining and intriguing questions in neuroscience to improve the understanding of brain processes. However, fMRI data itself is confounded by various factors, which make the analysis challenging. Due to the low SNR of the fMRI signal, data analysis is often carried out with appropriate assumptions or exploiting prior information about the experimental conditions. To increase the reliability of the statistical analysis in terms of specificity and sensitivity, task-related studies rely on the presentation of multiple trials of each experimental condition, which are repeated within and across sessions. The traditional analysis techniques often overlook possible non-stationary patterns in the BOLD signal, such as spontaneous or transient activities, or learning, habituation, anticipation processes, delays in responses, mental chronometry, pharmacological effects, etc [79, 84, 85, 169–171]. Specifically, the timing of the events cannot be anticipated or modelled in advance, possibly due to neurological and psychiatric disorders (e.g., interictal epileptic discharges and schizophrenic hallucinations) or absence of consciousness (e.g., ongoing activations during sleep) [172, 173]. In these cases, TA constitutes as a promising tool for analyzing these spontaneous and highly non-stationary dynamics of the fMRI signal.

TA is designed primarily as a denoising problem; i.e., the activity-related signals are "clean" fMRI signal satisfying spatial and temporal priors. Consequently, another potential use of TA is for preprocessing the fMRI BOLD signals prior to further analysis with other methods. Lately, there has been a debate on whether deconvolution is necessary for functional and effective connectivity analyses [145,174–178]. In one of the recent works, the Granger causality and dynamic causal modelling are compared for effective connectivity analysis and deconvolution has been suggested as a necessary step to obtain adequate inferences [178]. Even though limited number of works exist in the literature to draw any conclusions, deconvolution might potentially constitute a necessary step for fMRI analysis.

6.2 Event-Related Visual Experiment

We evaluated our method using fMRI data acquired from three healthy subjects engaged in an event-related experiment. Subjects were presented 10 (unexpected) visual stimuli of 8Hz flickering checkerboard of duration 1 sec. with onsets following a uniform distribution (see Fig. 6.1 for the

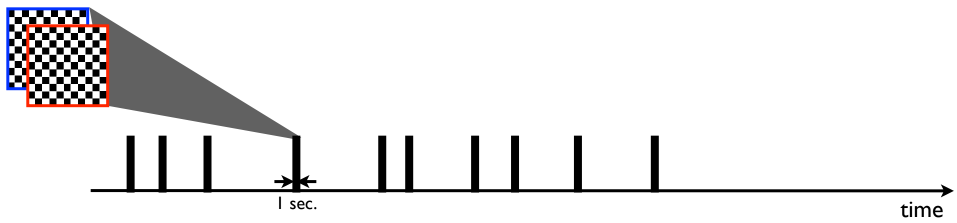


Figure 6.1: Visual stimuli. The paradigm consists of 10 visual stimuli of 8Hz flickering checkerboard of duration 1sec. The experiment lasted for 5-6 minutes with TR=2 sec.

paradigm). When no visual stimuli were present, subjects were instructed to maintain visual fixation on a cross in the center of the screen.

6.2.1 Acquisiton and Preprocessing

The experiment was conducted in a Siemens TIM Trio 3T MR scanner with a 32-channel head coil. The fMRI data comprised $N = 160$ (subjects 1 and 3) and $N = 190$ (subject 2) T_2^*-weighted gradient echo-planar volumes (TR/TE/FA=2s/30 ms/85^o, voxel size: $3.25 \times 3.25 \times 3.5\text{mm}^3$, matrix=$64 \times 64$). A T1-weighted MPRAGE anatomical image was also acquired during the MR session (192 slices, TR/TE/FA: 1.9s/2.32ms/9^o, voxel size: $0.45 \times 0.45 \times 0.9\text{mm}^3$, matrix = 512×512) to aid anatomical localization of the functional maps.

The preprocessing steps included realignment of the datasets to the first scan to correct for head motion of each subject and then spatial smoothing with a Gaussian smoother (FWHM = 5mm). The spatial smoothing was not an essential step since TA also included spatial regularization. However, the temporal regularization parameter was tuned for each voxel with respect to the (estimated) noise level. Therefore, spatial smoothing enabled reliable estimattion (less variations) of the noise level. Both steps were performed in the functional space of the subjects using SPM8 (FIL, UCL, UK). The anatomical automatic labelling (AAL) atlas, which is an automated parcellation of a single subject's structural MR image consisting 90 regions without the cerebellum, was mapped onto each subject's functional space using the IBASPM toolbox [179, 180]. The voxels' time courses labelled within the atlas were detrended using a first-degree polynomial (i.e., linear trend) and slow oscillations (i.e., DCT basis function up to cut-off frequency of 1/250Hz), and finally scaled to have unit variance.

Algorithm Setting

Next, datasets were analyzed with the TA algorithm using the following specifications. The temporal regularization parameter was adjusted automatically, as discussed in Chapter 5.2.3, for each voxel within the inner temporal regularization problem such that the residual noise level converges to the pre-estimated noise level [31]. Spatial regularization parameter was empirically selected to be 5, which compensate well between temporal and spatial priors. Increasing λ_2 forces the smoothness leaving no room for local differences, especially for large brain regions, whereas small λ_2 results in high variance of the activity-related signal in the regions. The algorithm is implemented in Matlab 7.14 (Mathworks, Natick, MA) on a 64-bit, 4-core computer with 16 GB RAM, operating Linux. The total allocated time was around 5-7 hours.

6.2.2 Network Analysis

After applying TA, we obtain three spatiotemporal datasets per subject: (1) the innovation signals $\mathbf{U_s}$; (2) the activity-inducing signal $\mathbf{U}$; (3) the activity-related signal $\mathbf{X}$. The innovation signal is the driver of the others, which can be derived through convolution. To summarize the rich amount of information available in these datasets, we computed the average of the activity-inducing signals within each anatomical region, and then obtained the Spearman correlation matrix between the averaged timecourses. Correlations were Fisher z-transformed, averaged over three subjects and fed into a Ward's hierarchical clustering algorithm, already implemented in Matlab (Mathworks, Natick, MA) to reveal the network structures in activity-inducing signals. We selected two different levels to cut the dendrogram in order to show the evolution of clusters with respect to the inconsistency criterion that measures the deviation in each cluster. We extracted functionally distinct clusters at coarse (high) and detailed (low) levels. In other words, going down from the highest level in the dendrogram (whole brain) the consistency in the hierarchy gradually increments until a first group of clusters is defined (high-level), further increasing the consistency splits the clusters into subclusters which are meaningful segregations (low-level). Fig. 6.2 depicts the average correlation matrix and dendrogram as a result of hierarchical clustering. At the high-level hierarchy, the brain was segregated into 9 global clusters (represented in different colors in the dendrogram); at the low-level hierarchy, 17 local networks (subclusters pinned from (1a) to (9) in the dendrogram) were revealed. Fig. 6.3 illustrates the high-level networks overlaid on the anatomical atlas. The extended anatomical descriptions in each (sub)cluster are listed in Table 1. We detail these clusters according to the order of the dendrogram.

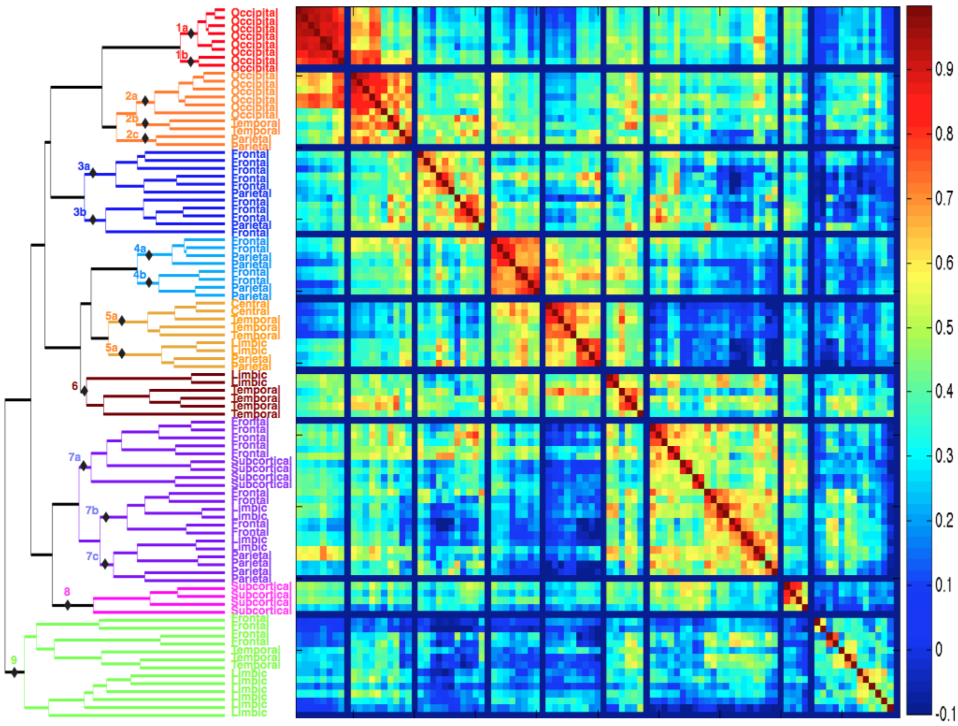


Figure 6.2: Correlation matrix and corresponding clusters for TA activity-inducing signal. The dendrogram that reflects the hierarchical organization is shown on the left. Each color is described with the regions correspond to a different cluster in high level clusters (total 9 clusters) which is evaluated via inconsistency measure. Note that low level clusters (marked with black pins in the dendrogram from (1a) to 9) subdivides the clusters resulting 17 clusters. The anatomical descriptions in the clusters are detailed in Table 1.

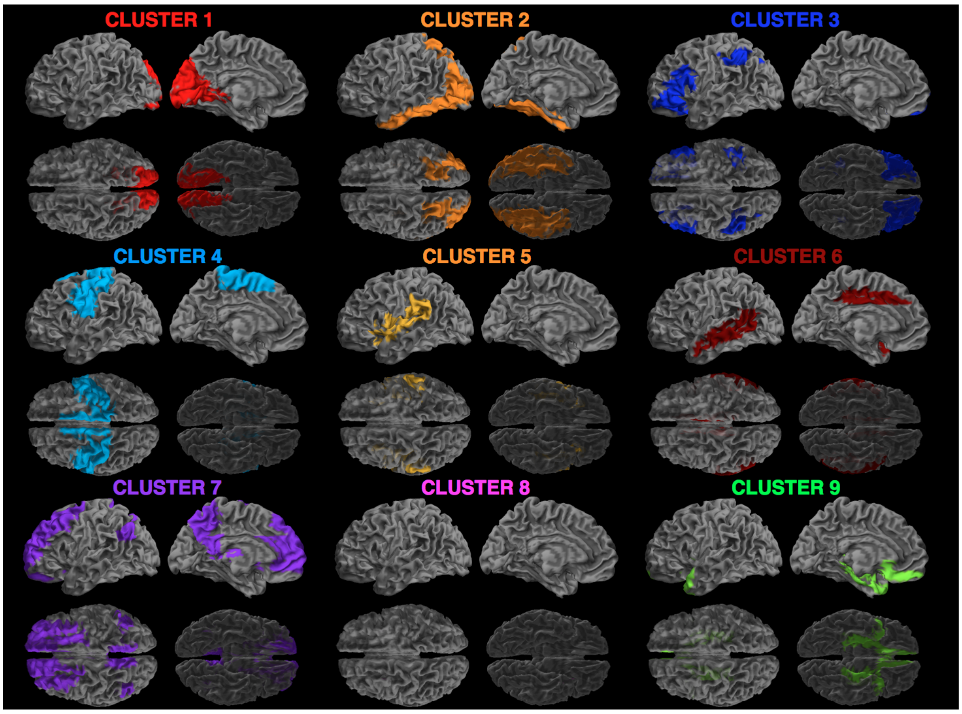


Figure 6.3: Brain maps for the 9 high-level hierarchy clusters viewed from sagittal left (top left), sagittal cross-section in the middle (top right), top view (bottom left) and bottom view (bottom right). The regions are generated using anatomical atlas in MNI space corresponding to the anatomical descriptions in Table 1. We recover the activity-related networks; i.e., primary and late visual networks in clusters 1 and 2, respectively. Additionally, the fronto-parietal network (cluster 3), motor and somatosensory regions (cluster 4) and auditory network (cluster 5) as well as the default-mode network (cluster 7), subcorticals (cluster 8) and limbic system (cluster 9) are observed. The clusters are nicely organized bilaterally.

CLUSTER	LOBE	ANATOMICAL DESCRIPTION
1a	Occipital	Calcarine Fissure Left
	Occipital	Calcarine Fissure Right
	Occipital	Lingual Gyrus Left
	Occipital	Lingual Gyrus Right
	Occipital	Cuneus Left
	Occipital	Cuneus Right
1b	Occipital	Superior Occipital Gyrus Left
	Occipital	Superior Occipital Gyrus Right
2a	Occipital	Middle Occipital Gyrus Left
	Occipital	Inferior Occipital Gyrus Left
	Occipital	Middle OccipitalGyrus Right
	Occipital	Fusiform Gyrus Left
	Occipital	Fusiform Gyrus Right
	Occipital	Inferior Occipital Gyrus Right
2b	Temporal	Inferior Temporal Gyrus Left
	Temporal	Inferior Temporal Gyrus Right
2c	Parietal	Superior Parietal Gyrus Left
	Parietal	Superior Parietal Gyrus Right
3a	Frontal	Superior Frontal Gyrus (Orbital) Right
	Frontal	Inferior Frontal Gyrus (Orbital) Right
	Frontal	Middle Frontal Gyrus (Orbital) Right
	Frontal	Inferior Frontal Gyrus (Opercular) Right
	Frontal	Inferior Frontal Gyrus (Triangular) Right
	Parietal	Inferior Parietal Gyrus Right
3b	Frontal	Middle Frontal Gyrus (Orbital) Left
	Frontal	Inferior Frontal Gyrus (Opercular) Left
	Frontal	Inferior Frontal Gyrus (Triangular) Left
	Parietal	Inferior Parietal Gyrus Left
	Frontal	Inferior Frontal Gyrus (Orbital) Left
4a	Frontal	Precentral Gyrus Left
	Frontal	Precentral Gyrus Right
	Parietal	Postcentral Gyrus Left
	Parietal	Postcentral Gyrus Right
4b	Frontal	Supplementary Motor Area Left
	Frontal	Supplementary Motor Area Right
	Parietal	Paracentral Lobule Left
	Parietal	Paracentral Lobule Right
5a	Central	Rolandic Operculum Left
	Central	Rolandic Operculum Right
	Temporal	Superior Temporal Gyrus Left
	Temporal	Heschl Gyrus Right
	Temporal	Heschl Gyrus Left
5b	Limbic	Insula Left
	Limbic	Insula Right
	Parietal	SupraMarginal Gyrus Left
	Parietal	SupraMarginal Gyrus Right
6	Limbic	Medial Cingulate Cortex Left
	Limbic	Medial Cingulate Cortex Right
	Temporal	Superior Temporal Gyrus Right
	Temporal	Middle Temporal Gyrus Left
	Temporal	Middle Temporal Gyrus Right
	Temporal	Temporal Pole (Superior) Right
7a	Frontal	Superior Frontal Gyrus (Orbital) Left
	Frontal	Middle Frontal Gyrus Left
	Frontal	Superior Frontal Gyrus (Dorsolateral) Left
	Frontal	Superior Frontal Gyrus (Dorsolateral) Right
	Frontal	Middle Frontal Gyrus Right
	Subcortical	Caudate Nucleus Left
	Subcortical	Caudate Nucleus Right
	Subcortical	Thalamus Left
	Subcortical	Thalamus Right
7b	Frontal	Superior Frontal Gyrus (Medial) Left
	Frontal	Superior Frontal Gyrus (Medial) Right
	Limbic	Anterior Cingulate Cortex Left
	Limbic	Anterior Cingulate Cortex Right
	Frontal	Superior Frontal Gyrus (Medial-Orbital) Left
	Frontal	Superior Frontal Gyrus (Medial-Orbital) Right
7c	Limbic	Posterior Cingulate Cortex Left
	Limbic	Posterior Cingulate Cortex Right
	Parietal	Precuneus Left
	Parietal	Precuneus Right
	Parietal	Angular Gyrus Left
	Parietal	Angular Gyrus Right
8	Subcortical	Putamen Left
	Subcortical	Pallidum Left
	Subcortical	Putamen Right
	Subcortical	Pallidum Right
9	Frontal	Olfactory Cortex Left
	Frontal	Olfactory Cortex Right
	Frontal	Gyrus Rectus Left
	Frontal	Gyrus Rectus Right
	Temporal	Temporal Pole (Superior) Left
	Temporal	Temporal Pole (Middle) Left
	Temporal	Temporal Pole (Middle) Right
	Limbic	Hippocampus Left
	Limbic	ParaHippocampal Gyrus Left
	Limbic	Hippocampus Right
	Limbic	ParaHippocampal Gyrus Right
	Limbic	Amygdala Left
	Limbic	Amygdala Right

Table 1: The list of regions in the clusters. The clustering algorithm delineates 9 and 17 clusters in the high and low-level hierarchies (also presented in dendrogram in Figure 6.2). The first two clusters correspond to the visual networks. Note that cluster 3 (fronto-parietal network) is subdivided into its right (3a) and left (3b) compartments in the higher hierarchy. Likewise, cluster 7 (default mode) is divided into its anterior (7a, 7b) and posterior (7c) components.

