AALBORG UNIVERSITY
STUDENT REPORT

Evaluation of Basic Preprocessing, aCompCor and FSL FIX for Optimal fMRI Preprocessing in an Individual Differences Study Perspective

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

Project group: 19gr9411

Christian Korfitz Mortensen, Martin Alexander Garenfeld

ii

Preface

Morbi luctus, wisi viverra faucibus pretium, nibh est placerat odio, nec commodo wisi enim eget quam. Quisque libero justo, consectetuer 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 consectetuer. 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

1 | Introduction

To experience and live with pain can be immensely debilitating for any individual, and the consequences of experiencing 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 pain, manifesting an extraordinary amount of human suffering along with additional large economic societal expenditures [3]. Compared to other major health conditions pain outnumber the combined total of heart disease and stroke, diabetes, and cancer combined [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 [4]. The phenomena of pain is a very complex composition of both sensory, reactive and cognitive components making up a self defense mechanism to warn and protect the human organism from any potential or further harm [5, 6]. It still remains incredibly challenging to define and treat pain, as the brain mechanisms during the experience of pain is not yet fully understood [7, 8].
Multiple studies [9, 10, 11] have shown the perception and sensitivity to pain are disposed to a great extent of subjective variability, suggesting that a pivotal key in understating the brain mechanism during pain is found in the individual differences between subjects. Additionally, individual's sensitivity can vary substantially from day to day, despite being opposed to the same stimuli, and the psychophysical rating given by the patient is rather looked upon as artifact of scale, than actually reflecting an unbiased measure of the experienced pain [9]. It can furthermore be difficult for the physician to evaluate the patient's experienced pain from a third-person perspective. To examine if there is an objective measure of pain sensitivity and if it then correlates with the reported experience is therefore of great interest [9].
A potential objective measure of pain sensitivity could be utilizing the method of functional Magnetic Resonance Imaging (fMRI), as it depicts the physiological brain response to noxious stimulus. A prior study [9] demonstrated that individuals reporting different pain sensitivity when exposed to the same stimuli showed correlated brain activation in areas of SI, ACC and prefontal cortex .

Transition to our problem at hand

End up with something presenting our comparison of preprocessing method and this work would contribute to the scientific community.

2 | Problem Analysis

The problem analysis has the purpose of providing a theoretical overview of essential topics in this project such as a deeper investigation of the physiology behind pain, and the difficulties in assessing the brain response through the modality of fMRI. Thereby laying the foundation for understanding the problems presented in the introduction in order to derive methods of solving these. With the main goal of finding the optimal method for denoising functional Magnetic Resonance Images, the background chapter will introduce the following topics: TBC

2.1 Processing of Pain in the Nervous System

Pain is a complex subjective phenomenon, where sensory and reactive components along with cognitive elements are activated to refrain an individual from further and future similar damage. 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 process of transmission of a damaging event related to the brain is called nociception. It is mediated by certain receptors called nociceptors attached to myelinated Aδ and unmyelinated C fibers that terminate at the dorsal horn of the spine. Perception of pain occurs when stimulation of nociceptors is high enough to activate Aδ 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. [6]

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 spine. 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. [12] These structures are what process somatosensory input and is outputting neural activation that affects nociception and pain perception. [6] 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 spine 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. [12] 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 The Hemodynamic Response

As specified in the prior section, certain brain regions are activated when the body is subjected to a noxious stimuli. To further understand the regions activated 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 underlaying physiologic response of a brain region activation is therefore of great importance to understand.

The activation of a brain region starts with a neurological input containing information about the noxius stimuli [12]. 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 yes 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 se found in figure 2.1. [13, 14]

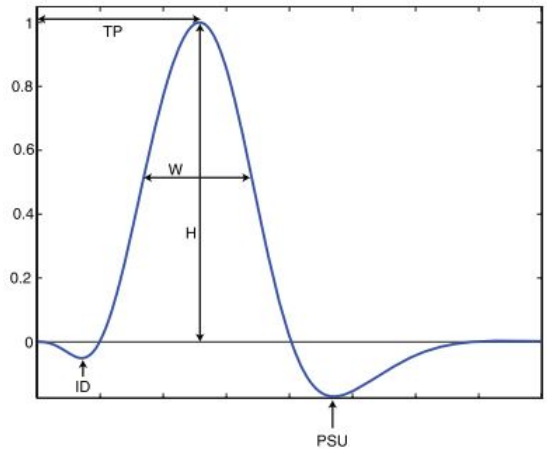


Figure 2.1: A depiction of a single hemodynamic response curve. ID is the initial dip as less oxygen will be present as the metabolic demand increases, TP is time from stimulus until peak, W and H is the width and height of the response and PSU is a post stimulus undershoot. [14]

Figure 2.1 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. [14]

2.3 Functional Magnetic Resonance Imaging

As established the BOLD contrast is effected by the neural activity producing changes in the local blood flow, blood volume and blood oxygenation. These metabolic changes associated with different neurological tasks can be accentuated through the use of fMRI [13]. 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 to why fMRI can 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 MRI signal. [15] Changes in susceptibility arise with the hemodynamic response as oxygenated hemoglobin (HbO_2) is diamagnetic, and deoxygenated 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 metabolic

utilization of oxygen resulting in high (HbO_2) to (Hb) ratio, the BOLD contrast. [13, 14, 16] Figure 2.2 illustrates how the BOLD contrast is dependent on the amount of oxygenated hemoglobin.

