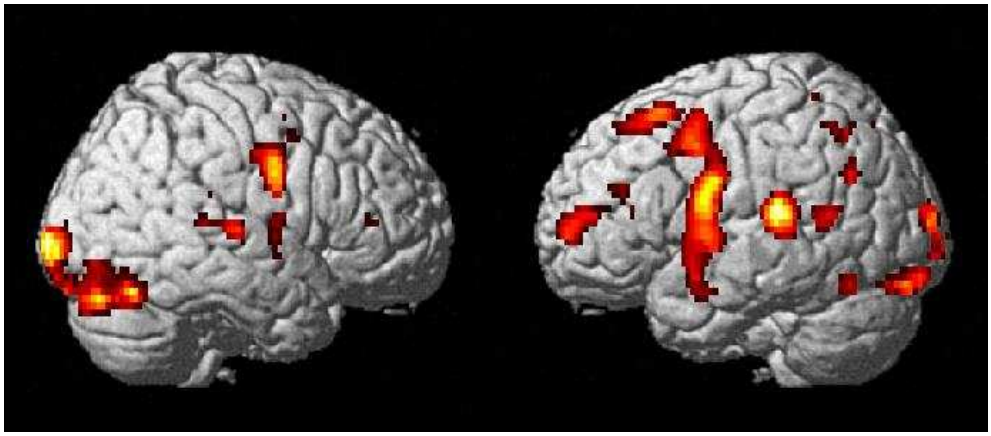
# The mathematical background of SPM99

*Eini Niskanen*



Master's thesis
28th April 2004
University of Kuopio
Department of Applied Physics

UNIVERSITY OF KUOPIO, Faculty of Natural Sciences and Environmental Sciences
Department of Applied Physics
Medical Physics

NISKANEN EINI IRENE: The Mathematical background of SPM99

Master's Thesis, 63 pages
Supervisors: Docent Pasi Karjalainen
Department of Applied Physics
University of Kuopio

MSc Mervi Könönen
Department of Clinical Neurophysiology, Department of Clinical Radiology
Kuopio University Hospital

---

Magnetic Resonance Imaging (MRI) is a powerful tool to investigate human anatomy, physiology and function. Being noninvasive and non-ionizing method MRI is also safe for the patient. Using MRI it is possible to obtain accurate anatomical images of human brain and also examine the function of brain. This examination of function is called functional MRI (fMRI). One method to detect the functional areas in brain is based on the so called blood oxygenation level dependent (BOLD) method. The stimuli causes a small but detectable change in MR signal intensity in active cortical areas. The increase in signal intensity is only few per cents and, therefore, it is usually not possible to detect the shange by naked eye just by examining the MR images. To reveal the active cortical areas statistical and mathematical methods are needed.

Numerous computational programs have been developed for analyzing fMRI data and revealing the secrets of the brain. One of these programs is SPM99 (Statistical Parametric Mapping) developed by the Wellcome institute in London, UK. The program is a suite of MATLAB® (Mathworks Inc.) functions and subroutines (with some externally compiled C routines) enabling the user to modify it, and also ensure the that the program works as it should work. The mathematical methods that are used in SPM99 are published in several articles but there is no comprehensive monograph.

In this thesis the mathematical background of SPM99 is investigated. SPM99 utilizes the General Linear Model (GLM) and Least Squares (LS) -method and some statistical tests e.g. t-test, F-test and Analysis of variance (Anova) in it's computation. The fMRI data is firstly preprocessed in order to reduce several artifacts (e.g. motion), and alter the data to better fulfil the assumptions of the statistics. After the preprocessing, a combination of basis functions, depending on the paradigm, is fitted into the time-series of each voxel. The better the fit, the more likely that voxel was active during the paradigm. Statistical tests are then needed to decide whether the fit of the basis functions was truly a consequence of brain activity but not just noise.

The mathematical background of SPM99 is studied by simulation, where artificial activation is located in real fMRI-data and the analysis is carried out using both SPM99 and MATLAB®. The main purpose of this simulation is to verify the theory of SPM99 explained in this thesis.

# Abbreviations

| | |
|---|---|
| BOLD | Blood oxygenation level dependent |
| CBF | Cerebral bloof flow |
| CBV | Cerebral blood volume |
| $CMRO_2$ | Cerebral metabolic rate of oxygen |
| CSF | Cerebro-spinal fluid |
| DCT | Discrete cosine transform |
| EC | Euler characteristic |
| ECC | Entropy correlation coefficient |
| EEG | Electroencephalography |
| EPI | Echo planar imaging, echo planar image |
| ERP | Event-related potential |
| FDR | False discovery rate |
| FID | Free induction decay |
| fMRI | Functional magnetic resonance imaging |
| FWE | Family wise error |
| FWHM | Full width of half maximum |
| GLM | General linear model |
| GM | Grey matter |
| GRF | Gaussian random field |
| hrf | hemodynamic response function |
| IRF | Impulse response function |
| LS | Least squares |
| MAP | Maximum *a posteriori* |
| MEG | Magnetoencephalography |
| MNI | Montreal neurological institute |
| MR | Magnetic resonance |
| MRI | Magnetic resonance imaging |
| MRS | Magnetic resonance spectroscopy |
| NMR | Nuclear magnetic resonance |
| PET | Positron emission tomography |
| RF | Radio frequency |
| SNR | Signal to noise ratio |
| SPET | Single photon emission tomography |
| SPM | Statistical parametric mapping |
| TE | Time to echo |
| TR | Time to repeat |
| WM | White matter |

# Contents

Chapter 1

# Introduction

Direct cortical stimulation studies in the middle of the $20^{th}$ century as well as some studies of head injuries have indicated that the different brain functions can be localized to certain cortical areas. These functional areas have been of interest since they were detected. During the second half of the $20^{th}$ century many non-invasive methods to study brain functions were invented and developed. One of these methods was called functional magnetic resonance imaging (fMRI) [59].

FMRI is a special technique of the normal magnetic resonance imaging (MRI). With MRI it is possible to differentiate soft tissues from each other. This ability is essential in brain imaging. MRI is used mostly for obtaining tomographic anatomical images of human body whereas fMRI is used to reveal the function of the brain and different cortical areas.

Nowadays, the use of fMRI has spread all over the world. There are many reasons for its popularity. First, the method does not use any ionizing radiation. It is thus safe imaging method and enables repeated studies with adults and children. Secondly, the required equipment is already widely available. MRI equipment used for normal anatomical imaging can be used to fMRI studies. Third reason is that the spatial and temporal resolution of the fMRI data is better than that of other systems based on hemodynamics e.g. positron emission tomography (PET).

FMRI signal is based on the local blood flow and blood oxygenation changes as response to neural activity. These hemodynamic changes are detected by using MRI technology. In MR imaging the contrast between different tissues can be enhanced by using contrast agents, that normally are paramagnetic substances. Blood deoxyhemoglobin (i.e. hemoglobin with no oxygen) is paramagnetic and can be considered as contrast agent in fMRI. Deoxyhemoglobin causes inhomogeneities in the local magnetic field which dephazes the spins thus decreasing the signal intensity in $T_2^*$-weighted images. In active cortical areas there is increase in blood flow and, therefore, also in oxygenated hemoglobin. However, the relative oxygen consumption is decreased which leads to a relative decrease in the amount of deoxyhemoglobin. This leads to a small, only few per cent, increase in the signal intensity. The increase is so small that it is not possible to detect with naked eye without any computational methods for data analysis.

For the data analysis many methods have been developed and there are many software packages available for that purpose. In this thesis one of these softwares is studied. This software is SPM99. SPM99 is developed by the Wellcome Department of Cognitive Neurology, London, UK[1] to analyze data achieved by fMRI, positron emission tomography (PET), or single photon emission computed tomography (SPECT). SPM99 software is a suite of MATLAB®[2] functions and subroutines with some externally compiled C routines. The abbreviation SPM comes from words *statistical parametric mapping*, which refers to the statistical processes used to test hypothesis about neuroimaging data measured to reveal the active cortical areas.

The analysis of SPM99 can roughly be divided in two parts: preprocessing and statistical analysis. In preprocessing the data is processed to better correspond the assumptions of the statistical tests. Some major artifacts are also removed from the data before analysis. First step in

[1] Homepage of SPM-software: `http://www.fil.ion.ucl.ac.uk/spm/`

[2] Homepage of Mathworks: `http://www.mathworks.com`

preprocessing is usually realigning, which removes the motion artifacts. The differences in acquiring time of slices in one whole head is corrected with slice timing -step. Then it is possible to normalize the data to reduce the anatomical differences between different subjects and to ease the localization of different functional cortical areas. Before normalization it is possible to coregister the anatomical and functional images so that the brain areas in both of these modalities would match. Finally the data is spatially smoothed in order to get the data match better the assumptions of the estimation and statistics.

The analysis of SPM99 is calculated in voxel by voxel bases and is based on the general linear model (GLM) and Least Squares (LS) estimation. The stimulations of the study determine an ideal time series of a voxel in active cortical area. This ideal time series is then fitted into the actual voxel time series by using LS-estimation as well as possible. The quality of the fitting determines if a particular voxel has been active during the stimulations or not. The quality of the fitting is evaluated with statistical tests e.g. t-test, F-test, Anova or AnCova.

In the Chapter 2 of this thesis the theory of magnetic resonance imaging is studied. The theory is presented in a way that combines some quantum mechanics with the classical mechanics. There is also presented the basic terms of the imaging as well as some imaging sequences. The Chapter 3 is about fMRI. The phenomenon behind the increase of the signal intensity is explained as well as the basics of the data analysis. Also some most severe artifacts and some methods how to prevent them are introduced. In the Chapter 4 the theory of LS-estimation and the equations related to it as well as statistical testing (t-test) are derived. The Chapter 5 establishes the mathematics behind the preprocessing steps including realigning, slice timing, coregistration, normalization and spatial smoothing. The statistical analysis and related equations are derived in the Chapter 6. These equations are verified in the Chapter 7 with simulations. In these simulations actual fMRI data with artificial activations is used. The data is then analyzed using SPM99 and by using the equations of Chapter 6. The results of these analysis are compared in the final chapter, Chapter 7.

Chapter 2

# Magnetic Resonance Imaging (MRI)

Nuclear magnetic resonance (NMR) was invented in 1946 by two independent groups (Bloch *et al.* and Purcell *et al.*) concurrently [27] and has since been used in physics and in chemistry to study construction and movement of molecules. Over the years the research of NMR has split into two branches: magnetic resonance imaging (MRI) and magnetic resonance spectroscopy (MRS). MRS is used for studying metabolism and MRI for obtaining images to study anatomy or function *in vivo*.

The most important nucleus for MRI is hydrogen nucleus $^1$H, that is, a proton. If a static external magnetic field is applied to a sample containing protons, their interactions with the external field cause the precession of the proton along the direction of the field. MRI is based on the manipulation of the precession frequency of protons in water, fat and other tissues [32].

The basic idea of imaging is to correlate a series of signal measurements with their spatial locations. If a sample is placed in a uniform magnetic field, then most of the spins have the same precession frequency. By spatially changing the magnetic field across the sample, a signal with spatially varying frequency components is produced. Using Fourier-analysis this signal can be converted to a representational image of the object.

Nowadays, MRI is a widely used method of imaging especially soft tissues of human body. One reason for the popularity of MRI is that the technology is completely non-invasive and does not necessitate use of ionizing radiation. MRI has also high spatial resolution and contrast even between different soft tissues which is essential e.g. for brain research.

## 2.1 The NMR signal

Atomic nuclei consisting of protons and neutrons have intrinsic angular momentum $\vec{S}$. This can only have certain discrete values determined by a quantum number $s$. The magnitude $S$ of angular momentum $\vec{S}$ is given by

$$S = \hbar\sqrt{s(s+1)}, \tag{2.1}$$

where constant $\hbar$ is equal to $h/2\pi$ where $h$ is the Planck's constant. Every elementary particle has a specific and constant value of $s$ which is called the spin of that particular particle. Protons and neutrons have spin $\frac{1}{2}$ [30]. For nuclei the quantum number $s$ may only have integer or half-integer values as follows [27]:

1. $s$ is integer for nuclei with even mass number
2. $s$ is zero for nuclei with even numbers of both neutrons and protons
3. $s$ is half-integer for nuclei with odd mass numbers

The angular momentum vector can only have discrete orientations with respect to any given direction i.e. it is quantized. The $z$-component of angular momentum along the $z$-axis is given by [33]

$$S_z = m_s\hbar, \tag{2.2}$$

where the $m_s$ is a spin quantum number and it can have only certain values defined by the quantum number $s$

$$m_s = -s,\ -s+1, \ldots, s-1,\ s. \tag{2.3}$$

Hydrogen nucleus $^1$H has spin $\frac{1}{2}$ and so from eq. (2.2) its $z$-component of angular momentum is $S_z = \pm\frac{1}{2}\hbar$.

A spinning charged particle constitutes a magnetic dipole. The magnetic dipole moment $\vec{\mu}$ is related to the intrinsic magnetic momentum $\vec{S}$

$$\vec{\mu} = \gamma\vec{S}, \tag{2.4}$$

where $\gamma$ is a proportionality constant called the gyromagnetic ratio of the nucleus. The $z$-component of the magnetic dipole moment is

$$\mu_z = \gamma S_z = \gamma m_s \hbar. \tag{2.5}$$

The magnitude of the gyromagnetic ratio $\gamma$ is specific to different nuclei and it cannot be predicted from classical physics. For most nuclei $\gamma$ is positive, but for some, it is negative [27].

If a particle is placed in an external magnetic field $\vec{B}_0$, it acquires energy $E$

$$E = -\vec{\mu} \cdot \vec{B}_0. \tag{2.6}$$

In $z$-direction this is

$$E_z = -\mu_z B_0 = \gamma m_s \hbar B_0. \tag{2.7}$$

This energy is quantized and for a nucleus having a spin $\frac{1}{2}$, it can have two values. Hence, the nucleus can have two different energy states. The energy difference $\Delta E$ between these states is

$$\Delta E = \gamma \hbar B_0. \tag{2.8}$$

The NMR signal is based on energy transitions between the different energy states. By using the Bohr's equation of electro-magnetic radiation of frequency $\nu$

$$\Delta E = h\nu \tag{2.9}$$

and equation (2.8) we find that

$$h\nu = \gamma \hbar B_0. \tag{2.10}$$

By substituing the $\nu$ with angular frequency $\omega$ we get

$$\omega = \gamma B_0. \tag{2.11}$$

The frequency $\omega$ is also called the Larmor frequency and radiation with this frequency can cause transitions between the different energy states.

By applying an additional oscillating magnetic field $B_1$ perpendicular to field $B_0$, it is possible to cause transitions between these energy states, if the frequency of field $B_1$ is the same as nucleus' Larmor frequency given in eq. (2.11). This oscillation enables spins on lower state to be transfered onto the higher state by emitting the needed energy $\Delta E$. Concurrently the nuclei start to precess around the orientation of the field $B_1$. The Larmor frequency of $^1$H is in radio frequency range, so the additional magnetic field $B_1$ is often called radio frequency pulse or RF-pulse. After the RF-pulse is switched off, the spins slowly recover to the lower energy state and return to precess around the orientation of the magnetic field $B_0$. The receiver coil will have current generated within during this recovery of the spins because of the changing magnetic field. The current is the raw signal detected in an MRI scanner [59].

## 2.2 Relaxation times

When a sample containing nuclei is placed into an external magnetic field the sample becomes magnetized, the net component of magnetization being along the direction of the external field. Usually the frame of reference is defined in a way that this direction is also the z-direction. The direction of the net magnetization can be disturbed by applying an RF-pulse. RF-pulse that is used to tilt the net magnetization into the plane perpendicular the main field $B_0$ (into the $xy$-plane) is called 90° RF-pulse. After the RF-pulse is switched off, the nuclei's net magnetization slowly recovers into its initial orientation along the direction of field $B_0$. These processes define the relaxation times $T_1$ and $T_2$.

### 2.2.1 Spin-lattice relaxation, $T_1$

Spin-lattice relaxation defines the recovery of the magnetization in the $z$-direction $M_z$ after the 90° RF-pulse is switched off. The energy the nuclei absorbed from the RF-pulse is consumed by interactions between the nuclei and their molecular environment. The return of the magnetization $M_z$ to its equilibrium magnetization $M_0$ with a time constant $T_1$ is expressed as

$$\frac{dM_z}{dt} = \frac{M_0 - M_z}{T_1}. \tag{2.12}$$

The time constant $T_1$ is also called the spin lattice relaxation time and it describes how $M_z$ returns to its equilibrium. By solving the differential equation (2.12) we get the equation for the magnetization in $z$-direction as a function of time

$$M_z(t) = M_0 \left(1 - e^{-t/T_1}\right). \tag{2.13}$$

From this equation it can be seen that after time $T_1$ the magnetization in the $z$-direction has recovered 63% of equilibrium ( $M_z(T_1) = M_0 \left(1 - e^{-1}\right) = 0.63 M_0$ ). This is illustrated in Fig. 2.1 (a).

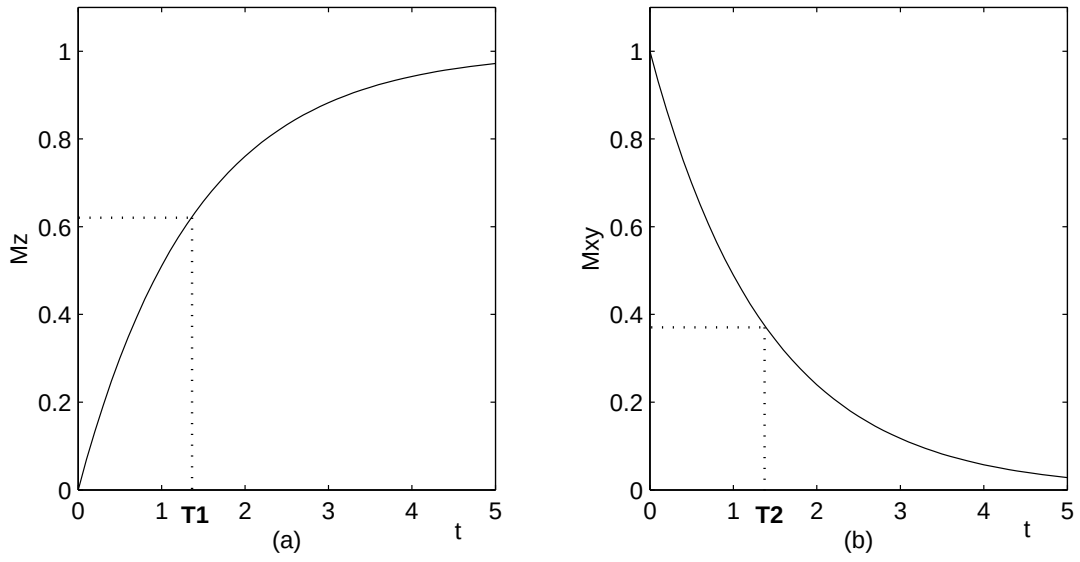



Figure 2.1: The magnetizations $M_z$ in (a) and $M_{xy}$ in (b) are illustrated as a functions of time $t$ (s) as are given in eq. (2.13) and (2.15).

Nuclei in different tissues have different molecular environment. Tissue containing mostly water have long $T_1$ because the energy transfer between the nuclei and the lattice is slow due to small

water molecules in the lattice. On the other hand, tissues containing medium size molecules have short $T_1$. This is because the medium size molecules have fluctuating magnetic fields near the nuclei's Larmor frequency. Hence, the energy exchange is faster and $T_1$ shorter. Tissues containing fat molecules have also short $T_1$, because the carbon bonds at the end of fatty acids have frequencies near the Larmor frequency, thus enabling fast energy transform [60]. Usually $T_1$-times are measured with inversion recovery pulse sequence or with saturation recovery pulse sequence. For further information about these pulse sequences see [32].

### 2.2.2 Spin-spin relaxation, $T_2$

The equilibrium state for the net magnetization $xy$-component $M_{xy}$ of the nuclei is zero when the RF-pulse is off. The $M_{xy}$ obtains its maximum value when it is tilted by the 90° RF-pulse. Intrinsic magnetic field in different tissues is not totally homogeneous. In fact there are always some local inhomogeneities. These inhomogeneities cause the nuclei precess in slightly different frequencies based on eq. (2.11). Thereby the nuclei get out of phase and $M_{xy}$ decreases. The exponential decay of $M_{xy}$ to zero with a time constant $T_2$ is expressed as

$$\frac{dM_{xy}}{dt} = -\frac{M_{xy}}{T_2}. \tag{2.14}$$

The time constant $T_2$ is called spin-spin relaxation time and it decribes how $M_{xy}$ returns to its equilibrium. We get the rate of decrease of the magnetization $M_{xy}$ as a function of time by solving the eq. (2.14):

$$M_{xy}(t) = M_0\, e^{-t/T_2}. \tag{2.15}$$

After time $T_2$ the magnetization $M_{xy}$ has decreased 37% of its maximum value as can be seen from eq. (2.15) with value $t = T_2$. This is illustrated in Fig. 2.1 (b).

Spin-spin relaxation time also differs between tissues. Based on intrinsic magnetic field inhomogeneities, homogeneous tissues have long $T_2$ whereas tissues containing large impurity molecules have short $T_2$.

### 2.2.3 Relaxation time $T_2^*$

Relaxation time $T_2^*$ is also related to the transversal relaxation. The difference between $T_2^*$ and $T_2$ is that $T_2$ takes into account only the dephazing effect caused by intrinsic magnetic field inhomogeneities whereas $T_2^*$ consists of different dephazing mechanisms: relaxation caused by intrinsic inhomogeneities (relaxation time $T_2$), and inhomogeneities of the external magnetic field $B_0$ and diffusion (relaxation time $T_2'$). The relaxation time $T_2^*$ can be calculated by using these different sources of relaxation and their corresponding relaxation times as follows [27]:

$$\frac{1}{T_2^*} = \frac{1}{T_2} + \frac{1}{T_2'}. \tag{2.16}$$

In practice the dephazing effect caused by the inhomogeneities of the externeal magnetic field $B_0$ is the only effect that can be manipulated by RF-pulses. The $T_2^*$ relaxation is illustrated in Fig. 2.5.

$T_2^*$ relaxation time is important in functional MRI -studies because it is sensitive to field distortions caused by deoxyhemoglobin. This sensitivity enables the detection of active cortical areas in brain.

## 2.3 Creating an image

The NMR signal from the object is collected by a RF coil. The RF coil is sensitive to frequency of the signal but not to the location of the precessing nucleus, so it cannot distinguish two nuclei at different locations that are precessing at the same frequency. For imaging, locating the nuclei with different precessing frequencies is essential. Therefore, it is necessary to make the nucleus' precessional frequency $\omega$ depend on the location of the nucleus. This is done by using magnetic field gradients. They are additional magnetic fields that are linearly altered along a selected direction.

As can be seen from eq. (2.11), the precession frequency $\omega$ is related to the magnetic field strength $B_0$. Thus, the location information of the precessing nuclei is provided by varying the magnetic field strength with location.

Creating a two-dimensional image requires three steps: slice selection, frequency encoding and phase encoding. In slice selection the nuclei in the slice of interest are excited. The direction of slice selection is usually denoted as $z$-direction. The nuclei within that slice are then localized by using frequency encoding as in $x$-direction and phase encoding as in $y$-direction.

### 2.3.1 Slice selection

Slice selection is implemented by superimposing a slice selecting gradient $G_z$ along the $z$-axis. This relates the precession frequencies of the nuclei to positions along the $z$-axis. The exciting RF-pulse is applied at the same time as the gradient $G_z$ thus exciting only the nuclei that are precessing with the same frequency. The RF-pulse contains not only one frequency but a narrow bandwidth of frequencies. The thickness of the slice $\Delta z$ is determined by the magnitude of the slice selecting gradient $G_z$ and the frequency bandwidth of the RF-pulse $\Delta f$ [14]:

$$\Delta z = \frac{\Delta f}{\gamma G_z} \tag{2.17}$$

By increasing the frequency bandwidth thicker slice is selected. This is achieved also by decreasing the strength of the gradient field $G_z$. The frequency bandwidth of the RF-pulse is inversely proportional to the duration of the pulse: the longer the pulse, the narrower the bandwidth. With narrower frequency bandwidth the thickness of the slice is more precise because the frequency range that excites the nuclei is smaller.

