Modelling of resting state fMRI

Mette Pejtersen

Kongens Lyngby 2013
IMM-MSc-2013-2

Technical University of Denmark
Informatics and Mathematical Modelling
Building 321, DK-2800 Kongens Lyngby, Denmark
Phone +45 45253351, Fax +45 45882673
reception@imm.dtu.dk
www.imm.dtu.dk IMM-MSc-2013-2

Summary

Multiple sclerosis is a complex disease that can affect all parts of the brain. The symptoms are varying from patient to patient and the disease is difficult to diagnose. Changes in the default mode network has been observed in patients with multiple sclerosis.
In this thesis resting state networks are detected by using the infinite relation model. SVM and KNN is used to classify subjects into two groups: A healthy group and a group of patients with multiple sclerosis. The elements in the η (link density) matrix are used as feature vector for classification. In the end the correlation between link densities and the progression in the disease is evaluated. The participants in the resting state fMRI study were 30 healthy subjects and 42 patients with multiple sclerosis. The two groups were matching in sex and gender.
The highest mean classification rate was 0.65 when using SVM and 0.61 for KNN. A large variation between the runs were found. For one run the classification rate was 0.73 when using SVM and 0.66 when using KNN. These results are comparable to classification results represented in the literature. 32 communities were detected by using the infinite relation model, and some of the communities were comparable with the default mode network, primary motor network, and the frontal network presented in the. One community seems to be comparable with both the primary visual and the extra-striate visual network. No significant correlation between EDSS and link density was found.
The results show that commonly represented resting state networks in the literature can be detected when using the infinite relation model. The best classification rates for using SVM and KNN were comparable with previous results, but the large variability in the classification rate is not optimal.

Resumé

Multipel sclerose er en komplex sygdom, der kan påvirke alle dele af hjernen. Symptomer varierer fra patient til patient, og sygdommen er svær at diagnosticere. Ændringer i default mode netværket er blevet observeret hos patienter med multipel sclerose.

I denne afhandling anvende infinite relation model til at detektere resting state netværk. SVM og KNN anvendes til at klassificere test personer, således at de opdeles i to grupper med henholdsvis raske og syge. Elementerne i η (link densitet) matricen anvendes som feature vector til brug i klassifikationen. Korrelationen mellem link densiteter og udviklingen af multipel sclerose bliver evalueret.

Resting state fMRI blev foretage på en gruppe testpersoner bestående af 30 raske og 42 patienter med multipel sclerose. De grupper var sammenlignelig i forhold til køn og alder. Den bedste middel klassifikationsrate var på 0.65 for SVM og på 0.61 for KNN. Der blev observeret en stor variation imellem de enkelte runs. Den bedste klassifikationsrate for et enkelt run var på 0.73 for SVM og på 0.66 for KNN. Disse resultater er sammenlignelige med med lignende klassifikationsresultater fra literaturen. 32 communities blev detekteret ved at bruge infinite relation model. Nogle af disse communities var sammenlignelige med default mode netværket, primary motor netværk samt frontalnetværket, som er beskrevet i litteraturen. En enkelt community var sammenlignelig med både det primary visuelle og the extra-striate visuelle netværk. Ingen significant korrelation blev fundet mellem EDSS og link densiteten.

Resultaterne viser at ofte beskrevne resting state netværk fra literaturen kan detekteres med brug af infinite relation model. Den bedste klassifikationsrate for anvendelsen af SVM og KNN var sammenlignelig tidligere resultater, men

den store variation mellem klassifikationsraterne for enkelte runs er ikke optimal.

Preface

This thesis was prepared at DTU Informatics the department of Informatics and Mathematical Modelling at the Technical University of Denmark as a partial fulfillment of the requirements for acquiring an M.Sc. in Medicine and Technology.

The thesis was written in coorporation with Danish Research Centre for Magnetic Resonance.

The work on this thesis was carried out from September 3, 2012 to February 4, 2013.

Lyngby, 04-February-2013

Mette Pejtersen

Acknowledgements

I would like to thank my supervisors Professor Lars Kai Hansen and Kasper Winther Andersen for their guidance and support throughout the project period.

I would also like to thank my family for their support.

List of abbreviations

AAL	Automated Anatomical Labeling
ACC	Anterior Cingulate Cortex
AUC	Area Under Curve
BCD	Bayesian Community Detection
BOLD	Blood Oxygenation Level Depend
CNS	Central Nervous System
CRP	Chinese Restaurant Process
DMN	Default Mode Network
EDSS	Expanded Disability Status Score
EPI	Echo Planar Imaging
fMRI	functional Magnetic Resonance Imaging
FWHM	Full Width Half Maximum
HC	Healthy Control
ICA	Independent Component Analysis
IDM	Infinite Diagonal Model
IHW	Infinite Hofman-Wiggins model
IRM	Infinite Relation Model

KNN	K Nearest Neighbour
MAP	Maximum A posteriori Probability
MI	Mutual Information
MNI	Montreal Neurological Institute
MRI	Magnetic Resonance Imaging
MS	Multiple Sclerosis
NMI	Normalized Mutual Information
rs-fMRI	resting state functional Magnetic Resonance Imaging
RSN	Resting State Network
SVD	Singular Value Decomposition
SVM	Support Vector Machine

Contents

Chapter 1

Introduction

1.1 The central nervous system

The central nervous system (CNS) is part of the nervous system and consists of the brain and the spinal cord. The brain is a complex network, which consists of different regions with specific tasks and functions. The brain can be divided into the cerebrum, cerebellum, diencephalon and the brainstem. Cerebrum is the largest part of the brain, and it consists of a left and a right hemisphere. Each hemisphere is divided into the frontal lobe, parietal lobe, occipital lobe and temporal lobe, see Figure 1.1. The primary motor cortex is a part of the posterior frontal lobe, and the primary sensory area is a part of the post central gyri.

The cerebral cortex is gray matter, which mainly consist of neuronal cell bodies. Small areas of gray matter are found deep inside the brain and these clusters are called nuclei. The white matter is the central part of the cerebrum, between the cortex and the nuclei. It mainly contains myelinated axons, which connect different areas in the cerebral cortex with each other and connect the cerebral cortex with other parts in the CNS [36].
There are connections between neurons inside the regions but also connections between neurons in different regions [45]. The functions in the frontal lobe are problem solving, executive function, motor movements and thinking. In the

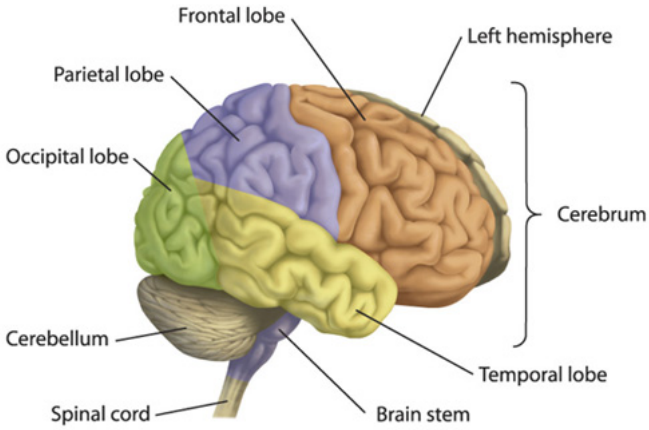

Figure 1.1: The different areas of the brain are shown. The figure is copied from [1].

parietal lobe the function is visuospatial processing. The functions in the temporal lobe is processing auditory and language. The functions in the occipital lobe is visual processing. The functions in cerebellum is controlling motor movements [36]. The brain can be divided into different areas. In this thesis the brain was divided into 116 regions covered in Automated Anatomical Labeling (AAL) [44]. The AAL will be described in Chapter 3.3.

The brain has a very high energy demand. During rest about 17 % of the cardiac output goes to the brain. This is a very large part because the brain only makes up about 2 % of the body weight. About 25 % of the oxygen carried in the blood are consumed by the resting brain. Because the energy reserves are very small and the anaerobic capacity in the brain tissue is very limited, continuous and adequate blood supply to the brain is necessary to sustain the functions[36].

1.2 Multiple sclerosis

Multiple Sclerosis (MS) is a degenerative disease in the CNS and is characterized by acute focal inflammation, demyelination and loss of axons, and chronic multifocal plaques will occur [9]. MS affects women two times as often as men

and is often diagnosed in young people between the ages of 20 and 40. In the industrialized societies it is the most common cause of neurological disability in young adults [48]. The cause of MS is unknown, but the environment and genetic predisposition influence the susceptibility to the disease [43] [9].
The neurological symptoms depend on the location and size of the focal lesion and vary from patient to patient. Some of the symptoms are spasticity, cognitive impairment, visual disruptions and weakness in the extremities [9].
In the early stages of MS the relapsing-remitting pattern is often seen. This means that the patient loses function, for example vision or motor function but the function recovers during remission. After 10-15 years with MS approximately 50 % of the patients has reached a more chronic stage. In this stage the neurological deficits occur cumulative without remission [48] [43].
Myelin is essential for axonal signal conduction. Demyelination and degeneration of axons influences the transmission of the neuronal signal along cortico-cortical and cortico-subcortial connections by delay or disruption. This can induce the neuronal synchronization and functional connection between regions in the brain [48] [43].
Plaques in the white matter have been shown with structural Magnetic Resonance Imaging (MRI). Modern functional Magnetic Resonance Imaging (fMRI) and pathological studies have shown widespread demyelination in the white matter of the brain and spinal cord in patients with MS. These findings are done in MS patients before it is possible to detect plaques in the same area with structural MRI. This indicates that the demyelization affect the brains functional connectivity [16]. To classify the disability in patients with multiple sclerosis the Expanded Disability Status Score (EDSS)[28] is used. The score range from 0 - 10. Where 0 is when a normal neurological examination are done and 10 is when the person is death due to MS. The score can be used to see the progression in the disability for a patient with MS, but it can also be used to group patients with MS when clinical trials are done.

1.3 Functional magnetic resonance imaging

Hydrogen atom nuclei in the body are slightly magnetic and magnetic properties are used when making MRI. A magnetic field B_0 is induced to the subject in the scanner, which partially aligns the nuclei. By inducing radio waves orthogonal to B_0 the magnetic field is rotated. Two different relaxing time scales are used. T_1 is he relaxation time for the longitudinal magnetization, which is the time it takes the magnetic field B_0 to return to equilibrium after the radio waves are induced. The transversal relaxation time T_2, is the time it takes to loss the induced magnetic field. This is due to the difference in spin frequency. The

transverse relaxation time depend on both the phase difference measured in T_2 and the inhomogeneity in the magnetic field. T_2^* is the time constant describing how fast the dephasing happens caused by the inhomogeneity [31].
In fMRI the Blood Oxygenation Level Depend (BOLD) contrast is measured. It is an indirect measure of the neural activity [37] .
The energy demand is increased during neuronal activity in the brain, which results in an increased hemodynamic response. When the supply of oxygenated hemoglobin in the blood is larger than the demand of oxygen, the concentration of oxygenated hemoglobin increases and the concentration of deoxygenated hemoglobin decrease [33]. Deoxyhemoglobin is paramagnetic and distort the magnetic field. Increased Inhomogeneity in the magnetic field, caused by deoxygenated hemoglobin, will make the dephasing faster. The measured signal is proportional to the transversal net magnetization, so an increased inhomogeneity in the magnetic field will attenuate the MRI signal [31].

1.4 Resting state networks

A Resting State Network (RSN) is defined as anatomically separated brain regions, which have a high level of functional connectivity during rest[45]. Known functional networks are represented in most of the RSNs.
In the literature different methods have been used when analyzing resting state rs-fMRI data, with the purpose of detecting brain regions with high connectivity. Seed-based correlation and Independent Component Analysis (ICA) have been the most common methods in fMRI studies[50] but also clustering based analysis have been used [45]. The first method used was seed based analysis. With this method connectivity between the region of interest and the other voxels have been evaluated. When using ICA and clustering based analysis the connectivity between all the regions are analysed. With ICA the independents components must be assigned to known networks [50].
The Default Mode Network (DMN) was the first detected RSN. It was first identified from PET and fMRI data by Raichle et al. [40] in 2001. They found decreased oxygen extraction fraction compared to baseline in DMN regions in the brain during specific tasks. In 2003 Greicius et al. [22] were the first to identify DMN using fMRI by showing significant connectivity in the areas corresponding to DMN.
With task based fMRI it has been shown, that the connectivity in the DMN is decreased compared to resting state. This have not been seen in other reported RSN [45]. The functional meaning of the DMN is related to remembering the past, envisioning the future and mind wandering[7] .
Abnormal activity in the DMN have been shown in patients with Alzheimer's

disease [23] and patients with MS [6].
The DMN during resting state have been reported by Damoiseaux et al. [11], Deluca et al. [12], Pizoli et al. [38] and Van Dijk et al. [47], Beckmann et al. [4] using ICA. The DMN has been confirmed by seed based methods (Greicius et al. [22], Fox et al. [19] and Raichle et al. [39]).
Formations of significant functionally linked brain regions during resting state have been reported from group resting state studies. The regions are referred to as resting state networks. RSN have been found in studies made by Raichle et al. [39], Damoiseaux et al. [11], Greicius et al. [22], Salvador et al. [41], Beckmann et al. [4], Fox et al. [19], Vandijk et al. [47], Laird et al. [29], van den Heuvel et al. [46], Deluca et al. [12] and Pizoli et al. [38]. An overview over these studies is found in Table A.1 and A.2 in Appendix A, where the applied method, the detected resting state networks, how the RSN are described and the material for the experiment are listed. An overlap between the reported RSNs are seen in the literature even though the studies have used different analysis methods [45].
Van den Heuvel et al. [45] have described eight consistently reported resting state networks: primary motor, primary visual, insular-temporal/anterior cingulate cortex (ACC), left parietal-frontal, right parietal-frontal, default mode network and frontal. A graphical illustration of the eight RSN in the brain is shown in Figure 1.2.

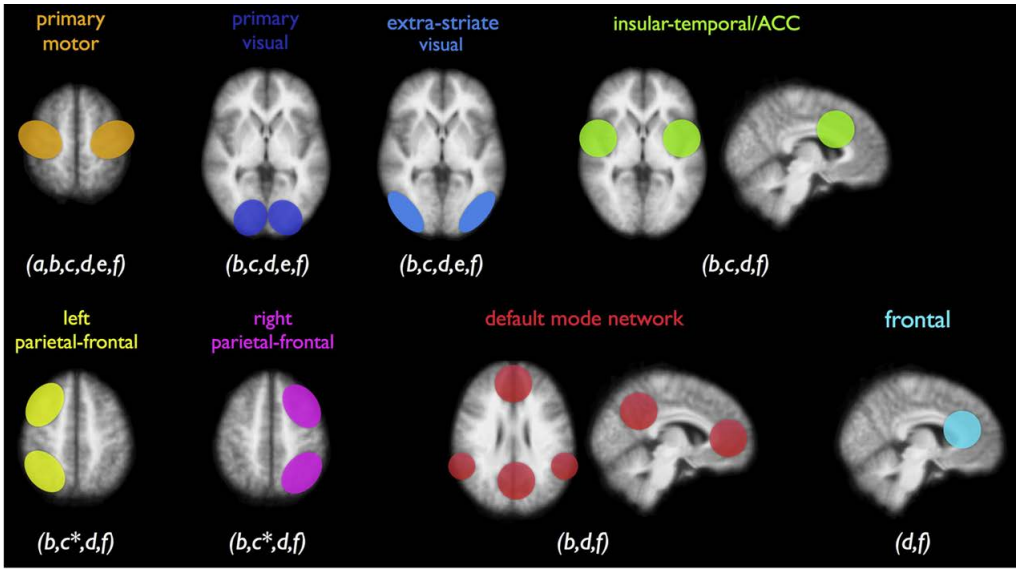

Figure 1.2: Eight resting state networks, which are consistently reported in the literature. The figyre is copied from[45].

1.5 Motivation

MS is a complex disease that can affect all parts of the brain. The symptoms varies from patient to patient. It is a difficult and time consuming job to diagnose a patient with MS using the current diagnosis methods. Early diagnosis makes it possible to start treatment of the disease at an earlier stage, and thereby improve the life quality of the patient. The current methods for diagnosis includes: Clinical examination, MRI, lumbar puncture and blood tests [32]. More knowledge about how the disease affects the brain is required in order to optimize the diagnosis and treatment.
Resting state fMRI studies [6] have shown abnormal activity in the default mode network of MS patients. These results indicates that further investigations of the DMN of patients with MS and healthy subjects can lead to better understanding of how MS affects the communication pattern in the human brain.
The DMN and other RSNs can be detected using the IRM. A method that classifies subjects into patients with MS and healthy subjects could be useful for speeding up the diagnosis process. It is interesting to investigate if the link density data detected by the IRM could be used as a feature vector for classification of subjects.

1.6 This thesis

This thesis seeks to detect RSNs with the IRM and to investigate classification of subjects using link density between communities as a feature vector. The classification seeks to group subjects into two groups: Healthy subjects and subjects with MS.

Chapter 2: The theoretical background of the IRM is described, and it is compared with other blockmodels. Evaluation criteria and classification methods are also described in this chapter.

Chapter 3: Describes the concept of rs-fMRI, the participants and the acquisition and pre-processing of the experimental data. Furthermore the method for making the graph of a brain network is described.

Chapter 4: The detected community structures are presented as a result of in-

ference with the IRM based on correlation. The subjects are classified using the link density between communities as feature vector and correlation with EDSS is estimated.

Chapter 5: The obtained results are discussed and evaluated i comparison with results from the literature.

Chapter 2

Theory

2.1 Networks models

Networks can be described as a system which contain nodes that are linked. The brain is a complex network which consist of regions with different functions. Information are continuously shared between different regions in the brain [45]. Complex networks can also be found in sociology, biology and computer science [17]. Graph theory can be used to examine the properties of complex networks [45].
A network can be represented by a graph G(N,E), where N is the nodes and E is the connection between two nodes, defined as edges or links. In graph theory the adjacency matrix, A represents the connections in the network. The adjacency matrix is binary, and it has nodes in the rows and columns. $A_{i,j} = 1$ if there is a link between node i and j and $A_{i,j} = 0$ if there are no link between the nodes [42] .
Real networks contain patterns of connections between nodes. These patterns can be defined as community structures, which are a relevant feature for graphs representing real systems. Nodes are organized in communities, where the number of links within a community is high and the number of links between communities is low [17].

A network could describe the connection between airports in USA. The nodes in the graph will represent the airports in USA and a link between two airports is defined if there is an airline route between the two airports. This is an undirected graph. The graph is directed if an airline is flying between airport a and b, but no airline flies between b and a. When links are defined as ones in the adjacency matrix the graph is unweighted. For a weighted graph the link between two airports will define that there is a flight between the two airports, and at the same time tells the number of scheduled flights between the two airports. In situations with a weighted graph the adjacency matrix will not be binary.
In this thesis rs-fMRI data is analyzed and the graphs are undirected and unweighted.

2.1.1 Stochastic blockmodel

The optimal partition of nodes into blocks has been a focus area in the literature the last 40 years [21]. This is known as blockmodeling or as detection of community structures.

A stochastic blockmodel is a generative model, where structures in the network are detected[21]. It is also known as a relation model [34]. In a simple stochastic blockmodel the number of communities is predefined as K, and each node is assigned to one of the communities. The links are undirected and independent. The link probability is a function of the community assignment of the nodes. η is a matrix with size K x K where each element is the link probability between two communities, and it is a function of community assignment,. The link probability between two nodes are given by the link probability between the two communities the nodes are assigned to. Z is a matrix of size K x N and it tells which community each of the nodes is assigned to. The column z_i in Z indicates which community the i'th node belong to. A is the adjacency matrix, where links between the nodes in the network is defined. A is a symmetric NxN, undirected and unweighted. The diagonal elements are zero because self-links does not have any interest [26] [42]. In the adjacency matrix the nodes are reordered, so the blocks are shown down the diagonal and the off diagonal is more sparse compared to the diagonal [21].
The probability of the network given the parameters Z and η is known as the likelihood for the network, and is defined as [42]

$$p(A \mid \theta) = \prod_{i,j} p(A_{i,j}|z_i, z_j, \eta) \tag{2.1}$$

where θ is the parameters z_i, z_j and η.
The parameters z_i, z_j and η that maximizes the likelihood, can be determined by integration [8].

2.1.1.1 Bayesian stochastic blockmodel

Bayes theorem is given by [8]

$$p(Y \mid X) = \frac{p(X \mid Y)p(Y)}{p(X)} \tag{2.2}$$

where $p(Y \mid X)$ is the posterior probability distribution, $p(X \mid Y)$ is the likelihood function, p(Y) is the prior, and the denominator is the normalization constant.
Inference is in Bayesian modelling described by probabilities. In this way the degree of certainty about the parameters is quantified. Because of the lack of knowledge about the parameters, they are modelled as random variables. The parameters are assigned a prior probability distribution. This represents the certainty of the model parameters before observing any data. Parameters of the prior distribution are named hyper parameters [42].
In the Bayesian stochastic blockmodel the likelihood function is the same as the one in the stochastic blockmodel (2.1).
The conditional distribution of the parameters given the observed data, known as the posterior, see (2.3), is specified in the Bayesian model when the likelihood and the prior are both known [42].

$$p(z_i, z_j, \eta \mid A) = p(A \mid z_i, z_j, \eta) \cdot p(\eta) \cdot p(Z) \tag{2.3}$$

The likelihood function $p(A \mid z_i, z_j, \eta)$ defines the link probability between nodes, and it is defined as a Bernoulli distribution [42]. The Bernoulli distribution is a special case of the binominal distribution, but with a biased probability. The Bernoulli distribution is given by [8]

$$\begin{aligned} p(A_{i,j} \mid z_i, z_j, \eta) &= \text{Bernoulli}(A_{i,j} \mid z_i \eta z_j) \\ &= (z_i \eta z_j)^{A_{i,j}} (1 - z_i \eta z_j)^{1 - A_{i,j}} \end{aligned} \tag{2.4}$$

$p(\eta)$, in (2.3) is the prior for the parameter η. It tells something about the probability of having a link between two nodes assigned to communities. The prior distribution for the parameter η is the Beta distribution for each pair of communities [42].

$$\begin{aligned} p(\eta_{l,m}) &= Beta(a,b) \\ &= \frac{1}{B(a,b)}(\eta_{l,m}^{a-1})(1-\eta_{l,m}^{b-1}) \end{aligned} \tag{2.5}$$

a and b are the hyper parameters for η. $\eta_{l,m}$ is the link probability between community l and m. When a and b both are set to 1, Beta(1,1) is a uniform distribution [8].

The prior for the parameter Z, $p(Z)$, in (2.3), tells something about how probable it is for a given node to be assigned to all of the K communities. The prior is the Dirichlet distribution, where the probability of being assigned to a community varies from community to community [42].

$$p(Z) = Dirichlet(\alpha). \tag{2.6}$$

α is the hyper parameter for Z.

2.1.1.2 Infinite Bayesian stochastic blockmodel

Until now the models have been based on a finite number of communities. In real complex networks the number of communities are unknown. Thus an infinite model is required to analyze these networks. If the number of communities are larger than the number of nodes, some of the communities will be empty. So in reality the number of communities can maximally be equal to the number of nodes. During the process where nodes are assigned to communities, the number of communities are infinite, but when all the nodes are assigned to a community the number of communities is finite [42].
To make the stochastic blockmodel infinite the prior p(Z) is chosen to be the Chinese Restaurant Process (CRP)

$$p(Z) = CRP(\alpha) \tag{2.7}$$

α is the hyper parameter for Z. In CRP the size of a community affects the probability for assigning a new node to this community. This means that a node has a larger probability for being assigned to a community with many nodes compared to a community with fewer nodes or an empty community.

2.1.2 Infinite relation model

The IRM is a Bayesian generative model. In this thesis it is used to detect community structures in rs-fMRI data [34][27]. IRM is defined by

$$\begin{aligned} Z \mid \alpha &\sim CRP(\alpha) \\ \eta_{l,m}^{(n)} \mid \beta^+, \beta^- &\sim Beta(\beta_{l,m}^+, \beta_{l,m}^-) \\ A_{i,j}^{(n)} \mid Z, \eta &\sim Bernoulli(z_i \eta^{(n)} z_j^T) \end{aligned} \tag{2.8}$$

Z defines which community a given node is assigned to. CRP is the prior for assigning nodes to communities, with α as the hyper parameter. η is a K x K symmetric matrix. Each element, $\eta_{l,m}$ is the probability of having a link between the two communities l and m. It also gives the probability of having a link between node $n_{i,l}$ (node i assigned to community l) and node $n_{j,m}$. The Beta distribution is the prior for the link probability, having $\beta_{l,m}^+$ and $\beta_{l,m}^-$ as the hyper parameters. $\beta_{l,m}^+$ is the pseudo count of links within a community and $\beta_{l,m}^-$ is the pseudo count of nonlinks between communities l and m. β^+ and β^- are vectors with two elements. The first element is the pseudo counts of links and nonlinks respectively within a community and the second element is between the communities. $A^{(n)}$ is the adjacency matrix for graph n and each term in the likelihood function is iterated from the Bernoulli distribution. The Bernoulli distribution depends on the link probability between two communities to which the nodes are assigned [2] [34] [27].

2.1.2.1 Chinese restaurant process

When using CRP as the prior distribution for Z the number of possible communities is infinite. Once all nodes are assigned to a community the number of communities becomes finite. A small number of communities is favored by CRP, and it only creates as many communities as the data warranted [34] [27][35].

The distribution over the communities for node i conditioned on the communities assignment for the rest of the nodes is given by [27]

$$p(z_i = l \mid Z_{\setminus i}) = \begin{cases} \dfrac{n_l}{i-1+\alpha} & n_l > 0 \\ \dfrac{\alpha}{i-1+\alpha} & \text{l is a new community} \end{cases} \tag{2.9}$$

n_l is the number of nodes already assigned to the given community l. It is seen in (2.9) that the probability of assigning a node to a community is dependent on the number of nodes already assigned to the specific community. α describes the probability of assigning a node to an empty community. The probability of assigning a node to an empty community is increased when the value of α increases [34] [27].

For different values of α and β_+ and β_- value set to [1 1] and [1 1] respectively, IRM has been applied a synthetic network generated. The synthetic network contains of 100 nodes, divided into 4 communities with 25 nodes in each. The link probability within the communities is 0.4 and the link probability is 0.1. The density of the graph is 0.17. The adjacency matrix and the Z matrix for the synthetic network is shown in Figure 2.1 and Figure 2.2.

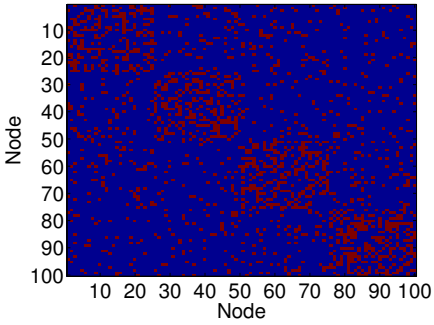

Figure 2.1: The adjacent matrix A for the synthetic network. The number of nodes is 100 and it have four clusters with 25 node in each.

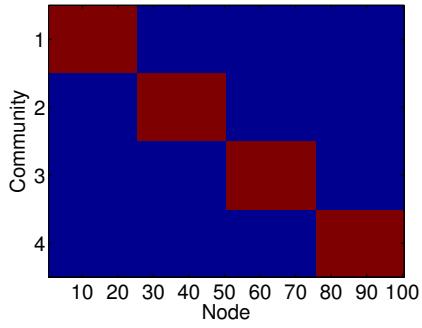

Figure 2.2: The true Z matrix for the synthetic network.

The estimated η matrix for the synthetic network when α is set to 1 is shown in Figure 2.4 and for α set to 40 is shown in Figure 2.3. When α is set to 1

and β is set to [1 1] and [1 1] four communities were detected, which correspond to the true number of communities. For α set to 40, the number of detected communities was 5, which is one more than the true number of communities. So for high values of α the probability of being assigned to an empty community increases, which means that too many communities is detected.

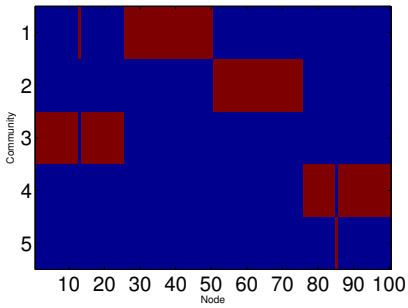

Figure 2.3: The estimated η matrix for the synthetic network when Inference was made with IRM. $\alpha = 40$, $\beta_+ =$ [1 1] and $\beta_- =$ [1 1].

