



AALBORG UNIVERSITY
STUDENT REPORT

Does Task-related FMRI Preprocessing Need a FIX?

3nd semester Masters, Biomedical
Engineering & Infomatics - Fall 2018

Project group: 19gr9411

Christian Korfitz Mortensen, Martin Alexander Garenfeld

Preface

Morbi luctus, wisi viverra faucibus pretium, nibh est placerat odio, nec commodo wisi enim eget quam. Quisque libero justo, consectetur a, feugiat vitae, porttitor eu, libero. Suspendisse sed mauris vitae elit sollicitudin malesuada. Maecenas ultricies eros sit amet ante. Ut venenatis velit. Maecenas sed mi eget dui varius euismod. Phasellus aliquet volutpat odio. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Pellentesque sit amet pede ac sem eleifend consectetur. Nullam elementum, urna vel imperdiet sodales, elit ipsum pharetra ligula, ac pretium ante justo a nulla. Curabitur tristique arcu eu metus. Vestibulum lectus. Proin mauris. Proin eu nunc eu urna hendrerit faucibus. Aliquam auctor, pede consequat laoreet varius, eros tellus scelerisque quam, pellentesque hendrerit ipsum dolor sed augue. Nulla nec lacus.

Contents

Del I Problem Analysis	1
1 Introduction	2
2 Problem Analysis	3
2.1 Processing of Pain in the Nervous System	3
2.2 Metabolic Response to Brain Activation	5
2.3 Imaging of Brain Activation	6
2.4 Individual differences in pain perception and fMRI	7
2.5 Artifacts Associated with BOLD fMRI	8
2.6 State of the Art in MRI Preprocessing	9
3 Study Objective	12
Del II Problem Solution	13
4 Methods	14
4.1 Study Protocol	14
4.2 Brain Extraction and Registration	17
4.3 Dataset Structuring	18
4.4 Standard Preprocessing	19
4.5 FSL FIX	21
4.6 Intensity normalization	39
4.7 Statistical analysis	39
Del III Results	41
Del IV Synthesis	42
A Appendices	49
A.1 MRI physics	49
A.2 MR image reconstruction	52
A.3 Stimuli Design	54
A.4 Independent Component Analysis	55
A.5 Principal Component Analysis	58
A.6 General Linear Model	60

Part I

Problem Analysis

1 | Introduction

To experience and live with chronic pain can be immensely debilitating for any individual, and the consequences of experiencing chronic pain has been linked to numerous physical and mental conditions: restrictions in mobility and daily activities, dependency of therapeutic drugs, anxiety and depression and a reduction in quality of life [1, 2].

More than 100 million Americans are estimated to live with chronic pain, manifesting an extraordinary amount of human suffering along with additional large economic societal expenditures, as cost of medical care and loss of wages and productivity are escalating [3, 4]. Compared to other major health conditions the number of patients suffering from chronic pain outnumber the combined total of patients suffering from heart disease and stroke, diabetes and cancer, and these conditions often include pain as a contributing component [2].

“An unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage”, is how the International Association for the Study of Pain defines pain [5]. The phenomenon of pain is a very complex composition of both psychological factors, personality traits and states, and cognitive, emotional, motivational, contextual and cultural variables. Pain can furthermore be defined either acute or chronic. Acute pain is a healthy response making up a self defense mechanism to warn and protect the human organism from any potential or further harm. [6, 7, 4] Contrary to acute pain, chronic pain is not linked to organ damage of any kind, and therefore does not serve any useful purpose to the human organism [8].

It still remains incredibly challenging to assess and treat pain, as the mechanisms during the experience of pain is not yet fully understood [9, 10]. This includes all neurological steps in the processing of pain - from the peripheral detection of a stimulus to the spinal cord transmission and brain processing, including the further relay of the pain experience [11]. A step towards getting a higher understanding of chronic pain is to examine the mechanisms involved in acute pain. A justification for this reasoning is that chronic pain often is transitioned from repetitive acute nociceptive stimulations that leads to neuroplastic changes. [11, 12]

Multiple studies [13, 14, 15, 4] have shown that the perception and sensitivity to acute pain are disposed to a great extent of subjective variability, suggesting that a pivotal key in understating the brain mechanism during noxious stimuli is found in the individual differences between subjects. Additionally, an individual’s sensitivity can vary substantially from day to day, despite being exposed to the same stimuli, and the psychophysical rating given by the patient can rather looked upon as artifact in the various scales used for reporting, than actually reflecting an unbiased measure of the experienced pain. It can furthermore be difficult for the physician to evaluate the patient’s experienced pain from a third-person perspective. [13]

To examine if there is a difference between individuals in brain activation during the experience of pain and if this activation is correlated with the self-reported experience of pain is therefore of great interest. [13]

2 | Problem Analysis

In the problem analysis chapter a theoretical overview of essential topics encapsulating the study of pain will be presented. First, the mechanisms involved in the experience of pain will be described, and how the metabolic response of brain activation can be translated into images. Subsequently, differences in the experience of pain between individuals will be outlined, along with the possible difficulties in interpreting the corresponding images. Lastly, different methods to counter these difficulties will be presented, thereby laying the foundation for understanding the problems faced when studying the differences in the individual's experience of pain.

2.1 Processing of Pain in the Nervous System

Pain is a complex subjective phenomenon, where sensory and affective components along with cognitive elements are activated to refrain an individual from further and future similar damage [4].

When the body is exposed to noxious stimuli, whether it originates from external or internal afflictions, information about the damaging stimuli is transmitted through neural pathways and through the peripheral nervous system to the autonomic and central nervous system. This relay to the brain of a damaging event is called nociception. It is mediated by certain receptors called nociceptors, which are attached to myelinated A δ and unmyelinated C fibers that terminate at the dorsal horn of the spinal cord. Perception of pain occurs when stimulation of nociceptors is high enough to activate A δ fibers, which results in an acute experience of prickling pain. When the stimulation increases, C fibers are activated, which results in a more intense pain experience that remains after the stimulus has ceased. These two phases of pain experience are associated with acute noxious stimulus, where the first phase is referred to as fast pain, and is of moderate intensity and appears immediately after the stimulus. The second phase, which is known as slow pain, is not as localizable, appears after a longer delay and is more painful. Activation of nociceptors can happen from sufficiently intense mechanical, chemical or thermal stimulation. Nociceptor activation is also modulated by inflammatory and bio-molecular influences. However, it is also possible for individuals to experience pain without a measurable noxious stimulus and for individuals to undergo a damaging trauma without suffering any pain sensation. [7]

As stated in The International Association for the Study of Pain's definition of pain, not only somatosensory components are involved in the experience of pain, but also cognitive and emotional elements. Psychological processes are major elements in the perception and expression of pain and involve the individual's attention to the painful infliction, the cognitive appraisal of pain and emotional and behavioral reactions. These factors may have increasing or decreasing effects on the perception of pain.

As mentioned, the activation of nociceptors is transmitted along axons in peripheral nerves to the dorsal horn of the spinal cord. From here the information is sent up the spinal cord

via the spinothalamic tract to the thalamus. The thalamus functions as a major transmission location for sensory information to the cerebral cortex. The neural pathways terminate in subdivisions of the thalamic nuclei, from where the information is relayed to different cortical and subcortical divisions, including regions of the cerebral cortex, amygdala, hypothalamus, periaqueductal grey and basal ganglia. Worth remarking is that the insula and anterior cingulate cortex seem to almost always be activated when nociceptors are stimulated. [16, 4] Figure 2.1 depicts the anatomy of pain pathway propagation of neurological signal from nociceptor to spinal cord to the brain centers affiliated with pain reception and modulation.

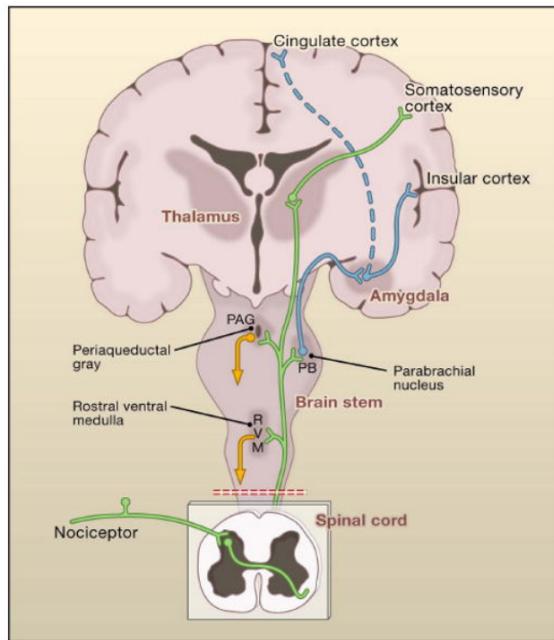


Figure 2.1: Coronal view of the mechanisms and regions associated with the experience of pain. The pain pathway from nociceptive input propagating through the dorsal horn continuing up through spinal cord to the affiliated brain regions. [17]

These structures are what process somatosensory input and is outputting neural activation that affects nociception and pain perception. [7] This means that the brain does not only passively receive sensory input, but in addition actively regulates the sensory transmission by influencing the dorsal horn of the spinal cord through a descending modulation. The brain structures involved in this regulation include the same as mentioned along with the rostral ventromedial medulla and dorsolateral pons/tegmentum. [16] Thus, when determining the intensity of pain perception during noxious stimulus of an individual these brain structures are of great importance to examine.

2.2 Metabolic Response to Brain Activation

To further understand the origin of brain activation during the experience of pain and how this activity is modulated by the brain, a measure of brain activation is therefore needed. To get a measure of brain activation, the underlying physiologic response is therefore of great importance to understand.

The activation of a brain region starts with a neurological input containing information about the noxious stimuli [16]. The increased neurological activity effects local metabolism as processing of the signal requires adenosine triphosphate (ATP) consumption during e.g. the reception and reformation of the action potential. Thus, ATP starts to be processed, leading to a decrease in oxygen concentration and increase in waste products. Thereby, the metabolic need for oxygen increases. Subsequently, these factors present in the local tissue of the corresponding brain region activate a vasodilation, increasing the blood flow to that region to reestablish the local homeostasis. During this regulation a not yet fully understood phenomenon occurs as more oxygenated blood than needed to compensate for the offset is delivered. Thereby flooding the local region with oxygenated blood. This response to the increased neural activity is known as the hemodynamic response. The overall increase in neural activity in that specific region following the need for metabolic regulation thereby permits the measure of the hemodynamic response, hence becoming an indirect measure of the neural activity. An example illustrating the hemodynamic response can be seen in figure 2.2. [18, 19].

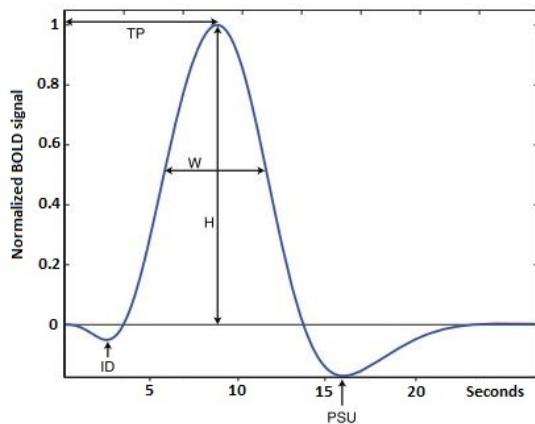


Figure 2.2: A depiction of a single hemodynamic response curve. ID is the initial dip as less oxygen will be present as the metabolic demand increases, TP is time from stimulus until peak, W and H is the width and height of the response and PSU is a post stimulus undershoot. Modified from [19]

Figure 2.2 should be considered the perfect noiseless hemodynamic response curve, to a brief stimuli, though the reality is that the response is effected by physiologic noise parameters and would be delayed in time compared to the stimuli onset. The peak height of the curve is commonly the most interesting feature of the response, as it portrays the amount of neural activity. The time to peak will occur 4-6 seconds after stimulus onset.

The duration of a response is around 20 seconds. There will further be a noticeable initial dip of 1-2 seconds duration, when the initial oxygen reserves are used up and a 20 second poststimulus undershoot as homeostasis is reestablished. [19]

2.3 Imaging of Brain Activation

The metabolic changes associated with the hemodynamic response can be accentuated through the use of fMRI [18]. The following section will introduce the concept of fMRI and why it can be used to study brain activation. Before reading this section, it is assumed that the reader is familiar with basic physics of nuclear magnetic resonance imaging and image reconstruction. If not, an overview can be found appendices A.1 and A.2.

The fundamental reason behind why fMRI can be used to study brain activity relies on the indirect measure of the magnetic properties of blood. MRI is dependent on susceptibility, which is the extend to which a material can be affected by magnetization. Local changes in susceptibility results in changes in the MR signal. [20] Changes in susceptibility arise with the hemodynamic response as oxygenated hemoglobin (HbO_2) is diamagnetic, and de-oxygenated hemoglobin (Hb) is highly paramagnetic due to its four unpaired electrons, resulting in what is known as the Blood Level Oxygen Dependent (BOLD) contrast. Thus, as presented in the prior section, the increase in neural activity increases the blood flow to an extend greater than the metabolic utilization of oxygen, which results in a high (HbO_2) to (Hb) ratio. This makes up the difference in BOLD contrast. [18, 19, 21] Figure 2.3 illustrates how the BOLD contrast is dependent on the amount of oxygenated hemoglobin.

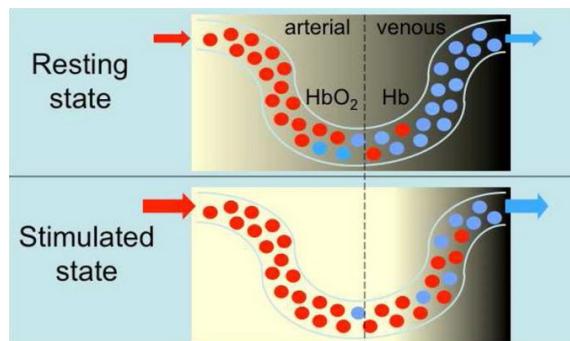


Figure 2.3: Illustration of the concentration difference in oxygenated (HbO_2) and de-oxygenated hemoglobin (Hb) during resting state and stimulated state. The diamagnetic properties of oxygenated blood changes the local magnetic substitutability facilitating a greater contrast change. [18]

The changes in the BOLD contrast can be recorded by using a T_2^* sequence, which is sensitive in detecting changes on the magnetic field [21, 22]. The areas highly filled with oxygenated blood will result in a higher signal in the T_2^* -weighted sequence making these appear brighter in the reconstructed image. To capture the changes in blood flow over time the acquisitions has to be relatively fast. To achieve fast sequences, the result is a sacrifice of spatial resolution for temporal resolution, making it possible to acquire a

whole brain volume in a couple of seconds. [21]

Other methods, as the arterial spin labeled (ASL) method exist for accentuating the functionality of the brain, but BOLD is favored, because it offers a high contrast to noise ratio and it is relatively simple to implement. [22]

As fMRI can capture the neurological response to pain through the hemodynamic response a measure to study the individual's response to pain is thereby derived. To know more of how the individual's perception of pain is measured in general and by fMRI, including the current difficulties, is of great significance.

2.4 Individual differences in pain perception and fMRI

The study of individual differences is basically to examine how each individual respond to the same intervention. This has been of great interest in medical fields ranging from psychology to pain research, as an understanding of each individual would shape more precise treatments given the individual's unique characteristics. However, understanding why individuals react differently across sensory modalities has for long been a perplexing problem. In pain research, studies have shown that individuals report a large difference in pain sensitivity on a Visual Analogue Scale (VAS), when exposed to identical noxious stimuli [9, 13]. This could indicate that individuals have different neurological responses to the same noxious stimuli, but on the other hand raise questions about the individual: is the subject over/underplaying the pain experience or engaging in drug seeking behavior? Thus, in the treatment of pain it has been of great interest to examine if the subjectively reported pain experience correlates with the actual neurological response in the individual. [10] Verifying a correlation between self-reports and brain activity during evoked pain stimulus would function as a important step in improving physician's understanding and treatment of patients suffering from chronic pain. Brain imaging could on sight be an assistive tool to self-reports and behavioral evidence when assessing a patient's claim regarding pain and physical condition. Such an addition would serve an especially valuable function when assessing patients unable to communicate verbally (e.g. young children and patients with dementia), patients with self-reports and behavioral evidence that are conflicting and for personalized pain management. In perspective, as brain imaging gets more accepted in personalized pain management it will gain more interest in legal practices. Using it as proof or disproof on whether a patient is actually experiencing pain might affect financial outcome in insurance cases. However, due to ethical and jurisdictional circumstances it would be inappropriate to use it as a tool in such cases until it has been sufficiently verified. [4]

An indirect measure of functional brain response can be achieved through a BOLD fMRI scanning, and is often used to get a generic understanding of brain function in research. Averaging data across individuals is commonly applied to raise the signal-to-noise ratio (SNR). However, while providing a general knowledge on brain function, this practice excludes the possibility of observing brain function in the individual. The interest of examining individual differences in brain activity has been present for several years, but

the technology has only during recent years been advanced enough for it to be carried out, due to higher magnetic field strength and faster acquisition time. [23]

In 1999 and 2003 Coghill et al. [24, 13] published studies showing correlation between intensity of activation of specific brain regions associated with processing of noxious stimuli and subjects' subjective pain sensitivity. Positron Emission Tomography (PET) scans and BOLD fMRI in a 1.5 tesla scanner was used respectively in the two studies for functional brain mapping. However, as Dubois et al. [23] emphasize only during the recent years the MRI technology has been advanced enough to create images with sufficient SNR to use as means in studying individual differences. A review study by Wood et al. [25] comparing 1.5 tesla and 3.0 tesla scanners, showed that using the BOLD fMRI technique can produce better susceptibility contrast sensitivity and due to the naturally higher SNR in 3.0 tesla scanners a 40 percentage increase in activation detection can be generated.

In that relation, a validation of the 2003 study by Coghill et al. [13] with data from a 3.0 tesla scanner would be relevant, as a 1.5 tesla scanner was used for BOLD fMRI acquisition in [13]. Further reasoning for a validation study is that clinics tend to get higher tesla scanners installed, when desiring to get finer images in fMRI [25]. Thus, the future of fMRI lies within the use of higher tesla scanners. In [13], a total number of 17 subjects were included. To raise the reproducibility and verification in a validation study the number of participants included should be increased to examine if statistically significant result can be found in a higher sample size [23, 26].

However, a downside of 3.0 tesla scanners compared to 1.5 tesla scanners is that artifacts related to motion, respiration, bloodflow, pulsation of cerebrospinal fluid and air-tissue interfaces are more prominent [25].