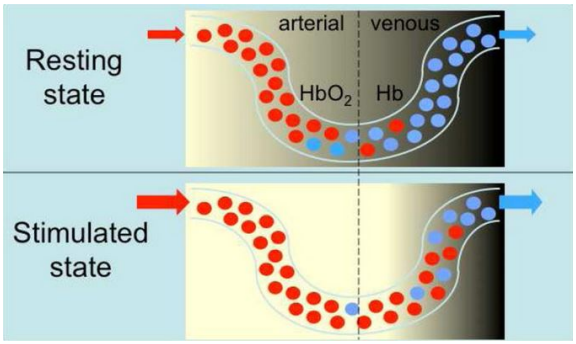


Figure 2.2: Illustration of how the difference in oxygen concentration in the hemoglobin change the magnetic properties, resulting in a higher measurable contrast [13].

The changes in the BOLD contrast can be recorded by using a $T2^*$ sequence, which is sensitive in detecting changes on the magnetic field [16, 17]. The areas highly filled with oxygenated blood will result in bigger signal in the $T2^*$ wighted sequence making these are 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. [16]
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. [17]

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 [7, 9]. 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. [8]
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. [18]
In 1999 and 2003 Coghill et al. [19, 9] 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. [18] 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. [20] 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. [9] with data from a 3.0 tesla scanner would be relevant. Further reasoning for a validation study is that clinics tend to get higher tesla scanners installed, when desiring to get finer images in fMRI [20]. Thus, the future of fMRI lies within the use of higher tesla scanners. Furthermore, to raise the reproducibility in a validation study the number of participants included should be raised to examine if statistically significant result can be found in a higher sample size [18, 21].
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 [20]. In that relation, examining different types of preprocessing methods to denoise the images would be of high interest.

2.5 State of art MRI preprocessing

In section **??** the various types of corrupting noise was presented. Continuous efforts have been made to avoid, remove and limit the influence of artifactual contributors before the quantitative fMRI analysis. Regular preprocessing methods of motion correction, spatial smoothing and temporal filtering, have for long been a part of the preprocessing toolbox [14]. But as studies grow more complex and seeking greater SNR along with the perplexity of greater noise sources associated with newer scanners, the need for more complex and fitting cleanup tools to add to the preprocessing toolbox is present. [20, 22]

In noise removal of fMRI datasets, there are mainly two widely used approaches: one data driven and one model based [23, 24]. In the latter, acquired data is compared to a predefined model, such as in the general linear model. These methods include additional physiological recordings, e.g. heart rate and breathing cycle, and use these regressors of non interest to regress/filter out their contribution to the recorded fMRI dataset. [23, 24, 25] Data driven models instead draw upon the data within the dataset utilizing methods of principal component analysis (PCA) and independent component analysis (ICA). Data driven methods are ideal when no

good existing model fits the data as they can analyze the data in more flexible way. Hence, they are able to identify new and unexpected noise components, which then can be made to fit the model, thereby laying the foundation for a model based noise removal. [24]

One of the most often referred to articles making use of a model based method is the RETROspective Image CORrection (RETROICOR), introduced by Glover et al. [26]. 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. [26]

A limitation of the model based RETROICOR was its dependency on physiologic data being acquired from external devices. A study by Behzadi et al. [27] surpassed this restraining need 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 were no neurological activity is presumed, such as in white matter and cerebrospinal fluid (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 a general linear model as nuance 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 were 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 the factors, limiting its use. [27]

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. [28]. Since then, multiple studies [29, 30, 31, 32, 33] have used and developed the use of ICA further. ICA has proven a powerful tool in separating the different sources which summate to be the complete fMRI scan. By maximizing spatial independence, the data can be decomposed into components each consisting of an activation map and its corresponding time course. [23] This facilitated the prospect of by visually inspecting the components, a discrimination between artifactual activation and task-related could be made. Tools for easy employment of the ICA has been presented by groups like the Oxford University Centre for Functional MRI of the Brain (FMRIB), who presented a program called MELODIC for fast ICA implementation. [34] 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 [23, 35], have proposed ways to recognize different artifactual components, but these can be hard to utilize since every component will be different from another. Subsequently, automatic labeling algorithms have been made to increase the sensibility, reliability and reproducibility of fMRI studies. These have been made by training classifiers in recognizing and separating artifactual components. [33]

In 2014 Salimi-Khorshidi et al. [23] introduced the FMRIB's ICA-based X-noiseifier (FIX), for automatic denoising of resting state and task-related fMRI data. This methods employ the use an initial standard processing followed by a spatial-temporal fastICA, hand labeling of training set, feature extraction of 180 features and a heracial classifier for classification. [23]
The use of pattern recognition is an interesting proposal, as it makes way for a standardized approach, and reduces the labor intense work of labeling all components.

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] Harold Merskey and Nikolai Bogduk. “Classification of Cronic Pain”. In: *International Associasion for the Study of Pain* 2 (1994).

[5] Peter Brook, Tony Pickering, and Jayne Connell. *Oxford Handbook of Pain Management*. 1st. Oxford University Press, 2011, p. 463.

[6] 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.

[7] Christopher S. Nielsen et al. “Individual differences in pain sensitivity: Genetic and environmental contributions”. In: *Pain* 136.1-2 (2008), pp. 21–29.

[8] Robert C. Coghill. “Individual Differences in the Subjective Experience of Pain: New Insights into Mechanisms and Models”. In: 50.9 (2011), pp. 1531–1535.