The visual network made up the first and second cluster, which was expected due to the stimulation and its strong coherence in resting-state. Cluster 1 contained primary visual areas such as calcarine fissure, lingual gyrus and cuneus. Cluster 2 included higher level visual areas extending towards ventral and dorsal visual pathways, inferior temporal gyrus and superior parietal lobule, which were subclusters 2b and 2c, respectively. In Fig. 6.7 (bottom right), the region-averaged activity-inducing signal in the visual network confirmed that the timing of the visual stimuli (red bars) was well recovered without any prior knowledge. Cluster 3 revealed a fronto-parietal network extending bilateral middle frontal gyrus, inferior frontal gyrus and inferior parietal lobule, which mimics the dorsal attention network [181] and involves in attentional mechanisms, especially for "salient and unattended events" [182]. Subclusters 3a and 3b represent the right and left lateralized fronto-parietal regions similar to [183, 184], respectively.

Cluster 4 revealed sensory-motor areas including primary motor cortex, primary somatosensory cortex as subcluster 4a, and supplementary motor areas as subcluster 4b. Cluster 5 maps the auditory network where speech and language processing occur, including the Heschl gyrus, superior temporal gyrus (Wernicke's area) and inferior frontal gyrus. Cluster 6 involved bilateral midcingulate cortex, middle temporal gyrus as well as the right superior temporal gyrus. Cluster 7 consisted of superior and middle frontal gyrus, anterior-posterior cingulate cortex (PCC) representing the default-mode network (DMN) including thalamus [185]. The hierarchical clustering suggested that cluster 7 was segregated into its anterior (7a, 7b) and posterior (7c) components, which are known to be part of saliency and executive control networks [181, 186], respectively. Similar subdivisions of the DMN have also been reported recently using real-time fMRI neurofeedback [187]. Subcortical regions, putamen and pallidum, were engaged in cluster 8 bilaterally. Cluster 9 involved bilateral limbic regions, parahippocampal gyrus, hippocampus and amygdala, as well as olfactory bulb, gyrus rectus and temporal poles.

The network analysis was performed over region averaged time series instead of individual voxel time courses for two reasons:

1. Coping with high dimensional data. Images with a whole-brain coverage comprise about 10'000-20'000 voxels, computing the correlation matrix voxel-wise and performing hierarchical clustering is not computationally feasible and roboust.

2. Presenting results at the group-level. TA was exploited in the subjects' native functional spaces, that is to say, the images were not normalized to a common space prior to TA analysis. Hence, incorporating an atlas enabled us to extract the time courses according to a common anatomical prior; i.e., Region 1 always referred to the same cortical area.

TA's spatial regularizer incorporates the anatomical atlas, which is a large-scale cortical parcellation of the brain with 90—rather course—distinct regions. In order to investigate the spatial segregation and elaborate how TA handles the spatial smoothing inside each brain region, we further exploited hierarchical clustering at the voxel level in a pivot region, insula, which is known to split functionally into anterior-posterior parts [188]. Fig. 6.4 depicts the hierarchical clustering, the dendrograms, and the color-coded maps of the right and left insula of Subject 2. We observe that TA revealed the expected the anterior-posterior partition inside insula.

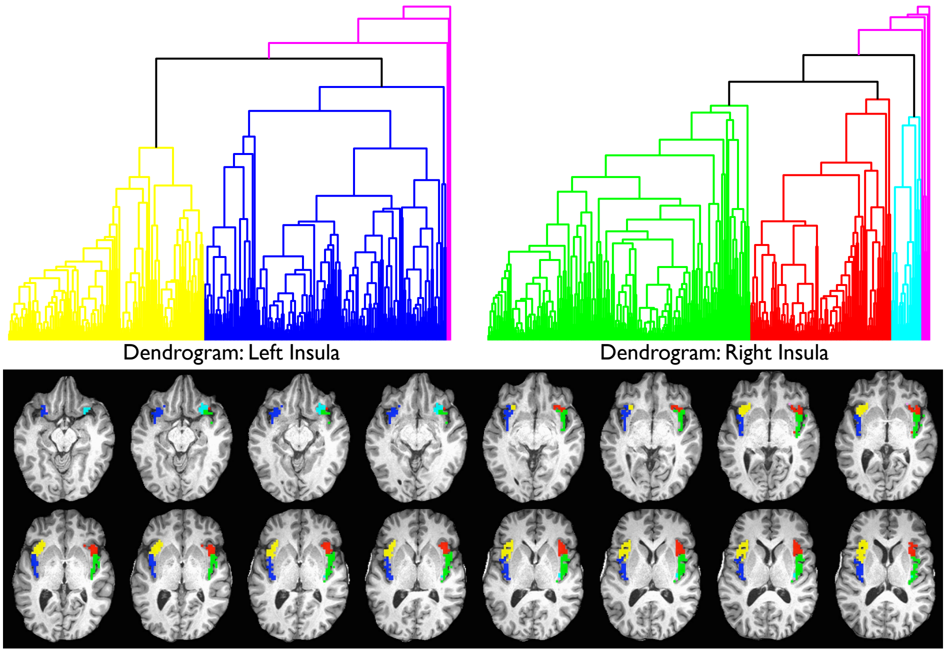


Figure 6.4: Hierarchical clustering using the voxel-wise correlations in insula. The hierarchical clustering inside the lateral regions provide two main anterior and posterior partitions. The result suggests that spatial smoothing in a region still allows for further functional segregation inside each parcel.

Network Analysis without TA

In order to illustrate the advantage of TA-processed data for capturing consistent networks, we performed the same network analysis on the preprocessed data without applying TA. The average correlation matrix and corresponding dendrogram following the hierarchical analysis are depicted in Fig. 6.5. Furthermore, the recovered networks are illustrated in 6.6 and corresponding regions are listed in Table 6.2. We find that visual, motor and auditory networks were also identified, however, they were given different preferences in dendrogram (auditory network was cluster 8 instead on 5, motor is 7 instead of 4). Moreover, the most prominent right-left lateralized fronto-parietal network and anterior-posterior segregation of the default-mode network was lost in the hierarchy.

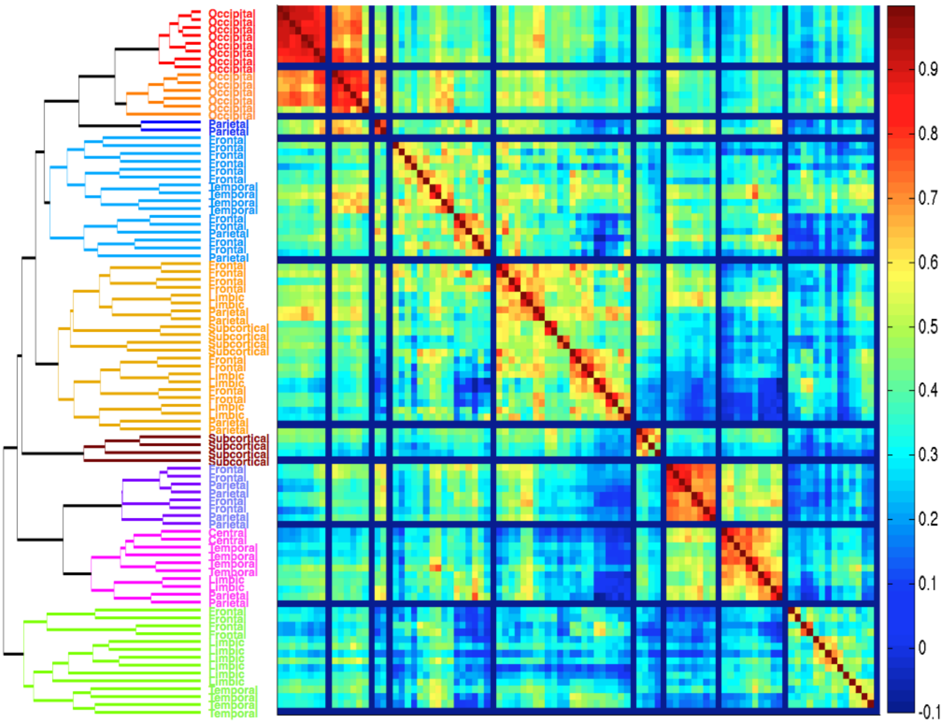


Figure 6.5: The average correlation matrix, dendrogram and corresponding clusters are depicted when only detrended data was analyzed (without TA). The correlation matrix still detected the boosted visual network, similarly motor and auditory networks possibly due to dominant activations in these regions. However, the hierarchy was lacking prominent organized networks detected by TA, such as left-right lateralization of fronto-parietal network.

6.2.3 Time Series Analysis

Having identified temporally coherent networks through clustering of the activity-inducing signals, now we can try to represent the dynamics for brain regions revealed by TA. Fig. 6.7 depicts the average activity-inducing signals rearranged according to the clusters. While the stimulus timings were well detected mainly in the clusters corresponding to visual areas, we observed spontaneous activity in the visual network which did not correspond to visual stimuli (e.g., subject 2, cluster 1, around 300 sec). Fig. 6.8 shows the dynamic activity-inducing maps of subject 2. Two time courses were picked randomly from cuneus and PCC in order to track the temporal evolution of the task-related and spontaneous events. The positive and negative activations in PCC lagging the stimulus reflected the alternating structure of functional reorganization in the brain.

Finally, since the temporal prior of TA favors block-like activity-inducing

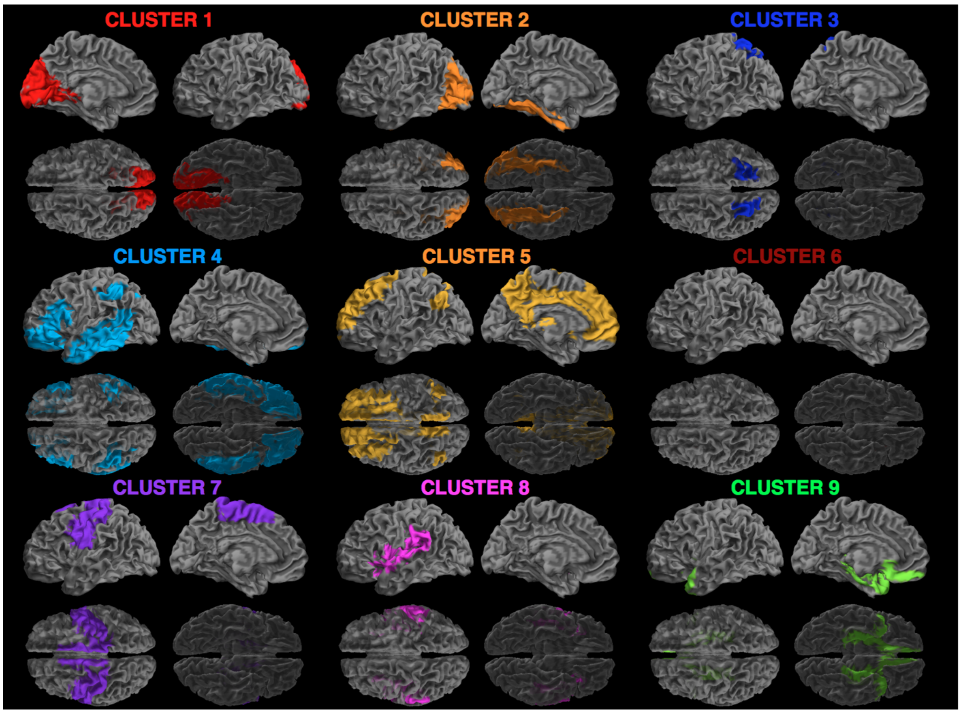


Figure 6.6: The recovered networks from clustering analysis of the data without applying TA. The primary visual network was perfectly recovered in cluster 1, however, contrary to TA-recovered clusters, secondary visual network was segregated into two; cluster 2 included mainly the ventral visual pathway and cluster 3 included regions of dorsal pathways. Cluster 4 shows the fronto-parietal network jointly with temporal gyrus; the left-right lateralization was lost. Cluster 5 was the default-mode network together with middle cingulate cortex. Despite being in a different order in the dendrogram compared to TA, motor and auditory networks were reconstructed successfully in clusters 7 and 8, respectively.

signals, we computed the average block-length, which was not set a priori in our analysis, per region as the 4th quartile of the activity duration, see Fig. 6.9. From the duration map, we clearly observe that the regions in the visual cortex had shorter duration whereas fronto-parietal regions had relatively the longest duration.

6.2.4 Discussion and Summary

We illustrated the application of TA on fMRI dataset acquired during an event-related visual stimuli. The results proved that both block-type and spike-type activity could be recovered successfully without prior knowledge

CLUSTER	LOBE	REGION
1	Occipital	Calcarine Fissure Left
	Occipital	Calcarine Fissure Right
	Occipital	Lingual Gyrus Left
	Occipital	Lingual Gyrus Right
	Occipital	Cuneus Left
	Occipital	Cuneus Right
	Occipital	Superior Occipital Gyrus Left
	Occipital	Superior Occipital Gyrus Right
2	Occipital	Middle Occipital Gyrus Left
	Occipital	Middle Occipital Gyrus Right
	Occipital	Inferior Occipital Gyrus Left
	Occipital	Fusiform Gyrus Left
	Occipital	Fusiform Gyrus Right
	Occipital	Inferior Occipital Gyrus Right
3	Parietal	Superior Parietal Gyrus Left
	Parietal	Superior Parietal Gyrus Right
4a	Frontal	Superior Frontal Gyrus (Orbital) Left
	Frontal	Middle Frontal Gyrus (Orbital) Left
	Frontal	Superior Frontal Gyrus (Orbital) Right
	Frontal	Middle Frontal Gyrus (Orbital) Right
	Frontal	Inferior Frontal Gyrus (Orbital) Left
	Frontal	Inferior Frontal Gyrus (Orbital) Right
	Temporal	Middle Temporal Gyrus Left
	Temporal	Middle Temporal Gyrus Right
	Temporal	Inferior Temporal Gyrus Left
	Temporal	Inferior Temporal Gyrus Right
4b	Frontal	Inferior Frontal Gyrus (Opercular) Left
	Frontal	Inferior Frontal Gyrus (Triangular) Left
	Parietal	Inferior Parietal Gyrus Left
	Frontal	Inferior Frontal Gyrus (Opercular) Right
	Frontal	Inferior Frontal Gyrus (Triangular) Right
	Parietal	Inferior Parietal Gyrus Right
5a	Frontal	Superior Frontal Gyrus Left
	Frontal	Middle Frontal Gyrus Left
	Frontal	Superior Frontal Gyrus Right
	Frontal	Middle Frontal Gyrus Right
	Limbic	Middle Cingulate Cortex Left
	Limbic	Middle Cingulate Cortex Right
	Parietal	Precuneus Left
	Parietal	Precuneus Right
	Subcortical	Caudate Nucleus Left
	Subcortical	Caudate Nucleus Right
	Subcortical	Thalamus Left
	Subcortical	Thalamus Right
5b	Frontal	Superior Frontal Gyrus(Medial) Left
	Frontal	Superior Frontal Gyrus (Medial) Right
	Limbic	Anterior Cingulate Cortex Left
	Limbic	Anterior Cingulate Cortex Right
	Frontal	Superior Frontal Gyrus (Medial-Orbital) Left
	Frontal	Superior Frontal Gyrus (Medial-Orbital) Right
	Limbic	Posterior Cingulate Cortex Left
	Limbic	Posterior Cingulate Cortex Right
	Parietal	Angular Gyrus Left
	Parietal	Angular Gyrus Right
6	Subcortical	Putamen Left
	Subcortical	Pallidum Left
	Subcortical	Putamen Right
	Subcortical	Pallidum Right
7a	Frontal	Precentral Gyrus Left
	Frontal	Precentral Gyrus Right
	Parietal	Postcentral Gyrus Left
	Parietal	Postcentral Gyrus Right
7b	Frontal	Supplementary Motor Area Left
	Frontal	Supplementary Motor Area Right
7c	Parietal	Paracentral Lobule Left
	Parietal	Paracentral Lobule Right
8a	Central	Rolandic Operculum Left
	Central	Rolandic Operculum Right
	Temporal	Heschl Gyrus Right
	Temporal	Heschl Gyrus Left
8b	Temporal	Superior Temporal Gyrus Left
	Temporal	Superior Temporal Gyrus Right
8c	Limbic	Insula Left
	Limbic	Insula Right
	Parietal	SupraMarginal Gyrus Left
	Parietal	SupraMarginal Gyrus Right
9a	Frontal	Olfactory Cortex Left
	Frontal	Olfactory Cortex Right
	Frontal	Gyrus Rectus Left
	Frontal	Gyrus Rectus Right
9b	Limbic	Hippocampus Left
	Limbic	ParaHippocampal Gyrus Left
	Limbic	Hippocampus Right
	Limbic	ParaHippocampal Gyrus Right
	Limbic	Amygdala Right
	Limbic	Amygdala Left
9c	Temporal	Temporal Pole (Superior) Left
	Temporal	Temporal Pole (Superior) Right
	Temporal	Temporal Pole (Middle) Left
	Temporal	Temporal Pole (Middle) Right

Table 6.2: Clustering analysis of detrended fMRI time courses before TA. List of regions in all (sub)clusters.

of the experimental paradigm. Further network analysis with hierarchical clustering showed that the activity-inducing signals revealed by TA contained information about meaningful task-related and resting-state networks, demonstrating good abilities for the study of non-stationary dynamics of brain activity.