2.5 Artifacts Associated with BOLD fMRI

When BOLD fMRI sequences are used to detect brain activity during in vivo experiments, not only task related activation will be present. Unfortunately, in most experiments sources of noise will inevitably impact the scan. [27]

As BOLD fMRI relies on very precise temporal and spatial placement, even millisecond or millimeter fluctuations caused by motion can have a large impact on the quality of the acquired signal. Motion does not only distort the brain in space but also disrupts the formation of the magnetic gradients that enables the BOLD signal to be detected correctly. Bulk motion of the head will cause the content of each voxel to change. As the maximum value of net magnetization in the direction of the magnetic field (M_0) is directly proportional to the number of spins in that voxel the BOLD signal will consequently be altered. In tissue interfaces such as white/grey matter boundaries, at the edge of the brain and around large vessels this is especially an issue. Furthermore, bulk motion will change the uniformity of the magnetic field as is has been tuned for a particular head position. This also directly affects M_0 and results in dropouts, reductions in signal intensity, in the BOLD signal and thus in the subsequent readout. Lastly, movement will change steady state magnetization, by changing the time between excitations in the tissues that have moved from a slice to another. M_0 is affected until steady state has been reached again, and is referred to as the spin history effect. The spin history effect can cause the intensity

Chapter 2. Problem Analysis

of the BOLD signal to be detected as twice the level of the expected signal. [28] This will often result in an image where the intensities change in a striped pattern [19].

A typical protocol in pain research involves inducing noxious stimuli on the participant while recording the brain activity. These stimuli often causes the participant to move even more. During the movement some brain regions might also show activation associated with stimulus. Therefore it is easy to mistake brain activation with stimulus correlated movement when analyzing the data, resulting in a weaker or even false statistical analysis. [19]

Motion due to cardiac and respiratory cycles will cause the same artifacts as during bulk motion of the head. Additionally, cardiac pulsation and respiratory cycles will cause the brain stem to push the surrounding brain tissue. This will cause deformation and movement of cerebrospinal fluid that will form changes in M_0 . Further artifacts can be found related to the frequency content of cardiac and respiratory cycles. The frequency of cardiac and respiratory cycles at rest are around 1 Hz and 0.3 Hz respectively. Relatively high-frequent compared to the BOLD signal of < 0.1 Hz. As the repetition time in a BOLD fMRI usually is 2-3 seconds, parts of the cardiac and respiratory cycles will be aliased into the BOLD signal's frequency bandwidth. Thus, these artifacts can be mistaken for the BOLD signal when examining the frequency content. Another respiratory factor that affects the BOLD signal is the arterial level of CO₂. It works as a vasodilator and facilitates a global increase in the cerebral blood flow and thereby the BOLD signal. Thus, when a participant holds the breath hypercapnia arises and the BOLD signal increases, and during hyperventilation hypocapnia arises causing a decrease in the BOLD signal. [28] This type of artifact might occur more when the participant is being subjected to noxious stimuli to attenuate the pain experienced.

Furthermore, artifacts can form in areas where air and tissues meet, e.g. sinuses and the ear canals, and is caused by the main magnetic field inhomogeneities the air-tissue interface produce. It will be visualized as dropouts in the brain region adjacent to the air-tissue interface. [19]

It can be summarized that many sources of noise can alter and disguise the signal of interest. Thus, being able to separate signal sources from noise is of great importance before analysing BOLD fMRI data, to avoid making any conclusions based on spurious data.

2.6 State of the Art in MRI Preprocessing

In section 2.5 the various types of corrupting noise was presented. Continuous efforts have been made to avoid, remove and limit the influence of artifactual contributors present in the BOLD signal before initiating statistical analyses. Standard preprocessing methods of rigid-body motion correction, spatial smoothing and temporal filtering, have for long been a part of the preprocessing toolbox [19, 27]. But as studies grow more complex, seek greater SNR and introduce greater noise sources associated with the newer scanners, the need for more complex and fitting cleanup tools to add to the preprocessing toolbox is present. [25, 29]

In noise removal of fMRI datasets, there are mainly two widely used approaches: one

data-driven and one model-based [27, 30]. In the latter, mainly physiologic fluctuations of non interest, recorded with additional devices during MRI acquisition are modeled and controlled for as nuisance regressors during analysis in the general linear model. These are e.g. heart rate and breathing cycle. Furthermore models representing fluctuations in cerebral spinal fluid (CSF) and white matter, is also incorporated as regressors of non interest. [27, 30, 31] Data-driven models instead draw upon the data within the MRI dataset utilizing methods of principal component analysis (PCA) and independent component analysis (ICA). Data-driven methods are ideal when no good modeling of noise can be made into regressors of non interest as they can analyze the data in more flexible way. Hence, they are able to identify new and unexpected noise components, as well as describing the know more precise. These can then can be made into regressors of non interest in the general linear analysis model, thereby laying the foundation for a model based noise removal. [30]

One of the most used model-based method is the RETROspective Image CORrection (RETROICOR), introduced by Glover et al. [32]. By measuring the phases of respiratory and cardiac cycles and comparing these to the fourier terms of the fMRI scan, they showed that it was possible to find these physiologic noise contributions which was synchronized with the scan. These could then be out-filtered showing considerably lowered peaks in the frequency spectrum at 0.8 Hz for cardiac influence and 0.15 Hz for respiratory influence. [32]

A limitation of the model-based RETROICOR is its dependency on physiologic data being acquired from external devices and thereby not reliant to counter for unexpected noise sources. A study by Behzadi et al. [33] surpassed this restriction by utilizing a proposed data-driven model of Component based noise Correction (CompCor). Instead of relying on data from external measurement of physiological noise, the CompCor method uses a PCA derived from noise regions of interest (ROI) to describe the physiological noise. Two methods of determining noise ROI's were introduced: one using anatomical data to identify voxels where no neurological activity is presumed, such as in the white matter and the CSF, the other defining noise ROI's from voxels with high temporal standard deviation (tSTD) as these were found to correspond to ventricles, edge-regions, and vessels. The first five components from the PCA were hypothesized to describe the variance of the noise, and were used to fit in a general linear model as nuisance regressors and thereby remove the noise influence. Both CompCor methods produced a significant reduction in physiological noise fluctuations compared to the RETROICOR for BOLD fMRI. In addition, the second CompCor method was able to reduce subject motion artifacts. Though in cases of both severe motion artifacts and physiological fluctuations the PCA might only be able to describe one of these factors, limiting its use. [33]

Model based noise cleanup methods like RETROICOR depend on external physiological measurements, and both the PCA based method CompCor and RETROICOR mainly focuses on physiological noise. This suggest that methods which can incorporate a higher level of motion correction along with correction for MRI scanner artifacts might be of greater use.

ICA is another data-driven model, building on the blind source separation paradigm, for source analysis of fMRI datasets and was first introduced in 1998 by McKeown et al. [34].

Chapter 2. Problem Analysis

Since then, multiple studies [35, 36, 37, 38, 39] have used and developed the use of ICA further, for exploratory analysis of sources in the recorded fMRI signal. ICA has proven to be a powerful tool in separating the different sources which in summation comprise the complete fMRI scan. By maximizing spatial independence and non-gaussianity, the data can be decomposed into components each consisting of an activation map and its corresponding time course and power spectrum. [27] Having these three parameters made it possible to both qualitatively and quantitatively assess the components, making it possible to discriminate between artifactual activation and task-related activation. Tools for easy use of the ICA have been presented by groups like the Oxford University Centre for Functional MRI of the Brain (FMRIB), who presented a program called Multivariate Exploratory Linear Optimized Decomposition into Independent Components (MELODIC) for fast ICA implementation. MELODIC is a newer and sophisticated method of ICA incorporating a probabilistic approach in the estimation of independent sources. [40] The possibility of visually inspecting and labeling components as either artifactual or of interest, lead to discussions on standardization and comparability, as labeling of components could be subjected to operator biasing. Studies in the likes of [27, 41], have proposed ways to recognize different artifactual components, but these can be hard to utilize since every component will express different characteristics. Subsequently, automatic labeling algorithms have been made as guidelines to increase the sensibility, reliability and reproducibility of fMRI results. These have been made by training classifiers in recognizing and separating artifactual components. [39]

In 2014 Salimi-Khorshidi et al. [27] introduced the FMRIB's ICA-based X-noiseifier (FIX), for automatic denoising of resting state and task-related fMRI data. This methods employ an initial standard preprocessing followed by a probabilistic ICA, hand labeling of a training set, extraction of 180 features and training and employment of a stacking classifier. FIX was developed and tested on resting state fMRI and yielded an overall accuracy of 95%, and 99% on high-quality resting state data. [27]

The use of probabilistic ICA and classifiers is an interesting proposal as it makes way for a standardized approach, and reduces the labor intense work of labeling all components in studies with large datasets.

3 | Study Objective

In summary it would be beneficial for the understanding of chronic pain to gain further knowledge on the brain mechanisms associated with acute pain. As different individuals report a wide variety in pain sensitivity when given identical pain stimuli, an examination on if the subjective reports correlate with intensity in brain activation in brain regions associated with pain would be of high interest. When using 3.0 tesla MRI scanners to indirectly measure brain activity through BOLD fMRI sequences, more of the activity will be captured compared to using 1.5 tesla scanners. The images will on the other hand be more prone to contain noise. For this reason proper preprocessing is needed to separate noise from signal of interest before analysing the images. During the recent years more preprocessing methods have been developed in the field of fMRI as alternatives to the standard method. These methods include the CompCor and the FSL FIX. A comparison of these newer methods and the standard method (In this project we define standard preprocessing as motion, slice timing, spatial and temporal) on a large set of subjects would be favorable in the field of pain research using fMRI, as this could provide insight on the advantages/disadvantages of each preprocessing method. This introduces the study objective:

Compare and evaluate the performance of standard preprocessing and FSL FIX for noise removal in images acquired with BOLD fMRI studying individual subjects' brain activation to noxious heat stimuli.

Part II

Problem Solution

4 | Methods

The methods section serves the purpose of presenting the study design, how the test data was acquired and document the implementation of the two preprocessing methods used for noise removal. A flowchart that introduces a general overview of the different procedures can be seen in figure 4.1. First, the data acquisition protocol and the subjects demographics will be presented. Subsequently the implementation of the standard preprocessing method will be introduced, followed by the FSL FIX method. Finally, the statistical framework used for testing the performance of each preprocessing method is presented.

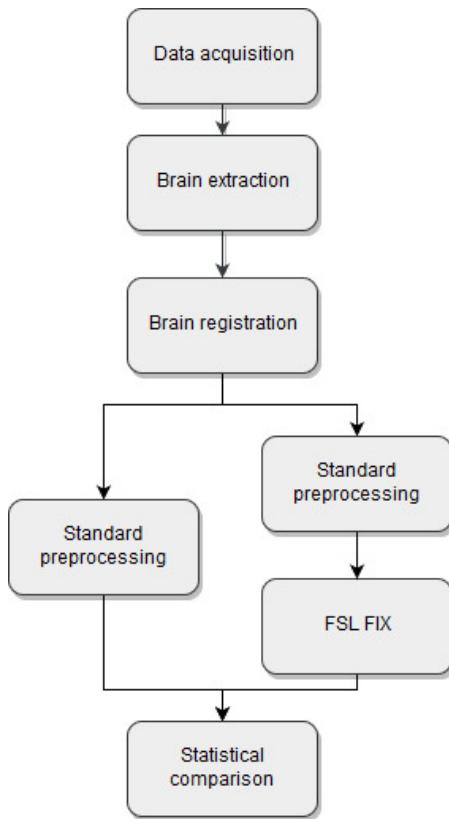


Figure 4.1: Flowchart presenting an overview of the processing pipeline. Initially the data is acquired and the brain is segmented and extracted from both a structural and functional scan. Afterwards the extracted brains are registered to each other and to standard space. Then the brain from the functional scan is run through the pipeline of each of the two preprocessing methods, and lastly the results of noise removal from each pipeline is compared.

4.1 Study Protocol

The data used in this study originates from an ongoing Individual Differences Project (IDP) at the Department of Anesthesia at Cincinnati Children's Hospital and Medical

Chapter 4. Methods

Center. In this section, information on the study design will be outlined, i.e. which criteria that were set for in/exclusion, what the stimuli design was and how the data was acquired. The IDP study design involved subjecting the participant to several types of noxious stimuli. However, as only the noxious heat stimuli will be analyzed in this project, only the heat stimuli design will be described.

139 healthy subjects between the age of 14 and 44 were recruited with no discrimination of ethnicity. In table 4.1 is an overview of essential subject population demographics presented.

Table 4.1: Overview of subject population demographics.

	Age, mean(std)	Gender n(%)	
		Female	Male
Total (n = 139)	28.40(7.31)	82(58.99)	57(41.01)
Training (n = 34)	28.00(6.10)	22(64.71)	12(35.29)
Test (n = 105)	28.53(7.68)	60(57.14)	45(42.86)

As the study involved MRI scans, subjects that suffered from claustrophobia or had any electronic or metallic implants were not able to participate. Further exclusion criteria were pregnancy, skin conditions or past skin damage on or near the site of the heat stimuli delivery and intake of medications that might have altered pain sensitivity or brain activation. The participants were able to withdraw from the study at any time. Additionally, the study investigators were able to withdraw participants if they declined to comply with the experiment protocol, if a drug or pregnancy test showed positive, if there was an identification of severe brain, if neurological or psychological abnormalities or if the investigator assessed that it would be in the participant's best interest.

Before enrollment in the study, the participants received a consent document (including the guardians/parents for participants under 18 years). First when full comprehension of the information was demonstrated the participant and guardians/parents would be allowed to sign the document.

Data acquisition

The heat stimuli was delivered on the left calf with a Pathway ATS, 16 x 16 mm, temperature stimulator, while the participant was in a MRI scanner for recording of brain activity using a T_2^* image sequence. For visualization of brain structures and spatial normalization

of functional data high resolution T_1 -weighted structural scan were first obtained. Both scan sequences were acquired with a Phillips Achieva 3.0 T x-series MRI scanner with 32 head coils. MRI acquisition sequence specific information for both the structural and functional scan can be found in table 4.2.

Table 4.2: MRI specification for the T_1 -weighted structural scan and the T_2^* -weighted BOLD scan. RL, AP and FH denote right-to-left, anterior-to-posterior and foot-to-head, respectively.

	Structural (T_1)	BOLD
TR:	10 ms	2000 ms
TE:	1.8 ms (delta 2 ms)	35 ms
Echo Sequence:	Spin Echo	Gradient Echo
Number of echoes:	4	1
Flip angle:	8°	90°
Slice orientation:	Sagittal	Transverse
Field of View:	200 mm RL 224 mm AP 256 mm FH	240 mm RL 240 mm AP 136 mm FH
Slice Thickness:	1 mm RL	4 mm FH
Acquisition voxel size:	1 mm AP x 1 mm FH	3 mm RL x 3 mm AP
Number of slices:	200	34
Acquisition Matrix:	256 x 224	80 x 80

Each participant went through three different block designs consisting of three different heat stimuli runs with seven noxious heat stimuli in each. A theoretical overview of stimuli designs and their effect on the hemodynamic response can be found in section A.3. Each heat run had a unique sequence of 10 seconds of noxious stimuli temperatures of 47°C or 48°C interleaved with resting periods of skin temperature (35°C) stimuli. The design of the three heat runs can be seen in figure 4.2. Following each noxious stimulus the participant was instructed in rating the pain sensitivity and unpleasantness using a trackball/response device and a computer projected VAS, after which the participant could rest until the subsequent stimulus was delivered. The time course of each heat run followed the same pattern starting with a 20 second initiating resting period followed by seven stimulus and rating/resting periods of 2 second increasing ramp stimulus, 10 second plateau stimulus, 3 second decreasing ramp stimulus and 37-38 second rating/resting. This means that the interval of stimuli onset is approximately every 40 seconds resulting in a frequency of 0.025 Hz. Thereby the stimuli frequency is above the range of 0 to 0.015

Chapter 4. Methods

Hz, which is a frequency range where scanner noise is present [19].

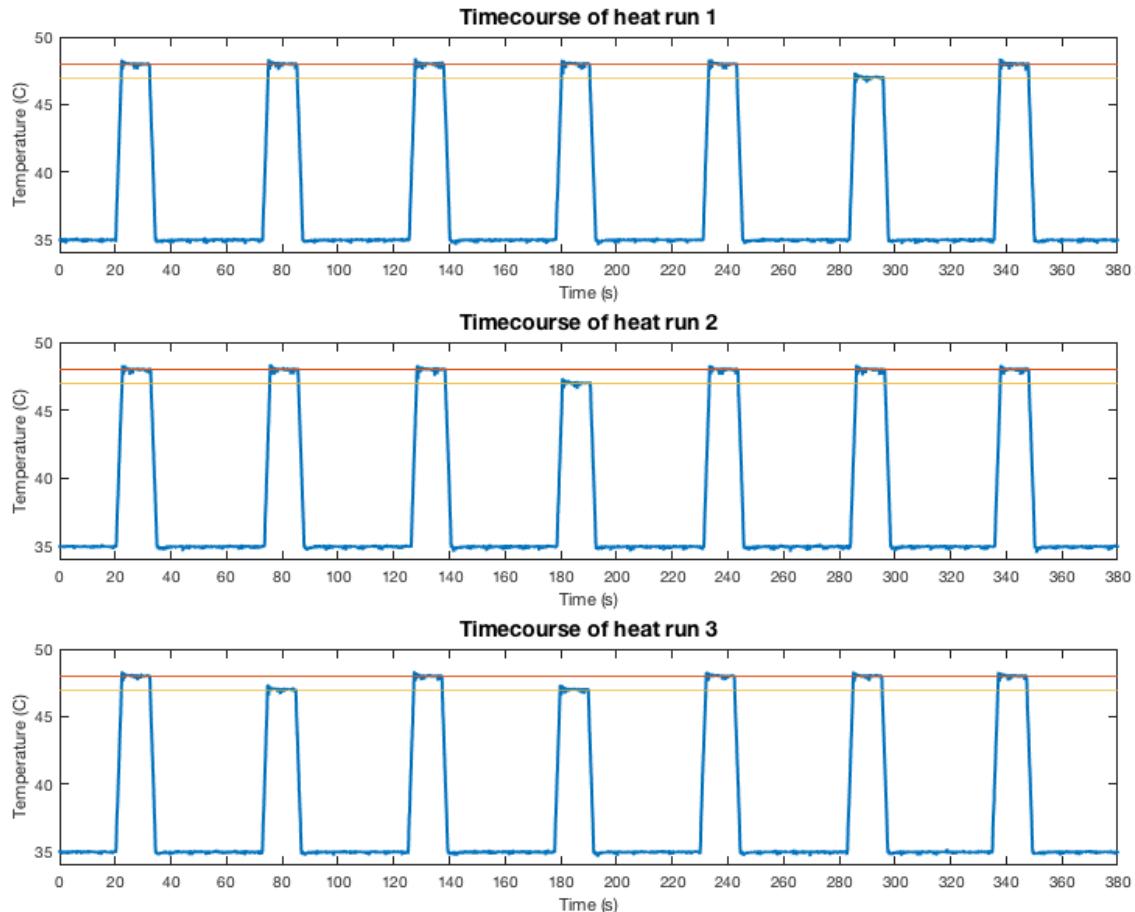


Figure 4.2: Illustrations of the time courses of heat stimuli for all three heat runs. The blue graphs depict the temperature of the heat stimuli and the yellow and red lines are used to distinguish between 47 and 48°C. In the first heat run the 47°C stimulus was delivered as the 6th stimulus, in the second as the 4th stimulus and in the third as the 2nd and 4th stimuli. The remaining stimuli were of 48°C.