[9] Robert C Coghill, John G Mchaffie, 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.

[10] 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.

[11] 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.

[12] Irene Tracey and Patrick W. Mantyh. “The Cerebral Signature for Pain Perception and Its Modulation”. In: *Neuron* 55.3 (2007), pp. 377–391.

[13] Gary H Glover. “Overview of functional magnetic resonance imaging”. In: *Neurosurg Clin N Am* 22.2 (2011), pp. 133–139.

[14] Thomas Poldrack A, Russell; Mumford A, Jeanette; Nichols E. *Handbook of functional MRI data analysis*. 2011.

[15] Mushabbar A. Syed, Subha V. Ramen, and Orlando P. Simonetti. *Basic Principles of Cardiovascular MRI*. 2015, pp. 1–338.

[16] Nishanth Khanna et al. "Functional neuroimaging: fundamental principles and clinical applications". In: *Neuroradiology Journal* 28.2 (2015), pp. 87–96.

[17] 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.

[18] J Dubois and R Adolphs. "Building a science of individual differences from fMRI". In: *Trends in Cognitive Sciences* (2016), pp. 425–443.

[19] 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.

[20] 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.

[21] 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.

[22] 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.

[23] 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.

[24] 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.

[25] 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.

[26] 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.

[27] Yashar Behzadi et al. "A Component Based Noise Correction Method (CompCor) for BOLD and Perfusion Based fMRI". In: 6.8 (2013), pp. 90–101.

[28] Martin J Mckeown et al. "Analysis of fMRI Data by Blind Separation Into Independent Spatial Components". In: 188.June 1997 (1998), pp. 160–188.

[29] 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.

[30] 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.

[31] 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.

[32] 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.

[33] Jussi Tohka et al. “Automatic independent component labeling for artifact removal in fMRI”. In: *NeuroImage* 39.3 (2008), pp. 1227–1245.

[34] Oxford Centre for Functional Magnetic Resonance and Imaging of the Brain (FMRIB); *MELODIC/FAQ*. 2016.

[35] Ludovica Griffanti et al. “Hand classification of fMRI ICA noise components”. In: *NeuroImage* 154.June 2016 (2017), pp. 188–205.

[36] A. A. Bharath. *Introductory Medical Imaging*. Vol. 3. 1. 2008, pp. 1–186.

[37] John C Edwards. “Principles of NMR”. In: ().

[38] 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.

[39] 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.

[40] 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.

[41] A Hyvärinen, Juha Karhunen, and Erkki Oja. “Independent Component Analysis”. In: (2001).

[42] 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.

[43] John L Semmlow. *Biosignal and Biomedical Image Processing*. 2004.

[44] CIS 520. *PCA*. 2018.

[45] Y Murali Mohan Babu. “PCA based image denoising”. In: *Signal & Image Processing : An International Journal* 3.2 (2012), pp. 236–244.

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

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. [36] The rate of precession can be calculated by the Lamour frequency:

$$f = y * B_0 \tag{A.1}$$

f is precession frequency, y 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. [36] 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.

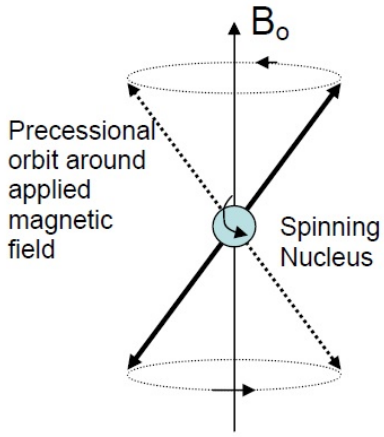


Figure A.1: The figure illustrates how the nucleus precess and spin in relation to the applied magnetic field B_0 surrounding it. The vectors, indicating the precessing, can go opposite or along the magnetic field depending on the nucleus energy state. [37]

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.

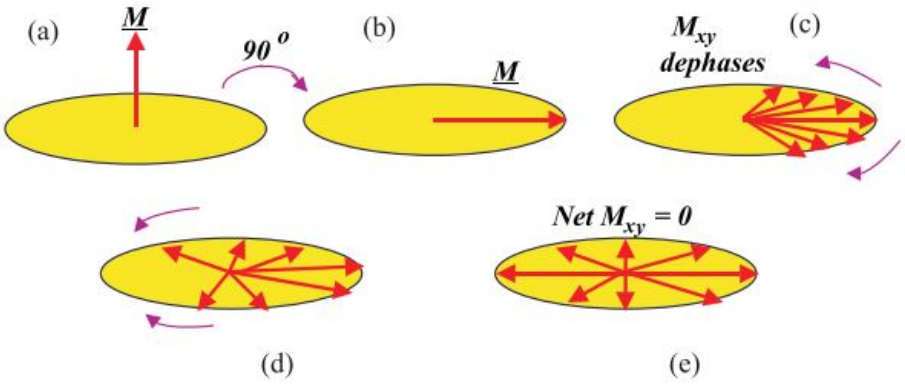


Figure A.2: An illustration of the T_2 relaxation. Before the RF pulse is emitted the majority of the nuclei are precessing out of phase in the low energy state and the net magnetization is pointing in the direction of the longitudinal magnetization, (a). The nuclei absorb the RF pulse, the magnetization 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. [36]