### 2.3.2 Frequency encoding

The spatial location in the $x$-direction is encoded with a frequency encoding gradient $G_x$. It is also called the read gradient because it is applied at the same time as the NMR signal is acquired. As the field gradient $G_z$ changes the magnitude of the effective magnetic field in $z$-direction and thus selects a slice, the read gradient $G_x$ changes the magnitude of the effective magnetic field in $x$-direction in a selected slice. According to the Larmor relation in eq. (2.11) the nuclei in stronger magnetic field precess faster, thus, having smaller precessing frequency than nuclei in weaker magnetic field. Since the applied gradient $G_x$ is spatially linear, the different frequencies corresponds to the locations on the $x$-axis. The frequency encoding procedure specifies the image size, the field of view (FOV), along the $x$-axis [14]:

$$FOV_x = \frac{\Delta f_r}{\gamma G_x}, \tag{2.18}$$

where $\Delta f_r$ is the frequency bandwidth of the receiver. The frequency bandwidth $\Delta f_r$ is the effective range of frequencies that can be properly detected and it is controlled by the number of points $N_x$ on the signal to be digitized and the signal acquisition time $t_x$:

$$\Delta f_r = \frac{N_x}{t_x}. \tag{2.19}$$

The pixel size along the $x$-axis can be derived from the equations (2.18) and (2.19):

$$d_x = \frac{FOV_x}{N_x} = \frac{1}{\gamma G_x t_x}. \tag{2.20}$$

### 2.3.3 Phase encoding

The $y$-coordinate of a single point in the cartesian coordinate system of MRI device is received using phase encoding gradient $G_y$. The gradient $G_y$ changes the magnitude of the effective magnetic field in $y$-direction. Hence, the nuclei in different locations along the $y$-axis precess at different rate according to their $y$-position, and after a time $t_y$ undergo a phase shift $\Delta\phi$ [27]:

$$\Delta\phi = \gamma G_y y t_y. \tag{2.21}$$

The gradient $G_y$ is on only a short time and after it is turned off, the nuclei revert to the precession frequency determined by the main magnetic field. However, the difference in the phase of nuclei in different locations in $y$-direction remains after the gradient $G_y$ is turned off. The phase encoding must be repeated because the application of just one $y$-gradient does not provide enough information to generate an image. Hence, the phase encoding gradient is applied a number of times gradually stepping from $G_{y\max}$ through 0 to $-G_{y\max}$. The number of steps depends on the desired resolution and the size of the region of interest ranging typically from 128 to 512 [68].

The acquired NMR signal consists of different frequencies and phases, related to the locations of the nuclei. The phase accumulated by a nucleus at position $\vec{r} = x\vec{i} + y\vec{j}$ in the $x-y$-plane in the selected slice is [18]:

$$\phi(\vec{r}) = \gamma \left( x G_x t_x + y G_y t_y \right). \tag{2.22}$$

The magnetization $M_r$ at that particular location depends on the $\phi$ and the transversal magnetization $M_{xy}$ [14]:

$$M_r = M_{xy}(\vec{r}) e^{-i\phi}. \tag{2.23}$$

The receiver coil integrates this magnetization $M_r$ over the entire volume of the selected slice. Hence, the NMR signal $\zeta$ can be expressed as

$$\zeta = \int_{\text{slice}} M_{xy}(\vec{r}) e^{-i(xG_x t_x + yG_y t_y)} dx\, dy \tag{2.24}$$

This signal is aFourier transform of the magnetization at a single point in a spatial frequency space known as the $k$-space. The image reconstructing reduces to accumulating data over a grid in $k$-space, by varying combinations of $G_x t_x$, $G_y t_y$ and then performing a discrete two-dimensional inverse Fourier transform [18]. The $k$-space has two dimensions, $k_x$ and $k_y$, corresponding to the read gradient $G_x$ and the phase encoding gradient $G_y$, respectively. The $k$-space is illustrated in Fig. 2.2. High frequencies (high $k$-values) draw fine details in the image whereas low frequencies (low $k$ values) give shape and contrast to the images. The $k$-space is symmetrical along the $k_x$-axis. In theory only the half of the $k$-space is needed to collect for one MR image. This shortens the total imaging time. In practise, however, over 50% of $k$-space is collected because the symmetry of the $k$-space is not perfect. There are errors in the phase direction of the $k$-space and collecting more than 50% reduces these artefacts in an image. Collecting more 50% of the $k$-space also improves the signal-to-noise ratio (SNR).

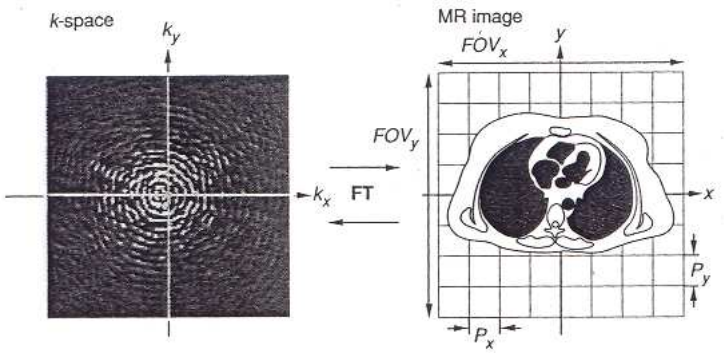



Figure 2.2: The $k$-space used in MRI and corresponding reconstructed image. (Figure reprinted from [64])

### 2.3.4 $T_1$-weighted image

Using MRI it is possible to obtain different kind of images depending on the imaging parameters and the used pulse sequence. When the contrast in an image is mainly achieved by using the spin-lattice relaxation time the obtained image is called $T_1$-weighted image.

The difference between the $T_1$ times of different tissues causes the contrast in $T_1$-weighted image with appropriate pulse sequence. The simplest pulse sequence, illustrated in Fig. 2.3, can also be called as 'pulse and acquire' [27]. The 90° pulse is applied and the response is collected. To increase SNR the pulse sequence is repeated. The time between 90° pulses is called the time to repetition $TR$. The contrast between different tissues depends on the ratio $TR/T_1$. Using long $TR$ (i.e. $TR >> 5T_1$) there is no contrast due to totally recovered longitudinal magnetizations in all tissues [14]. By optimizing $TR$ it is possible to obtain sufficient contrast between different tissues.

The basic form of received NMR-signal is the so-called free induction decay (FID) -signal. The FID-signal is created by applying an 90° pulse and letting the signal to disappear in time. The FID-signal is illustrated in Fig. 2.3.

In $T_1$-weighted image of human brain the cerebral spinal fluid (CSF) looks darker than the tissues and the gray matter looks darker than white matter because tissue containing more water looks darker than tissue containing fat. CSF contains more water than solids and gray matter contains more water than white matter.

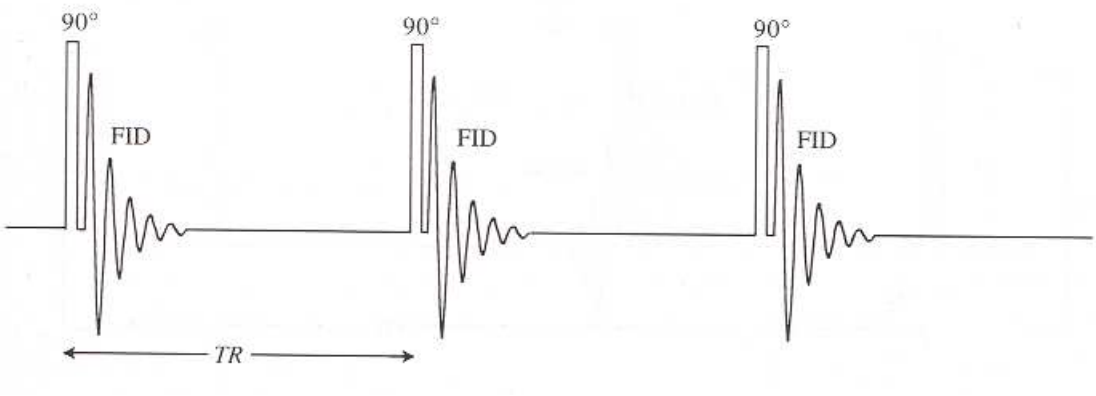



Figure 2.3: A simple pulse sequence used to obtain $T_1$-weighted images (Figure reprinted from [27, p. 247]).

### 2.3.5 $T_2$-weighted image

The pulse sequence used to obtain $T_2$-weighted image is called the spin-echo sequence and is illustrated in Fig. 2.4. In its simplest form it consists of a 90° pulse followed, after a delay time of $TE/2$, by a 180° pulse [27], where the abbreviation $TE$ means Time to Echo. The 90° pulse tilts the magnetization in $xy$-plane. Due to the local inhomogeneities ($T_2^*$ -effect) the nuclei slowly get out of phase. By applying a 180° pulse at the time $TE/2$ the coherence between the nuclei is retrieved again and the received signal at the time $TE$ is called an echo signal. The 180° pulse, however, rephases the nuclei that have undergone $T_2^*$ decay. The gradual decline in signal from subsequent echoes reflects $T_2$ decay. This process is illustrated in Fig. 2.5. The $TE$ signifies the time between the 90° pulse and the echo.

In order to diminish the $T_1$ effect in the image the used $TR$ should be quite long ($TR >> 5T_1$). Therefore, all tissues have enough time to recover their z-direction magnetization. To reveal the $T_2$ effect long $TE$ ($TE \approx T_2$) is used so that the differences in $T_2$ between different tissues have enough time to become distinguished.[14]

In $T_2$-weighted image of human brain the CSF looks brighter than the tissues and the gray matter looks brighter than the white matter.

$T_2^*$-weighted images are obtained by using e.g. gradient echo pulse sequence.

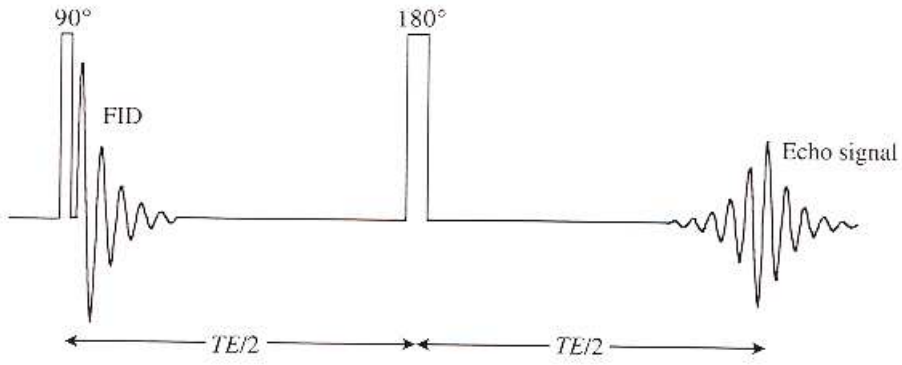



Figure 2.4: The spin-echo sequence used to obtain $T_2$-weighted images. (Figure reprinted from [27, p. 250])

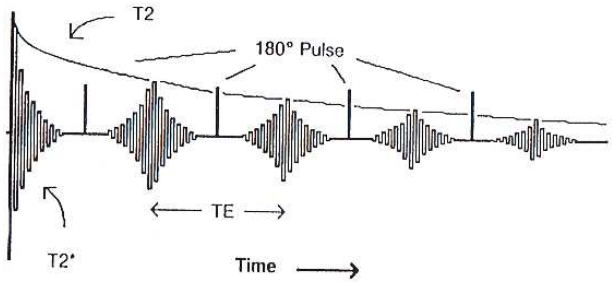



Figure 2.5: The pulse sequence used to obtain an echo signal. The envelope curve illustrates the $T_2$ effect. (Figure reprinted from [63, ch. 5,p. 7])

### 2.3.6 Proton density -weighted image

In proton density -weighted image the contrast depends on the amount of protons in different tissues. To obtain a proton density -image it is essential to minimize the $T_1$ and $T_2$ effects. To minimize the $T_1$ effect a long $TR$ is used and to minimize the $T_2$ effect a short $TE$ ($TE < T_2$) is used [14]. The received signal is, therefore, mainly influenced by differences in proton density.

In proton density -weighted image of human brain the CSF looks dark but as a difference to $T_1$-weighted image the white matter looks darker than gray matter. This is due to larger proton density in more water containing gray matter than less water containing white matter.

### 2.3.7 Echo Planar Imaging (EPI)

The conventional imaging sequences are too slow for imaging the changes in signal intensity caused by cortical activity. Long imaging time also adds motional artifacts caused by subjects movements, breathing and blood flow. To decrease these artifacts and to increase image quality faster imaging techniques are needed. The technique called Echo Planar Imaging (EPI) was invented in 1977 [47]. In principle it made possible to obtain images in fast scan time of 30 ms to 100 ms. However the technical demands were too high to MRI equipment of those days. EPI technique did not become common until in the early 1990's when the present form of EPI technique was introduced. [64]

In standard two-dimensional Fourier transform imaging methods only one line in $k$-space is acquired with each $TR$ interval. In contrast, the speed of EPI is based on the possibility to

measure the whole $k$-space applying only one excitation pulse. This difference in measuring the $k$-space is illustrated in Fig. 2.6. Hence, only one radio frequency excitation is needed for acquiring

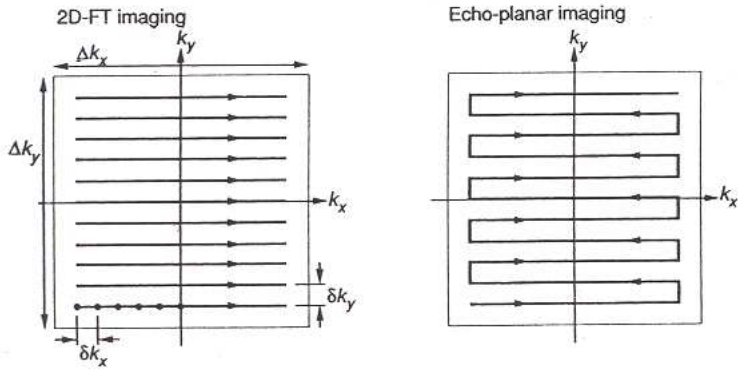



Figure 2.6: The difference between normal pulse sequences and EPI sequence in spanning the $k$-space. (Figure reprinted from [64])

one slice. The gradient-echo EPI pulse sequence is illustrated in Fig. 2.7. Instead of keeping

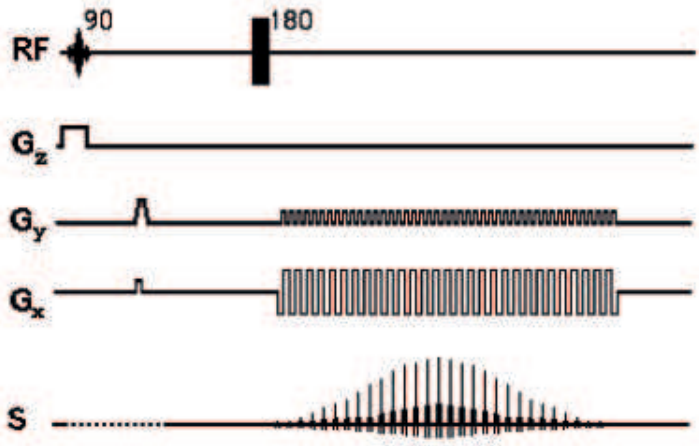



Figure 2.7: The gradient echo EPI sequence. $RF$ is the exciting radio frequency pulse, $G_z$ is the slice selecting gradient, $G_y$ is the phase encoding gradient and $G_x$ is the frequency encoding gradient. $S$ is the received NMR signal (Figure reprinted from [35])

the frequency encoding gradient $G_x$ constant, as in standard two-dimensional Fourier transform imaging method, the frequency encoding gradient is rapidly oscillated during the build-up and decay of the echo signals. A series of gradient echoes is thereby produced, each one of which is separately phase encoded by application of a brief phase encoding gradient $G_y$ [14]. The contrast in gradient echo EPI images is mainly affected by the spin density and the $T_2^*$. The $T_1$-effect is excluded due to the use of only one RF-pulse.

The capability of measuring the whole $k$-space with one RF-pulse speeds up the imaging time enormously. With EPI one slice can be collected in less than 100 ms with adequate in-plane resolution. This enables to use EPI to study human brain function where the changes in brain function happens quite fast. The imaging time cannot be reduced infinitely. Limitations arise from

the gradient duty cycle capacity and the limited data input-output rate.

Although EPI is fast imaging sequence, it has its own weaknesses. The rapid switching of the gradients cause a high frequency mechanical pulse which generates loud noise. Thus patients must always use hearing protectors. If the rate of change of the magnetic field is too rapid it can induce currents in the body which can interference cardiac functions [59]. This can also stimulate the nerves and make the muscles twitch. At 1,5 T magnetic field this is, however, almost never an issue. EPI images suffer also from many kind of artifacts due to fast spanning of the $k$-space (see Ch. 3.6.1). To reduce these artifacts there are numerous variations of the basic EPI pulse sequence e.g. using different routes to measure the $k$-space. More detailed explanation of these different EPI sequence variations can be found in [32].

Chapter 3

# Functional Magnetic Resonance Imaging (fMRI)

Functional neuroimaging is a general term for several different brain imaging methods such as positron emission tomography (PET), single photon emission tomography (SPET), electroencephalography (EEG), magnetoencephalography (MEG) and functional magnetic resonance imaging (fMRI) (for overview, see [36]). Although the bases of these methods are different they are all used to reveal the function of human brain.

Using MRI to study brain function is relatively new invention. First fMRI studies were taken place in the early 1990's. During the last decade fMRI has evolved with great speed. The main reason for this development have been the use of EPI sequences and increased computer speed. FMRI is one of the most remarkable neuroimaging methods because one can visualize functionally activated regions in the brain completely non-invasively and without the use of radioactivity. Also the spatial and temporal resolution of fMRI is usually better than with other techniques. Thereby fMRI has spread all over the world and become popular among the neuroimaging specialists. Several neuroimaging groups have used fMRI to study vision, motor, language, memory, emotion and pain as well as stroke, epilepsy, amputation and phantom limbs and psychiatric disorders. It has also been used in presurgical planning [37].

## 3.1 Blood oxygenation level dependent (BOLD) contrast

The most common fMRI technique in brain function studies relies on the oxygenation of hemoglobin. In the oxygenated state hemoglobin is diamagnetic and has no effect to the intrinsic inhomogeneities in the tissue. Deoxygenated hemoglobin is paramagnetic and can thus act as an contrast agent accelerating the transversal magnetization by increasing inhomogeneities, which affects the intensity in $T_2^*$-dependent EPI-images. This phenomena is called the blood oxygenation level dependent (BOLD) effect and the contrast it creates in functional images is called the BOLD contrast. Changes in BOLD contrast were demonstrated first time using rats inhaling different gases [52]. It was observed that under normoxic conditions arterial blood was totally oxygenated and had no effect to the BOLD contrast, while venous blood was deoxygenated and affected the BOLD contrast. It was also shown that the BOLD contrast relied on an intrinsic contrast agent and image acquisition could be precisely synchronized to external stimuli with good time resolution. This encouraged many research groups to continue studying BOLD contrast.

PET-studies have revealed that local oxygenation state of blood is affected by neuronal activity. In active cortical area there is increase in cerebral blood flow (CBF) and in cerebral blood volume (CBV). The activity, however, does not remarkably increase the relative cerebral metabolic rate of oxygen ($CMRO_2$) and thus the venous oxygenation increases i.e. the relative venous deoxygenation decreases. This causes a small but detectable change in $T_2^*$-weighted signal intensity. In the early 1990's several studies related the increase of the signal intensity to neuronal activity [43, 11]. In these studies changes in image intensity were observed on visual cortex by using flashing checkerboard pattern stimulation [11] or patterned-flash photic stimulation [43]. The activity was also detected on the motor cortex with hand squeezing task [43].

Because the BOLD contrast does not measure absolute blood flow but relative changes in

venous blood oxygenation, the exact localization of the active cortex may be disrupted by the large blood vessels. The larger the blood vessel the stronger the signal due to the bigger contrast agent volume. The signal from large vessel has however a larger time delay than signal from the cortical area [44]. These time delays could be used to differentiate signals that originate from large vessels from signal that originates from the cortical area, and hence improve the acquired spatial resolution.

## 3.2 The BOLD response

Functional MRI improved fast in the 1990's but the exact phenomena behind the BOLD i.e. the correlation between the deoxygenation concentration and neuronal firing is not yet completely understood. The neural response caused by a stimulus is related to the CBF, CBV and $CMRO_2$. The disproportional changes of CBF, CBV and $CMRO_2$ generate the BOLD response [50]. This is illustrated in Fig. 3.1. The characteristics of the steps in the transformation from stimulus to

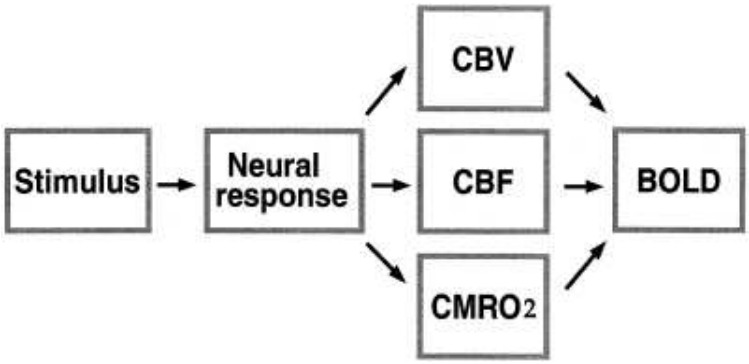



Figure 3.1: Diagram of the transformation from stimulus presentation to BOLD response. (Figure reprinted form [50])

response in MRI signal define the data analysis methods suitable for different study paradigms. An important aspect is whether any of these steps is linear and time-invariant. A system $L\{\cdot\}$ is linear if

$$L\left\{\sum_j a_j f_t^{(j)}\right\} = \sum_j a_j L\left\{f_t^{(j)}\right\}. \tag{3.1}$$

In fMRI this means that BOLD responses to short-duration stimuli could be used to linearly predict the response to long-duration stimuli. Previous studies have demonstrated that the relationship between stimulus and BOLD response is not purely linear. Nonlinearities of the responses have been observed in the visual cortex [67, 34] and motor and auditory cortex [29, 49]. The nonlinear behaviour has been reported to vary accross the cortex [10]. There are studies that have reported that the BOLD response begins to behave linearly when the short-duration response is longer than some threshold duration [67, 12] but also long stimuli were found to behave in a nonlinear way producing smaller response amplitude than the linear model predicts. The nonlinearity of the BOLD response was found to be sum of neural response adaptation and BOLD saturation [50]. The nonlinear behaviour also varies depending on the used experimental design [48]. These studies, however, concentrated on the stimulus-BOLD response step forgetting the steps in the middle. The step from stimulus to neural activity was found to be nonlinear, the step from neural activity to CBF change was found to be linear and the last step from CBF change to BOLD signal change was found to be nonlinear [50].

The shape of the hemodynamic response has been under investigation for a long time although the nonlinearity complicates the studies considerably. One model of the hemodynamic response is

the so called gamma function [12]:

$$h(t) = \frac{(t/\tau)^{n-1}\ e^{-(t/\tau)}}{\tau\,(n-1)\,!}, \tag{3.2}$$

in which $h(t)$ is the hemodynamic response as a function of time $t$, $\tau$ is the time constant and $n$ is an integer representing the phase delay. In addition it is possible to allow a delay between stimulus onset and the beginning of the fMRI response. Another model for hemodynamic response is the Gaussian function model [56]

$$h(t) = G(t;\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}, \tag{3.3}$$

where $\mu$ represents the lag and $\sigma^2$represents the dispersion. The hemodynamic response has been modelled also with impulse response functions [72] or truncated Gaussian functions [31]. One possibility is to use Volterra kernels in hemodynamic response modeling [25].

The shape of the hemodynamic response varies person to person. The form of the response does not even stay constant in one person if the same paradigm is performed several times in different days. Also different cortical areas produce different responses. In [1] it was found that a significant variability exists in the shape of the hemodynamic response between subjects and also within subject if collected at different times. Immediately after stimulus some subjects with certain stimuli show a small initial dip of about 1 s duration [20]. The main BOLD signal increases after that with peak between 4 s and 8 s after the stimulus onset. After the peak the signal decreases with a small undershoot [1] and then returns slowly back to the baseline. Typical shape of the hemodynamic response is illustrated in Fig. 3.2.

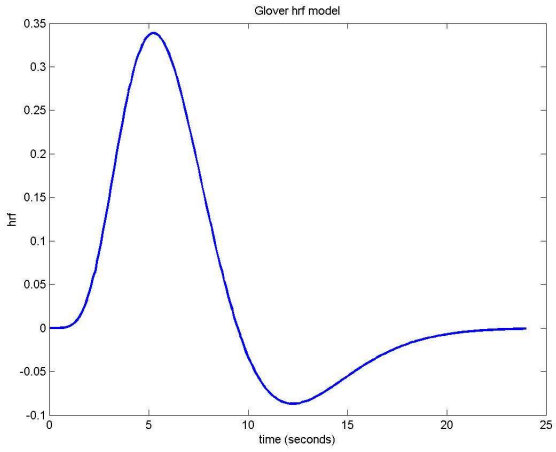



Figure 3.2: A rough estimation of the shape of the hemodynamic response for a brief stimulus. This estimation is a difference of two gamma density functions (Figure reprinted from [69]).

The statistical power (i.e. the capability of finding the actual active voxels) of few different hemodynamic models have been compared in [53]. The result of the study was that there is no clearly superior model for hemodynamic response. All the models in the study had strengths and weaknesses. The nonlinearity of the hemodynamic response and the lack of thorough understanding of the phenomena sets the limits to the inventing the suitable model of hemodynamic response.

## 3.3 Data acquisition

Functional MR images are usually collected using gradient echo EPI imaging protocoll because it is fast and sensible to the BOLD contrast. Usually the data acquisition is performed as follows: The

subject lies in the bore of the magnet head in a head coil. At first some high-resolution (spatial resolution of $1 \times 1 \times 1$ mm) $T_1$-weighted anatomical images are collected. The anatomical images are used later to relate the active areas detected to the anatomical landmarks in the brain. After the anatomical images are collected begins the functional part of the examination. The patient may perform some motor tasks e.g. finger tapping, visual tasks, memory tasks or even auditory tasks. While performing the tasks the functional images are collected.