To minimize bias all participants received the same stimuli in a single-blind fashion. Furthermore, the stimuli was induced in a pseudo-random order to confine the effects of habituation and expectation.

4.2 Brain Extraction and Registration

In order to study activation in the brain, the brain has to be segmented and extracted from the surrounding tissue. To achieve an image only containing the brain, there was made use of the FSL Brain Extraction Tool (BET). BET was used for extracting the brain in both the structural T₁-weighted image and in the functional T₂^{*}-weighted image sequence. As this is a standard applied method and is done before the implementation of preprocessing methods, it is seen non essential for the focus of this project. Hence, no

further documentation of this step will be presented. For further documentation of BET specifications we refer to Smith et al. [42].

The next step of preparing the functional images for analysis is to obtain the same coordinate system for these and the structural images. This step is called registration and is also known as spatial normalization. There are two levels of registration, intra-subject and inter-subject. Intra-subject are registering different scan types within the same subject and is done to spatially localize the activation. As the functional images contain very low resolution the localization of activation is hard to assess and therefore activation localization can more easily be studied by registering the functional with a structural image. Inter-subject is to obtain the same brain localization in the coordinate system, called common space, for all subjects. This is done to achieve population analysis facilitating result to be interpreted and reported objectively and consistently across studies. Inter-subject registration can be fulfilled by registering the brain for each subject to the often used Montreal Neurological Institute (MNI) template. [43]

In this project a two step registration was implemented. First a intra-subject registration was completed by registering the functional to the structural, by utilizing the FSL FM-RIB'S Linear Image Registration Tool (FLIRT), which uses affine linear transformation. [44] Secondly, a inter-subject registration was implemented, registering the structural for all subjects to the MNI template. This was achieved by implementing both FLIRT and FMRIB'S Non linear Image Registration Tool (FNIRT), which uses non linear transformation. [45]

4.3 Dataset Structuring

As presented earlier the total number of subjects included in this study was 139. Each subject underwent 3 heat runs where, for each, a functional scan was acquired, summited to a total of 417 scans. In section 2.6 it was introduced that the FIX preprocessing method utilizes a classification algorithm for separating noise sources from signal sources. The FIX software package comes with predefined training sets for training the classifier, but none of these have been made from fMRI scans consisting of signal related to the pain response from a noxious stimulus block design. Therefore, it was chosen to split the total dataset into a training dataset for training the classifier for this specific application and a test dataset for evaluating the performance of both preprocessing methods.

The training data set consisted of 34 subjects making a total of 102 scans. However, during the initial analyses two heat scans were found to have missing data resulting in the exclusion of these. This meant that the test dataset consisted of the remaining 105 subjects. However four subjects were found to not having the cerebellum fully scanned, which resulted in additional exclusion of these four subjects. In the remaining test dataset 7 scans were found to have a missing heat run scan. Therefore, the test dataset consisted of a total of 296 scans.

Approximately one fourth of the entire dataset was to be used as training data. In Salimi-Khorshidi et al. [27] it was stated that a minimum of 10 subjects were required to train the classifier. With consideration to the hypothesized higher complexity of the acquired data, a higher number of subjects was implemented. However, it was prioritized that the

data comprised of more than 100 subjects to retain high statistical power.

4.4 Standard Preprocessing

Following the three initial steps, the data needs to undergo multiple preprocessing steps before the multidimensional images are ready to undergo statistical analysis. These steps involves correction methods, which are often referred to as preprocessing. There are multiple steps in preprocessing fMR images depending on the apparent application and intended outcome. However, there is a standard set of methods that is usually used across all applications. [46] The standard approach to preprocessing of task-related fMRI, which defines a baseline for noise removal, will be described in the following section. The procedures included in the standard preprocessing pipeline can vary depending on the defined study design. In this project the standard preprocessing steps include motion correction, slice timing correction, spatial smoothing and temporal filtering. Figure 4.3 provides a chronological overview of the implemented steps in the standard preprocessing pipeline.

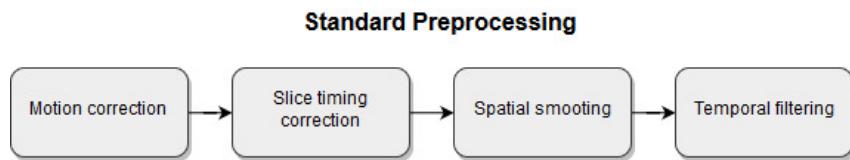


Figure 4.3: Flowchart illustrating the chronology of steps in the standard preprocessing pipeline.

4.4.1 Motion correction

Correcting for motion artifacts when working with fMRI is nearly inevitable, since even the best subjects will not be able to hold completely still during acquisition. Even subtle movements as swallowing will be visible in the acquired image. [19] To correct for these movement induced artifacts, a realignment of all volumes within a scan has been carried out using the FSL software program FMRIB's Motion Correction Linear Image Registration Tool (MCFLIRT). [47] The tool was used to realign the series of images to a reference image. The volume acquired halfway through the scan was used a reference image. This choice of reference image is often justified by the middle volume being the closets to the average as well as the scanner at that time would have achieved maximum stability, as the magnetization would have reached steady state. [19] The displacement for each voxel in each volume to the voxel in the reference image is calculated and a linear 6 degree of freedom transformation was used to realign the images. The optimal realignment was found by optimizing a the normalized correlation cost function. Due to MCFLIRT only using linear transformation for motion correction only bulk motion is corrected for. [47]

4.4.2 Slice timing correction

When acquiring fMRI scans the slices in each volume are recorded one by one. This can either be in an ascending, descending or interleaved order. Interleaved order is sequentially skipping every either odd or even slice and then afterwards acquiring the skipped slices. Descending and Ascending are acquired from top to bottom or bottom to top, respectively. Regardless of which order the slices were acquired, a misrepresentation of the hemodynamic response in each slice will be present due to the time difference in which each slice was acquired with. A representation of this effect can be seen in figure 4.4. The difference in timing for each slice constitutes a problem in the later statistical analyses of the data. The statistical model assumes that all slices are acquired at the same time point, thereby the measured signal and the estimated in the statistical model creates a mismatch. [19]

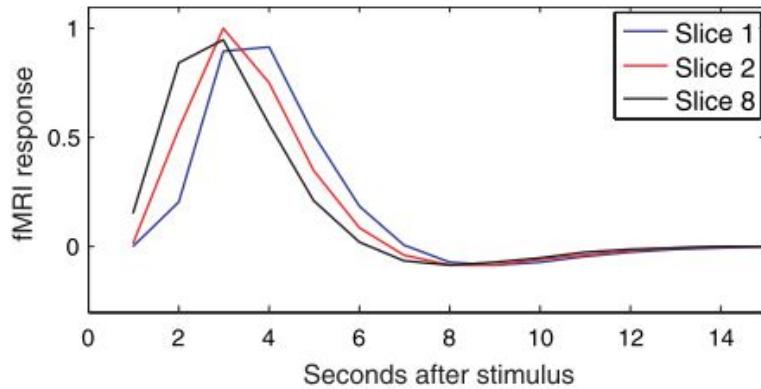


Figure 4.4: An illustration of the slice-wise acquisition effects the representation of the hemodynamic response. The response will seem to appear earlier in slice one compared to the others, as it always will be the first acquired slice. [19]

As mentioned in section 4.1 the scan time for one heat run was 386 seconds, and with a TR period of 2 seconds, this resulted in 193 volumes per scan, which needed undergo slice timing correction. To account for the mismatch in the acquisition time of each slice, slice timing correction was implemented using FSL software [48]. This was by supplying the algorithm with acquisition order and the TR period. As the acquisition order was ascending and the TR-period was two seconds, this was used as input. The slice timing worked by shifting the time series for the hemodynamic response curve for each slice in time to fit the middle of the TR period. This was achieved using Hanning windowed sinc interpolation¹. Choosing the middle slice as the reference, introduces the least amount of interpolation, and was therefore seen as the most accurate method. [19]

¹FIXme Note: ELABORATE

4.4.3 Spatial smoothing

Some of the image pixels will likely be contaminated with scattered noise occurring as higher pixel intensities. These can be removed by averaging the misrepresented pixel with its surrounding neighbors. Hence, this allows for the possibility of gaining a higher signal to noise ratio, however, with the consequence of a decrease in spatial resolution as the image gets more blurred and smaller areas of activation gets smeared together. The operation can be justified by the closely neighboring voxel being correlated in effect to the hemodynamic response. Thereby some of the higher-frequency information is removed by spatial smoothing. Furthermore spatial smoothing has the advantage that it reduces the difference between subjects, which improves the between subject comparison. [19]

The implementation of spatial smoothing on the functional brain image was done using the FSL Smoothing over Univalue Segment Assimilating Nucleus (SUSAN) noise reduction filter. The three dimensional spatial smoothing was carried out on each voxel of the fMRI data set separately. Thereby reducing noise without reducing valid activation, which should be achieved using a default mask size of 5 mm full width half maximum (FWHM), thus resulting in a trade off of smoothing both bigger and smaller areas of activation. An advantage of the SUSAN algorithm is its ability to distinguish between tissue types as e.g. grey matter and white matter and thereby only including intensity value from neighboring voxels which consist of the same tissue type as the voxel being smoothed. Thereby preserving both anatomical structures and avoiding errors intensity levels in different tissue types influencing the averaging. [49]

4.4.4 Temporal filtering

A common noise component in the fMRI data is the presence of a low-frequency drift. The drift is characterized as a slow increasing trend in the BOLD magnitude, when assessing the signal in the time domain. As mentioned in section 4.1, the frequency of noise was 0 Hz to 0.015 Hz and the frequency band of activation is approximately 0.025 Hz. To dampen some of the low-frequency artifactual content a high pass filter with a cutoff frequency of 0.01 Hz was implemented. [48]

4.5 FSL FIX

The following section will describe the implementation of the automated de-noising system FIX, which introduces a faster and more standardized approach to ICA-based noise cleanup [27]. The FIX method is a more extensive approach, which adds another layer of analyses to the standard preprocessing. This means that before initiating FIX the data needed to undergo the exact same steps as in the standard preprocessing presented in section 4.4. Following the standard preprocessing, the functional scan underwent an ICA to separate the sources in the acquired signal and ideally splitting signal of interest and noise. Afterwards an analysis of optimal number components in which the signal should be split was performed, as the initial results from the default setting were assessed not

to sufficiently separate the signal sources. The FIX method uses a classifier to separate noise components from signal components. Hence a classifier made and trained in doing this separation. In figure 4.5 a flowchart showing the steps for both the training data set and the test data set in the FIX pipeline is illustrated.

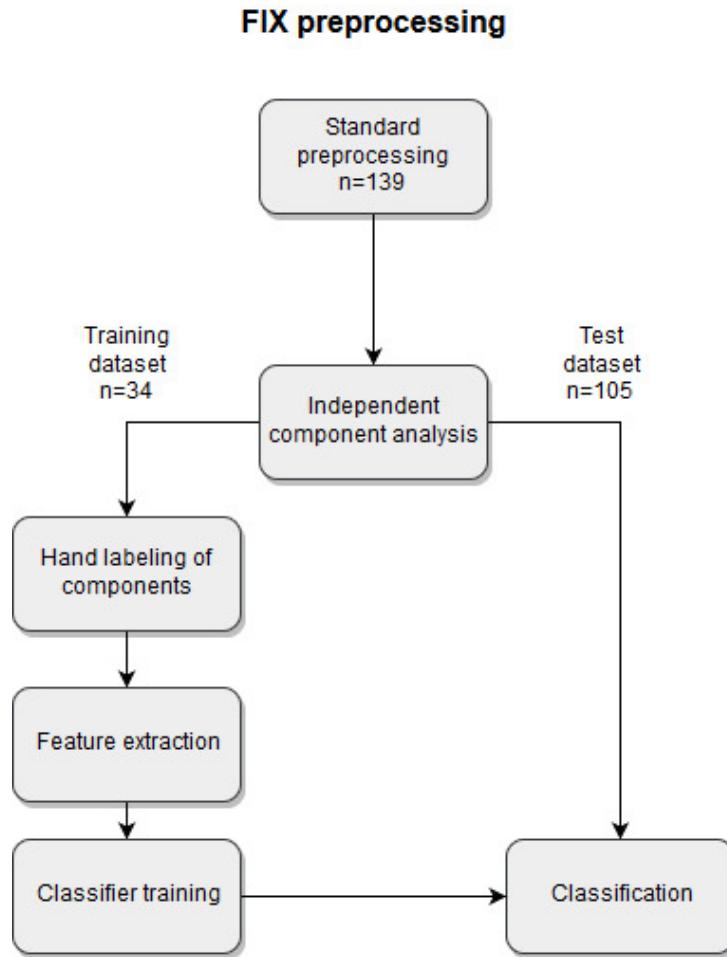


Figure 4.5: Flowchart illustrating the two datasets and the corresponding steps that was applied to each dataset in the FIX pipeline.

4.5.1 MELODIC ICA

After the implementation of the standard preprocessing, the subsequent step of the FIX implementation was to apply an ICA to the functional scan. ICA has for many years been a favored exploratory analysis tool in fMRI, as it allows for a decomposition of the signal information into estimated sources. Hence, it would be possible to differentiate between sources of brain activation, physiologic noise and other artifact contributors. ICA employs blind source separation techniques expressed as having a random set of variables (the acquired observations), as a linear combinations of statistically independent component variables (the sources). [50]

Chapter 4. Methods

To achieve a decomposition of the BOLD signal into components each consisting of a source signal the FSL ICA tool MELODIC was utilized. The following section will document how the ICA was performed on fMRI data through the use of MELODIC. For a introduction of general ICA we refer to section A.4.

MELODIC is similar to the general ICA, but includes a probabilistic approach, forming what is called a Probabilistic ICA (PICA). It stills expresses the assumption that a vector of p observations is generated from a set of q statistically independent non-Gaussian sources via a linear instantaneous mixing process. In addition it is assumed that this process corrupted is by Gaussian noise $n(t)$ [50]. This can be expressed as,

$$x_i = As_i + \mu + n_i \quad \forall i \in V. \quad (4.1)$$

x_i denotes a p -dimensional column vector of each measurement in a specific voxel i , which result in a column vector for each voxel V . s_i denotes a q -dimensional column vector of the independent source signal contained in the data. μ denotes the mean of the observations x_i . A is the $p \times q$ mixing matrix and n_i is the Gaussian noise, which is estimated using a probabilistic principal component analysis (PPCA). Visualized in matrix form the input observations for each entire functional scan was set up as,

$$\begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,217600} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,217600} \\ \cdots & \cdots & \cdots & \vdots \\ x_{193,1} & x_{193,2} & \cdots & x_{193,217600} \end{bmatrix} \quad (4.2)$$

Thereby, for each of the 217600 voxels 193 measurements were made, resulting in a 193×217600 input matrix. In order to estimate the independent sources, the solution lies in finding a linear mixing matrix W , which makes:

$$\hat{s} = Wx \quad (4.3)$$

Such that \hat{s} is the optimal approximation of independent source signals.

PPCA

Before finding the mixing matrix PPCA is employed to estimate the "true" number of sources in the signal. A generel PCA is an often used method of reducing the dimensionality of a dataset. A general introduction to PCA can be found in section A.5. The PPCA is used to find a usable linear subspace of sources. PPCA incorporates a probabilistic model, which is used to asses the number of true sources and random observed

noise. Before employing the PPCA the input data was variance normalized. Finding the subspace of sources was done by using Singular Value Decomposition (SVD), which leaves the subspace with a set of sources which are orthogonal to each other and thereby being uncorrelated. Through a probabilistic Bayesian framework the number of true sources is estimated from the covariance matrix incorporating the information of unstructured noise that is found in the weakest eigenvectors of the PCA. Hence, through the PPCA a linear subspace of "true" signals was estimated. Furthermore, the PPCA estimated subspace is spatially whitened, which is beneficial as it simplifies the later ICA estimation by lowering the parameters to be estimated.

In this project there was not made use of the dimensionality estimation for reasons explained later in section 4.5.2. Hence, only a fixed number of sources were estimated being the first 25 "true" components describing the most of the variance.

PICA

The next step was to solve for the linear subspace, by estimating the mixing matrix, in order to achieve independent source signals. There are multiple approaches in which this can be achieved. The main idea of the PICA is the importance of the source distributions being assumed to be non-Gaussian with an additive Gaussian noise. Hence, estimating the unmixing matrix W is based on maximizing non-Gaussianity. Non-Gaussianity was maximized by approximating neg-entropy.

The result of solving the mixing matrix was a 193×25 matrix containing 25 independent components with a time course for each. In order to get a spatial map visualizing where the activation of a specific source was present, each voxel's time course was projected onto the columns of the unmixing matrix. These maps were then transformed into Z score maps using the standard deviation of the prior estimated noise, which enables differentiating between source signal and background noise. Finally, a single spatial IC map for each component was made by fitting a Gaussian Mixture model to the individual IC maps to find voxel locations that were significantly activated through the IC time course. This made a probability map for each component, which was then thresholded to only capture meaningful activation. The resulting output of employing the MELODIC ICA was a 193×25 matrix of temporal components and a 25×217600 matrix of spatial components. A summarization of the input and output of MELODIC is illustrated in equation (4.4), where the left-hand side was the input and the right-hand side was the two output matrices.

$$\begin{bmatrix} x_{1,1} & \cdots & x_{1,217600} \\ x_{2,1} & \cdots & x_{2,217600} \\ \vdots & \ddots & \vdots \\ x_{193,1} & \cdots & x_{193,217600} \end{bmatrix} = \begin{bmatrix} x_{1,1} & \cdots & x_{1,25} \\ x_{2,1} & \cdots & x_{2,25} \\ \vdots & \ddots & \vdots \\ x_{193,1} & \cdots & x_{193,25} \end{bmatrix} \otimes \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,217600} \\ \vdots & \ddots & \ddots & \vdots \\ x_{25,1} & x_{25,2} & \cdots & x_{25,217600} \end{bmatrix} \quad (4.4)$$

4.5.2 Analysis of optimal number of ICs

After having undergone the initial standard preprocessing presented in section 4.4, the next step was to extract the ICs using MELODIC. But before the data could be subjected to the ICA, several considerations were made in order to increase ICA performance. The separation of sources is highly dependent on the number of ICs (n-ICs) chosen and therefore crucial for the FIX end performance. Consequently an investigation of the optimal number of components was conducted. The risk of underestimating the number of components, would result in a loss of information and suboptimal signal extraction [50]. Contrarily, overestimating the number of components would result in a false representation, as the different informative sources would be split leaving fractional sources, which would be very hard to identify and utilize [50, 51].

Data used for investigating the optimal number of components was derived from ten randomly chosen subjects having three heat runs each. This made for 30 IC analyses each completed utilizing the MELODIC tool.