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. [36] 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. [36]
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 inhomogenities the nuclei dephase more rapidly. Applying a second RF pulse of 180°after the nuclei have dephased, the nuclei will rephase. Again the constrast in the image is expressed as a result of the relaxation time of the nuclei, which is depends on the surrounding environment. . [36]
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. [36]

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. [38]

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 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. [36]

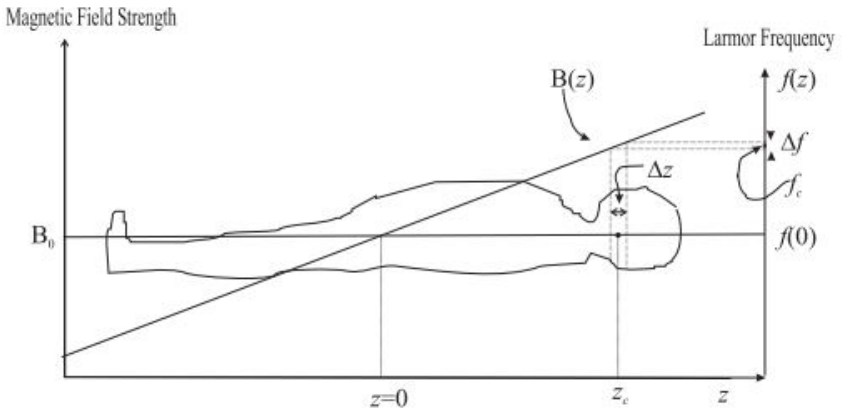


Figure A.3: The position of the slice is specified through the direction of the frequency encoding gradient ($B(z)$) and though the central frequency of the emitted RF pulse (f_c). The thickness of the slice is dependent on the steepness slope of $B(z)$, and on the bandwith of the emitted RF pulse (δf). [36]

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. [36]
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. [36] Figure A.4 shows the acquired signal in k-space and the reconstructed image after the Fourier Transform.

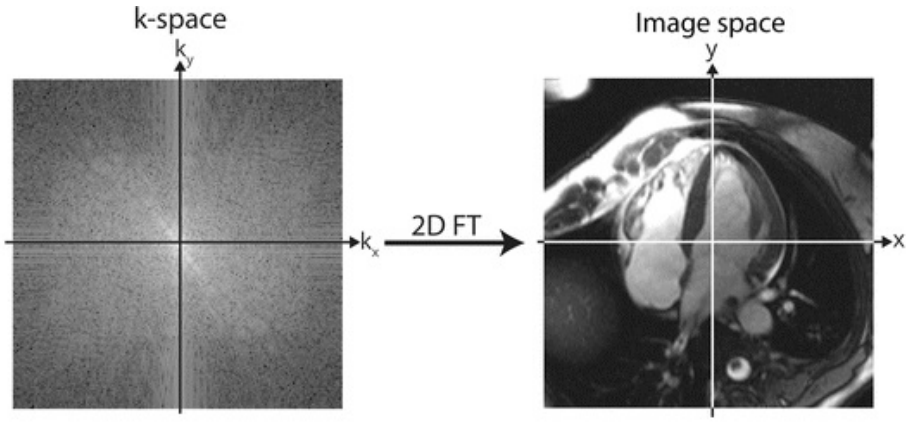


Figure A.4: A depiction of the acquired signal represented in k-space and the resulting reconstructed image after the inverse 2D Fourier Transform [15].

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, adding to the knowledge gained in section **??**.

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. [39]

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. [40]
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 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. [40]

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. [39]

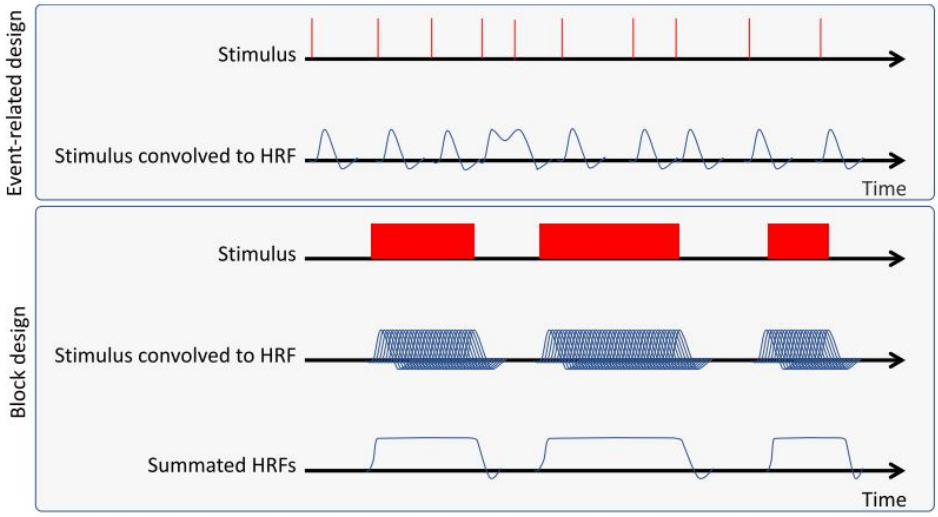


Figure A.5: The top image show the event-related stimuli design and the stimulus convolved with the hemodynamic response function. The lower image depicts a block design, the stimulus convolved with hemodynamic response function and the summated HRF response. [39]

A.4 General MRI Preprocessing

The following section will present some preprocessing methods, which generally are applied to remove noise and artifacts. For each method the artifacts/noise problem will be described followed by how the method solves this problem. The section will thereby present some of the common issues to be aware of in fMRI acquisition and what methods to implement in limiting artifact/noise influence.