Dynamics of Activity-Inducing Maps

When visualizing activity-inducing signals obtained by TA as dynamic brain maps, we could easily recognize the presence of the visual stimuli. However, it was also clear that the data were much richer and many spontaneous events were captured as well; e.g., we observed strong activity in the visual network of subject 2 during the final resting period (Fig. 6.7). Interestingly, activity-inducing signals revealed some non-stationary relationships between the different brain regions; e.g., as could be seen from Fig. 6.8, the correla-

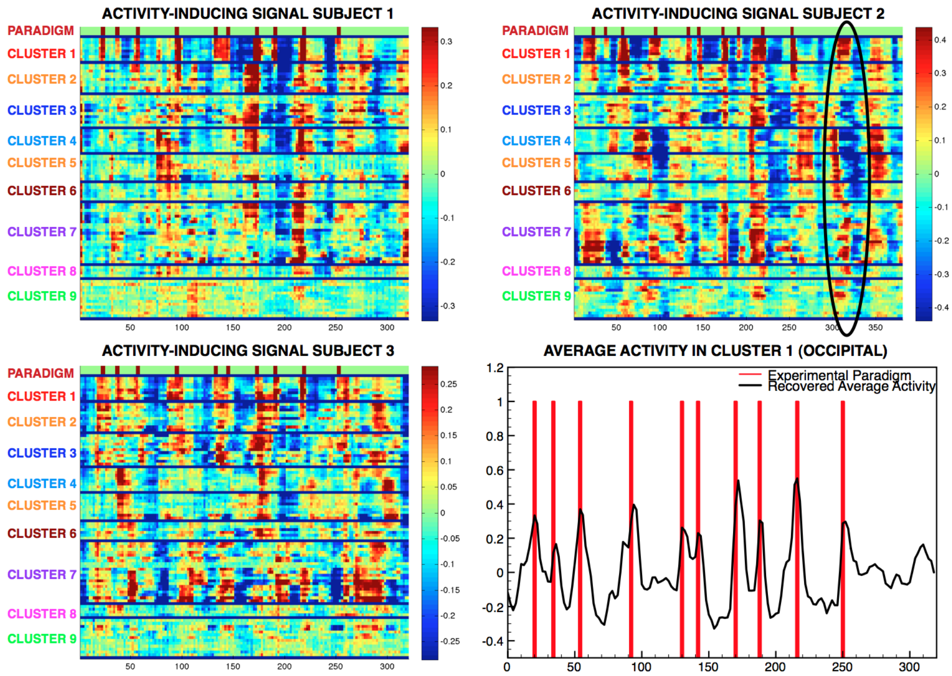


Figure 6.7: Activity-inducing signals per region and subject. Clearly, the activity-inducing signal in the visual regions (clusters 1 and 2) followed the visual paradigm closely. Moreover, we observe the intrinsic brain activity, for example, a spontaneous event (in black contour) occurs around 300s in subject 2 which was followed by negative activation in clusters 4-7 (posterior default-mode network). The average activity in the occipital lobe (bottom right) matched with the visual stimulation.

tion sign of signals from PCC and visual cortex were alternating. Similar non-stationary behavior has also been noted in a time-frequency (wavelet) coherence analysis of fMRI data [189]. Moreover, in recent work, Smith *et. al* exploited temporal and spatial ICA on high resolution data to reveal the temporally-independent and spatially overlapping activity maps called "temporal functional modes" [190]. The authors showed that different networks share common subcomponents of each other, that is, one brain region does not necessarily belongs to a distinct functional network.

Hierarchical Clustering

Clustering TA-recovered activity-inducing signals leads to a better understanding of the data. We obtained functionally plausible networks (many

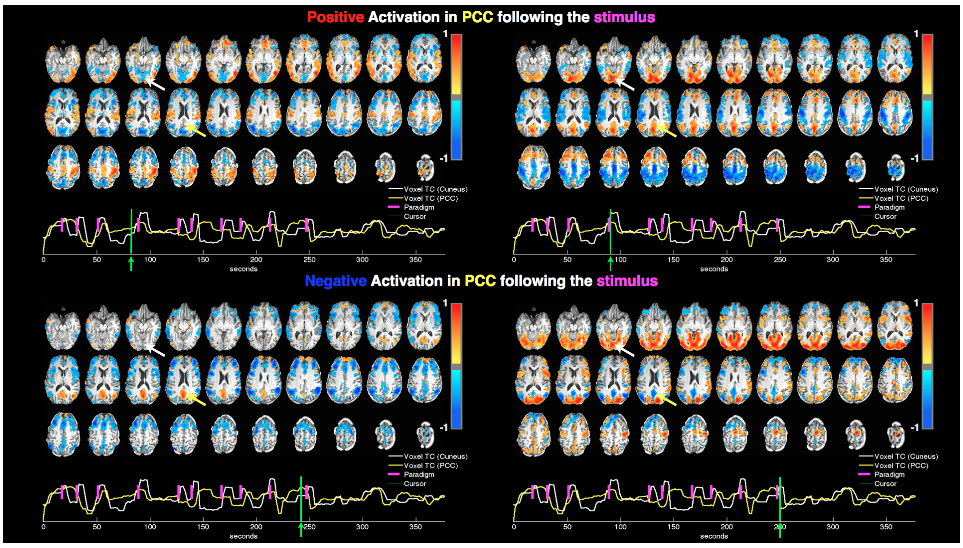


Figure 6.8: Reconstructed dynamic activity-inducing maps for subject 2. Time courses from cuneus and PCC are plotted in white and yellow, respectively. The stimulus time course is shown in magenta. In the first frame, the activity maps are illustrated for two instances (top row around 90 s and bottom row around 250s). The left and right columns show the activation maps before and after the stimulus, respectively. PCC lagged the stimuli with positive (top row) or negative response (bottom row).

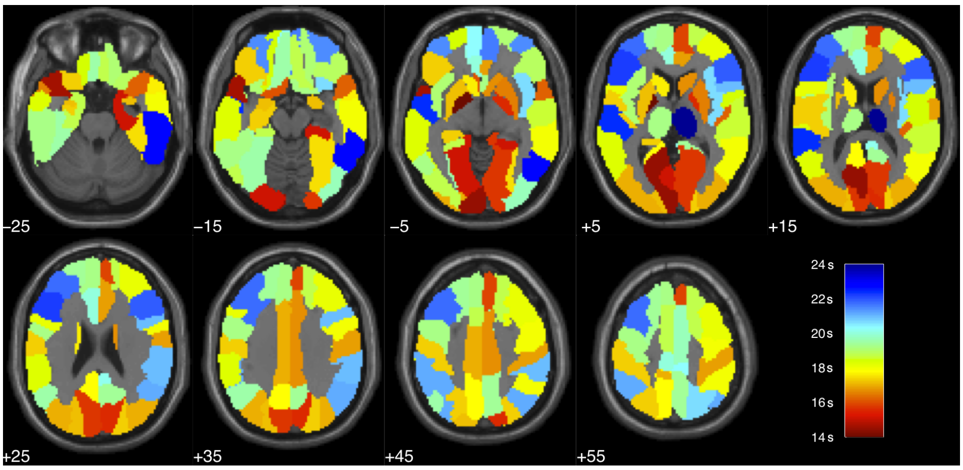


Figure 6.9: Average block length for the region-averaged activity-inducing signals. Regions in the visual network had relatively shorter average activity than other brain regions.

bilateral) reflecting both task-related and task-negative[2] activity. It is somewhat intriguing that structure of both task-related and resting-state networks were so well captured. While it is known that spontaneous activity is intertwined with task-related activity and networks very similar to resting-state networks are formed [169], it also means that our model for block-like activity-inducing signals is well suited for both types of activity. This raises the interesting hypothesis for future studies whether resting-state activity is rather block-like (with long durations on average) versus low frequency sinusoidal fluctuations as it is commonly assumed in resting-state studies.

6.3 Simultaneous EEG-FMRI Data Analysis of Epilepsy Patients

In the previous section, we validated TA through an event-related experiment; the activity-inducing signals in the visual regions confirmed the experimental paradigm. In this section, we exploit TA to explore RS-fMRI data acquired with simultaneous recording of electroencephalography (EEG) data in pharmaco-resistant epileptic patients with the goal of detecting and localizing unpredicted interictal epileptic discharges (IEDs).

6.3.1 Epilepsy Monitoring with Simultaneous EEG-FMRI

Epilepsy is a major neurological disorder that causes recurrent episodes of seizures affecting cognitive and physical functions of the patients. Interictal epileptic discharges (IEDs) are transient discharges that occur between the seizures. EEG is one of the most prominent methods to diagnose and monitor epilepsy noninvasively [191, 192]. EEG measures the electric potential induced by neuronal using several electrodes on the scalp. It offers a valuable research and clinical tool with a time resolution of milliseconds. However, EEG alone lacks the spatial resolution to localize the sources, especially in deep brain structures; i.e., EEG source localization is an ill-posed inverse problem that requires additional assumptions on the source model. Therefore, to profit from fMRI's high spatial resolution and overcome the source localization problem of EEG, simultaneous recording of EEG and FMRI is often proposed for presurgical exploration. Importantly, the EEG signals must be cleaned from MR artifacts [191]. Combined with the electrophysiological measurements, BOLD signal variations due to the EEG-driven IEDs can be identified and mapped on structural MR image of the patients. The spatial maps are then used to (1) delineate the surgical resection areas or (2) localize the target areas of the intracranial EEG (icEEG) electrodes for

[2]Task-negative networks, the prominent one being the default-mode network, are active during rest

further investigation [193–195]. However, significant variations in BOLD signal due to EEG-driven IEDs have been reported only for around 50% of the analyzed patients [173, 196, 197]. The possible explanations constitute;

1. No IEDs are observed during EEG-fMRI.
2. The presumed BOLD model might not be sensitive to some IEDs during the scanning due to modifications in the neurovascular coupling.
3. The electric potentials related to epileptic seizures may not be observable on the surface if these occur in deep-brain structures [198].

Therefore, despite being an invasive procedure, icEEG is still acknowledged as the "gold-standard" before proceeding towards epilepsy surgery. It is also possible that EEG recordings might miss the IEDs that can be detected with icEEG [199]. Indeed, experiments with simultaneous icEEG-fMRI show that the BOLD signal seems to localize the epileptogenic regions based on precise icEEG-driven IEDs while EEG might miss them due to spatial blurring [200]. Yet, icEEG-fMRI is not a common procedure as it carries various risks and many concordant results support the profound benefits of simultaneous EEG-fMRI for presurgical evaluation [201–203]. All these studies suggest that further investigation of the BOLD correlates of IEDs is still needed for understanding the physiology of epilepsy and improving the presurgical assessments. For a review on imaging studies of patients with epilepsy we refer the reader to [196, 202, 203].

Typically, EEG is the primary imaging method that drives the analysis in simultaneous EEG-fMRI studies of IEDs. EEG-derived IED onsets are used to set up the regressors of general linear model (GLM) [194, 198, 204]. Fig. 6.10(a) illustrates the schematic diagram of the conventional analysis. A temporal indicator function of IED onsets pinned by an neurophysiologist is convolved with HRF to be fed into GLM analysis as regressors. Topographic mapping (TM) approach is proposed for finding the BOLD correlate of epileptogenic activity when no IED is detected during simultaneous EEG-fMRI [198], shown in Fig. 6.10(b). Instead of detecting IEDs during the EEG-fMRI, the topographic map of epileptic activity is extracted from long-term EEG outside the MR and correlated with the intra-MRI EEG. The resulting spatial similarity time course is then plugged into the GLM analysis.

Alternative methods for the investigation of IED with simultaneous EEG-fMRI include independent component analysis [205–208], activelets [141], and mutual information [209]. Here, we apply TA to data of patients with epilepsy and localize the epileptogenic regions by measuring the similarity between the EEG-driven onsets, either from simultaneous EEG-fMRI or long-term EEG.

6.3.2 Acquisition and Preprocessing

The data consist of five pharmacoresistant epilepsy patients, scanned during simultaneous EEG-fMRI. Long-term EEG was also recorded. The fMRI data was acquired by a Siemens 3T TIM Trio MR scanner with gradient echo EPI while resting (eyes-closed). The acquisition parameters were TR/TE/FA = 1.5s/35ms/85^o, voxel size= $3.75\times3.75\times5.5$mm^3, 25 slices and N=1100 scans (subjects 1-4) and 600 volumes (subject 5). T1- and T2-weighted (pre & post-operation) images were also acquired. The preprocessing of the fMRI volumes included an initial realignment to the first volume to correct for head motion, and then spatial smoothing with Gaussian filter (FWHM=5mm) using SPM8 (FIL,UCL,UK). The anatomical AAL atlas [179] (90 regions without the cerebellum) was mapped onto the subject's functional space using the IBASPM toolbox [180]. The first 10 volumes were discarded so that the fMRI signal achieves steady-state magnetization. Voxels' time series labelled within the atlas were detrended for slow oscillations using a first-degree polynomial and DCT basis function up to cut-off frequency of 1/125 Hz, and finally scaled to have unit variance.

EEG signals were recorded with a 64 MR-compatible EEG cap (EasyCaps, FalkMinnow Services, Herrsching, Germany) according to the 10-20 system. Electrodes were equipped with an additional 5k resistance and impedances were kept as low as possible. EEG was acquired at 5kHz using 2 BrainAmp MR compatible amplifiers (Brain Products, Munich, Germany) and recordings were synchronized with the MR clock. MR gradient and cardioballistic artefacts were removed from the EEG using Vision Analyzer (Brain Products, Munich, Germany) using average artifact subtraction methods [210]. EEG data was subsequently downsampled to 250Hz, and IED were visually marked by an experienced neurophysiologist and averaged. The EEG map at the maximum of the GFP was selected as the epileptic map [198,211]. The maps were correlated with the intra-MRI EEG recordings and the absolute value of the correlation yielded the long-term EEG-driven IEDs [198].

Clinical details of patients are listed in Table1. Patients 1-3 had significant IEDs during simultaneous EEG-fMRI whereas patients 4-5 did not; therefore, only the topographic maps could be utilized for these patients. Three patients had undergone icEEG and all patients were seizure-free for more than one year after the resection surgery. Recently, (three years after the surgery) Patient 3 experienced new epileptic seizures with a different semiology.

Patient	Focus Localization	Cause	Scalp EEG focus	icEEG	Resection	IEDs	Outcome
1	Left frontal	Tuberous sclerosis	Left frontotemporal	+	Left prefrontal tuber	EEG-fMRI	SF (¿ 1 Y)
2	Left fronto-temporal	Gliosis bacterial abcess	Left frontal	+	Left parieto-temporal cortectomy	EEG-fMRI	SF (¿ 1 Y)
3	Left parieto-occipital	DNT	Bilateral parieto-occipital	-	Lesionectomy	EEG-fMRI	SF(¿1 Y)*
4	Left temporal	HS	Left temporal	-	Left anterior temporal lobe	Long-term EEG	SF (¿ 1 Y)
5	Left parieto-temporal	Tuberous sclerosis	Left temporal	+	Left parieto-temporal tuber	Long-term EEG	SF (¿ 1 Y)

Table 1: The clinical details of epilepsy patients. DNT: Dysembryplastic Neuroepithelial Tumor; icEEG: intracranial EEG; HS: hippocampal sclerosis; FCD: focal cortical dysplasia; SF: Seizure-free; $*$: new seizures occurred three years after the resection; Y: year.

6.3.3 Spatial Mapping of Interictal Epileptic Discharges with TA

We applied TA and recovered activity-inducing signals, which accommodated not only epileptic activity but also spontaneous brain activity. Therefore, a robust measure to reflect the brain regions with significant epileptic activity was needed. In order to extract epilepsy related activity, we correlated the activity-inducing signals with the EEG-driven IEDs. Fig. 6.11 illustrates our method step by step. Specifically, we performed the following steps;

1. We did not know the size of the epileptic focus a priori, however, we presumed that the regions of the anatomical atlas could be oversized. Indeed, TA could be able to delineate small areas. To that end, we further segregated large regions in the atlas into regions with a maximum of 100 voxels. Subsequently, K-means clustering was performed on the activity-inducing signals (K clusters for regions with number of voxels more than K $\times$ 100 voxels).

2. The IED indicator signals; i.e., the topography map correlations and EEG-fMRI IEDs, were convolved with a Gaussian filter (FWHM= 3s) and downsampled to the fMRI temporal resolution TR= 1.5s. Patient 1 had IEDs during simultaneous EEG-fMRI and both EEG-fMRI IED regressors and topographical correlations were used as IED reference signals. In contrast, patients 2-5 were analyzed exploiting only the correlation signal driven by topographical analysis.