Default setting

Initially no limit was selected and the MELODIC algorithm per default estimated the optimal number of components, as recommended by the MELODIC developers. [40] This was done on the first 30 scans in the training dataset to assess the MELODIC performance. The MELODIC tool provided analyses with the estimated number of component to represent the variance ranging from 9-64 components. Due to the substantial amount of variability in the number of components estimated, the degree to how much the sources had been separated varied substantially between the scans. Inspecting the output from the ICA, which produced nine components, it became clear that the separation of sources was not sufficient, as the components were mixtures of multiple sources of noise and signal. Furthermore sparse non localized task related activation were seen in any of the components indicating that variance of noise accounted for the first nine components. Qualitatively inspecting the analyses with the highest number of components, it became evident that the algorithm had overestimated the number of ICs, leaving empty and fractured activation in the components unrecognizable for labeling. It was assessed that using MELODIC with default setting would be overestimating and underestimating the number of components. Therefore, a case specific analysis for the optimal number of components was conducted.

n-IC analysis

Seeking guidelines in the literature, various articles have worked with estimating the optimal number of components. In the case of Majeed et al. [52], results indicate that a n-IC lower than 5 components would lead to an underestimation and more than 50 components would result in overestimation. The optimal n-IC was found to be 10. Another study limited the number of components to 20, as this was found to preserve much of the information without overestimating [53]. Furthermore other studies limited the number to 25 components [54, 55]. Building on this knowledge, MELODIC was run on the 30

scans again limiting the n-ICs to 15, 20, 25, 30 to assess which would give the optimal separation between signal and noise, and subsequently separation of each source. The qualitative assessment was made by counting the following outcomes:

- The number of components containing signal activation from multiple sources or unrecognizable due to fragmentation
- The number of components containing signal activation from a single source
- The number of components containing no activation

Examples of spatial maps containing multiple sources, fragmented sources and a single source can be seen in figure 4.6, figure 4.7 and figure 4.8, respectively. Identification of sources was supported by utilizing the guidelines presented in [41], showing how the different noise sources would be present in the spatial map and using the information of which brain regions would be activated during pain stimuli as presented in section 2.2. Both investigators conducted the assessment. One found 25 components to be optimal, whilst the other found that 30 components was more favorable. The assessment was assisted by incorporating an expert's opinion from an expert on field of fMRI. With assistance from expert, 25 components was identified to provide the best separation of sources in the data. This decision was supported by the results of the prior studies of [54, 55].

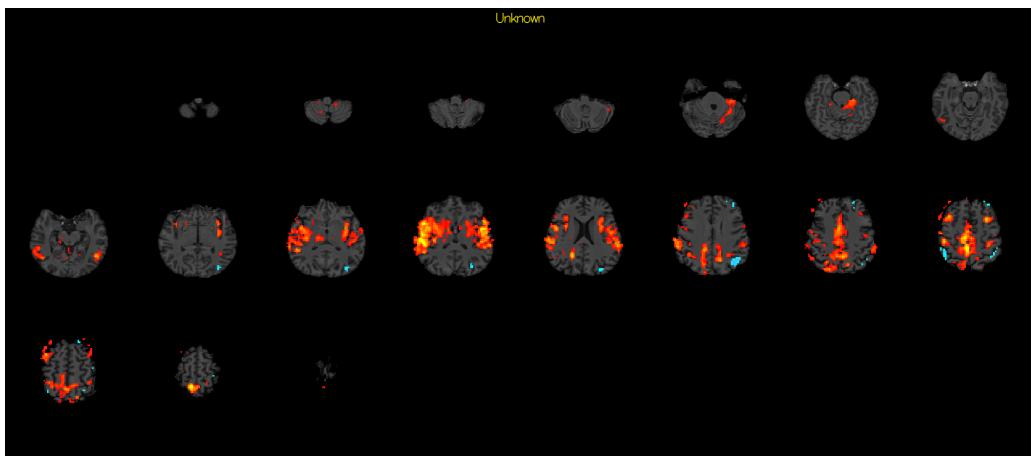


Figure 4.6: An example of a component with a spatial map for a component showing activation from multiple sources. Activation could not be localized to single brain regions, and was mixture of noise and signal.

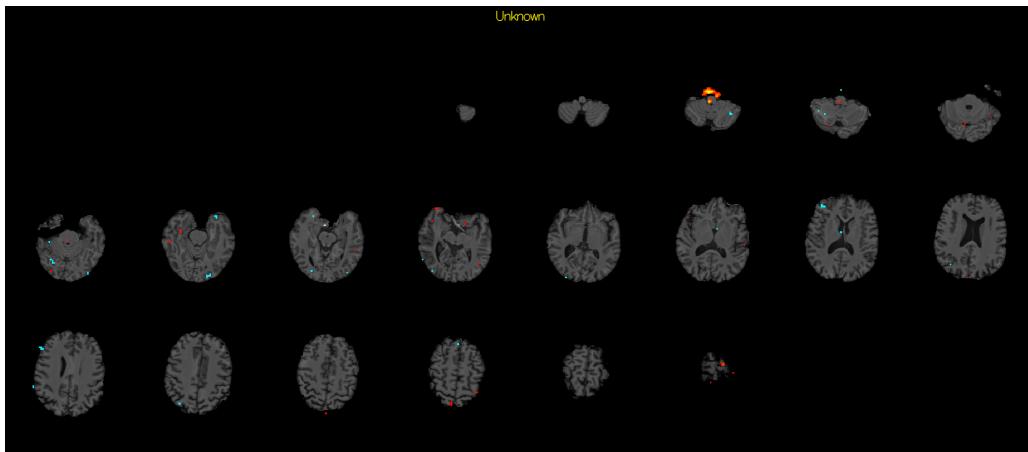


Figure 4.7: An example of a component with a fragmented spatial map illustrating that no brain activation or noise can be localized.

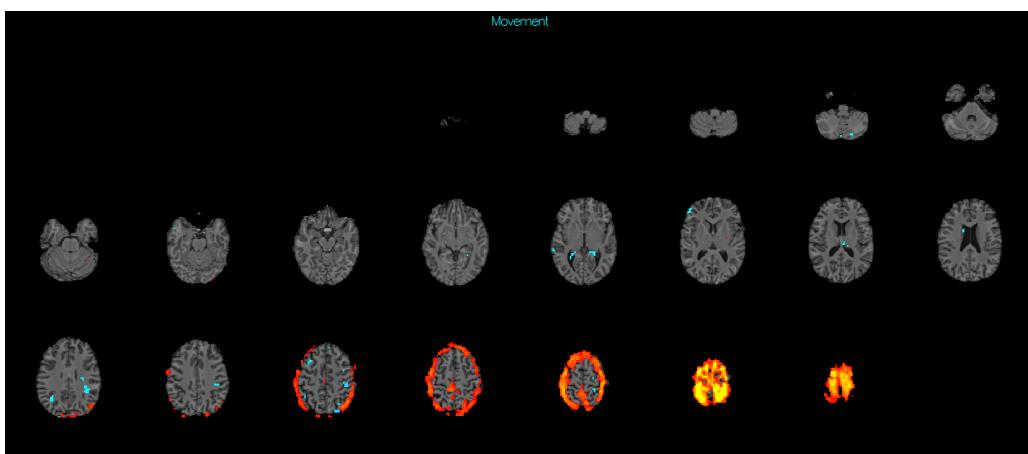


Figure 4.8: An example of a spatial map showing a clearly recognizable single source artifact in the form of movement dominating the component.

4.5.3 Hand labeling of independent components

Before the automatic FIX classifier could be used to identify components of noise and interest it needed to be trained with a set of hand labeled components [27]. For this purpose 25 components were extracted from all heat runs for 34 participants using the MELODIC tool, after which each component was manually hand labeled.

The process of the hand labeling followed that the investigators labeled half of the components each, where after the labeled components were cross-checked and discussed by the investigators. The investigator assessment relied heavily on information with regards to hand labeling provided by Griffanti et al. [41] and Salimi-Khorshidi et al. [27]. As a final assessment an expert in the field of fMRI did a control and corrected labels if they did not fit the expert opinion.

The MELODIC program proposed a set of predefined classes to label the components

as. The classes suited for artifacts were movement, susceptibility-motion, MRI, cardiac, respiratory, white matter, sagittal sinus and unclassified noise. It was possible that one component could contain more than one artefact class. If all artefact classes of the component were identifiable, more classes were added to the component. The unclassified noise class was proposed for components containing a mixture of different artifact sources, which were less clear to pinpoint. A class named unknown was intended for labeling components that contained both noise and signal. For any signal of interest the signal class was proposed, as there was no discrimination between different types of signal of interest. The following illustrations figure 4.9, figure 4.10, figure 4.11, figure 4.12, figure 4.13, figure 4.14, figure 4.15, figure 4.16 and figure 4.17 show transverse spatial examples of components related to the various artifact classes, the signal class and the unknown class.

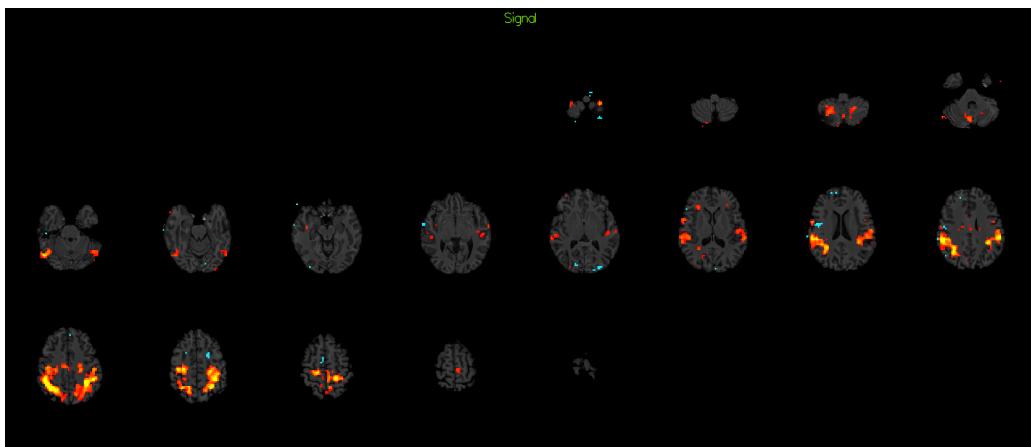


Figure 4.9: An example component of signal, which is recognizable, isolated, and not corrupted by a substantial amount of noise. In this example, the signal is characterized as a strong and fairly isolated activation mainly in the parietal lobe and some in the temporal lobe.

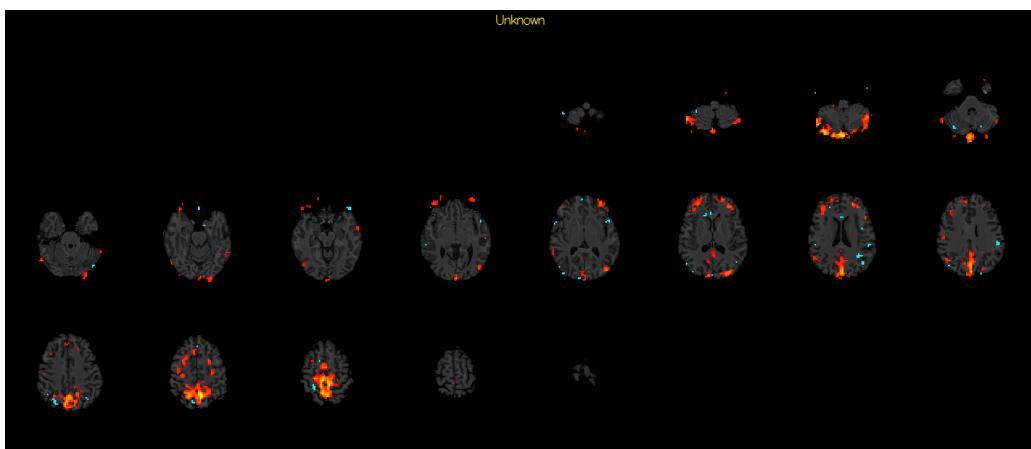


Figure 4.10: An example component of a mix of signal of interest and noise. In these cases there are some areas, which seem to have a signal related activation, the activation is either hard to localize, not seen in an expected area or the remaining spatial map is corrupted but an substantial amount of noise.

Chapter 4. Methods

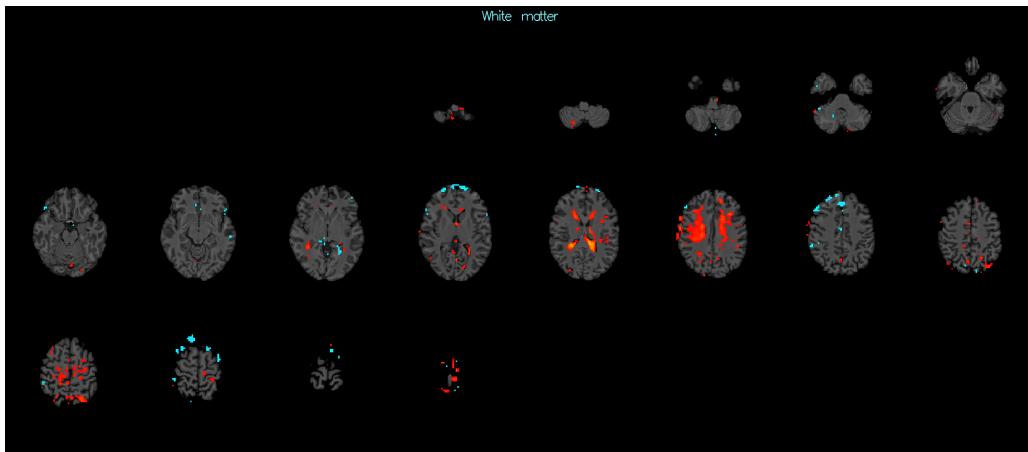


Figure 4.11: An example component of the white matter artefact. White matter artefacts are visible as activation in the white matter area of the brain, while no or low activation of areas of interest or noise are visible.

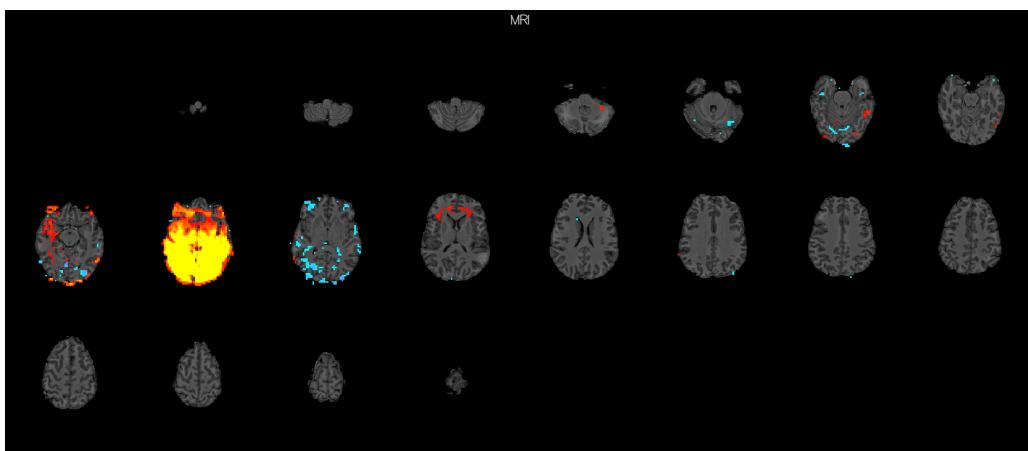


Figure 4.12: An example component of the MRI artefact. This artefact is usually visualized as fully activated (or deactivated) single slices in separated temporal levels. In the sagittal or coronal plane, this is reflected as clear stripes in the images, which can not be explained as physiologically meaningful. Another example is slices containing large portions of alternating activation and deactivation, which too can not be physiologically explained.

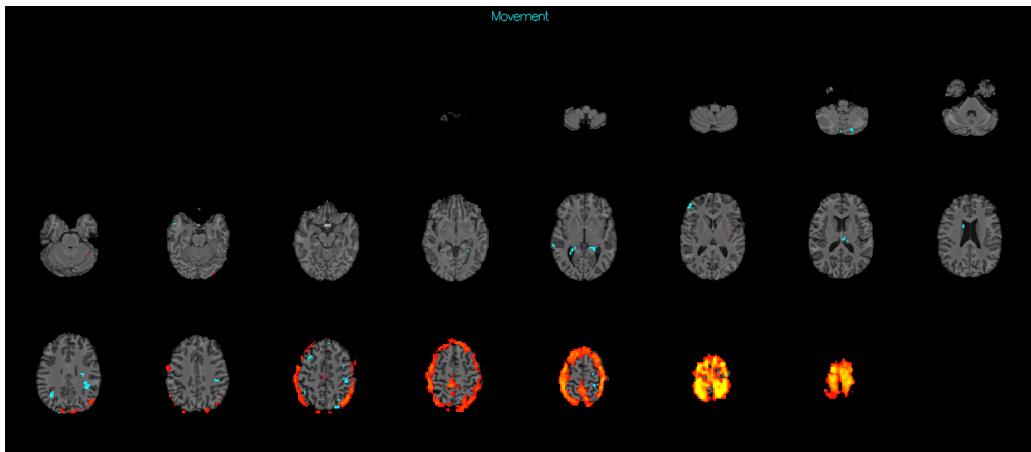


Figure 4.13: An example component of the movement artefact. Movement noise will usually be seen as a ring or lines on the periphery of the brain. Depending on the nature of the movement causing the noise, the voxels will be located on the inside or outside of the brain.

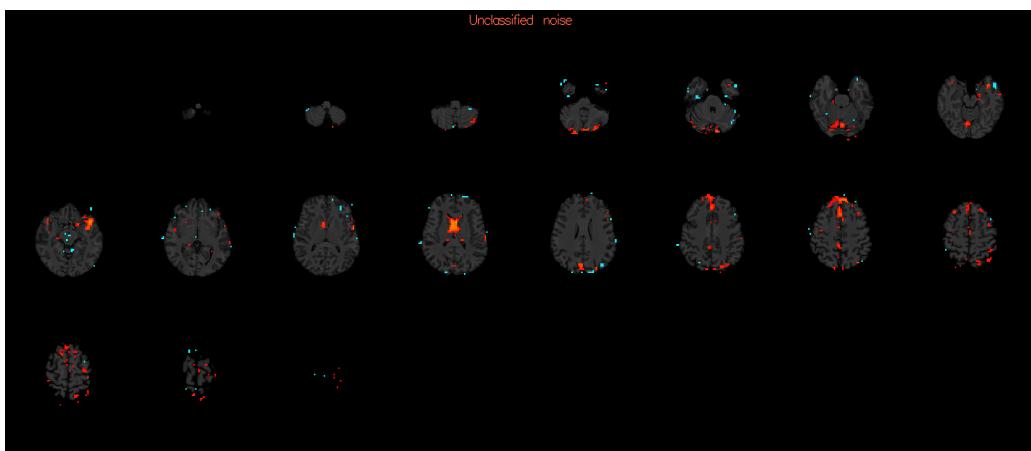


Figure 4.14: An example component of unclassified noise. A component was given this label when no clear noise source could be recognized and/or when the component was a mix of multiple noise sources.