When MRI data is acquired there are several steps, which need to be taken, before the multi-dimensional images are ready to undergo statistical analysis. These steps involves correction methods, which are often referred to as preprocessing. [39] There are multiple steps in preprocessing fMR images depending on the apparent application and outcome intended. However, there is a standard set of methods that is usually used across all applications. An example of some of the general processing steps for MR imaging can be seen in figure A.6. [14]

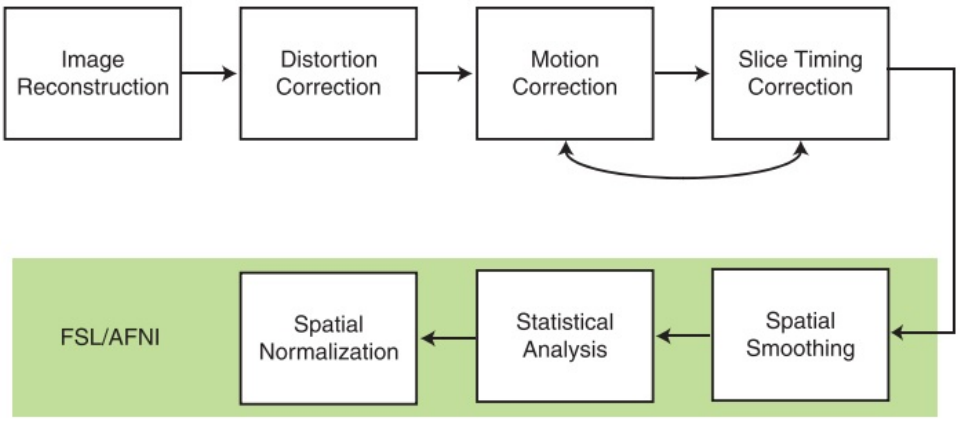


Figure A.6: The general pipeline for MRI processing done in either FSL or AFNI, showing the different processing steps considered before final statistical analysis. Modified from [14].

A.4.1 Quality control

Conducting a continuous quality control is highly recommended after each performed correction step. Various scanner artifacts can occur while acquiring an MRI series. Before performing any common correction steps, one should consider to look for spike or ghosting artifacts. Spike artifacts are seen as a regular pattern of change in brightness across the entire image. This problem can occur due to instability inside the scanner deriving from e.g. electrical discharges. The artifact called ghosting occur mainly due to two reasons. One being an offset in phase between different lines in K-space and the other due to periodic motion as in heartbeat and respiration. Ghosting can be seen as light copies of the object appearing to either side of the object. Both types of artifacts can corrupt the information contained in the images. However, artifacts of this kind rarely present themselves in newer scanners, nevertheless it is still recommended to perform a quality control of the raw input from the scan as well as the result of the correction steps. [14]

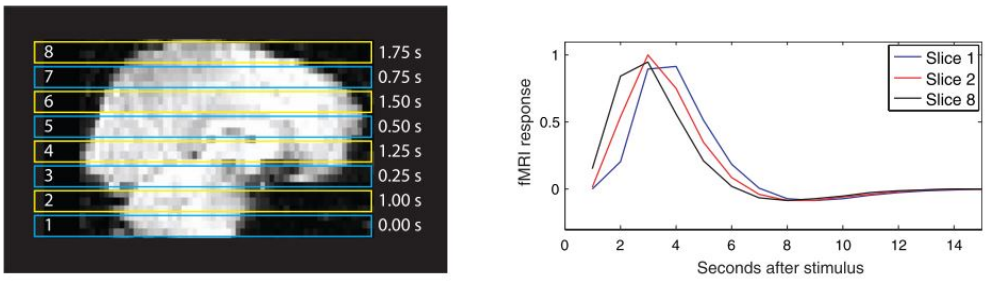
A.4.2 Distortion correction

Some fMRI acquisition methods suffers from artifacts at regions where air and tissue meet. The ear canal and sinuses are areas especially vulnerable. Inhomogeneity in the magnetic field in these areas can cause two types of artifacts: dropout and geometric distortion. A dropout will result in a reduced signal intensity in regions close to the air to tissue transition. When a dropout during an acquisition occurs, the lost signal cannot be restored and the damage is permanent. Therefore it is wise to consider the appropriate acquisition method taking the area of interest into addition. Air to tissue passages can also be subject to spatial distortion due to inhomogeneity created in the magnetic field. This will lead to structures not being located correctly in the captured image. This distortion makes is difficult to align two different scans, as done when aligning fMRI images with structural images. The spatial distortion can partially be corrected by employing field maps. In order to do a field map, the pulse sequence from the scan needs

to be known. The process involves acquiring images at two different echo times. This results in images with two different phases, which can be used to compute the field inhomogeneity. Thereby it becomes possible to calculate the relative distance each voxel has shifted. This makes up a map for the distance shift for each voxel, and by inverting the map the original image can be restored. [14]

A.4.3 Slice timing correction

Acquiring fMRI scans is nearly always done in two-dimensions, where the slices are taken one by one. This can either be in an ascending, descending or interleaved order. Interleaved order [1] is sequentially skipping every either odd or even slice and then afterwards acquiring the skipped slices. Regardless of which order the slices are acquired, a difference in effect in each slice to the same hemodynamic response will be present due to the time difference in the slices. The method and result of interleaved MRI acquisition order can be seen figure A.7. The difference in time between slices can range up to a couple of seconds depending on the acquisition protocol.