3. The region averaged activity-inducing signals were correlated (Spear-

man) with the reference signals.

4. Finally, non-parametric hypothesis testing was performed to localize the epileptogenic regions. Correlations with 999 surrogates (random shuffling) were computed to establish the null-hypothesis distribution and fifth highest value of maximum statistics is selected as a threshold ($p < 0.05$ (corrected)) [209, 212].

The summary of results and comparison with topographic mapping is listed in Table 1. The target area was defined as the resection area and its proximity (¡15 mm margin). The results were designated as concordant (+) only if the regions survived the non-parametric test $p < 0.05$ (corrected); otherwise designated as discordant (-). The Patients 1-3, who had significant IEDs during the EEG-fMRI, showed concordant results with the clinical findings whereas no conclusive results were obtained form the analysis of the patients 4-5. For patient 5, the TM analysis results were found to be concordant [198]. We further discuss our results in detail for each patient.

Patient 1: The patient had tuberous sclerosis with two epileptogenic tubers. Fig. 6.12(a)-(b) depicts detected regions with TA using IEDs driven by intra-MRI IEDs and topography-related correlation, respectively. Both maps were concordant with the first target area (i.e., anterior frontal). Nevertheless, long-term EEG also provided localization around the second target area (inferior frontal).

Patient 2: The patient had left hemispheric epilepsy symptomatic of a large abscess gliotic scar. The focus areas confirmed by icEEG were left fronto-temporal-parietal areas. Fig. 6.13(a) depicts the significant epileptogenic regions. TA delineated the target areas and some remote areas on the left hemisphere; e.g., right hippocampus and inferior occipital, and negative correlations were revealed.

Patient 3: No icEEG were recorded for patient 3; the scalp EEG focused on bilateral parieto-occipital regions and resected area was localized on the left hemisphere. TA analysis showed negative correlations in the bilateral parietal-occipital regions. The epilepsy episodes reoccurred three years after the resection surgery with a different semiology.

Patient 4: The patient had hippocampal sclerosis ans was operated for the resection of the left anterior temporal lobe. Both TA and TM found significant diffuse bilateral regions which were not conclusive. The patient had significant head jerks during EEG-fMRI recordings, and it was substantially degraded by movement artifacts [198].

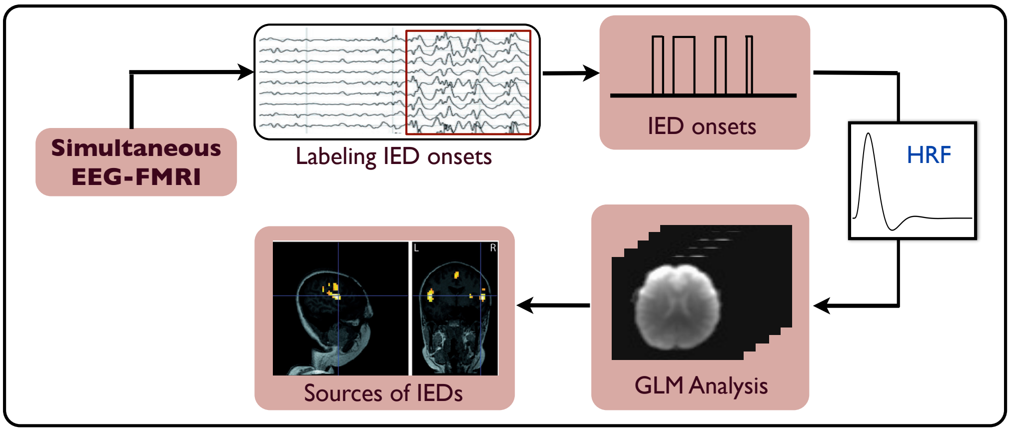
Patient 5: Patient 5 had tuberous sclerosis localized in the left temporal lobe. The resected area was the left parieto-temporal tuber. TA found diffuse bilateral regions, located more on the right hemisphere. TM also found diffuse bilateral regions but endorsing the left parietal target area. Fig. 6.13(b) depicts the spatial map obtained by TA analysis. We remark that the detected area constituted the default-mode network (DMN), which is characterized as the task-negative network in RS-fMRI. Indeed, the involvement of the posterior regions in temporal lobe epilepsy has been previously noted in various studies in the literature [173, 200].

Patient	**Concordance: TA/TM**	**IED onsets**	**TA (in target)**	**TA (remote)**	**TM (in target)**	**TM (remote)**
1	+/+	EEG-fMRI	Left ant frontal	Left medial frontal, Left insula	Left ant frontal, Left inf frontal	Left medial frontal, Right cerebellum, Bilat parietal, Right frontal
1	+/+	Long-term EEG	Left ant frontal	Left medial frontal, Left insula, Left temp sup, Right rolandic operculum	Left ant frontal, Left inf frontal	Left medial frontal, Right cerebellum, Bilat parietal, Right frontal
2	+/+	Long-term EEG	Left parietal, Left frontoparietal	Right hippocampus, Right parietal, Right inf occipital	Left parietal, Left frontoparietal	bilat cingulate, Right temporal, Right cerebellum, Basal ganglia
3	+/+	Long-term EEG	Left parieto-occipital	Right parieto-occipital , Mid occipital	Left parieto-occipital	Left cerebellum, Basal ganglia, Bilat orbito-frontal, Right parietal
4	-/-	Long-term EEG	Diffuse bilateral	Scattered Mid occipital, Right temporal, Bilat mid frontal	Diffuse bilateral	Scattered bilateral sup frontal
5	-/+	Long-term EEG	–	Bilat temporal superior, Ant frontal, Post cingulate, Right parietal	Left parietal	Diffuse bilateral

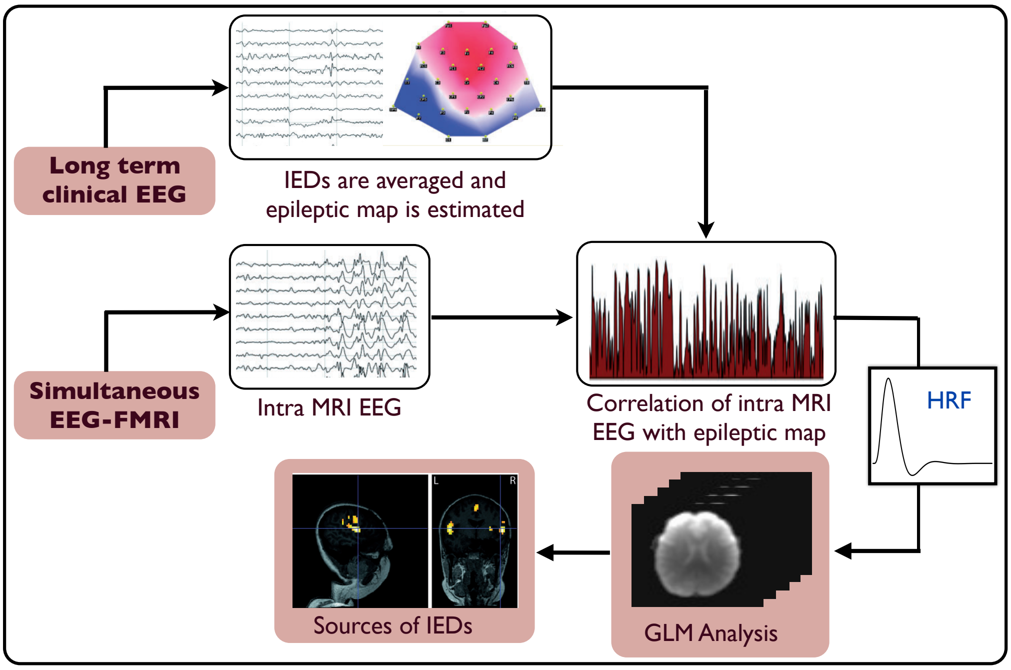
Table 1: The results of TA and topographic map (TM) analysis, ant/post/inf/sup/mid/bilat = anterior/posterior/inferior/superior/middle/bilateral. Concordance is defined by p-value ($p < 0.05$ corrected). Two IED onset signals were used; (1) driven from simultaneous EEG-fMRI, (2) driven from topographic map correlations. Both TA and TM found in target areas, defined as the resection areas with 15 mm proximity margin, and out of focus (remote) areas.

6.4 Discussion and Summary

We have applied TA to simultaneous EEG-fMRI data of five patients with epilepsy. The activity-induced signals were correlated with the EEG-drived IED signals (either simultaneous EEG-fMRI or long-term EEG). TA was able to localize the epileptogenic regions especially when IEDs were detected during simultaneous EEG-fMRI. Even though TA did not provide conclusive results for the patients who did not have IEDs during EEG-fMRI (Patients 4-5), in Patient 5 the TA analysis revealed negative correlations with the reference signals in the areas of captured default-mode network (see Fig. 6.13(b)). This is an intriguing finding which motivates us to further investigate the use of TA for the study of dynamics of the BOLD correlates associated with IED by following these possible paths; (1) studying the network organization during the epileptic activity and outside the epileptic activity: “Are there any suspended or persistent network structures?”, (2) investigating the latencies of the BOLD response in relation to IEDs, (3) elaborating the effect of IED occurrence frequency on BOLD signals; especially in terms of BOLD deviations and scattered activity patterns.



(a) Conventional analysis for detecting epileptogenic regions from simultaneous EEG-fMRI



(b) Topographic map correlation analysis

Figure 6.10: Schematic representation of state-of-the-art methods for localizing epileptogenic regions from simultaneous EEG-fMRI. (a) IEDs driven from simultaneous EEG-fMRI is fed into GLM analysis. (b) Topographic mapping uses the long term EEG recordings to find the epileptogenic mapping when no IEDs are observed during simultaneous EEG-fMRI. The epileptic map is correlated with the simultaneous EEG recordings and the correlation time course is fed into GLM analysis (Courtesy of Grouiller *et al.* [198]).

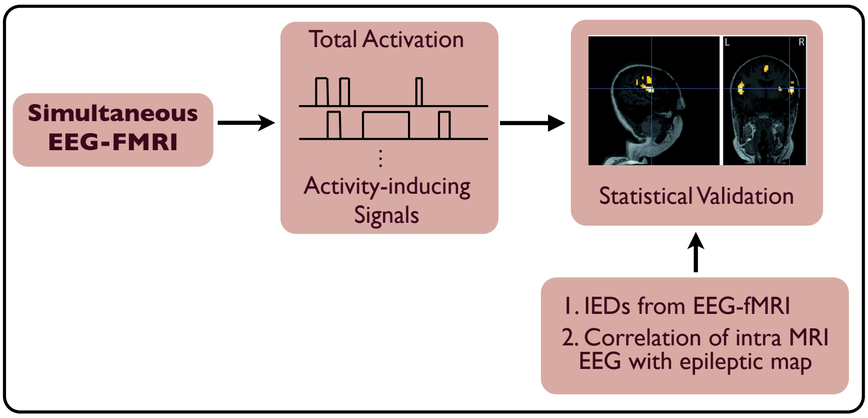
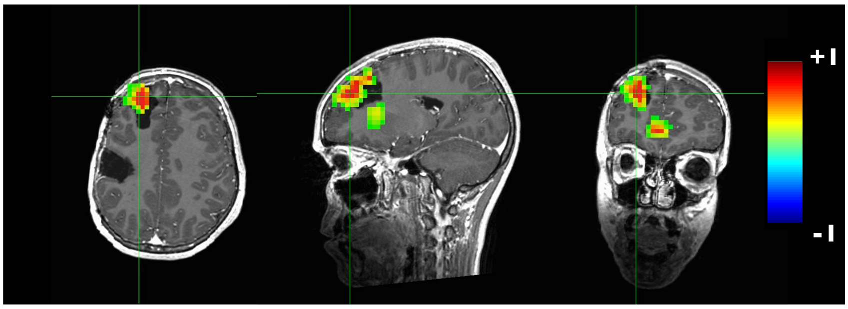
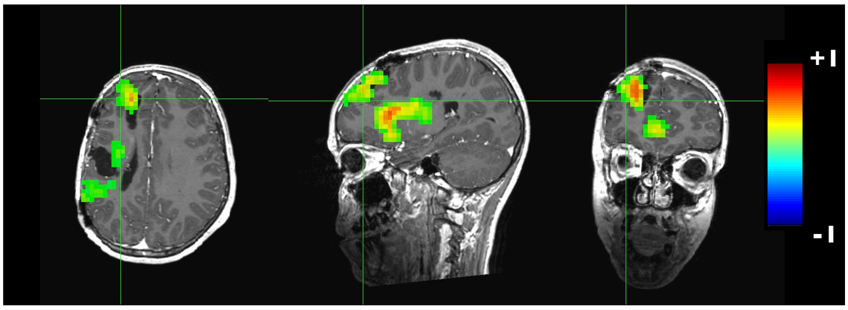


Figure 6.11: Total activation analysis in patients with epilepsy. The region averaged activity inducing signals are correlated with the IEDs driven either from EEG-fMRI or long tern EEG. Non-parametric hypothesis testing revealed epileptogenic brain regions.

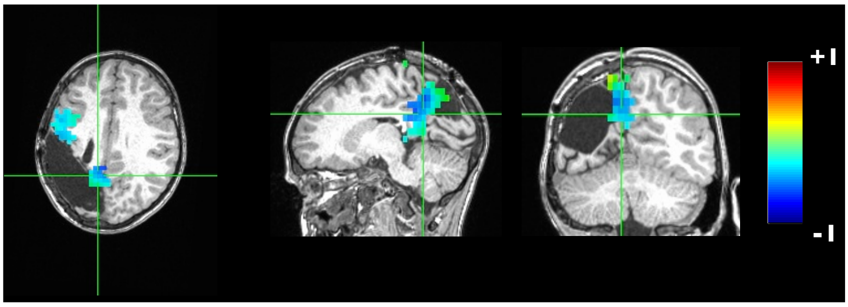


(a) Localization using IEDs driven by simultaneous EEG-fMRI (Patient 1)

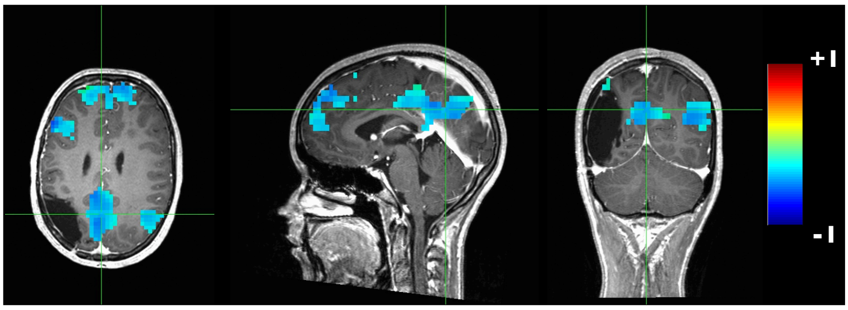


(b) Localization using IEDs driven by topographic map (Patient 1)

Figure 6.12: Spatial mapping of estimated epileptogenic brain regions of Patient 1. (a) IEDs were driven from simultaneous EEG-fMRI. (b) IEDs were driven by topographic map correlation. Both results were concordant with the target regions.



(a) Accordance with the target area (Patient 2)



(b) Discordance with the target area (Patient 5)

Figure 6.13: Spatial mapping of detected epileptogenic brain regions of Patient 2 and 5 using IEDs driven by topography-related correlation. (a) Patient 2 had IEDs during EEG-fMRI; the map includes left parietal and fronto-parietal regions, and are concordant with target regions. (b) Patient 5 had no IED during simultaneous EEG-fMRI; the map shows the negatively correlated DMN.

Chapter 7

Discussion and Outlook

In this dissertation, we have introduced a novel framework, total activation (TA), which opens new avenues for the analysis of fMRI data. Our contributions are two-folds:

1. From a signal processing perspective, we extended the total variation (TV) regularization to incorporate a linear system.
2. We further developed and applied the generalized TV as a new spatiotemporal regularization for fMRI data analysis.

In what follows, we discuss our main achievements, explore the potential future research directions, and identify possible extensions of our method.

7.1 Summary

Generalized *L*-Total Variation We extended the TV regularization concept, which is favoring piecewise constant signals. More complex signals, such as those composed out of the Green's functions of differential operators other than the first-order derivative, could be handled as well. The regularization was expressed as an analysis prior, thereby, acting directly on the system's driving signal through sparsity-promoting ℓ_1-norm in terms of deconvolution an facilitating interpretation. In particular, we guarantee the sparsity of the "*innovation*" signal obtained after applying the differential operator and access both denoised and deconvolved signals easily. Simulation results and real audio signal examples highlighted the improvement of generalized L-TV over existing methods.