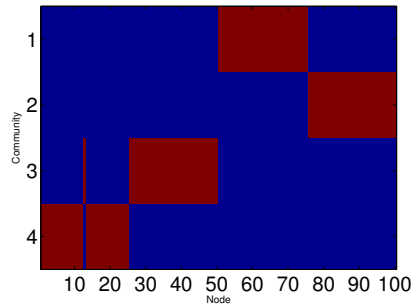

Figure 2.4: The estimated η matrix for the synthetic network when Inference was made with IRM. $\alpha = 1$, $\beta_+ =$ [1 1] and $\beta_- =$ [1 1]

2.1.2.2 Beta distribution.

As mentioned above, the prior for the link probability is the Beta distribution. The hyper parameters β_+ and β_- control the shape of the function.
IRM was applied the synthetic network for different values of β when α was set to 1. With β_+ set to [50 3] and β_- to [3 50] 6 communities were detected. The estimated η matrix is shown in Figure 2.5.

2.1.2.3 Inference

Inference with IRM is based on Gibbs sampling in combination with split-merge sampling[25]. By inference the assignment of nodes to a community and the number of communities is determined.

The Beta distribution is the conjugated prior for the Bernoulli distribution, so η can be integrated out analytically. Which means that the relations are

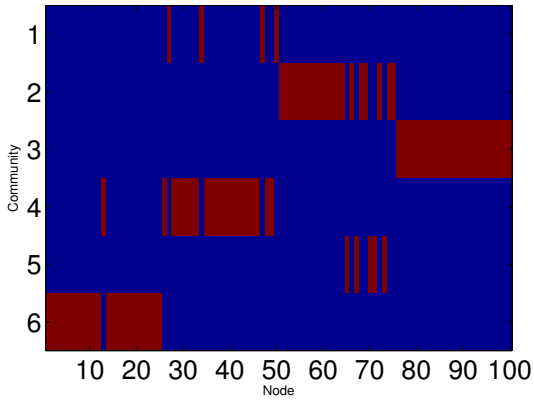

Figure 2.5: The estimated η matrix for the synthetic network when Inference was made with IRM. $\beta_+ = [50\ 3]$, $\beta_- = [3\ 50]$ and $\alpha = 1$

determined by the community assignment in Z.
The derivation is not shown in this thesis, but the analytically solution is given by

$$\begin{aligned} p(A^{(n)} \mid Z, \beta_+, \beta_-) & \\ &= \int P(A^{(n)} \mid \eta^{(n)}, Z) p(\eta^{(n)} \mid \beta_+, \beta_-) d\eta^{(n)} \\ &= \prod_n \prod_{l \geq m} \frac{Beta(n_{lm}^+ + \beta_{lm}^+, n_{lm}^- + \beta_{lm}^-)}{Beta(\beta_{lm}^+, \beta_{lm}^-)} \end{aligned}$$

where n_{lm}^+ is the number of links between community l and m, and n_{lm}^- is the number of nonlinks between community l and m.

The posterior likelihood is then given by

$$p(Z \mid A^{(1)} ... A^N \mid Z, \beta_+, \beta_-, \alpha) \propto (\prod_n p(A^(n) \mid Z, \beta_+, \beta_-)) p(Z \mid \alpha) =$$

$$CRP(\alpha) \prod_n \prod_{l \geq m} \frac{Beta(n_{lm}^{+(n)} + \beta_{lm}^{+(n)}, n_{lm}^{-(n)} + \beta_{lm}^{-(n)})}{Beta(\beta_{lm}^{+(n)}, \beta_{lm}^{-(n)})} \tag{2.10}$$

In Gibbs sampling [8] the assignment of the i'th node is iteratively drawn from the conditional distribution, see (2.11), given the assignment of all the remaining nodes

$$p(z_i = l \mid Z_{\setminus n}, A) \tag{2.11}$$

$Z_{\setminus n}$ is the assignment of all the nodes except z_n.

z_i is replaced by the value sampled from the conditional distribution. This is done for all nodes in each iteration. The first drawn samples will be dependent of the initial state distribution, but after many iterations the samples will be independent of the initial state distribution.

The conditional distribution can be found by evaluation with changes in the likelihood and the prior will happens when node n is assigned for different communities. For n assigned for different communities the posterior is given by

$$p(Z_{il} = 1 \mid Z\_Z_{i_r}, A^1, ..., A^N) \propto \begin{cases} m_a \prod_n \prod_b \frac{Beta(n^+_{lm} + \beta^+_{lm}, n^-_{lm} + \beta^-_{lm})}{Beta(\beta^+_{lm}, \beta^-_{lm})} & if m_l > 0 \\ \alpha \prod_n \prod_b \frac{Beta(n^+_{lm} + \beta^+_{lm}, n^-_{lm} + \beta^-_{lm})}{Beta(\beta^+_{lm}, \beta^-_{lm})} & otherwise \end{cases} \tag{2.12}$$

z_i is the i'th row of Z, N is the number of graphs and m_a is the size of the a'th cluster $m_a = \sum_{j \neq i} Z(j, a)$. The posterior depends on the community size and the number of links and nonlinks between communities, and the β parameter.

To detect community structures in rs-fMRI data inference with IRM is made in this thesis. An adjacency matrix for each subject is given as input and one Z matrix for all subjects have been detected, at the same time an η matrix for each subject is estimated.

2.1.2.4 Other blockmodels IDM, IHW and BCD

Other Bayesian networks models are the simpler Infinite Diagonal Model (IDM) and an Infinite extension of the Hofman-Wiggins model (IHW) proposed by Hofman et al. [24] and the more complex model Bayesian Community Detection (BCD) [35].

The difference between the four models is in the definition of the link probability between communities. In IRM the link probability varies in the upper triangle, see (2.13). In IDM the link probability within a community is different for each community but only one link probability between communities is defined, see (2.14). IHW has two parameters one describing the link probability within communities and one describing the link probability between communities, see (2.15) [35].

$$\eta_{IRM} = \begin{bmatrix} \eta_{1,1} & \eta_{1,2} & \cdots & \eta_{1,m} \\ \eta_{2,1} & \eta_{2,2} & \cdots & \eta_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ \eta_{m,1} & \eta_{m,2} & \cdots & \eta_{m,m} \end{bmatrix} \tag{2.13}$$

$$\eta_{IDM} = \begin{bmatrix} \eta_1 & \eta_0 & \cdots & \eta_0 \\ \eta_0 & \eta_2 & \cdots & \eta_0 \\ \vdots & \vdots & \ddots & \vdots \\ \eta_0 & \eta_0 & \cdots & \eta_m \end{bmatrix} \tag{2.14}$$

$$\eta_{IHW} = \begin{bmatrix} \eta_1 & \eta_0 & \cdots & \eta_0 \\ \eta_0 & \eta_1 & \cdots & \eta_0 \\ \vdots & \vdots & \ddots & \vdots \\ \eta_0 & \eta_0 & \cdots & \eta_1 \end{bmatrix} \tag{2.15}$$

BCD is described in the paper by Mørup et al. [35]. The link probability within a community varies as in IRM. A community gap between 0 and 1 is generated for each community. The maximum of the inter community link probability is defined by multiplying the community gap with the link probability within the community. Next the inter community link probability is generated, so the

value is less than the maximum link probability between communities for each of the two communities [35].

2.1.2.5 Model verification with synthetic data

The IRM model has been evaluated against the models IDM, IHW and BCD. A study like this has been done by Andersen et al.[2], where IRM was evaluated against IDM and IHW.

Synthetic data from each of the four generative models IRM, IDM, IHW and BCD were generated. Datasets containing 100 nodes and 20 adjacency matrices were generated for each model. 500 iterations were made. The parameter α for the CRP was set to $\alpha = 5$, the parameters for the Beta distribution were set to $\beta_+ = [2\ 1]$ and $\beta_- = [3\ 5]$.
The data was randomly split into two sets with equal size, and inference with the four models were made for the two splits separately. This was done 20 times for each model.
To evaluate the models the predictability and reproducibility were estimated.
The test log likelihood was used as a measure for the predictability. To estimate the test log likelihood the Maximum A posteriori Probability (MAP) solution was used. The method is described in Section 2.2.
Mutual Information (MI) was used as a measure for the reproducibility of the identified community structures. The method for estimation MI is described in Section 2.2

The Figures 2.6, 2.8, 2.10 and 2.12 shows the predictability as a function of reproducibility for the four models IRM, IDM, IHW and BCD. Data was generated respectively with IRM, IDM, IHW and BCD. The dotted line in the top of the plots is the log likelihood for the model which has generated the data (the true model). The dotted line in the bottom of the plot is the log likelihood for a random model, which have identical elements in the η matrix (the density in the adjacency matrix). The vertical line in the right side of the plot is the entropy. Entropy is a measure of reproducibility of the model, and it is defined as log(true number of communities) [8].
The Figures 2.7, 2.9, 2.11 and 2.13 shows the histograms of the number of communities estimated by the four models. The dotted vertical line is the true number of communities in the dataset.
In Figure 2.6 data were generated from IRM. The models IDM and IHW have low predictability and reproducibility and BCD have high predictability and re-

producibility. BCD has a higher reproducibility compared with IRM and IRM has the highest predictability compared to BCD.
Figure 2.7 shows that IRM is close to detecting the correct number of communities, BCD detect a few communities more than the correct number. IHW and IDM both detect too few communities. With data generated with IDM, IRM, IHW and BCD all have a high predictability close to the log likelihood for the true model, see Figure 2.8. IDM has the highest reproducibility of the three models. It seems like IRM and BCD almost have the same reproducibility but the scattering in the 20 points are larger with BCD compared to IRM. IHW has a lower predictability compared to the three other models but the highest reproducibility. The reproducibility for IHW is higher than the entropy, for some of the points. In Figure 2.9 it can be seen that too many communities are detected with IHW compared to the true number of communities in the data. The number of detected communities influences the mutual information. Small number of detected communities gives a small mutual information and vice versa.
In Figure 2.10 it can be seen, that all four models have a high predictability when data is generated with IHW. IDM and IHW both have values around the predictability for the true model. IRM has the lowest reproducibility and IHW has the highest reproducibility. The largest scatter in the reproducibility is seen for BCD. IDM and IHW almost detect the true number of communities, IRM detects too few and BCD has a large spread in number of detected communities, but too few communities are detected in most cases.
For data generated with BCD the results for the four models is seen in Figure 2.12. IDM and IHW have the lowest predictability. The predictability for BCD is a little bit higher compared to the values for IRM, which also has a larger spread. One of the IHW data points has a low reproducibility and a relative low predictability compared to the other points. It seems like the IHW model can be unstable. IHW and BCD have the largest reproducibility, and IDM has the lowest. In Figure 2.13 it is seen that BCD detect the correct number of communities most of the time. IRM and IDM detect too few communities and IHW detect too many communities.
The complex models IRM and BCD generates more complex data compared to the model IDM and IHW. It is seen that for data generated with the most complex models IRM and BCD, IRM and BCD have the highest predictability and reproducibility compared the IDM and IHW. For the more simple data generated with IDM and IHW the four models almost have the same predictability, and IDM and IHW have the highest reproducibility.
The results seems comparable with the results presented by Andersen et al. [2], even only 20 number of splits have been made in this thesis.
It seems like IRM and BCD are the best models when analysing complex data. Resting state fMRI data consists of complex networks, and IRM is used to evaluate rs-fMRI data in this thesis.

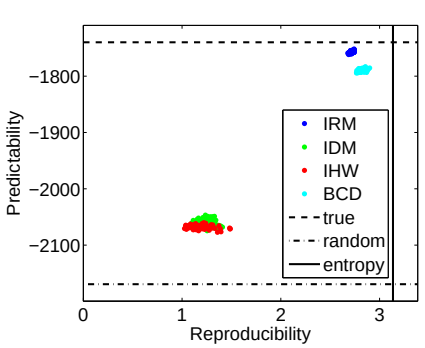

Figure 2.6: Predictability as function of reproducibility. Data was generated with IRM.

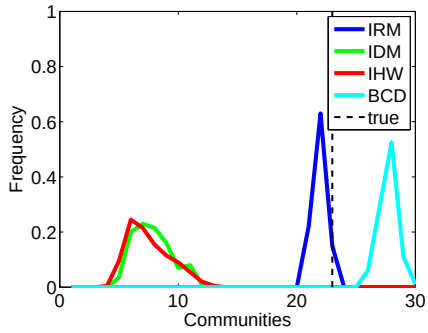

Figure 2.7: Histogram of the estimated number of communities. Data was generated with IRM.

2.2 Evaluation criteria

In this thesis log likelihood, Mutual Information (MI), Normalised Mutual Information (NMI) and Area Under Curve (AUC) of the receiver operating characteristic have been used as evaluations criteria.
The evaluation parameters have been estimated when the data have been divided into two splits (S1, S2) and each split have been inferred with the model independently. It is the MAP solution of the inference which is used.

2.2.0.6 log likelihood

Log likelihood is an estimate for how likely the adjacency matrix from S1 is when the adjacency matrix for S2 is used for inference with the model. The log likelihood is given by [2]

$$logL(Z,\eta \mid A^{S2,(1)},...,A^{S2,(N)}) = \frac{1}{N}\sum_{n=1}^{N}\sum_{j>i} A_{ij}^{S2,(n)} log(z_i \eta z_j^T) + (1 - A_{ij}^{S2,(n)}) log(1 - z_i \eta z_j^T) \quad (2.16)$$

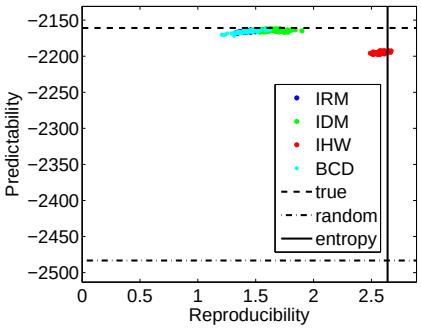

Figure 2.8: Predictability as function of reproducibility. Data was generated with IDM.

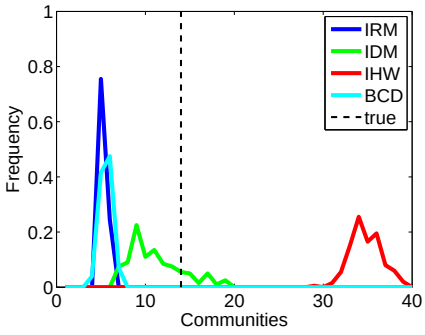

Figure 2.9: Histogram of the estimated number of communities. Data was generated with IDM.

2.2.0.7 Mutual information

MI is used as a measure for the reproducibility of the detected community structure between Z^{S1} and Z^{S2}. MI is defined as [34] [2]

$$MI(Z^{(S1)}, Z^{(S2)}) = \sum_{i=1} \sum_{j=1} p(z_i^{S1}, z_j^{S2}) log \left(\frac{p(z_r^{S1}, z_j^{S2})}{p(z_i^{S1}) p(z_j^{S2})} \right) \tag{2.17}$$

NMI is defined by [26]

$$NMI(Z^{(S1)}, Z^{(S2)}) = \frac{2 * MI(Z^{(S1)}, Z^{(S2)})}{MI(Z^{(S1)}, Z^{(S1)}) + MI(Z^{(S1)}, Z^{(S2)})} \tag{2.18}$$

2.2.0.8 Area under curve

AUC of the Receiver Operating Characteristic (ROC) curve is a measure for how good the model is to predict links [15]. It can be used when the true data is unknown.
The ROC curve is made by plotting the true positive rate vs. the false positive rate. These two performance can be reduced to one value, which represent the expected performance, by estimating the AUC. The value of AUC is between 0

and 1. A random classifier has an area of 0.5.

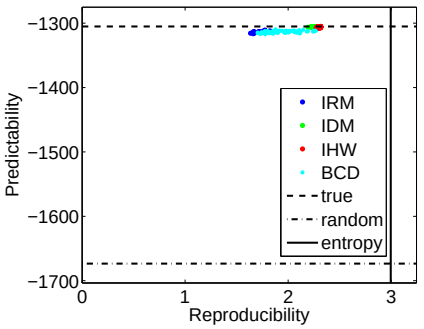

Figure 2.10: Predictability as function of reproducibility. Data was generated with IHW.

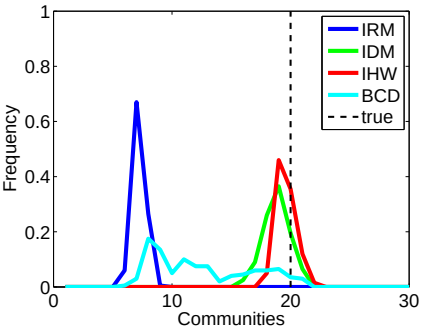

Figure 2.11: Histogram of the estimated number of communities. Data was generated with IHW.

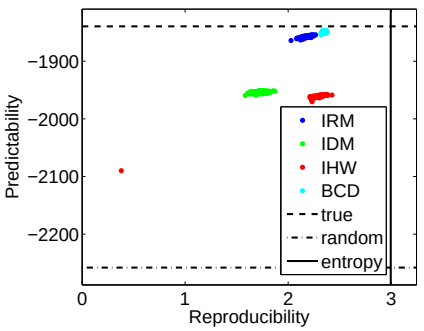

Figure 2.12: Predictability as function of reproducibility. Data was generated with BCD

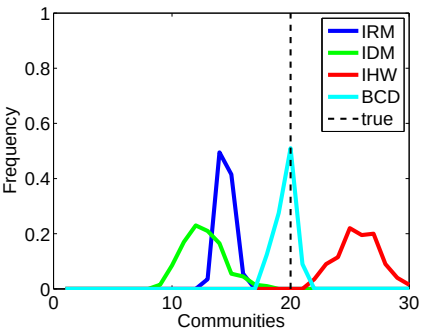

Figure 2.13: Histogram of the estimated number of communities. Data was generated with BCD

2.3 Classification methods

2.3.1 Singular value decomposition

Singular Value Decomposition (SVD) is used to factorize a matrix. SVD is applied for feature extraction. The decomposition of a matrix X (m x n) is given by [14]

$$X = USV' \tag{2.19}$$

Where U is a unitary matrix of size m x m. V is a unitary matrix of size n x n. The components are in the columns. The components are ordered, so the first is the one with the largest variance. S is a diagonal matrix (m x n) with singular values in the diagonal.

In this thesis SVD is used for feature extraction of the correlation and η matrices. Both the correlation matrix and the η matrix are symmetric so only the elements in the upper diagonal have been used as features. In the η matrix the diagonal have also been used. The upper triangle of the matrix have been converted to a vector, and SVD is applied to the vector of each subject.

2.3.2 K nearest neighbours

K Nearest Neighbours (KNN) is a simple algorithm used to classify objects in feature space. The number of classes, K is given, and the distance in feature space is calculated as the Euclidean distance. The test data is classified to the group with the highest representation among the K nearest training data points [8].
In this thesis the η matrix for each subject have been used for the classification. To reduce the number of features SVD is applied to the η matrices before KNN have been used for classification. The first 10 components in V have been used as features to classify the subjects by using KNN.

2.3.3 Support vector machine

Support Vector Machine (SVM) can be used for classification and regression. The parameters are determined by optimization. In this thesis SVM has been used to classify in to two groups: A healthy group and a MS group.
A two class classification problem using linear models can be defined as [8]

$$y(x) = w^T \phi(x) + b \tag{2.20}$$

$\phi(x)$ is the fixed feature space transformation, w is the normal to a hyper plane, b is the bias parameter and x is the training data. Corresponding to the training data a target vector t is given. The values in t are -1 or 1, and tells which class the given object belongs to. The test data is classified according to the sign of $y(x)$. A correct classification satisfy $t_n y(x_n) > 0$ [8].
It is assumed that the class distributions are overlapping in feature space, so it is not possible to make a correct classification for all training data points. Misclassification of some of the training data points have to be allowed [8].

The best separation is found by maximizing the margin and at the same time penalizing points that lie on the wrong side of the margin boundary softly. This is done by minimizing

$$C \sum_{n=1}^{N} \xi_n + \frac{1}{2} \mid\mid w \mid\mid^2 \tag{2.21}$$

ξ_n is the slack variables. It is defined as $\xi_n = 0$ for points on the correct side of the margin boundary and $\xi_n =\mid t_n - y(x_n) \mid$ for other points. C is a soft margin parameter, that controls the trade-off between the slack variable penalty and the margin [8].

Chapter 3

Experiments

3.1 Method of resting state fMRI

With task based fMRI the activated brain area during a given task is measured, from the BOLD constant. In rs-fMRI no task is made. The patient has to lie totally still in the scanner with closed eyes, in some studies the eyes have to be fixed at a cross in the top of the scanner, during the rs-fMRI. The subject is instructed before the scanning in being awake and not doing any kind of voluntary cognitive task or movement [18].
In rs-fMRI it is the spontaneous fluctuation in the BOLD signal which is measured [18]. The regional BOLD-signal oscillates in synchrony and thus provides an index of functional connectivity [5].

In 1995 Biswal et al. [5] found a high correlation between fMRI BOLD time-series in the right and left motor cortex during resting state. Afterwards other studies[10] [11] have replicated these results by showing high functional connectivity between the right and left motor cortex. Also high correlation between other regions in the brain are found [41] [4]. Cordes et al. found high correlation in the auditory and visual cortices [10].
These results show that the spontaneous BOLD activity is not random noise but organized activity in the resting brain [18].

Cordes et al. [10] have shown that physiological noise is low frequent, the respiratory frequency is 0.1 - 0.5 Hz and the cardiac frequency is 0.6 - 1.2 Hz. The fluctuation frequencies of interest in the BOLD signal is below 0.1 Hz, which is below the physiological noise [10] [5].

3.1.1 Participants

The participants in the rs-fMRI study were 42 patients with MS (20 men and 22 women) and 30 Healthy Control (HC) subjects (15 men and 15 women) . The two groups were matching in sex and age. The EDSS were made for each patient to rate the clinical disability. The EDDS range from 0 to 7 with a mean of 4.3 range (0-7). The median age was 45 years in both the HC group and MS group. Only clinical stable patients with more than 3 month since last experienced relapse were included in the study. In the healthy control group only persons without neurological or psychiatric history were included in the study. The patients were recruited from The Danish Multiple Sclerosis Center, Copenhagen, Denmark. The scans were made by Anne-Marie Dogonowski and Kristoffer H Madsen from the Danish Research Centre for Magnetic Resonance at Hvidovre Hospital [13].

3.1.2 Data Acquisition

20 minutes rs-fMRI were performed for each subject, followed by structural MRI at a 3.0 Tesla Magnetom Trio scanner. The rs-fMRI data was recorded with a standard single-channel birdcage head-coil. A T2*-weighted echo planer imaging (EPI) sequence was used with TR = 2490 ms, TE = 30 ms and a 90^o flip angle. 480 whole brain volumes were acquired over 20 minutes (2 contiguous axial slices, slice thickness of 3 mm, FOV = 192x192 mm, 64x64 acquisition matrix, voxel size 3x3x3 mm).
The subjects were instructed to rest with eyes closed and not falling asleep or thinking about anything, and they had to lie still. The patients were asked not to consume caffeine or alcohol and smoking cigarettes six hours before the scanning session. During the scanning the cardiac cycles and the respiratory frequency were monitored. An infra red pulse oximeter were placed at the subjects index finger for monitoring cardiac cycles and pneumatic thoracic belt was used to monitor respiration frequency [13].

3.2 Preprocessing resting state fMRI data

The SPM8[1] software toolbox was used for preprocessing the rs-fMRI data. The steps in the preprocessing were: realignment, slice wise, co-registration, normalization and smoothing, and they will be described briefly in the following. The preprocessing was done independently for each subject.

Realignment and reslice

During a scanning movement of the body will happen. The aim with realignment was to remove movement artefacts [3] so the mean squared difference between the images was minimized.
Before estimation of the transformation parameters the Gaussian Full Width Half Maximum (FWHM) 5 mm, was applied to the images [3].
The translation in the x, y and z direction and the rotation (pitch, roll and yaw) was estimated for each image. A least squares approach and the 6 rigid body spatial transformation parameters were used to realign the time series of the images [3]. The mean image was used as a reference, to which the rest of the images were realigned to. B-spline was chosen for interpolation during the transformation because the error is low. Higher degree interpolation often give a better results but they are slower compared to low degree because more neighbouring voxels are used [3].
After realignment the transformation parameters were saved so they could be used as regressor in the general linear model in the next step slice wise, and the images were resliced so they match the reference image in each voxel [3].

Slice wise

Slice wise[2] is a toolbox for SPM8. The aim with slice wise was to remove the nuisance noise from respiration, pulse and movements. The nuisance noise were identified and defined as nuisance variable regressors. Physical recordings for pulse and respiration were measured during the scan and the movement was identified and the transformation parameters were saved after the realignment was done. The design matrix for the data were made. Figure 3.1 shows the design matrix for subject HC1. It contains one row for each scan and one

[1] http://www.fil.ion.ucl.ac.uk/spm/

[2] Authors: Kristoffer H. Madsen Danish Research Centre for MR: kristoffer m@magnet.drcmr.dk and Torben E. Lund Center for Functionally Integrative Neuroscience (CFIN): torbenelund@mac.com

column for each regressor. The regressors are the time series that were due to the nuisance noise.

The general linear model (GLM) was used to filter the fMRI data. GLM is given by [20]

$$Y = X * \beta + \varepsilon \tag{3.1}$$

Y is the measured fMRI data, ε is noise, β is the real fMRI data and X is the explanatory variables also known as the design matrix [20].
In this thesis the data was rs-fMRI, so the study did not contain any task, and the design matrix only contained nuisance regressors. When subtracting the design matrix from the measured data the residuals are left. The residuals can be defined as the resting state data.

Co-registration

In co-registration the functional images were aligned to the structural image. The mean of the functional images was estimated and the mean functional image and the structural image were co-registered. This was done by estimating parameters which best describe the spatial transformation. By maximizing mutual information the parameters for the best transformation was found [49] [3].

Normalization

The images were normalized to Montreal Neurological Institute (MNI) space. When this is done it is possible to compare different subjects [3]. A Bayesian framework is used to choose the parameters having the maximized posterior probability. [20]
The anatomic image was chosen as source image and normalized to the MNI template. This was done by first fitting the image to the template by a linear transformation and afterword by a nonlinear transformation (wrapping). Afterwards the co-registration functional images were transformed with the same parameters. The voxel size was set to 2x2x2 when the normalized images were written [3].

Smoothing

The aim with smoothing was to minimize the noise and improve the signal-to-noise ratio. An disadvantage was, that the spatial resolution of the image is reduced [3].
The Gaussian FWHM was used to smooth the functional images. The default value of 8 mm in each direction was chosen [3].

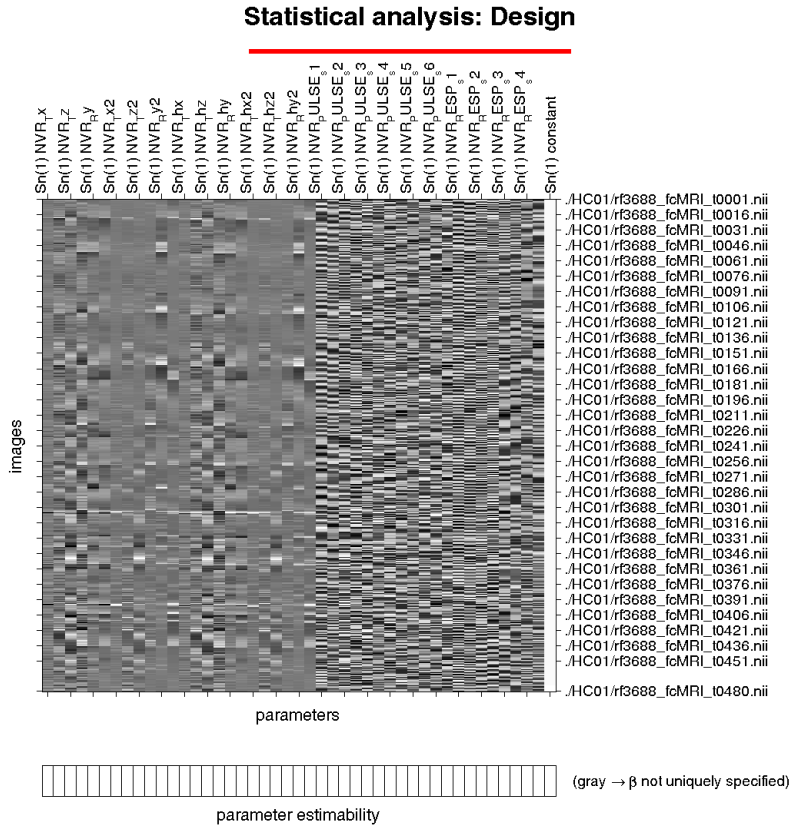

Figure 3.1: The design matrix for subject HC1.