(a) (b)

Figure A.7: Figure (a) illustrates an example of an MRI acquisition using interleaved order, where initially every odd slice is acquired followed by every even. The information about the hemodynamic response in each slice, and thereby also the difference at each time-point is shown in figure b. Figure modified from [14].

The difference in slice timing constitutes a problem when analyzing the data. The data is formed into statistical model, but since this model assumes that all slices are acquired at the same time point, the actual signal and the statistical model creates a mismatch. To counter this problem slice timing correction has been introduced. The common approach of this method is to choose a reference slice. Usually the slice acquired at T/2, where T is the total scan time, is used to interpolate the others. Linear interpolation can be used for simplicity, but most often sinc interpolation is used as it imposes less smoothing to the signal. [14]

A.4.4 Motion correction

Correcting for motion artifacts when doing fMRI is inevitable, since even the best subjects will not be able to hold still. Even subtle movements as swallowing will be visible in the acquired

[1] FiXme Note: think it is done to avoid crosstalk

image. [14]

Multiple internal and external factors can cause a subject to move. Internal factors are non-avoidable physiologic motion. The heartbeat causes a pulsating movement, which makes the brain move. Additionally, motion created during respiration can cause small changes in the magnetic field around the head. External factors like imposed stimulus might also cause the subject to make sudden movements. Often when doing fMRI the brain activation is measured while the subject is subjected to some kind of stimulus. The stimulus would make the patient move, while 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. [14]

Motion during image acquisition can result in two primary artifact effects: bulk motion and spin history. Bulk motion refers to the movement of the head as a whole and requires standard correction methods, e.g. the images throughout the series to be realigned to a reference image. The effect of bulk motion can be visual in the entire image of the brain, but the effect will be most predominant at the edges of the brain. Here the artifact will be noticeable as either a drop or increase in intensity as a voxel would switch from containing brain tissue to suddenly not. Spin history is head movement interfering with the MRI signal itself. The interference occur during acquisition when a voxel of excited protons is moved in to a neighboring slice. The scanner will thereby receive a different signal than expected which do not correctly represent the actual local properties. This results in an image where the intensities change in a striped pattern. The standard motion correction methods cannot cope with this type of artifact, but Independent Component Analysis (ICA) might give opportunities to correct for this artifact. [14]

As mentioned earlier motion correction is to realign the series of images to a reference image trying to minimize cost in an introduced cost function. The reference image is usually the one taken midway into the series, justified by it 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. The images are thereafter realigned utilizing an image registration method as it register the brain in each image. The general methods for motion correction treats the brain as a rigid objects, thus only performing rigid body transformations. Subsequently, the brain can either translate or rotate along the three axis, but the shape of the head cannot change. This method is therefore only applicable for bulk motion. [14]

A.4.5 Spatial smoothing

In some cases, introducing spatial smoothing in the preprocessing pipeline proves to be beneficial. Spatial smoothing allows the possibility of gaining a higher signal to noise ratio within the image, though with the consequence of a decrease in spatial resolution as the image gets 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. Spatial smoothing removes the higher-frequency information. This might wash out some of the less

significant features in the image, but this is favorable if the signal is increased for the more significant features. Especially when acquiring small voxels spatial smoothing is favorable, as it reduces the overall noise. Smoothing can also be applied to lessen the anatomical variability in images when doing studies with multiple subjects. [14]

Smoothing is done by applying a kernel, also called filter, to the image. The three dimensional image is convoluted with a three dimensional filter. The most commonly used is a Gaussian filter, where the extent of smoothing is controlled by the width of the distribution. The filter works such that for each voxel a new value is calculated based on a weighted average of the neighboring pixels, where the ones closest contribute the most and those further away contribute the least. In statistical terms this can be put as the width of the distribution is described by the full width at half maximum (FWHM). Using standard deviation this would be:

$$FWHM = 2\sigma\sqrt{2ln(2)} \tag{A.2}$$

, where an increase in FWHM, would result in a greater smoothing. The amount of smoothing needed to be implemented highly depends on the application and purpose. When smoothing fMRI signals for noise, the width of the filter distribution should not be bigger than the area of activation signals of interest. The effect of smoothing is shown on figure A.8, where it is seen that as width increases smaller activation areas get removed and bigger areas of activation gets smeared together. Using smoothing to reduce the effect of anatomical variability, the optimal distribution width depends on the amount of variability in the subject population. [14]

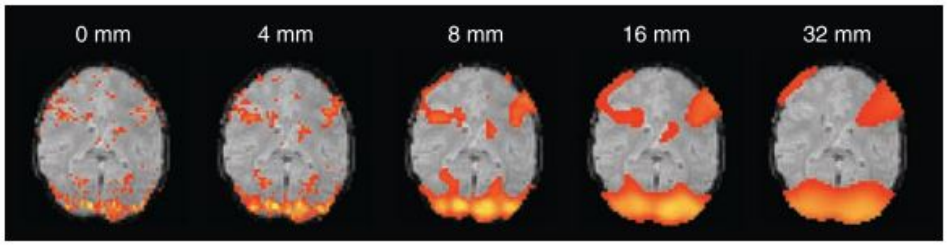


Figure A.8: Illustration showing the impact of using different distribution width on the activated areas. An increase in width results in greater areas of activation smearing together and the removal of smaller. [14]