Total Activation for fMRI We developed a spatiotemporal regularization method for the recovery the activity-inducing signals in fMRI without requiring prior knowledge of the onsets and durations of the events. TA overcame the lack of existing methods by incorporating the hemodynamic system and an anatomical prior. The variational formulation included generalized L-TV as a temporal regularization term that inverts the fMRI's hemodynamic blur. By that means, specific tuning of the differential operator enforced sparse "innovation" signals and, consequently, block-type activity-inducing signals. Furthermore, a spatial regularization was added to account for voxel dependencies; i.e., the $\ell_{2,1}$-norm favored smooth activations inside anatomically defined regions. The multi-term regularization problem was solved efficiently, with automatic calibration of temporal regularization parameter, exploiting state-of-the-art convex optimization techniques. We proved the flexibility of the method to model perturbations, which, in turn, encouraged us to go forward with real experimental data.

FMRI Experimental Results FMRI data acquired during event-related experiment (visual stimuli) were analyzed with TA. The results showed convincingly that TA recovered not only the activation patterns for visual stimuli whose timing information was unknown to the model, but also other plausible resting-state networks, thereby, suggesting great use for studying non-stationary dynamics in fMRI. Finally, using simultaneous EEG-fMRI recordings, we obtained promising results for localizing the brain regions that related (unknown) interictal epileptic discharges in patients with epilepsy.

7.2 Outlook

TA provided compelling results that potentially can lead to new insights into exploration of brain organization and temporal dynamics. Here, we discuss future considerations and some emerging directions concerning the techniques we developed in this thesis.

Model Selection Generalized L-TV framework currently takes in a fixed differential operator to invert the degradation effect of the underlying linear system. In its formal definition, a linear differential operator is defined by poles and zeros from which a discrete filter can be implemented. In this thesis, we built the simplest (i.e., minimum support) discrete filter of the continuous differential operator. One future goal would be to improve the filter implementation better taking into account the frequency spectrum of the operator. Another aspect worth elaborating in the future is using this scheme for system identification; i.e., model selection. Specifically, we could optimize the characterization of the differential operator from model parameters and sparsity of the innovation signal.

Continuous-Domain Interpretation Despite the fact that generalized L-TV is inspired by the continuous domain formulation, future research is needed to tighten the mathematical link between the proposed signal-processing approach (in the discrete domain) and proper generalization of

TV in the continuous domain. Our approach is promising in this respect because in recent work it was shown that the signal-processing approach for conventional TV (that is, ℓ_1-norm of finite differences) can be linked to proper continuous-domain modeling of stochastic processes [213, 214].

Higher-Dimensional Extensions Generalized L-TV can be extended for image denoising and deconvolution. Specifically, the operators role would result in efficient representation of the undetermined systems in analysis formulations [215].

FMRI Deconvolution TA for fMRI took the (linearized) balloon model for granted in its temporal regularization. It is, however, well known that this model is limited since high variability of the hemodynamic response function (HRF) within and across subjects has been noted, and the HRF identification problem has been well acknowledged [85, 158]. Therefore, a fundamental step in terms of fMRI analysis is to be able to account for the hemodynamic variability. In its original form, it is possible to incorporate different HRF models per regions or voxels as long as they are defined a priori. Instead, the substantial contribution would be to perform HRF identification and activity-inducing signal estimation simultaneously within the TA framework.

Continuous Domain Solution TA adopts the underlying continuous domain definition of differential operators, however, instead of considering the explicit analytical problem, the sampled formulation, which requires discretized operators, is solved. Recently, in signal processing community, a new framework, named finite rate of innovation (FRI), that could go beyond the Nyquist sampling limit was introduced [216]. According to this theory, the analytical problem can be driven by spike-type signals and solution can be expressed in continuous domain since HRF is expressed explicitly in a compact analytical form. To that end, FRI would be a potential candidate for handling the problem in the continuous domain.

Anatomical Prior TA used an anatomical atlas that is not optimal; i.e., not subject-specific and contains course structures (90 regions). Several interesting ideas could be adapted in the future for improving the spatial processing: (1) estimating a (functional) data-adaptive atlas [160, 217]; or (2) incorporating source separation methods (e.g., ICA) to define the regions instead of an anatomical atlas.

Spatiotemporal Dynamics The recovered activity-inducing signals revealed intriguing properties in terms of spontaneous activity and dynamical brain organization. Current state-of-the-art methods, especially in functional network connectivity analysis, are moving away from static analysis and increasingly exploring non-stationary dynamical organization [190, 218–220]. In that direction, TA-regularized activity-inducing signals would constitute a substantially better starting point for dynamical analysis of the data. Indeed, the activity-inducing signals have an increased temporal "crispness" and are also cleaned from noise.

Clinical Applications We showed that simultaneous EEG-fMRI recordings could give an insight into localization of epileptogenic brain regions for pre-surgical planning. The literature contains a few prominent examples of fMRI-driven analyses with compelling results [141, 209]. We presented our preliminary results on detection and localization of interictal epileptic discharges, which are encouraging for future investigations to be confirmed with more subjects. The underlying dynamics in the BOLD revealed by TA also enables to study the stimuli's effect in task-based experiments. One other example is the study of fMRI data for exposure of olfactory stimuli, where both the precise timing if the onsets (due to breathing) and habituation are unknown.

Appendix A

Appendices

A.1 Proof of Proposition 1

Proof. We make the proof by construction. For the first-order differential operator $L = (D - \alpha_1)$ for which $N = 1$ and $M = 0$, the corresponding discrete operator, Δ_L, becomes $\Delta_L \{\mathbf{x}\}[n] = \mathbf{x}[n] - e^{\alpha_1}\mathbf{x}[n-1]$, see [221].

For the differential operator $L = \prod_{i=1}^{N}(D - \alpha_i I)$ of order $N > 0$ and $M = 0$, we can obtain the filter Δ_L by successive convolutions (leading to support of $N+1$); the z-transform of $\widehat{\Delta}_L(z)$ is then

$$\widehat{\Delta}_L(z) := \sum_{n=0}^{N} \Delta_L[n] z^{-n} = \prod_{i=1}^{N}(1 - e^{\alpha_i} z^{-1}) = \prod_{i=1}^{N} p_i(z^{-1}), \tag{A.1}$$

where p_i is a polynomial with 2 coefficients $p_{i,k} = (-e^{\alpha_i})^k$ with $k \in [0, 1]$. Note that we can express (A.1) benefiting the polynomial multiplication which leads to the convolution as

$$\prod_{i=1}^{N} p_i(z^{-1}) = \sum_{n=0}^{N} \mathbf{P}_N[n] z^{-n}, \tag{A.2}$$

where $\mathbf{P}_N = p_1 * p_2 * \ldots * p_N$ and $*$ is the convolution operator. Therefore, we can express the filter $\Delta_L[n] = \mathbf{P}_N[n]$ as

$$\mathbf{P}_N[n] = \sum_{k_{N-1}=0}^{n} \cdots \sum_{k_2=0}^{k_3} \sum_{k_1=0}^{k_2} (-e^{\alpha_1})^{k_1} (-e^{\alpha_2})^{k_2 - k_1} \\ \ldots (e^{\alpha_N})^{n - k_{N-1}}, \quad \{k_1, \ldots, n - k_{N-1}\} \in [0, 1]^N,$$

with a change of variables $m_1 = k_1, m_i = k_i - k_{i-1}, m_N = n - k_{N-1}$ we have

$$\Delta_L[n] = (-1)^n \sum_{|\mathbf{m}|=n} (e^{\boldsymbol{\alpha}})^{\mathbf{m}}, \quad \mathbf{m} \in [0, 1]^N, 0 \leq n \leq N. \tag{A.3}$$

Similarly, for the general differential operator, L, with $M > 0$ we have in z-domain

$$\widehat{\Delta}_L(z) = \widehat{\Delta}_{L_n}(z)\widehat{\Delta}_{L_d}^{-1}(z) = \frac{\prod_{i=1}^{N}(1 - e^{\alpha_i} z^{-1})}{\prod_{i=1}^{M}(1 - e^{\gamma_i} z^{-1})}, \tag{A.4}$$

where $\Delta_{L_n}[n]$ is represented explicitly in (A.3). Note that the filter $\widehat{\Delta}_{L_d}^{-1}$ in (A.4) has infinite support in time, therefore stability should be assured. Depending on the poles $\boldsymbol{\gamma}$ of the operator, we find a combination of causal ($\boldsymbol{\gamma}'$) and anti-causal ($\boldsymbol{\gamma}''$) filters that guarantee stability; e.g., for $M = 1$ and $N = 0$ we have either $\Delta_L[n] = e^{\gamma' n}u[n]$ or $\Delta_L[n] = -e^{\gamma'' n}u[-n-1]$, where $u[n]$ is the unit step function. In practice the filter with input $\mathbf{x}$ and output $\mathbf{y}$ can be reformulated and implemented as in a recursive way by providing the realization of causal and anti-causal parts separately. To this aim, we represent the inverse filter

$$\widehat{\Delta}_{L_d}^{-1}(z) = \widehat{\Delta}_{L_d'}^{-1}(z)\, \widehat{\Delta}_{L_d''}^{-1}(z)\, \mathbf{e}^{-\boldsymbol{\gamma}''}(-z)^{m_2} \tag{A.5}$$

by the causal and anti-causal filters

$$\begin{aligned} \widehat{\Delta}_{L_d'}(z) &= \prod_{i=1}^{m_1}(1 - e^{\gamma_i'} z^{-1}) \\ \widehat{\Delta}_{L_d''}(z) &= \prod_{i=1}^{m_2}(1 - e^{-\gamma_i''} z), \end{aligned} \tag{A.6}$$

where $\mathbf{e} = (e, \ldots, e)$ is a vector of length m_2. Then the corresponding recursive algorithm can be easily derived from the z-domain representation. Here we will concentrate on the anti-causal part (the derivation for the causal part is similar). To obtain $\mathbf{y} = \Delta_{L_d''}^{-1}\left\{\mathbf{x}[n + m_2](-1)^{m_2}\, \mathbf{e}^{-\boldsymbol{\gamma}''}\right\}$, we consider

$$\widehat{\mathbf{y}}(z)\widehat{\Delta}_{L_d''}(z) = \mathbf{e}^{-\boldsymbol{\gamma}''}(-z)^{m_2}\, \widehat{\mathbf{x}}(z),$$

from which we find

$$\sum_k \mathbf{y}[n-k]\, \Delta_{L_d''}[k] = \mathbf{x}[n + m_2](-1)^{m_2}\, \mathbf{e}^{-\boldsymbol{\gamma}''}.$$

From (A.3), we can derive the explicit time domain expression for the anti-causal filter $\Delta_{L_d''}$ as

$$\Delta_{L_d''}[n] = (-1)^n \sum_{|\mathbf{m}|=-n} (e^{-\boldsymbol{\gamma}''})^{\mathbf{m}},\ \mathbf{m} \in [0,1]^{m_2}, -m_2 \leq n \leq 0. \tag{A.7}$$

Therefore, we obtain

$$\mathbf{y}[n] = \mathbf{x}[n + m_2]\Delta_{L_d''}[-m_2] - \sum_{k=-m_2}^{-1} \mathbf{y}[n-k]\Delta_{L_d''}[k],$$

where we used $\Delta_{L_d''}[0] = 1$.

$\square$

Let us give an example for the third-order differential operator $L = \prod_{i=1}^{3}(D - \alpha_i I)$. The FIR filter $\Delta_L[n]$ then becomes

$$\Delta_L[n] = [1, -(e^{\alpha_1} + e^{\alpha_2} + e^{\alpha_3}), e^{(\alpha_1+\alpha_2)} + e^{(\alpha_1+\alpha_3)} \\ + e^{(\alpha_2+\alpha_3)}, -e^{(\alpha_1+\alpha_2+\alpha_3)}], \quad 0 \leq n \leq 3.$$

Bibliography

[1] A. M. Owen, M. R. Coleman, M. Boly, M. H. Davis, S. Laureys, and J. D. Pickard, "Detecting awareness in the vegetative state," *Science*, vol. 313, no. 5792, pp. 1402, 2006.

[2] N.K. Logothetis, J. Pauls, M. Augath, T. Trinath, and A. Oeltermann, "Neurophysiological investigation of the basis of the fMRI signal," *Nature*, vol. 412, no. 6843, pp. 150–157, 2001.

[3] G. H. Glover, "Deconvolution of impulse response in event-related BOLD fMRI," *NeuroImage*, vol. 9, no. 4, pp. 416 – 429, 1999.

[4] M. Elad, P. Milanfar, and R. Rubinstein, "Analysis versus synthesis in signal priors," *Inverse Problems*, vol. 23, no. 3, pp. 947–968, June 2007.

[5] S.G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3397–3415, 1993.

[6] Y.C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition," in *Proceedings of Record of The Twenty-Seventh Asilomar Conference on Signals, Systems and Computers, 1993.*, 1993, pp. 40–44 vol.1.

[7] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, pp. 33–61, 1998.

[8] T. Hastie, R. Tibshirani, and J. H. Friedman, *The elements of statistical learning: data mining, inference, and prediction: with 200 full-color illustrations*, New York: Springer-Verlag, 2001.

[9] M. Yuan and Y. Lin, "Model selection and estimation in regression with grouped variables," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 68, no. 1, pp. 49–67, 2006.

[10] L. Jacob, G. Obozinski, and J.-P. Vert, “Group LASSO with overlap and graph lasso,” in *Proceedings of the 26th Annual International Conference on Machine Learning*, New York, NY, USA, 2009, ICML ’09, pp. 433–440, ACM.

[11] R. Tibshirani, M. Saunders, S. Rosset, J. Zhu, and K. Knight, “Sparsity and smoothness via the fused LASSO,” *Journal of the Royal Statistical Society Series B*, vol. 67, no. 1, pp. 91–108, 2005.

[12] R. Chartrand, “Exact reconstruction of sparse signals via nonconvex minimization,” *IEEE Signal Processing Letters*, vol. 14, no. 10, pp. 707–710, 2007.

[13] E. J.. Candes, M. B.. Wakin, and S. P. Boyd, “Enhancing sparsity by reweighted ℓ_1 minimization,” *Journal of Fourier Analysis and Applications*, vol. 14, no. 5-6, pp. 877–905, 2008.

[14] R. Chartrand, “Fast algorithms for nonconvex compressive sensing: MRI reconstruction from very few data,” in *Proceedings of IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2009 (ISBI ’09).*, 2009, pp. 262–265.

[15] Z. Lyu, Q.and Lin, Y. She, and C. Zhang, “A comparison of typical minimization algorithms,” *Neurocomputing*, 2013, in press.

[16] L. I. Rudin, S. Osher, and E. Fatemi, “Nonlinear total variation based noise removal algorithms,” in *Proceedings of the eleventh annual international conference of the Center for Nonlinear Studies on Experimental mathematics : computational issues in nonlinear science.* 1992, pp. 259–268, Elsevier North-Holland, Inc.

[17] T. F. Chan, S. Osher, and J. Shen, “The digital TV filter and nonlinear denoising,” *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 231–241, Feb. 2001.

[18] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*, Springer, first edition, 2010.

[19] G. Steidl, J. Weickert, T. Brox, P. Mrázek, and M. Welk, “On the equivalence of soft wavelet shrinkage, total variation diffusion, total variation regularization, and SIDEs,” *SIAM Journal on Numerical Analysis*, vol. 42, pp. 686–713, 2004.

[20] I. W. Selesnick and M. A. T. Figueiredo, “Signal restoration with overcomplete wavelet transforms: Comparison of analysis and synthesis priors,” in *Proceedings of the SPIE, Wavelets XIII*, Aug 2-4 2009, vol. 7446.

[21] I. Khalidov and M. Unser, "From differential equations to the construction of new wavelet-like bases," *IEEE Transactions on Signal Processing*, vol. 54, no. 4, pp. 1256–1267, Apr. 2006.

[22] C. Vonesch, T. Blu, and M. Unser, "Generalized Daubechies wavelet families," *IEEE Transactions on Signal Processing*, vol. 55, no. 9, pp. 4415–4429, September 2007.

[23] M. Elad, J. Starck, P. Querre, and D. Donoho, "Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA)," *Applied and Computational Harmonic Analysis*, vol. 19, no. 3, pp. 340–358, November 2005.

[24] J.-C. Baritaux, K. Hassler, M. Bucher, S. Sanyal, and M. Unser, "Sparsity-driven reconstruction for FDOT with anatomical priors," *IEEE Transactions on Medical Imaging*, vol. 30, no. 5, pp. 1143–1153, May 2011.

[25] R. Tibshirani, "Regression shrinkage and selection via the LASSO," *Journal of the Royal Statistical Society, Series B*, vol. 58, pp. 267–288, 1994.

[26] R. Efron, T. Hastie, I.M. Johnstone, and R. Tibshirani, "Least angle regression," *The Annals of Statistics*, vol. 32, no. 2, pp. 407–499, 2004.

[27] I. Daubechies, R. DeVore, M. Fornasier, and C. S. Güntürk, "Iteratively reweighted least squares minimization for sparse recovery," *Communications on Pure and Applied Mathematics*, vol. 63, no. 1, pp. 1–38, 2010.

[28] J. J. Moreau, "Proximitéet dualité dans un espace hilbertien," *Bull. Soc. Math. France*, vol. 93, pp. 273–299, 1965.