3.3 Automated anatomic labeling regions

In this thesis the Automated Anatomic Labeling (AAL) is used to divide the brain into regions. In 2002 Tzourio-Mazoyer et al. [44] published the Automated Anatomic Labeling (AAL), which divides the brain into 116 regions. Each hemisphere is divided into 45 regions, cerebellum into 18 regions and vermis into 8 regions. The 116 regions predominantly lies in the areas with gray matter. Table 3.1 shows first 78 AAL regions and Table 3.2 shows the last 38 AAL regions.

3.3.1 Default mode network defined in AAL regions

In these thesis the brain has been divided into the 116 AAL regions and the detected communities are defined by the AAL regions. To make it possible to compare the DMN found in the literature [12] [47] [22] [38] [29] [39] [19] [46] with the results in this thesis the DMN found in the literature have been converted to AAL regions. The AAL regions is defined in MNI space, so the given DMN have been converted to MNI space before the AAL regions are defined. In the study by Beckmann et al. [4] the DMN was only described in anatomic areas so it was not possible to convert it to AAL regions.

Table B.1 - B.4 in Appendix B shows the 8 DMN from the literature defined in AAL regions. An AAL region is defined as being a part of the DMN if the region is represented in minimum 4 of the 8 studies. The definition of the DMN can be seen in table 3.3.

3.4 Making the graph

When analyzing rs-fMRI data it is the connections between the regions in the brain that is of interest. So before the rs-fMRI data can be analyzed, a graph has to be made for each subject. Correlation matrices have been used when communities have been found in the literature [41]. In this thesis the graph for each brain network is generated from the correlation matrices.

Number	AAL region	Number	AAL region
1	Amygdala_L	40	Frontal_Inf_Tri_R
2	Amygdala_R	41	Frontal_Med_Orb_L
3	Angular_L	42	Frontal_Med_Orb_R
4	Angular_R	43	Frontal_Mid_L
5	Calcarine_L	44	Frontal_Mid_Orb_L
6	Calcarine_R	45	Frontal_Mid_Orb_R
7	Caudate_L	46	Frontal_Mid_R
8	Caudate_R	47	Frontal_Sup_L
9	Cerebelum_10_L	48	Frontal_Sup_Medial_L
10	Cerebelum_10_R	49	Frontal_Sup_Medial_R
11	Cerebelum_3_L	50	Frontal_Sup_Orb_L
12	Cerebelum_3_R	51	Frontal_Sup_Orb_R
13	Cerebelum_4_5_L	52	Frontal_Sup_R
14	Cerebelum_4_5_R	53	Fusiform_L
15	Cerebelum_6_L	54	Fusiform_R
16	Cerebelum_6_R	55	Heschl_L
17	Cerebelum_7b_L	56	Heschl_R
18	Cerebelum_7b_R	57	Hippocampus_L
19	Cerebelum_8_L	58	Hippocampus_R
20	Cerebelum_8_R	59	Insula_L
21	Cerebelum_9_L	60	Insula_R
22	Cerebelum_9_R	61	Lingual_L
23	Cerebelum_Crus1_L	62	Lingual_R
24	Cerebelum_Crus1_R	63	Occipital_Inf_L
25	Cerebelum_Crus2_L	64	Occipital_Inf_R
26	Cerebelum_Crus2_R	65	Occipital_Mid_L
27	Cingulum_Ant_L	66	Occipital_Mid_R
28	Cingulum_Ant_R	67	Occipital_Sup_L
29	Cingulum_Mid_L	68	Occipital_Sup_R
30	Cingulum_Mid_R	69	Olfactory_L
31	Cingulum_Post_L	70	Olfactory_R
32	Cingulum_Post_R	71	Pallidum_L
33	Cuneus_L	72	Pallidum_R
34	Cuneus_R	73	ParaHippocampal_L
35	Frontal_Inf_Oper_L	74	ParaHippocampal_R
36	Frontal_Inf_Oper_R	75	Paracentral_Lobule_L
37	Frontal_Inf_Orb_L	76	Paracentral_Lobule_R
38	Frontal_Inf_Orb_R	77	Parietal_Inf_L
39	Frontal_Inf_Tri_L	78	Parietal_Inf_R

Table 3.1: The AAL regions element number in the adjacency matrix.

Number	AAL region	Number	AAL region
79	Parietal_Sup_L	98	Temporal_Inf_R
80	Parietal_Sup_R	99	Temporal_Mid_L
81	Postcentral_L	100	Temporal_Mid_R
82	Postcentral_R	101	Temporal_Pole_Mid_L
83	Precentral_L	102	Temporal_Pole_Mid_R
84	Precentral_R	103	Temporal_Pole_Sup_L
85	Precuneus_L	104	Temporal_Pole_Sup_R
86	Precuneus_R	105	Temporal_Sup_L
87	Putamen_L	106	Temporal_Sup_R
88	Putamen_R	107	Thalamus_L
89	Rectus_L	108	Thalamus_R
90	Rectus_R	109	Vermis_10
91	Rolandic_Oper_L	110	Vermis_1_2
92	Rolandic_Oper_R	111	Vermis_3
93	Supp_Motor_Area_L	112	Vermis_4_5
94	Supp_Motor_Area_R	113	Vermis_6
95	SupraMarginal_L	114	Vermis_7
96	SupraMarginal_R	115	Vermis_8
97	Temporal_Inf_L	116	Vermis_9

Table 3.2: The AAL regions element number in the adjacency matrix.

3.4.1 Correlation matrix

To find the connectivity between the 116 regions in the brain, the correlation between each pair of regions have been calculated. This was done by extracting the mean time-series from each of the regions and then calculating the correlation using (3.2).

$$\rho_{ij} = \frac{1}{N}\sum_{n=1}^{N}\frac{(x_{ni} - x_i^-)(x_{nj} - x_j^-)}{(\sigma_i\sigma_j)} \tag{3.2}$$

where ρ_{ij} is the correlation between the time series in the i and j region. N is the length of the time series, x_{ni} is a element in the time series from region i and σ_i is the standard deviation in the i'th time-series [8].

The correlation between a given time-series and it self is 1 because they are perfectly correlated. If two of the time-series were uncorrelated the correlation coefficient would be 0. The diagonal in the correlation matrix was set to 0 [8].

DMN
Frontal_Sub_R
Frontal_Sup_Medial_L
Cingulum_Post_L
Angular_R
Angular_L
Temporal_Mid_R

Table 3.3: Result when evaluate the DMN detected in different studies. The AAL region were defined as a part of the DMN if 4 of more of the studies has the given region in their DMN.

The correlation matrix is symmetric because $\rho_{ij} = \rho_{ji}$. Each element in the correlation matrix is identical to one of the AAL regions. In the Tables 3.1 - 3.2 the element number and the matching AAL region are shown.
The correlation matrix for subject HC1 is shown in Figure 3.2. The plot shows, that the correlation matrix is symmetric and have a diagonal with elements of zeros. A high correlation is seen between the elements 15-26 but also between the elements 27 - 106 with out the elements 69-72 and 87-90. All so the elements 1-6 have a high correlation with the elements 29 - 106.

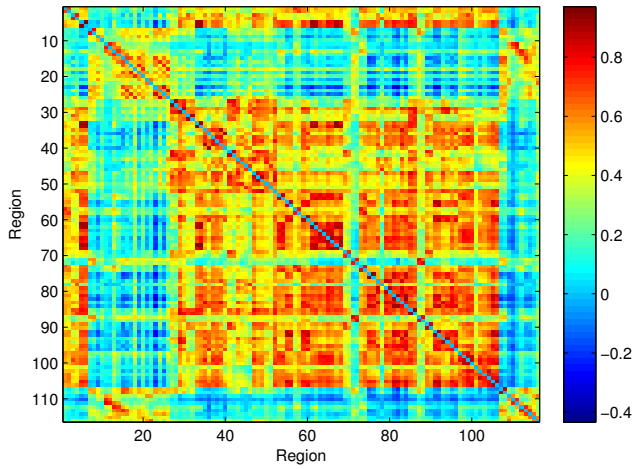

Figure 3.2: The correlation matrix for subject HC1.

To see the variation in the time-series during time for one subject, two correlation matrices for each of the 30 HC subject were made. This was done by

splitting the volumes for one subject (480 volumes) in to two parts, and estimate a correlation matrix for each of the two parts. The two correlation matrices for subject HC1 is shown in Figures 3.3 and 3.4. The pattern seen in Figure 3.2 is also seen in the Figures 3.3 and 3.4. A higher correlation is found in the correlation matrix for the last part of the volumes compared to the correlation matrix for the first part.

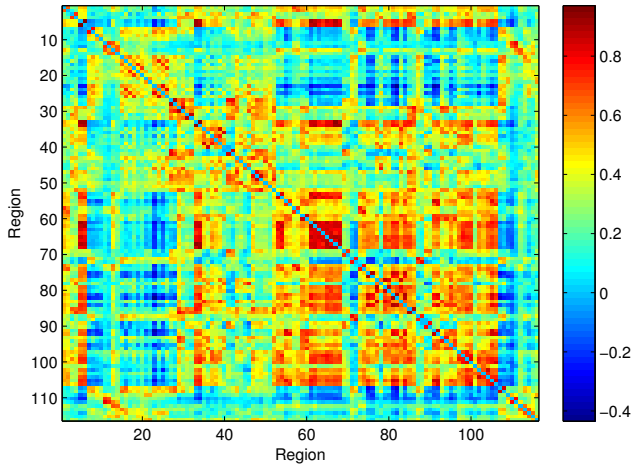

Figure 3.3: The correlation matrix for the first half of the volumes. Subject HC 1.

3.4.1.1 Adjacency matrix

In this thesis the focus was on the fact whether a link between the pairs of regions exists or not, so the graph to generate was undirected and unweighted. This means that the adjacency matrix A was binary and symmetric and has the size 116 x 116. The diagonal in the matrix was zero because we are not interested i self-links.
For each density the number of links in the the Adjacency matrix was estimated. The total number of links are shown in Table 3.4. When generating an adjacency matrix with density 32 % for subject HC1 the number of links in the full matrix has to bee 4268, which is the same as 2134 links in the upper matrix. Because the adjacency matrix was symmetric the upper triangle was created first. A threshold was set, so the 2134 elements with the highest correlation was found,

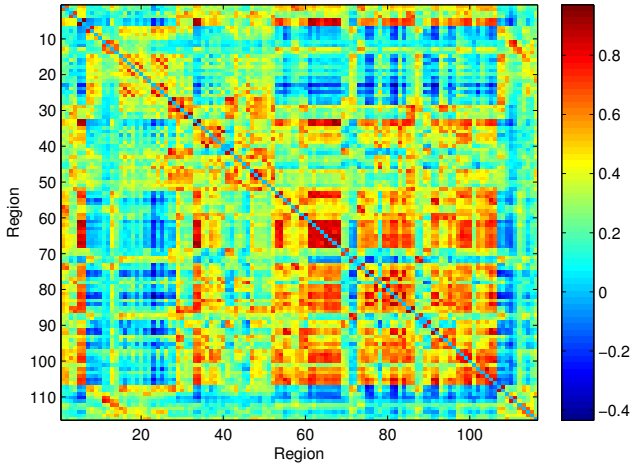

Figure 3.4: The correlation matrix for the second half of the volumes. Subject HC1.

we only look for the positive correlation. The threshold for subject HC1 was set to 0.4978. Figure 3.6 shows the elements in the upper triangle with a correlation value above or equal to 0.4978. These elements were set to one and the rest of the elements were set to zero in the upper triangle with out the diagonal. This gave a binary upper triangular matrix, which is reflected to the lower triangular and the adjacency matrix for subject HC1 was generated, see Figure 3.7. The same was done for the HC and MS subjects for all the six densities. The Adjacency matrices were made with a density of 2 %, 4 %, 8 %, 16 %, 32 % and 50 % for all the subjects. The mean threshold set for the HC and MS group at different densities and the standard deviations are shown in Figure 3.5. As expected it can bee seen, that the threshold value decreases when the density increases. The mean threshold for the HC group and MS group were relatively equal for all the density.
For the six densities two adjacency matrices were also estimated using the two correlation matrices for each of the HC subjects.

density	number of links
2	266
4	534
8	1068
16	2134
32	4268
50	6670

Table 3.4: The number of links in the adjacency matrix for different densities.

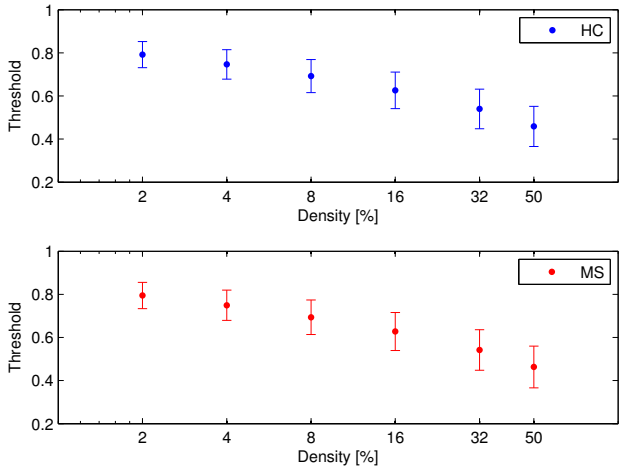

Figure 3.5: The mean threshold set for the HC and MS group at different densities when generating A and the standard deviation.

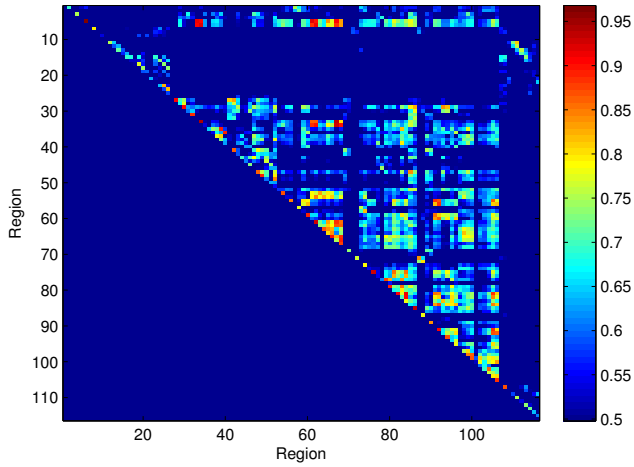

Figure 3.6: The elements in the upper triangle of the correlation matrix with a correlation value above or equal to 0.4978. The density was 32 % and the correlation matrix was for subject HC1.

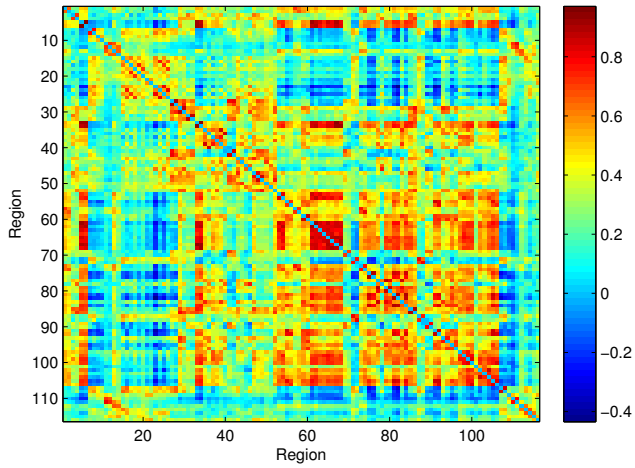

Figure 3.7: The adjacency matrix for subjects HC1 with density 32.

Chapter 4

Data analysis

4.1 Correlation matrices

A correlation matrix for each of the 72 subjects have been estimated. The mean of the correlation matrices for the HC subjects and the standard deviation in each element are shown in Figure 4.1 and Figure 4.2. Figure 4.3 and Figure 4.4 shows the mean of the correlation matrices for the MS subjects and the standard deviation in each element.
The mean correlation matrices for the HC subjects and the MS subjects almost look alike. In both plots a high correlation is seen between the regions 15-26, 61-68 and 109-116. Regions 15-26 are parts of cerebellum, region 61-68 is in the occipital lobe. 109-116 is vermis, which is a part of cerebellum. A general low correlation with other regions is found for region 87 and 88, which are the right and left putamen, region 89 and 90, which are the right and left part of rectus and region 109-11 which are parts of vermis.
In both plots of the standard deviation it can be seen that the highest values are found in areas with a low correlation. The plot of the standard deviation between MS subjects, Figure 4.4, is a little bit brighter compared to the plot for the HC subjects, Figure 4.2, so the variation between the subjects seems to be higher for the MS subjects compared to the HC subjects.

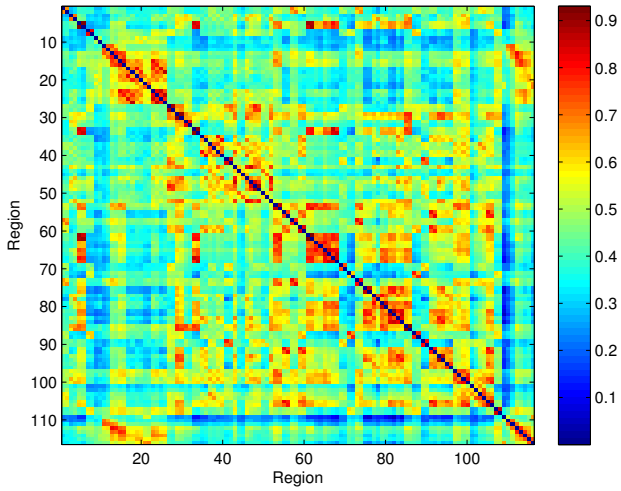

Figure 4.1: The mean of the correlation matrices for the HC subjects.

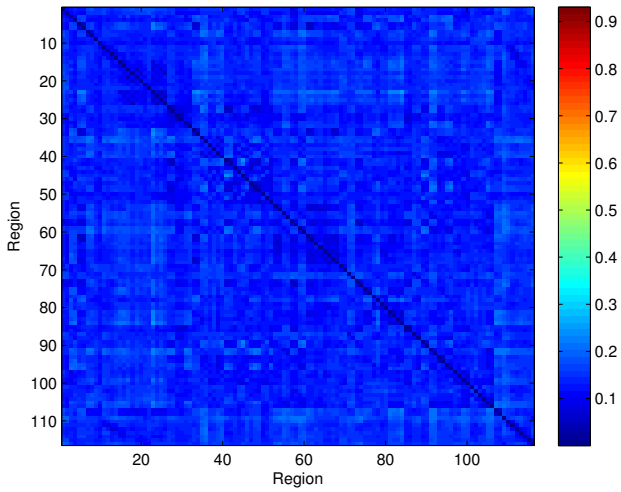

Figure 4.2: The standard deviation for each element in the mean correlation matrix for the HC subjects.

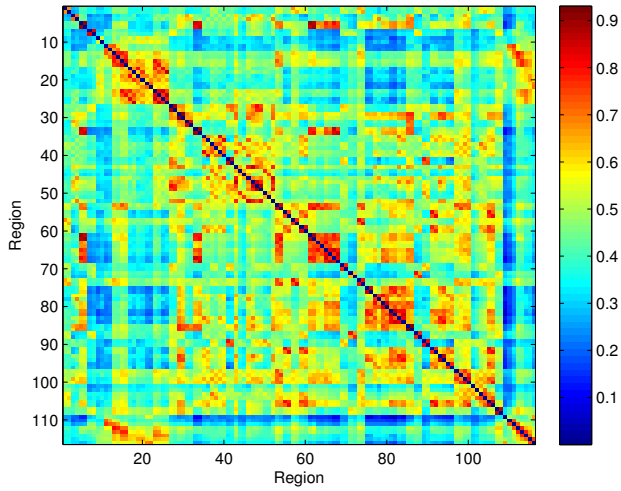

Figure 4.3: The mean of the correlation matrices for the MS subjects.

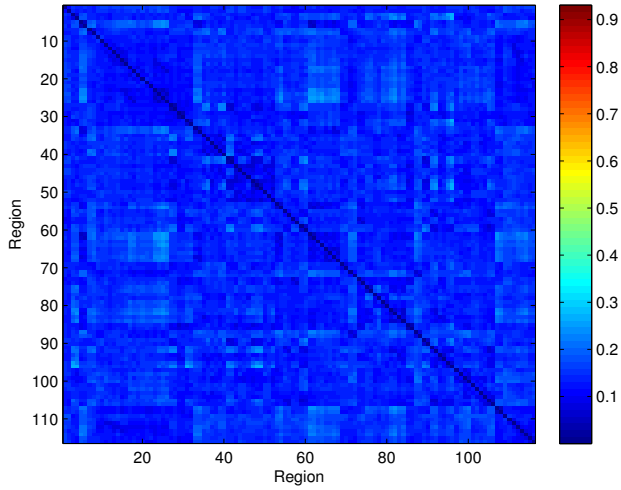

Figure 4.4: The standard deviation for each element in the mean correlation matrix for the MS subjects.

4.1.1 Singular value decomposition for the correlation matrices

The rs-fMRI data for each subject consists of 480 volumes (20 min). To examine the variation across the volumes, SVD was applied to the two correlation matrices for each of the 30 HC subjects.
The correlation matrix for the first part of the volumes are named corr1 and the correlation matrix for the last part of the volumes are named corr2. Figure 4.5 shows the first and second component plotted against each other from the decomposition of the two correlation matrix for each subject. Each dot is numbered, and the number refer to the subject. The distance between the red and blue dots from the same subject is small. It is not possible to separate the corr1 matrices from corr2 matrices, when looking at the first component plotted against the second. The same is seen for the other components, which are not shown. The distance between the two correlation matrices from one subject is in general smaller than the distance between correlation matrices from two subjects. It seems like the robustness for the two correlation matrices for one subject is high.

In this thesis it is examined if it is possible to discriminate the HC and MS subjects. SVD was applied to the correlation matrices from the HC and MS subjects, to see if it was possible to distinguish the subjects from each other using the correlation matrices as features. The correlation matrices is symmetric with zeroes in the diagonal, so only the elements in the upper triangle without the diagonal were converted to a vector for each subject and used for the SVD. The decomposition was made for a matrix, which consist of the vectors for each subject.
The results are shown in Figure 4.6 where the first and second component, the first and third component, the first and fourth component and the second and third component are plotted against each other respectively. The HC and MS subjects form a complex pattern and it is not possible to separate the two groups from each other by apply SVD to their correlation matrices.

4.2 Evaluation of IRM when varying hyper parameters

IRM has been implemented in MATLAB by Morten Mørup[1], and those scripts have been used in this thesis. It was necessary to understand the implementa-

[1] http://www.mortenmorup.dk/MMhomepageUpdated_files/Page327.htm

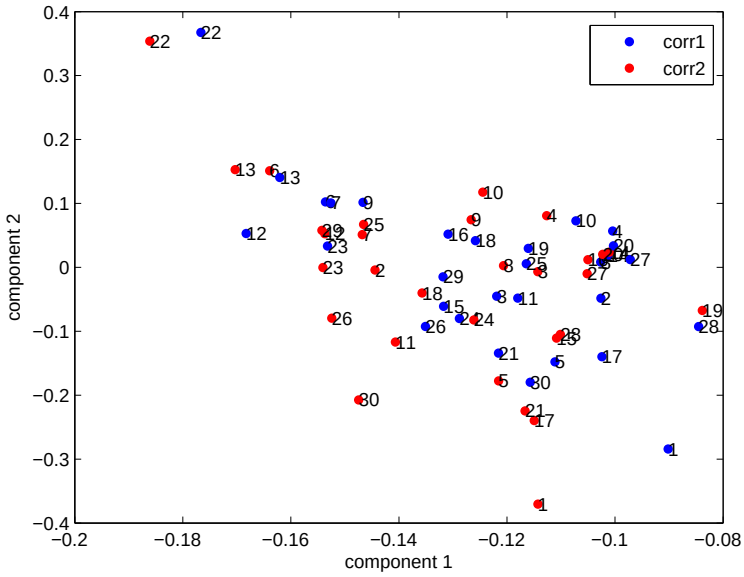

Figure 4.5: SVD was applied to the two correlation matrices for each of the 30 HC subjects. The first and second component, from the decomposition are plotted against each other. The blue dots are the correlation matrices for the first half of the volumes in the dataset and the red dots are the correlation matrices for the last part of the volumes. Each dot is numbered and refer to the HC subject.

tion before the scripts have been used.

The hyper parameters α and β had to be set before IRM could be used. It was unknown which values would be a good choice when analyzing rs-fMRI data. Three different values of α and β were tested. The hyper parameters were varied independently of each other. The α values were 0.1, 1 and 10 and the β values were [1 1], [1 0.1] and [5 1] for β_+ and [1 1], [0.1 1] and [1 5] for β_-. When α was varied the β value was set to $\beta_+ = [1\ 1]$ and $\beta_- = [1\ 1]$, and when β was varied α was set to 1. When the hyper parameters are set to 1 they are neutral. The data used to evaluate IRM for different hyper parameters were the two adjacency matrices for the HC subjects. By random the adjacency matrices were divided into two splits S1 and S2. Each split contained 30 adjacency matrices, one from each subject. Half of the adjacency matrices were from the first part

of the volumes and the other half were from the last part of the volumes.
Inference with IRM was made for S1 and S2 separately. The number of iterations was 500, the initial guess of number of communities was set to 50. Ten splits were made for each value of α and β and it was done for the densities 2, 4, 8, 16, 32, 50 %. One Z matrix and a η matrix for each subject are estimated for each split.

Reproducibility, predictability, link prediction, and the number of detected communities are used as evaluation parameters.
The reproducibility is given by the mutual information and normalised mutual information, which were estimated for the Z matrices from S1 and S2 in each of the 10 splits.
Log likelihood is the estimate for how good IRM is to predict data. The log likelihood for the adjacency matrix in S1 when the adjacency matrix for S2 is used for inference with IRM, is estimated for each of the adjacency matrices in S1 and S2 for each split.
AUC is a measure for link prediction. AUC is estimated for each adjacency matrix in S1 and S2 for each of the 10 splits.
The mean value of MI and NMI for the 10 splits for each value of the hyper parameters is estimated for each of the six densities and the results are shown in Figure 4.7 when varying α and Figure 4.11 when varying β.
The mean of the log likelihood for all the adjacency matrices in S1 and S2 respectively is estimated. The mean of this values for the 10 splits and the standard deviation are estimated and the results for each density are shown in Figure 4.8, when varying α and Figure 4.12, when varying β. The left column in the plots are the results from S1 and the right column are the results from S2. One value of the hyper parameter are represented in each of the rows.
AUC for each of the adjacency matrices in S1 and S2 respectively were estimated and the mean value was found. The mean of this mean value in each split and the standard deviation are estimated for each density and the results are shown in Figure 4.9, when varying α and Figure 4.13, when varying β. The mean of the number of clusters detected in S1 and S2 respectively for each of the 10 runs and the standard deviation for each density are shown in Figure 4.10, when varying α and in Figure 4.14, when varying β. The plots are divided as described for the plots of the log likelihood.

Varying the hyper parameter α

In Figure 4.7 the MI and NMI are shown. The MI increases when the density increases for all the α values. The NMI increases except between density 4 % and 8 % where it decreases. The NMI is higher for density 4 % compared to

density 16 %. This is seen for all values of α.
For all values of α the log likelihood decreases when the density increases. The difference between the log likelihood for different α values is small, and the results for the two splits are not consistent for all the densities, see Figure 4.8. In Figure 4.9 it can be seen that the highest AUC is found for the densities 8, 16, 32 and 50 % for all values of α. The smallest value is found for density 2 %. Only a small difference in AUC values are seen for the different values of α.
In Figure 4.10 it can been seen that the number of detected communities is correlated to the density. When the density increases the number of communities increases.
The model is stable for the three different α values. An α value of 1 is chosen.