Chapter 4. Methods

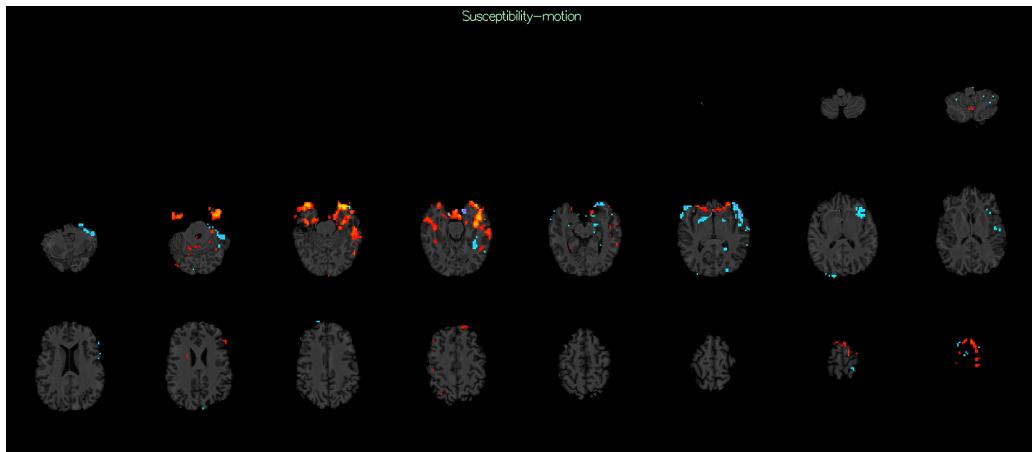


Figure 4.15: An example component of the susceptibility-motion artefact. Susceptibility-motion noise is mainly due to air-tissue interfaces, and thus usually spatially expressed as a mix of activation and deactivation at the sinuses of the skull.

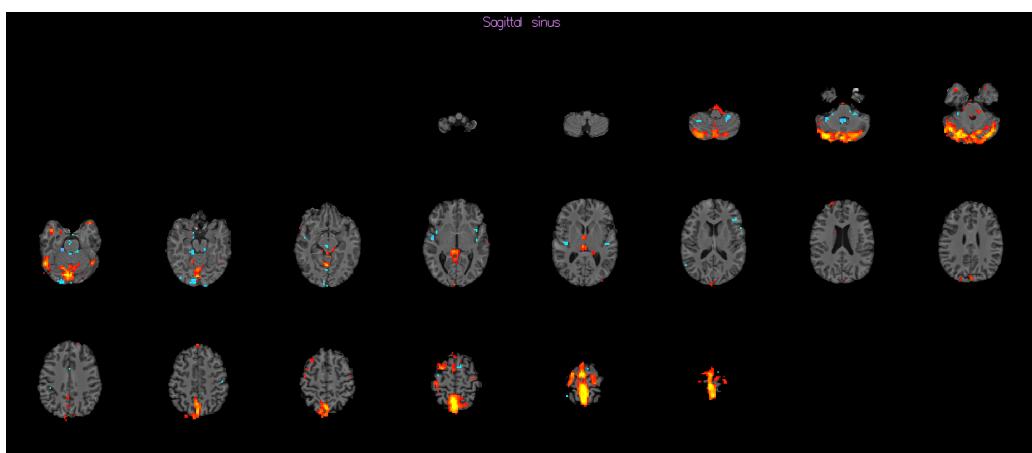


Figure 4.16: An example component of the sagittal sinus artefact. Sagittal sinus artefacts are visible as activation of the sagittal sinus veins, while no or low activation of areas of interest or noise are visible.

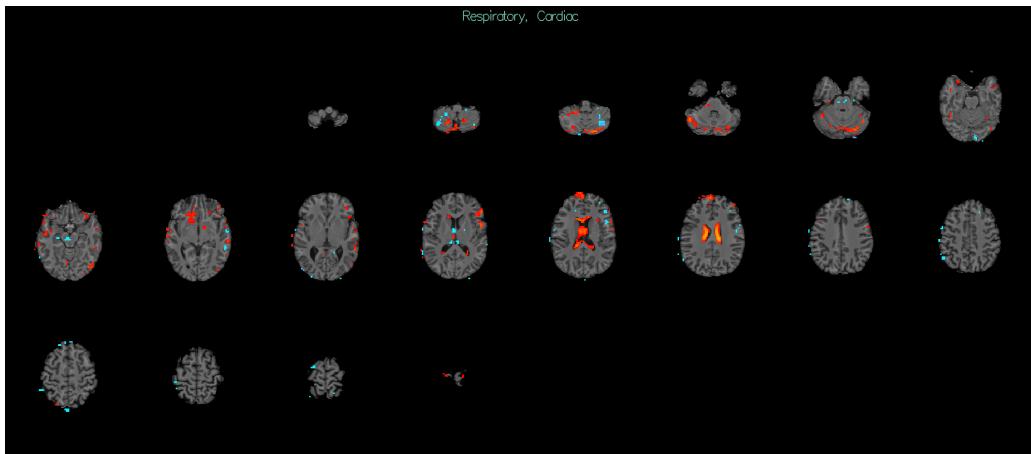


Figure 4.17: An example component of cardiac and respiratory artefacts. Cerebrospinal fluid pulsation is usually due to respiratory and cardiac cycles. Thus, activation of the ventricular system of the brain will be an indicator of respiratory and/or cardiac artefacts.

It should be noted that a coronal and sagittal view were additionally available, along with the time course and frequency content of each components for shaping the assessment of the labeling. The corresponding structural images were used to underlay the components to assess in which brain regions the components were located. Through the component labeling the scheme illustrated in figure ?? was used to decide whether the component was either signal, unknown or noise.

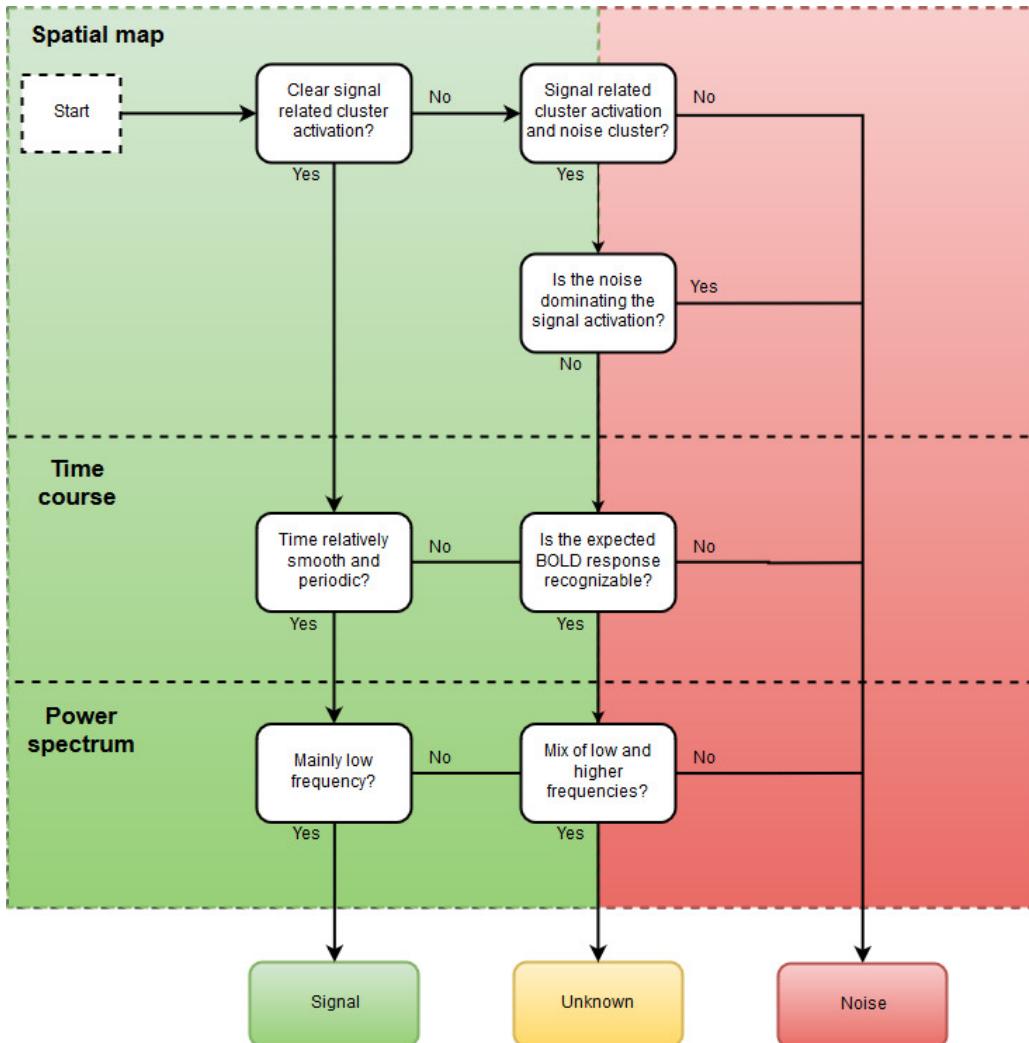


Figure 4.18: A flowchart describing the assessment of a component. First, is the component assessed in the spatial map, secondly the time course and lastly in the power spectrum. Note that only the assessment to either signal, unknown or noise is presented in the flowchart. When a component was found to be noise a secondary assessment into which type of noise would begin and is therefore not accounted for in this flowchart.

4.5.4 Feature extraction

An important step in establishing a robust classification model is extracting effective features to feed the classifier. Extraction of reasonably independent features that correlate with the targeted class will establish a more robust classifier when training it. In the FIX algorithm 186 features with temporal or spatial characteristics are extracted. The first feature extracted is, however, considered of spatial-temporal origin and consists of the number of IC computed by MELODIC. It will be expected that the noise present in the data will affect the number of ICs extracted. [27] However, in this study the number of IC calculated for each heat run was fixed on 25, and that feature would therefore

have no influence on the expectation of a component containing noise or signal. The following sections aim to explain the idea behind extracting these exact temporal and spatial features and to give examples of selected features. Thus, not all 186 features will be covered in depth. For a further explanation of the features we refer the reader to the 2014 study by Salimi-Khorshidi et al. [27] The result of the feature extraction for one from one heat run from one subject consisted of a 25 x 186 matrix - one feature value for each feature of each component.

Temporal features

Autoregressive models Some temporal features extracted were based on autoregressive (AR) properties of the time course. It was expected that the temporal smoothness derived from AR models, could help distinguish signal from particular artefact components. For an AR(n) model, where n is the order up to $n = 6$, and a_k denotes the AR parameters. Let e_p denote the variance of the residual of up to AR(6). The first AR property features derived were the slope and intercept of the straight line which explained e_p as a function of n , where an increase in n would output a better fit and thus a smaller residual variance. It was expected that an improvement in goodness of fit would decrease as more noise was present in the time course. The other AR features extracted were simply a_1 , a_2 , e_1 and e_2 . These features encapsulated the autocorrelation estimated from the low order AR models. It was expected that components containing signal would have a higher temporal autocorrelation and lower residual variance compared to components containing unstructured noise.

Distribution information It was expected that signal components would be normally distributed, while noise components would vary in how they were distributed (e.g. long-tailed distribution and multi-modal), e.g. due to sudden spikes in the time course caused by rapid movements and/or scanner artefacts. Features that gave information on distribution were therefore extracted, and these included kurtosis, skewness, mean-median difference, entropy and negentropy.

Jump amplitudes The extent of jumps in the amplitude of the time course is an important feature in distinguishing signal from noise. Signal components are expected to be fairly smooth, while noise components usually will contain large jump and fluctuations. To describe this characteristic features based on division calculations that include the standard deviation (std), mean or maximum of the differential of the time course and the std, mean and/or maximum of the time course were extracted.

Fourier transform Taking the fast Fourier transform (fft) of the time course will capture the frequency content, which can be utilized to differentiate between signal and noise. It is expected for signal components to almost exclusively have high power amplitude in low frequencies due to the block design of the stimuli task. Noise components may on the other hand contain frequency in the whole spectrum. Therefore several fft-based features were extracted. Examples of features derived were the total power above 0.1, 0.15, 0.2 and 0.25 Hz respectively and the percentage of power that lied in the following intervals: 0:0.01, 0.01:0.025, 0.025:0.05, 0.05:0.1, 0.1:0.15, 0.15:0.2 and 0.2:0.25 Hz.

Correlation It is expected that time courses of signal components are strongly associ-

ated with grey matter (GM), while noise components will correlate with time courses of white matter (WM) and cerebrospinal fluid (CSF) along with the motion correction time courses calculated in the initial pre-processing. The last temporal features extracted were therefore based on correlation between the component time course and reference GM-, WM-, CSF-derived time courses and the motion correction time courses. The reference time courses were calculated using FSL's tissue segmentation tool, from which GM, WM and CSF masks were extracted. Each type of tissue's time course was then computed as the average of the time courses that corresponded to that given tissue. [27]

Spatial features

Clusters' size and spatial distribution The distribution of activated and deactivated cluster sizes is an important indicator of a component containing either noise or signal. It was expected that signal components contain small number of relatively large clusters, while noise components contain a large number of small clusters. To capture this information, a list of a spatial map's clusters, \mathbf{c} was formed, where only clusters of at least 5 connected activated or deactivated voxels were kept, and listed in descending order. Feature extracted to summarize \mathbf{c} are: $\text{length}(\mathbf{c})$, $\text{mean}(\mathbf{c})$, $\text{median}(\mathbf{c})$, $\text{max}(\mathbf{c})$, $\text{var}(\mathbf{c})$, $\text{skewness}(\mathbf{c})$, $\text{kurtosis}(\mathbf{c})$, $\mathbf{c}[1]$, $\mathbf{c}[2]$, and $\mathbf{c}[3]$, where the last three features were the first to third elements of \mathbf{c} . Looking at the distribution of clusters in individual slices can help detect scanner artefacts. Let \mathbf{V} and \mathbf{U} contain slice specific information of the ICA spatial map \mathbf{m} and slice specific information of voxels above 2.5 z-score \mathbf{m}_p^τ , respectively. Further features consist of $\text{max}(\mathbf{V})$ and $\text{max}(\mathbf{u})$ and slices containing above 15. It is not very likely for signal components to have large contents of both activation and deactivation in the spatial map. For this reason features that use the mean, standard deviation and entropy of \mathbf{m} to measure the amount of positive and negative voxels in the components were extracted.

Voxels overlaying dark/bright raw data voxels As mentioned signal of interest is associated with dark voxels where noise is usually found in bright voxels. The next features are based on multiplying and dividing the components' spatial map with the mean of the corresponding pre-processed time courses, forming two new images. It was expected that the intensity information contains information on which voxels were of interest and which are not. The features consisted of the 95th and 99th percentile of the new images.

Percent on brain boundary High activation/deactivation in the spatial map that overlaps the boundary of the brain and the area outside the brain will most likely be movement related. Segmenting the brain using the FSL BET tool and subtracting this mask with an eroded version of itself will output a mask of the brain's edge. Five masks with varying thickness were extracted. The features measured how large a percentage of the components spatial maps that was imbedded in the edges and how large a percentage of edges was covered by the component. The higher these values were, the higher probability the component was to have captured a movement artefact.

Mask-based features It might be needed to use spatially-specific masks to detect certain noise sources (Sagittal sinus, CSF and WM) more precisely. The spatial appearance of the major veins could for instance be mistaken for signal of interest and the cluster-like

activation pattern may look similar to signal of interest. To detect these structures more conservatively FIX utilizes three standard-space masks that consist of three major veins. From the three masks three different masks with varying thickness are derived, due to the anatomical variability between subjects. The same features as in the brain boundary features were extracted, where these new masks were utilized instead. The same procedure was executed for GM.

Other spatial features Other features detect spatial smoothness. Signal components are expected to have high spatial smoothness, meaning a fairly small amount of connected clusters, where noise components are expected to have a more patchy spatial map. Features that extract spatial smoothness in mm and voxels counts are therefore included. The last spatial features detect striped patterns of interleaved positive and negative activation. A largely striped pattern in the spatial map could imply that a component contained noise. [27]

4.5.5 Classification algorithm

The most important function of the classifier is to correctly separate the noise components from the signals of interest. However, it is highly probable that components do not purely contain either noise or signal, and that the characteristics of the various types of artefacts overlap. Thus, in the context of a classification model, noise and signal are not left as two well-defined clusters in both the temporal and spatial domain, and the decision boundaries separating noise from signal will be complex. [27]

A first step in simplifying the decision making process was to divide features into subsets, partially through a feature selection process. The reasoning for this was that some components might have more clear signal fluctuations shown only in either temporal or spatial features or other feature subsets. Let $S = S_t \cup S_s$ denote the full feature set, where S_t and S_s are the full temporal and spatial feature sets. On these subsets FIX applies feature selection based on F-scores for each feature to calculate signal-noise discrimination and logistic regression and a linear support vector machine (SVM) for ranking features. Then $S_{sel \supset S}$ denotes a feature subset of S , which contains both temporal and spatial features, denoted S_{t_sel} and S_{s_sel} respectively. The resulting subsets are then $S, S_t, S_s, S_{sel}, S_{t_sel}$ and S_{s_sel} , which were used to train the classifier (all the subsets have a column vector containing the label of each component).

To achieve a robust classification of components using these subsets of extracted features, there exist no ultimately best classifier. Each classifier has its weaknesses and strengths. The k-nearest neighbor algorithm (kNN) is a decent local classifier, but struggles with detecting patterns in a full dataset; SVMs is great at finding decision boundaries with maximum between cluster variance; decision trees are great at finding complex decision boundaries that can be described by a set of if-then rules. To compensate a classifier's weakness through another's strength, FIX employed an ensemble learning method called classifier stacking. Here the output of several single "lower level" classifiers becomes the input in a "higher-level" classifier - a fusion of different classifiers. The "lower level" classifiers consists of decision tree, kNN, SVM with radial basis function (*RBF*) kernel (SVM_r), SVM with polynomial kernel (SVM_p) and linear SVM (SVM_l). Another rea-

Chapter 4. Methods

soring for using these specific classifiers was that they could produce a probability output between 0 and 1, where 0 denoted perfect noise and 1 denoted perfect signal. Each of the six subsets were fed to each of the 5 classifiers, thus producing 30 probability values (5×6 matrix) between 0 and 1 which function as input to the “higher level” classifier. The “higher level” classifier consists of multiple classifiers: SVM_l , SVM_r , random forest, and conditional-inference tree, that learns how to best combine the inputs.