The quality of collected images depends on the chosen imaging parameters used for data acquisition. Different parameters and their effect to image quality have also been studied [38]. It was noticed that slice collection order has no effect to image quality but slice thickness has. The thicker the slices are the shorter the imaging time is and the better the sensitivity per unit time is. Thin slices have also benefits: using thin slices the areas which suffer from magnetic susceptibility artifacts have more signal. Also the spatial resolution is better when thinner slices are used. The disadvantage in using very thin slices is that the SNR is poorer than using thicker slices. Good SNR assures better and more reliable results due to lower disturbing noise component.

### 3.3.1 Time series

The functional images are $T_2^*$-weighted images with lower spatial resolution than in anatomical images. However, functional images are not considered as anatomical images of the object and they are not examined as ones. The collection of them in certain time rate constitutes a time series of images. Actually there are several time series since the intensity values of every voxel constitutes a time series. Each voxel's intensity value in each instant is one value in their time series. This is illustrated in Fig. 3.3 for a single voxel.

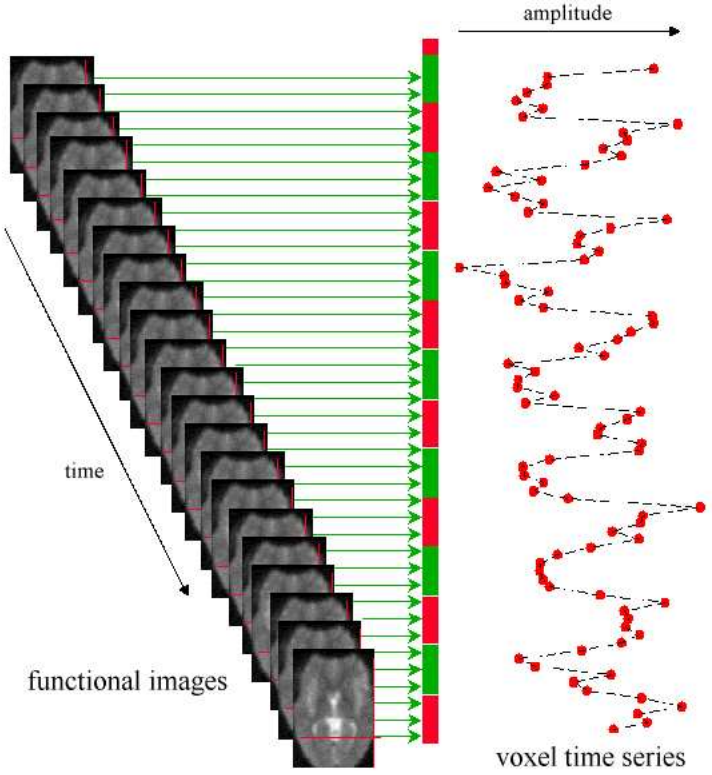



Figure 3.3: Voxel time series in fMRI. (Figure reprinted from [55])

## 3.4 Experimental designs

The choice of experimental design when setting up a fMRI research depends on the results wanted and the topic of the research. Two basic types of setups are used when concerning the appropriate and suitable designs: block design, event-related design [57, 40] or a combination of these two. All these designs have their weaknesses and strengths which should be considered when choosing the design.

### 3.4.1 Block design

The first fMRI studies used a similar study paradigms as had been used in PET and SPET studies. In these studies the activating stimuli are presented continuosly during some time interval that is called a block. The blocks of activating stimuli are usually alternating with the so called baseline or resting blocks. During the baseline block no stimuli are presented. One active block may be consisted of only one long stimulus or several similar stimuli presented rapidly. It is also possible to study different stimuli by presenting each stimuli type in its own block. Therefore, several type of blocks might belong in one study and the order of the blocks may alternate randomly. The duration of the blocks may also vary.

Block designs are still in use nowadays. One reason for this is probably better signal-to-noise ratio (SNR) due to bigger amount of data to be averaged. This also ensures better detection power, a measure of the ability to locate active cortical regions. As a weakness, block designs have poor estimation efficiency, a measure of the ability to estimate the hemodynamic response for a single stimulus [45]. This is basically due to fast presenting rate of the stimuli so the responses overlap with each other. This overlapping is proved to be nonlinear, which complicates the estimation of the shape of the hemodynamic response. Block designs are also experimentally less demanding than more flexible desings. The possible inaccuracies in scanner synchronization with the stimulus equipment, stimulus timing, or stimulus randomization are less serious in block designs than in event-related designs because the responses in one block are averaged.

### 3.4.2 Event-related design

In SPET studies only block designs can be assessed because of the relatively long half-life of the used radioactive tracers. In fMRI, however, the origin of the response to a stimulus can be related to the cerebral hemodynamic changes, and these changes are detectable within seconds of the stimulus onset. The relatively fast response to stimuli enables the use of brief stimuli in studies of brain function. In the early 90's a stimulus duration of 2 seconds was shown to produce detectable signal changes [11]. Later, the signal changes produced by even shorter stimuli were detected.

A design of brief stimuli presented randomly is called an event-related design. The term event-related derives from electrophysiology and measuring the event-related potentials (ERPs). The design and the presentation of stimuli in fMRI is quite similar to technique used in measuring the ERPs. The stimuli are no longer presented in blocks of similar stimuli but one type of stimuli can be randomized [19] so that different types of stimuli alter with each other and with baseline. The presentation rate may also vary i.e. a stimulus may occur twice a second or twice a minute.

Event-related design has many virtues compared to block design. When the stimuli are presented in blocks, the subject's cognitive behaviour may disrupt the response because the subject can quess when the next stimulus is presented and what kind of stimulus it is. The randomization of the stimuli prevents this kind of problems and also habituation. The responses can be post hoc categorized according to subjects performance and hence it is possible to study the difference between different responses caused by similar stimuli. Another advance of the event-related designs compared to block design is the ability to use the so called oddball-paradigm and study unpredicted stimuli. The advantages of the event-related design over block design encouraged research groups to study and compare the results obtained with both these design types [13, 15]. The results of these studies showed that the results of event-related studies and block studies were comparable.

## 3.5 Resolution

One reason why fMRI is so popular is its good spatial and temporal resolution. Spatial resolution in fMRI means the smallest activated area that can be reliably detected. It is basically limited by imaging time and by reasonable SNR. Smaller voxels indicate smaller SNR but also improve spatial resolution by enabling the detection of smaller structures and smaller activated areas. The majority of echo-planar images collected with 1.5 T have a resolution of $3 \times 3 \times 3$ mm or larger. However, this is not the final limit. With higher magnetic field strength (3 T or more) the spatial resolution will be better. It has been reported of studies at higher field strength with successful BOLD contrast detection using $1 \times 1 \times 3$ mm resolution. The final limit is determined by physiological and hemodynamical limits, i.e. by the size of the smallest vessels that show local changes with neural activity. This limit is somewhere between 0.1 and 1 mm. [59]

Temporal resolution is defined as the shortest time between two stimuli in the same cortical area that produce distinguishable responses. MR scanner sets the limit of the speed the images are collected and therefore also the temporal resolution. In the echo planar imaging it is possible to collect even 20 volumes per second but because of the heating of the gradients it is possible to maintain this rate much less than a minute. When the image collection rate is 7 volumes per second, it is possible to carry on the imaging without the heating problem. [59]

The MR scanner is not the main limiting factor for the temporal resolution. The principal reason for fMRI not achieving as good temporal resolution as for example EEG or MEG is the hemodynamic response. The assumed typical shape of the hemodynamic response for a fast stimulus is illustrated in Fig. 3.2. As can be seen, the duration of the response is several seconds and there is a delay before its initial rise. This relatively slow hemodynamic response acts as an filter that disperses and delays the neuronal response for a stimulus. This is the main limitation of the temporal resolution in fMRI based on BOLD contrast.

The limitations of temporal resolution in fMRI has been studied in [42]. It was discovered that if there are two events in the same area less than four seconds apart they cannot be distinguished without deconvolution. If two events occur in different cortical areas less than four seconds apart, they can be distinguished. It was also noticed that the time it takes the hemodynamic response to arise differs from subject to subject. There were also differences in arising times of hemodynamic response within a subject in different parts of the brain. It was argued that the reason to this is the different vascular environment in different parts of the brain.

## 3.6 Artifacts

Functional MRI studies necessitate an equipment capable to achieve high quality MR images rapidly enough. For this the EPI sequence is usually used. This sequence however is rather demanding for the MRI scanner and this causes artifacts in images. The main problems in fMRI studies are the magnetic susceptibility differences, Nyquist ghosts and motion.

### 3.6.1 Magnetic susceptibility differences

Functional magnetic resonance imaging is based on $T_2^*$ -relaxation, which is sensitive to microscopic and macroscopic field inhomogeneities in order to generate contrast in an image. Different substances such as tissue, air and bone have different magnetic susceptibilities. These differences establish local field gradients typical to structure of brain. The extra field gradients disturb the $T_2^*$ -relaxation and thus cause signal loss. This signal loss is severe in particular areas where there are major interfaces between air, bone and/or tissue e.g. in frontal lobe (the air-filled sinuses), and in inferior temporal lobe (the petrous bone and the air passages of the ear) [39].

Differences in magnetic susceptibility may also cause geometric distortion. This is a cause of spin dephasing in tissue depending on the extra field gradient.

It is possible to decrease the magnetic susceptibility artifacts by using thinner slices and planning the slice orientation carefully. In most cases, coronal slices show less loss than axial slices in the frontal and temporal lobes [59].

### 3.6.2 Nyquist ghosts

Nyquist ghost is a typical problem for images collected with EPI sequence. This low intensity ghost image occurs on the image in halfway across the field of view. In EPI adjacent lines in $k$-space are sampled under read gradients with opposite directions. If there are misalignment in sampling, eddy currents, nonlinearities in receiving filters or differences in positive and negative gradients, there is an alternate line modulation in k-space, which leads to a 'ghosting' of the image.

Nyquist ghost may be reduced by calibration of the gradients in the magnet. Because there are intensity changes during time in Nyquist ghost, it is difficult to remove it using any standard post processing procedure.[39]

### 3.6.3 Motion

Head motion is the biggest problem in fMRI studies and the main reason in abandoning images from the functional study. By fixating the head firmly to the receiver coil it is possible to prevent large movements of the head, but movements of 1-2 mm are still common.

Although the head would be perfectly fixed, there exists still low-frequency motion in the brain caused by respiration and cardiac cycles. Cyclic blood flow in arteries causes movements in whole brain and especially in subcortical areas. This effect is even strengthened by the pressure changes in arteries.

Motion artifacts are possible to be reduced by realignment of the slices after the image acquisition, but it does not correct all the artifacts. Hence, it is essential to have as cooperative subjects as possible to prevent the head movements and also be careful with fixating the head to the coil.

Chapter 4

# Data-analysis

Functional MRI data-analysis is not image processing although the studied data is in format of images. The data is concerned as voxel time series (section 3.3.1) and the image gives only the spatial location for the time series.

## 4.1 Least Squares (LS) -estimation

Let us consider a simple linear regression example. Let a vector $\mathbf{z} = (z_1, z_2, \ldots, z_N)^T \in \mathbb{R}^N$ denote the observations, and a vector $\theta = (\theta_1, \theta_2, \ldots, \theta_m)^T \in \mathbb{R}^m$ the parameters to be estimated. The observations may be written in the form:

$$\mathbf{z} = \mathbf{H}\theta + \mathbf{v}, \tag{4.1}$$

where $\mathbf{H}$ represents the $N \times m$ known design matrix, and $\mathbf{v}$ is a $N \times 1$ column vector of the errors.

The columns of the matrix $\mathbf{H}$ span a $m$-dimensional subspace $R(\mathbf{H})$. This is illuistrated in Fig. 4.1. The least squares estimation is based on solving the parameters $\theta$ in a way that minimizes

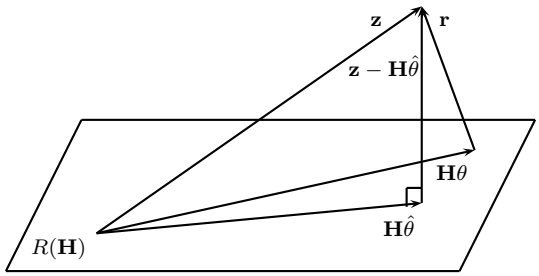



Figure 4.1: Subspace $R(\mathbf{H})$

the squared error norm

$$l(\theta) = \| \mathbf{z} - \mathbf{H}\theta \|^2 . \tag{4.2}$$

This happens as the residual $\mathbf{r} = \mathbf{z} - \mathbf{H}\theta$ is ortohogonal to the subspace $R(\mathbf{H})$. Then we can write

$$\langle \mathbf{z} - \mathbf{H}\hat{\theta}, \mathbf{H}\theta \rangle \; = \; 0, \; \forall\, \theta. \tag{4.3}$$

The inner product in eq. (4.3) may be written

$$(\mathbf{H}\theta)^T \left( \mathbf{z} - \mathbf{H}\hat{\theta} \right) \; = \; \theta^T \mathbf{H}^T \mathbf{z} - \theta^T \mathbf{H}^T \mathbf{H}\hat{\theta} \; = \; \theta^T \left( \mathbf{H}^T \mathbf{z} - \mathbf{H}^T \mathbf{H}\hat{\theta} \right). \tag{4.4}$$

This can be written as inner product

$$\langle \mathbf{H}^T \mathbf{z} - \mathbf{H}^T \mathbf{H}\hat{\theta}, \theta \rangle \; = \; 0, \; \forall\, \theta. \tag{4.5}$$

The above is true when

$$\mathbf{H}^T\mathbf{z} - \mathbf{H}^T\mathbf{H}\hat{\theta} \;=\; 0 \quad \Rightarrow \quad \mathbf{H}^T\mathbf{H}\hat{\theta} \;=\; \mathbf{H}^T\mathbf{z}. \tag{4.6}$$

From here we get the estimate for the parameters $\theta$

$$\hat{\theta} = \left(\mathbf{H}^T\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{z}. \tag{4.7}$$

This is the least squares (LS) solution for the parameters $\theta$. [41]

## 4.2 Statistics

In one EPI volume there are about 50 000 - 100 000 voxels depending on the imaging parameters, which leads to equal number of voxel time-series. The difference in voxel's intensity values between activated and non-activated state is only few per cents. Therefore, examining all voxel time series by hand or by subtracting mean images of active state and non-active state is not reasonable. To solve these problems statistical tests are needeed.

Basically statistics is used for revealing the active cortical areas. After LS-estimation statistical tests are used to decide whether the estimated parameters fit the data well enough, so that one can say that specific voxel was active during stimuli.

For statistical testing the question of interest is simplified into two competing hypothesis: a null hypothesis and an alternative hypothesis. In all cases one of these hypothesis must be true in order to achieve reasonable results from the used test. The hypothesis testing is always performed in a way that according to the result of the test it is possible either accept or reject the null hypothesis. The null hypothesis must be established such that when it is true the probability distribution of the question of interest is known. The choice of the test statistic will depend on the assumed probability distribution and the hypothesis under question. If the test value does not fit the chosen probability distribution, the null hypothesis is rejected. The decision of whether to reject the null hypothesis or not is made by using a threshold value. The threshold is the value to which the value of the test statistic is compared. If the test value exceeds the threshold, the null hypothesis is rejected. The threshold is decided by using significance levels and p-values. p-value is a probability of rejecting the null hypothesis if it in fact is true, and, therefore, it contains the amount of mistakenly rejected tests. In fMRI the statistical test is performed to every voxel. If the chosen significance level is e.g. 0.05 ($p < 0.05$), it means that five per cent of the voxel denoted as active are actually not active i.e. they are false positives. Small p-value indicates also that the null hypothesis is unlikely to be true and the smaller it is, the more convincing is the rejection. The conclusion of the test statistic is usually presented with the p-values because the p-value contains the information about how significant the statistical result is.

### 4.2.1 T-test

In the early $20^{th}$ century W. Gosset showed that the mean of a sample from a normal distribution with unknown variance does not follow the normal distribution. W. Gosset called the distribution the *t distribution* and later it was been referred as *Student's t distribution* because W. Gosset published his studies under the name 'Student'. The properties of the population from where the samples were collected can be described by using the t distribution. This is done by using the t-test, which assumes that the data is normally distributed with equal variance. One form of t-test is a *one sample t-test*, where the null hypothesis is that the data are sampled from a population with specific hypothesized mean $k$. The value of $t$ is calculated as [2]:

$$t \;=\; \frac{\overline{x} - k}{s/\sqrt{n}} \;=\; \frac{(\overline{x} - k)\sqrt{n}}{s}, \tag{4.8}$$

where $\overline{x}$ and $s$ are the mean and standard deviation of the sample of size $n$, respectively. If one wants to compare two independent groups of data, the *two sample t-test* is used [66]:

$$t \;=\; \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} - \frac{\sigma_2^2}{n_2}}}, \tag{4.9}$$

where $\overline{x}_1$ and $\overline{x}_2$ are the means of the groups, $\sigma_1$ and $\sigma_2$ the standard deviations and $n_1$ and $n_2$ the sizes of the groups, respectively.

Chapter 5

# Spatial preprocessing in SPM99

Functional MRI data usually contain many types of artifacts as well as noise. The intensity change of BOLD is relatively small, only few per cents. Preprocessing steps are performed to enhance the SNR and reduce the confounding effects of artifacts. In SPM99 there are five preprocessing steps that are needed in fMRI data analysis: the realignment, the slice timing, the coregistration, the normalization and the smoothing. All preprocessing steps are not necessary for all studies, but realignment is a step that usually is performed. The realignment corrects the most severe artifacts, the movement related artifacts. The slice timing corrects the differences in acquisition times between the first and the last slice in one functional image volume. In order to get the functional images and the anatomical image into the same space the coregistration step is performed. The normalization transforms the images into a standard space defined by a template image. This also enables the group comparisons. The last preprocessing step is the spatial smoothing. It is performed in order to enhance the SNR and to transform the data to fit better the assumptions of the statistics used in the analysis and to diminish the interindividual differences in neuroanatomy.

## 5.1 Realignment

One of the strongest artifacts are the movement related artifacts (section 3.6.3). The large movements may cause so much confounding effects that the data must be discarded from the study. Because the head movement in the scanner cannot be completely eliminated, it is necessary to correct the movement mathematically afterwards. Movement affects the statistical results in two ways: *i*) movements add the residual variation causing the activation to vary between adjacent voxels thus decreasing the estimation efficiency, and *ii*) the movements may be considered as activations if they correlate the stimuli. For example the subject listens to different sounds and is instructed to press the button every time he or she hears certain different sound. If the pressing the button causes movements of the head, it correlates also with the stimuli and may confound the results.

The difference between two images $f$ and $g$ consists of two components [24]: *i*) voxel value or intensity differences, when the two images are in the exactly same space and *ii*) misalignment or differences in shape or size of the objects. The second component of the difference is in SPM99 taken into account via rigid body transformations, that assumes head to be a rigid body. Rigid body transformations are special cases of more general affine transformations. In three dimensions six parameters are needed: three translations $x_\mathrm{t}$, $y_\mathrm{t}$, $z_\mathrm{t}$ and three rotations $\theta$, $\phi$, $\omega$ about the $x$, $y$, and $z$ axes in radians, respectively. These parameters form a vector $\mathbf{p} = [x_\mathrm{t},\ y_\mathrm{t},\ z_\mathrm{t},\ \theta,\ \phi,\ \omega]^T$ [3].

A matrix that implements the translations is:

$$\mathbf{M_t} = \begin{bmatrix} 1 & 0 & 0 & x_t \\ 0 & 1 & 0 & y_t \\ 0 & 0 & 1 & z_t \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{5.1}$$

The matrices that implement the rotations are:

$$\mathbf{M}_\theta = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \mathbf{M}_\phi = \begin{bmatrix} \cos\phi & 0 & \sin\phi & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\phi & 0 & \cos\phi & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{5.2}$$

$$\text{and } \mathbf{M}_\omega = \begin{bmatrix} \cos\omega & \sin\omega & 0 & 0 \\ -\sin\omega & \cos\omega & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

[3]

Voxel sizes of images may differ and this must be taken into account. For simplicity the Euclidian space, which defines distances in millimeters rather than in voxels, is used. The transformation into Euclidian space is carried out by simple affine transformation matrix. A transformation matrix $\mathbf{M_e}$ for image $e$ with matrix size $a \times b \times c$ and voxel size $x_{mm} \times y_{mm} \times z_{mm}$ can be defined as:

$$\mathbf{M_e} = \begin{bmatrix} x_{mm} & 0 & 0 & -\frac{a}{2}x_{mm} \\ 0 & y_{mm} & 0 & -\frac{b}{2}y_{mm} \\ 0 & 0 & z_{mm} & -\frac{c}{2}z_{mm} \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{5.3}$$

Here the axes of the Euclidian space are parallel to those of the image and the origin is in the middle of the image.

For rigid body transformations a position $\mathbf{x} = [x_1, x_2, x_3, 1]^T$ in image $f$ is mapped to a position $\mathbf{y} = [y_1, y_2, y_3, 1]^T$ in image $g$ by rigid body transformation:

$$\mathbf{y} = \mathbf{M}\mathbf{x}. \tag{5.4}$$

Here the matrix $\mathbf{M}$ is a transformation matrix that is formed by using the Euclidian space transformation matrix (from eq. 5.3) and translation (eq. 5.1) and rotation (eq. 5.2) matrices [3]:

$$\mathbf{M} = \mathbf{M_f^{-1}}\,\mathbf{M_t}\mathbf{M}_\theta\mathbf{M}_\phi\mathbf{M}_\omega\mathbf{M_g}, \tag{5.5}$$

where $\mathbf{M_f}$ and $\mathbf{M_g}$ are the transformation matrices into Euclidian space using eq. (5.3) for images $f$ and $g$ that are to be registered together.

The images that are to be realigned may have differences in their intensity level. This difference in voxel intensity is taken into account via an extra parameter $s$. Now the parameter vector is $\mathbf{p} = [x_t, y_t, z_t, \phi, \theta, \omega, s]^T$. The best possible fit for all $j$ voxels between the two images $f$ and $g$ is obtained by minimizing the following equation and solving the parameters $\mathbf{p}$

$$l(\mathbf{p}) = \sum_j \left(\mathbf{M}\,\mathbf{x}_j - s\mathbf{y}_j\right)^2. \tag{5.6}$$

Let us denote vector $\mathbf{b_a}$ such that

$$\mathbf{b_a} = \mathbf{M}\mathbf{x} - s\,\mathbf{y}.$$

By using Taylor's theorem in the neighbourhood of some $\mathbf{p}^\star$ $l(\mathbf{p})$ gets a form

$$l(\mathbf{p}) = l(\mathbf{p}^\star) + \mathbf{J_l}(\mathbf{p} - \mathbf{p}^\star), \tag{5.7}$$

where $\mathbf{J_l} = \frac{\partial l(\mathbf{p}^\star)}{\partial \mathbf{p}}$ and $\mathbf{p}^\star$ is. To solve the parameters $\mathbf{p}$ the equation (5.6) is therefore written as:

$$\begin{aligned} l(\mathbf{p}) = \| \mathbf{M}\mathbf{x} - s\mathbf{y} \|^2 &\approx \| \mathbf{M}^\star\mathbf{x} - s^\star\mathbf{y} - \mathbf{J_l}(\mathbf{p} - \mathbf{p}^\star) \|^2 \\ &= \| \mathbf{M}^\star\mathbf{x} - s^\star\mathbf{y} - \mathbf{J_l}\mathbf{p} + \mathbf{J_l}\mathbf{p}^\star \|^2. \end{aligned} \tag{5.8}$$

By denoting $\mathbf{h} = \mathbf{M}^\star\mathbf{x} - s^\star\mathbf{y} + \mathbf{J_l}\mathbf{p}^\star$ we can write the equation (5.8) in a linear form:

$$l(\mathbf{p}) = \| \mathbf{h} - \mathbf{J_l}\mathbf{p} \|^2. \tag{5.9}$$

This is now easy to solve by using LS-estimation (Ch.4.1, eq. 4.7):

$$\hat{\mathbf{p}} \;=\; \left(\mathbf{J_l}^T\mathbf{J_l}\right)^{-1}\mathbf{J_l}^T\mathbf{h}. \tag{5.10}$$

By solving $\mathbf{h}$, the equation above gets the form

$$\begin{aligned}\hat{\mathbf{p}} &= \left(\mathbf{J_l}^T\mathbf{J_l}\right)^{-1}\mathbf{J_l}^T\left(\mathbf{M}^\star\mathbf{x} - s^\star\mathbf{y} + \mathbf{J_l}\mathbf{p}^\star\right)\\ &= \left(\mathbf{J_l}^T\mathbf{J_l}\right)^{-1}\mathbf{J_l}^T\mathbf{b_a}^\star + \left(\mathbf{J_l}^T\mathbf{J_l}\right)^{-1}\mathbf{J_l}^T\mathbf{J_l}\mathbf{p}^\star \;.\\ &= \mathbf{p}^\star + \left(\mathbf{J_l}^T\mathbf{J_l}\right)^{-1}\mathbf{J_l}^T\mathbf{b_a}^\star\end{aligned} \tag{5.11}$$

This can be written in iterative form for improving the estimation of the parameters $\mathbf{p}$. For iteration $n+1$ the parameters $\mathbf{p}$ are updated as:

$$\mathbf{p}^{(n+1)} \;=\; \mathbf{p}^{(n)} + \left(\mathbf{J_l}^T\mathbf{J_l}\right)^{-1}\mathbf{J_l}^T\mathbf{b_a}^{(n)}. \tag{5.12}$$

This form of iteration is Gauss-Newton optimization strategy.