[29] P. L. Combettes and J.-C. Pesquet, "Proximal splitting methods in signal processing," in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, H. H. Bauschke, R. S. Burachik, P. L. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz, Eds., Springer Optimization and Its Applications, pp. 185–212. Springer New York, 2011.

[30] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 613 –627, May 1995.

[31] A. Chambolle, "An algorithm for total variation minimization and applications," *Journal of Mathematical Imaging and Vision*, vol. 20, no. 1, pp. 89–97, January 2004.

[32] R. T. Rockafellar, *Convex Analysis*, Princeton Press, 1996.

[33] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Transactions on Image Processing*, vol. 18, no. 11, pp. 2419–2434, Nov. 2009.

[34] P. L. Combettes and V. R. Wajs, "Signal recovery by proximal forward-backward splitting," *Multiscale Modeling and Simulation*, vol. 4, no. 4, pp. 1168–1200, 2005.

[35] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Communications on Pure and Applied Mathematics*, vol. 57, no. 11, pp. 1413–1457, 2004.

[36] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 183–202, 2009.

[37] B. Wohlberg and P. Rodriguez, "An iteratively reweighted norm algorithm for minimization of total variation functionals," *IEEE Signal Processing Letters*, vol. 14, no. 12, pp. 948–951, Dec. 2007.

[38] P. Rodriguez and B. Wohlberg, "Efficient minimization method for a generalized total variation functional," *IEEE Transactions on Image Processing*, vol. 18, no. 2, pp. 322–332, 2009.

[39] J.M. Bioucas-Dias, M. A T Figueiredo, and J.P. Oliveira, "Total variation-based image deconvolution: a majorization-minimization approach," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, 2006. (ICASSP'06)*, 2006, vol. 2, pp. II–II.

[40] M. A. T. Figueiredo, J. M. Bioucas-Dias, and R. D. Nowak, "Majorization Minimization Algorithms for Wavelet-Based Image Restoration," *IEEE Transactions on Image Processing*, vol. 16, pp. 2980–2991, Dec. 2007.

[41] Y. Wang, J. Yang, W. Yin, and Y. Zhang, "A new alternating minimization algorithm for total variation image reconstruction," *SIAM Journal on Imaging Science*, pp. 248–272, 2008.

[42] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jan. 2011.

[43] H. Raguet, J. Fadili, and G. Peyrè, "Generalized forward-backward splitting," 2012, Available [Online], arxiv.org/abs/1108.4404v3.

[44] P.L. Combettes, "Iterative construction of the resolvent of a sum of maximal operators," *Convex Analysis*, vol. 16, pp. 727–748, 2009.

[45] F. I. Karahanoglu, I. Bayram, and D. Van De Ville, "A signal processing approach to generalized 1-D total variation," *IEEE Transactions on Signal Processing*, vol. 59, no. 11, pp. 5265 –5274, nov. 2011.

[46] T. Chan, S. Esedoglu, F. Park, and A. Yip, "Recent developments in total variation image restoration," *Handbook of Mathematical Models in Computer Vision*, pp. 17–30, 2010.

[47] W. Stefan, R. Renaut, and A. Gelb, "Improved total variation-type regularization using higher order edge detectors," *SIAM Journal on Imaging Sciences*, pp. 232–251, Jan 2010.

[48] A. Chambolle and P. L. Lions, "Image recovery via total variation minimization and related problems," *Numerical Mathematics*, vol. 76, pp. 167–188, 1997.

[49] J. Yuan, C. Schnörr, and G. Steidl, "Total-variation based piecewise affine regularization," in *Proceedings of the Second International Conference on Scale Space and Variational Methods in Computer Vision 2009, (SSVM '09)*, Berlin, Heidelberg, 2009, pp. 552–564, Springer-Verlag.

[50] T. F. Chan, A. Marquina, and P. Mulet, "High-order total variation-based image restoration," *SIAM Journal on Scientific Computing*, vol. 22, no. 2, pp. 503–516, 2000.

[51] Z. Dogan, S. Lefkimmiatis, A. Bourquard, and M. Unser, "A second-order extension of tv regularization for image deblurring," in *Proceedings of 18th IEEE International Conference on Image Processing (ICIP), 2011.*, 2011, pp. 705–708.

[52] T. Chan, S. Esedoglu, and S. Park, "Image decomposition combining staircase reduction and texture extraction," *Journal of Visual Communication and Image Representation*, pp. 468–486, Jan 2007.

[53] F. Knoll, K. Bredies, T. Pock, and R. Stollberger, "Second order total generalized variation (TGV) for MRI," *Magnetic Resonance in Medicine*, vol. 65, no. 2, pp. 480–491, 2011.

[54] A. Chambolle, "Total variation minimization and a class of binary MRFs models," *Lecture Notes in Computer Science*, vol. 3757, pp. 136–152, 2005.

[55] J.-L. Starck, F. Murtagh, and J. M. Fadili, *Sparse Image and Signal Processing: Wavelets, Curvelets, Morphological Diversity*, Cambridge University Press, first edition, 2010.

[56] S. Ramani, T. Blu, and M. Unser, "Monte-Carlo SURE: A black-box optimization of regularization parameters for general denoising algorithms," *IEEE Transactions on Image Processing*, vol. 17, no. 9, pp. 1540–1554, September 2008.

[57] J.-F. Aujol, G. Gilboa, T. Chan, and S. Osher, "Structure-texture image decomposition – modeling, algorithms, and parameter selection," *International Journal of Computer Vision*, vol. 67, no. 1, pp. 111–136, Apr. 2006.

[58] V. Solo, "Selection of regularisation parameters for total variation denoising," in *IEEE International Conference Acoustics, Speech, and Signal Processing*, march 1999, vol. 3, pp. 1653 –1655.

[59] S.D. Babacan, R. Molina, and A.K. Katsaggelos, "Parameter estimation in TV image restoration using variational distribution approximation," *IEEE Transactions on Image Processing*, vol. 17, no. 3, pp. 326 –339, 2008.

[60] Y. Lin, B. Wohlberg, and H. Guo, "UPRE method for total variation parameter selection," *Signal Processing*, vol. 90, pp. 2546–2551, August 2010.

[61] F. Jaillet and B. Torrésani, "Time-Frequency Jigsaw Puzzle: adaptive multiwindow and multilayered Gabor expansions," *International Journal of Wavelets Multiresolution and Information Processing*, vol. 5, no. 2, pp. 293 – 315, 05 2007.

[62] G. Gilboa and S. Osher, "Nonlocal operators with applications to image processing," *SIAM Multiscale Modeling and Simulation*, vol. 7, no. 3, pp. 1005–1028, 2007.

[63] 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 USA*, vol. 87, no. 24, pp. 9868–9872, Dec 1990.

[64] S. A. Huettel, A. W. Song, and G. McCarthy, *Functional Magnetic Resonance Imaging*, Sinauer Associates, Inc., second edition, 2009.

[65] G. E. Sarty, *Computing Brain Activity Maps from fMRI Time-Series Images*, Cambridge University Press, 2006.

[66] D. M Cole, S. M. Smith, and C. F Beckmann, "Advances and pitfalls in the analysis and interpretation of resting-state FMRI data," *Frontiers in Systems Neuroscience*, vol. 4, no. 8, 2010.

[67] I. I. Rabi, J. R. Zacharias, S. Millman, and P. Kusch, “A new method of measuring nuclear magnetic moment,” *Physical Review*, vol. 53, no. 4, pp. 318, 1938.

[68] E. M. Purcell, H. C. Torrey, and R. V. Pound, “Resonance absorption by nuclear magnetic moments in a solid,” *Physical Review*, vol. 69, no. 1-2, pp. 37–38, 1946.

[69] F. Bloch, “Nuclear induction,” *Physical Review*, vol. 70, pp. 460–473, 1946.

[70] R. V. Damadian, “Tumor detection by nuclear magnetic resonance,” *Science*, vol. 171, pp. 1151–1153, 1971.

[71] P. C. Lauterbur, “Image formation by induced local interactions: examples employing nuclear magnetic resonance,” *Nature*, vol. 242, pp. 190–191, 1973.

[72] P. Mansfield and A. Maudsley, “Line scan proton spin imaging in biological structures by nuclear magnetic resonance,” *Physics in Medicine and Biology*, vol. 21, pp. 847–852, 1976.

[73] M. E. Haacke, R. W. Brown, M. R. Thompson, and R. Venkatesan, *Magnetic Resonance Imaging: Physical Principles and Sequence Design*, Wiley-Liss, 1999.

[74] M. A. Bernstein, K. F. King, and Z. J. Zhou, *Handbook of MRI pulse sequences*, Elsevier, Amsterdam, Boston, Heidelberg, 2004.

[75] K. J. Friston, A. Mechelli, R. Turner, and C. J. Price, “Nonlinear responses in fMRI: The balloon model, Volterra kernels, and other hemodynamics,” *NeuroImage*, vol. 12, no. 4, pp. 466 – 477, 2000.

[76] R. B. Buxton, E. C. Wong, and L. R. Frank, “Dynamics of blood flow and oxygenation changes during brain activation: The Balloon model.,” *Magnetic Resonance in Medicine*, vol. 39, no. 6, pp. 855–864, Jun 1998.

[77] D. Malonek and A. Grinvald, “Interactions between electrical activity and cortical microcirculation revealed by imaging spectroscopy: Implications for functional brain mapping,” *Science*, vol. 272, no. 5261, pp. 551–554, 1996.

[78] D. J. Heeger and D. Ress, “What does fMRI tell us about neuronal activity?,” *Nature Reviews Neuroscience*, vol. 3, pp. 142–151, 2002.

[79] G.D. Iannetti and Richard G. Wise, “BOLD fMRI in disease and pharmacological studies: room for improvement?,” *Magnetic Resonance Imaging*, vol. 25, no. 6, pp. 978 – 988, 2007.

[80] D.A. Feinberg, S. Moeller, S.M. Smith, E. Auerbach, S. Ramanna, M.F. Glasser, K.L. Miller, K. Ugurbil, and E. Yacoub, "Multiplexed echo planar imaging for sub-second whole brain fMRI and fast diffusion imaging.," *PLoS One*, vol. 5, no. 12, pp. e15710, 2010.

[81] K. J. Friston, S. Williams, R. Howard, R. S. J. Frackowiak, and R. Turner, "Movement-related effects in fMRI time-series," *Magnetic Resonance in Medicine*, vol. 35, no. 3, pp. 346–355, 1996.

[82] K. J. Friston, P. Fletchera, O. Josephs, A. Holmes, M. D. Rugg, and R. Turner, "Event-Related fMRI: Characterizing Differential Responses," *NeuroImage*, vol. 7, no. 1, pp. 30–40, Jan. 1998.

[83] T. E. Lund, K. H. Madsen, K. Sidaros, W. L. Luo, and T. E. Nichols, "Non-white noise in fMRI: Does modelling have an impact?," *NeuroImage*, vol. 29, no. 1, pp. 54–66, Jan 2006.

[84] G.K. Aguirre, E. Zarahn, and M. D'Esposito, "The variability of human, BOLD hemodynamic responses," *NeuroImage*, vol. 8, no. 4, pp. 360 – 369, 1998.

[85] D. A. Handwerker, J. M. Ollinger, and M. D'Esposito, "Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses," *NeuroImage*, vol. 21, no. 4, pp. 1639 – 1651, 2004.

[86] R.N.A Henson, "Chapter 40 - analysis of fMRI time series: Linear time-invariant models, event-related fMRI, and optimal experimental design," in *Human Brain Function (Second Edition)*, Richard S.J. Frackowiak, Karl J. Friston, Christopher D. Frith, Raymond J. Dolan, Cathy J. Price, Semir Zeki, John T. Ashburner, and William D. Penny, Eds., pp. 793 – 822. Academic Press, Burlington, second edition edition, 2004.

[87] A. M. Smith, B. K. Lewis, U. E. Ruttimann, Ye. F. Q., T. M. Sinnwell, Y. Yang, J. H. Duyn, and J. A. Frank, "Investigation of low frequency drift in fMRI signal," *NeuroImage*, vol. 9, no. 5, pp. 526 – 533, 1999.

[88] E. Zarahn, G.K. Aguirre, and M. D'Esposito, "Empirical analyses of BOLD fMRI statistics," *NeuroImage*, vol. 5, no. 3, pp. 179 – 197, 1997.

[89] P. L. Purdon, V. Solo, R. M. Weisskoff, and E. N. Brown, "Locally regularized spatiotemporal modeling and model comparison for fMRI," *NeuroImage*, vol. 14, no. 4, pp. 912 – 923, 2001.

[90] R. M. Birn and P.A. Bandettini, "The effect of stimulus duty cycle and "off" duration on BOLD response linearity," *NeuroImage*, vol. 27, pp. 70–82, 2005.

[91] A. L. Vazques and D. C. Noll, "Nonlinear aspects of the BOLD response in fMRI," *NeuroImage*, vol. 7, no. 108-118, 1998.

[92] J. B. Mandeville, J. J.A. Marota, C. Ayata, M. A. Moskowitz, R. M. Weisskoff, and B. R. Rosen, "MRI measurement of the temporal evolution of relative CMRO2 during rat forepaw stimulation," *Magnetic Resonance in Medicine*, vol. 42, no. 5, pp. 944–951, 1999.

[93] S Ogawa, R S Menon, D W Tank, S G Kim, H Merkle, J M Ellermann, and K Ugurbil, "Functional brain mapping by blood oxygenation level-dependent contrast magnetic resonance imaging. a comparison of signal characteristics with a biophysical model," *Journal of Biophysics*, vol. 64, no. 3, pp. 803–12, Mar 1993.

[94] J. L. Boxerman, P. A. Bandettini, K. K. Kwong, J. R. Baker, T. L. Davis, B. R. Rosen, and R. M. Weisskoff, "The intravascular contribution to fMRI signal change: Monte Carlo modeling and diffusion-weighted studies in vivo," *Magnetic Resonance in Medicine*, vol. 34, no. 1, pp. 4–10, 1995.

[95] R Buxton and L Frank, "A model for the coupling between cerebral blood flow and oxygen metabolism during neural stimulation," *Journal of Cerebral Blood Flow & Metabolism*, Jan 1997.

[96] N. Zhang, E. Yacoub, X.-H. Zhu, K. Ugurbil, and W. Chen, "Linearity of blood-oxygenation-level dependent signal at microvasculature," *NeuroImage*, vol. 48, no. 2, pp. 313 – 318, 2009.

[97] Z. Chen and V. Calhoun, "Volumetric BOLD fMRI simulation: from neurovascular coupling to multivoxel imaging," *BMC Med Imaging*, vol. 12, pp. 8, Jan 2012.

[98] O. Arthurs and S. Boniface, "How well do we understand the neural origins of the fMRI BOLD signal?," *Cell Trends in Neurosciences*, vol. 25, no. 1, pp. 27–31, Jan 2002.

[99] N. K. Logothetis, "The neural basis of the blood-oxygen-level-dependent functional magnetic resonance imaging signal.," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 357, no. 1424, pp. 1003, Aug 2002.

[100] R. D. Hoge, J. Atkinson, B. Gill, G. R. Crelier, S. Marrett, and G. B. Pike, "Linear coupling between cerebral blood flow and oxygen consumption in activated human cortex," *Proceedings of the National*

Academy of Sciences of the United States of America, vol. 96, no. 16, pp. 9403, Aug 1999.

[101] I. Khalidov, J. Fadili, F. Lazeyras, D. Van De Ville, and Michael Unser, "Activelets: Wavelets for sparse representation of hemodynamic responses," *Signal Processing*, vol. 91, no. 12, pp. 2810–2821, 2011.

[102] B. Biswal, F. Z. 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, Oct. 1995.

[103] D. Zhang and M. E. Raichle, "Disease and the brain's dark energy.," *Nature reviews. Neurology*, vol. 6, no. 1, pp. 15–28, January 2010.

[104] M. E. Raichle, A. M. MacLeod, A. Z. Snyder, W. J. Powers, D. A. Gusnard, and G. L. Shulman, "A default mode of brain function.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 2, pp. 676–682, January 2001.

[105] G. Shulman, J. Fiez, M. Corbetta, and R. Buckner, "Common blood flow changes across visual tasks: II. decreases in cerebral cortex," *Journal of Cognitive Neuroscience*, Jan 1997.

[106] M. D. Fox, A. Z. Snyder, J. L. Vincent, M. Corbetta, D. C. Van Essen, and M. E. Raichle, "The human brain is intrinsically organized into dynamic, anticorrelated functional networks," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 27, pp. 9673, 2005.

[107] M. E. Raichle, "The brain's dark energy," *Science*, vol. 314, no. 5803, pp. 1249–1250, 2006.

[108] S. Strother, "Evaluating fMRI preprocessing pipelines," *Engineering in Medicine and Biology Magazine*, Jan 2006.

[109] N. W. Churchill, A. Oder, H. Abdi, F. Tam, W. Lee, C. Thomas, J. E. Ween, S. J. Graham, and S. C. Strother, "Optimizing preprocessing and analysis pipelines for single-subject fMRI. I. standard temporal motion and physiological noise correction methods," *Human Brain Mapping*, vol. 33, no. 3, pp. 609–27, Mar 2012.