Thus, training this ensemble learner algorithm consisted of selecting the data subsets, training the “lower-level” level classifiers, whose output was the input of the “higher level” classifier, whose output was the probability of a component being noise or signal. A graphical illustration of the classifier design can be seen in figure 4.19.²

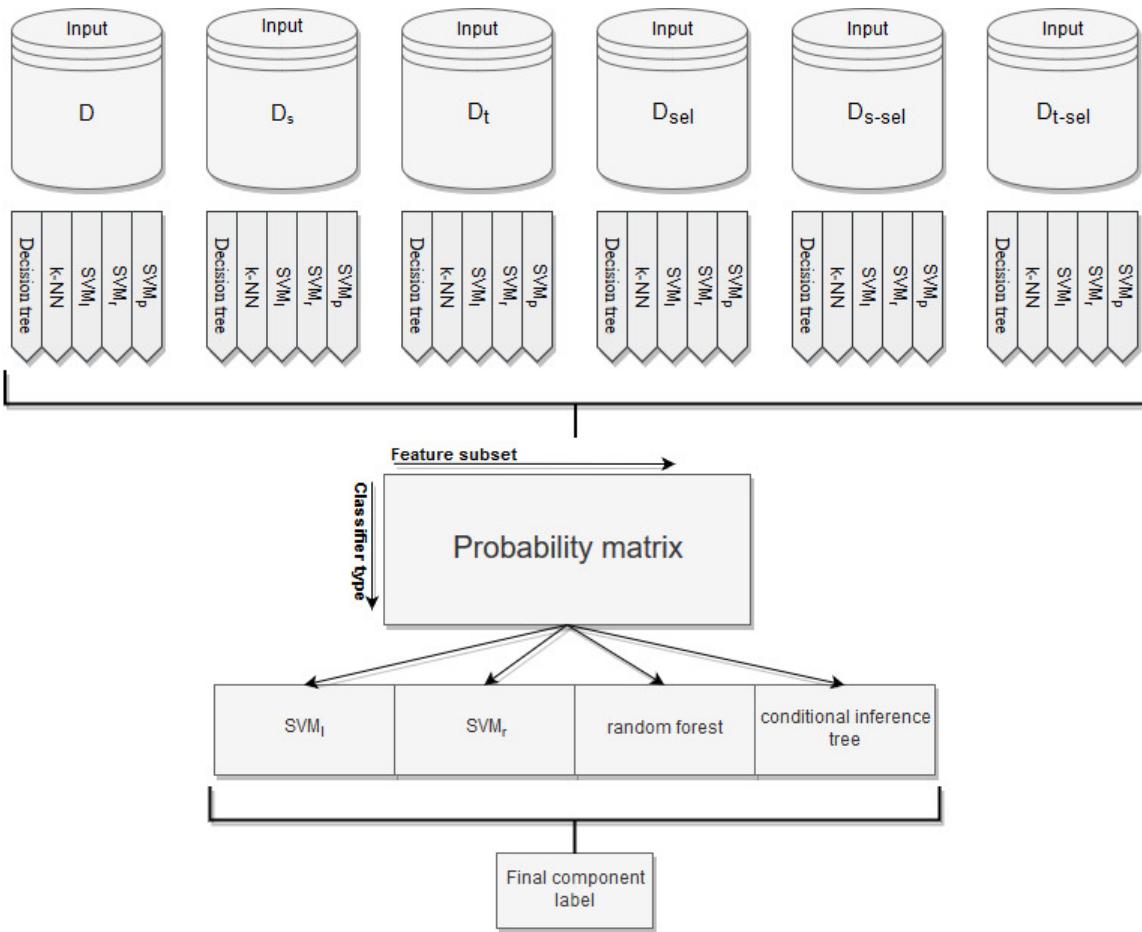


Figure 4.19: An illustration of the classifier design, where the six feature subsets can be seen feed to five different classifiers in the lower level classifier. The output from each of 30 classifiers was a 5×6 probability matrix, containing the probability for each. This matrix was then feed to the higher level classifier consisting of four classifiers, deciding the final component label. Modified from [27].

²FiXme Note: (elaborate on the structure of the higher level classifier.)

4.5.6 Estimating component label threshold

The classifier in FIX required setting a threshold on how soft or aggressive the classifier should be at including potentially noisy signal components. This threshold scale varied from 1-100, where 1 was the most soft, meaning that almost no components containing some portion of signal of interest are labeled as noise.

To be able to perform a quantitative evaluation of which threshold to choose, the FIX software provided the leave-one-out(LOO) classification option. On a data set containing n outputs (e.g. one for each heat run per subject), a cross-validation was performed, where each fold uses a $n-1$ data set for training and tests the obtained decision boundary on the data set left out. This LOO was carried out using the following thresholds: 1, 2, 5, 10, 20, 30, 40, 50. For each heat run and each threshold the following was additionally calculated: true-positive ratio (TPR), true-negative ratio (TNR) and a weighted ratio (WR), $\frac{3*TPR+TNR}{4}$, which favored the TPR. Summarizing these values as the mean across all heat runs yielded an accuracy matrix, from which a decision on which threshold to choose could be based upon.

After performing the LOO classification with the training data from this project, it was observed that a TPR of NaN for all thresholds was calculated for two of the heat runs; possibly due to the heat runs not containing any signal components. This resulted in the mean TPR and the WR being NaN for all thresholds. These two heat runs were therefore excluded from the project. The resultant accuracy matrix of the LOO classification can be seen in table 4.3.

Table 4.3: Accuracy matix from the LOO classifier. TPR denotes the mean percentage of true signals correctly classified. TNR denotes the mean percentage of true noise correctly classified. WR is a weighted ratio that favors TPR. The more the threshold increases the lower the TPR and the higher the TNR.

Threshold:	1	2	5	10	20	30	40	50
TPR:	99.1	98.6	98.4	97.3	96.3	94.3	92.4	91.4
TNR:	78.1	80.9	85.5	90.2	93.7	95.5	96.6	97.7
WR:	93.9	94.2	95.2	95.5	95.7	94.6	93.4	93.0

For this project it was important that much signal was kept, while not including too much noise. This demand was well summarized by the WR. Thus, the threshold with the highest WR (20) was initially chosen. However, it was recommended by the FIX developers to validate the most preferred thresholds quantitatively, before making a final decision. The test data set was therefore run on the trained classifier using the thresholds with a WR above 95 % (5, 10 and 20). Checking the classified labels on three subjects per threshold, left the initial decision on choosing a threshold of 20 to stand.

4.6 Intensity normalization

Before analysis of the two noise cleaned datasets, intensity normalization was applied to both. Intensity normalization was applied due to the assumption of a global signal change within a heat scan. The global signal change is believed to occur due to physiologic fluctuations related noxious stimuli, which creates blood pressure and perfusion changes. The global signal change is of no interest in the statistical analysis and therefore needs to be removed beforehand. [Macey2004]

Intensity normalization was implemented such that each volume in the heat scan was set to have the same mean intensity. This was done by calculating the mean intensity of each volume and the mean intensity for a whole scan. Then each volume is scaled such that each volume in the whole scan has the same mean intensity, forcing the global signal change to become a preset constant.

4.7 Statistical analysis

The purpose of the statistical analysis was to detect the areas of activation associated with the noxious heat stimuli of 48°. A statistical analysis was performed to show statistically significant activation within single heat runs, within single participants and within groups. General linear models (GLM) were used in all analyses. A more detailed explanation of the GLM can be found in the appendix A.6.

4.7.1 Within heat run analysis

On a participant level the aim of the fMRI analysis was to analyze the time courses of every voxel to find out if the BOLD signal changed in response to some perturbation. In the context of this project, we were interested in finding the response to the repeated block designed heat stimulus delivered to the participants. GLM was used for analysing this through the FSL tool FEAT (fMRIB's Expert Analysis Tool) using a multivariate linear regression approach, as shown in figure A.9, section A.6. The BOLD signal time courses made up the dependent variable and the expected BOLD signal response for the 48° stimuli, 47°stimuli and the rating period made up the independent variables. As only the BOLD signal response of the 48°stimuli was the signal of interest, the 47°stimuli and the rating period BOLD signal response regressors were set as nuisance regressors. The nuisance regressors are used to account for the variability in the BOLD signal response of no interest.

The BOLD signal response is not instantaneous with the delivered stimuli, but delayed and blurred as a result of the hemodynamic response. It was therefore modeled by convoluting each of the regressors with a hemodynamic response function (HRF). As a default function FEAT uses a gamma function, which was also applied for this project.

The intercept and estimation parameters were estimated through FEAT by applying the ordinary least square method that minimizes distances between dependent and independent variables. The estimation parameters from the regressor of interest for each voxel

were used to test the hypothesis about whether the activation was associated with the 48°stimuli. A t-test was applied to test if the estimation parameters were significantly different from 0, thus showing an effect. This resulted in a spatial map of activation for a single heat run for one participant, along with the within participant variance σ_W^2 . This process was performed for every heat run for each participant on the data from standard preprocessing and FIX preprocessing, respectively.

4.7.2 Within participant analysis

To combine the results obtained for the individual heat runs for a single participant a fixed effect model was used. This fixed effect model amounted a weighted average of the single heat run effect. This was done through the FIXED algorithm in FEAT that uses a weighted linear regression. Here the weights were the inverse of the within participant variances σ_W^2 estimated for each of the three heat runs and the input was the spatial maps obtained. This was performed for both the standard preprocessed data and the FIX preprocessed data, providing a full spatial map of each subject for both preprocessing methods.

4.7.3 Within group analysis

The aim of the group-level analysis was to combine the results obtained from each participant. This was done by using a mixed effect model, which treats the group as a random sample from a total population. This mixed effect model uses an addition of the within participant variance σ_W^2 and the between participant variance σ_B^2 in a weighted linear regression approach, which takes the spatial maps estimated in the fixed effect model as input for the independent variable. The within participant variances were estimated in the participant-level analysis, while σ_B^2 was estimated through the FLAME (FMRIB's Local Analysis of Mixed Effects) algorithm in the mixed effect model option in FEAT, which uses an iterative process for the estimation. This is all done on a voxel level to provide more sensitive estimations. To achieve a full spatial map of all subjects combined for both the standard preprocessed data and the FIX preprocessed data respectively, a t-test was used to test if the estimated parameters differed significantly from 0 for the data from each of the two preprocessing methods respectively. To test the difference in activation in the data of the two preprocessing methods, the contrast of the hypothesis was simply changed. This provided two spatial maps: one where the activation was significantly higher in the FIX preprocessed data compared to the standard preprocessed data, and one where the activation was significantly higher in the standard preprocessed data compared to the FIX preprocessed data.

Part III

Results

Part IV

Synthesis

Bibliography

- [1] James Dahlhamer et al. "Prevalence of Chronic Pain and High-Impact Chronic Pain Among Adults - United States, 2016." In: *MMWR. Morbidity and mortality weekly report* 67.36 (2018), pp. 1001–1006.
- [2] National Center for Health Statistics Health. "Health, United States, 2006 With Chartbook on Trends in the Health of Americans". In: (2006), p. 559.
- [3] Institute of Medicine. *Relieving Pain in America: a Blueprint for Transforming Prevention, Care, Education, and Research*. Vol. 26. 2. 2011, pp. 197–198.
- [4] Karen D. Davis et al. "Brain imaging tests for chronic pain: Medical, legal and ethical issues and recommendations". In: *Nature Reviews Neurology* 13.10 (2017), pp. 624–638.
- [5] Harold Merskey and Nikolai Bogduk. "Classification of Cronic Pain". In: *International Associasion for the Study of Pain* 2 (1994).
- [6] Peter Brook, Tony Pickering, and Jayne Connell. *Oxford Handbook of Pain Management*. 1st. Oxford University Press, 2011, p. 463.
- [7] Eric L Garland and D Ph. "Pain Processing in the Human Nervous System: A Selective Review of Nociceptive and Biobehavioral Pathways Eric". In: *Vickrey BG, Shekelle P, Morton S, et al. Prevention and Management of Urinary Tract Infections in Paralyzed Persons: Summary. 1999 Jan. In: AHRQ Evidence Report Summaries. Rockville (MD): Agency for Healthcare Research and Quality (US); 1998-2005. 6. Avai* 39.3 (2013), pp. 561–571.
- [8] Robert F. Schmidt. *Fundamentals of Sensory Physiology*. Third edit. 1986.
- [9] Christopher S. Nielsen et al. "Individual differences in pain sensitivity: Genetic and environmental contributions". In: *Pain* 136.1-2 (2008), pp. 21–29.
- [10] Robert C. Coghill. "Individual Differences in the Subjective Experience of Pain: New Insights into Mechanisms and Models". In: 50.9 (2011), pp. 1531–1535.
- [11] A Feizerfan. "Transitions from acute to chronic pain". In: *Acute Pain Management* 15.2 (2015), pp. 98–102.
- [12] Kai McGreevy, Michael M. Bottros, and Srinivasa N. Raja. "Preventing Chronic Pain following Acute Pain: Risk Factors, Preventive Strategies, and their Efficacy". In: *European Journal of Pain Supplem* 5.2 (2011), pp. 365–372.
- [13] Robert C Coghill, John G McChaffie, and Y Yen. "Neural correlates of interindividual differences in the subjective experience of pain". In: *Proceedings of the National Academy of Sciences* 114.48 (2003), E10507–E10507.
- [14] Hyungsuk Kim et al. "Genetic influence on variability in human acute experimental pain sensitivity associated with gender, ethnicity and psychological temperament". In: *Pain* 109.3 (2004), pp. 488–496.

- [15] Nichole M. Emerson et al. “Pain Sensitivity is Inversely Related to Regional Grey Matter Density in the Brain Nichole”. In: 155.3 (2014), pp. 566–573.
- [16] Irene Tracey and Patrick W. Mantyh. “The Cerebral Signature for Pain Perception and Its Modulation”. In: *Neuron* 55.3 (2007), pp. 377–391.
- [17] Allan I. Basbaum et al. “Cellular and Molecular Mechanisms of Pain”. In: 139.2 (2009), pp. 1–6.
- [18] Gary H Glover. “Overview of functional magnetic resonance imaging”. In: *Neurosurg Clin N Am* 22.2 (2011), pp. 133–139.
- [19] Thomas Poldrack A, Russell; Mumford A, Jeanette; Nichols E. *Handbook of functional MRI data analysis*. 2011.
- [20] Mushabbar A. Syed, Subha V. Ramen, and Orlando P. Simonetti. *Basic Principles of Cardiovascular MRI*. 2015, pp. 1–338.
- [21] Nishanth Khanna et al. “Functional neuroimaging: fundamental principles and clinical applications”. In: *Neuroradiology Journal* 28.2 (2015), pp. 87–96.
- [22] Sang Pil Lee, Afonso C. Silva, and Seong Gi Kim. “Comparison of diffusion-weighted high-resolution CBF and spin-echo BOLD fMRI at 9.4 T”. In: *Magnetic Resonance in Medicine* 47.4 (2002), pp. 736–741.
- [23] J Dubois and R Adolphs. “Building a science of individual differences from fMRI”. In: *Trends in Cognitive Sciences* (2016), pp. 425–443.
- [24] Robert C. Coghill et al. “Pain Intensity Processing Within the Human Brain: A Bilateral, Distributed Mechanism”. In: *Journal of Neurophysiology* 82.4 (1999), pp. 1934–1943.
- [25] R Wood et al. “1.5 Tesla magnetic resonance imaging scanners compared with 3.0 Tesla magnetic resonance imaging scanners: systematic review of clinical effectiveness”. In: *CADTH Technology Overviews* 2.2 (2012), e2201.
- [26] Katherine S. Button, John P. A. Ioannidis, and Claire Kokrysz. “Power failure: why small sample size undermines the reliability of neuroscience Katherine”. In: *Nature Reviews Neuroscience* 33.1 (2013), pp. 365–376.
- [27] Gholamreza Salimi-Khorshidi et al. “Automatic denoising of functional MRI data: Combining independent component analysis and hierarchical fusion of classifiers”. In: *NeuroImage* 90 (2014), pp. 449–468.
- [28] Kevin Murphy, Rasmus M. Birn, and Peter A. Bandettini. “Resting state fMRI confounds and cleanup”. In: *NeuroImage* 80 (2013), pp. 349–359.
- [29] Chia Shang J Liu et al. “Spatial and temporal characteristics of physiological noise in fMRI at 3T”. In: *Academic Radiology* 13.3 (2006), pp. 313–323.
- [30] Armin Iraji et al. “The connectivity domain: Analyzing resting state fMRI data using feature-based data-driven and model-based methods”. In: (2016), pp. 494–507.

Bibliography

- [31] Martin Monti. “Statistical Analysis of fMRI Time-Series: A Critical Review of the GLM Approach”. In: *Frontiers in Human Neuroscience* 5.March (2011), pp. 1–13.
- [32] Gary H. Glover, Tie Qiang Li, and David Ress. “Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR”. In: *Magnetic Resonance in Medicine* 44.1 (2000), pp. 162–167.
- [33] Yashar Behzadi et al. “A Component Based Noise Correction Method (CompCor) for BOLD and Perfusion Based fMRI”. In: 6.8 (2013), pp. 90–101.
- [34] Martin J McKeown et al. “Analysis of fMRI Data by Blind Separation Into Independent Spatial Components”. In: 188.June 1997 (1998), pp. 160–188.
- [35] V. D. Calhoun et al. “fMRI activation in a visual-perception task: Network of areas detected using the general linear model and independent components analysis”. In: *NeuroImage* 14.5 (2001), pp. 1080–1088.
- [36] Johnathan Deslauriers et al. “Increase of posterior connectivity in aging within the Ventral Attention Network: A functional connectivity analysis using independent component analysis”. In: *Brain Research* 1657 (2017), pp. 288–296.
- [37] Linden Parkes et al. “An evaluation of the efficacy, reliability, and sensitivity of motion correction strategies for resting-state functional MRI”. In: *NeuroImage* 171.July 2017 (2018), pp. 415–436.
- [38] Yuhui Du et al. “Identifying functional network changing patterns in individuals at clinical high-risk for psychosis and patients with early illness schizophrenia: A group ICA study”. In: *NeuroImage: Clinical* 17.May (2018), pp. 335–346.
- [39] Jussi Tohka et al. “Automatic independent component labeling for artifact removal in fMRI”. In: *NeuroImage* 39.3 (2008), pp. 1227–1245.
- [40] Oxford Centre for Functional Magnetic Resonance and Imaging of the Brain (FM-RIB); *MELODIC/FAQ*. 2016.
- [41] Ludovica Griffanti et al. “Hand classification of fMRI ICA noise components”. In: *NeuroImage* 154.June 2016 (2017), pp. 188–205.
- [42] Stephen M. Smith. “Fast robust automated brain extraction”. In: *Human Brain Mapping* 17.3 (2002), pp. 143–155.
- [43] Joseph V. Hajnal, Derek L.G. Hill, and David J. Hawkes. *Medical Image Registration*. Mecheal R. Vol. 39. 5. The BIOMEDICAL ENGINEERING Series, 2001, p. 394.
- [44] Mark Jenkinson and Stephen Smith. “A global optimisation method for robust affine registration of brain images”. In: *Medical Image Analysis* 5.2 (2001), pp. 143–156.
- [45] JLR Andersson, M Jenkinson, and S Smith. “Non-linear registration aka Spatial normalization. FMRIB Tech Rep TR07JA2.” In: June (2007).
- [46] Massieh Moayedi, Tim V. Salomons, and Lauren Y. Atlas. “Pain Neuroimaging in Humans: A Primer for Beginners and Non-Imagers”. In: *Journal of Pain* 19.9 (2018), 961.e1–961.e21.