SPM99 performs the realignment by registering the whole time-series of images to image picked first. The realignment parameters (parameters in vector $\mathbf{p}$) estimated for every image, are shown in a graphical form after the realignment is performed.

## 5.2 Slice timing

In normal fMRI-studies the scan time of individual slice differs within one $TR$ depending on the study paradigm and imaging protocol. The longer the scan time, the bigger is the difference in imaging time between the first and the last slice in the same echo planar image volume. The slice timing -preprocessing step in SPM99 makes the data in each slice to correspond the same time instant. This is done by a shift of the phase of the sines that make up the signal. The signal is convolved with a shifter filter to apply the phase shift and the correction is computed using sinc interpolation. The result is time-series having the values that would have been obtained if the slice had been acquired at the beginning of each scan.

Let us now consider a case where fMRI data consists of $k$ echo planar images ($k$ time points). Each image volume has $q$ slices with $m \times n$ pixels. For each column $c = 1 \ldots n$ in each slice $s = 1 \ldots q$ the data is reshaped to $k \times m$ matrix $\mathbf{A_{cs}}$:

$$\mathbf{A_{cs}} \;=\; \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & & & \vdots \\ a_{k1} & \cdots & & a_{km} \end{pmatrix} \tag{5.13}$$

where an element $a_{ij}$ represents the intensity value of $j^{\text{th}}$ row in column $c$ of slice $s$ at a time instant $i$.

To perform a slice timing correction a shifter vector of range of phases up to the Nyqvist frequency is created. The time shifting of the data is performed by adding a constant, that depends on the number of slices, the scan time and the slice acquisition order, to each phase. Next these phases are transformed into the frequency domain.

The correction of the differences in slice timing is performed by convolving the columns of the matrix $A_{cs}$ one by one with the complex shifter vector. The convolution is performed in frequency domain by using the convolution theorem (convolution in time domain equals multiplication in frequency domain). The transformations between time and frequency domains are performed by using the discrete Fourier transformation and its inverse transformation. After the convolution the columns of the matrix $A_{cs}$ are re-inserted into a matrix formation.

## 5.3 Coregistration

Functional MRI requires fast imaging method (such as EPI) to be able to measure the changes in fMRI signal caused by brain activity. Usage of EPI sequence, however, sets limits to image quality and spatial resolution. Typically the in-plane resolution in functional images is $2 \times 2\text{mm}^2 - 4 \times 4\text{mm}^2$ or even bigger and the acquired slices may be even 6 mm thick. Weak spatial resolution blurs the small anatomical details, that would be helpful in locating the estimated activations correctly. However, with technique depended on BOLD contrast smallest details are about $3 \times 3\text{mm}^2$ in size. The solution to this problem is to acquire a high-resolution image with the functional part of the study. This high-resolution image is often called the anatomical image. The estimated activations from functional images can then be superimposed onto this high-resolution image.

The coregistration is a preprocessing step that realigns the functional EPI volumes and the high-resolution anatomical image, in the way that the anatomical structures match. SPM99 has two methods to do the coregistration: *i*) method based on the use of template images of both of the modalities of the images to be coregistered, and *ii*) method based on mutual information. If the functional images have low spatial resolution and/or they do not cover the whole head of the subject, there might be some problems in the method based on template images. In these cases it is recommended to use the method based on mutual information.

### 5.3.1 Coregistration based on templates

The coregistration based on templates involves segmentation of the images into gray and white matter (GM and WM) and cerebrospinal fluid (CSF). Because of the low resolution of the functional images and small difference in intensity between GM and WM, the segmentation may be unsuccessful and the results of the coregistration incorrect. Hence this method is explained here quite roughly and with no details. More detailed derivation is found in [4] and [3].

This method relies on the use of the so-called template images, which are of the same modality as the images that are to be coregistered. SPM99 uses also template images of GM, WM and CSF that are used in segmentation step. All these template images are in same stereotactic space.

The first step is to define a transformation between the images and corresponding templates. The transformation is similar to rigid body transformation (Ch. 5.1) but it also applies zooming with three zooming parameters $x_z$, $y_z$, and $z_z$. These paremeters form a zooming matrix $\mathbf{M_z}$:

$$\mathbf{M_z} = \begin{bmatrix} x_z & 0 & 0 & 0 \\ 0 & y_z & 0 & 0 \\ 0 & 0 & z_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{5.14}$$

The transformation between images and templates is computed as rigid body transformation in section 5.1. The rigid body transformation matrix $\mathbf{M}$ must be reformed to take into account the scaling parameters. Hence, the eq. (5.5) is now:

$$\mathbf{M} = \mathbf{M}_f^{-1}\mathbf{M_t}\mathbf{M}_\theta\mathbf{M}_\phi\mathbf{M}_\omega\mathbf{M_z}\mathbf{M}_g \tag{5.15}$$

The transformation is determined by minimizing the sum of squared differences between the templates and images as in section 5.1.

The next step is to segment the images into grey and white matter and cerebrospinal fluid. A MR image is assumed to consist of number of clusters of distinct tissue types e.g. GM, WM, CSF, scalp, background, etc. The intensities of the voxels belonging into these clusters follow a multivariate normal distribution described by a mean vector, a covariance matrix and the number of voxels belonging to the distribution. The transformations determined in the previous step are used as an initial values in separating the voxels to appropriate cluster. For every cluster the mean value, variance and the number of voxels are computed to fully determine the normal distribution of each cluster of different tissue type.

The template images of GM, WM and CSF have voxel values in the range of 0 to 1 representing the prior probability of a voxel belonging into one of these three possibilities. The template images for scalp and background are estimated from these image. The assignment of a probability for a

voxel is based on determining the probability of it belonging to certain cluster. The calculation is based on Bayesian statistics using the intensity values of template images as *a priori* information.

The steps of segmentation are repeated until adequate convergence or prespecified number of iterations is reached. The result is images of GM and WM of the original images that were to be coregistered.

The third and final step is to coregister the image partitions produced in previous step. For this six parameter rigid body transformation is needed to get the final solution. The starting estimates for these six parameters are the results from the first step of coregistration. This last step is to fine-tune the rough realigning from the first step.

### 5.3.2 Coregistration based on Mutual information

Mutual information measures the statistical dependence between two random variables or the amount of information that one variable contains about the other. In order to register different images suitable well together, this mutual information is maximized. The method requires no *a priori* information or model of the relationship of these images, it only assumes that when the images are perfectly realigned, the images provide the most information about each other. The registration method used in SPM99 is based on the work described in [17, 46] with some changes in interpolation methods.

Let us consider two images, which are to be registered together to be random variables $A$ and $B$ with corresponding intensity values $a$ and $b$. Mutual information $I(A, B)$ of the random variables $A$ and $B$ is defined in terms of entropy in the following way [46]:

$$I\,(A,\,B)\;=\;h(A)+h(B)-h(A,B), \tag{5.16}$$

where $h(A)$ and $h(B)$ are the entropies of $A$ and $B$, respectively, and $h(A, B)$ their joint entropy. These entropies are defined as follows:

$$\begin{array}{rcl} h(A) & = & -\sum_a p_A(a)\log p_A(a) \\ h(A,B) & = & -\sum_{a,b} p_{AB}(a,b)\log p_{AB}(a,b). \end{array} \tag{5.17}$$

Here $p_{AB}(a, b)$ is the joint density function of random variables $A$ and $B$ and $p_A(a)$ and $p_B(b)$ are their marginal density functions. Combining eq. (5.17) and (5.16) the mutual information gets the form:

$$I(A,B)\;=\;\sum_{a,b} p_{AB}(a,b)\log\frac{p_{AB}(a,b)}{p_A(a)p_B(b)}. \tag{5.18}$$

The registration of images $A$ and $B$ depends on a registration parameter $\alpha$. All the density functions in eq. (5.18) depend on $\alpha$. The mutual information criterion is that the images $A$ and $B$ are geometrically aligned when the $I(A, B)$ is maximal. The optimal registration parameter $\hat{\alpha}$ is found from:

$$\hat{\alpha}\;=\;\arg\max_{\alpha} I(A,B). \tag{5.19}$$

This method has been changed slightly in SPM99. The images are smoothed slightly, as is the histogram. The smoothing minimizes the cost function faster. SPM99 uses *entropy correlation coefficient* ECC [8], which can be defined as

$$ECC(A,B)\;=\;\frac{2I(A,B)}{h(A)+h(B)}. \tag{5.20}$$

The ECC is negated as the optimization function that minimizes the cost function in order to maximize the mutual information.

## 5.4 Normalization

Spatial normalization as a preprocessing step is performed for two reasons: firstly, it enables to report the locations of activation according to a well-known, meaningful coordinates within a standard space, and secondly, it enables group comparisons.

In neuroimaging community it is common to report the locations of activations in the stereotactic space of e.g. Talairach & Tournoux [65]. To be able to do this, the subjects brain must be mapped into a standard space i.e. normalized. SPM99 does not, however, use the standard space of Talairach & Tournoux, but a space determined by Montreal Neurological Institute (MNI, `http://www.bic.mni.mcgill.ca/`). The MNI space is defined by using a large group of normal MRI scans, therefore, being more representative of the population whereas the Talairach & Tournoux atlas is determined based on one person's brain only. These two atlases are very similar although there are also some differences. Still, a coordinate defined in one atlas does not differ much from the other.

The second reason why normalization is used, is to be able to perform group analysis. Every brain is slightly different in size, shape and in location of functional areas. In order to find same functional areas from each subject of the study, everyone's brain must be standardized into same space.

Spatial normalization is performed by using a two-step approach. The first step involves determination of the optimum 12-parameter affine registration between the template and the image to be normalized. The second step is to estimate nonlinear deformations defined by a linear combination of three dimensional discrete cosine transform (DCT) basis functions.

### 5.4.1 Determination of affine transformation

For the determination of the optimum affine transformation a maximum *a posteriori* (MAP) approach is used. The plane affine transformation is not always sufficient in moving the homologous regions of different brain close enough together. By using some *a priori* information the iteration gives usually better results.

The MAP estimator is based on the Bayesian rule, which for random variables $\theta$ and $z$ is described as follows:

$$f(\theta|z) \propto f(z|\theta)f(\theta), \tag{5.21}$$

where $f(\theta|z)$ (posterior density) and $f(z|\theta)$ (likelihood) are the conditional density functions of $z$ given $\theta$, and $\theta$ given $z$ respectively, and $f(\theta)$ (the priori) is the density function of $\theta$. In MAP estimation the optimum result is achieved by maximizing the posterior density.

The MAP approach for defining the optimum affine transformation of image registration is derived precisely in [7]. The main results are explained as follows.

In normalization a image $f$ is normalized into a standard space defined by a template image $g$. The affine transformation is similar to transformation described in section 5.3.1 but it also takes into account shears. Three shearing parameters $x_s$, $y_s$ and $z_s$ are added into the parameter vector $\mathbf{p}$:

$$\mathbf{p} = [x_t, y_t, z_t, \phi, \theta, \omega, x_z, y_z, z_z, x_s, y_s, z_s, s]^T \tag{5.22}$$

The shearing parameters form a shearing matrix $M_s$:

$$\mathbf{M_s} = \begin{bmatrix} 1 & x_s & y_s & 0 \\ 0 & 1 & z_s & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{5.23}$$

The final transformation matrix $\mathbf{M}$ is now

$$\mathbf{M} = \mathbf{M}_f^{-1}\mathbf{M_t}\mathbf{M}_\theta\mathbf{M}_\phi\mathbf{M}_\omega\mathbf{M_z}\mathbf{M_s}\mathbf{M_g}. \tag{5.24}$$

For the purpose of image normalization the MAP estimate for the affine transformation parameters $\mathbf{p}$ (eq. 5.22) is the mode of $p(\mathbf{p}|f)$, the conditional probability that $\mathbf{p}$ is observed given that $f$ is true, $p(\mathbf{p})$ represents a known prior probability distribution from which the parameters are drawn, and $p(f|\mathbf{p})$ is the likelihood of obtaining the parameters $\mathbf{p}$ given the data $f$. All the probability functions are assumed to be jointly normal, being, therefore, described by a mean vector and a covariance matrix. The parameters $\mathbf{p}$ are assumed to be drawn from an normal distribution of mean $\mu_\mathbf{p}$ and covariance $\mathbf{C_p}$. The errors of the fitted parameters $\mathbf{p}$ are approximated by a jointly normal distribution with covariance matrix $\mathbf{C}$

The parameters $\mathbf{p}$ are estimated by using *a priori* probability density function of the parameters as a weighted average of the mean $\mu_{\mathbf{p}}$ and $\mathbf{p}$:

$$\hat{\mathbf{p}} = \left(\mathbf{C_p^{-1}} + \mathbf{C^{-1}}\right)^{-1} \left(\mathbf{C_p^{-1}}\mu_{\mathbf{p}} + \mathbf{C^{-1}}\mathbf{p}\right). \tag{5.25}$$

For the MAP solution the estimated covariance matrix is

$$\mathbf{C}_{\hat{\mathbf{p}}} = \left(\mathbf{C_p^{-1}} + \mathbf{C^{-1}}\right)^{-1}.$$

In order to employ the MAP solution, the covariance matrix of the errors of the fitted parameters $\mathbf{C}$ must be computed. If the sampled locations of the image $f$ are independent and each has unit standard deviation, $\mathbf{C} = \left(\mathbf{J_l}^T\mathbf{J_l}\right)^{-1}$, where $\mathbf{J_l}$ is the Jacobian matrix as in section 5.1. Because the standard deviations are unknown, they are assumed to be equal for all the sampled locations. The estimate of the standard deviation is

$$\sigma^2 = \sum_i \left(\mathbf{M}\mathbf{x}_i - \mathbf{s}\mathbf{y}_i\right)^2. \tag{5.26}$$

This gives a covariance matrix $\mathbf{C}$:

$$\mathbf{C} = \frac{\left(\mathbf{J_l}^T\mathbf{J_l}\right)^{-1}\sigma^2}{\nu}, \tag{5.27}$$

where $\nu$ is the number of the sampled locations in the images minus the number of the parameters $\mathbf{p}$.

By combining equations (5.12), (5.25) and (5.27) the MAP Bayesian approach for determining optimum affine transformation parameters is

$$\hat{\mathbf{p}}^{(n)} = \left(\mathbf{C_p^{-1}} + \alpha\right)^{-1} \left(\mathbf{C_p^{-1}}\mu_{\mathbf{p}} + \alpha\hat{\mathbf{p}}^{(n-1)} - \beta\right), \tag{5.28}$$

where $\alpha = \mathbf{J_l}^T\mathbf{J_l}\nu/\sigma^2$, $\beta = \mathbf{J_l}^T f\nu/\sigma^2$.

### 5.4.2 Estimation of nonlinear deformations

The affine registration is followed by estimating nonlinear deformations as described in [6, 3]. To map a brain in one space to a brain in another, some nonlinear distortions are needed to match the brain. These distortions are modeled by a linear combination of basis functions, which in this case are the lowest-frequency components of discrete cosine transform. The normalization also means changing the coordinates of the object image $f$ to match the coordinates of the template image $g$. A transformation from known coordinates $\mathbf{x}$ in image $f$ to unknown coordinates $\mathbf{y}$ in template image $g$ by using a linear combination of basis functions is

$$\mathbf{y}_i = \mathbf{x}_i - \sum_j t_j b_j(\mathbf{x}_i), \tag{5.29}$$

where $t_j$ is the $j$th coefficient describing translation for each three dimensions, and $b_j(\mathbf{x}_i)$ is the $j$th basis function at spatial position $a_i$.

The optimization involves minimizing the sum of squared differences between the object image $f$ and the template image $g$. Now the minimized function is

$$\sum_i (\mathbf{y}_i - w\mathbf{x}_i)^2,$$

where $w$ is an additional scaling parameter. Elements of vector $\mathbf{b_a}$ (in eq. 5.12) contains $\mathbf{y}_i - w\mathbf{x}_i$. In order to compute matrix $\mathbf{J_l}$ the derivatives of the function $\mathbf{y}_i - w\mathbf{x}_i$ must be computed by using the chain rule.

The matrices $\mathbf{J_l}^T\mathbf{J}$ and $\mathbf{J}^T\mathbf{b_a}$ could be computed by using the Gauss-Newton algorithm (Ch. 5.1). This might, however, be time consuming. In [6] also a fast algorithm for performing this is described.

## 5.5 Spatial smoothing

Functional MRI-data contains noise. Some of this noise is independent of each voxel thus forming high spatial frequencies. The actual signal produced by hemodynamic changes is expressed over spatial scale of several millimeters [23]. To improve the signal-to-noise ratio the data is spatially smoothed. This is done by low-pass filtering the data with a Gaussian kernel.

The analysis of fMRI data demands making of statistical inferences about the regionally specific effects and in SPM99 this requires the theory of Gaussian Random Fields (GRF). The fMRI data however does not accomplish the requirements of the GRF theory, which are that the autocorrelation function should be twice differentiable and the spatial correlations should be stationary. By smoothing the images with gaussian kernel the GRF requirements are approximately assured.

The smoothing is performed as a discrete convolution with a gaussian kernel [3]. The amplitude of the gaussian $A$ at $j$ units away from the center is defined by:

$$A_j = \frac{e^{-\frac{j^2}{2s^2}}}{\sqrt{2\pi s^2}}, \tag{5.30}$$

where the parameter $s = \frac{FWHM}{\sqrt{8\ln 2}}$ and the $FWHM$ is the *full width at half maximum* of the gaussian kernel. The recommended FWHM for fMRI data is 2-3 times the voxel size of the functional images. The discrete convolution of function $h$ with $g$ to give the convolved function $t$ is performed as:

$$t_i = \sum_{j=-d}^{d} h_{i-j} g_j. \tag{5.31}$$

The value of $d$ is the kernel length of about six FWHMs. The convolution is performed in all three directions of fMRI volume.

Chapter 6

# Parameter estimation and statistics in SPM99

The purpose of the preprocessing steps is to remove some confounding effects e.g. movement-related effects, and make the data better meet the assumptions of the Gaussian Random Field (used in statistical tests) and General Linear Model. The General Linear Model explains the variation in some random variable in terms of a linear combination of explanatory variables and an error term. In fMRI the number of random variables is huge, because each and every voxel in MRI scan is treated as a separate random variable. The explanatory variables used are the so called basis functions. Linear combination of them is designed in a way that it spans the space of possible fMRI responses revealing voxels with suitable activity. The errors are assumed to be independent and identically distributed with zero mean.

## 6.1 Basis functions in fMRI time series

Basis functions are used to define the time intervals were the activation should occur in the activation time series according to the paradigm. The different stimuli and their responses can be distinguished from each other by defining them all in separate columns in design matrix. For example, if one paradigm consists two different type of stimuli like visual stimuli and movement, they are defined as two different tasks. The timing of these stimuli are defined in the design matrix by placing the appropriate basis functions in correct place in time scale. Hence the columns of the design matrix can perform also as a design of the paradigm denoting the order and timing of different stimuli.

In fMRI-analysis the form of the basis functions depends on the experimental design. If the design is a block design (Ch. 3.4.1), the basis functions may be box-car functions. A typical box-car functionis illustrated in Fig. 6.1. To get this kind of basis function more realistic, it can be convolved with the hemodynamic response function (hrf) as is illustrated in Fig. 6.1. The shape typical to hrf is illustrated in Fig. 3.2. In the case of event-related designs, hrfs are used as basis functions, because a hrf resembles the form of BOLD response. Other functions have also been used as basis functions in event-related desings e.g. shifted impulse response functions (IRFs) [72], trigonometric functions or combination of gamma functions.

## 6.2 Parameter estimation

Functional MRI time series are modeled by using a linear model

$$\mathbf{z} = \mathbf{H}\theta + \mathbf{v}, \tag{6.1}$$

where $\mathbf{z}$ is a vector containing the observations i.e. the fMRI time series, $\mathbf{H}$ is a design matrix formed by the basis functions, $\theta$ is a column vector of parameters adjusting the model to fit the time series as well as possible, and $\mathbf{v}$ is a vector of errors, which are assumed to be independent and identically normally distributed $\mathbf{v} \sim N(0, \sigma^2)$ [22]. The columns of the design matrix represent how the signal is thought to be varying in active areas and the estimated parameters $\hat{\theta}$ defines how well the model included in design matrix actually fits the time-series of each voxel.

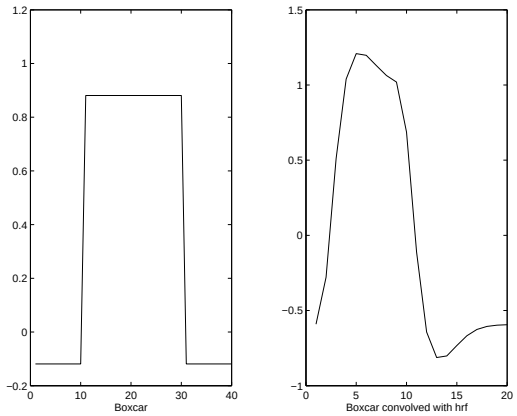



Figure 6.1: The box-car function on the left and the same box-car convolved with the hemodynamic response function (hrf) on the right.

The fMRI time series contain noise from measurement (white noise), and physiological noise from cardiac and respiratory effects aliased to low frequencies. These confounding effects are taken into account by temporally smoothing the data and the desing matrix, with a matrix $\mathbf{K}$ [21], so the eq. (6.1) gets a form:

$$\mathbf{Kz} = \mathbf{KH}\theta + \mathbf{Kv} \quad \Rightarrow \quad \mathbf{z}^* = \mathbf{H}^*\theta + \mathbf{v}^*. \tag{6.2}$$

Parameters $\theta$ are estimated by using the LS-method (see Ch. 4.1 and eq. 4.7):

$$\hat{\theta} = \left(\mathbf{H}^{*T}\mathbf{H}^*\right)^{-1}\mathbf{H}^{*T}\mathbf{z}^*. \tag{6.3}$$

## 6.3 Obtaining the results in fMRI analysis

The aim of fMRI studies is usually to define which areas in brain are active during some stimuli and do these areas vary between different populations (i.e. healthy controls and diseased patients). Statistical methods are used to answer these questions.

Statistical significance of a voxel activating during the stimuli can be calculated many different ways depending on the paradigm and the questions of interest. In SPM99 the test, that is used most often, is a simple t-test, because the results obtained using t-test can be passed on to the second level analysis i.e. the group analysis. It is possible to use also a F-test in SPM99, but the results obtained from F-test are not suitable for the second level analysis.

### 6.3.1 Comparisons with t-test

A t-test is used when one wants to see which cortical areas are more active during one stimuli than another (two sample t-test) or which cortical areas are active during some specific stimuli (one sample t-test).

#### Computing the $t$-value

$T$-value is computed for every voxel using the t-test. In the t-test a contrast of estimated parameters is divided by a square root of the variance estimate. Stimuli to be compared are selected by using a contrast, which is a vector defining the parameters to be compared using statistical tests. The values of the elements of the contrast vector depend on the questions to be answered. If the paradigm consists of two different stimuli and one wants to define the areas that are more active during the first stimuli than during the second one, a contrast with elements 1 and -1 is used. The contrast of the estimated parameters is computed by multiplying the parameter estimate vector

$\hat{\theta}$ by the appropriate contrast vector. Depending on the question of interest and the comparison made, the test is similar to one sample or two sample t-test (Ch. 4.2.1).