[110] G. H. Glover, T.-Q. Li, and D. Ress, "Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR," *Magnetic Resonance in Medicine*, vol. 44, no. 1, pp. 162–167, 2000.

[111] M. M. Monti, "Statistical analysis of fMRI time-series: A critical review of the glm approach," *Frontiers in Human Neuroscience*, vol. 5, no. 28, 2011.

[112] K.J. Worsley, C.H. Liao, J. Aston, V. Petre, G.H. Duncan, F. Morales, and A.C. Evans, "A general statistical analysis for fMRI data," *NeuroImage*, vol. 15, no. 1, pp. 1 – 15, 2002.

[113] K. Friston, J. Ashburner, A., S. Kiebel, T. Nichols, and W. Penny, *Statistical Parametric Mapping: The Analysis of Functional Brain Images*, Academic Press, London, 2007.

[114] C. R. Genovese, D. C. Noll, and W. F. Eddy, "Estimating test-retest reliability in functional MR imaging I: Statistical methodology," *Magnetic Resonance in Medicine*, vol. 38, no. 3, pp. 497–507, 1997.

[115] C. R. Genovese, N. A. Lazar, and T. Nichols, "Thresholding of statistical maps in functional neuroimaging using the false discovery rate," *NeuroImage*, vol. 15, no. 4, pp. 870 – 878, 2002.

[116] J. K. Worsley, "Detecting activation in fMRI data," *Statistical Methods in Medical Research*, vol. 12, no. 5, pp. 401–418, 2003.

[117] D. A. Gusnard and M. E. Raichle, "Searching for a baseline: Functional imaging and the resting human brain," *Nature Reviews Neuroscience*, vol. 2, pp. 685 – 694, 2001.

[118] R Baumgartner, L Ryner, W Richter, R Summers, M Jarmasz, and R Somorjai, "Comparison of two exploratory data analysis methods for fMRI: fuzzy clustering vs. principal component analysis," *Magnetic Resonance Imaging*, vol. 18, no. 1, pp. 89 – 94, 2000.

[119] Y. Liu, J.-H. Gao, H.-L. Liu, and P. T. Fox, "The temporal response of the brain after eating revealed by functional MRI," *Nature*, vol. 405, pp. 1058–1062, 2000.

[120] V. L. Morgan, Y. Li, B. Abou-Khalil, and J. C. Gore, "Development of 2dTCA for the detection of irregular, transient BOLD activity," *Human Brain Mapping*, vol. 29, no. 1, pp. 57–69, 2008.

[121] C. F. Beckmann and S. M. Smith, "Probabilistic independent component analysis for fMRI," *IEEE Transactions on Medical Imaging*, vol. 23, no. 2, pp. 137 –152, feb. 2004.

[122] V.D. Calhoun and T. Adali, "Unmixing fMRI with independent component analysis," *IEEE Engineering in Medicine and Biology Magazine*, vol. 25, no. 2, pp. 79 – 90, march-april 2006.

[123] A.R. McIntosh, F.L. Bookstein, J.V. Haxby, and C.L. Grady, "Spatial pattern analysis of functional brain images using partial least squares," *NeuroImage*, vol. 3, no. 3, pp. 143 – 157, 1996.

[124] A. Krishnan, L. J. Williams, A. R. McIntosh, and H. Abdi, "Partial least squares (PLS) methods for neuroimaging: A tutorial and review," *NeuroImage*, vol. 56, no. 2, pp. 455 – 475, 2011, Multivariate Decoding and Brain Reading.

[125] B. Afshin-Pour, G.-A. Hossein-Zadeh, S. C. Strother, and H. Soltanian-Zadeh, "Enhancing reproducibility of fMRI statistical maps using generalized canonical correlation analysis in NPAIRS framework," *NeuroImage*, vol. 60, no. 4, pp. 1970 – 1981, 2012.

[126] J.W. Evans, R.M. Todd, M.J. Taylor, and S.C. Strother, "Group specific optimisation of fMRI processing steps for child and adult data," *NeuroImage*, vol. 50, no. 2, pp. 479 – 490, 2010.

[127] M. D. Greicius, B. Krasnow, A. L. Reiss, and V. Menon, "Functional connectivity in the resting brain: a network analysis of the default mode hypothesis," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 1, pp. 253–8, Jan 2003.

[128] A. R. McIntosh, M. N. Rajah, and N. J. Lobaugh, "Interactions of prefrontal cortex in relation to awareness in sensory learning," *Science*, vol. 284, no. 5419, pp. 1531–1533, 1999.

[129] A. R. McIntosh and N. J. Lobaugh, "Partial least squares analysis of neuroimaging data: applications and advances," *NeuroImage*, vol. 23, Supplement 1, pp. S250 – S263, 2004, Mathematics in Brain Imaging.

[130] R. L. Buckner, J. R. Andrews-Hanna, and D. L. Schacter, "The brain's default network," *Annals of the New York Academy of Sciences*, vol. 1124, no. 1, pp. 1–38, 2008.

[131] J. V. Stone, *Independent Component Analysis and Projection Pursuit: A Tutorial Introduction*, The MIT Press, 2004.

[132] M. J. McKeown, S. Makeig, G. G. Brown, T. P. Jung, S. S. Kindermann, A. J. Bell, and T. J. Sejnowski, "Analysis of fMRI data by blind separation into independent spatial components," *Human Brain Mapping*, vol. 6, no. 3, pp. 160–88, Jan 1998.

[133] R. M. Birn, J. B. Diamond, M. A. Smith, and P. A. Bandettini, "Separating respiratory-variation-related fluctuations from neuronal-activity-related fluctuations in fMRI," *NeuroImage*, vol. 31, no. 4, pp. 1536 – 1548, 2006.

[134] A. J. Bell and T. J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural Computation*, vol. 7, no. 6, pp. 1129–1159, Nov. 1995.

[135] M. Gaeta and J.-L. Lacoume, "Source separation without a priori knowledge: The maximum likelihood solution," in *Proceedings of European Signal Processing Conference EUSIPCO*, 1990, pp. 621–624.

[136] A. Hyvärinen, "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Transactions on Neural Networks*, vol. 10, no. 3, pp. 626–634, 1999.

[137] J. Himberg, A. Hyvarinen, and F. Esposito, "Validating the independent components of neuroimaging time series via clustering and visualization," *NeuroImage*, vol. 22, no. 3, pp. 1214 – 1222, 2004.

[138] X-N. Zuo, C. Kelly, J. A. Adelstein, D. F. Klein, F. X. Castellanos, and M. P. Milham, "Reliable intrinsic connectivity networks: Test-retest evaluation using ICA and dual regression approach," *NeuroImage*, vol. 49, no. 3, pp. 2163 – 2177, 2010.

[139] V. Kiviniemi, T. Starck, J. Remes, X. Long, J. Nikkinen, M. Haapea, J. Veijola, I. Moilanen, M. Isohanni, Y.-F. Zang, and O. Tervonen, "Functional segmentation of the brain cortex using high model order group pica," *Human Brain Mapping*, vol. 30, no. 12, pp. 3865–3886, 2009.

[140] D.R. Gitelman, W.D. Penny, J. Ashburner, and K.J. Friston, "Modeling regional and psychophysiologic interactions in fMRI: the importance of hemodynamic deconvolution," *NeuroImage*, vol. 19, pp. 200–207, 2003.

[141] R. Lopes, J.M. Lina, F. Fahoum, and J. Gotman, "Detection of epileptic activity in fMRI without recording the EEG," *NeuroImage*, vol. 60, no. 3, pp. 1867 – 1879, 2012.

[142] C. Caballero-Gaudes, N. Petridou, S. T. Francis, V. Dryden, and P. A. Gowland, "Paradigm free mapping with sparse regression automatically detects single-trial fMRI BOLD responses," *Human Brain Mapping*, vol. 34, no. 3, pp. 501–518, 2013.

[143] C. Caballero-Gaudes, N. Petridou, I. L. Dryden, L. Bai, S. T. Francis, and P. A. Gowland, "Detection and characterization of single-trial fMRI BOLD responses: Paradigm free mapping," *Human Brain Mapping*, vol. 32, no. 9, pp. 1400–1418, 2011.

[144] L. Hernandez-Garcia and M. O. Ulfarsson, "Neuronal event detection in fMRI time series using iterative deconvolution techniques," *Magnetic Resonance Imaging*, vol. 29, no. 3, pp. 353 – 364, 2011.

[145] G.-R. Wu, W. Liao, S. Stramaglia, J.-R. Ding, H. Chen, and D. Marinazzo, "A blind deconvolution approach to recover effective connectivity brain networks from resting state fMRI data," *Medical Image Analysis*, vol. 17, no. 3, pp. 365–74, Apr 2013.

[146] K J Friston, L Harrison, and W Penny, "Dynamic causal modelling," *NeuroImage*, vol. 19, no. 4, pp. 1273–302, Aug 2003.

[147] J. J. Riera, J. Watanabe, I. Kazuki, M. Naoki, E. Aubert, T. Ozaki, and R. Kawashima, "A state-space model of the hemodynamic approach: nonlinear filtering of BOLD signals," *NeuroImage*, vol. 21, no. 2, pp. 547 – 567, 2004.

[148] K.J. Friston, N. Trujillo-Barreto, and J. Daunizeau, "DEM: A variational treatment of dynamic systems," *NeuroImage*, vol. 41, no. 3, pp. 849 – 885, 2008.

[149] K J Friston, "Variational filtering," *NeuroImage*, vol. 41, no. 3, pp. 747–66, Jul 2008.

[150] K. J. Friston, K. Stephan, and J. Daunizeau, "Generalised filtering," *Mathematical Problems in Engineering*, 2010.

[151] M. Havlicek, K. J. Friston, J. Jan, M. Brazdil, and V. D. Calhoun, "Dynamic modeling of neuronal responses in fMRI using cubature kalman filtering," *NeuroImage*, vol. 56, no. 4, pp. 2109 – 2128, 2011.

[152] K. Bush and J. Cisler, "Decoding neural events from fMRI BOLD signal: A comparison of existing approaches and development of a new algorithm," *Magnetic Resonance Imaging*, Apr 2013.

[153] C Goutte, F A Nielsen, and L K Hansen, "Modeling the haemodynamic response in fMRI using smooth FIR filters," *IEEE Transaction on Medical Imaging*, vol. 19, no. 12, pp. 1188–201, Dec 2000.

[154] P. Ciuciu, J.-B. Poline, G. Marrelec, J. Idier, C. Pallier, and H. Benali, "Unsupervised robust nonparametric estimation of the hemodynamic response function for any fMRI experiment," *IEEE Transaction on Medical Imaging*, vol. 22, no. 10, pp. 1235–51, Oct 2003.

[155] T. D. Wager, A. Vazquez, L. Hernandez, and D. C. Noll, "Accounting for nonlinear BOLD effects in fMRI: parameter estimates and a model for prediction in rapid event-related studies," *NeuroImage*, vol. 25, no. 1, pp. 206–18, Mar 2005.

[156] M. W. Woolrich, M. Jenkinson, J. M. Brady, and S. M. Smith, "Fully bayesian spatio-temporal modeling of fMRI data," *IEEE Transactions on Medical Imaging*, vol. 23, no. 2, pp. 213–31, Feb 2004.

[157] T. Vincent, L. Risser, and P. Ciuciu, "Spatially adaptive mixture modeling for analysis of fMRI time series," *IEEE Transactions on Medical Imaging*, vol. 29, no. 4, pp. 1059 –1074, april 2010.

[158] L. Chaari, T. Vincent, F. Forbes, M. Dojat, and P. Ciuciu, "Fast joint detection-estimation of evoked brain activity in event-related fMRI using a variational approach," *IEEE Transactions on Medical Imaging*, vol. PP, no. 99, pp. 1, 2013.

[159] S. Makni, J. Idier, T. Vincent, B. Thirion, G. Dehaene-Lambertz, and P. Ciuciu, "A fully Bayesian approach to the parcel-based detection-estimation of brain activity in fMRI," *NeuroImage*, vol. 41, no. 3, pp. 941 – 969, 2008.

[160] L. Chaari, F. Forbes, T. Vincent, and P. Ciuciu, "Adaptive hemodynamic-informed parcellation of fMRI data in a variational joint detection estimation framework," in *Proceedings of Medical Image Computing and Computer Assisted Intervention MICCAI, LNCS Springer Verlag*, Oct. 2012, vol. 7512, pp. 180–188.

[161] B. Thirion, G. Flandin, P. Pinel, A. Roche, P. Ciuciu, and J.-B. Poline, "Dealing with the shortcomings of spatial normalization: Multi-subject parcellation of fMRI datasets," *Human Brain Mapping*, vol. 27, no. 8, pp. 678–693, 2006.

[162] F. I. Karahanoglu, C. Caballero-Gaudes, F. Lazeyras, and D. Van De Ville, "Total activation: fMRI deconvolution through spatio-temporal regularization," *NeuroImage*, vol. 73, pp. 121 – 134, 2013.

[163] A. Kirsch, *An introduction to the mathematical theory of inverse problems*, vol. 120 of *Applied Mathematical Sciences*, Springer, 2nd edition, 1996.

[164] M. Zibulevsky and M. Elad, "L1-l2 optimization in signal and image processing," *IEEE Signal Processing Magazine*, vol. 27, no. 3, pp. 76 – 88, 2010.

[165] M. Yuan and Y. Lin, "Model selection and estimation in regression with grouped variables," *Journal of the Royal Statististical Society*, vol. 68, pp. 49–67, 2006.

[166] D. L. Donoho and J. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.

[167] M. W. Schmidt, N. Le Roux, and F. Bach, "Convergence rates of inexact proximal-gradient methods for convex optimization," in *Advances in Neural Information Processing Systems (NIPS)*, 2011, pp. 1458–1466.

[168] F. I. Karahanoglu, F. Grouiller, C. Caballero-Gaudes, M. Seeck, S. Vulliemoz, and D. Van De Ville, "Spatial mapping of interictal epileptic discharges in fMRI with total activation," in *Proceedings of IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2013 (ISBI '13)*, 2013, pp. 1492–1495.

[169] M. D. Fox and M. E. Raichle, "Spontaneous fluctuations in brain activity observed with fMRI," *Nature Review Neuroscience*, vol. 8, no. 9, pp. 700–711, 2007.

[170] R. S. Menon, D. C. Luknowsky, and J. S. Gati, "Mental chronometry using latency-resolved fMRI," *Proceedings of the National Academy of Sciences*, vol. 95, no. 18, pp. 10902–10907, 1998.

[171] D. Gonzalo, T. Shallice, and R. Dolan, "Time-dependent changes in learning audiovisual associations: A single-trial fMRI study," *NeuroImage*, vol. 11, no. 3, pp. 243 – 255, 2000.

[172] H. Laufs, A. Kleinschmidt, A. Beyerle, E. Eger, A. Salek-Haddadi, C. Preibisch, and K. Krakow, "Eeg-correlated fMRI of human alpha activity," *NeuroImage*, vol. 19, no. 4, pp. 1463 – 1476, 2003.

[173] A. Salek-Haddadi, B. Diehl, K. Hamandi, M. Merschhemke, A. Liston, K. Friston, J. S. Duncan, D. R. Fish, and L. Lemieux, "Hemodynamic correlates of epileptiform discharges: An EEG-fMRI study of 63 patients with focal epilepsy," *Brain Research*, vol. 1088, no. 1, pp. 148 – 166, 2006.

[174] K. Friston, "Dynamic causal modeling and granger causality comments on: the identification of interacting networks in the brain using fMRI: model selection, causality and deconvolution," *NeuroImage*, vol. 58, no. 2, pp. 303–5; author reply 310–1, Sep 2011.

[175] P. A. Valdes-Sosa, A. Roebroeck, J. Daunizeau, and K. Friston, "Effective connectivity: influence, causality and biophysical modeling," *NeuroImage*, vol. 58, no. 2, pp. 339–61, Sep 2011.

[176] A. Roebroeck, E. Formisano, and R. Goebel, "The identification of interacting networks in the brain using fMRI: Model selection, causality and deconvolution," *NeuroImage*, vol. 58, no. 2, pp. 296–302, Sep 2011.

[177] S. M. Smith, K. L. Miller, G. Salimi-Khorshidi, M. Webster, C. F. Beckmann, T. E. Nichols, J. D. Ramsey, and M. W. Woolrich, "Network modelling methods for fMRI," *NeuroImage*, Sep 2010.

[178] O. David, I. Guillemain, S. Saillet, S. Reyt, C. Deransart, C. Segebarth, and A. Depaulis, "Identifying neural drivers with functional MRI: an electrophysiological validation," *PLoS Biol*, vol. 6, no. 12, pp. 2683–97, Dec 2008.

[179] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, and M. Loliot, "Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain," *NeuroImage*, vol. 15, pp. 273–289, 2002.

[180] Y. Alemán-Gómez, L. Melie-Garćia, and P. Valdés-Hernandez, "IBASPM: Toolbox for automatic parcellation of brain structures," *12th Annual Meeting of the Organization for Human Brain Mapping*, vol. 27, 2006.