- [47] Mark Jenkinson et al. “Improved optimization for the robust and accurate linear registration and motion correction of brain images”. In: *NeuroImage* 17.2 (2002), pp. 825–841.
- [48] FMRIB. *FEATUserGuide*. 2018.
- [49] Stephen M. Smith and J. Michael Brady. “SUSAN - A new approach to low level image processing”. In: *International Journal of Computer Vision* 23.1 (1997), pp. 45–78.
- [50] Christian F. Beckmann and Stephen M. Smith. “Probabilistic Independent Component Analysis for Functional Magnetic Resonance Imaging”. In: *IEEE Transactions on Medical Imaging* 23.2 (2004), pp. 137–152.
- [51] Yi-Ou Li, Tülay Adalı, and Vince D. Calhoun. “Estimating the number of independent components for functional magnetic resonance imaging data”. In: *Human Brain Mapping* 28.11 (2007), pp. 1251–1266.
- [52] Waqas Majeed and Malcolm J. Avison. “Robust data driven model order estimation for independent component analysis of fMRI data with low contrast to noise”. In: *PLoS ONE* 9.4 (2014).
- [53] V.D. Calhoun et al. “Spatial and temporal independent component analysis of functional MRI data containing a pair of task-related waveforms”. In: *Human Brain Mapping* 13.1 (2001), pp. 43–53.
- [54] Jieun Kim et al. “Sustained deep-tissue pain alters functional brain connectivity”. In: *Pain* 154.8 (2013), pp. 1343–1351.
- [55] Nathalie Erpelding and Karen D. Davis. “Neural underpinnings of behavioural strategies that prioritize either cognitive task performance or pain”. In: *Pain* 154.10 (2013), pp. 2060–2071.
- [56] A. A. Bharath. *Introductory Medical Imaging*. Vol. 3. 1. 2008, pp. 1–186.
- [57] John C Edwards. “Principles of NMR”. In: () .
- [58] Govind B. Chavhan et al. “Principles, Techniques, and Applications of T2*-based MR Imaging and Its Special Applications”. In: *RadioGraphics* 29.5 (2009), pp. 1433–1449.
- [59] Michael W L Chee et al. “Comparison of block and event-related fMRI designs in evaluating the word-frequency effect”. In: *Human Brain Mapping* 18.3 (2003), pp. 186–193.
- [60] A Hyvärinen, Juha Karhunen, and Erkki Oja. “Independent Component Analysis”. In: (2001).
- [61] A Hyvärinen and Erkki Oja. “Independent component analysis: algorithms and applications.” In: *Neural networks : the official journal of the International Neural Network Society* 13.4-5 (2000), pp. 411–30.
- [62] John L Semmlow. *Biosignal and Biomedical Image Processing*. 2004.
- [63] CIS 520. *PCA*. 2018.

Bibliography

- [64] Jean Baptiste Poline and Matthew Brett. “The general linear model and fMRI: Does love last forever?” In: *NeuroImage* 62.2 (2012), pp. 871–880.

Bibliography

A | Appendices

A.1 MRI physics

Magnetic Resonance Imaging (MRI) is a non-invasive imaging technology, which does not involve potentially damaging ionizing radiation as in other scanners, eg. CT and X-ray. MRI is especially suited for representing soft tissue portions of the body, which makes it a widely used technology in brain imaging, both for clinical and research purposes, as it can depict the anatomical structure at a millimeter resolution. This section will provide information on the physics behind MRI and which physiological properties that can be exploited to create an image of the body.

Magnetic resonance imaging (MRI) is founded on the principle of nuclear magnetic resonance (NMR), which exploits the magnetic properties of the hydrogen nucleus that contains a single proton. The proton is not static, but rotates around its own axis. As the proton is positively charged it creates a magnetic moment in the direction described by the thumb rule, and can interact with an external magnetic field. The human body consists of approximately 10% hydrogen atoms, but as the hydrogen nuclei spins are randomly orientated, the net magnetic moment equals zero, as the nuclei cancel each other out. Placing the body in a strong magnetic field will align the nuclei. A property of the hydrogen nucleus is its quantum spin rate, which can either be $\frac{1}{2}$ or $-\frac{1}{2}$ either in the direction or the opposite direction of the main magnetic field. Most will align in the direction of the magnetic field, while the rest align in the opposite direction, possibly as a result of heat radiation absorbed by the nuclei. The direction of the nucleus is determined by its energy level, leaving the former in a low energy state and the latter in a high energy state. The nuclei do not simply point in the direction or opposite the direction of the magnetic field, but precess. [56] The rate of precession can be calculated by the Lamour frequency:

$$f = \gamma * B_0 \quad (\text{A.1})$$

f is precession frequency, γ is gyroscopic ratio and B_0 is magnetic field strength. The equation states that the precession frequency is proportional to the strength of the magnetic field. After canceling out all opposing precessing nuclei, the net magnetization, or longitudinal magnetization, will point in the direction of the external magnetic field. However, the longitudinal magnetization can not be detected directly as it points in the direction of the strong external magnetic field. Additional techniques are therefore used in NMR, to facilitate a detectable signal. [56] A depiction of how the nucleus precessing can either align along or opposite to the magnetic field depending on its energy state, can be found in figure A.1.

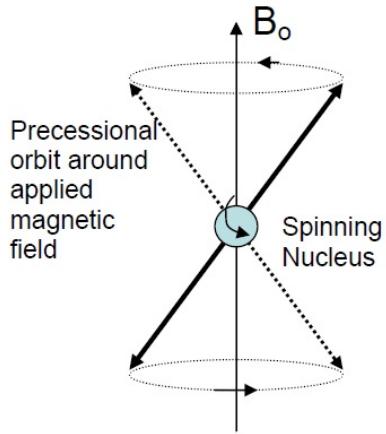


Figure A.1: The figure illustrates how the nucleus precess and spin in relation to the applied magnetic field B_0 surrounding it. The vectors, indicating the precessing, can go opposite or along the magnetic field depending on the nucleus energy state. [57]

A radio frequency pulse (RF pulse) tuned to the precession of the nuclei is transmitted in the vicinity of the nuclei. The RF pulse is absorbed by the nuclei and more, favorably half of the targeted nuclei population, will enter the high energy state, leaving the longitudinal magnetization to equal zero. The number of nuclei that flip is determined by the amount of energy the RF pulse injects, and the nuclei only exchange energy efficiently if the frequency of the energy from the RF pulse matches the precession rate. The RF pulse furthermore shifts the precession of the nuclei into same phase angle, which creates resonance, and a net magnetization pointing 90° to the longitudinal magnetization. This magnetization is called the transverse magnetization. The coherent nuclei produce a radio signal, or free induction decay signal (FID signal), that can be detected by a radio antenna. After the RF pulse is removed, the nuclei will relax into baseline state. Firstly, the spins of the nuclei will repel each other, as they are positively charged, and thus shift phase. The net magnetization will return to zero. This relaxation is called T_2 or “spin-spin” relaxation, as the energy exchange between the nucleus spins is causing the relaxation. An illustration of the T_2 relaxation can be seen in figure A.2.

Appendix A. Appendices

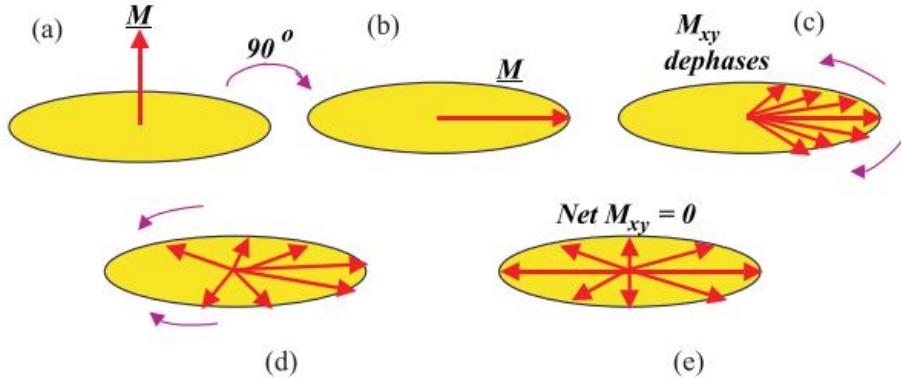


Figure A.2: An illustration of the T_2 relaxation. Before the RF pulse is emitted the majority of the nuclei are precessing out of phase in the low energy state and the net magnetization is pointing in the direction of the longitudinal magnetization, (a). The nuclei absorb the RF pulse, the magnetization shifts 90° and the nuclei precess in phase, which leaves a high transverse magnetization and FID-signal, (b). Due to interaction between the nuclei spins, the nuclei gradually dephase, until the net magnetization is 0, which is referred to as T_2 relaxation. [56]

A second relaxation appears as the high energy nuclei returns to the low energy state. The energy that was previously absorbed by the nuclei is dissipated in to the surrounding lattice in the form of heat. During this relaxation the longitudinal magnetization is regrown. This relaxation is called T_1 or “spin-lattice” relaxation, as the spins transfer energy to the surrounding lattice. [56] The hydrogen nuclei are located in different local environments in the body. Some are for instance associated with free-floating water molecules, while others are associated with structural and storage molecules such as proteins and lipids, and thus more fixed in position. The nuclei have different T_1 and T_2 relaxation characteristics, depending on the local environment or tissue they are associated with. This can be accentuated and measured in NMR. [56]

The chosen pulse sequence is key to how the tissue will be portrayed in an image, and is described by the T_{echo} , time before the FID signal is measured, and T_{rep} , time before a new RF pulse is applied. In a case of nuclei associated with lipids and water molecules, the nuclei in lipids are fixed and will have a fast T_1 relaxation after exposure to a RF pulse. Meanwhile the nuclei in the water molecules will maintain being in a synchronized phase. At T_{echo} , the nuclei associated with the lipids will have a low amplitude FID signal, as the transverse magnetization is weak, and the nuclei associated with the water molecules will have a high amplitude FID signal, as the transverse magnetization is strong. The water molecules will be assigned a white color on a greyscale image and the lipids as dark grey/black. In this case there is a long T_{echo} and a long T_{rep} , and is referred to as T_2 -weighted MRI. A commonly used T_2 pulse sequence is the spin echo (SE). Due to magnetic field inhomogeneities the nuclei dephase more rapidly. Applying a second RF pulse of 180° after the nuclei have dephased, the nuclei will rephase. Again the contrast in the image is expressed as a result of the relaxation time of the nuclei, which depends on the surrounding environment. . [56]

In case of T_1 -weighted MRI, the T_{echo} and T_{rep} are short. As in T_2 -weighted MRI a RF pulse is applied and the nuclei associated with lipids will quickly return to baseline state

and the water molecule nuclei will remain a strong transverse magnetization. At this time point a second RF pulse will be induced, referring to the short T_{rep} . Now the lipid nuclei will return to a strong transverse magnetization state and excite a high FID signal. More low energy state nuclei of the water molecules will absorb the RF pulse and shift to a high energy state, leaving a majority of nuclei in a high energy state. The water molecule nuclei now has a weak transverse magnetization and 180 degrees longitudinal magnetization, thus producing a low-amplitude FID signal. A short T_{echo} after the second RF pulse then shows lipids as white and the water molecules as dark grey/black in a greyscale image. [56]

T_2 relaxation was defined as dephasing due to energy exchange between nuclei. In practice the T_2 relaxation time happens much faster than would be predicted by these natural atomic interactions. Another factor in the dephasing of nuclei is inhomogeneities in the magnetic field. This observed T_2 relaxation is referred to as T_2^* relaxation. In a T_2^* - weighted MRI, a gradient echo (GRE) is used instead of a SE. In a GRE pulse sequence only one RF pulse is emitted with a low flip angle, and the echo time is therefore usually shorter. A gradient is applied after the initiating RF pulse, which enhances the dephasing. When the net magnetization is zero a rephasing gradient with opposite polarity of the dephasing gradient is turned on, which reverses the phase shift. A FID-signal is produced as a GRE. The gradient only reverses the phase shifts that have been affected by the gradient itself, and not those affected by magnetic field inhomogeneities. Contrast in tissues is therefore not decided through natural T_2 relaxation but by T_2^* . Thus, T_2^* -weighted MRI only works well in scanners that do not lack magnetic field homogeneity. Due to the fast acquisition time T_2^* -weighted MRI is widely used in functional MRI to image brain activity. [58]

A.2 MR image reconstruction

Different pulse sequences and certain physiological properties that can be exploited with certain pulse sequences, have been laid out in the previous sections. This section aims to describe how the corresponding echo signals are reconstructed as a MR image.

Following the Lamour frequency equation (A.1), the main magnetic field causes all hydrogen nuclei to precess with the same frequency. Without any specification of spatial localization a MRI of a human body would consist of a single number. To prevent this, separate coils in the x, y and z directions are introduced. These coils can be adjusted in position, and thus produce gradient magnetic fields with a varying strength depending on position. According to the Lamour frequency the nuclei will precess with different frequencies when in a magnetic field with varying strength. The gradients can be turned on in combination to create any direction in space. These varying frequencies can be exploited to separate parts of the anatomy and ultimately illustrate a desired area. As mentioned, the nuclei only exchange energy efficiently if the frequency of energy, or RF pulse, matches the precession rate. Thus, by altering the magnetic field along the body in one direction, z-direction for the sake of the example, the nuclei will have slightly different

Appendix A. Appendices

precession rates, and the RF pulse will only efficiently affect a desired slice of the nuclei. The nuclei of that slice now precess at the same rate. To get an image with a spatial resolution, the voxels that make out the image needs to be discriminated between. By turning on the gradient of the x-direction the lines in the y-direction are now encoded with a particular frequency. This gradient functions as a frequency encoding gradient, and is illustrated in figure A.3. [56]

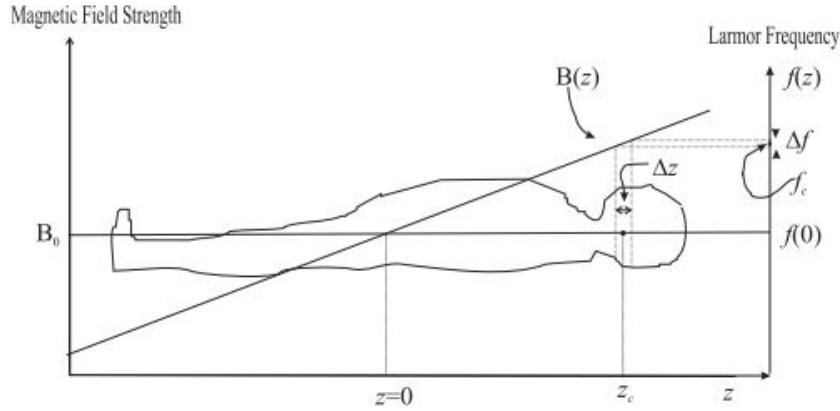


Figure A.3: The position of the slice is specified through the direction of the frequency encoding gradient ($B(z)$) and though the central frequency of the emitted RF pulse (f_c). The thickness of the slice is dependent on the steepness slope of $B(z)$, and on the bandwidth of the emitted RF pulse (δf). [56]

Turning the y-gradient on and quickly off, will de-phase the nuclei while still remaining the same frequency as before. This gradient functions as a phase encoding gradient. When comparing two locations approximately one voxel apart in the x-direction, then based on the amount of gradient strength difference, there will be a certain amount of change in phase between the spins spread across that distance. The farther away from isocenter, where the magnetic field strength is B_0 , the higher the change in phase will be. This notion is used to assign the correct spatial location of each voxel, when reconstructing the FID signals into an image. This phase encoding procedure is done in different gradient strengths in iterations to assign unique phases to the nuclei in the both directions. One iteration of a certain strength of the phase encoding gradient followed by a measurement is performed at a time. The only change per iteration is the phase encoding gradient strength. These iterations are then series of measurements acquired at different points in time, where each entry of the slice then represent a certain signal intensity. This time domain measurement is referred to at the raw data. [56]

The next step is to Fourier Transform (FT) the raw data, which will yield frequency information to the acquired signal intensities. This step gives a summation of the signal intensities at the different frequencies produced by the frequency encoding gradient. This is called the k-space, as the k-numbers of a signal describes its relative orientation and frequency. The k-space image contains the contrast in the center and the resolution in the periphery, as there is low or no phase encoding at the center and increasing towards the periphery, giving more brightness in the center and dimmer tones in the periphery. To

allocate the voxels in correct spatial localization an inverse 2D discrete FT is performed on the k-space image. This provides the desired image of the anatomy slice. [56] Figure A.4 shows the acquired signal in k-space and the reconstructed image after the Fourier Transform.

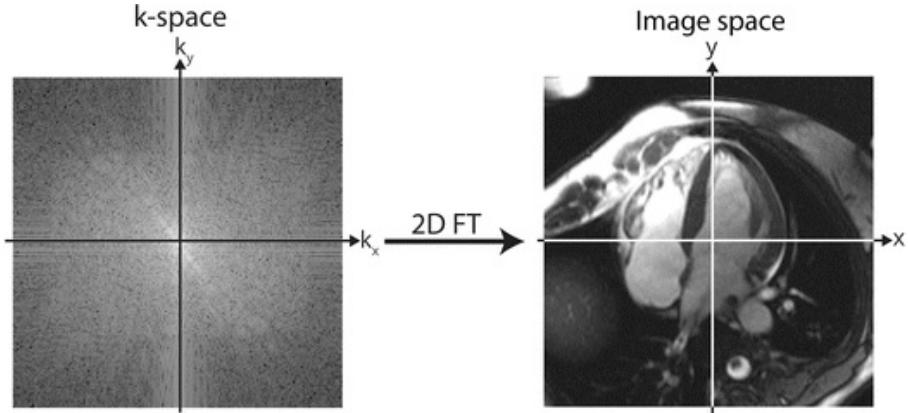


Figure A.4: A depiction of the acquired signal represented in k-space and the resulting reconstructed image after the inverse 2D Fourier Transform [20].

A.3 Stimuli Design

The following section will describe different standards of designing experiments, where the impact of a given stimuli is used to assess the subsequent brain activation. The impact of stimuli on the hemodynamic response and how it is transformed into the hemodynamic response function (HRF) used analysis will be further explained.