For the variance estimate the covariance of $\hat{\theta}$ must be computed. Covariance for a random variable is defined using the expectation $E\{\cdot\}$:

$$C_{\hat{\theta}} = E\left\{(\hat{\theta} - E\{\hat{\theta}\})(\hat{\theta} - E\{\hat{\theta}\})^T\right\}. \tag{6.4}$$

If $\hat{\theta}$ is an unbiased estimate of $\theta$, the expectation of $\hat{\theta}$ is $\theta$. This is proved by defining an estimation error $\tilde{\theta}$ and it's expectation:

$$\tilde{\theta} = \theta - \hat{\theta} = \theta - \left(\mathbf{H}^{*T}\mathbf{H}^*\right)^{-1}\mathbf{H}^{*T}\mathbf{z}^*. \tag{6.5}$$

Inserting $\mathbf{z}^*$ from eq. (6.2) and simplifying we get:

$$\begin{aligned}\tilde{\theta} &= \theta - \left(\mathbf{H}^{*T}\mathbf{H}^*\right)^{-1}\mathbf{H}^{*T}\left(\mathbf{H}^*\theta + \mathbf{v}^*\right) \\ &= -\left(\mathbf{H}^{*T}\mathbf{H}^*\right)^{-1}\mathbf{H}^{*T}\mathbf{v}^*.\end{aligned} \tag{6.6}$$

The expectation of the estimation error is (after substituting $\mathbf{v}^*$ by $\mathbf{K}\mathbf{v}$ and rearranging):

$$E\{\tilde{\theta}\} = -\left(\mathbf{H}^{*T}\mathbf{H}^*\right)^{-1}\mathbf{H}^{*T}\mathbf{K}\underbrace{E\{\mathbf{v}\}}_{=0} = 0. \tag{6.7}$$

Therefore, $\hat{\theta}$ is an unbiased estimator of $\theta$ and $E\{\hat{\theta}\} = \theta$. Hence, the covariance of $\hat{\theta}$ is:

$$\begin{aligned}C_{\hat{\theta}} &= E\left\{(\hat{\theta} - \theta)(\hat{\theta} - \theta)^T\right\} \\ &= E\left\{\hat{\theta}\hat{\theta}^T - \hat{\theta}\theta^T - \theta\hat{\theta}^T + \theta\theta^T\right\}.\end{aligned} \tag{6.8}$$

Now inserting here $\hat{\theta}$ from eq. (6.3) and $\mathbf{z}^*$ from eq. (6.2):

$$\begin{aligned}C_{\hat{\theta}} = E\Big\{&(\mathbf{H}^{*T}\mathbf{H}^*)^{-1}\mathbf{H}^{*T}(\mathbf{H}^*\theta + \mathbf{v}^*)\left[(\mathbf{H}^{*T}\mathbf{H}^*)^{-1}\mathbf{H}^{*T}(\mathbf{H}^*\theta + \mathbf{v}^*)\right]^T \\ &-(\mathbf{H}^{*T}\mathbf{H}^*)^{-1}\mathbf{H}^{*T}(\mathbf{H}^*\theta + \mathbf{v}^*)\theta^T - \theta\left[(\mathbf{H}^{*T}\mathbf{H}^*)^{-1}\mathbf{H}^{*T}(\mathbf{H}^*\theta + \mathbf{v}^*)\right]^T + \theta\theta^T\Big\}.\end{aligned} \tag{6.9}$$

After simplification the previous equation gets a form:

$$\begin{aligned}C_{\hat{\theta}} &= E\left\{(\mathbf{H}^{*T}\mathbf{H}^*)^{-1}\mathbf{H}^{*T}\mathbf{v}^*\mathbf{v}^{*T}\mathbf{H}^*(\mathbf{H}^{*T}\mathbf{H}^*)^{-1}\right\} \\ &= (\mathbf{H}^{*T}\mathbf{H}^*)^{-1}\mathbf{H}^{*T}E\left\{\mathbf{v}^*\mathbf{v}^{*T}\right\}\mathbf{H}^*(\mathbf{H}^{*T}\mathbf{H}^*)^{-1} \\ &= (\mathbf{H}^{*T}\mathbf{H}^*)^{-1}\mathbf{H}^{*T}\mathbf{K}\,E\left\{\mathbf{v}\mathbf{v}^T\right\}\mathbf{K}^T\mathbf{H}^*(\mathbf{H}^{*T}\mathbf{H}^*)^{-1}.\end{aligned} \tag{6.10}$$

The covariance of the error $\mathbf{v}$ is

$$\begin{aligned}C_{\mathbf{v}} &= E\left\{(\mathbf{v} - \eta_{\mathbf{v}})(\mathbf{v} - \eta_{\mathbf{v}})^T\right\} \\ &= E\left\{\mathbf{v}\mathbf{v}^T - \mathbf{v}\eta_{\mathbf{v}}^T - \eta_{\mathbf{v}}\mathbf{v}^T + \eta_{\mathbf{v}}\eta_{\mathbf{v}}^T\right\} \\ &= E\left\{\mathbf{v}\mathbf{v}^T\right\}\end{aligned} \tag{6.11}$$

because $\eta_{\mathbf{v}} = E\{\mathbf{v}\} = 0$. The covariance of $\mathbf{v}$ is $\sigma^2$. Substituting this into eq. (6.10) the covariance of $\hat{\theta}$ gets a form [70]:

$$C_{\hat{\theta}} = \sigma^2(\mathbf{H}^{*T}\mathbf{H}^*)^{-1}\mathbf{H}^{*T}\mathbf{K}\mathbf{K}^T\mathbf{H}^*(\mathbf{H}^{*T}\mathbf{H}^*)^{-1}. \tag{6.12}$$

Normally the variance of the errors is unknown, so it must be estimated and use this estimate in eq. (6.12) instead of $\sigma^2$. Let us examine the residual $\mathbf{r} = \mathbf{z}^* - \hat{\mathbf{z}}^*$, where $\hat{\mathbf{z}}^*$ is the observations computed by using the estimated parameters i.e. $\hat{\mathbf{z}}^* = \mathbf{H}^*\hat{\theta}$:

$$\mathbf{r} = \mathbf{z}^* - \hat{\mathbf{z}}^* = \mathbf{z}^* - \mathbf{H}^*\hat{\theta} = \mathbf{z}^* - \mathbf{H}^*(\mathbf{H}^{*T}\mathbf{H}^*)^{-1}\mathbf{H}^{*T}\mathbf{z}^* = \mathbf{R}\mathbf{z}^*, \tag{6.13}$$

where $\mathbf{R} = \mathbf{R}^T = I - \mathbf{H}^*(\mathbf{H}^{*T}\mathbf{H}^*)^{-1}\mathbf{H}^{*T}$. The residual sum of squares is therefore

$$\begin{aligned}\mathbf{r}^T\mathbf{r} &= (\mathbf{R}\mathbf{z}^*)^T\mathbf{R}\mathbf{z}^* = \mathbf{z}^{*T}\mathbf{R}^T\mathbf{R}\mathbf{z}^* \\ &= (\mathbf{H}^*\theta + \mathbf{v}^*)^T(I - \mathbf{H}^*(\mathbf{H}^{*T}\mathbf{H}^*)^{-1}\mathbf{H}^{*T})(I - \mathbf{H}^*(\mathbf{H}^{*T}\mathbf{H}^*)^{-1}\mathbf{H}^{*T})(\mathbf{H}^*\theta + \mathbf{v}^*) \\ &= \mathbf{v}^{*T}\mathbf{v}^* - \mathbf{v}^{*T}\mathbf{H}^*(\mathbf{H}^{*T}\mathbf{H}^*)^{-1}\mathbf{H}^{*T}\mathbf{v}^* \\ &= \mathbf{v}^{*T}\mathbf{R}\mathbf{v}^*.\end{aligned} \tag{6.14}$$

The expectation of the residual sum of squares is [62]

$$\begin{aligned}E\{\mathbf{r}^T\mathbf{r}\} &= E\{\mathbf{v}^{*T}\mathbf{R}\mathbf{v}^*\} \\ &= \eta_{\mathbf{v}^*}^T\mathbf{R}\eta_{\mathbf{v}^*} + \mathrm{trace}(\mathbf{R}C_{\mathbf{v}^*}).\end{aligned} \tag{6.15}$$

The first term of this is zero, because $\eta_{\mathbf{v}} = 0$. The required covariance of $\mathbf{v}^*$ is derived as follows:

$$\begin{aligned}C_{\mathbf{v}^*} &= E\left\{(\mathbf{v}^* - \eta_{\mathbf{v}^*})(\mathbf{v}^* - \eta_{\mathbf{v}^*})^T\right\} \\ &= E\{\mathbf{v}^*\mathbf{v}^{*T}\} \\ &= \mathbf{K}E\{\mathbf{v}\mathbf{v}^T\}\mathbf{K}^T \\ &= \sigma^2\mathbf{K}\mathbf{K}^T\end{aligned} \tag{6.16}$$

for $E\{\mathbf{v}^*\} = \mathbf{K}E\{\mathbf{v}\} = 0$. By inserting this in the brackets of trace in eq. (6.15) the expectation of the residual sum of squares becomes

$$E\{\mathbf{r}^T\mathbf{r}\} = \sigma^2\mathrm{trace}(\mathbf{R}\mathbf{K}\mathbf{K}^T) \quad \Rightarrow \quad \sigma^2 = \frac{E\{\mathbf{r}^T\mathbf{r}\}}{\mathrm{trace}(\mathbf{R}\mathbf{K}\mathbf{K}^T)}. \tag{6.17}$$

Because $\hat{\sigma}^2$ is an unbiased estimate of $\sigma^2$, $E\{\hat{\sigma}^2\} = \sigma^2$. Therefore, the estimate for $\sigma^2$ is obtained from eq. (6.17) [70]:

$$\hat{\sigma}^2 = \frac{\mathbf{r}^T\mathbf{r}}{\mathrm{trace}(\mathbf{R}\mathbf{K}\mathbf{K}^T)}. \tag{6.18}$$

The variance estimate needed for the t-test is derived in eq. (6.12). Here $\sigma^2$ must be replaced by $\hat{\sigma}^2$ from eq. (6.18). The formula of two sample t-test (eq. 4.9) contains the variances of the random variables to be compared only. Therefore, the variance estimate in eq. (6.12) must also be modified in a way that it only contains the variances of the parameters to be compared. In SPM99 this is done by multiplying the eq. (6.12) with the contrast vector and its transpose. Hence, the formula for t-test is [70]:

$$T = \frac{c\,\hat{\theta}}{\sqrt{c\,\hat{\sigma}^2(\mathbf{H}^{*T}\mathbf{H}^*)^{-1}\mathbf{H}^{*T}\mathbf{K}\mathbf{K}^T\mathbf{H}^*(\mathbf{H}^{*T}\mathbf{H}^*)^{-1}c^T}}. \tag{6.19}$$

Computing the effective degrees of freedom

The effective degrees of freedom is computed using the Satterthwaite's approximation [58, 54]:

$$\nu = \frac{2E\{\hat{\sigma}^2\}^2}{\mathrm{var}\{\hat{\sigma}^2\}}, \tag{6.20}$$

where $E\{\hat{\sigma}^2\} = \sigma^2$. The variance of $\hat{\sigma}^2$ is

$$\begin{aligned}\mathrm{var}\{\hat{\sigma}^2\} &= E\{(\hat{\sigma}^2 - \eta_{\hat{\sigma}^2})^2\} \\ &= E\{\left(\hat{\sigma}^2\right)^2\} - \sigma^4\end{aligned} \tag{6.21}$$

Inserting here $\hat{\sigma}^2$ from eq. (6.18) we get

$$\mathrm{var}\{\hat{\sigma}^2\} = \frac{E\{(\mathbf{r}^T\mathbf{r})^2\}}{\mathrm{trace}(\mathbf{R}\mathbf{K}\mathbf{K}^T)^2} - \sigma^4. \tag{6.22}$$

By using eq. (6.14) the term $(\mathbf{r}^T\mathbf{r})^2$ gets a form

$$(\mathbf{r}^T\mathbf{r})^2 \;=\; (\mathbf{v}^{*T}\mathbf{R}\mathbf{v}^{*T})^2 \;=\; (\mathbf{v}^T\mathbf{K}^T\mathbf{R}\mathbf{K}\mathbf{v})^2. \tag{6.23}$$

The expectation of the eq. (6.23) is calculated using the following equation [61]: Let a matrix $\mathbf{Y}$ be normally distributed with variance $\sigma_Y$. If $\mathbf{A}$ is any symmetric matrix, then

$$E\{(\mathbf{Y}^T\mathbf{A}\mathbf{Y})^2\} \;=\; \sigma_Y^2\left(\mathrm{trace}(\mathbf{A})^2 + 2\mathrm{trace}(\mathbf{A}^2)\right). \tag{6.24}$$

Therefore, the expectation of $(\mathbf{r}^T\mathbf{r})^2$ is

$$\begin{aligned} E\{(\mathbf{r}^T\mathbf{r})^2\} &= \sigma^4\left(\mathrm{trace}(\mathbf{K}^T\mathbf{R}\mathbf{K})^2 + 2\mathrm{trace}(\mathbf{K}^T\mathbf{R}\mathbf{K}\mathbf{K}^T\mathbf{R}\mathbf{K})\right) \\ &= \sigma^4\left(\mathrm{trace}(\mathbf{R}\mathbf{K}\mathbf{K}^T)^2 + 2\mathrm{trace}(\mathbf{R}\mathbf{K}\mathbf{K}^T\mathbf{R}\mathbf{K}\mathbf{K}^T)\right). \end{aligned} \tag{6.25}$$

Inserting this into eq. (6.22) we get the variance of $\hat{\sigma}^2$:

$$\mathrm{var}\{\hat{\sigma}^2\} \;=\; \frac{2\sigma^4\mathrm{trace}\left(\mathbf{R}\mathbf{K}\mathbf{K}^T\mathbf{R}\mathbf{K}\mathbf{K}^T\right)}{\mathrm{trace}\left(\mathbf{R}\mathbf{K}\mathbf{K}^T\right)^2}. \tag{6.26}$$

By adding this and the expectation into the eq. (6.20) the equation for the effective degrees of freedom gets its final form [70]

$$\nu \;=\; \frac{\mathrm{trace}\left(\mathbf{R}\mathbf{K}\mathbf{K}^T\right)^2}{\mathrm{trace}\left(\mathbf{R}\mathbf{K}\mathbf{K}^T\mathbf{R}\mathbf{K}\mathbf{K}^T\right)}. \tag{6.27}$$

### 6.3.2 Determination of the significance level

The statistical test is performed to every voxel in the brain. Therefore, every voxel has its own statistical value. Replacing the intensity values that form the magnetic resonance images with these statistical values, a statistical map is formed. This map is created to localize the areas activated by the used stimuli.

To determine which voxels have been active a critical threshold is chosen. A threshold is a limit of statistical values: for a voxel to be concerned as active, its statistical value must exceed the threshold. If it does not, it is concerned as not being active. The limit is chosen in a way that the number of false positives would be as low as possible. At the same time, however, the limit cannot be so strict that it rejects the active voxels. The threshold is chosen by using the $p$-values. Usually if no multiple comparison corrections is used, the threshold is $p < 0.001$, which means that the false positives rate the test performer is prepared to accept is 0.1 %. This rate of false positives one is prepared to accept is usually denoted by $\alpha$. By using the $p$-value and the degrees of freedom the corresponding threshold value is computed and compared to the values in the statistical map.

When multiple independent statistical comparisons are performed simultaneously (as it is the case in fMRI studies), the chosen significance level is not appropriate for all comparisons although it might be appropriate for each individual comparison. In typical statistical map there may be even 200 000 statistical values. When the number of independent tests is this high, with a standard statistical threshold for the individual statistical values, there are always values that appear to be significant in rejecting the null hypothesis even though the null hypothesis is everywhere true. This is called the multiple comparison problem. There are several corrections for this problem and by examining the SPM99 source code (files *spm_P.m and spm_P_Bonf.m*) became evident that SPM99 uses two of them: a Bonferroni correction or a Gaussian Random Field correction. SPM99 performs both of these correction methods and chooses the method that gives $t$-value threshold. This is called the Family Wise Error (FWE) -correction.

### Bonferroni correction

Bonferroni correction was created by Carlo Emilio Bonferroni in 1930's. It is the simplest and most conservative correction for multiple correction problem. The method sets the chosen alpha value for the entire set of comparisons equal to $\alpha$, which means setting the alpha value for each individual comparison equal to $\alpha/n$ when there are $n$ individual comparisons. So for each individual comparison the probability of the statistical value (denoted as $t_i$ in the following equation) being significant under the assumption of the null hypothesis $H_0$ is [51]

$$P(t_i \text{ passes} \mid H_0) \leq \frac{\alpha}{n}. \tag{6.28}$$

Hence, for the entire group of $n$ comparisons the probability of some statistical values being significant under the null hypothesis $H_0$ is

$$P(\text{some } t_i \text{ passes} \mid H_0) \leq \alpha. \tag{6.29}$$

If the alpha value is 0.05 (as it often is) and there are 200 000 voxels, the Bonferroni corrected p-value is 0.05/200 000. By using such a threshold one can be sure that all the remaining significant statistical values are very unlikely to be false positives.

The problem of the Bonferroni correction is that sometimes it is too strict and many active voxels are not detected as active after the multiple comparison correction. Therefore, other multiple comparison correction methods are developed e.g. the Gaussian Random Field theory (see the following section) or the False Discovery Rate (FDR) correction method [28].

### Gaussian Random Field correction

Gaussian Random Field (GRF) correction is based on the utilization of the expected Euler Characteristics (EC) of the voxels where the statistical value exceeds the fixed threshold $t$. The EC is basically the number of clusters (a bundle of voxels, which all have their statistical value above the threshold) in certain threshold, although it depends also on the shape of the concerned volume $V$. For high thresholds near the global maximum $M$ of the statistical field, the EC counts one if $M \geq t$ and zero otherwise, so the EC is a good approximation of the probability of observing one or more clusters above that threshold $t$. The formulas for the EC densities for different thresholds and statistics are given in [71].

The GRF correction also uses a quantity called a *resel count* [71]. It is a unitless quantity which depends only on certain $d$-dimensional features of the concerned volume $V$. For defining the resel counts, the $V$ must be transformed into resel space by dividing the coordinates of the voxels by the FWHMs of the smoothness of the image in the corresponding directions. The resel counts for a three dimensional volume $V$ are *i*) the volume of the $V$ in resel space, *ii*) the resel surface area, which is the half surface area of $V$, *iii*) the resel diameter, which is, if $V$ is convex, twice the average caliper diameter of $V$, and *iv*) the EC of $V$. Approximations of these for voxel data are given in [71].

For the GRF correction the statistical maps are concerned as a four dimensional statistical field, where time is the fourth dimension. The activation peak is supposed to occur somewhere in the brain at unknown time in an interval of length $T$. The time interval needs also to be in time resels, so the time interval $T$ must be divided by the FWHM of the hemodynamic response function $\omega$ i.e. $\tau = T/\omega$. The GRF-corrected p-value is then [71]

$$P(M \geq t) = \sum_{d=0}^{3} R_d(V)[\tau \rho_{d+1}(t) + \rho_d(t)], \tag{6.30}$$

where $R_d(V)$ is the resel count (depending on the volume $V$) and $\rho_d(t)$ the EC density depending on the threshold $t$.

Chapter 7

# Verification of the mathematics with simulations

To verify the mathematics in this thesis concerning SPM99 we performed two simulations: one was an event-related paradigm and the other was a block paradigm. In the simulations real fMRI-data was used with artificial activations inserted afterward in certain voxels in the cortex of the subject. The artificial activations enabled the thorough examination of the correctness of the obtained results. The time-series of these voxels were analyzed by using the equations presented in the previous chapters. The whole dataset was also analyzed with SPM99.

## 7.1 Data acquisition and preprocessing

For the simulations a young healthy right-handed volunteer was studied. The imaging was performed in the Department of Clinical Radiology in Kuopio university hospital[1] with a Siemens Magnetom Vision 1.5 T MRI scanner. Both $T_1$-weighted anatomical images and approximately 700 $T_2^*$-weighted gradient-echo EPI images were acquired. During the imaging of the EPI-series the volunteer had an instruction to do nothing in order to acquire data as "clean" as possible with no intentional actual activations. All images were acquired using $TR$ of 2.5 s. The $T_1$-weighted volume comprised of 180 1 mm thick sagittal slices with in-plane resolution of $1 \times 1$ mm covering the whole head. The EPI images comprised of 16 5 mm thick slices of orientation -10° from transaxial slices to coronal slices. The in-plane resolution within these slices was $4 \times 4$ mm. Between these slices was a 1 mm gap to prevent the artifact caused by the cross-talking between adjacent slices.

The data was preprocessed by using the preprocessing method recommended for the data that is to be analyzed with SPM99. The preprocessing involved correction of movements by using the *Realign*-option (see Ch. 5.1). The differences in acquisition times between slices in one whole head realigned EPI was then corrected with the *Slice Timing*-option (see Ch. 5.2). The EPIs and the anatomical image were then coregistered in order to get the anatomy match in both image modalities. The coregistration was performed by using the method based on templates (see Ch. 5.3.1). The EPIs and the anatomical image were then normalized into MNI co-ordinates (see Ch. 5.4), although this step was not obligatory concerning there was only one subject. This step was however performed, because in actual studies there are usually more than one subject and the normalization is essential. The final preprocessing step was the smoothing (see Ch. 5.5). The EPIs were smoothed by using a gaussian kernel with FWHM of 8 mm. The analysis was performed by using both the unsmoothed and smoothed images.

The artificial activations were added to the normalized but unsmoothed data. The normalized data was used in order to be able to know precisely which voxels should be active and which should not. After the activations had been added the data was smoothed.

## 7.2 Simulation with event-related activations

To simulate event-related activations two cubes of $3 \times 3 \times 3$ voxels were chosen, one cube from postcentral gyrus and the other cube from medial frontal gyrus. The locations of the cubes are

[1] Homepage: `www.kuh.fi`

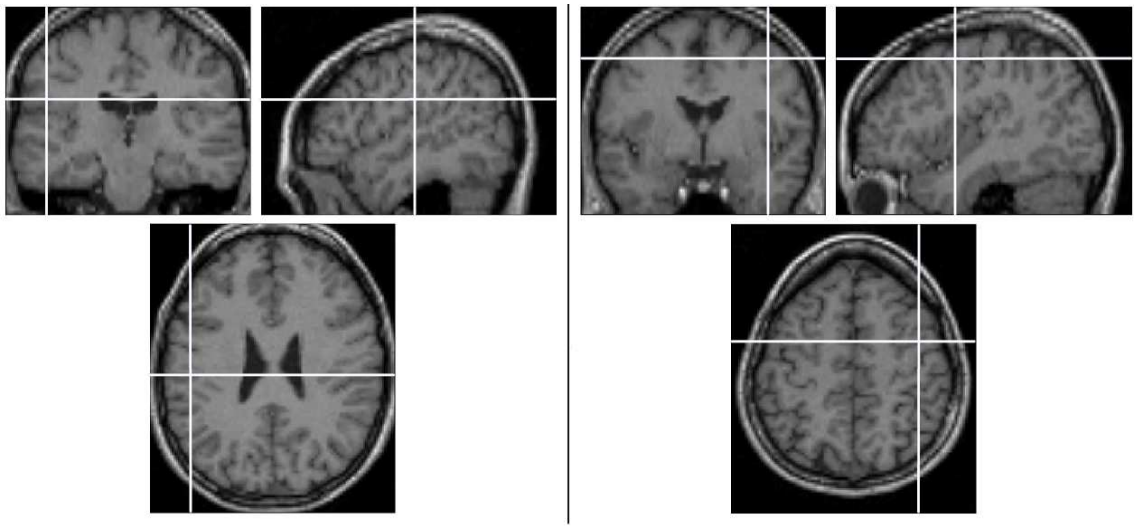


Figure 7.1: The locations of the artificial activations in event-related simulation shown in coronal (up left), sagittal (up right) and transversal (down) view. The postcentral gyrus is on the left and the medial frontal gyrus on the right.

shown in Fig. 7.1. All these voxels' time-series were manipulated by adding there some artificial activation. The artificial BOLD responses were created by multiplying the voxels' time series with an gamma-variate function which represents an ideal hemodynamic response:

$$I(t) = A\,t^{8.60}e^{-t/0.547} \tag{7.1}$$

The parameters were taken from [16]. They were empirically determined from a one second duration visual stimulation. This same model has also been used for simulations in [9]. For this simulation the constant $A$ was one and $t$ was a vector from 1 to 20. The gamma-variate function used in these simulations is illustrated in Fig. 7.2.

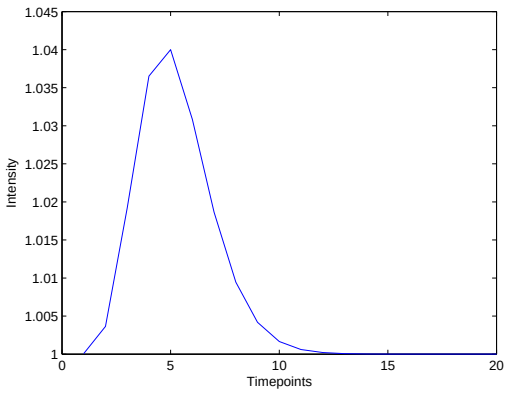



Figure 7.2: The gamma-variate function (see eq. 7.1) used in creating the artificial activations for the event-related case.

For the event-related simulation only a period of 51 scans were involved in the analysis. The paradigm for event-related simulations was not similar for both of the cubes. For the cube in postcentral gyrus there were one stimulus at the very beginning, the following stimuli were placed in $10^{th}$ scan, $22^{nd}$, $30^{th}$ and $44^{th}$ scan. Let this paradigm be called *task 1* later on in this thesis.

For the other cube in medial frontal gyrus the paradigm (*task 2*) was as the following: one stimulus in $5^{th}$, $13^{th}$, $20^{th}$, $24^{th}$ and $36^{th}$ scan. The rest of the time series was left unchanged. The intensity of the simulated BOLD response was not the same in both of the cubes. In postcentral gyrus the peak value was three per cent higher than the original value of the time series and in medial frontal gyrus it was five per cent higher. The mean of unmanipulated time series versus the mean of simulated activation in postcentral gyrus and in medial frontal gyrus are shown in Fig. 7.3. The time-series of all the 27 voxels in both areas before and after the addition of the activation are shown in Fig. 7.4. As it can be seen, the time series of the postcentral gyrus has more activity even before adding any artificial activation whereas the time series in the medial frontal gyrus is smoother. This may arise from the fact that the postcentral gyrus is a part of the somatosensor cortex which handles different senses of touch and the medial frontal gyrus is a part of premotor cortex which masterminds the movements. Because the subject was instructed to do nothing while she laid in the scanner tube, there are no significant activity on the premotor cortex, whilst it is not possible to turn off the sense of touch. Therefore, there are some small activation in the postcentral cortex.