Varying the hyper parameter β

When varying the hyper parameter β it can bee seen in Figure 4.11 that the MI increases when the density increases for the values $\beta_+ = [1\ 1]$ and $\beta_- = [1\ 1]$. For $\beta_+ = [1\ 0.1]$ and $\beta_- = [0.1\ 1]$ the highest MI value is achieved for density 16 % and for $\beta_+ = [5\ 1]$ and $\beta_- = [1\ 5]$ the highest value is achieved at a density of 32 %. The results for $\beta_+ = [1\ 1]$ and $\beta_- = [1\ 1]$ is the same as for $\alpha = 1$. The NMI has a maxima at 0.94 with a density of 16 % for the parameter values $\beta_+ = [1\ 0.1]$ and $\beta_- = [0.1\ 1]$. For the parameter values $\beta_+ = [5\ 1]$ and $\beta_- = [1\ 5]$ the NMI curve has two maxima, one with a density of 4 % with a value of 0.91 and one with a density of 32 % with a value of 0.90.
Figure 4.12 shows a correlation between the log likelihood and density. The log likelihood decreases when the density decreases for all the β values. The standard deviation is largest for the high densities 32 % and 50 %. The standard deviation is larger in the first split compared to the second split. The highest log likelihood values are found for the densities 2 % - 16 % with $\beta_+ = [1\ 0.1]$ and $\beta_- = [0.1\ 1]$ and for densities 32 % and 50 % with the parameter values $\beta_+ = [5\ 1]$ and $\beta_- = [1\ 5]$.
Figure 4.13 shows that for $\beta_+ = [1\ 1]$ and $\beta_- = [1\ 1]$ the greatest AUC values are found for the three largest densities, and the values are in S1 0.915, 0.918 and 0.917 and in S2 0.9197, 0.9196 and 0.9178. For $\beta_+ = [1\ 0.1]$ and $\beta_- = [0.1\ 1]$ the AUC values are quite equal for the densities 2 %, 4 % and 8 %. For the higher densities the values decrease. This is seen in both S1 and S2. For $\beta_+ = [5\ 1]$ and $\beta_- = [1\ 5]$ the maximum AUC value are found with a density of 16 % in S1 and 8 % in S2. The highest AUC values are found for $\beta_+ = [1\ 0.1]$ and $\beta_- = [0.1\ 1]$ for the densities 2 % - 32 % and for $\beta_+ = [1\ 1]$ and $\beta_- = [1\ 1]$ with a density of 50 %. The largest standard deviation is found with density 2 % for all the values of β.
Figure 4.14 shows the number of detected communities. For the values $\beta_+ = [1$

1] and $\beta_- = [1\ 1]$ there is a positive correlation between the number of communities and the density. For the values $\beta_+ = [1\ 0.1]$ and $\beta_- = [0.1\ 1]$ the number of communities has a maxima at the densities 8 % and 16 %, where the mean value is 34.5 and 34.9 in the S1. In S2 the number of communities have the highest value with a density of 16 %. For $\beta_+ = [5\ 1]$ and $\beta_- = [1\ 5]$ the highest number of communities are detected with the densities 16 % and 32 % in both splits.
The hyper parameter values for β are chosen to $\beta_+ = [1\ 0.1]$ and $\beta_- = [0.1\ 1]$, because IRM has a high reproducibility and predictability with these values. Also the link predictability is high for these parameter values. With the choice of β values the priori favoured a higher within community link density compared to the between link density.

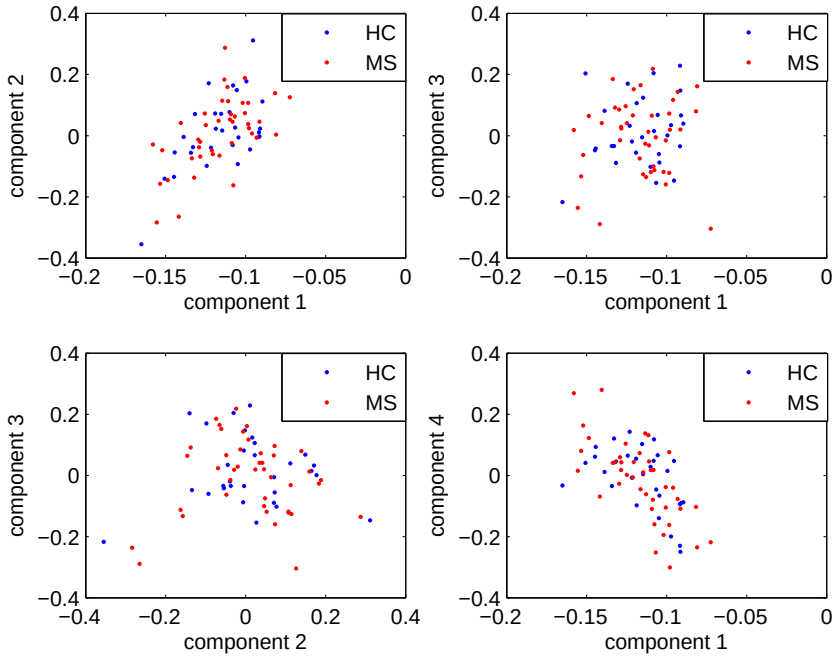

Figure 4.6: The result from SVD applied the correlation matrices for HC and MS subject. Top left: The first component plotted against the second component. Top right: The first component plotted against the third component. Bottom left: The second component plotted against the third component. Bottom right: The first component plotted against the fourth component. The blue dots are the HC subjects and the red dots are the MS subjects.

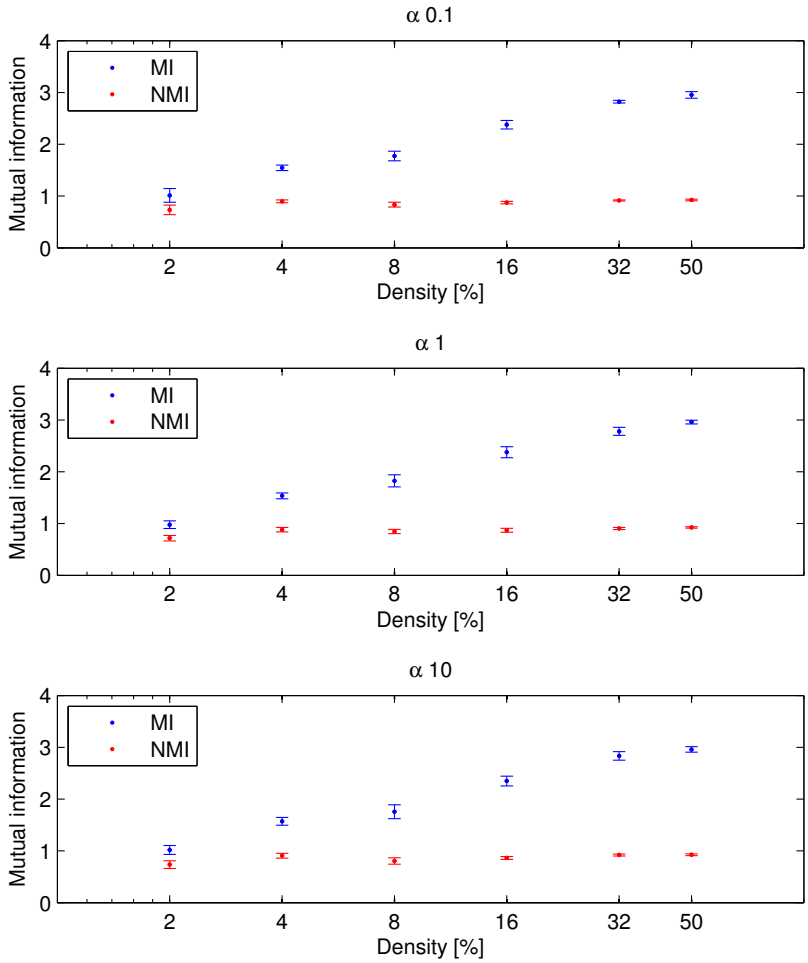

Figure 4.7: Mean value of mutual information and normalized mutual information were estimated for the Z matrices from S1 and S2 in each of the 10 splits. When α was varied between 0.1 1 and 10. β was set to 1. The mean of mutual information and normalised mutual information of the 10 splits and the standard deviation for each density is shown.

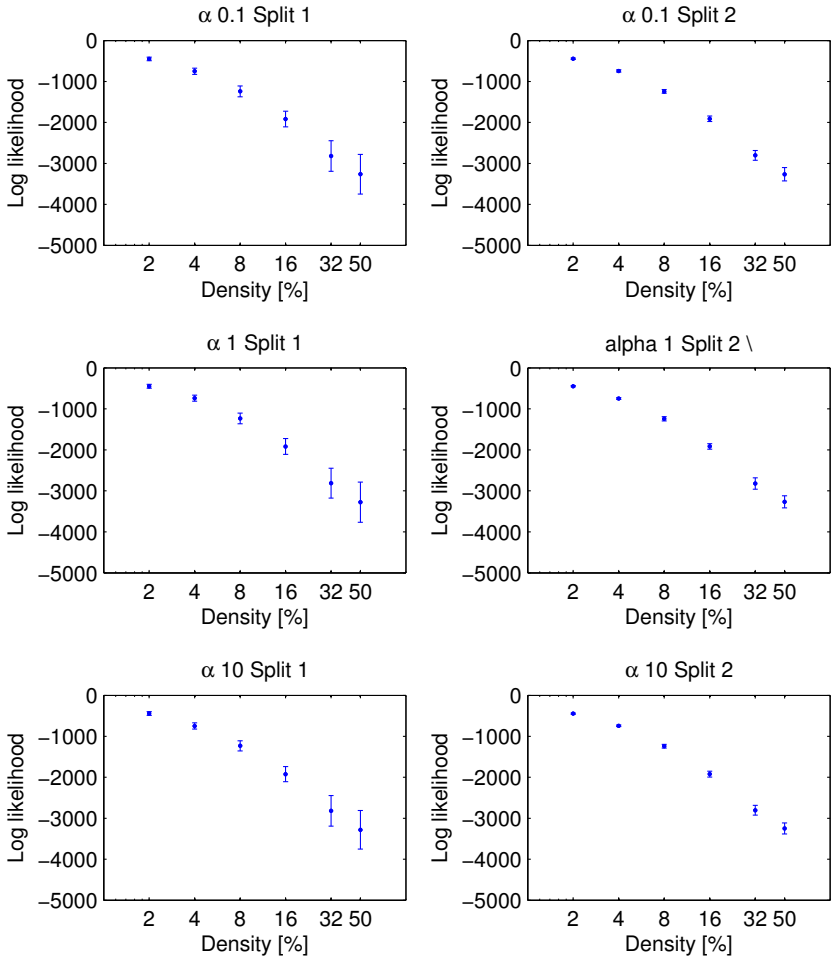

Figure 4.8: The mean and the standard deviation in S1 and S2. The mean over the subjects mean log likelihood over the 10 run for each densities and the standard deviation are shown, when α was varied between 0.1, 1 and 10. β was set to 1. The results from split 1 is in the left column and for split 2 in the right column.

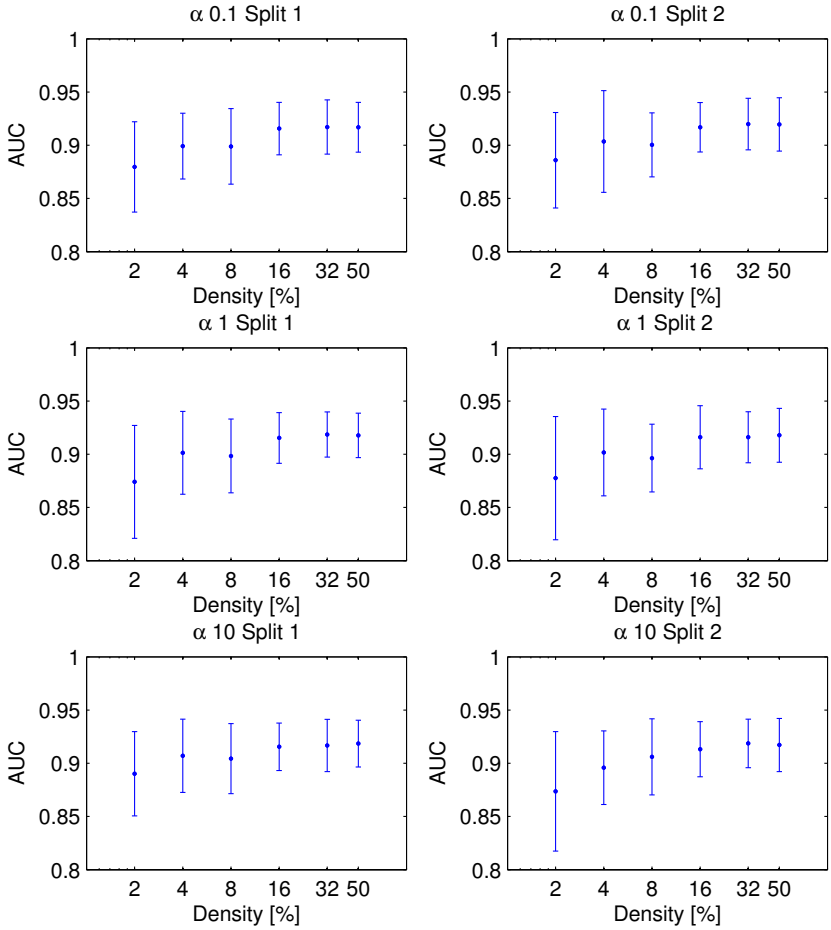

Figure 4.9: AUC for the two splits are calculated for each of subjects in the 10 inference with IRM. The mean over the subjects mean AUC over the 10 run, for each densities and the standard deviation are shown, when α was varied between 0.1, 1 and 10. β was set to 1. The results from split 1 is in the left column and for split 2 in the right column.

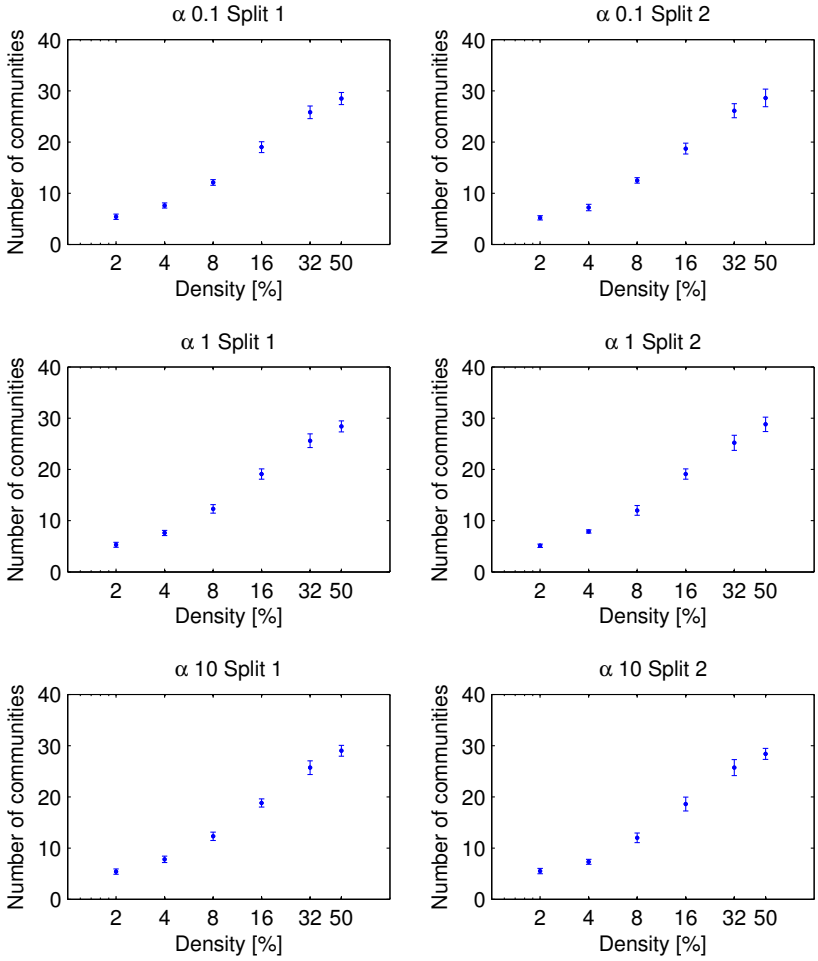

Figure 4.10: Number of communities detected when making inference with IRM 10 times. The mean number of communities over the 10 runs, fore each density and the standard deviation are shown. α was varied between 0.1, 1 and 10. β was set to 1. The results from split 1 is in the left column and for split 2 in the right column.

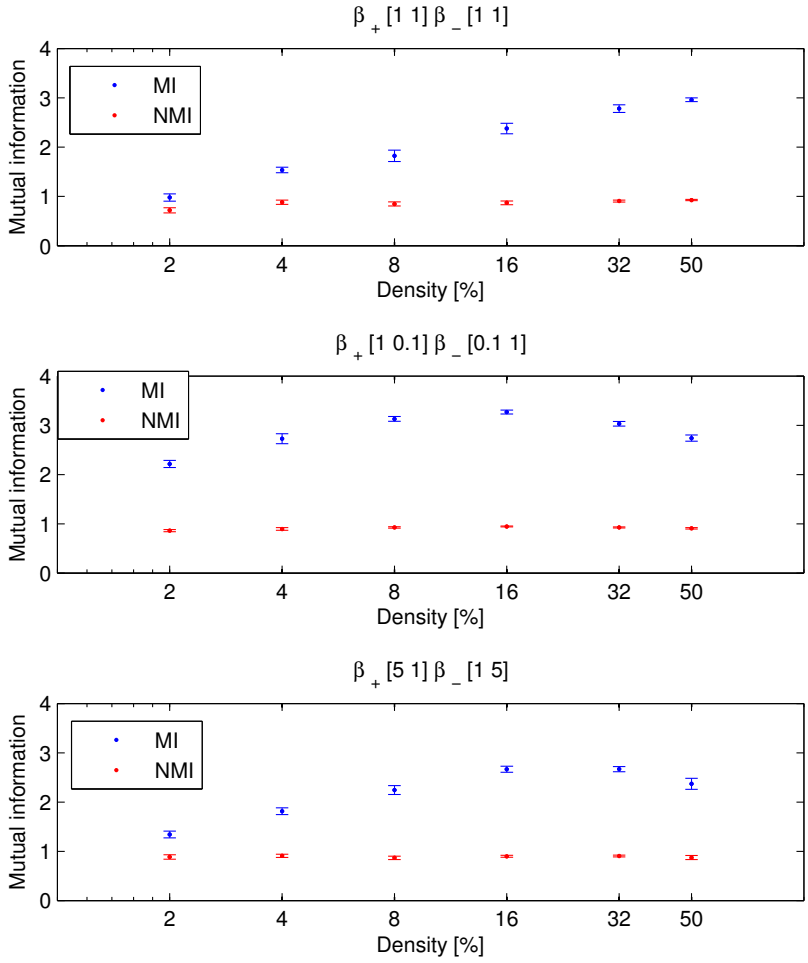

Figure 4.11: Mutual information and normalized mutual information were estimated for the Z matrices from the two splits in each of the 10 inference runs with IRM. When the hyper parameter β_+ was varied between [1 1], [1 0.1] and [5 1] and β_- was varied between [1 1], [0.1 1] and [1 5]. α was set to 1.

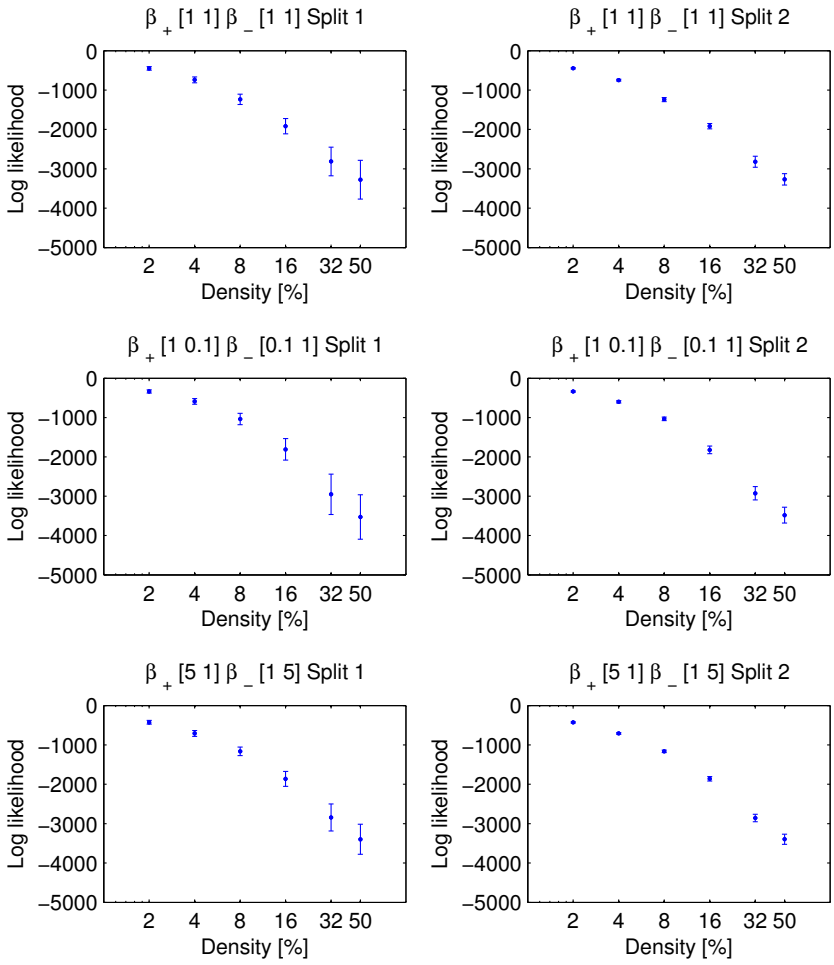

Figure 4.12: The mean over the subjects mean log likelihood over the 10 runs for each density and the standard deviation are shown. β_+ was varied between [1 1], [1 0.1] and [5 1] and β_- was varied between [1 1], [0.1 1] and [1 5]. α was set to 1. The results from split 1 is in the left column and for split 2 sin the right column.

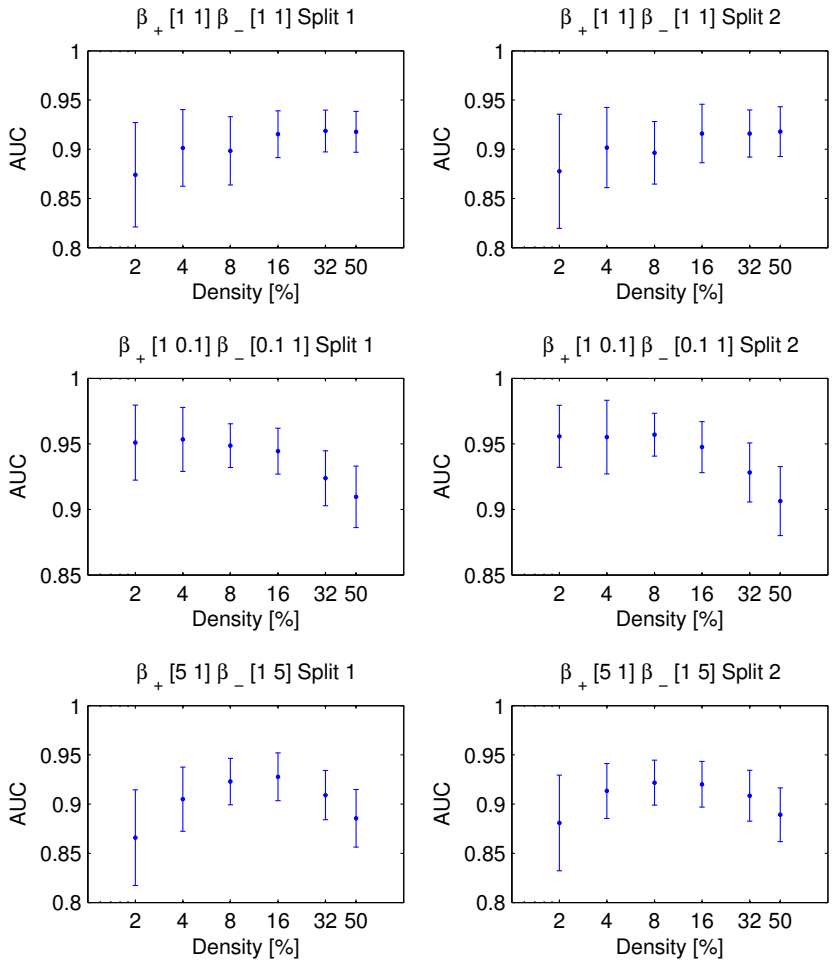

Figure 4.13: AUC for the two splits are calculated for each subjects in the 10 inference runs with IRM. The mean over the subjects mean AUC over the 10 runs for each density and the standard deviation are shown. β_+ was varied between [1 1], [1 0.1] and [5 1] and β_- was varied between [1 1], [0.1 1] and [1 5]. α was set to 1. The results from split 1 is in the left column and for split 2 sin the right column.

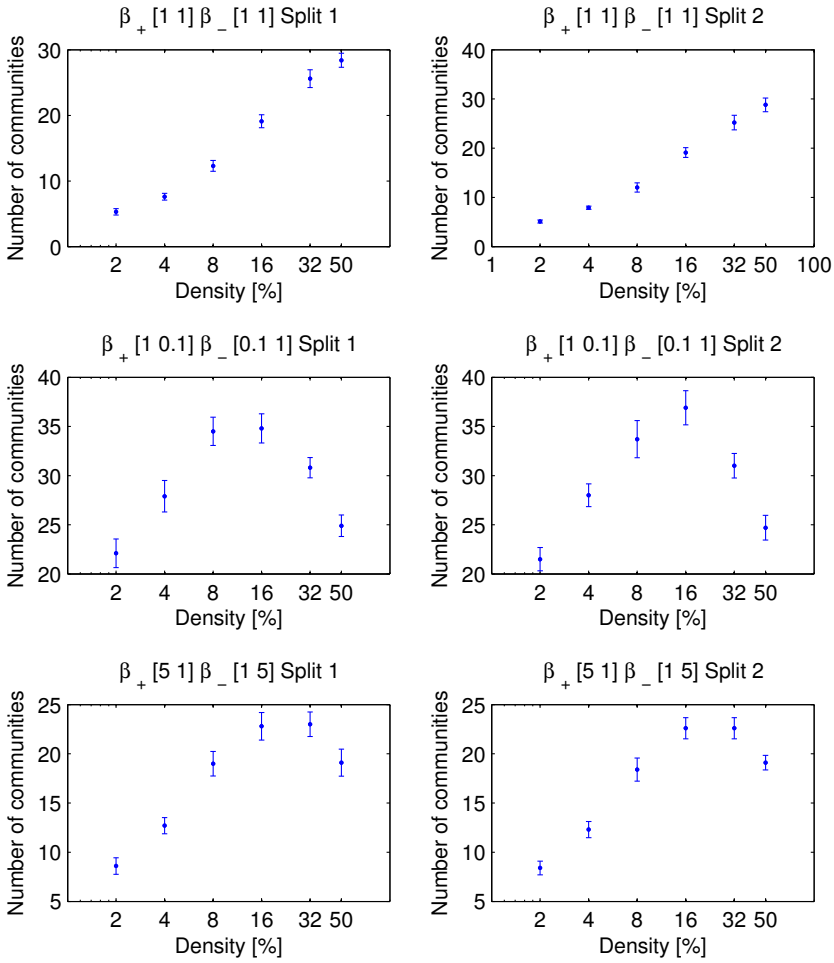

Figure 4.14: Number of communities detected after 10 inference runs with IRM. The mean over number of clusters detected in the 10 runs for each density and the standard deviation are shown. β_+ was varied between [1 1], [1 0.1] and [5 1] and β_- was varied between [1 1], [0.1 1] and [1 5]. α was set to 1. The results from split 1 is in the left column and for split 2 sin the right column.

4.3 Inference with IRM

When inference with IRM is made in the rest of this thesis, the hyper parameters are set to $\alpha = 1$,$\beta_+ = [1\ 0.1]$ and $\beta_- = [0.1\ 1]$. The values of the hyper parameters are found in the section above where IRM was evaluated for different values of the hyper parameters. The number of iterations is set to 500 and the initial guess of number of communities is set 50. The iteration with the highest log likelihood is chosen as the maximum a posteriori probability (MAP) solution. One community assignment matrix Z and an adjacency matrix for each subject are estimated for each run.
Inference with IRM is made for the adjacency matrices from the subjects. 20 runs are made for each of the six densities 2, 4, 8, 16, 32 and 50 % of the adjacency matrices, and the run with the highest log likelihood for each density is defined as the MAP run for the given density.