In order to accomplish a well designed experiment, the researcher must consider the multiple types of stimuli delivered, including the form and duration of these. The researcher must be aware of the timing of events in the scanning session and any responses provoked. Additionally, the researcher should have general knowledge of where in the brain activation is seen and how the hemodynamic response will be presented. [46]

Doing cognitive experiments using fMRI, two main design types are utilized, by either using a block- or event-related design. Event-related design is inducing a series of very short lasting stimuli used to investigate single hemodynamic response. A characteristic of this method is that it permits the possibility of increasing and decreasing the interval between stimuli. Thereby the theoretical likelihood of subject confounds should be reduced as the interval would not become predictable. Event-related stimuli design further allows more temporals characteristics to be inspected, compared to a block design. Characteristics could be hemodynamic response in duration and amplitude. [59]

Block design works by performing a series of less but longer stimuli. Block designs are ideal for experiments involving detection of small differences in BOLD signal across various test conditions where its statistical power is superior. Furthermore, if there is artifacts present they are more easily detected in the signal time course, because of the signals temporal structure. A block design is easier to design than an event-related, as randomization of

Appendix A. Appendices

intervals of stimuli is not required. The design instead focuses on the total number of stimuli used, block length, inter stimulus interval, block length and TR. An illustration of both design types can be found in figure A.5. [59]

To enable use of the hemodynamic response in fMRI analysis, it needs to be transformed in order to represent the ideal physiologic response to the stimulus. Thus, portraying the biological delay from stimuli to response. Therefore stimulus in the design is combined with the hemodynamic response function through convolution. Thus making a function that models how the BOLD signal would be represented if the voxel activity increased in a given area each time stimuli is induced. [46]

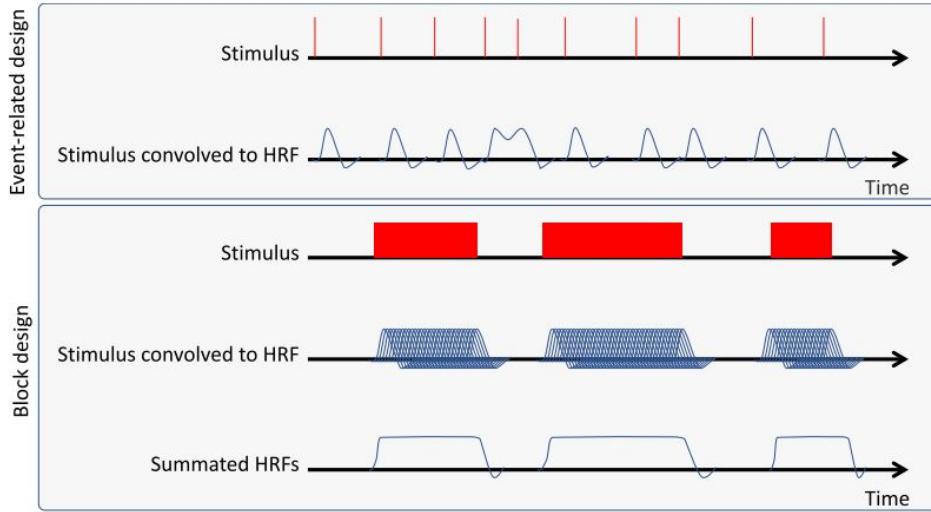


Figure A.5: The top image show the event-related stimuli design and the stimulus convolved with the hemodynamic response function. The lower image depicts a block design, the stimulus convolved with hemodynamic response function and the summated HRF response. [46]

A.4 Independent Component Analysis

Independent component analysis has proven to be a very useful tool in separating noise from wanted signal in many applications. The following section will seek to elucidate the general basic theoretical background behind the blind source separation analysis.

The basic idea about Independent Component Analysis (ICA) is to recover m signal sources, which are mixed in n observed signals. The observed signal is given by [60]:

$$\mathbf{x} = \mathbf{As} \quad (\text{A.2})$$

where \mathbf{x} is a observed signal vector containing the mixed signal elements x_1, x_2, \dots, x_n , \mathbf{s} is the source signal vector with the elements s_1, s_2, \dots, s_m and \mathbf{A} is a mixing matrix with the dimension $n \times m$. Note that the dimensions of the mixing matrix can be equal to each other, meaning that it is required to have at least the number of observed signals

as the source signals. The observed signal is assumed to be a linear mix of independent source signals. When performing ICA the goal is to find an inverted mixing matrix A^{-1} that recovers the source signal [60]:

$$\mathbf{s} = \mathbf{A}^{-1}\mathbf{x} \quad (\text{A.3})$$

This can easily be achieved if the mixing matrix is known. However, this is rarely the case, as both the mixing matrix and source signals are unknown. There is no reliable way, to fully determine \mathbf{s} , thus a set of assumptions are required, which the ICA method is based on [60]:

- The independent components (IC)'s are assumed to be statistically independent.
- The IC's must be non-Gaussian distributions.

Regarding the assumption of independence, a set of variables y_1, y_2, \dots, y_n is not allowed to share mutual information so that $i \neq j$. This can be expressed as the joint probability of the variables is equal the product of each marginal probability of the variables [60]:

$$p(y_1, y_2, \dots, y_n) = p(y_1) \cdot p(y_2) \cdot \dots \cdot p(y_n) \quad (\text{A.4})$$

Satisfying this condition assures independency of the variables. A graphical example of independence in two dimensions is shown in figure A.6 [60, 61].

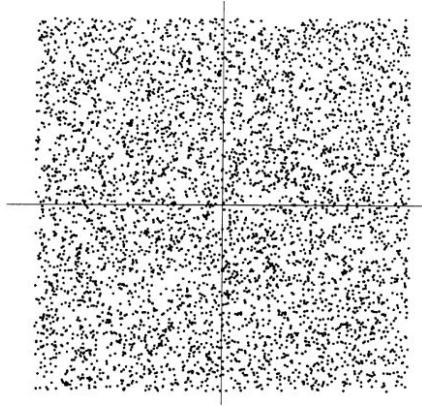


Figure A.6: Figure illustrating the joint distribution of two independent components having uniform distribution. [61].

It is clear to see that information about a variable on the horizontal axis does not give any information about variables at vertical axis and vice versa. Note that uncorrelatedness does not equal independency. However, whitening of the observed signals and thus ensuring uncorrelatedness is very helpful in solving the ICA problem. Whitening of the data is archived by performing a linear transformation that transforms the components

Appendix A. Appendices

so that the covariance matrix equals the identity matrix, thus having unit-variance. This can be expressed through performing a linear transformation of \mathbf{x} into a random vector \mathbf{z} :

$$\mathbf{z} = \mathbf{V}\mathbf{x} = \mathbf{V}\mathbf{A}\mathbf{s} \quad (\text{A.5})$$

A graphical example of whitened data in two dimensions is shown in figure A.7 [60, 61].

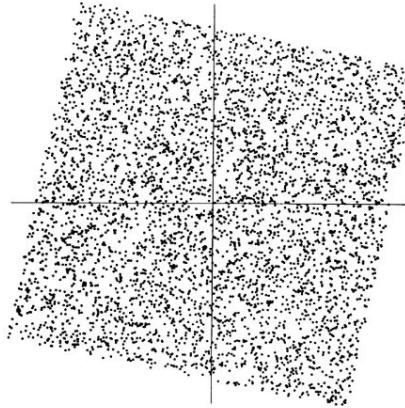


Figure A.7: Figure illustrating the impact of whitening on the joint distribution of two components [61].

The squared distribution is clearly a rotated form of the independent data. What is left to ensure independence and solving the ICA problem is to estimate an angle that gives the correct rotation. Concerning the second assumption on non-Gaussianity, the joint distribution of uncorrelated Gaussian distributions are not necessarily independent, and the distribution will be symmetrical and no information on the direction of the distribution can be derived. Thus, no transformation that allows independency can be performed, and the mixing matrix \mathbf{A} can not be estimated from the mixtures. This can be shown in a two dimensional example with two sources s_1 and s_2 and the mixing matrix \mathbf{A} , the joint probability density function (pdf) of Gaussian distributions is calculated as [60]:

$$\begin{aligned} p(s_1, s_2) &= \frac{1}{2\pi} \exp\left(-\frac{s_1^2 + s_2^2}{2}\right) \\ &= \frac{1}{2\pi} \exp\left(-\frac{\|\mathbf{s}\|^2}{2}\right) \end{aligned} \quad (\text{A.6})$$

For an orthogonal mixing matrix \mathbf{A} the inverse can be written as $\mathbf{A}^{-1} = \mathbf{A}^T$, and then $\mathbf{s} = \mathbf{A}^T \mathbf{x}$. The joint pdf can then be rewritten as:

$$p(x_1, x_2) = \frac{1}{2\pi} \exp\left(-\frac{\|\mathbf{A}^T \mathbf{x}\|^2}{2}\right) |\det(\mathbf{A}^T)| \quad (\text{A.7})$$

Because $\| \mathbf{A}^T \mathbf{x} \|^2 = \| \mathbf{x} \|^2$ and $| \text{Det}(\mathbf{A}^T) | = 1$, hence the orthogonality of \mathbf{A} , the joint pdf is:

$$p(x_1, x_2) = \frac{1}{2\pi} \exp\left(-\frac{\| \mathbf{s} \|^2}{2}\right) \quad (\text{A.8})$$

The joint pdf is not changed when choosing an orthogonal mixing matrix as well as the property of independence. No further information about the mixing matrix can thus be revealed when the source signals are from a Gaussian distribution. [60]

A.4.1 ICA approaches

As the mixing matrix and source signal most often are unknown, the IC's must be approximated in an iterative process. There are two main ICA iteration approaches: minimizing mutual information and maximizing non-Gaussianity. The former approach seeks to minimize mutual information by maximizing independence between components. In practice this is done by minimizing the difference between the joint density distribution and the product of the marginal density functions, the left-hand side and right-hand side of equation (A.4). This can for instance be done through a Kullback-Leibler divergence between the $p(y_1, y_2)$ and $p(y_1) * p(y_2)$ in a two dimensional case. The second approach seeks to maximize non-Gaussianity of the components. The central limit theorem states that if sources are mixed, the mix tend to get more Gaussian than the individual sources. The strategy here is to find the directions in the data that is as far away from Gaussian as possible through a linear transformation. That direction will most likely be independent components. To find these directions different ICA approaches use fourth order moments and negentropy of the data. [60]

A.5 Principal Component Analysis

The use of principal component analysis (PCA) in denoising algorithms has been exploited in various studies, showing that the properties of the PCA might be useful [33]. In this section the theory and common use of PCA is presented, including a small part on its use in images.

Principal Component Analysis (PCA) is a well renowned and widely used analysis tool, capable of finding the most defining variables in a dataset. This facilitates finding the components that are the most saying for the dataset. This introduces the possibility of dimensionality reduction by lowering the amount of redundant information. The PCA is used to transform a set possibly correlated variables into a set of uncorrelated components, called principle components. Each principal component (PC) is orthogonal on the former and are uncorrelated and have zero covariance. They each define the largest variance in an axis, such that PC 1 describes the direction of the maximum variance of the dataset. Each following PC describes the next highest variance of the dataset, with the constraint that it is orthogonal and has zero covariance with any of the former PCs. PCA is the orthogonal projection of data onto a lower dimension linear space. A PC is found by minimizing the variance by projecting the feature values (blue dots) onto the line (now

Appendix A. Appendices

red dots) describing the highest variance in the data set (black line) as seen on figure A.8. The PC is found by minimizing the mean square distance between the data points. [62]

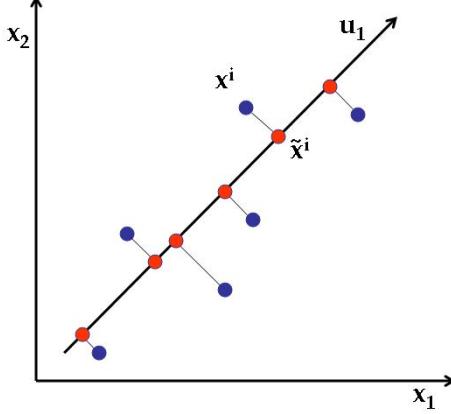


Figure A.8: Two-dimensional example of projection of data variables (blue dots) onto a PC axes (black line). (u_1) indicates the direction of the eigenvector. [63]

The algebraic method of calculating the PCs can be done by using Singular Value Decomposition (SVD). The first step is to compute the squared cross product matrix of variances and covariances among every pair of the variables in the data set, where the diagonals are the variances and the off-diagonals are the covariances, as done in the following equation:

$$S = X'X \quad (\text{A.9})$$

Where S is the cross product and X is the dataset matrix. When finding the PCs it includes an eigen-analysis of S. The eigenvalues of are solutions to the following equation:

$$|S - \lambda I| = 0 \quad (\text{A.10})$$

Where λ is the variances of each PC and I is the identity matrix. After solving for λ the eigenvectors can be solved through the following equation:

$$\det|S - \lambda I| b_i = 0 \quad (\text{A.11})$$

Where b_i is used to calculate the eigenvectors as in:

$$u_i = \frac{b_i}{\sqrt{b_i' b_i}} \quad (\text{A.12})$$

Where u_i is the i number of eigenvectors that contain a contribution to the principal components. The SVD orders the eigenvalues by size $\lambda_1 > \lambda_2 \dots > \lambda_i$. The scores for each PC is equal to the corresponding eigenvalue for that exact axis. The eigenvalues describe how much of the variance is accounted for by the associated PC. Summation of all eigenvalues accounts for the total variance of the data set; this is called the trace. To find how much the each PC accounts for, the eigenvalue of that PC is divided by the total variance: $\% \text{ of total variance} = \frac{\lambda_i}{\text{Trace}}$. This can be used for deciding how many components are significant and by how much the dataset can be reduced. [62]

A.6 General Linear Model

In order to determine if a task or stimuli has made a significant contribution to the measured BOLD signal, each voxel in the scan needs to be evaluated. The following section will therefore seek to explain a method of how to determine if any significant change is found. In this project the general linear model (GLM) will be explained and applied as it is the most used and recognized tool for fMRI analysis during the past 20 years [64].

The GLM is a well considered and used analysis tool constituting a simple way of doing standard statistical analysis on fMRI. The overall goal of the GLM model is to determine how well the time course of the scan corresponds to the known used experimental interference. In an fMRI case, that means, how well the BOLD signal over time fits the time course of the predicted signal given by the imposed stimuli, that causes brain activity. This statistical test is carried out for each voxel independently of its neighbor across the scan resulting in thousands of statistical tests in one scan. [46, 31] To explain the implementation of a GLM we can use the example presented by [31]. An acquired BOLD signal will throughout a time series, of n images, have a varying output signal. The signal in a voxel Y , at any time point, can be seen as a summation of predictor variables (regressors of interest) X and nuisance regressors (regressors of none-interest) that model noise ϵ , a scaling parameter for each regressor β , and an error term ϵ . This can be presented in the GLM's basic form as [31]:

$$Y = X\beta + \epsilon \quad (\text{A.13})$$

Expanding this formulation into a complete scan, including multiple regressors, can be presented on a matrix form. A depiction of a GLM design matrix incorporating the factors making up the BOLD signal can be seen in figure A.9.

Appendix A. Appendices

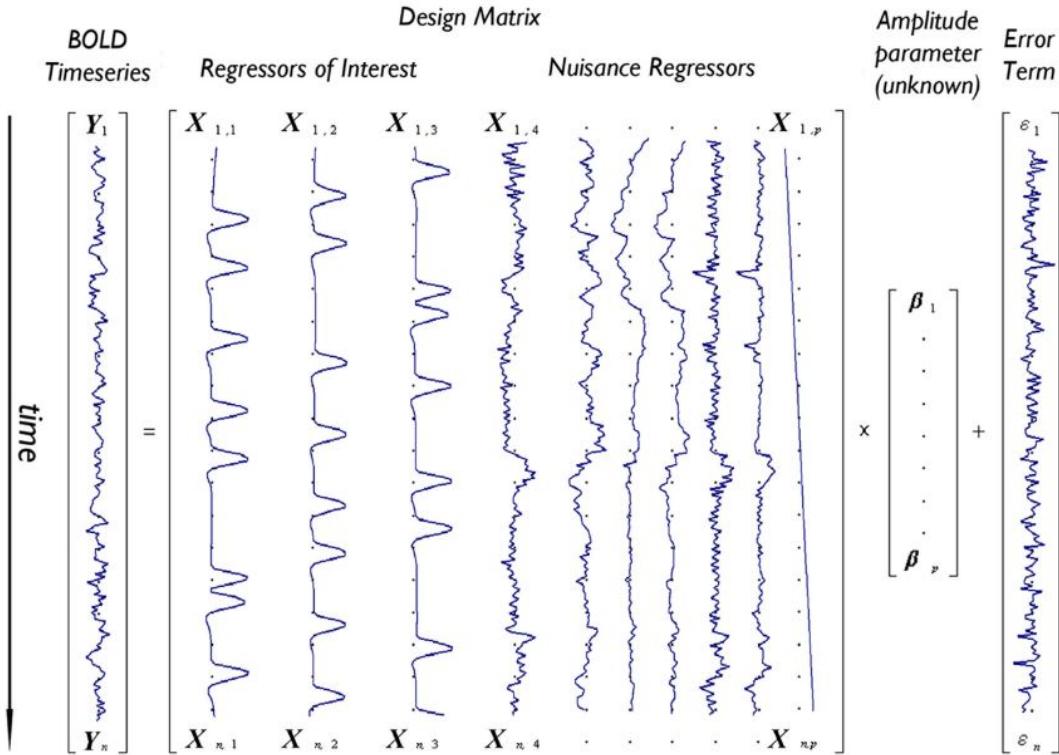


Figure A.9: Illustration of a design matrix for a given GLM model, describing the BOLD time series Y_n , the regressors of interest $X_{1,1}, X_{1,2}, X_{1,3}$, the nuisance regressors $X_{4..p}$, the amplitude parameter $\beta_{1..p}$ and the error terms $\epsilon_{1..n}$. [31]

The predictor variables are found by modeling what is known or predicted as output. The predictors of interest would come from hemodynamic response curve convoluted with stimuli design as presented in section A.3. GLM models will often introduce nuisance predictors as well used to model variables such low frequency drift and motion to make a more robust model. These are non interest regressors, as they do not resemble the wanted signal, being the stimuli induced. The amplitude parameter is the unknown weight, scaling the magnitude of resemblance between each predictor value and the output. They describe the strength of the relationship between that regressor and the voxel’s BOLD signal course of activation. The error term contains the value for each observation, which can not be explained by the weighted sum of the amplitude parameter. [46, 31]

The goal is to estimate the value of the scaling parameter β for all regressors and afterwards determine if any regressor significantly account for the variance found in the BOLD signal. A regressor associated with the measured BOLD signal should hypothetically show a greater value for the voxels in the brain area corresponding to a given task. E.g. finger tapping would show greater β values in the motor cortex. A method for estimating β values for the regressors is the ordinary least squares (OLS). The OLS method estimates β by minimizing the sum of squared residuals [31]:

$$\sum_{i=1}^n (Y_i - X_i \times \hat{\beta})^2 \quad (\text{A.14})$$

Y is the observed signal, X is the predicted signal which is scaled by β . β and the variance of β can be estimated by:

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (\text{A.15})$$

$$var(\hat{\beta}) = \sigma^2 (X^T X)^{-1} \quad (\text{A.16})$$

The application of this method rests on the following assumptions being fulfilled: Error terms are independently and Gaussian distributed with zero mean, the regressors in the matrix are independent of error, non stochastic and known and no regressor is a linear transformation of another regressor. [31]