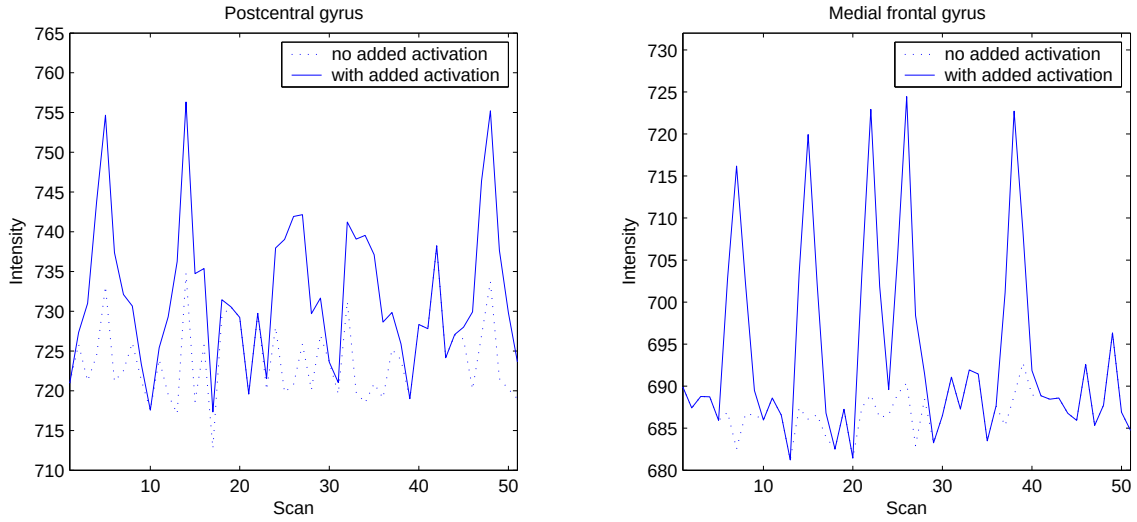



Figure 7.3: The simulated activation on event-related simulation. This figure illustrates the mean of the voxels before and after the insertion of artificial activation for both chosen clusters. The peak value of the artificial BOLD response was +3% in the postcentral gyrus (on the left) and +5% in the medial frontal gyrus (on the right).

The first step of the analysis is the definition of the basis functions for the design matrix. For this simulation the hemodynamic response function (hrf) created by SPM99 and its first time derivative were used as basis functions. These are illustrated in Fig. 7.5. The complete design matrix, illustrated in Fig. 7.6, was created with SPM99 and it was used in both analysis, with SPM99 and with MatLab. The first column of this design matrix represents the imaginary stimulation that activates the postcentral gyrus (task 1) and the second column is its first time derivative. The third and fourth columns are the hrf and its derivative for the imaginary stimulation that activates the medial frontal gyrus (task 2), respectively. The last column is a constant that models the baseline of the time series.

SPM99 analysis includes filtering the data in time space. There is a possibility to both high- and low-pass filter the data but for these analysis only the high-pass filtering was used. The filtering is performed by including the filter matrix $\mathbf{K}$ into the least squares estimation, and into the equations of $t$-value (eq. 6.19) and the degrees of freedom (eq. 6.27). The matrix $\mathbf{K}$ needs to be chosen to suite the time series. Actually the matrix $\mathbf{K}$ is interrelated with the autocorrelation matrix for the time series such that in the equation (eq. 6.16) of the covariance of the smoothed

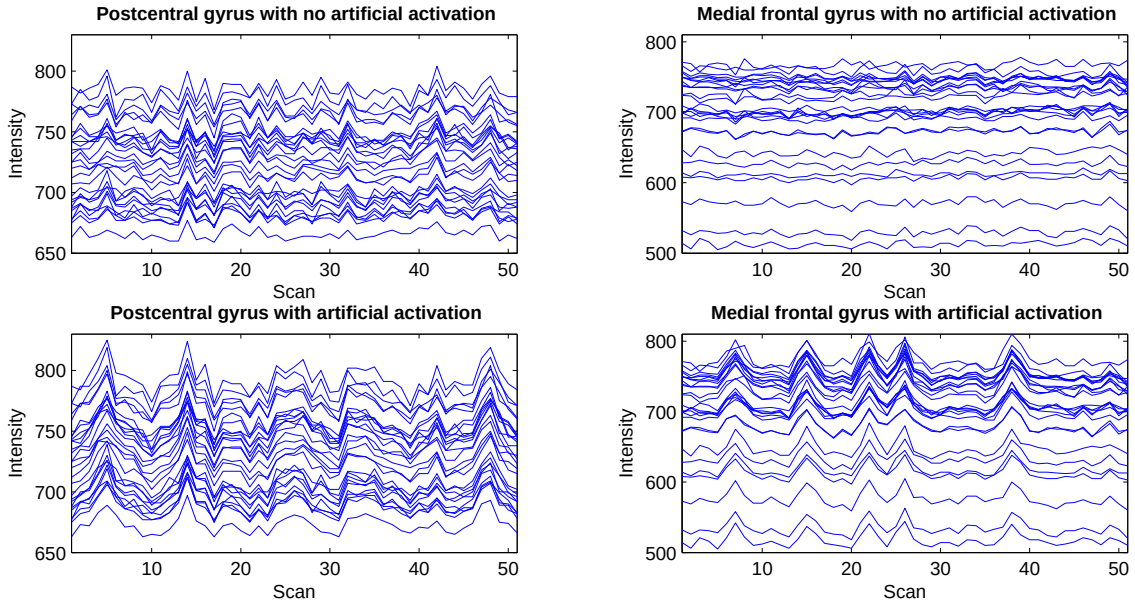



Figure 7.4: Time series of all the 27 voxels in both areas with and without artificial activations.

error term $C_{\mathbf{v}^*}$ the term $\mathbf{K}\mathbf{K}^T$ is the autocorrelation matrix for the time series i.e. $\mathbf{V} = \mathbf{K}\mathbf{K}^T$. The matrix $\mathbf{K}$ should be square and nonsingular [61]. In practice the autocorrelation matrix $\mathbf{V}$ needs to be estimated somehow in order to be able to determine the matrix $\mathbf{K}$.

In SPM99 the estimation of the autocorrelation matrix and the matrix $\mathbf{K}$ is solved by determining a filtering matrix $\mathbf{S}$ [26], which is used to filter the data and the design matrix i.e. to produce the matrices $\mathbf{z}^*$ and $\mathbf{H}^*$ in eq. (6.2). If only high-pass filter is used, these filtered matrices are computed with equation [26]:

$$\mathbf{a}^* = \mathbf{a} - \mathbf{S}(\mathbf{S}^T\mathbf{a}). \tag{7.2}$$

Here $\mathbf{a}$ is the matrix to be filtered and $\mathbf{S}$ is the filtering matrix compounded of a discrete cosine transform set for time points $t = 1, \ldots, N$, where $N$ is determined by the size of the data matrix. The equation for the filtering matrix $S$ is [5]:

$$S(t) = \sqrt{\frac{2}{N}} \cos\left(r\pi\frac{t}{N}\right). \tag{7.3}$$

The integer index $r = 1, \ldots, R$, where $R$ is determined as

$$R = \frac{2N(T_{\text{scan}})}{d_{\text{cut}} + 1}, \tag{7.4}$$

where $T_{\text{scan}}$ is the scan time of the EPI sequence and $d_{\text{cut}}$ defines the high-pass filter cutoff frequency as $f_{\text{cut}} = 1/d_{\text{cut}}$ Hz. For event-related designs the default $d_{\text{cut}}$ is twice the longest interval between two appearances of the most frequently occurring event. In this simulation the $d_{\text{cut}} = 2 \cdot TR \cdot 14 = 70$. Therefore, the high-pass filter cutoff was 1/70 Hz.

For low-pass filtering it is possible to use either hemodynamic response function (hrf) as the filter or gaussian filter. It is also possible to use auto-regressive model for low-pass filtering. If also low-pass filter is used, the high-pass filtered data matrix is additionally multiplied by the low-pass filtering matrix. More information about filtering in fMRI studies is found in [26].

The autocorrelation matrix $\mathbf{V}$ is estimated by using this filtering matrix $\mathbf{S}$ as follows:

$$\mathbf{V} = I - 2\mathbf{S}\mathbf{S}^T + \mathbf{S}\mathbf{S}^T\mathbf{S}\mathbf{S}^T, \tag{7.5}$$

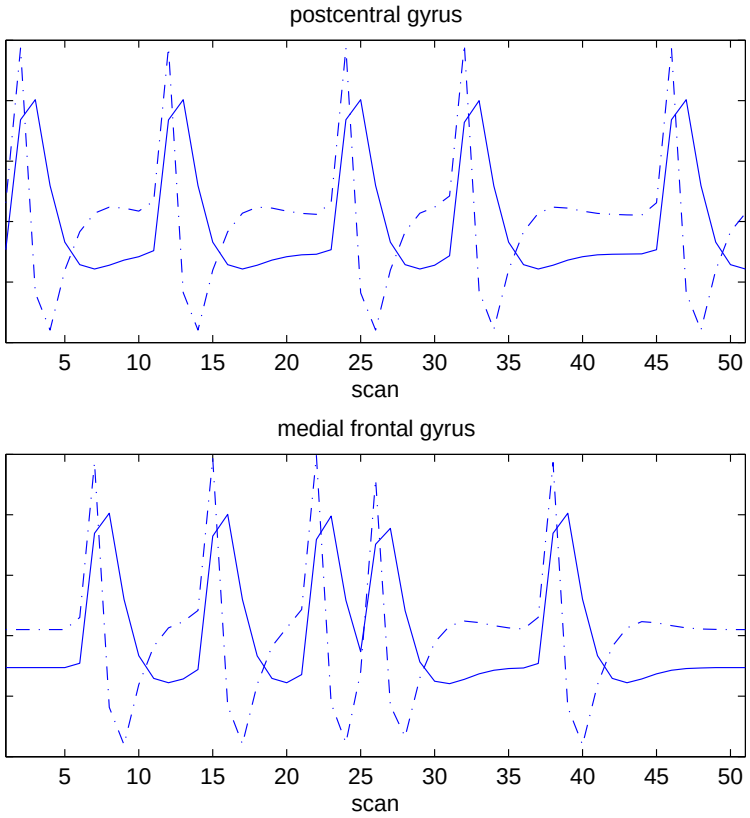



Figure 7.5: The basis functions used in event-related simulation for task 1 (upper) and task 2 (lower). These are actually the first four columns of the design matrix used in analysis. The solid line represents the hrf and the dashdot line hrf's first time derivative.

where $I$ is an identity matrix. This estimated autocorrelation matrix $\mathbf{V}$ is used in computing the estimator of the $\sigma^2$ (eq. 6.18), the $t$-value (eq. 6.19) and the effective degrees of freedom (eq. 6.27) instead of the multiplication $\mathbf{KK}^T$.

The final preprocessing step, spatial smoothing with gaussian kernel, is used to better the SNR. The default FWHM of the gaussian kernel in SPM99 is two or three times the original size of the voxel. In these stimulations the FWHM was 8 mm. The analysis with SPM99 was carried out both with unsmoothed and smoothed data. This was done because the smoothing affected the artificial activations and literally smoothed them away. In Fig. 7.7 is illustrated the time series of the cubes in postcentral gyrus and medial frontal gyrus after the smoothing. As can be seen, there is not much remaining. Hence, it was not a surprise that SPM99 analysis did not locate these areas nor anything else as active when the smoothed data was analyzed not even with threshold $p \leq 0.001$ uncorrected. Obviously the spatial filtering kernel was too wide for the activations of this size. The analysis was, however, performed by using the default values of the SPM99 analysis and, therefore, the FWHM of the spatial filtering kernel was 8 mm.

The same analysis was then performed with unsmoothed data. With the same threshold as with smoothed data SPM99 found some active voxels. In Fig. 7.8 the results of both tasks (1 & 2) with the same threshold $p \leq 0.001$ uncorrected are illustrated. The activations of task 1 is illustrated with red and the activation of task 2 with green. As can be seen, SPM99 categorized not only the cubes of voxels in postcentral gyrus and medial frontal gyrus but also some random voxels as active. The threshold can also be adjusted by defining how many voxels must pass the threshold in one cluster in order to be shown as active. For the Fig. 7.8 this was not restricted so even clusters of one active voxel are shown. The results section in SPM99 tables the $t$-values and

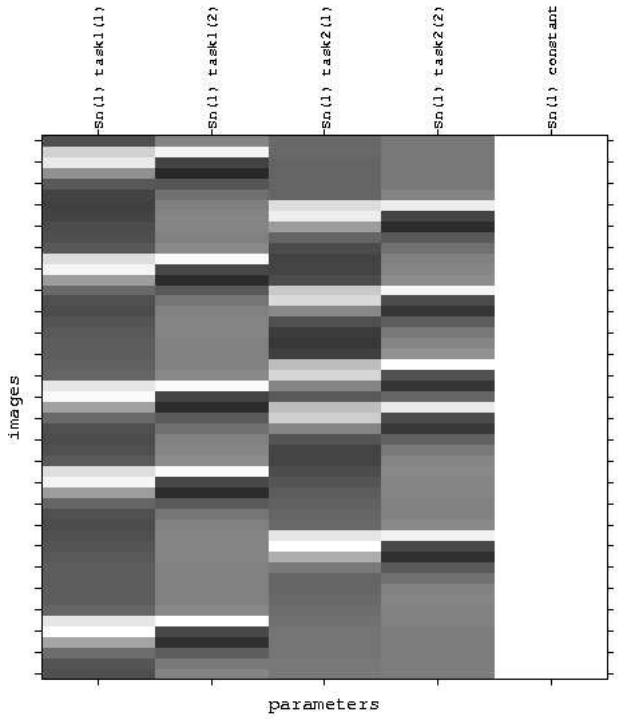



Figure 7.6: The design matrix for the simulation of event-related responses. The columns represents the basis functions, hemodynamic response function and its first time derivative, for both tasks and a constant.

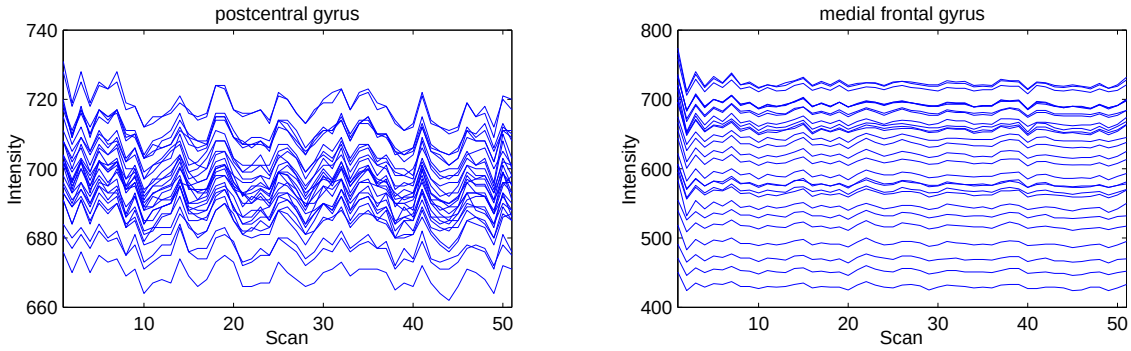



Figure 7.7: The time series of the event-related simulation after smoothing with gaussian kernel (FWHM of 8 mm).

some other statistical parameters for the active clusters. These tables for both tasks are illustrated in Fig. 7.9. In these tables the column $K_B$ shows the number of voxels in the particular cluster. As it can be seen for task 1, the cube in postcentral gyrus is not entirely detected as active, only 22 voxels of 27 are found. For task 2 all 27 voxels in the cube in medial frontal gyrus are found. This is a consequence of the better signal to noise ratio in the original time series of the voxels in medial frontal gyrus. The columns on *voxel level* show the $t$- and $Z$-value and the uncorrected, FWE-corrected (see Ch. 6.3.2) and FDR-corrected (this correction method is not an original SPM99 correction method, for detailed description see [28]) $p$-values for the voxel that has the maximum $t$-value in each cluster. We call this voxel *the maximum of the cluster*. On the far right in the

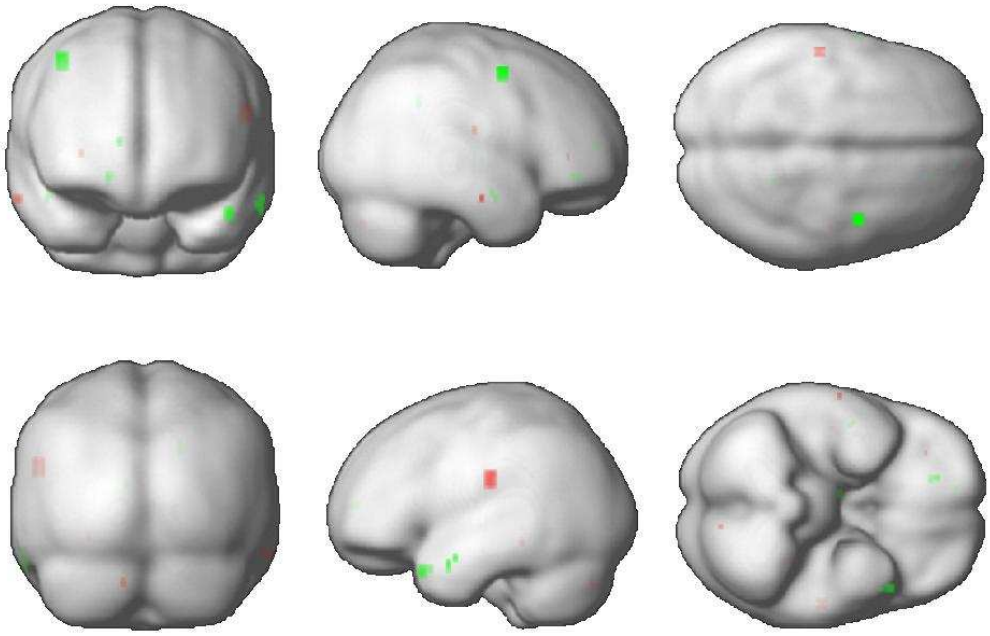


Figure 7.8: The results of SPM99 analysis. The red corresponds to the task 1 (artificial activations located in the postcentral gyrus) and the green task 2 (artificial activations located in the medial frontal gyrus).

table of the results (Fig. 7.9) there are also the co-ordinates for the maximum of the cluster. If the active cluster is very large i.e. it contains a large number of voxels, there might also be information of the voxels of local maximum of the cluster, not only the global maximum.

The voxels that SPM99 did not detect as active for task 1 had a $t$-value less than 3.29 (see Table 7.1 for $t$-values for every voxel in both cubes (postcentral gyrus and medial frontal gyrus)). With a threshold ($p \leq 0.05$ corrected, which corresponds to $t$-value 5.71) the voxels in medial frontal gyrus were still defined as active (in task 2) but the voxels in postcentral gyrus did not survive the threshold (in task 1).

For the Matlab analysis the data and the design matrix had to be filtered before the computation by using the eq. (7.2). The filtering was done with the same filtering matrix that SPM99 uses in its calculations. The filter consisted only of high-pass filtering with cutoff 1/70 Hz. The analysis was carried out for the voxels belonging to the prespecified cubes, not other voxels. The fitting of the basis functions is illustrated for randomly chosen voxels in both areas in Fig. 7.10. The fitting is plotted with dotted line and the original, although filtered, data with solid line. It can be seen that the fitting of the basis functions in the voxel in medial frontal gyrus is more successful than in the voxel in postcentral gyrus. The $t$-values for the voxels in postcentral gyrus were computed using eq. (6.19). The $t$-values for the voxels were identical to those computed with SPM99. The computed values are presented in table 7.1 as well as the coordinates of the voxels. The effective degrees of freedom for these were 43 (from eq. 6.27).

**Statistics:** *volume summary (p-values corrected for entire volume)*

| set-level | | cluster-level | | | voxel-level | | | | | x,y,z {mm} |
|---|---|---|---|---|---|---|---|---|---|---|
| p | c | $p_{corrected}$ | $\hat{k}_E$ | $p_{uncorrected}$ | $p_{FWE-corr}$ | $p_{FDR-corr}$ | T | ($Z_=$) | $p_{uncorrected}$ | |
| 1.000 | 6 | 0.430 | 22 | 0.018 | 0.986 | 0.977 | 4.19 | ( 3.81) | 0.000 | -52 -24 27 |
| | | 1.000 | 2 | 0.445 | 1.000 | 0.977 | 3.78 | ( 3.50) | 0.000 | 64 -12 -21 |
| | | 1.000 | 1 | 0.600 | 1.000 | 0.977 | 3.35 | ( 3.14) | 0.001 | -28 -40 -12 |
| | | 1.000 | 1 | 0.600 | 1.000 | 0.977 | 3.30 | ( 3.10) | 0.001 | 46 -16 18 |
| | | 1.000 | 1 | 0.600 | 1.000 | 0.977 | 3.30 | ( 3.10) | 0.001 | -10 -80 -36 |
| | | 1.000 | 1 | 0.600 | 1.000 | 0.977 | 3.30 | ( 3.09) | 0.001 | 32 38 3 |

**Statistics:** *volume summary (p-values corrected for entire volume)*

| set-level | | cluster-level | | | voxel-level | | | | | x,y,z {mm} |
|---|---|---|---|---|---|---|---|---|---|---|
| p | c | $p_{corrected}$ | $\hat{k}_E$ | $p_{uncorrected}$ | $p_{FWE-corr}$ | $p_{FDR-corr}$ | T | ($Z_=$) | $p_{uncorrected}$ | |
| 1.000 | 11 | 0.269 | 27 | 0.010 | 0.000 | 0.000 | 9.16 | ( 6.78) | 0.000 | 40 2 54 |
| | | 1.000 | 2 | 0.445 | 1.000 | 0.673 | 3.87 | ( 3.57) | 0.000 | -62 -2 -21 |
| | | 0.999 | 5 | 0.225 | 1.000 | 0.851 | 3.79 | ( 3.50) | 0.000 | 18 42 -9 |
| | | 1.000 | 2 | 0.445 | 1.000 | 0.999 | 3.66 | ( 3.39) | 0.000 | 8 -12 0 |
| | | 0.945 | 10 | 0.093 | 1.000 | 0.999 | 3.64 | ( 3.38) | 0.000 | -46 16 -30 |
| | | 1.000 | 3 | 0.346 | 1.000 | 0.999 | 3.61 | ( 3.36) | 0.000 | -62 2 -24 |
| | | 1.000 | 2 | 0.445 | 1.000 | 0.999 | 3.50 | ( 3.27) | 0.001 | 20 -48 36 |
| | | 1.000 | 1 | 0.600 | 1.000 | 0.999 | 3.50 | ( 3.26) | 0.001 | 50 -4 -21 |
| | | 1.000 | 2 | 0.445 | 1.000 | 0.999 | 3.49 | ( 3.26) | 0.001 | 12 54 9 |
| | | 1.000 | 1 | 0.600 | 1.000 | 0.999 | 3.49 | ( 3.26) | 0.001 | 48 -6 -18 |
| | | 1.000 | 1 | 0.600 | 1.000 | 0.999 | 3.37 | ( 3.16) | 0.001 | -8 -48 15 |

Figure 7.9: The statistical tables from SPM99. The upper table is for task 1 and the lower for task 2.

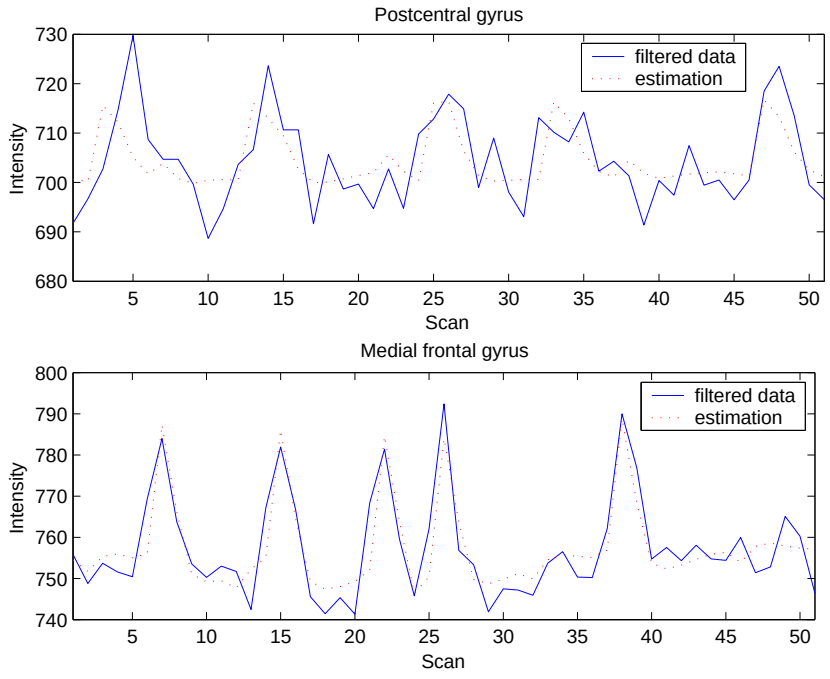



Figure 7.10: The fitting of the basis functions in a randomly selected voxel in postcentral gyrus and in medial frontal gyrus. The solid line is the filtered time series of the voxel and the dotted line is the fitting of the basis functions to the time series.