Table 4.1 shows the mean of the number of detected communities and the standard deviation. The largest number of communities is found when the density is 16 %. The log likelihood decreases when the density of the networks decreases.

The community structures are detected for the HC and MS subjects by inference with IRM. Figure 4.15 shows the log likelihood for density 32 % and run number 11. The burn in period is seen in the first 15 iterations, after the first 15 iterations the curve begin to flatten off. It seems like the log likelihood is relative stable when 500 iteration are made during inference.

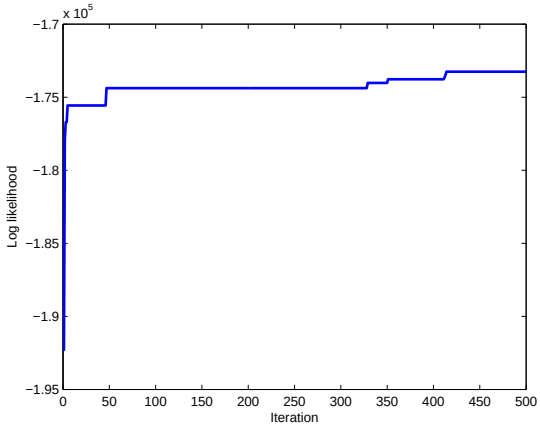

Figure 4.15: Log likelihood during inference with IRM in run number 11 for density 32 %.

	Number of communities	
Density [%]	mean	std
2	21.15	± 1.35
4	29.3	± 2.11
8	34.2	± 1.85
16	35.75	± 1.80
32	30.55	± 1.47
50	26.25	± 1.02

Table 4.1: The Mean of the detected number of communities for each density and the standard deviation.

For density 32 %, MAP run 11 the estimated η matrices for the HC subjects is shown in Figure 4.16 and the standard deviation in each element of the matrix is shown in Figure 4.17. The mean of the estimated η matrices for the MS subject is shown in Figure 4.18 and the standard deviation in each element is shown in Figure 4.19. Only small differences are seen between the two mean η matrices. Also the standard deviation in each element in the two η matrices seems relative equal.

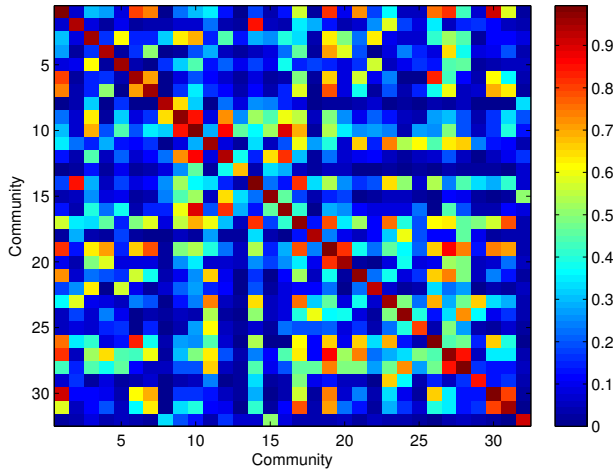

Figure 4.16: The mean of the estimated η matrices for the HC subjects. MAP run 11, density 32 %.

4.3.1 SVD applied the η matrices

SVD is applied the η matrices for the HC and MS subjects in order to see if it is possible to distinguish between the two groups. This is done for each of the 20 runs and for each of the six densities independently. The method is described in Chapter 2.3. The first and second component from the decomposition is shown in Figure 4.20 for density 32 %, MAP run 11. It is not possible to partition the subjects into two groups when looking at the first and second component from the SVD or some of the other components, which are not shown.

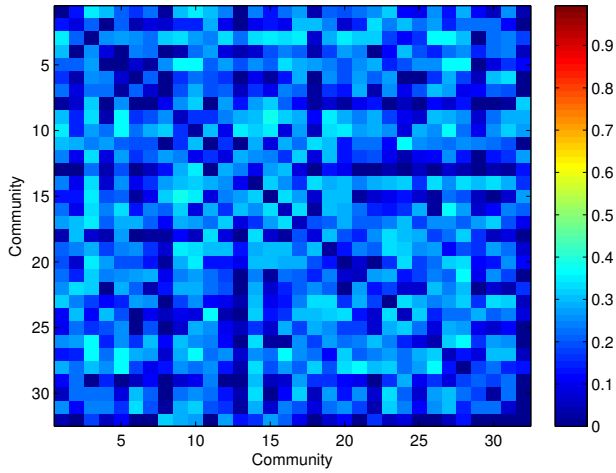

Figure 4.17: The standard deviation in each voxel of the mean HC η matrix. MAP run 11, density 32 %.

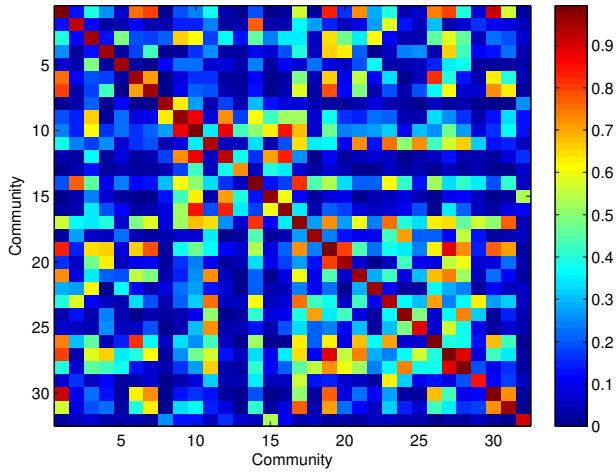

Figure 4.18: The mean of the estimated η matrices for the MS subjects. MAP run 11, density 32 %.

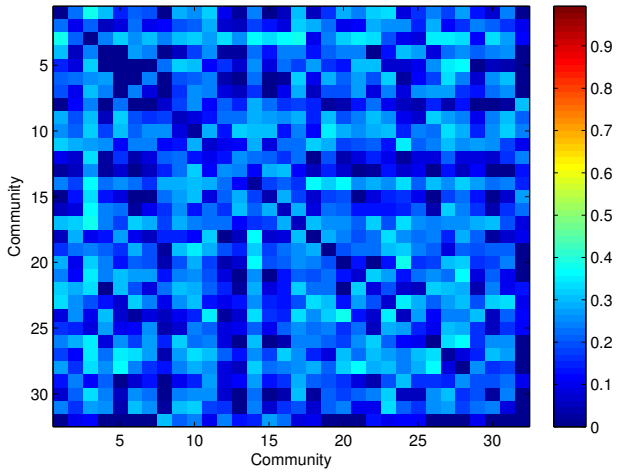

Figure 4.19: The standard deviation in each voxel of the mean MS η matrix. MAP run 11, density 32 %.

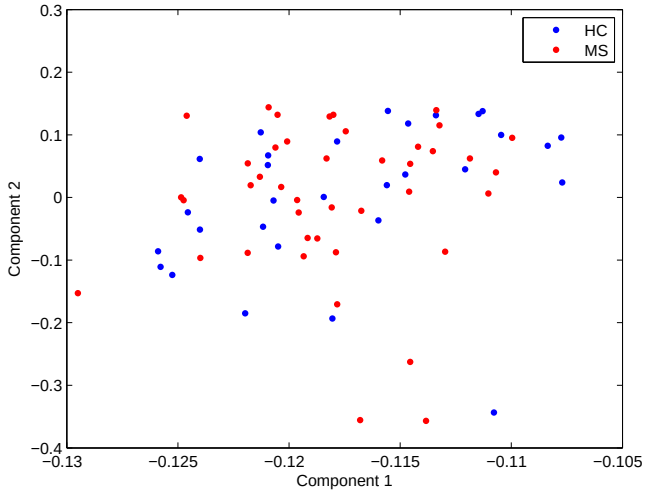

Figure 4.20: Result from SVD applied the η matrices from all the HC and MS subjects. The first and second component are plotted against each other. MAP run 11, density 32 %.

4.4 Community detection

The community structures in rs-fMRI is detected by IRM. IRM was applied the adjacency matrices from the 30 HC subjects and the 42 MS subjects. The MAP run 11 for density 32 % is used when analysing the detected communities.
The detected community structures are given by the cluster assignment matrix, and the link density between the regions are given by teh η matrix. 32 communities are detected and the size of the communities is varying from 1 - 12 regions. Table 4.2 shows the AAL regions in community 1-20 and Table 4.3 shows the AAL regions community 21-32. Figures C.1 - C.32 in Appendix C shows plots of the communities in the brain. Axial slices of the brain are shown in the Figures. The regions in a community is marked with different colors, and the name of the regions are shown in the lower right corner. The communities are symmetric, which means that they consist of the same region in both the right and the left side of the brain, if the region are represented in both sides.

The median of the 72 estimated η matrices is shown in Figure 4.21. Each element in the matrix is the median of the link density between two communities.

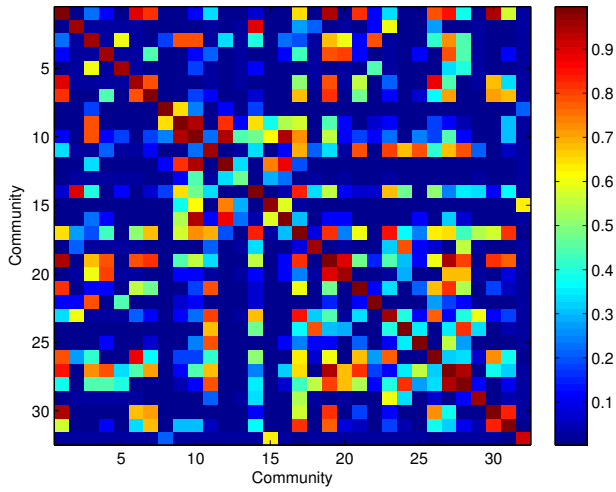

Figure 4.21: The median of the 72 η matrices. MAP run 11 for density 32 %.

The link densities within the communities are between 0.95 and 0.99 except com-

Community	AAL regions	Community	AAL regions
1	Supp_Motor_Area_R Supp_Motor_Area_L Precentral_R Precentral_L	11	Frontal_Inf_Orb_R Frontal_Inf_Orb_L
2	Amygdala_R Amygdala_L	12	Vermis_8 Vermis_7 Cerebelum_8_R Cerebelum_8_L
3	Thalamus_R Thalamus_L	13	Cerebelum_10_R Cerebelum_10_L
4	Angular_R Angular_L	14	ParaHippocampal_R ParaHippocampal_L Hippocampus_R Hippocampus_L
5	Caudate_R Caudate_L	15	Vermis_9 Cerebelum_9_R Cerebelum_9_L
6	SupraMarginal_R SupraMarginal_L	16	Cerebelum_Crus2_R Cerebelum_Crus2_L Cerebelum_7b_R Cerebelum_7b_L
7	Parietal_Sup_R Parietal_Sup_L Parietal_Inf_R Parietal_Inf_L	17	Temporal_Mid_R Temporal_Mid_L Temporal_Inf_R Temporal_Inf_L Fusiform_R Fusiform_L
8	Vermis_3 Vermis_1_2 Cerebelum_3_R Cerebelum_3_L	18	Olfactory_R Olfactory_L
9	Vermis_6 Vermis_4_5 Cerebelum_4_5_R Cerebelum_4_5_L	19	Precuneus_R Precuneus_L Cingulum_Mid_R Cingulum_Mid_L
10	Cerebelum_Crus1_R Cerebelum_Crus1_L Cerebelum_6_R Cerebelum_6_L	20	Cingulum_Post_R Cingulum_Post_L

Table 4.2: The AAL regions in community 1 - 20, which have been detected by using IRM. MAP run 11 for density 32 %.

Community	AAL regions	Community	AAL regions
21	Frontal_Inf_Tri_R Frontal_Inf_Tri_L Frontal_Inf_Oper_R Frontal_Inf_Oper_L	27	Frontal_Sup_R Frontal_Sup_L Frontal_Mid_R Frontal_Mid_L
22	Putamen_R Putamen_L Pallidum_R Pallidum_L	28	Frontal_Sup_Medial_R Frontal_Sup_Medial_L Cingulum_Ant_R Cingulum_Ant_L
23	Temporal_Pole_Sup_R Temporal_Pole_Sup_L	29	Temporal_Pole_Mid_R Temporal_Pole_Mid_L
24	Rectus_R Rectus_L Frontal_Med_Orb_R Frontal_Med_Orb_L	30	Postcentral_R Postcentral_L Paracentral_Lobule_R Paracentral_Lobule_L
25	Frontal_Sup_Orb_R Frontal_Sup_Orb_L Frontal_Mid_Orb_R Frontal_Mid_Orb_L	31	Occipital_Sup_R Occipital_Sup_L Occipital_Mid_R Occipital_Mid_L Occipital_Inf_R Occipital_Inf_L Lingual_R Lingual_L Cuneus_R Cuneus_L Calcarine_R Calcarine_L
26	Temporal_Sup_R Temporal_Sup_L Rolandic_Oper_R Rolandic_Oper_L Insula_R Insula_L Heschi_R Heschi_L	32	Vermis_10

Table 4.3: The AAL regions in community 21 - 32, which have been detected by using IRM. MAP run 11 for density 32 %.

munity 32, with a density of 0.91 and community 13 with 0.48 in link density. Community 13 consist of a part of the left and right cerebellum. Community 32 only consist of the region vermis_10, and it has very low link densities with the rest of the communities.
The link density between communities varying between the different communities. Community 1 consist of the regions Supp_Motor_Area right and left and the right and left precentral regions, which form the primary motor cortex. The link density between community 1 and 6 and 19 and 30 is above 0.89. Community 17 has a link probability around 0.5 with relative many of the other communities. Community 13 consists of regions in part of cerebellum. The highest link density with community 13 is found between community 13 and 10 and it has a value of 0.45

The communities 8, 9, 10, 12, 13, 15, 16 and 32 all consist of regions in cerebellum and vermis. When looking at the four highest link densities for these communities only community 8, 9, 10 and 32 have connections to other parts of the brain among these. Community 9 and 10 both have a link density of 0.78 to community 3 which consist of the right and left thalamus region. A communication structures between the communities in cerebellum is seen.

Community 11, 21, 25, 27 and 28 all contains regions in the frontal lobe. They are all connected to some of the other communities in the frontal lobe with a link density above 0.78. Community 11 has the highest link density to community 23 which is a part of the temporal lobe. Community 21 has the highest link density to community 1, which is in the motor cortex and community 27 has the highest link probability to community 19 which consist of regions in the right and left precunes and part of the right and left cingulum. The communities in the frontal lobe is connected.

The link densities between community 4 and community 19, 20 and 27 are above 0.78. The link density between community 19 and community 20 and 27 respectively is above 0.89. The link density between community 20 and community 4 and 19 respectively is above 0.80 and community 27 and 28 respectively are 0.67. Community 27 has a link density of 0.94 with community 19 and community 28. The link density between community 28 and 27 is 0.94 and 0.80 between community 28 and 19. It seems like there is a high connection between community the five communities 4, 19, 20, 27 and 28. Community 19, 27 and 28 are described above, community 4 is the right and left angular region and community 20 is a part of the cingulum.

Community 31 is the largest community, and it consists of 12 regions in the occipital lobe.

4.5 Classification

4.5.1 K nearest neighbours

Until now it has not been possible to differentiate the HC and MS subjects from each other after SVD was applied the η matrices. The classification methods KNN and SVM will now be tested.
For optimizing the classification rate for KNN different values of K were tested. The K values are the odd numbers between 1 and 30. Only the odd values are chosen because it can be difficult to classify when K is an eval number. An example could be when K is 4 and two of the nearest subjects is HC and two of them is MS.
SVM is evaluated for different values of C. The tested C values are integers between 1 and 30.
The η matrix for each subject was used for the classification. KNN and SVM were test by the method leave a HC and a MS subject out. The remaining 70 subjects were used as training data and the two subjects were used as test data. This was done randomly 1000 times for each of the 20 runs for each K and C value respectively. To make it possible to evaluate for which density KNN and SVM have the highest classification rate, the tests are made for each of the six densities.
Before classification is done with KNN, SVD is applied the η matrices, and then the 10 components with the largest variance was used as feature vector in KNN. The mean of the 20 classification rates for each vale of K was estimated and the results are shown in Figure 4.22 for the six densities. It is seen that the results from KNN depend of the K value. For the high densities a low value of K gets the best classification rate. The same is seen for density 2 %. For densities 4 % and 8 % the highest classification rate is found for a high value of K. The highest classification rate is 0.6079 and it was achieved for a density of 50 % and a K value of 7. The classification rate was below 0.5 for all values of K for the densities 2 %, 4 % and 8 %. Baseline for classification is 0.58, so only with the densities 32 % and 50 % it has been possible to achieve a classification rate above baseline.

4.5.2 Support Vector Machine

The η matrices are used for classification with SVM. The mean of the 20 classification rates for each value of C are estimated and Figure 4.23 shows the results for the six densities. The classification rate of SVM is relative stable when the C value varied, so the classification rate does not seems to depend on the C value. The highest classification rate is achieved
The Highest classification rate with SVM is achieved for C = 7 and a density of 32 %. The classification rate is higher for density 16 % compared to 50 %. The classification rate for the low densities 2 %, 4 % and 8 % is below 0.5, which is below the baseline rate.

SVM has a higher classification rate compared to KNN in the high densities 16 %, 32 % and 50 %. With both classification methods a gap in classification rate is seen between the low densities 2 %, 4 % and 8 % and the high densities 16 %, 32 % and 50 %.

Figure 4.24 shows the log likelihood vs. classification rate for SVM and KNN, for each of the 20 runs with density 32 %, C = 7 and K = 7. The classification rate between the runs for both methods are relative large. The highest classification rate for KNN is 0.66 and the lowest is 0.5. For SVM the highest classification rate is 0.73 and 0.59 for the lowest. It is seen, that the run with the highest classification rate for both KNN and SVM does not has the highest log likelihood.

The results of the classification shows, that with a density of 16 %, 32 % and 50 % it is possible to classify the subjects when using the elements in the η matrix as features. on the basis of this it is assumed that there is a difference in some of the elements in the η matrix between the HC subjects and the MS subjects. MAP run 11, density 32 % is used.
A permutation test is made to test the null hypothesis: No difference between an element in the mean of the HC η matrices and the mean of the MS η matrices. Each element in the η matrix is tested independently. Because a η matrix is symmetric only the elements in the upper triangle of the matrix and the diagonal is tested.
The mean of the HC η matrices and the mean of the MS η matrices are estimated and the absolute difference is found in each element. The 72 subjects were randomly divided into two groups one with 30 subjects and one with 42 subjects. The mean of the η matrices in each group are estimated and the difference between the two mean matrices in each element are estimated. This is

done 1000 times. It is the absolute values of the difference which have been used. The distribution over the 1000 differences in each element is made and a significance level of 0.05 is chosen. If the probability of getting the true absolute difference between the HC and MS is below 0.05 the null hypothesis is rejected. Figure 4.25 shows the elements in the η matrix with $p < 0.05$. Four of the elements had a p value below 0.005. These four elements, for which a difference between HC and MS have been found is (6,2), (25,6), (25,17) and (25,21). The different AAL regions in community 2 and 6 are shown in Table 4.2 and i community 17, 21 and 25 are shown in Table 4.3. The five different communities and their matching AAL regions are shown in axial slices of the brain in Appendix C. Community 2 is amygdala, community 6 is the primary motor cortex, community 17 is part of the temporal lobe and communities 21 and 25 are parts of the frontal lobe. The difference in link density is found between the regions in the frontal lobe and the primary motor cortex and the temporal lobe respectively. Difference is also found between amygdala and primary motor cortex. The expected error rate for $p<0.005$ is about three elements and five elements are found in this permutation test.

4.6 Correlation between the link density and the EDSS.

Inference with IRM was made for the 42 MS subject. 20 runs were made for a density of 32 %. The MAP run 12 is used, and 28 communities were detected in this run.
The correlation between each element in the η matrices for the MS subjects and theirs EDSS was estimated. Because the η matrix is symmetric only the element in the upper triangle and the diagonal are used. The null hypothesis: No correlation between the link density and the EDSS is tested against the alternative hypothesises: There is a non zero correlation. The significance level is set to $p<0.05$. Figure 4.26 shows the 29 elements with $p<0.05$ in the η matrix represented by their p value. Figure 4.27 shows the correlation values for the 29 elements. 14 of the elements have a positive correlation and 15 of the elements have a negative correlation. A positive correlation means that when EDSS increases the link density increases. A negative correlation means that the higher EDSS the lower link density between the elements.
6 of the elements in the η matrix have a p value < 0.005. The elements are (10,4), (10,5), (10,15), (10,21) and (13,4) which have a negative correlation and the element (22,18) which has a positive correlation. Table 4.4 shows the regions which represents the eight communities. The expected error rate for a significance level of 0.005 is about 4.

Community	AAL regions
10	Vermis_9
	Vermis_8
	Cerebelum_8_R
	Cerebelum_8_L
	Cerebelum_9_R
	Cerebelum_L
4	Caudate_R
	Caudate_L
5	Precuneus_R
	Precuneus_L
	Cingulum_Mid_R
	Cingulum_Mid_L
13	Cerebelum_Crus1_R
	Cerebelum_Crus1_L
	Cerebelum_Crus2_R
	Cerebelum_Crus2_L
15	Supp_Motor_Area_R
	Supp_Motor_Area_L
	Precentral_R
	Precentral_L
	Postcentral_R
	Postcentral_L
	Paracentral_Lobule_R
	Paracentral_Lobule_L
18	Frontal_Inf_Orb_R
	Frontal_Inf_Orb_L
21	Frontal_Sup_R
	Frontal_Sup_L
	Frontal_Mid_R
	Frontal_Mid_L
22	Frontal_Sup_Orb_R
	Frontal_Sup_Orb_L
	Frontal_Mid_Orb_R
	Frontal_Mid_Orb_L

Table 4.4: The communities which have significant correlation between the link densities The elements, in the η matrix, with a significant correlation between the link probability and the EDSS. Also the regions representing the eight communities are shown.

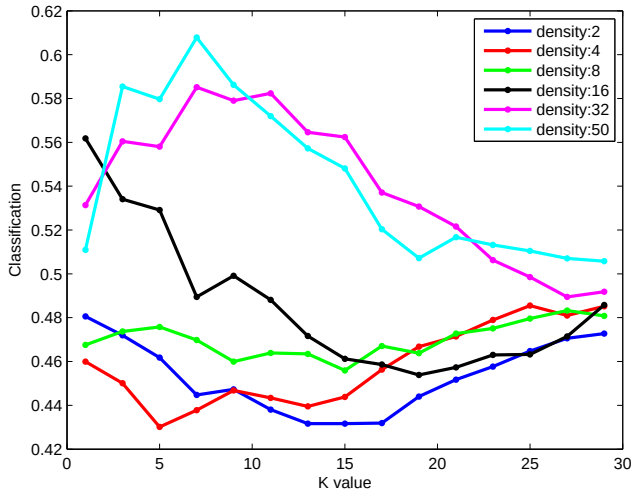

Figure 4.22: The mean classification rate of the 20 run made for different values of K, when using KNN for classification. The results are shown for the six densities.

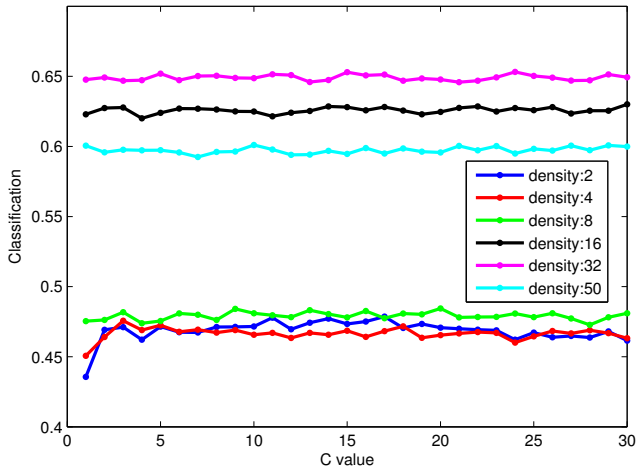

Figure 4.23: The mean classification rate of the 20 run made for different values of C, when using SVM for classification. The results are shown for the six different densities.

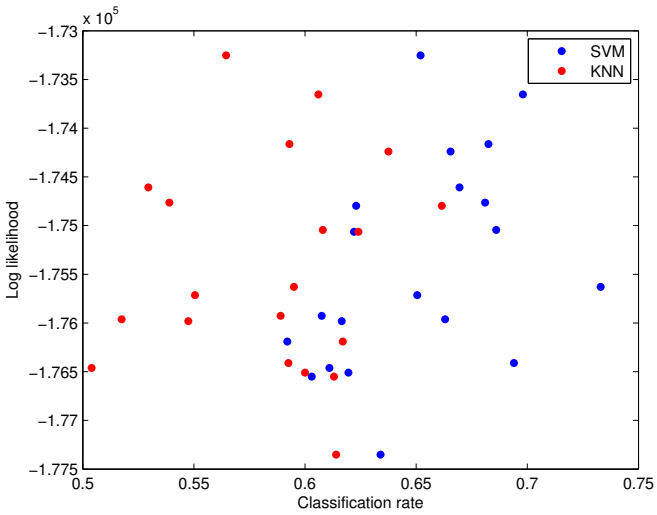

Figure 4.24: The likelihood plotted as a function of classification rate for the 20 runs. Density 32 %, C = 7 and K = 7 .

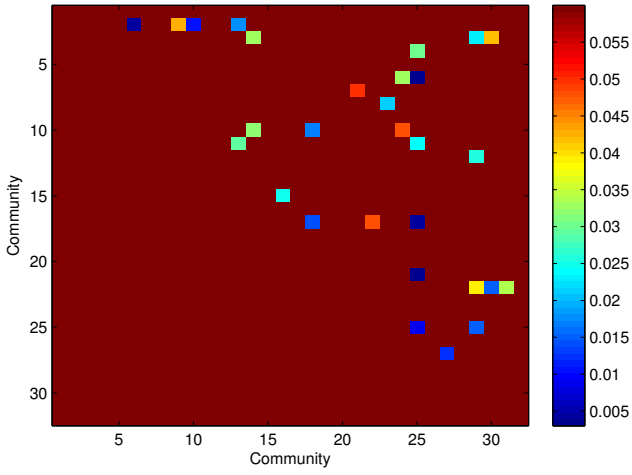

Figure 4.25: The elements in η with a p value ≤ 0.05 when comparing the mean values for the HC and MS group.

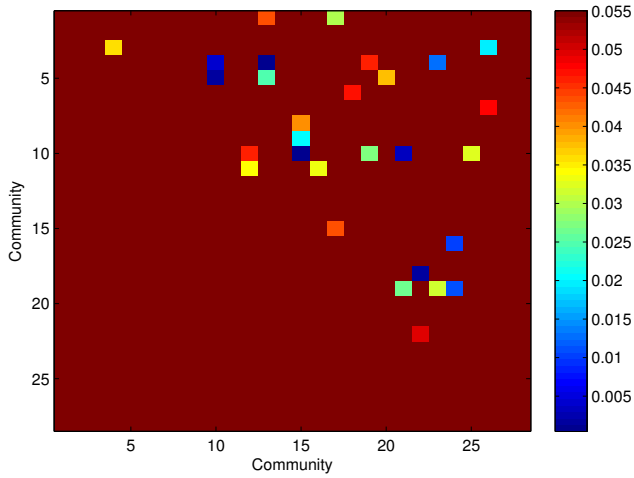

Figure 4.26: The elements in the η matrix for MS subjects, with significant ($p<0.05$) correlation between link density and EDSS, represented by their p values.