[181] M. D. Fox, M. Corbetta, A. Z. Snyder, J. L. Vincent, and M. E. Raichle, "Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems," *Proceedings of the National Academy of Sciences*, vol. 103, no. 26, pp. 10046–10051, 2006.

[182] M. Corbetta and G. L. Shulman, "Control of goal-directed and stimulus-driven attention in the brain," *Nature Reviews Neuroscience*, vol. 3, pp. 201–215, 2002.

[183] C. F. Beckmann, M. DeLuca, J. T. Devlin, and S. M. Smith, "Investigations into resting-state connectivity using independent component analysis," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 360, no. 1457, pp. 1001–1013, 2005.

[184] J. S. Damoiseaux, S. A. R. B. Rombouts, F. Barkhof, P. Scheltens, C. J. Stam, S. M. Smith, and C. F. Beckmann, "Consistent resting-state networks across healthy subjects," *Proceedings of the National Academy of Sciences*, vol. 103, no. 37, pp. 13848–13853, 2006.

[185] M. De Luca, C.F. Beckmann, N. De Stefano, P.M. Matthews, and S.M. Smith, "fMRI resting state networks define distinct modes of long-distance interactions in the human brain," *NeuroImage*, vol. 29, no. 4, pp. 1359 – 1367, 2006.

[186] W. W. Seeley, V. Menon, A. F. Schatzberg, J. Keller, G. H. Glover, H. Kenna, A. L. Reiss, and M. D. Greicius, "Dissociable intrinsic connectivity networks for salience processing and executive control," *The Journal of Neuroscience*, vol. 27, no. 9, pp. 2349–2356, 2007.

[187] D. Van De Ville, P. Jhooti, T. Haas, R. Kopel, K.-O. Lovblad, K. Scheffler, and S. Haller, "Recovery and spatio-temporal segregation of default mode subnetworks," *NeuroImage*, vol. 63, no. 4, pp. 1175–1781, 2012.

[188] A. D. (Bud) Craig, "How do you feel–now? the anterior insula and human awareness," *Nature Reviews Neuroscience*, vol. 10, pp. 59–71, 2009.

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

[190] S. M. Smith, K. L. Miller, S. Moeller, J. Xu, E. J. Auerbach, M. W. Woolrich, C. F. Beckmann, M. Jenkinson, J. Andersson, M. F. Glasser, D. C. Van Essen, D. A. Feinberg, E. S. Yacoub, and K. Ugurbil, "Temporally-independent functional modes of spontaneous brain activity," *Proceedings of the National Academy of Sciences*, vol. 8, pp. 3131–3136, 2012.

[191] J. Gotman, E. Kobayashi, A. P. Bagshaw, C.-G. Bénar, and F. Dubeau, "Combining EEG and fMRI: A multimodal tool for epilepsy research," *Journal of Magnetic Resonance Imaging*, vol. 23, no. 6, pp. 906–920, 2006.

[192] J. W. Miller and J. Gotman, "The meaning of interictal spikes in temporal lobe epilepsy: Should we count them?," *Neurology*, vol. 71, no. 6, pp. 392–393, 2008.

[193] S. Vulliemoz, R. Rodionov, D.W. Carmichael, R. Thornton, M. Guye, S.D. Lhatoo, C.M. Michel, J.S. Duncan, and L. Lemieux, "Continuous EEG source imaging enhances analysis of EEG-fMRI in focal epilepsy," *NeuroImage*, vol. 49, no. 4, pp. 3219 – 3229, 2010.

[194] S. Vulliemoz, R. Thornton, R. Rodionov, D.W. Carmichael, M. Guye, S. Lhatoo, A.W. McEvoy, L. Spinelli, C.M. Michel, J.S. Duncan, and L. Lemieux, "The spatio-temporal mapping of epileptic networks: Combination of EEG-fMRI and EEG source imaging," *NeuroImage*, vol. 46, no. 3, pp. 834 – 843, 2009.

[195] K. Groening, V. Brodbeck, F. Moeller, S. Wolff, A. Van Baalen, C. M. Michel, O. Jansen, R. Boor, G. Wiegand, U. Stephani, and M. Siniatchkin, "Combination of EEG-fMRI and EEG source analysis improves interpretation of spike-associated activation networks in paediatric pharmacoresistant focal epilepsies," *NeuroImage*, vol. 46, no. 3, pp. 827 – 833, 2009.

[196] J. S. Duncan, "Imaging in the surgical treatment of epilepsy," *Nature Review Neurology*, vol. 6, pp. 537–550, 2010.

[197] Jean Gotman, "Epileptic networks studied with EEG-fMRI," *Epilepsia*, vol. 49, pp. 42–51, 2008.

[198] F. Grouiller, R.C. Thornton, K. Groening, L. Spinelli, J.S. Duncan, K. Schaller, M. Siniatchkin, L. Lemieux, M. Seeck, C.M. Michel, and S. Vulliemoz, "With or without spikes: localization of focal epileptic activity by simultaneous electroencephalography and fMRI," *Brain*, vol. 134, no. 10, pp. 2867–2886, 2011.

[199] J. X. Tao, A. Ray, S. Hawes-Ebersole, and J. S. Ebersole, "Intracranial EEG substrates of scalp EEG interictal spikes," *Epilepsia*, vol. 46, no. 5, pp. 669–676, 2005.

[200] S. Vulliemoz, D. W. Carmichael, K. Rosenkranz, B. Diehl, R. Rodionov, M. C. Walker, A. W. McEvoy, and L. Lemieux, "Simultaneous intracranial EEG and fMRI of interictal epileptic discharges in humans," *NeuroImage*, vol. 54, no. 1, pp. 182 – 190, 2011.

[201] F Lazeyras, O Blanke, and S Perrig, "EEG-triggered fMRI in patients with pharmacoresistant epilepsy," *Journal of Magentic Resonance Imaging*, Jan 2000.

[202] M. Zijlmans, G. Huiskamp, M. Hersevoort, J.-H. Seppenwoolde, A. C. van Huffelen, and F. S. S. Leijten, "EEG-fMRI in the preoperative work-up for epilepsy surgery," *Brain*, vol. 130, no. 9, pp. 2343–2353, 2007.

[203] R. Thornton, H. Laufs, R. Rodionov, S. Cannadathu, D. W. Carmichael, S. Vulliemoz, A. Salek-Haddadi, A. W. McEvoy, S. M. Smith, S. Lhatoo, R. D. C. Elwes, Maxime Guye, M. C. Walker, L. Lemieux, and J. S. Duncan, "EEG correlated functional MRI and postoperative outcome in focal epilepsy," *Journal of Neurology, Neurosurgery and Psychiatry*, vol. 81, no. 8, pp. 922–927, 2010.

[204] D. Liston A, J. C. De Munck, K. Hamandi, H. Laufs, P. Ossenblok, J. S. Duncan, and L. Lemieux, "Analysis of EEG-fMRI data in focal epilepsy based on automated spike classification and signal space projection," *NeuroImage*, vol. 31, no. 3, pp. 1015 – 1024, 2006.

[205] R. Rodionov, F. De Martino, H. Laufs, D. W. Carmichael, E. Formisano, M. Walker, J. S. Duncan, and L. Lemieux, "Independent component analysis of interictal fMRI in focal epilepsy: comparison with general linear model-based EEG-correlated fMRI," *NeuroImage*, vol. 38, no. 3, pp. 488–500, Nov 2007.

[206] P. LeVan and J. Gotman, "Independent component analysis as a model-free approach for the detection of BOLD changes related to epileptic spikes: a simulation study," *Human Brain Mapping*, vol. 30, no. 7, pp. 2021–31, Jul 2009.

[207] Pierre LeVan, Louise Tyvaert, Friederike Moeller, and Jean Gotman, "Independent component analysis reveals dynamic ictal BOLD responses in EEG-fMRI data from focal epilepsy patients," *NeuroImage*, vol. 49, no. 1, pp. 366–78, Jan 2010.

[208] F. Moeller, P. LeVan, and J. Gotman, "Independent component analysis (ica) of generalized spike wave discharges in fMRI: Comparison with general linear model-based EEG-fMRI," *Human Brain Mapping*, vol. 32, no. 2, pp. 209–217, 2011.

[209] C. Caballero-Gaudes, D. Van De Ville, F. Grouiller, R. Thornton, L. Lemieux, M. Seeck, F. Lazeyras, and S. Vulliemoz, "Mapping interictal epileptic discharges using mutual information between concurrent EEG and fMRI," *NeuroImage*, vol. 68, pp. 248–62, Mar 2013.

[210] P. J. Allen, O. Josephs, and R. Turner, "A method for removing imaging artifact from continuous EEG recorded during functional MRI," *NeuroImage*, vol. 12, no. 2, pp. 230 – 239, 2000.

[211] C. Michel, G. Lantz, L. Spinelli, R. Peralta, T. Landis, and M. Seeck, "128-channel EEG source imaging in epilepsy: Clinical yield and localization precision," *Journal of Clinical Neurophysiology*, vol. 21, no. 2, 2004.

[212] T. E. Nichols and A. P. Holmes, "Nonparametric permutation tests for functional neuroimaging: A primer with examples," *Human Brain Mapping*, vol. 15, no. 1, pp. 1–25, 2002.

[213] M. Unser and P.D. Tafti, "Stochastic models for sparse and piecewise-smooth signals," *IEEE Transactions on Signal Processing*, vol. 59, no. 3, pp. 989–1006, Mar 2011.

[214] E. Bostan, U.S. Kamilov, M. Nilchian, and M. Unser, "Sparse stochastic processes and discretization of linear inverse problems," *IEEE Transactions on Image Processing*, vol. 22, no. 7, pp. 2699–2710, 2013.

[215] S. Hawe, M. Kleinsteuber, and K. Diepold, "Analysis operator learning and its application to image reconstruction," *IEEE Transactions on Image Processing*, vol. 22, no. 6, pp. 2138–2150, 2013.

[216] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *Signal Processing, IEEE Transactions on*, vol. 50, no. 6, pp. 1417–1428, 2002.

[217] P. Bellec, P. Rosa-Neto, O. C. Lyttelton, H. Benali, and A. C. Evans, "Multi-level bootstrap analysis of stable clusters in resting-state fMRI," *NeuroImage*, vol. 51, no. 3, pp. 1126 – 1139, 2010.

[218] E. A. Allen, E. Damaraju, S. M Plis, E. B. Erhardt, T. Eichele, and V. D. Calhoun, "Tracking whole-brain connectivity dynamics in the resting state," *Cerebral Cortex*, Nov 2012.

[219] V. D. Calhoun, T. Eichele, T. Adali, and E. A. Allen, "Decomposing the brain: components and modes, networks and nodes," *Cell Trends in Cognitive Sciences*, vol. 16, no. 5, pp. 255–6, May 2012.

[220] P. Ciuciu, G. Varoquaux, P. Abry, S. Sadaghiani, and A. Kleinschmidt, "Scale-free and multifractal time dynamics of fMRI signals during rest and task," *Frontiers in Fractal Physiology*, vol. 3, pp. 186, Jun 2012.

[221] M. Unser and T. Blu, "Generalized smoothing splines and the optimal discretization of the Wiener filter," *IEEE Transactions on Signal Processing*, vol. 53, no. 6, pp. 2146–2159, June 2005.

Fikret Işık Karahanoğlu

CONTACT INFORMATION

Medical Image Processing Laboratory — office: +41 21 693 7219
Faculty of Engineering, Institute of Bioengineering — *e-mail:* `isik.karahanoglu@epfl.ch`
Ecole Polytechnique Fédérale de Lausanne (EPFL) — *www:* `http://miplab.epfl.ch/karahanoglu/`

EDUCATION

Ecole Polytechnique Fédérale de Lausanne, Switzerland
Ph.D. Electrical Engineering — 09/2009 – 09/2013
Advisor: Prof. Dimitri Van De Ville

Ecole Polytechnique Fédérale de Lausanne, Switzerland
M.Sc. Communication Systems — 09/2007 – 07/2009
Advisor: Prof. Jean-Philippe Thiran
Specialization: Signals & Images

Middle East Technical University (METU), Turkey
B.S. Electrical & Electronics Engineering — 09/2003 – 06/2007

ACADEMIC EXPERIENCE

EPFL & University of Geneva (UNIGE), Switzerland

Research Assistant — **09/2009 – present**

- Joint position with UNIGE, Department of Radiology and Medical Informatics
- Experience on the development of sparse signal processing methods and its applications on functional magnetic resonance imaging
- Supervision of semester and master project students

EPFL, Switzerland

Master Thesis Project & Internship — **02/2009 – 09/2009**

- Master Thesis: Multimodal Image Fusion for Eye Radiotheraphy
- Worked on fusion of Computerized Tomography and Eye Ultrasound, and segmentation of Retinablastoma.

EPFL, Switzerland

Semester Project — **09/2008 – 01/2009**

- Project Focus: Design of an analytical phantom for parallel magnetic resonance imaging
- Worked and gained expertise on the analytical formulation of ground-truth k-space data of hearth phantom for parallel magnetic resonance imaging

HONOURS AND AWARDS

Neuroimage Journal Editors' Choice Award (in Methods and Modelling), 2013

Swiss Society of Neuroscience Travel Award, 2013

EPFL, Excellence Scholarship, 2007-2009

METU, Senior Design Project Award, 2007

METU, High Honor Rolls, 2003-2007

PUBLICATIONS

F.I. Karahanoğlu C. Caballero-Gaudes, F. Lazeyras, D. Van De Ville, *Total Activation: FMRI Deconvolution through Spatio-Temporal Regularization*, Neuroimage, vol 73, June 2013, pp. 121-134, June 2013. (Editors' choice award)

F.I. Karahanoğlu, I. Bayram, D. Van De Ville, *A Signal Processing Approach to Generalized 1D Total Variation*, IEEE Transactions on Signal Processing, vol.59, no.11, pp. 5265-5274, Nov 2011.

M. Bach Cuadra, S. Gorthi, **F.I. Karahanoğlu**, B. Paquier, A. Pica, H.P. Do, A. Balmer, F. Munier, J.-Ph. Thiran, *Model-based Segmentation and Fusion of 3D Computed Tomography and 3D Ultrasound of the Eye for Radiotherapy Planning*, Advances in Computational Vision and Medical Image Processing Methods and Applications: Computational Methods in Applied Sciences, Springer, 2011.

F.I. Karahanoğlu, F. Grouiller, C. Caballero Gaudes, M. Seeck, S. Vulliemoz, D. Van De Ville, *Spatial Mapping of Interictal Epileptic Discharges in fMRI with Total Activation*, Proceedings of the Tenth IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI'13), pp. 1492-1495.

C. Caballero Gaudes, **F.I. Karahanoğlu**, F. Lazeyras, D. Van de Ville, *Structured Sparse Deconvolution for Paradigm Free Mapping of Functional MRI Data*, (ISBI'12), Barcelona, pp. 322-325, May 2012.

M. Guerquin-Kern, **F.I. Karahanoğlu**, D. Van De Ville, K.P. Pruessmann, M. Unser, *Analytical Form of Shepp-Logan Phantom for Parallel MRI*, (ISBI'10), Rotterdam, The Netherlands, pp. 261-264, April 2010.

M. Bach Cuadra, S. Gorthi, **F.I. Karahanoğlu**, F. Salvador, A. Pica, H.P. Do, A. Balmer, F. Munier and J.- P. Thiran, *Model-based Segmentation and Image Fusion of 3D Computed Tomography and 3D Ultrasound of the Eye for Radiotherapy Planning*, Second ECCOMAS Thematic Conference on Computational Vision and Medical Image Processing, pp. 53-58, Sept 2009.

Abstracts and Talks

F.I. Karahanoğlu, F. Grouiller, C. Caballero Gaudes, M. Seeck, S. Vulliemoz, D. Van De Ville, *Localizing Sources of Interictal Epileptic Discharges using Total Activation Regularized BOLD FMRI*, Organization of Human Brain Mapping (HBM), 2013.

F.I. Karahanoğlu, *Total Activation: Recovery of fMRI Activity-Inducing Signals based on Sparsity Constraints*, Biomedical Imaging Center- Brain and Behavior Lab Research Day, Geneva, 2013.

F.I. Karahanoğlu, C. Caballero Gaudes, D. Van de Ville, *Total Activation: A New Spatiotemporal Regularization Technique to Analyze Spontaneous Brain Activity*, HBM, 2012.

F.I. Karahanoğlu, G. Cuendet, J. Britz, D. Van de Ville, C. Michel, *Multidimensional Random Walk Embedding to Reveal the EEG Microstates Dynamics*, Alpine Brain Imaging Meeting (ABIM), 2012.

F.I. Karahanoğlu, D. Van de Ville, *Generalizing Total Variation for Linear Differential Operators*, February Fourier Talks (FFT), 2011.

F.I. Karahanoğlu, C. Caballero Gaudes, D. Van de Ville, *Total-Activation Brain Mapping Reveal Brain Activity Without Paradigm Knowledge*, Alpine Brain Imaging Meeting (ABIM), 2011.

Teaching Experience

EPFL, *Image Processing I-II* 2010-2013
EPFL, *Signal Processing for Communications* 2009
EPFL, *Information Sciences* 2008