TABLE 7.1: The $t$-values for the voxels in postcentral gyrus and in medial frontal gyrus in event-related simulation for both task 1 and task 2.

| Postcentral gyrus | | | Medial frontal gyrus | | |
|---|---|---|---|---|---|
| voxel (x,y,z) | $t$-value (task 1) | $t$-value (task 2) | voxel (x,y,z) | $t$-value (task1) | $t$-value (task 2) |
| (-56, -24, 21) | 3.3608 | 0.5257 | (40, -2, 48) | 0.2858 | 8.4330 |
| (-56, -24, 24) | 3.4934 | 0.2107 | (40, -2, 51) | 0.1215 | 8.6736 |
| (-56, -24, 27) | 3.8990 | 0.3395 | (40, -2, 54) | -0.1508 | 7.5921 |
| (-56, -22, 21) | 3.3389 | 0.2981 | (40, 0, 48) | -0.0666 | 7.6825 |
| (-56, -22, 24) | 4.0456 | -0.1525 | (40, 0, 51) | -0.0326 | 8.9117 |
| (-56, -22, 27) | 3.6968 | -0.1999 | (40, 0, 54) | -0.1452 | 8.3495 |
| (-56, -20, 21) | 3.4495 | 0.3764 | (40, 2, 48) | -0.1911 | 6.2814 |
| (-56, -20, 24) | 3.8837 | 0.1385 | (40, 2, 51) | 0.3165 | 8.2772 |
| (-56, -20, 27) | 4.1329 | 0.2387 | (40, 2, 54) | 0.8608 | 9.1553 |
| (-54, -24, 21) | 3.2152 | 0.1158 | (42, -2, 48) | 0.7272 | 5.8691 |
| (-54, -24, 24) | 3.2415 | -0.1625 | (42, -2, 51) | 0.3659 | 7.9320 |
| (-54, -24, 27) | 3.5596 | -0.1133 | (42, -2, 54) | 0.0270 | 8.2981 |
| (-54, -22, 21) | 3.4520 | -0.3108 | (42, 0, 48) | 0.5236 | 5.8346 |
| (-54, -22, 24) | 3.4959 | -0.4135 | (42, 0, 51) | 0.2487 | 8.1393 |
| (-54, -22, 27) | 3.5091 | -0.3930 | (42, 0, 54) | -0.1181 | 8.3724 |
| (-54, -20, 21) | 3.0764 | -0.2100 | (42, 2, 48) | 0.2424 | 5.7580 |
| (-54, -20, 24) | 2.9379 | -0.3506 | (42, 2, 51) | 0.8793 | 8.6167 |
| (-54, -20, 27) | 3.1689 | -0.1092 | (42, 2, 54) | 1.3396 | 8.9811 |
| (-52, -24, 21) | 3.6676 | 0.0996 | (44, -2, 48) | 0.6171 | 6.5432 |
| (-52, -24, 24) | 3.7671 | -0.1313 | (44, -2, 51) | 0.7565 | 8.2686 |
| (-52, -24, 27) | 4.1858 | -0.0592 | (44, -2, 54) | 0.7342 | 8.1278 |
| (-52, -22, 21) | 3.9665 | -0.0865 | (44, 0, 48) | 0.4410 | 6.9215 |
| (-52, -22, 24) | 3.7069 | -0.1196 | (44, 0, 51) | 0.3819 | 8.7527 |
| (-52, -22, 27) | 3.8415 | -0.0053 | (44, 0, 54) | 0.2976 | 8.0075 |
| (-52, -20, 21) | 3.9732 | 0.3424 | (44, 2, 48) | -0.1724 | 6.9471 |
| (-52, -20, 24) | 3.8957 | 0.2189 | (44, 2, 51) | 0.5470 | 8.4722 |
| (-52, -20, 27) | 4.0276 | 0.4500 | (44, 2, 54) | 0.9749 | 7.8426 |

## 7.3 Simulation with block activations

For the simulation of block activations three different shapes of clusters of artificial activations were chosen, a cube of $3 \times 3 \times 3$ voxels (27 voxels), a plus sign of seven voxels and a three dimensional letter c of ten voxels. The cube was located in medial frontal gyrus as in event-related simulation, the plus sign was located in superior frontal gyrus and the letter c in cuneus. The locations of the plus sign and the letter c are illustrated in Fig. 7.11 and the location of the cube in Fig. 7.1. The time series of these voxels were all 101 scans long and the artificial activations in these clusters all followed similar paradigm of alternating blocks of baseline and activity each consisting of ten scans. The baseline block was the first block followed by an active block. In total there were five baseline blocks and five active blocks. The artificial activation was added to time series simply by raising the intensity level 3 % during the activity blocks so this corresponds the basic box-car -function (see Fig. 6.1) with no convolution with the hrf. The activations were added again to normalized unsmoothed data and the data was then spatially smoothed with gaussian kernel with FWHM of 8 mm. The mean of the time series for every cluster before and after the addition of the artificial activation is illustrated in Fig. 7.12.

As in event-related simulation, the design matrix and the filtering matrix were both created using tools of SPM99. The design matrix for the block simulation is illustrated in Fig. 7.13. The basis functions used in this block simulation were the box-car functions (Fig. 6.1) convolved with

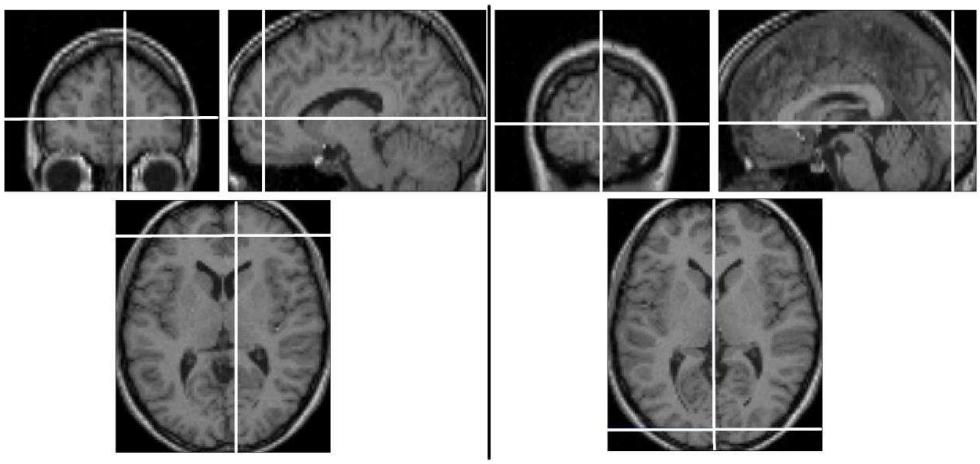


Figure 7.11: The locations of the artificial activations in block simulation (superior frontal gyrus and cuneus) shown in coronal (up left), sagittal (up right) and transversal (bottom) view. The superior frontal gyrus is on the left and the cuneus on the right.

the hrf to get the model more realistic. To take into account the differences in the timing of the raise of the signal intensity, the first time derivatives of the box-car function were also included into the model. The basis functions are illustrated in Fig. 7.14. In the design matrix there exists five columns. The resting block is modeled by box-car functions (the first column) and its first time derivative (the second column). The active block is also modeled by box-car and its first time derivative (third and fourth columns). The last column is the constant, that defines the baseline of the time series. In block simulation as in event-related simulation the data and the design matrix were filtered with high-pass filter. In block designs the $d_{\text{cut}}$ is defined as two times the experimental cycle. In this simulation the experimental cycle consisted of ten scans of activation and ten scans of baseline. Therefore, the $d_{\text{cut}} = 2 \cdot 2.5 \cdot (10 + 10) = 100$ and the high-pass filter cutoff was 0.01 Hz.

SPM99 analysis was performed again with spatially smoothed and with unsmoothed data. For the block simulation the smoothing did not remove the activations in all areas as it did with the event-related data. Activations in cube in medial frontal gyrus remained but the activations in all the other areas were removed. In Fig. 7.15 the mean of the time series of the voxels in the cube in medial frontal gyrus is illustrated before and after spatial smoothing. As it can be seen the smoothing has affected the active blocks by decreasing the intensity difference between the active state and the baseline, but there is still activity remaining. The results of SPM99 analysis are shown in Fig. 7.16. In the top of the figure the activations are shown in maximum intensity projection (MIP) from three orthogonal view (right, back, and top views). The table below the glass brain shows the statistical information SPM99 has computed. The quantity of the voxels in each cluster can be seen in this table. SPM99 found 11 active voxels in the cube in medial frontal lobe in analysis for the smoothed data. Other areas that contained artificial activation was not be recognized as active.The used threshold was $p \leq 0.001$, uncorrected which equates the minimum $t$-value of 3.18. The size of the active cluster was not restricted. With threshold $p \leq 0.05$, corrected no voxels were detected as active. The minimum $t$-value that was detected as active was 4.75.

The same analysis with the same threshold as with smoothed data was performed also with unsmoothed data. The results of this are shown in Fig. 7.17. SPM99 detected all seven voxels in plus sign and all 27 voxels in cube as active, but only 8 voxels of 10 in letter c. It also detected several other voxels as active, although the time series of these voxels were not manipulated by adding any artificial activation. Nevertheless the largest $t$-values were detected in the clusters of artificial activation. The lowest $t$-value SPM99 detected as active was 3.18.

With threshold $p \leq 0.05$, corrected only the plus sign (all seven voxels), a part of the cube (24

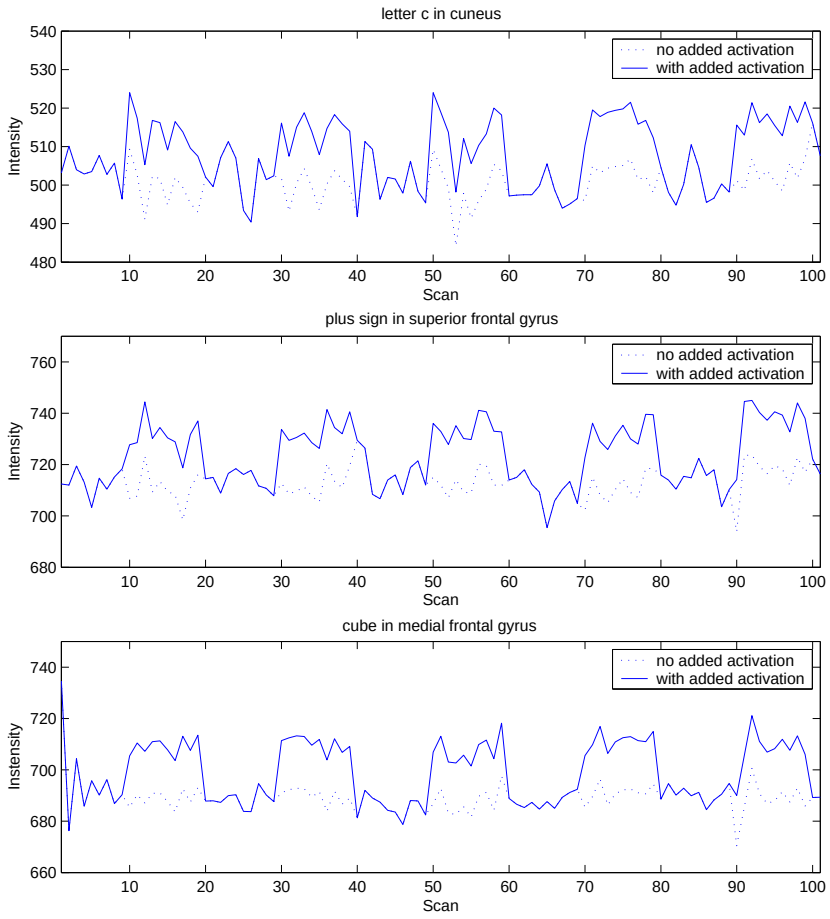



Figure 7.12: The mean of voxels' time series before and after the insertion of artificial activation for every cluster in block simulation. The uppermost box is for the activation in the cuneus, the box in the middle is for the activation in the superior frontal gyrus and the undermost box is for the activation in the medial frontal gyrus. The raise of intensity during the activation was 3 %.

voxels of 27) and three voxels from the letter c were detected as active. There were no other active voxels detected. The minimum $t$-value that was detected as active was 5.26.

For the analysis with Matlab the data and the design matrix were filtered with the filtering matrix from SPM99 analysis using the eq. (7.2). The Matlab analysis was performed with smoothed and unsmoothed data since SPM99 found some active voxels in the cube in medial frontal gyrus. The calculated $t$-values from eq. (6.19) for both smoothed and unsmoothed data with the voxel co-ordinates are shown in table 7.2 for the voxels in the plus sign and in the letter c and in table 7.3 for the voxels in the cube. The $t$-values for the smoothed data were systematically smaller than the $t$-values for the unsmoothed data as it can be seen in the tables.

The fitting of the basis functions to the filtered data is shown in the Fig. 7.18. The fitting is actually the estimation of the data based on the information of the basis functions. This is illustrated for both the unsmoothed and the smoothed data by using one randomly selected voxel from each area of interest. Because the active areas in the plus sign and in the letter c were only few voxels the smoothing with default FWHM of the gaussian kernel has removed the activations. The active area in the cube was much larger and, therefore, there are some activation still left.

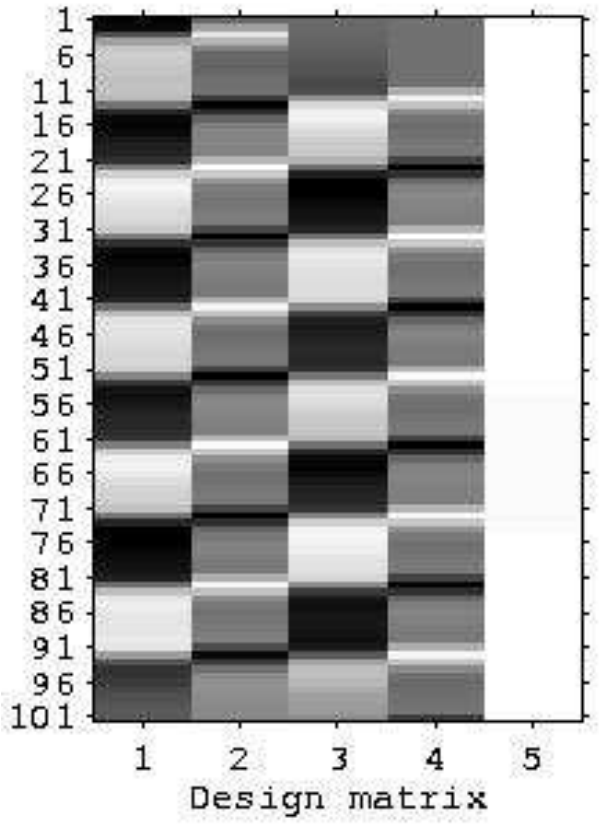



Figure 7.13: The design matrix of the block simulation. The first column represents the resting block, the second is its first time derivative, third is the active block, fourth its first time derivative and the last column is the constant.

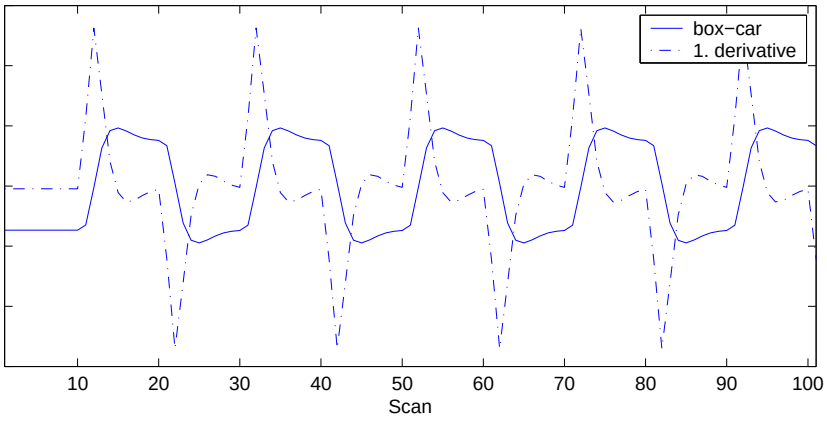



Figure 7.14: The basis functions of the active block in the block simulation. The solid line represents the box-car function and the dashdot line its first time derivative.

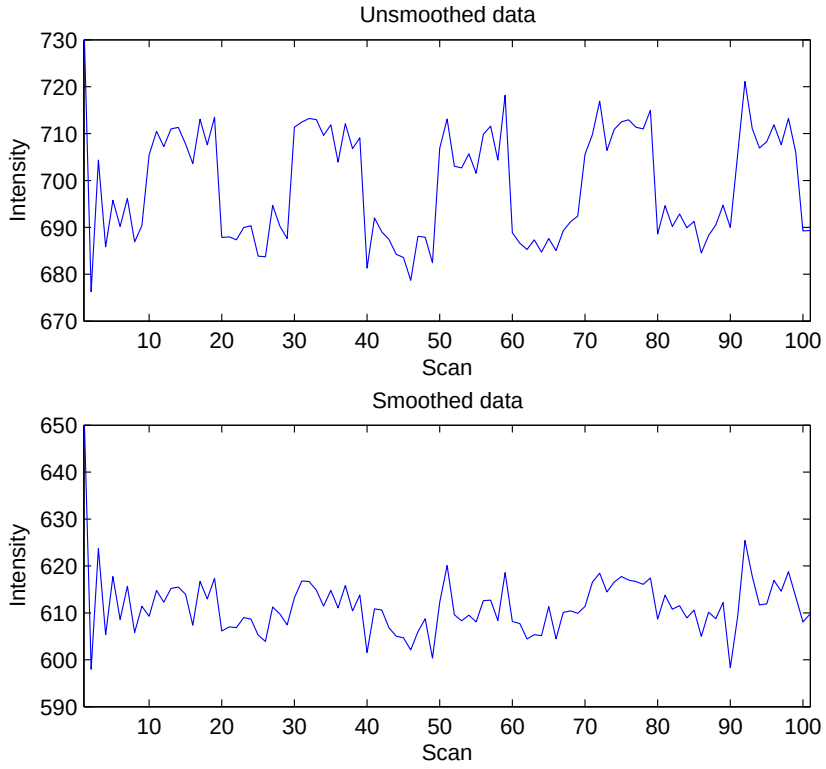
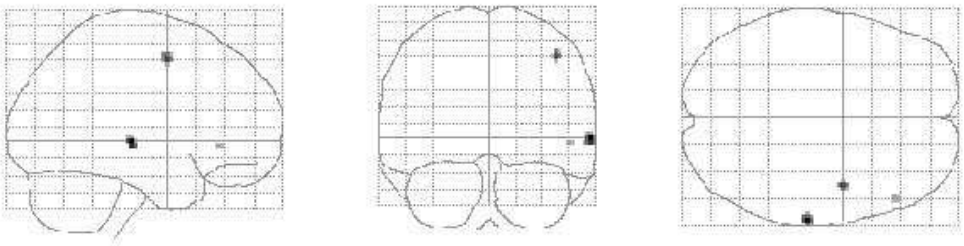


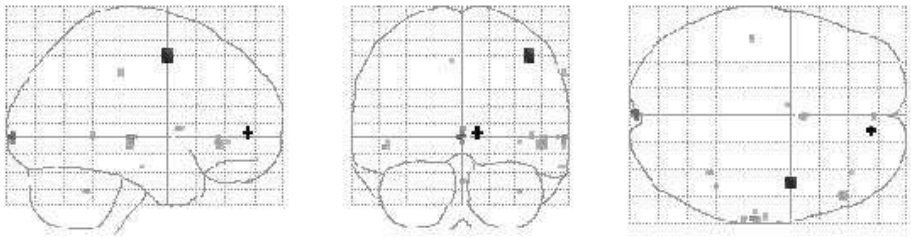

Figure 7.15: The effect of spatial smoothing (gaussian kernel, FWHM 8 mm) to block simulation in medial frontal gyrus. The upper window shows the mean of the time series before smoothing and the lower shows the same mean after smoothing.



**Statistics:** *volume summary (p-values corrected for entire volume)*

| set-level | | cluster-level | | | voxel-level | | | | | x,y,z {mm} |
|---|---|---|---|---|---|---|---|---|---|---|
| $p$ | $c$ | $p_{corrected}$ | $\hat{k}_E$ | $p_{uncorrected}$ | $p_{FWE-corr}$ | $p_{FDR-corr}$ | $T$ | $(Z_e)$ | $p_{uncorrected}$ | |
| 0.817 | 3 | 0.842 | 18 | 0.417 | 0.281 | 1.000 | 4.16 | ( 3.97) | 0.000 | 64 -22 0 |
| | | 0.905 | 11 | 0.533 | 0.611 | 1.000 | 3.80 | ( 3.65) | 0.000 | 42 0 51 |
| | | 0.959 | 4 | 0.726 | 0.966 | 1.000 | 3.30 | ( 3.20) | 0.001 | 50 32 -3 |

Figure 7.16: The results of SPM99 analysis for the smoothed block simulation data. The three orthogonal views of the brain shows the all activation of that particular view plotted one on the other. The table below this shows the statistical details of the active clusters SPM99 has detected.



**Statistics:** *volume summary (p-values corrected for entire volume)*

| set-level | | cluster-level | | | voxel-level | | | | | x,y,z (mm) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| p | c | $p_{corrected}$ | $\hat{k}_E$ | $p_{uncorrected}$ | $p_{FWE-corr}$ | $p_{FDR-corr}$ | T | ($Z_=$) | $p_{uncorrected}$ | | | |
| 0.999 | 15 | 0.991 | 7 | 0.153 | 0.000 | 0.000 | 8.28 | ( 7.13) | 0.000 | 10 | 50 | 3 |
| | | 0.266 | 27 | 0.010 | 0.000 | 0.000 | 7.20 | ( 6.39) | 0.000 | 44 | 0 | 48 |
| | | 0.982 | 8 | 0.129 | 0.019 | 0.001 | 5.49 | ( 5.09) | 0.000 | 0 | -96 | 0 |
| | | 0.982 | 8 | 0.129 | 0.801 | 0.061 | 4.29 | ( 4.09) | 0.000 | 64 | -24 | 0 |
| | | 0.795 | 14 | 0.051 | 0.993 | 0.207 | 3.91 | ( 3.75) | 0.000 | 52 | 32 | -3 |
| | | 1.000 | 2 | 0.444 | 0.999 | 0.300 | 3.80 | ( 3.65) | 0.000 | 36 | -50 | -33 |
| | | 0.999 | 5 | 0.224 | 0.999 | 0.304 | 3.79 | ( 3.64) | 0.000 | 0 | 6 | 6 |
| | | 1.000 | 3 | 0.345 | 1.000 | 0.352 | 3.73 | ( 3.59) | 0.000 | -48 | -24 | -6 |
| | | 1.000 | 2 | 0.444 | 1.000 | 0.522 | 3.58 | ( 3.45) | 0.000 | 0 | 52 | -27 |
| | | 0.999 | 5 | 0.224 | 1.000 | 0.546 | 3.55 | ( 3.42) | 0.000 | 66 | -30 | 39 |
| | | 1.000 | 2 | 0.444 | 1.000 | 0.754 | 3.43 | ( 3.31) | 0.000 | 44 | -46 | 3 |
| | | 1.000 | 2 | 0.444 | 1.000 | 0.818 | 3.38 | ( 3.27) | 0.001 | 62 | -16 | -18 |
| | | 1.000 | 1 | 0.599 | 1.000 | 0.921 | 3.31 | ( 3.21) | 0.001 | 60 | -22 | 0 |
| | | 1.000 | 1 | 0.599 | 1.000 | 1.000 | 3.21 | ( 3.11) | 0.001 | -6 | -2 | 48 |
| | | 1.000 | 1 | 0.599 | 1.000 | 1.000 | 3.20 | ( 3.10) | 0.001 | 36 | 38 | -3 |

Figure 7.17: The results of SPM99 analysis for the unsmoothed block simulation data. This MIP -visualization shows two of the artificial active clusters: the cube in medial frontal gyrus, and the plus sign in the superior frontal gyrus. The third cluster (the letter c in cuneus) is less visible. The table below shows the statistical details of the active clusters SPM99 has detected.