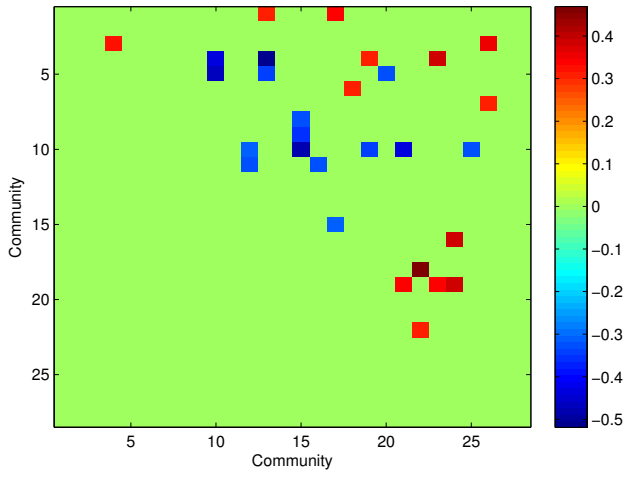

Figure 4.27: The elements in the η matrix for MS subjects, with significant ($p<0.05$) correlation between link density and EDSS, represented by their correlation values.

Chapter 5

Discussion

5.1 Resting state networks

In this thesis 32 communities have been detected during resting state, this is more than the 8 RSN presented by [45]. A higher level of details in the communication pattern are seen when the number of detected communities are high. The detected communities have been compared to the RSN defined by van den Heuvel et al. [45]. Communication patterns are found between communities, which is a part of cerebellum and also between communities, which is a part of the frontal lobe.

A high link density is found between the communities 4, 19, 20, 27 and 28. These communities consists of the brain regions angular, precuneus, cingulum and part of the frontal lobe. The detected regions are symmetric so both the left and right region is a part of the networks. The DMN defined in Table 3.3 (defined from the DMN described in the 8 papers.) is not symmetric. This could be caused by the converting from the given areas in the article to the AAL regions. The precuneus regions are part of community 19 but it is not represented in Table 3.3. In the literature the precuneus is described as a part of the DMN [45]. The angular, cingulum and the regions in the frontal lobe is a part of the default mode networks. In the communities detected in this thesis

all the regions in cingulum are a part of the network. The network found from the literature contain only one region of the cingulum. A region in the temporal lobe are not represented in the five communities detected in this thesis. A region in the tempral lobe is represented in the network shown in Table 3.3. Region in the temporal lobe is also described in the literature [45]. Community 17 consist of parts of the temporal area, and this community has a relative high link probability to community 19 and 27. It seems like the communities 4, 19, 20, 27 and 28 form a network which is comparable with the DMN defined from the 8 papers. It can be discussed if community 17 should be a part of this network.
Community 31 consists of regions in the occipital lobe, and it seems comparable to the with the primary visual and the extra-striate visual network.
Community 1 contains the regions left and right Supp_Motor_Area and left and right Precentral, these regions can be defined as the primary motor cortex. This network seems to be comparable with the primary motor network.
It has not been possible to detect a network which could be compared to the right and left parietal-frontal networks. Community 25 consist of part of the frontal lobe and community 7 is a part of the parietal lobe, but the link density between the two communities are 0.06, so only very low communication between these communities are found. With such a low link density it is not comparable with the parietal-frontal network in the literature.
The RSN in this thesis have been detected by using rs-fMRI data from both HC subjects and MS subjects. Changes in the connectivity in the DMN have been found in patients with MS [6]. To achieve an understanding of the RSN in HC subjects and MS subjects, studies have to be made separate for these two groups.
The communities detected with IRM is comparable with the DMN, primary motor, primary visual and extra-striate visual. The results indicates that it is possible to detect RSN when using IRM rs-fMRI data.

5.2 Classification

It was not possible to classify the HC and MS subjects after SVD were made for the feature vector of the elements in the correlation matrices and η matrices respectively. Instead KNN and SVM were used for classification.
The best classification rate was found for a density of 32 %. When using SVM and for a density of 50 % when using KNN. SVM and KNN were used to discriminate between the HC and MS subjects. The best mean classification rate for SVM was 0.65 and 0.61 for KNN. The baseline probability is 0.58 so the classification rate when using KNN was only above baseline for the density 0.5. For density 32 % the classification rate and baseline was almost the same and

for the low densities 2 %, 4 %, 8 % the calssification rates were below baseline. When using SVM the classification rate for the densities 16 %, 32 % and 50 % were above baseline. The classification rates for both methods for the low densities 2 %, 4 % and 8 % are significant lower compared to the classification rates for the three highest densities. The brain is divided into 116 regions, so the size of the graphs is relative small. This means, that for the low densities the number of links in the adjacency matrices are small. This could be an explanation for why the classification rate is so low for the low densities.
When using SVM the highest classification rate is found for density 32 %, and the classification rates are higher for density 16 % compared to the classification rate for density 50 %. With a density of 50 % the highest half of correlations between regions are defined as a link. This means that links are defined even between regions with a low correlation, so the noise in the graph can be high, and this can have an influence on the results. When using KNN for classification the highest rate is achieved for density 50 %, which deviate from the results with SVM.

It have been possible to discriminate between HC and MS subjects with a classification rate above baseline. This indicates that there is a difference between some of the elements in the estimated η matrices for the HC and the MS subjects. The permutation test shows, that in four of the elements a difference between HC and MS subjects were found with $p<0.005$. The accepted error rate for the permutation test for $p<0.005$ is about 3, so the four elements with $p<0.005$ is almost of the same size as the accepted error. So the result is not significant, but it indicates that difference between HC and MS subjects can be found in given communities.
The difference in link density is found between regions in the frontal lobe, between regions in the frontal lobe an regions in the temporal lobe and fusiform. Between regions in the frontal lobe and the supra marginal region and between amygdala and the supra marginal region. This is all regions with cognitive functions like long time memory in the frontal lobe, body and face recognition and word recognition in fusiform, storing new memory in the temporal lobe, handle defence reactions. In multiple sclerosis the hole brain can be affected, so cognitive deficits are also seen in patients with multiple sclerosis.

The classification rate is varying between the runs, with density 32. The lowest classification rate for SVM is 0.59 and it is just above the baseline. The lowest classification rate for KNN was 0.5, which is below the baseline. SVM has the highest classification rate of 0.73. The MAP run (11), for the same density, has a classification rate of 0.67 for C = 7. The same is seen when using KNN. The highest classification rate of 0.66 was achieved for run number 1 (K=7). The MAP run (11) has a classification rate of 0.64. The same is also seen for

the other densities. This means that the run with the highest log likelihood for the model fitting the data does not have the highest classification rate. The likelihood tells something about the communication patterns in the brain and the classification is made of the result of the changes in the brain due to the diseases, this can explain why the run with the highest likelihood do not achieved the highest classification rate.
Mørup et al.[34] analysed rs-fMRI data from HC and MS subjects, with the purpose of detecting community structures by using IRM. KNN and SVM were used to discriminate the HC subjects from the MS subjects. They achieved a classification rate of 0.67 with KNN and 0.72 for SVM. Mørup et al. achieved higher classification rates for SVM and KNN compared to the mean classification rate found in this thesis. When comparing the results by Mørup et al. with the highest classification rate for density 32% the classification rate for KNN and SVM the classification rates are almost the same. In both studies IRM was used, and the η matrices for each subjects were used as feature vector for the classification, but in this thesis SVD was made for the feature vectors and the first 10 components was used as features when classify with KNN. They create the adjacency matrix by thresholding the estimated mutual information graph and made classification for one run with IRM.
It seems like it could be difficult to classify subjects when using the η matrix as feature vector, because of the large variation between runs.

5.3 Correlation between EDSS and elements in η

Inference with IRM were made for the data from the MS subjects. The estimated η matrix for each subject was used when correlation between the EDSS for each subject and each of the elements in the η matrices. The null hypothesis that there was no correlation between the the element in η and the EDSS was tested against the alternative hypothesis that the correlation is different from zero between the element and EDDS. Six elements in the η matrix were correlated to EDSS with $p<0.005$. Community 10 was one of the communities represented in 4 of the elements. Each with a negative correlation with EDSS, which mean that the link density decreases when the EDSS increases. The communities 10 and 13 is part of the cerebellum, which in concert with motor cortex and prefrontal areas in the brain support motor and cognitive skills [30]. The link density between cerebellum and regions in the frontal part of the brain and cerebellum and the motor and sensory areas are negative correlated with EDSS. This results indicates that the communication paths decreases when the diseases progress which affects the patients cognitive and motor skills. An positive correlation is found between EDSS and the link density between regions in

the orbitaofrontal cortex. So when the EDSS score increases for patients with MS the link density between regions in the orbitofrontal cortex increases. This could be because new communication paths are made to compensate for the degenerative neurons in these regions.
Six elements in the η matrix have a $p<0.005$ this is just above the number of expected errors, so the results is not significant. But the results indicate that the communication between cerebellum and frontal and motor regions are reduced which have been shown in the literature.

more studies have to be made before it possible to make a conclusion.

Chapter 6

Conclusion

The IRM has been used to detect community structures in rs-fMRI data sets. These communities have been combined to form an equivalent resting state network.
It seems that the IRM can be used for detection and evaluation of resting state networks in the human brain.
The classification rate measured over 20 runs on the same data set varies between 59-73 % for SVM. This range is on par with baseline classification and is comparable with the best published results [34], with a classification rate of approximately 72 %.
Classification using KNN was found to have a mean value of 66 % compared with 67 % by Møeup et al.[34]. The mean value in this thesis have been obtained over 20 runs compared with a single run in [34].
Classification of the subjects is not stable, and the investigated method with link density as feature vector can not be used as a standalone classification method.
It has been found that there is no significant correlation between the likelihood and the classification rate.

Appendix A

Literature studies RSN

Studies	Method	Resting state networks	Description	Material
Biswal et al. [5]	Seed region correlation	Primary motor	Anatomic areas	resting state 11 healthy subjects
Greicius et al. [22]	Linear regression	DMN	Talairach space	resting state 14 healthy subjects
Salvador et al. [41]	Hierarchical cluster analysis	Primary motor primary visual extra-striate visual	regions from a template with 90 brain regions	resting state 12 healthy subjects
Beckmann et al. [4]	ICA	Primary visual extra-striate visual insular-temporal /ACC primary motor DMN parietal-frontal left parietal-frontal right	Anatomic areas	resting state 10 healthy subjects
Fox et al. [19]	Seed region correlation	DMN	Talairach space	resting state 10 healthy subjects
Damoiseaux et al. [11]	ICA tensor	extra-striate visual DMN parietal-frontal left parietal-frontal right Primary visual primary motor insular-temporal /ACC frontal	Brodmann space anatomic areas	resting state 10 healthy subjects
Raichle et al. [39]	seed region correlation	DMN Primary visual primary motor parietal-frontal	711-2B space	resting state 1 healthy subject

Table A.1: Overview over published studies, which have detected resting state networks. In the table the method, the detected resting sate networks, the way the resting state networks are presented and the material for the experiment are shown for each paper.

Studies	Method	Resting state networks	Description	Material
De Luca et al. [12]	ICA	primary motor DMN extra-striate visual Primary visual parietal-frontal left parietal-frontal right insular-temporal /ACC	Talairach space anatomic areas	resting state 10 healthy subjects
van den Heuvel et al. [46]	graph method normalized cut graph clustering	DMN parietal-frontal left parietal-frontal right insular-temporal /ACC frontal Primary visual primary motor extra-striate visual	Brodmann areas anatomic areas	resting state 26 healthy subjects
Larid et al. [29]	activation likelihood	DMN	Talairach space	resting state 840 healthy subjects from different studies
Pizoli et al. [38]	ICA correlation	DMN primary motor frontal	Talairach space	sedation 1 child with epileptic encephalopathy
Van Dijk et al. [47]	ICA seed region correlation	DMN extra-striate visual / primary motor	MNI space	resting state 48 healthy subjects

Table A.2: Overview over published studies, which have detected resting state networks. In the table the method, the detected resting sate networks, the way the resting state networks are presented and the material for the experiment are shown for each paper.

Appendix B

DMN described in literature

	DMN									
AAL region	De Luca et al. [12]	Van Dijk et al. [47]	Greicius et al. [22]	Pizoli et al. [38]	Larid et al. [29]	Raichle et al. [39]	Fox et al. [19]	van den Heuvel et al. [46]		
Precentral_L								x		
Precentral_R								x		
Frontal_Sup_L			x	x				x		
Frontal_Sup_R			x	x			x	x		
Frontal_Sup_Orb_L								x		
Frontal_Sup_Orb_R								x		
Frontal_Mid_L			x		x			x		
Frontal_Mid_R								x		
Frontal_Mid_Orb_L								x		
Frontal_Mid_Orb _R								x		
Frontal_Inf_Oper_L										
Frontal_Inf_Oper_R										
Frontal_Inf_Tri_L										
Frontal_Inf_Tri_R										
Frontal_Inf_Orb_L								x		
Frontal_Inf_Orb_R								x		
Rolandic_Oper_L								x		
Rolandic_Oper_R								x		
Supp_Motor_Area_L								x		
Supp_Motor_Area_R								x		
Olfactory_L										
Olfactory_R								x		
Frontal_Sup_Medial_L					x	x	x	x		
Frontal_Sup_Medial_R								x		
Frontal_Med_Orb_L		x						x		
Frontal_Med_Orb_R								x		
Rectus_L			x					x		
Rectus_R								x		
Insula_L										
Insula_R								x		

Table B.1: The DMN given by AAL regions from different studies. The AAL regions are found from coordinates in MNI space. In [47] the coordinates were given in MNI space. In [39] the coordinates were given in 711-2B space. In [46] the network was given in Brodmann areas. In the rest of the studies the coordinates were given in Talairach space.

	DMN									
AAL region	De Luca et al. [12]	Van Dijk et al. [47]	Greicius et al. [22]	Pizoli et al. [38]	Larid et al. [29]	Raichle et al. [39]	Fox et al. [19]	van den Heuvel et al. [46]		
Cingulum_Ant_L				x			x			
Cingulum_Ant_R			x	x	x					
Cingulum_Mid_L				x			x	x		
Cingulum_Mid_R			x					x		
Cingulum_Post_L	x	x	x			x		x		
Cingulum_Post_R				x				x		
Hippocampus_L										
Hippocampus_R										
ParaHippocampal_L	x	x					x			
ParaHippocampal_R		x					x			
Amygdala_L										
Amygdala_R										
Calcarine_L				x				x		
Calcarine_R				x				x		
Cuneus_L								x		
Cuneus_R								x		
Lingual_L			x							
Lingual_R										
Occipital_Sup_L								x		
Occipital_Sup_R								x		
Occipital_Mid_L				x	x					
Occipital_Mid_R				x						
Occipital_Inf_L										
Occipital_Inf_R										
Fusiform_L				x				x		
Fusiform_R				x				x		
Postcentral_L								x		
Postcentral_R								x		
Parietal_Sup_L								x		
Parietal_Sup_R								x		

Table B.2: The DMN given by AAL regions from different studies. The AAL regions are found from coordinates in MNI space. In [47] the coordinates were given in MNI space. In [39] the coordinates were given in 711-2B space. In [46] the network was given in Brodmann areas. In the rest of the studies the coordinates were given in Talairach space.

	DMN									
AAL region	De Luca et al. [12]	Van Dijk et al. [47]	Greicius et al. [22]	Pizoli et al. [38]	Larid et al. [29]	Raichle et al. [39]	Fox et al. [19]	van den Heuvel et al. [46]		
Parietal_Inf_L								x		
Parietal_Inf_R								x		
SupraMarginal_L					x			x		
SupraMarginal_R					x			x		
Angular_L		x	x	x		x	x	x		
Angular_R		x	x	x		x	x	x		
Precuneus_L				x	x			x		
Precuneus_R			x					x		
Paracentral_Lobule_L								x		
Paracentral_Lobule_R								x		
Caudate_L										
Caudate_R			x							
Putamen_L										
Putamen_R										
Pallidum_L										
Pallidum_R										
Thalamus_L										
Thalamus_R	x									
Heschl_L										
Heschl_R										
Temporal_Sup_L								x		
Temporal_Sup_R								x		
Temporal_Pole_Sup_L								x		
Temporal_Pole_Sup_R								x		
Temporal_Mid_L						x		x		
Temporal_Mid_R	x				x	x	x	x		
Temporal_Pole_Mid_L								x		
Temporal_Pole_Mid_R								x		
Temporal_Inf_L				x			x	x		
Temporal_Inf_R				x				x		

Table B.3: The DMN given by AAL regions from different studies. The AAL regions are found from coordinates in MNI space. In [47] the coordinates were given in MNI space. In [39] the coordinates were given in 711-2B space. In [46] the network was given in Brodmann areas. In the rest of the studies the coordinates were given in Talairach space.

DMN									
AAL region	De Luca et al. [12]	Van Dijk et al. [47]	Greicius et al. [22]	Pizoli et al. [38]	Larid et al. [29]	Raichle et al. [39]	Fox et al. [19]	van den Heuvel et al. [46]	
Cerebelum_Crus1_L									
Cerebelum_Crus1_R									
Cerebelum_Crus2_L						x			
Cerebelum_Crus2_R						x			
Cerebelum_3_L									
Cerebelum_3_R									
Cerebelum_4_5_L									
Cerebelum_4_5_R									
Cerebelum_6_L									
Cerebelum_6_R									
Cerebelum_7b_L									
Cerebelum_7b_R									
Cerebelum_8_L									
Cerebelum_8_R									
Cerebelum_9_L									
Cerebelum_9_R							x		
Cerebelum_10_L									
Cerebelum_10_R									
Vermis_1_2									
Vermis_3									
Vermis_4_5							x		
Vermis_6									
Vermis_7									
Vermis_8									
Vermis_9									
Vermis_10									

Table B.4: The DMN given by AAL regions from different studies. The AAL regions are found from coordinates in MNI space. In [47] the coordinates were given in MNI space. In [39] the coordinates were given in 711-2B space. In [46] the network was given in Brodmann areas. In the rest of the studies the coordinates were given in Talairach space.

Appendix C

Communities

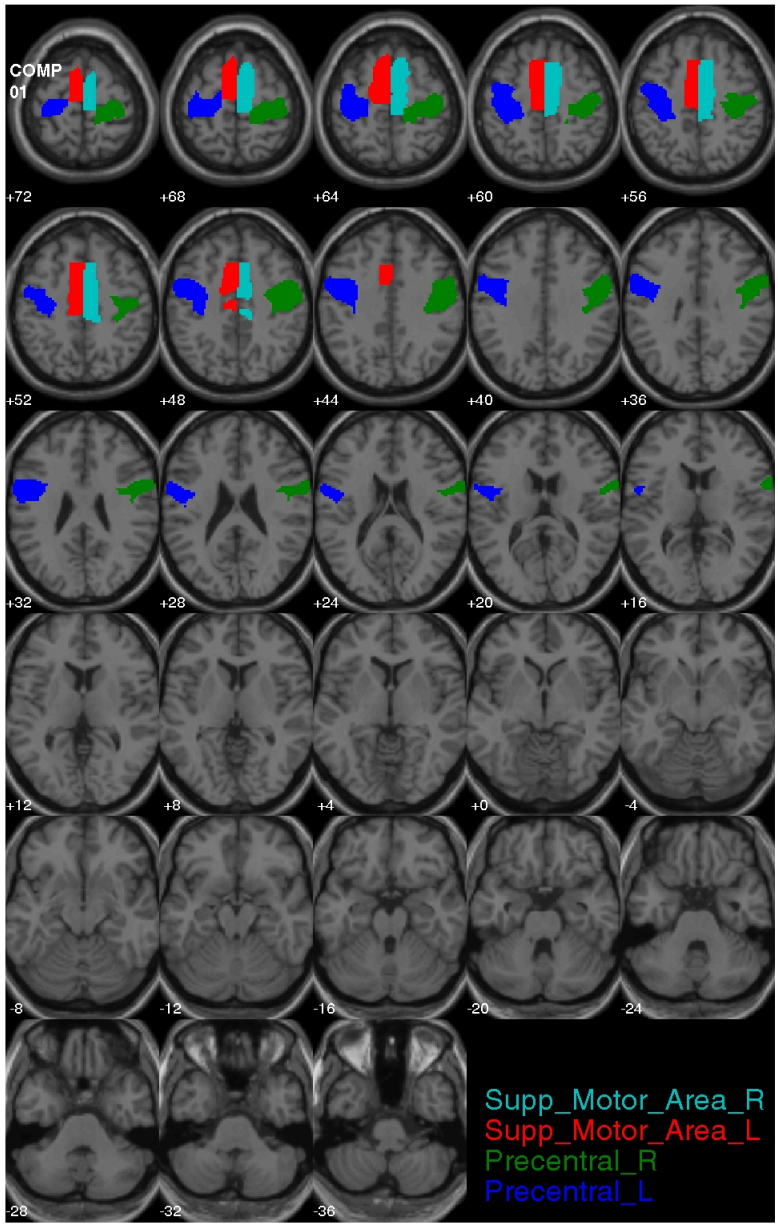

Figure C.1: The AAL regions of 1. community detected by IRM, shown in the brain.

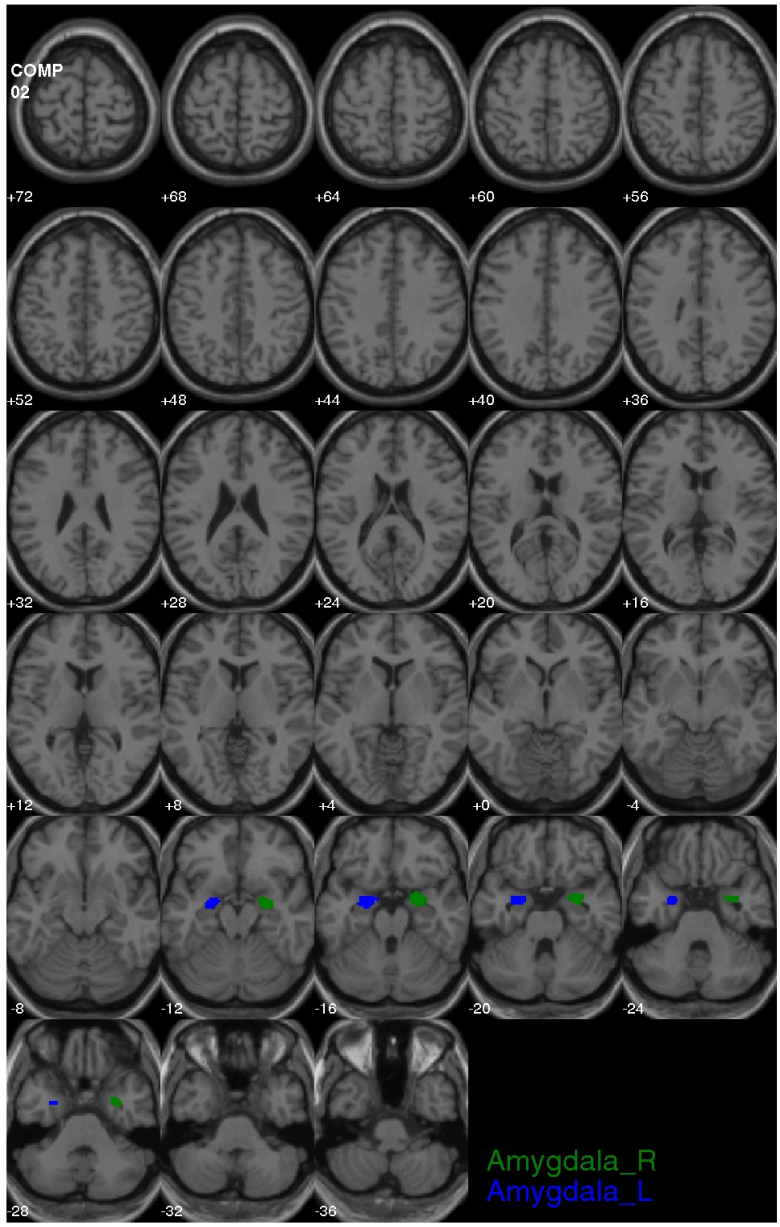

Figure C.2: The AAL regions of 2. community detected by IRM, shown in the brain.

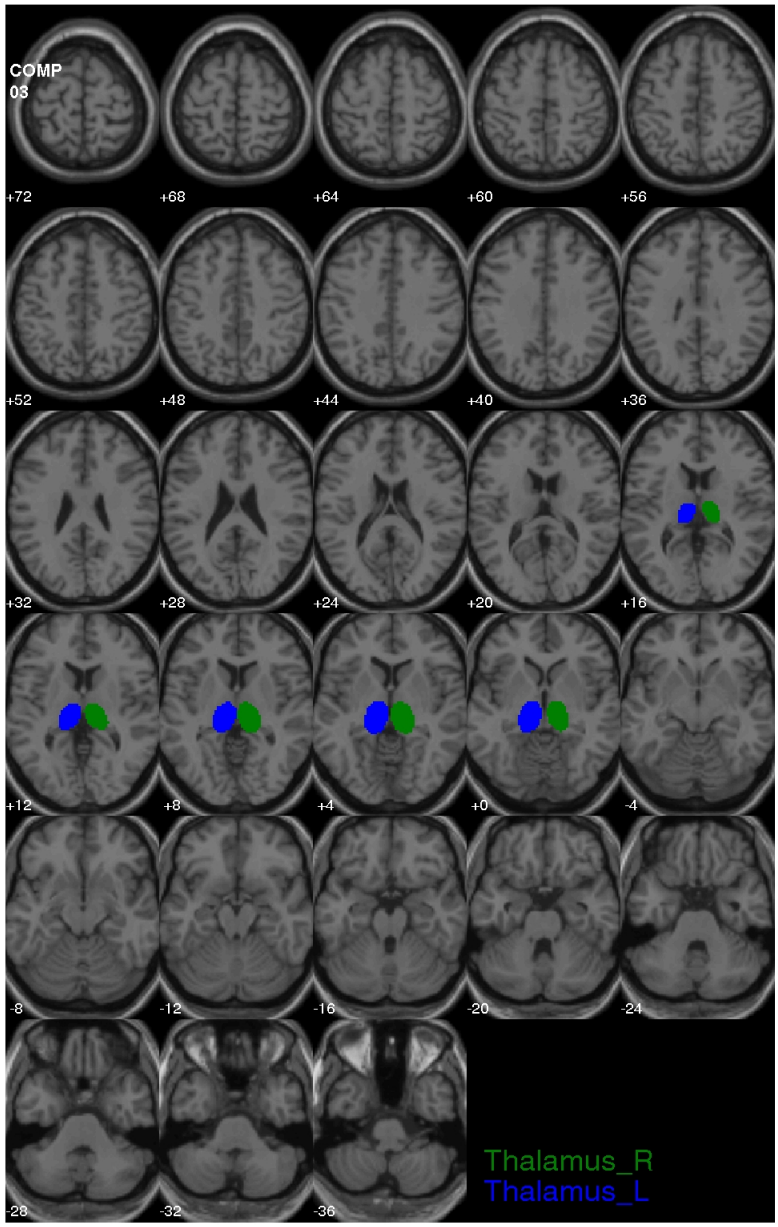

Figure C.3: The AAL regions of the 3. community detected by IRM, shown in the brain.

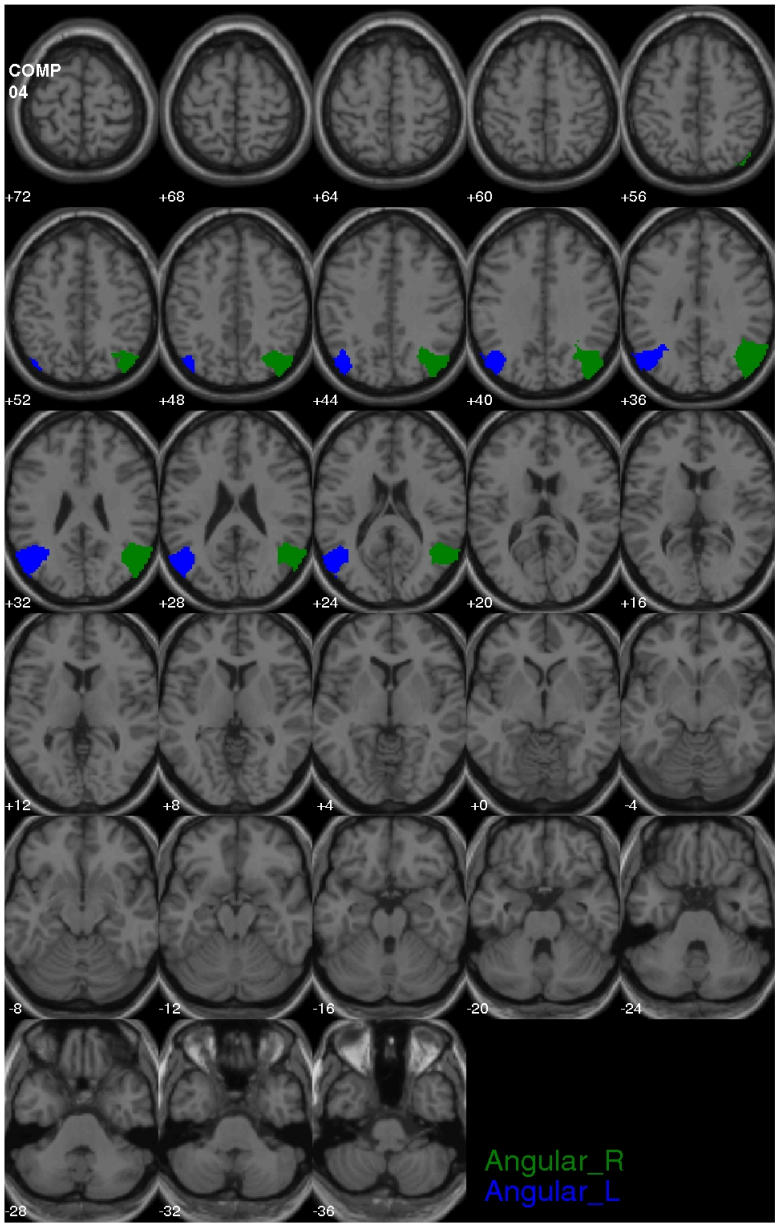

Figure C.4: The AAL regions of the 4. community detected by IRM, shown in the brain.