A.4.6 Temporal filtering

A characteristic noise which represents itself during 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. Doing a Fourier Transform, to analyze the signal in the power spectrum, would reveal low frequency contributers influencing the output. The reason for this type of noise contamination has been heavily investigated, and conclusions state that the noise originates from MRI scanner instability.[2] As this low-frequency drift will always be present, it is very crucial to consider the interval of which tasks or stimuli are performed to

[2] FiXme Note: heating of coils and such i believe

avoid the output being present in the noise range of 0 to 0.015 Hz. Therefore stimuli or tasks should be performed within intervals of 70 s or less. [14]
A two step approach is used to remove the low-frequency contribute from the signal. Firstly a high-pass filter is used to remove the drift trend in the data. Introducing the high-pass filter impose a correlation of the data making the time series correlated over time. This would violate some of the assumptions made in the General Linear Model (GLM), often used for statistical analysis, because the data is not temporally autocorrelated and the variance to be constant over observations, as presented in section A.7. Not attending to this problem might cause an elevation in false positive rate. Thus, the second step is to estimate the autocorrelation and undo the correlation structure of the data. This is typically done by pre-whitening the data. There are multiple ways to implement a high-pass filter ADD SOMETHING ABOUT LOWESS METHOD [14]

A.5 Independent Component Analysis

Independent component analysis has proven to be a very useful tool in separating noise from wanted signal in many application. The methods has been adapted to work in BOLD-signals, making it a very desirable tool in fMRI. [23] The following section will seek to elucidate the general basic theoretical background behind the analysis exposing it possible use in this project.

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 [41]:

$$\mathbf{x} = \mathbf{A}\mathbf{s} \tag{A.3}$$

where $\mathbf{x}$ is a observed signal vector containing the mixed signal elements $x_1, x_2, ..., x_n$, $\mathbf{s}$ is the source signal vector with the elements $s_1, s_2, ..., s_m$ 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 [41]:

$$\mathbf{s} = \mathbf{A}^{-\mathbf{1}}\mathbf{x} \tag{A.4}$$

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 [41]:

- 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 $y_1, y_2, ..., y_n$ 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 [41]:

$$p(y_1, y_2, \ldots, y_n) = p(y_1) \cdot p(y_2) \cdot \ldots \cdot p(y_n) \tag{A.5}$$

Satisfying this condition assures independency of the variables. A graphical example of independence in two dimensions is shown in figure A.9 [41, 42].

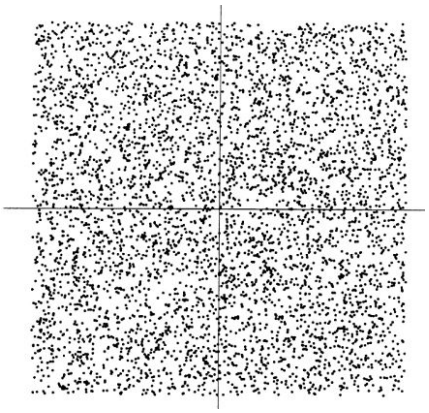


Figure A.9: Figure illustrating the joint distribution of two independent components having uniform distribution. [42].

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 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 $\mathbf{z}$:

$$\mathbf{z} = \mathbf{V}\mathbf{x} = \mathbf{V}\mathbf{A}\mathbf{s} \tag{A.6}$$

A graphical example of whitened data in two dimensions is shown in figure A.10 [41, 42].

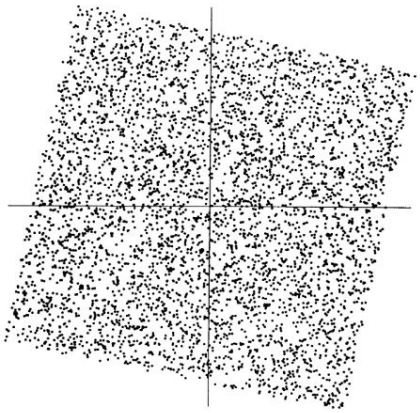


Figure A.10: Figure illustrating the impact of whitening on the joint distribution of two components [42].

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 [41]:

$$\begin{aligned} p(s_1, s_2) &= \frac{1}{2\pi} exp(-\frac{s_1^2 + s_2^2}{2}) \\ &= \frac{1}{2\pi} exp(-\frac{\| \mathbf{s} \|^2}{2}) \end{aligned} \tag{A.7}$$

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

$$p(x_1, x_2) = \frac{1}{2\pi} exp(-\frac{\| \mathbf{A^T x} \|^2}{2}) \mid Det(\mathbf{A^T}) \mid \tag{A.8}$$

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

$$p(x_1, x_2) = \frac{1}{2\pi} exp(-\frac{\| \mathbf{s} \|^2}{2}) \tag{A.9}$$

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. [41]

A.5.1 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.5). 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. [41]

A.6 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 [27]. 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 red dots) describing the highest variance in the data set (black line) as seen on figure A.11. The PC is found by minimizing the mean square distance between the data points. [43]