TABLE 7.2: The $t$-values from Matlab analysis for the voxels in the letter c in cuneus and in the plus sign in posterior frontal gyrus for both spatially smoothed and unsmoothed data in block simulation.

| the letter c in cuneus | | | the plus sign in posterior frontal gyrus | | |
|---|---|---|---|---|---|
| voxel $(x, y, z)$ | $t$-value (unsmoothed) | $t$-value (smoothed) | voxel $(x, y, z)$ | $t$-value (unsmoothed) | $t$-value (smooth) |
| ( 0, -96, 0) | 5.4912 | 0.8737 | (10, 50, 3) | 8.2827 | 2.5672 |
| (-2, -96, 0) | 5.3745 | 0.6833 | (10, 50, 0) | 7.7082 | 2.1106 |
| ( 0, -96, 3) | 4.8434 | 0.4502 | (10, 50, 6) | 7.8941 | 2.3076 |
| ( 2, -96, 0) | 5.0669 | 0.5937 | ( 8, 50, 3) | 6.7992 | 2.3712 |
| ( 0, -96, -3) | 4.2715 | 0.6555 | (12, 50, 3) | 7.4180 | 2.5944 |
| ( 0, -98, 0) | 4.2502 | 0.5237 | (10, 48, 3) | 7.9869 | 2.4343 |
| (-2, -98, 0) | 5.3045 | 0.2590 | (10, 52, 3) | 7.3912 | 2.2052 |
| ( 0, -98, 3) | 3.1623 | 0.0622 | | | |
| ( 2, -98, 0) | 2.9079 | 0.2900 | | | |
| ( 0, -98, -3) | 3.3242 | 0.3129 | | | |

TABLE 7.3: The $t$-values form Matlab analysis for the voxels in the cube in medial frontal gyrus for both spatially smoothed and unsmoothed data in the block simulation.

the cube in medial frontal gyrus

| voxel $(x, y, z)$ | $t$-value (unsmoothed) | $t$-value smoothed | voxel $(x, y, z)$ | $t$-value (unsmoothed) | $t$-value smoothed |
|---|---|---|---|---|---|
| (40, -2, 48) | 6.5503 | 2.4562 | (42, 0, 54) | 5.8621 | 3.5438 |
| (40, -2, 51) | 6.9843 | 3.0868 | (42, 2, 48) | 6.2316 | 2.9669 |
| (40, -2, 54) | 6.7382 | 3.0414 | (42, 2, 51) | 6.4449 | 3.5073 |
| (40, 0, 48) | 6.5431 | 2.7224 | (42, 2, 54) | 5.5932 | 3.2031 |
| (40, 0, 51) | 6.5680 | 3.1928 | (44, -2, 48) | 6.3970 | 2.4802 |
| (40, 0, 54) | 5.9896 | 3.0362 | (44, -2, 51) | 6.2211 | 3.1014 |
| (40, 2, 48) | 6.0744 | 2.5056 | (44, -2, 54) | 5.5764 | 2.9684 |
| (40, 2, 51) | 5.9020 | 2.8108 | (44, 0, 48) | 7.1973 | 3.1699 |
| (40, 2, 54) | 5.2501 | 2.7305 | (44, 0, 51) | 6.5068 | 3.5488 |
| (42, -2, 48) | 6.5205 | 2.7976 | (44, 0, 54) | 5.1653 | 3.2032 |
| (42, -2, 51) | 6.7626 | 3.4101 | (44, 2, 48) | 6.7405 | 2.9019 |
| (42, -2, 54) | 6.2392 | 3.4842 | (44, 2, 51) | 6.3690 | 3.2333 |
| (42, 0, 48) | 7.0486 | 3.2547 | (44, 2, 54) | 5.1476 | 2.8643 |
| (42, 0, 51) | 6.7661 | 3.7986 | | | |

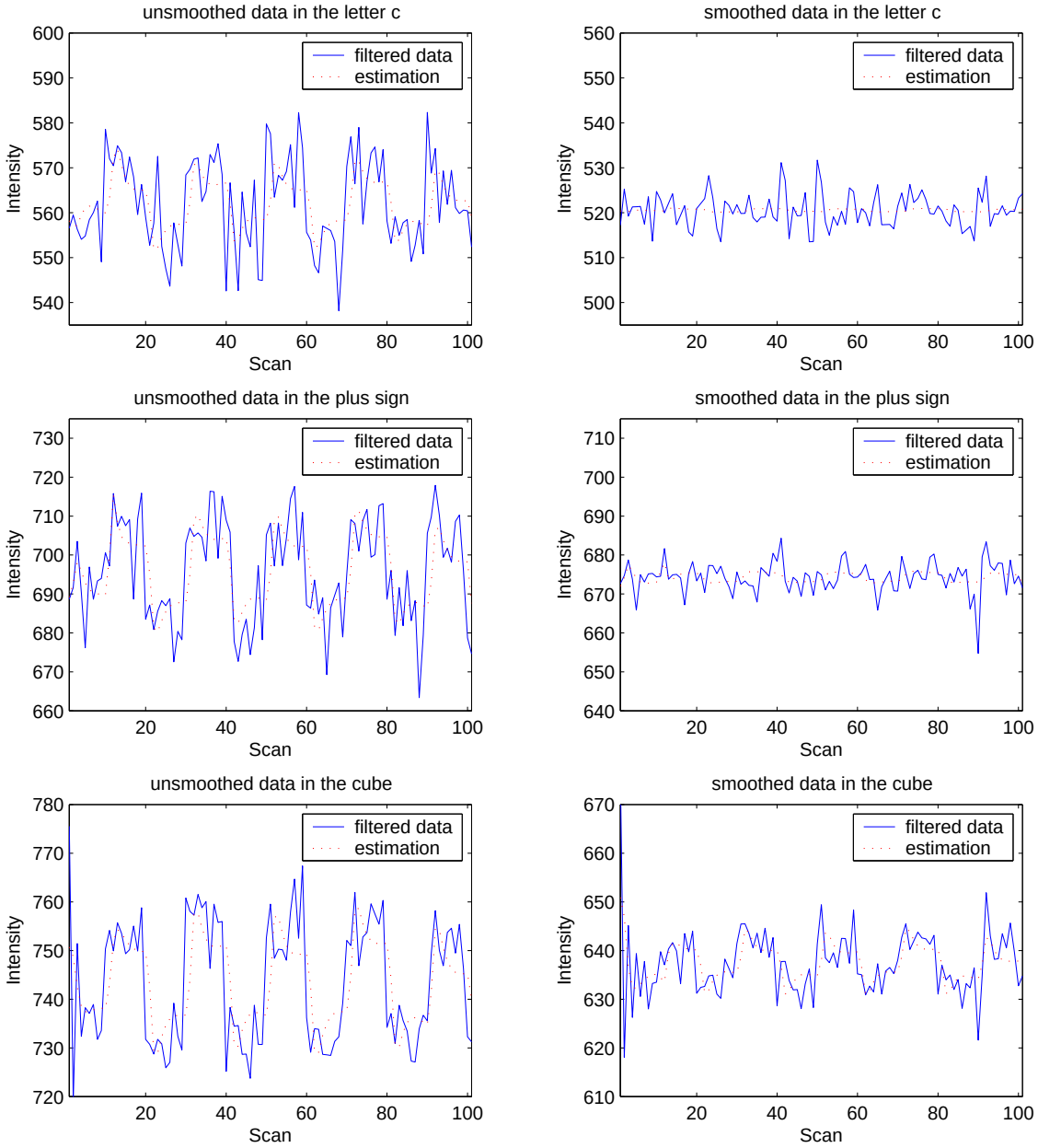



Figure 7.18: The fitting of the basis functions in areas of interested for both unsmoothed and smoothed data. For each area the example voxel was randomly selected. The co-ordinates of the selected voxel in the cuneus were (0,-96,-3), in the posterior frontal gyrus (12, 50, 3) and in the medial frontal gyrus (42, 0, 51). The unsmoothed data is on the left and the smoothed on the right. The solid line represents the filtered data and the dotted line the estimation of the filtered data based on the basis functions.

Chapter 8

# Discussion

SPM99 is one of the most common fMRI analysis tools in neuroimaging community, although a comprehensive manual of the mathematics it uses in estimation and statistics has not been published. That was the main reason for the topic of this thesis: to cather the information form articles and internet publications to a more comprehensive and simple form. The theory of SPM99 was also verified with simulation. The main interest was in the estimation and statistical methods, but also the preprocessing steps of SPM99 were examined.

The preprocessing steps were considered in Chapter 5. The preprocessing steps are essential part of the fMRI analysis, so the general idea of each step was explained. In SPM99 there are several choices in each step that affects the result e.g. the used interpolation method or some other defaults that can be altered. The effect of these was not examined in this thesis.

The estimation and statistics were studied in Chapter 6. In this thesis only the single level analysis was concerned. In fMRI studies it is common to perform the analysis as a two-stage procedure. At the first stage the analysis is carried out for every subject separately (first level analysis). At the second stage the results from the first stage are used in the multisubject level (second level analysis) in order to get more general results concerning the population the subjects were chosen. With SPM99 it is possible to perform also multisubject analysis after first level analysis has been performed. In first level analysis it is possible to perform either $t$-test or $F$-test in statistical evaluation. When the aim of the study is to get general results and, therefore, perform multisubject analysis, the $t$-test is the only choice. The statistical tests in the second level analysis use the contrasts defined in the first level. The contrast in $t$-test is defined as a vector and in $F$-test it is in a matrix form. The matrix form, however, is unsuitable for the statistical tests used in the second level. Therefore, the results from the $F$-test cannot be used in second level. Only the $t$-test is studied in this thesis, because it is suitable for first level analysis and also for the second level analysis. The second level analysis and the $F$-test is excluded from this thesis to simplify and restrict the topic.

The verification with simulations in Chapter 7 was valuable for the evaluation of the correctness of the equations in Chapter 6. The filtering in time space tend to create some problems. The original article about filtering and the filtering matrix in the equation of the $t$-value (eq. 6.19) did not correspond to the way the calculation was performed in SPM99. The actual way of calculation was then clarified after studying the Matlab code and more articles. The method in calculating the filtering matrix and performing the filtering is presented in the beginning of the Chapter 7.

The artificial activation in simulations in Chapter 7 was added to preprocessed data. Only the spatial smoothing was performed after the activations were inserted. Although in this simulation the used data was only from one subject, all the preprocessing steps that are recommended to be performed to data in multisubject study were also performed to this data. This was done because the data wanted to be as authentic as possible compared to data in real fMRI studies. The artificial activation was not inserted to raw unprocessed data, because the normalization step reshapes the EPIs so much that the precise localizing of the active voxels would be impossible. The precise locations of the active voxels was crucial to know for the Matlab analysis and for the comparison

of the results and the time-series of the voxels. There is a certain recommendation from the creators of SPM99 of the order of the preprocessing steps. It is recommended to perform the realignment at first, then slice timing, coregistration (if needed), normalization and finally smoothing. This order is not the only one right, but it is the most used. The data in the simulations was preprocessed in this way and, therefore, only smoothing was performed after the artificial activations were placed.

The results from SPM99 analysis and Matlab analysis matched perfectly after the problem with the filtering matrix was solved. The $t$-values and the effective degrees of freedom were the same, no matter which analysis method was used. The only difference was that SPM99 does not give the $t$-values for every voxel but only to those that survive the chosen threshold. The $t$-values for all of the voxels can however be found from the $t$-value images, but there is no possibility to get a table of $t$-values of certain voxels. The only way to get these values would be to display the $t$-value image and then to click with the mouse on each voxel. Hereby the $t$-value of that voxel is displayed on the screen. If there are several voxels this clicking might be little frustrating on the long run.

As a by-product of this simulation it was noticed that the spatial smoothing really affects the small activated areas. In these simulations only the large area of 27 voxels in block study survived after the spatial smoothing despite of the threshold. This is understandable because the spatial smoothing smoothess the differences of the intensities of the neighboring voxels. If only few voxels has been active and the increase in signal intensity has not been huge, these differences are lost in this preprocessing step. The spatial smoothing has been included in SPM99 because of the advantages. Spatial smoothing alters the data to be more Gaussian which is one assumption of the statistical $t$-test. It also improves the SNR by smoothing the noise spikes. Unfortunately, this affects also actual activations if they are spikelike. A general correct FWHM for every different data is difficult to determine because of the variation of the imaging methods and paradigms. Only some guidelines have been published. For the simulation in this thesis the FWHM could have been altered, but it was decided to use only the FWHM of 8 mm or no smoothing at all, because the FWHM of 8 mm to a data with original pixel size in plane of $4 \times 4$ mm was recommended by the creators of SPM99.

In the future the estimation method and the statistics is to be developed. The nonlinearity of the whole phenomenon from a stimulus to detected signal encourages to modify the analysis to more nonlinear direction. New estimation methods also may solve the problem of spatial smoothing.

# Bibliography


[1] G. K. Aguirre, E. Zarahn, and M. D'Esposito. The variability of human, BOLD hemodynamic responses. *NeuroImage*, 8(4):360–369, November 1998.

[2] D. G. Altman. *Practical statistics for medical research*. Chapman & Hall, 1991.

[3] J. Ashburner and K. J. Friston. Spatial transformation of images. SPMcourse, short course notes, May 1997.

[4] J. Ashburner and K. Friston. Multimodal image coregistration and partitioning - a unified framework. *NeuroImage*, 6(3):209–217, 1997.

[5] J. Ashburner, K. Friston, and W. Penny, editors. *Human Brain function, $2^{nd}$ edition*, chapter 7, The general linear model.

[6] J. Ashburner and K. J. Friston. Nonlinear spatial normalization using basis functions. *Human Brain Mapping*, 7(4):254–266, 1999.

[7] J. Ashburner, P. Neelin, D. L. Collins, A. Evans, and K. Friston. Incorporating prior knowledge into image registration. *NeuroImage*, 6(4):344–352, 1997.

[8] J. Astola and I. Virtanen. Entropy correlation coefficient a measure of statistical dependence for categorized data. In *Proceedings of the University of Vaasa, Discussion papers*, number 44. 1982.

[9] R. M. Birn, R. W. Cox, and P. A. Bandettini. RAPID COMMUNICATION, detection versus estimation in event-related fMRI: Choosing the optimal stimulus timing. *NeuroImage*, 15(1):252–264, 2002.

[10] R. M. Birn, Z. S. Saad, and P. A. Bandettini. Spatial Heterogeneity of the Nonlinear Dynamics in the FMRI BOLD Response. *NeuroImage*, 14(4):817–826, 2001.

[11] A. M. Blamire, S. Ogawa, K. Ugurbil, D. Rothman, G. McCarthy, J. M. Ellerman, F. Hyder, Z. Rattner, and R. G. Shulman. Dynamic mapping of the human visual cortex by high-speed magnetic resonance imaging. *Proc. Natl. Acad Sci. USA*, 89:11069–11073, November 1992.

[12] G. M. Boynton, S. A. Engel, G. H. Glover, and D. J. Heeger. Linear systems analysis of functional magnetic resonance imaging in human V1. *The Journal of Neuroscience*, 16(13):4207–4221, July 1996.

[13] R. L. Buckner, P. A. Bandettini, K. M. O'Craven, R. L. Savoy, S. E. Petersen, M. E. Raichle, and B. R. Rosen. Detection of cortical activation during averaged single trials of a cognitive task using functional magnetic resonance imaging. *Proc. Natl. Acad. Sci. USA*, 93:14878–14883, December 1996.

[14] J. T. Cacioppo, L. G. Tassinary, and G. G. Berntson, editors. *Handbook of Psychophysiology*. Cambridge University Press, 2. edition, 2000.

[15] V. P. Clark, J. M. Maisog, and J. V. Haxby. fMRI Study of Face Perception and Memory Using Random Stimulus Sequences. *Journal of Neurophysiology*, 79(6):3257–3265, June 1998.

[16] M. S. Cohen. Parametric Analysis of fMRI Data Using Linear Systems Methods. *NeuroImage*, 6(2):93–103, 1997.

[17] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal. Automated multi-modality image registration based on information theory. In Y. Bizais, C. Barillot, and R. D. Paola, editors, *Proc. 14th Int. Conf. Information processing in Medical Imaging*, pages 263–274. Kluwer Academic Publishers, June 1995.

[18] B. P. Cowan. *Nuclear Magnetic Resonance and Relaxation*. Cambridge University Press, Cambridge, 1997.

[19] A. M. Dale and R. L. Buckner. Selective averaging of rapidly presented individual trials using fMRI. *Human Brain Mapping*, 5(5):329–340, 1997.

[20] P. Fransson, G. Krüger, K.-D. Merbolt, and J. Frahm. Temporal characteristics of oxygenation-sensitive MRI responses to visual activation in humans. *Magn Reson Med*, 39:912–929, 1998.

[21] K. Friston, A. Holmes, J.-B. Poline, P. Grasby, S. Williams, R. Frackowiak, and R. Turner. Analysis of fMRI time-series revisited. *NeuroImage*, 2:45–53, 1995.

[22] K. J. Friston. Statistical parametric mapping and other analysis of functional imaging data. *Brain Mapping*, pages 363–386, 1996.

[23] K. J. Friston. Basic concepts and overview. SPMcourse, short course notes, May 1997.

[24] K. J. Friston, J. Ashburner, C. D. Frith, J.-B. Poline, J. D. Heather, and R. S. J. Frackowiak. Spatial registration and normalization of images. *Human Brain Mapping*, 2:165–189, 1995.

[25] K. J. Friston, O. Josephs, G. Rees, and R. Turner. Nonlinear event-related responses in fMRI. *Magn Reson Med*, 39(1):41–52, 1998.

[26] K. J. Friston, O. Josephs, E. Zarahn, A. P. Holmes, S. Rouquette, and J.-B. Poline. To smooth or not to smooth? Bias and efficiency in fMRI time-series analysis. *NeuroImage*, 12(2):196–208, 2000.

[27] D. G. Gadian. *NMR and its applications to living systems*. Oxford Science Publications, second edition, 1995.

[28] C. R. Genovese, N. A. Lazar, and T. Nichols. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage*, 15(4):870–878, 2002.

[29] G. H. Glover. Deconvolution of Impulse Response in Event-Related BOLD fMRI. *NeuroImage*, 9:416–429, 1999.

[30] D. J. Griffiths. *Introduction to quantum mechanics*. Prentice Hall, Inc, 1995.

[31] C. Gössl, L. Fahrmeir, and D. P. Auer. Bayesian Modeling of the Hemodynamic Response Function in BOLD fMRI. *NeuroImage*, 14(1):140–148, 2001.

[32] E. M. Haacke, R. W. Brown, M. R. Thompson, and R. Venkatesan. *Magnetic Resonance Imaging - Physical Principles and Sequence Desing*. A John Wiley & sons, Inc., Publication, 1999.

[33] R. Harris. *Nonclassical physics, Beyond Newton's view*. Addison Wesley Longman, Inc, 1999.

[34] R. D. Hoge, J. Atkinson, B. Gill, G. R. Crelier, S. Marrett, and G. B. Pike. Stimulus-Dependent BOLD and Perfusion Dynamics in Human V1. *NeuroImage*, 9(6):573–585, 1999.

[35] J. P. Hornak. The basics of MRI. World Wide Web, `http://www.cis.rit.edu/htbooks/mri/`.

[36] B. Horwitz, K. J. Friston, and J. G. Taylor. Neural modeling and functional brain imaging: an overview. *Neural Networks*, 13:829–846, 2000.

[37] A. M. Howseman and R. W. Bowtell. Functional magnetic resonance imaging: imaging techniques and contrast mechanisms. *Phil. Trans. R. Soc. Lond. B*, 354:1179–1194, 1999.

[38] A. M. Howseman, S. Grootoonk, D. A. Porter, J. Ramdeen, A. P. Holmes, and R. Turner. The effect of slice order and thickness on fMRI activation data using multislice echo-planar imaging. *NeuroImage*, 9(4):363–376, April 1999.

[39] A. M. Howseman, O. Josephs, G. Rees, and K. J. Friston. Special Issues in Functional Magnetic Resonance Imaging. SPM course, short course notes, May 1997.

[40] O. Josephs and R. N. A. Henson. Event-related functional magnetic resonance imaging: modelling, inference and optimization. *Phil. Trans. R. Soc. Lond. B*, 354:1215–1228, 1999.

[41] P. Karjalainen. Fysiikan data-analyysi, April 1998. University of Kuopio, Department of Applied Physics, ISSN 0788-4672.

[42] S.-G. Kim, W. Richter, and K. Uğurbil. Limitations of temporal resolution in functional MRI. *Magnetic Resonance in Medicine*, 37(4):631–636, April 1997.

[43] K. K. Kwong, J. W. Belliveau, D. A. Chesler, I. E. Goldberg, R. M. Weisskof, B. P. Poncelet, D. N. Kennedy, B. E. Hoppel, M. S. Cohen, R. Turner, H.-M. Cheng, T. J. Brady, and B. R. Rosen. Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proc. Natl. Acad Sci. USA*, 89:5675–5679, June 1992.

[44] A. T. Lee, G. H. Glover, and C. H. Meyer. Discrimination of large venous vessels in time-course spiral blood-oxygen-level-dependent magnetic-resonance functional neuroimaging. *Magnetic Resonance in Medicine*, 33(6):745–754, June 1995.

[45] T. T. Liu, L. R. Frank, E. C. Wong, and R. B. Buxton. Detection Power, Estimation Efficiency, and Predictability in Event-Related fMRI. *NeuroImage*, 13(4):759–773, 2001.

[46] F. Maes, A. Collignon, D. Vandermeulen, G. Matchal, and P. Suetens. Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging*, 16(2):187–198, 1997.

[47] P. Mansfield. Multi-planar image formation using NMR spin echoes. *J. Phys. C*, 10:L55–L58, 1977.

[48] A. Mechelli, C. J. Price, and K. J. Friston. Nonlinear coupling between evoked rCBF and BOLD signals: A simulation study of hemodynamic responses. *NeuroImage*, 14(4):862–872, 2001.

[49] F. M. Miezin, L. Maccotta, J. M. Ollinger, S. E. Petersen, and R. L. Buckner. Characterizing the hemodynamic response: Effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. *NeuroImage*, 11(6):735–759, June 2000.

[50] K. L. Miller, W.-M. Luh, T. T. Liu, A. Martinez, T. Obata, E. C. Wong, L. R. Frank, and R. B. Buxton. Nonlinear temporal dynamics of the cerebral blood flow response. *Human Brain Mapping*, 13:1–12, 2001.

[51] R. G. Miller. *Simultaneous statistical inference*. Springer-Verlag, New York, 1981.

[52] S. Ogawa, T. M. Lee, A. R. Kay, and D. W. Tank. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proc. Natl. Acad Sci. USA*, 87:9868–9872, December 1990.

[53] J. M. Ollinger, M. Corbetta, and G. L. Shulman. Separating Processes within a Trial in Event-Related Functional MRI. *NeuroImage*, 13(1):218–229, 2001.

[54] W. Pan. Approximative confidence intervals for one proportion and difference of two proportions. *Computational Statistics & Data Analysis*, 40:143–157, 2002.

[55] J.-B. Poline. Spm 2001 course slides, contrasts and classical inference. `http://www.fil.ion.ucl.ac.uk/spm/course/notes01.html`.

[56] J. C. Rajapakse, F. Kruggel, J. M. Maisog, and D. Y. von Cramon. Modeling Hemodynamic Response for Analysis of Functional MRI Time-Series. *Human Brain Mapping*, 6(4):283–300, 1998.

[57] B. R. Rosen, R. L. Buckner, and A. M. Dale. Event-related functional MRI: Past, present and future. *Proc. Natl. Acad. Sci. USA*, 95:773–780, 1998.

[58] F. E. Satterthwaite. An approximate distribution of estimates of variance components. *Biometric Bulletin*, 2:110–114, 1946.

[59] D. R. Savoy. Course material for fMRI-based experiments. AMI Center, Neuro HUT, March 2002.

[60] H. H. Schild. *MRI made easy (... well almost)*. Schering AG, 1990.

[61] G. A. F. Seber. *Linear Regression Analysis*. John Wiley & Sons, 1977.

[62] G. A. F. Seber and A. J. Lee. *Linear Regression Analysis*, chapter 1, page 9. Wiley Interscience, 2003.

[63] E. Soini and E. Hiltunen. Kuvantamismenetelmien fysikaaliset perusteet. Turun yliopisto, lääketieteellinen fysiikka, 1997.

[64] M. K. Stehling, R. Turner, and P. Mansfield. Echo-planar imaging: Magnetic resonance imaging in a fraction of a second. *Science*, 254:43–50, October 1991.

[65] J. Talairach and P. Tourneaux. *Co-Planar stereotaxic atlas of the human brain: 3-dimensional proportional system: An approach to cerebral imaging*. Thieme Medical Pub., 1988.

[66] M. Uhari and P. Nieminen. *Epidemiologia & biostatistiikka*. Kustannus Oy Duodecim, 1 edition, 2001.

[67] A. L. Vazquez and D. C. Noll. Nonlinear aspects of the BOLD response in functional MRI. *NeuroImage*, 7(2):108–118, 1998.

[68] A. B. Wolbarst. *Physics of Radiology*. Appleton & Lange, 1993.

[69] K. J. Worsley. fmristat, a general statistical analysis for fmri data. `http://www.math.mcgill.ca/keith/fmristat/`.

[70] K. J. Worsley and K. J. Friston. Analysis of fMRI time-series revisited - again. *NeuroImage*, 2:173–181, 1995.

[71] K. J. Worsley, S. Marrett, P. Neelin, A. C. Vandal, K. J. Friston, and A. C. Evans. A unified statistical approach for determining significant signals in images of cerebal activation. *Human Brain Mapping*, 4(1):58–73, 1996.

[72] E. Zarahn, G. Aguirre, and M. D'Esposito. A trial-based experimental design for fMRI. *NeuroImage*, 6:122–138, 1997.