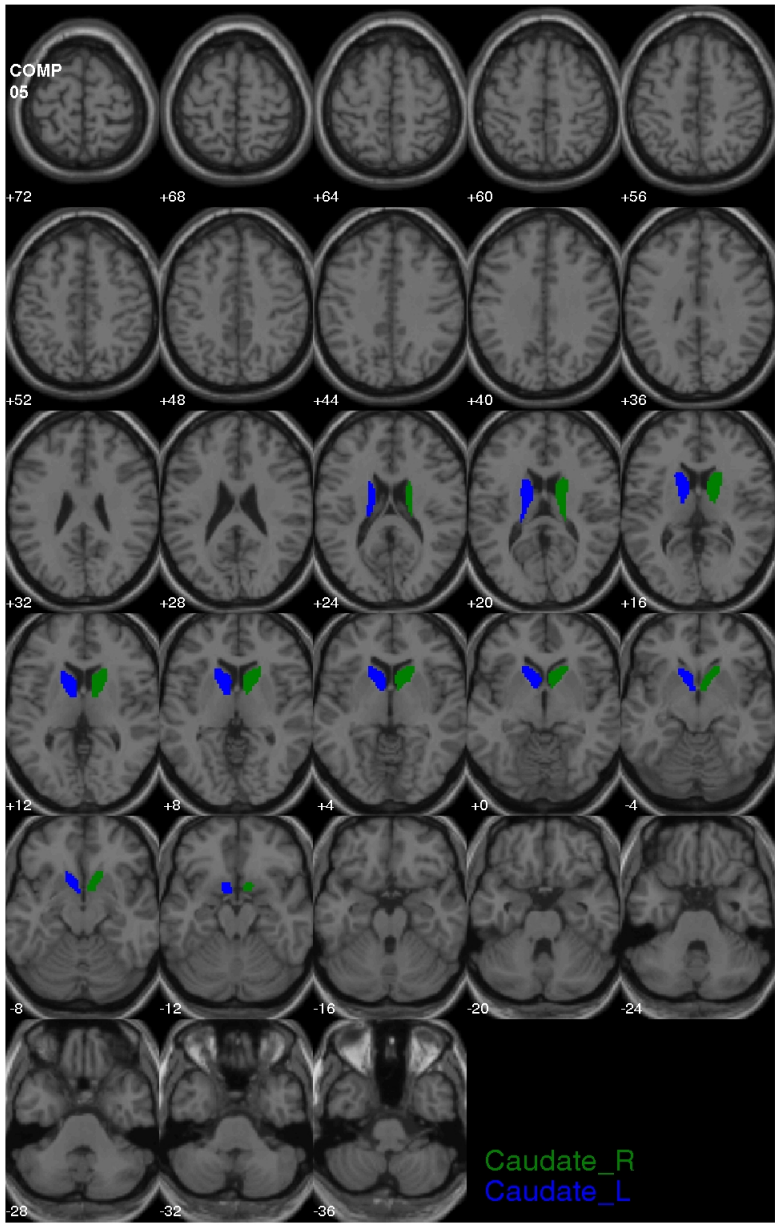

Figure C.5: The AAL regions of the 5. community detected by IRM, shown in the brain.

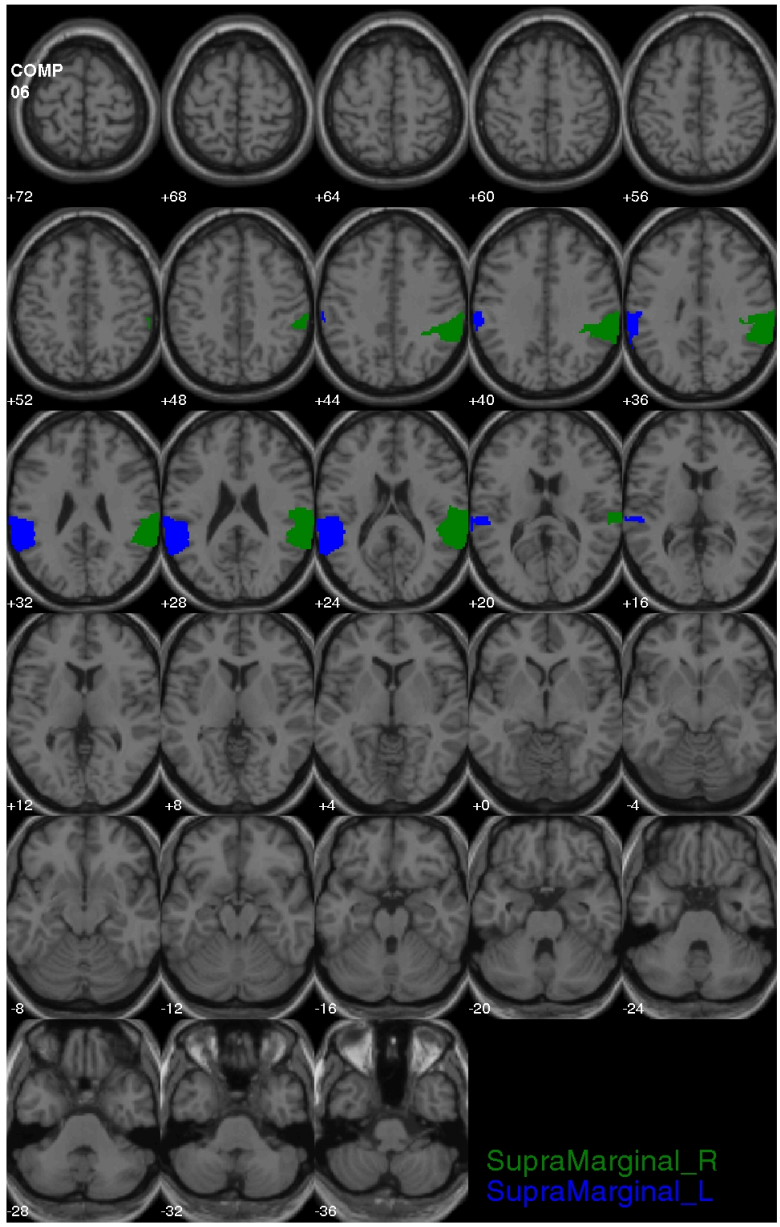

Figure C.6: The AAL regions of the 6. community detected by IRM, shown in the brain.

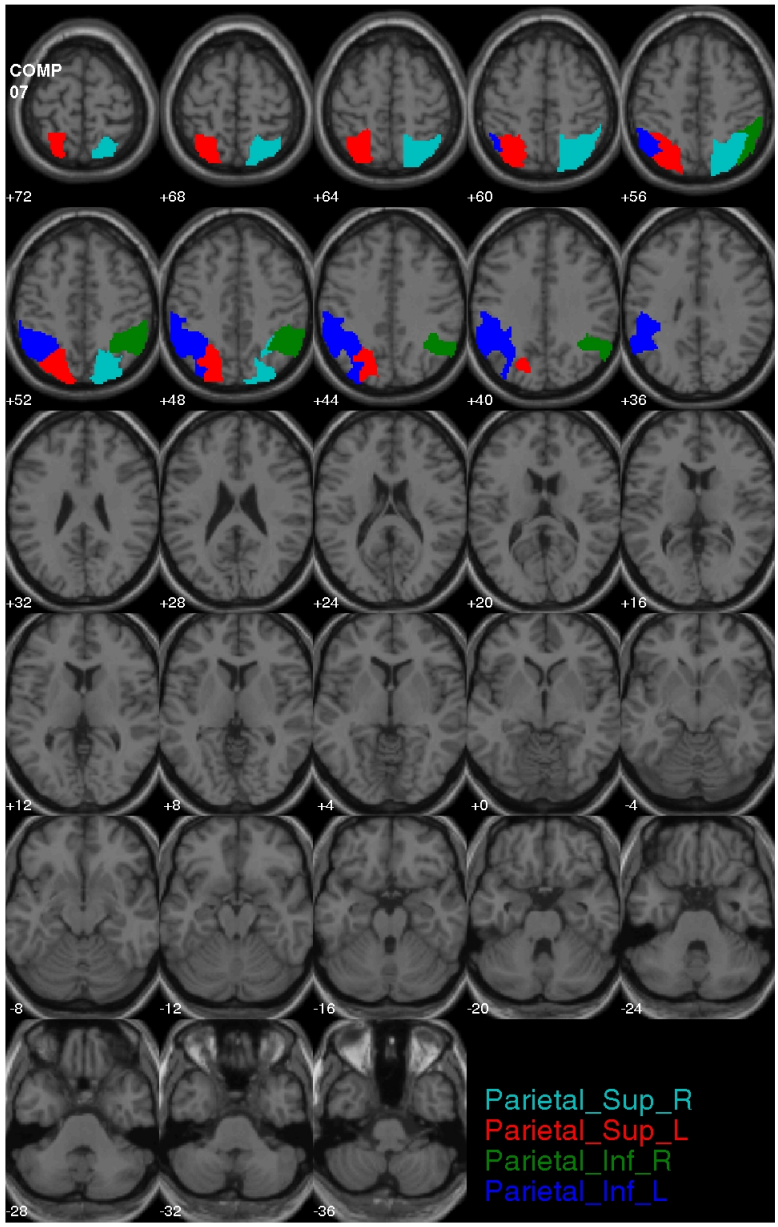

Figure C.7: The AAL regions of the 7. community detected by IRM, shown in the brain.

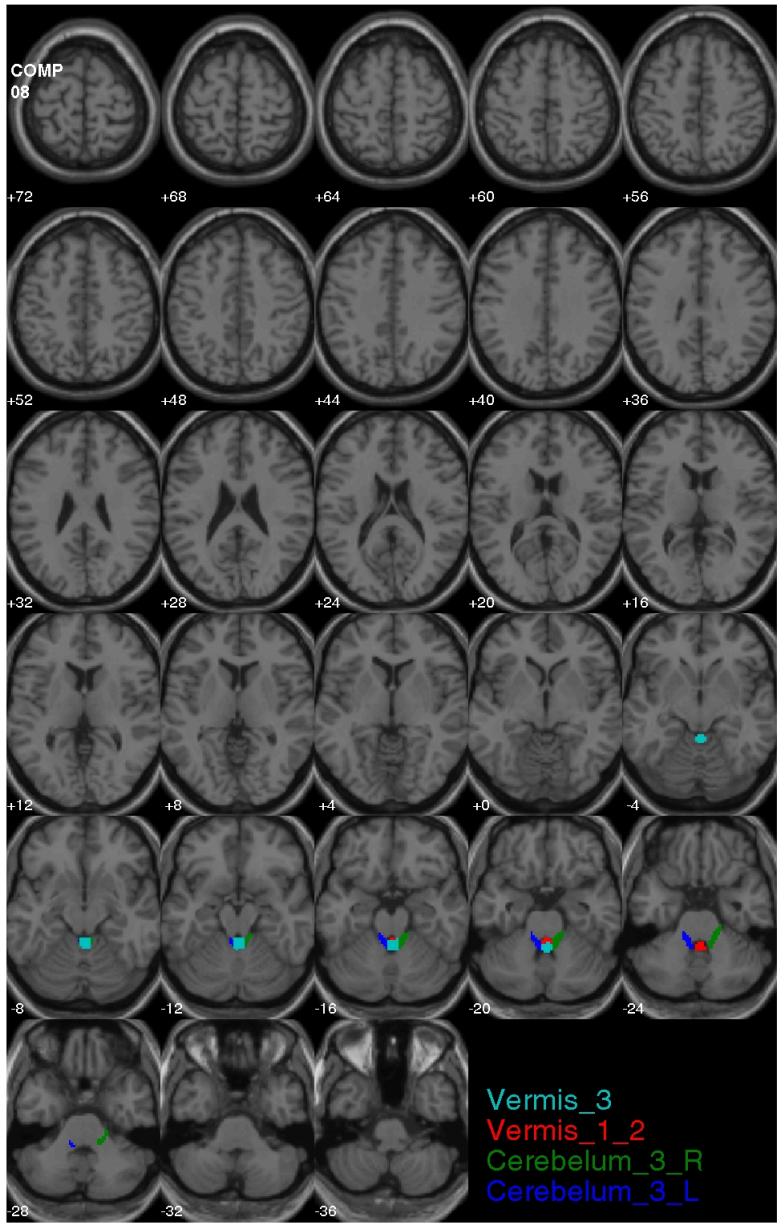

Figure C.8: The AAL regions of the 8. community detected by IRM, shown in the brain.

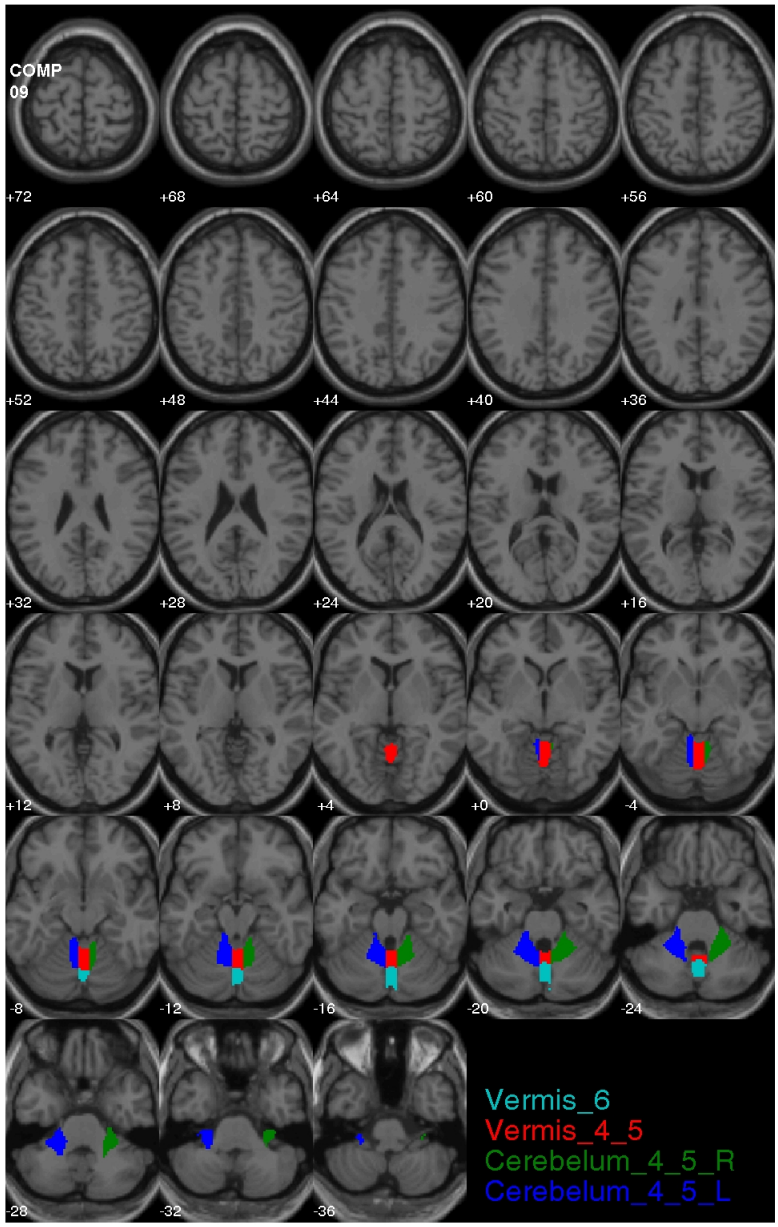

Figure C.9: The AAL regions of the 9. community detected by IRM, shown in the brain.

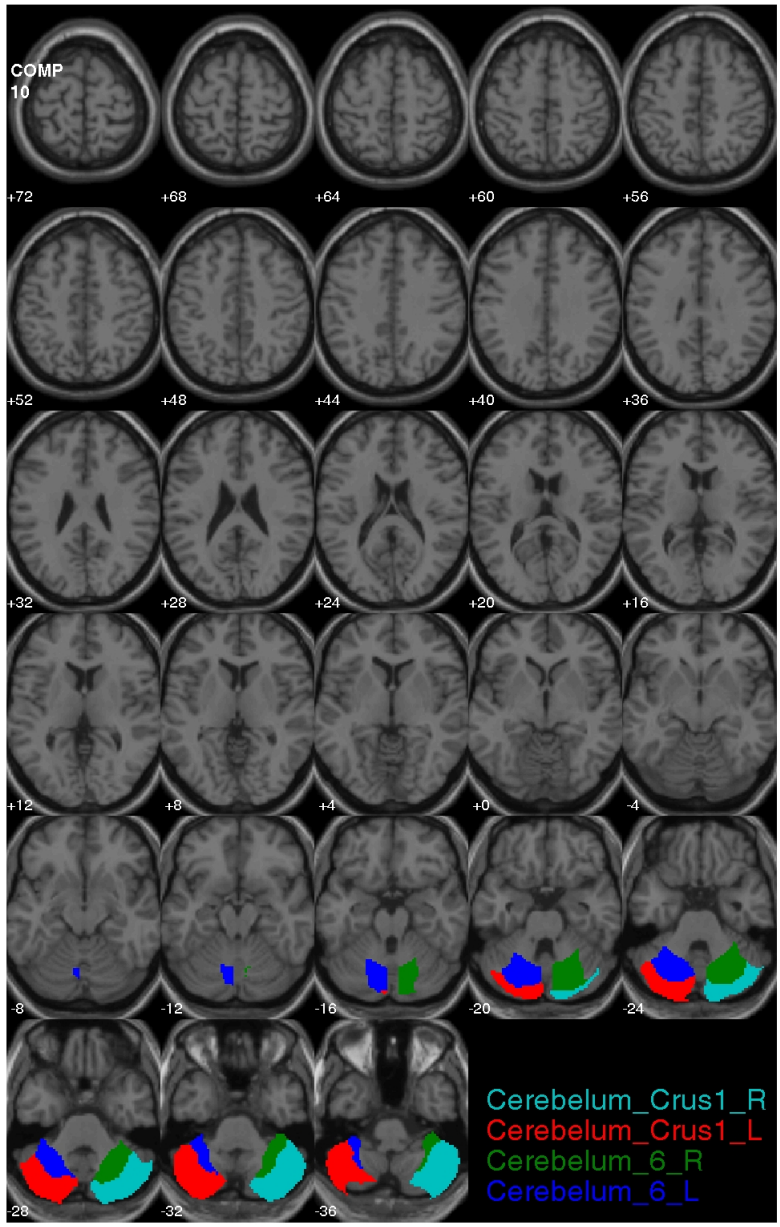

Figure C.10: The AAL regions of the 10. community detected by IRM, shown in the brain.

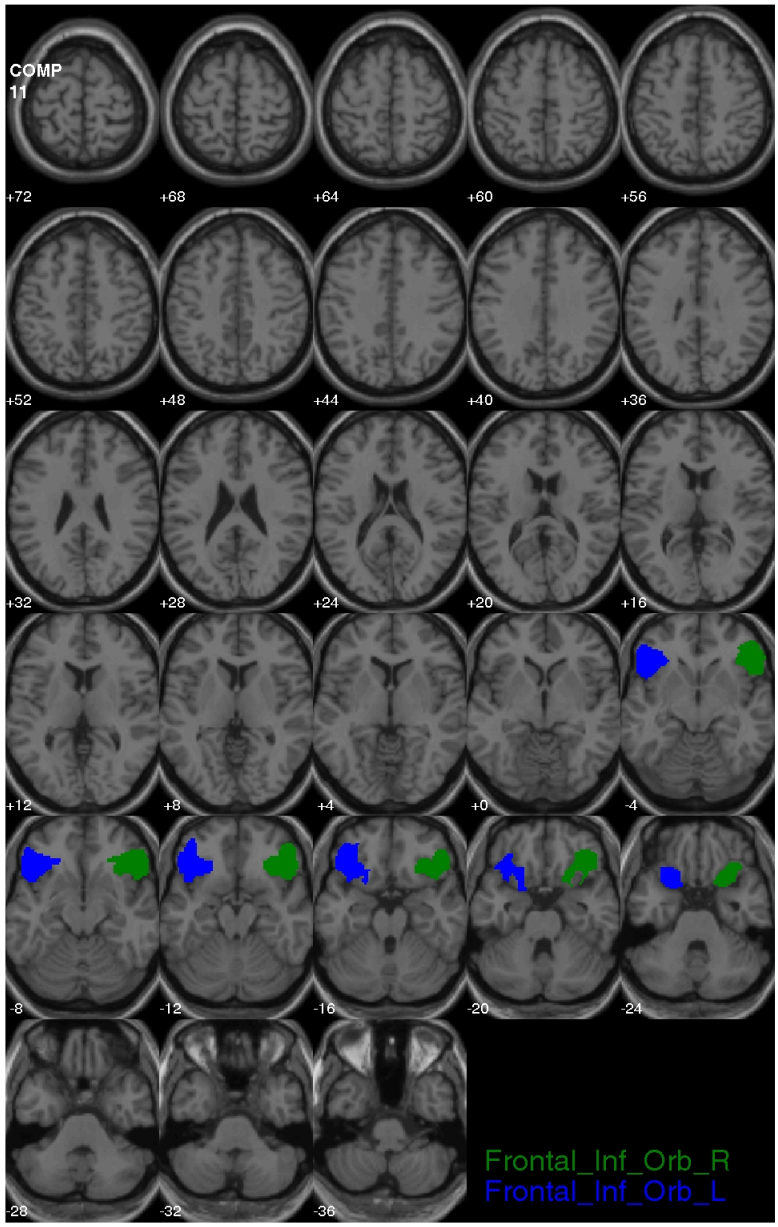

Figure C.11: The AAL regions of the 11. community detected by IRM, shown in the brain.

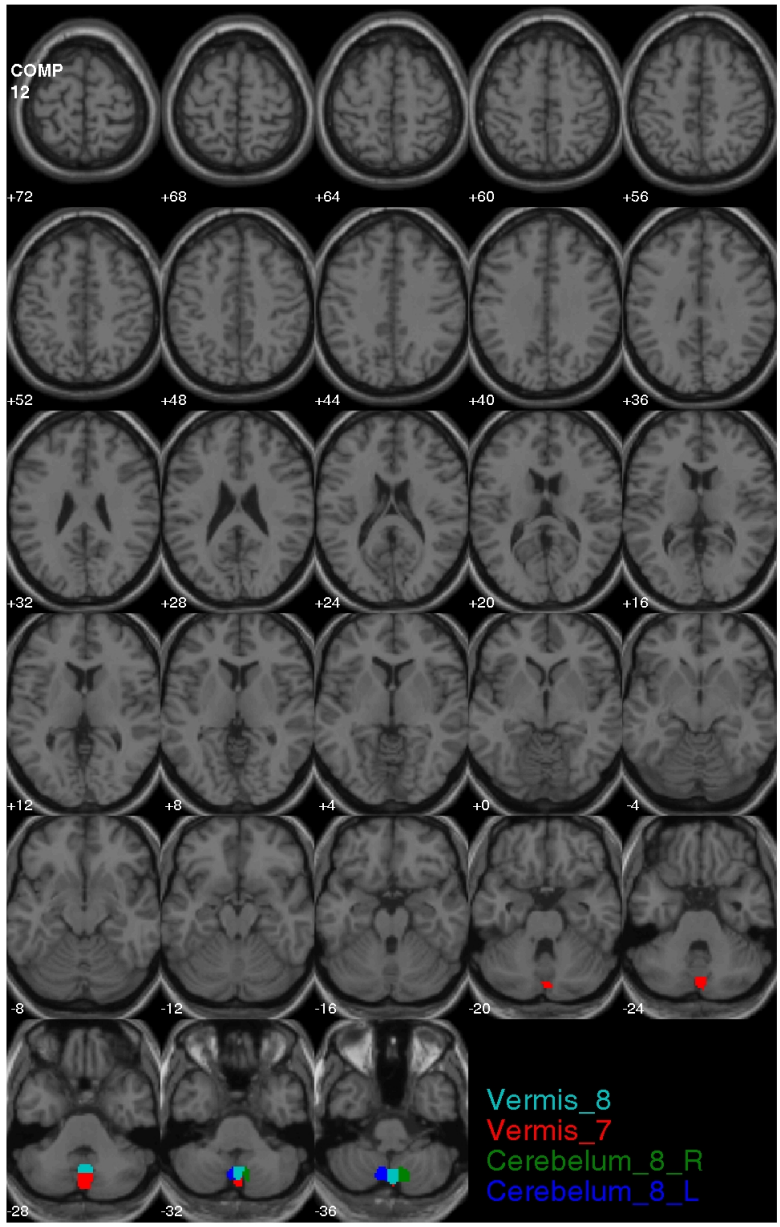

Figure C.12: The AAL regions of the 12. community detected by IRM, shown in the brain.

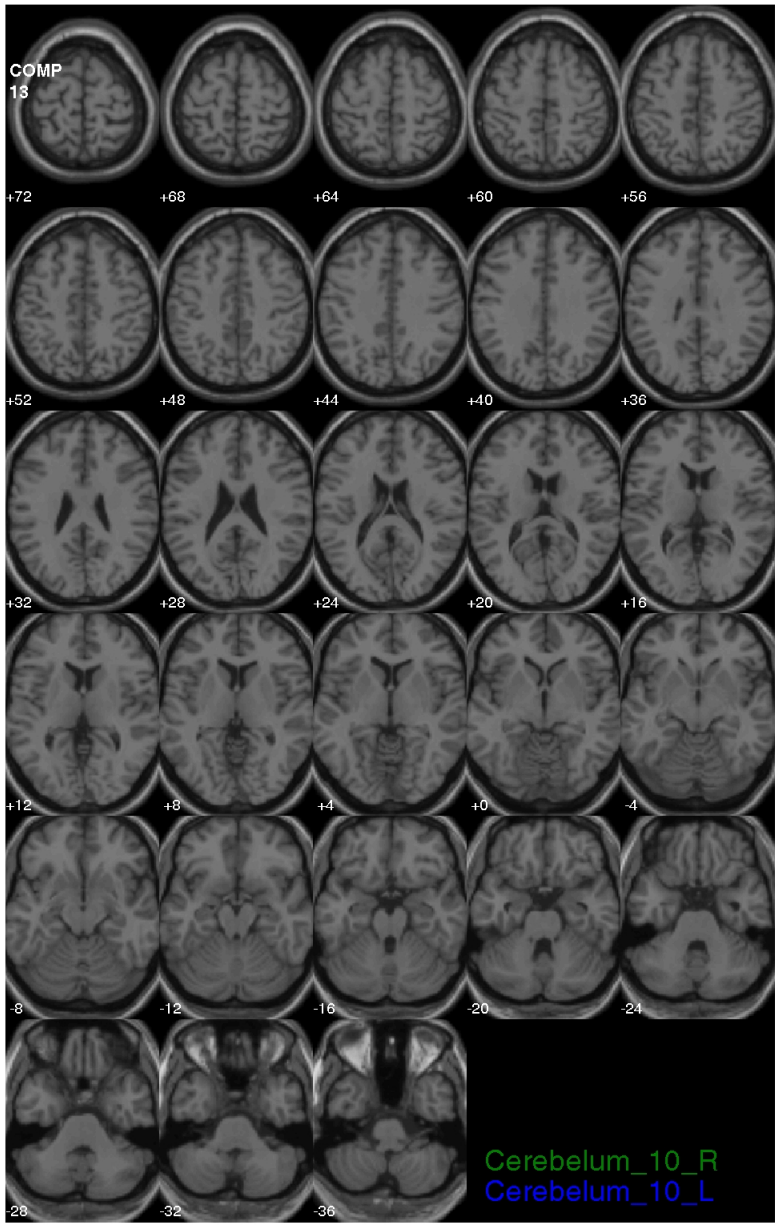

Figure C.13: The AAL regions of the 13. community detected by IRM, shown in the brain.