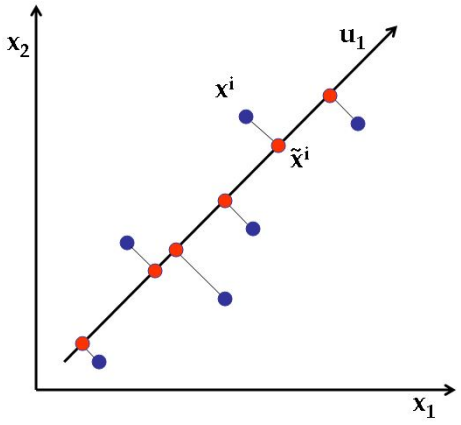


Figure A.11: Two-dimensional example of projection of data variables (blue dots) onto a PC axes (black line). (u_1) indicates the direction of the eigenvector. [44]

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:

$$S = X'X \tag{A.10}$$

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:

$$|S - \lambda I| = 0 \tag{A.11}$$

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:

$$det|S - \lambda I|b_i = 0 \tag{A.12}$$

Where b_i is used to calculate the eigenvectors as in:

$$u_i = \frac{b_i}{\sqrt{b_i' b_i}} \tag{A.13}$$

Where u_i is the i number of eigenvectors that contain a contribution to the principal components. The SVD orders the eigenvalues by size $\lambda_1 > \lambda_2 \ldots > \lambda_i$. 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: $\%\ of\ total\ variance = \frac{\lambda_i}{Trace}$. This can be used for deciding how many components are significant and by how much the dataset can be reduced. [43]

A.6.1 PCA in images

The PCA can also be implemented on images, where the dimensionality reduction principle is mostly used for image compression. Here images can be reconstructed using only very few principle components without much information loss. Similarly PCA can be used to de-noise images, as noise would be presented in some of the less saying components, and by removing these, noise would be removed from the image.
The PCA can be run on the entire image, but methods introducing local PCA in smaller windows of the image, have also been introduced for noise removal. [45]

A.7 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 [46].

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. [39, 25] To explain the implementation of a GLM we can use the example presented by [25]. 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 X, a scaling parameter for each regressor β, and an error term ε. This can be presented in the GLM's basic from as [25]:

$$Y = X\beta + \varepsilon \tag{A.14}$$

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.12.

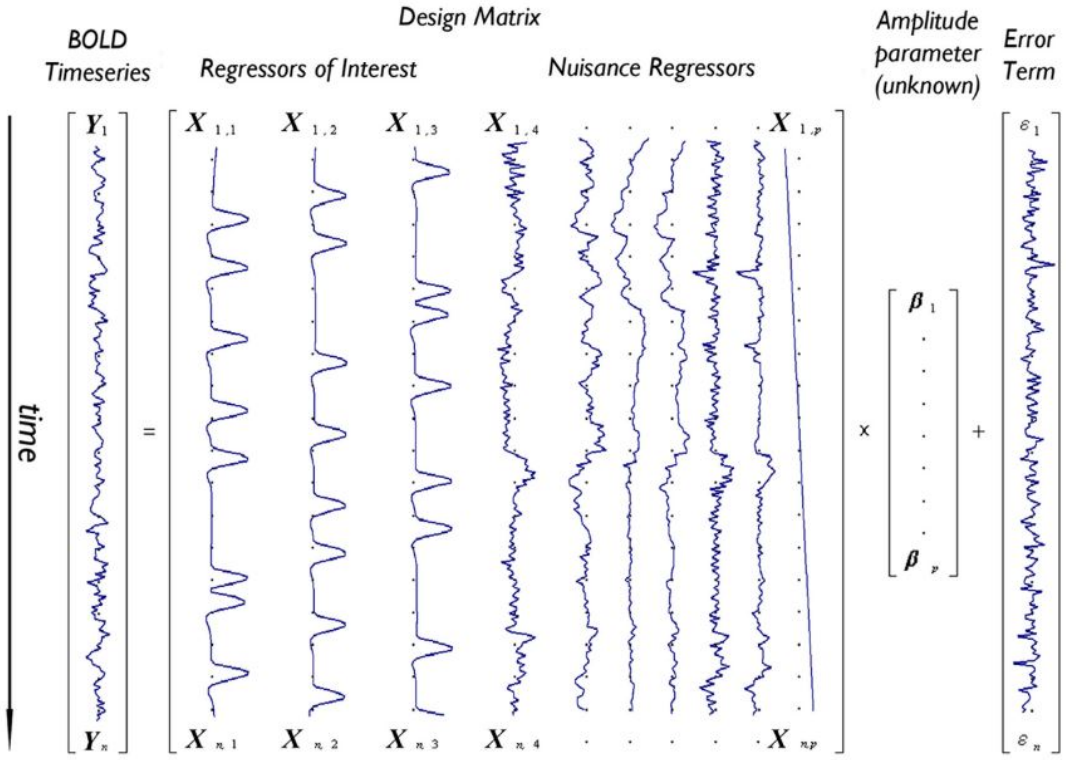


Figure A.12: Illustration of a design matrix for a given GLM model, describing the BOLD time series Y_n, the regressors of interest $X_{1,1}, X_{1,2}, X_{1,3}$, the nuisance regressors $X_{4..p}$, the amplitude parameter $\beta_{1..p}$ and the error terms $\varepsilon_{1..n}$. [25]

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. [39, 25]

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 [25]:

$$\sum_{i=1}^{n}(Y_i - X_i \times \hat{\beta})^2 \tag{A.15}$$

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:

$$\hat{\beta} = (X^T X)^{-1} X^T Y \tag{A.16}$$

$$var(\hat{\beta}) = \sigma^2 (X^T X)^{-1} \tag{A.17}$$

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. [25]