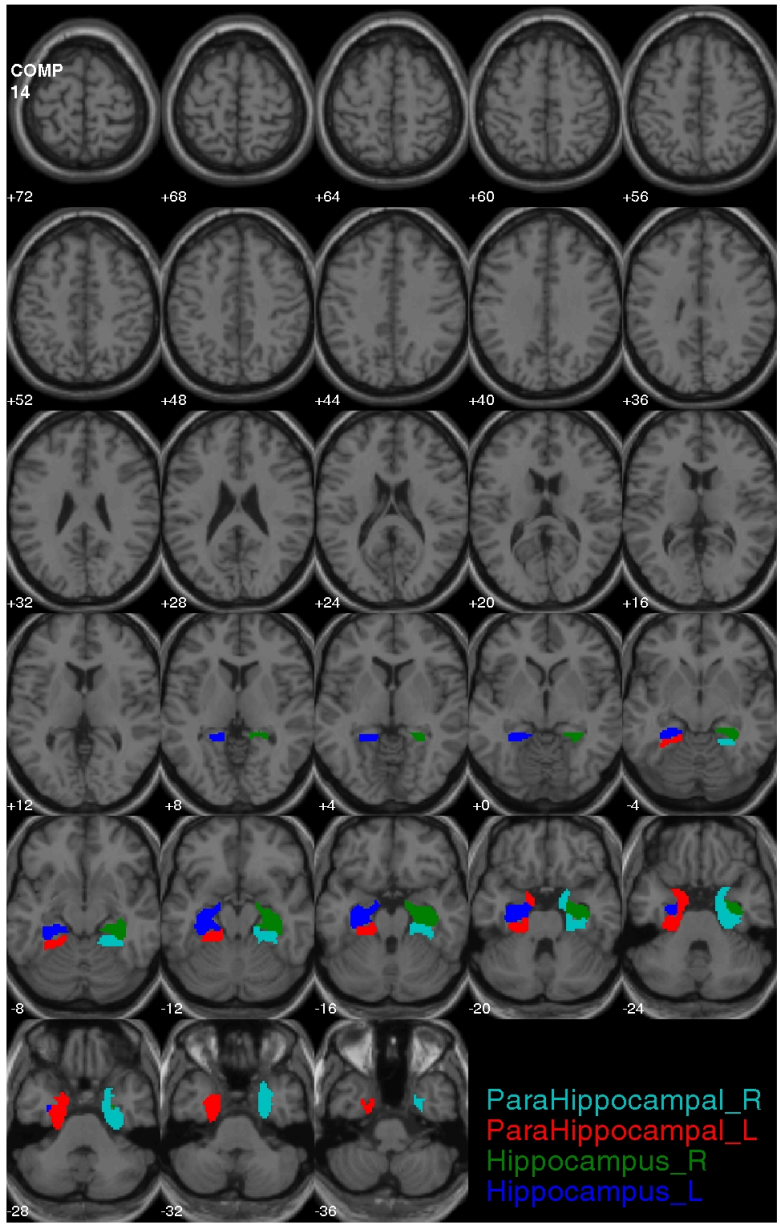

Figure C.14: The AAL regions of the 14. community detected by IRM, shown in the brain.

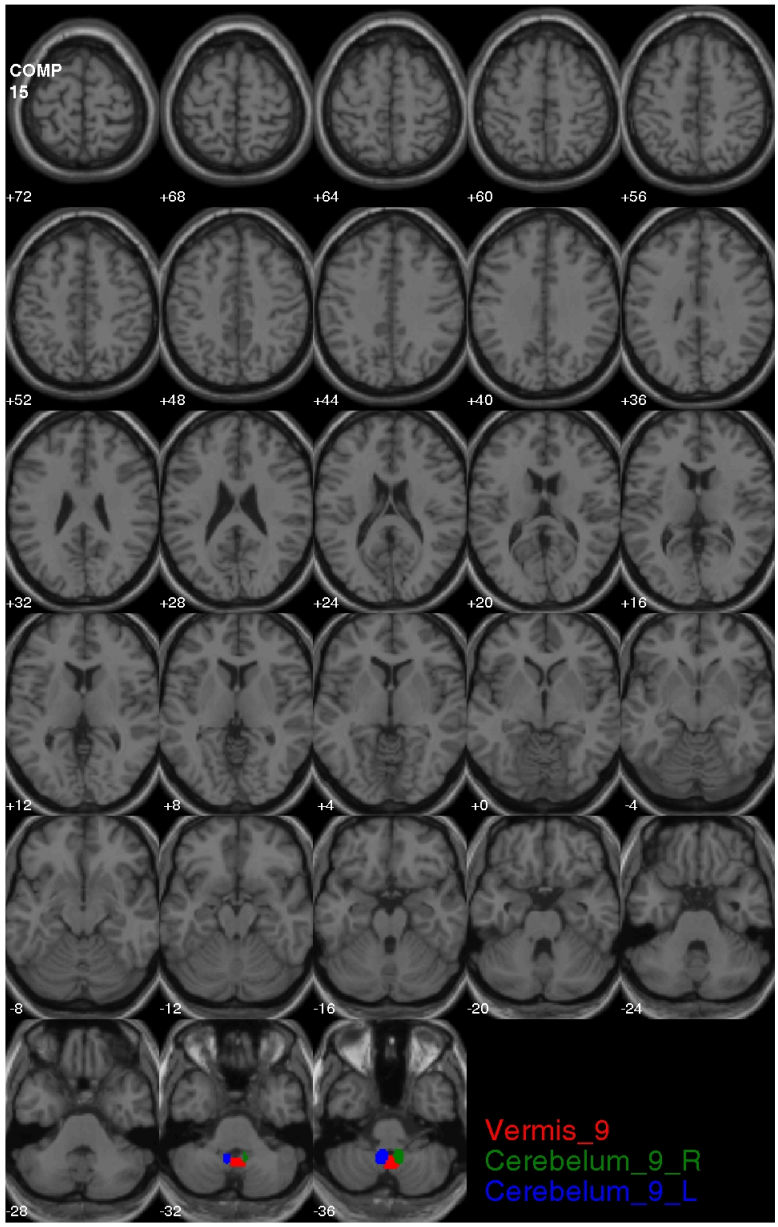

Figure C.15: The AAL regions of the 15. community detected by IRM, shown in the brain.

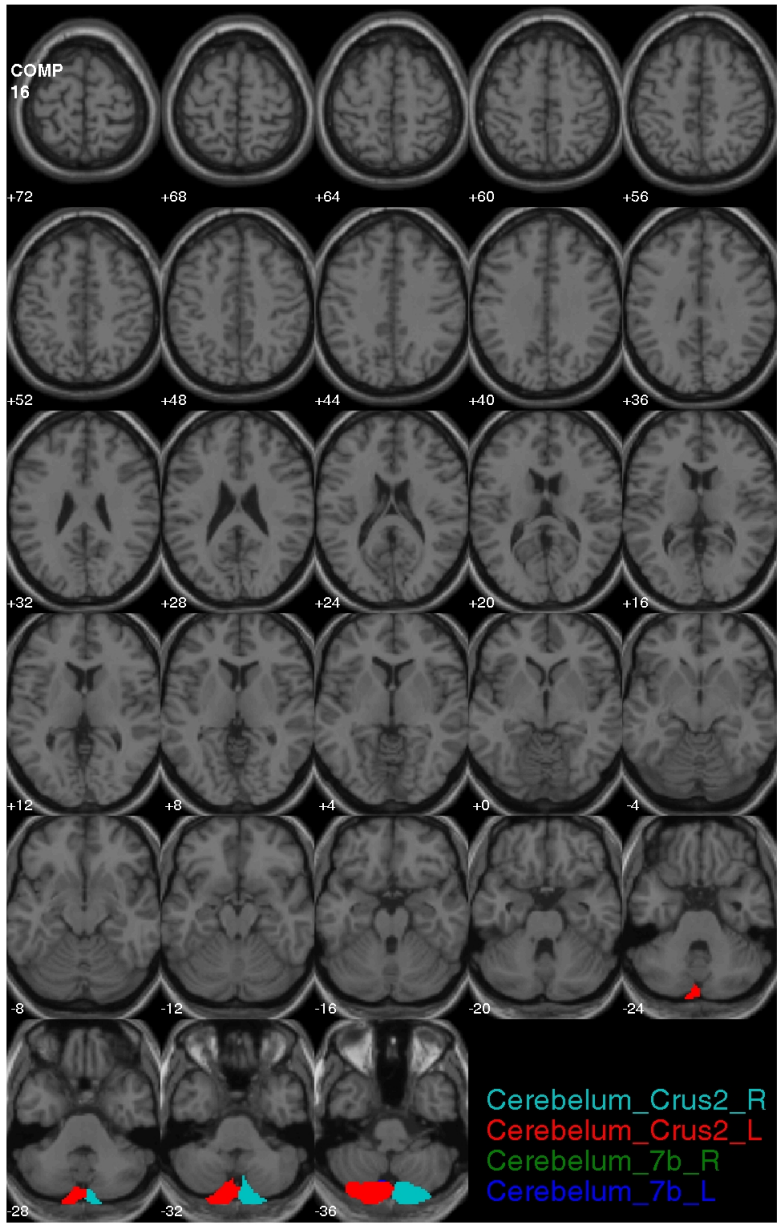

Figure C.16: The AAL regions of the 16. community detected by IRM, shown in the brain.

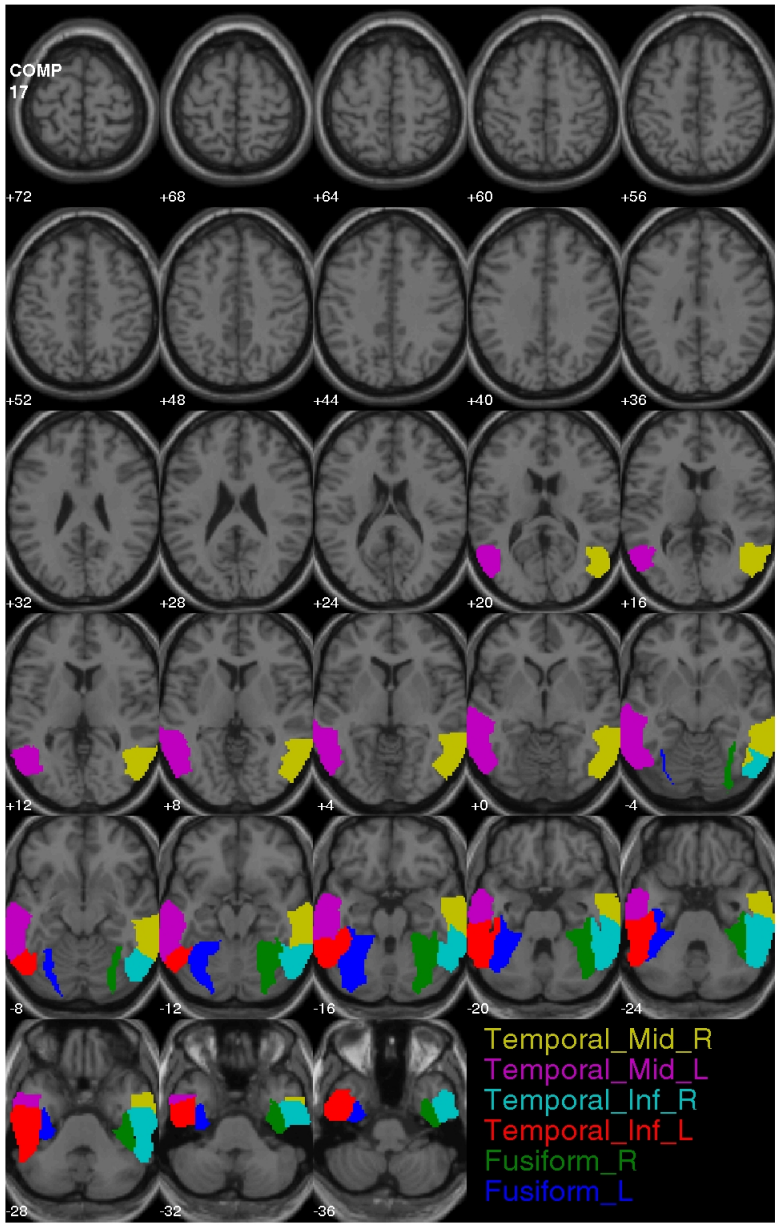

Figure C.17: The AAL regions of the 17. community detected by IRM, shown in the brain.

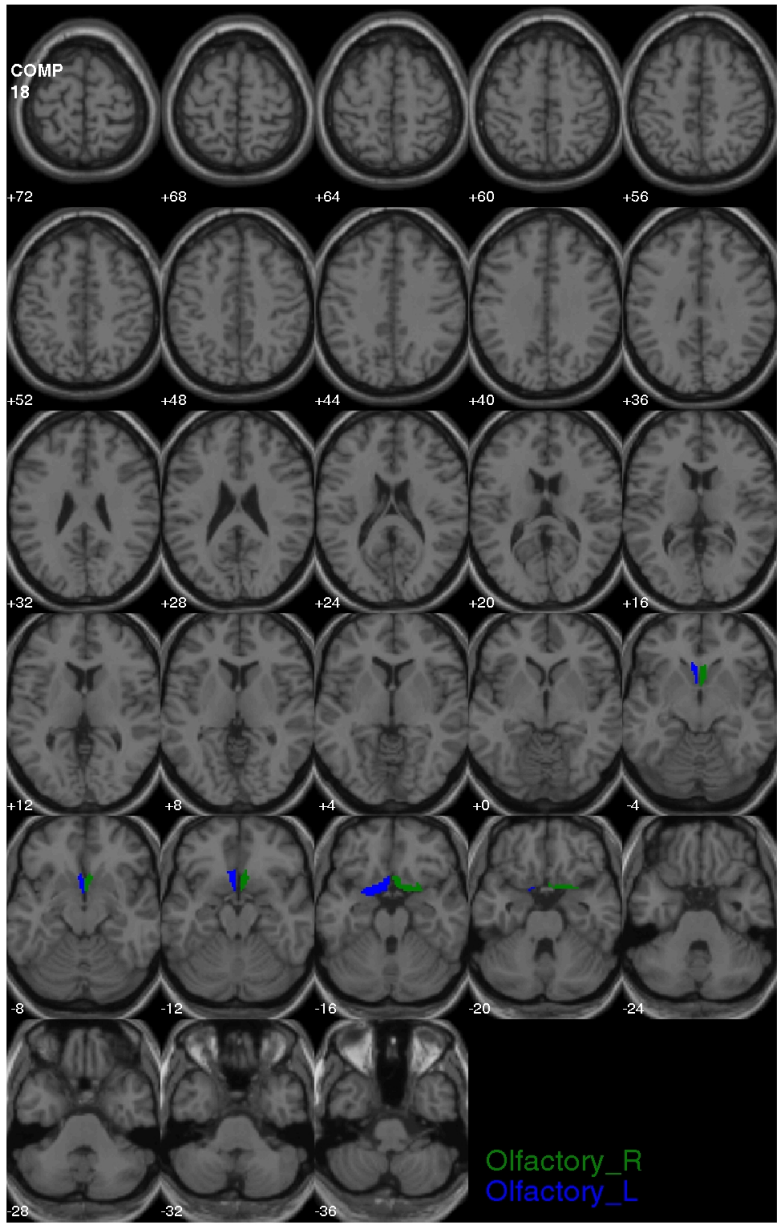

Figure C.18: The AAL regions of the 18. community detected by IRM, shown in the brain.

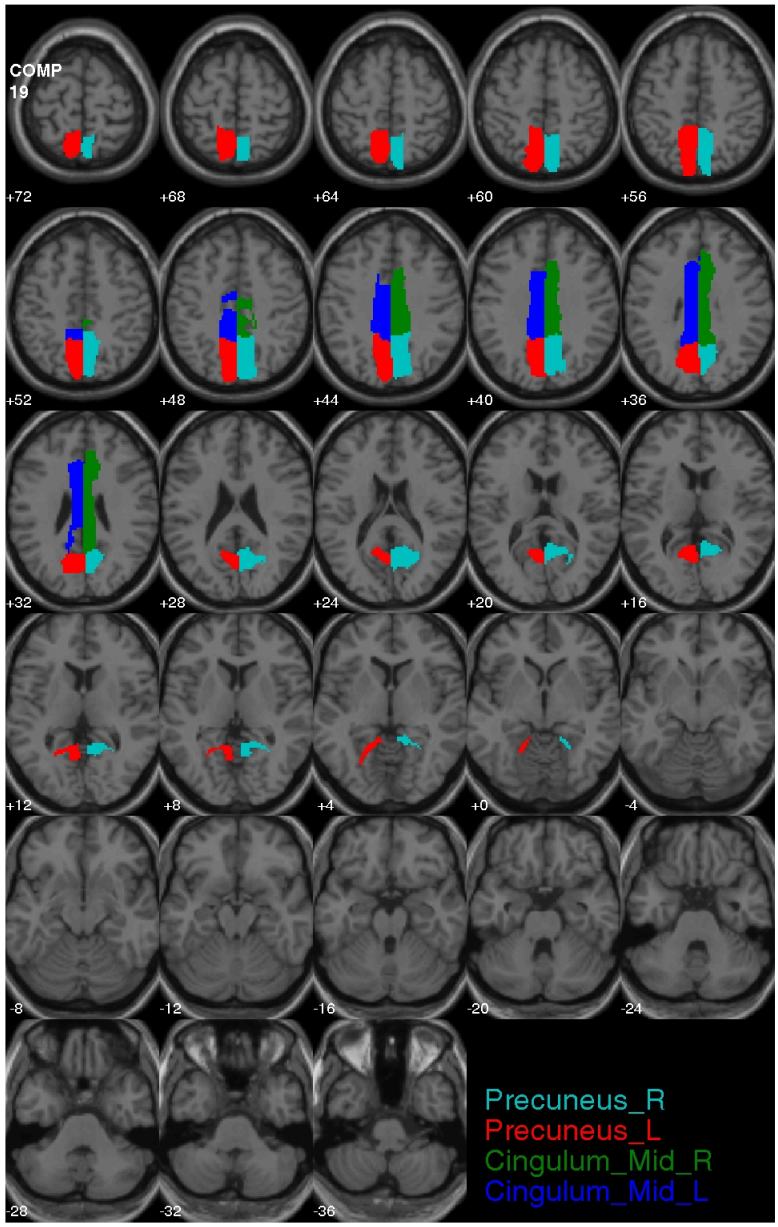

Figure C.19: The AAL regions of the 19. community detected by IRM, shown in the brain.

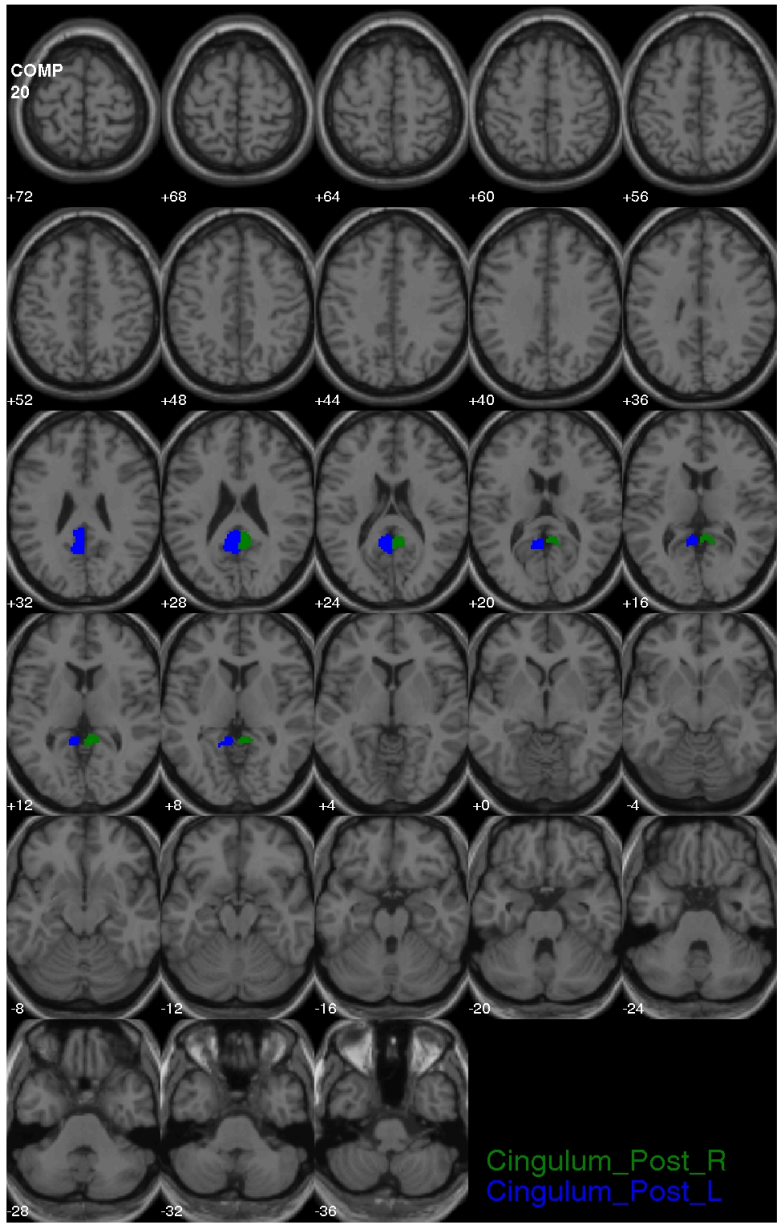

Figure C.20: The AAL regions of the 20. community detected by IRM, shown in the brain.

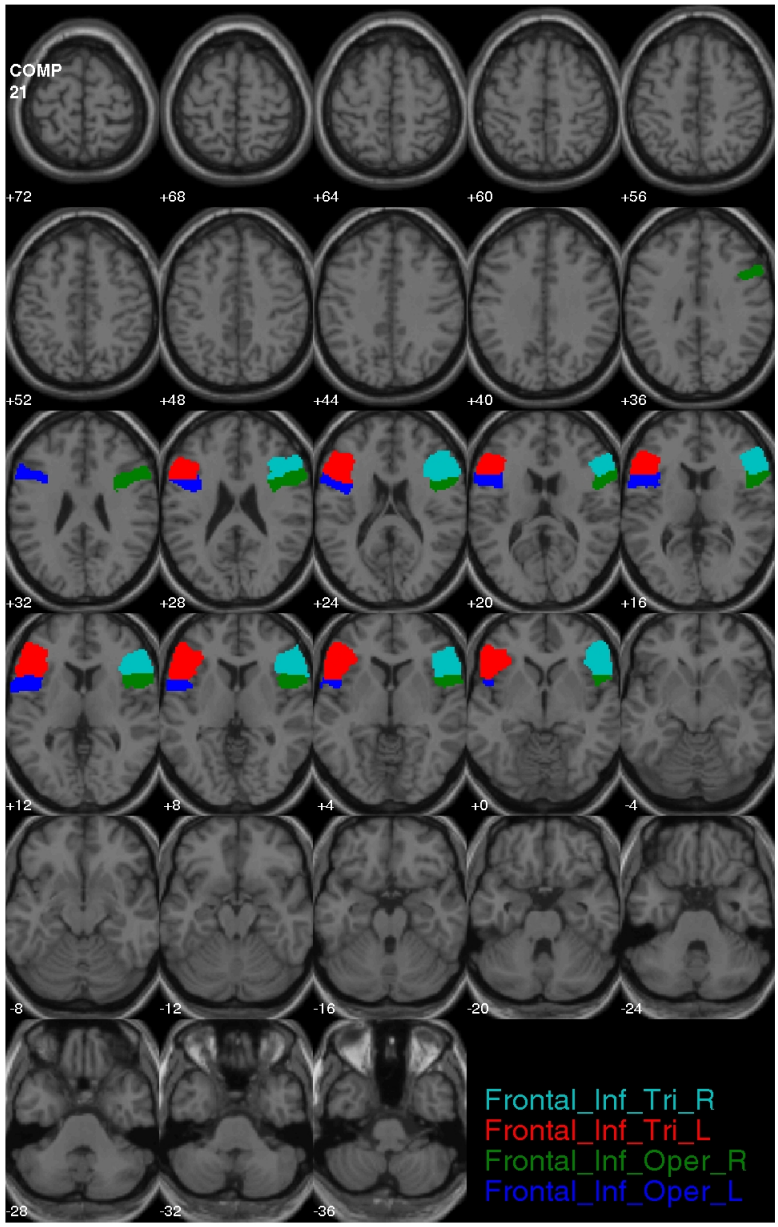

Figure C.21: The AAL regions of the 21. community detected by IRM, shown in the brain.

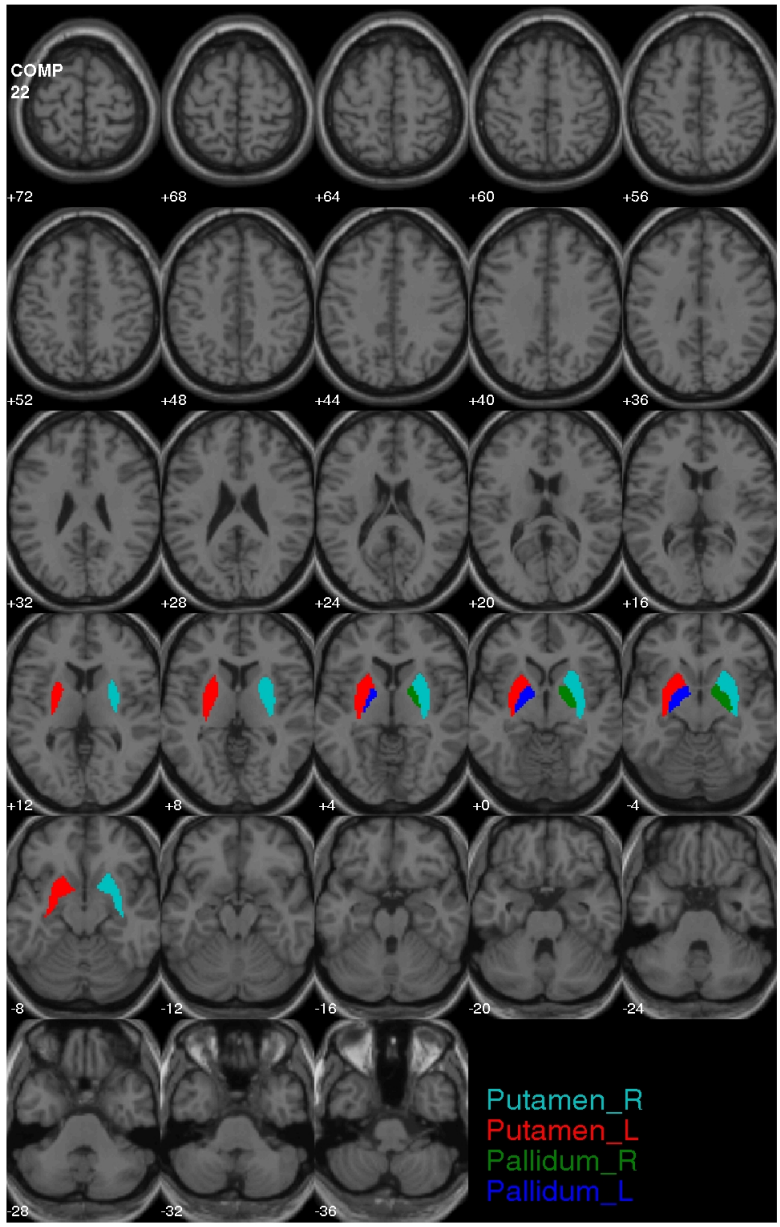

Figure C.22: The AAL regions of the 22. community detected by IRM, shown in the brain.

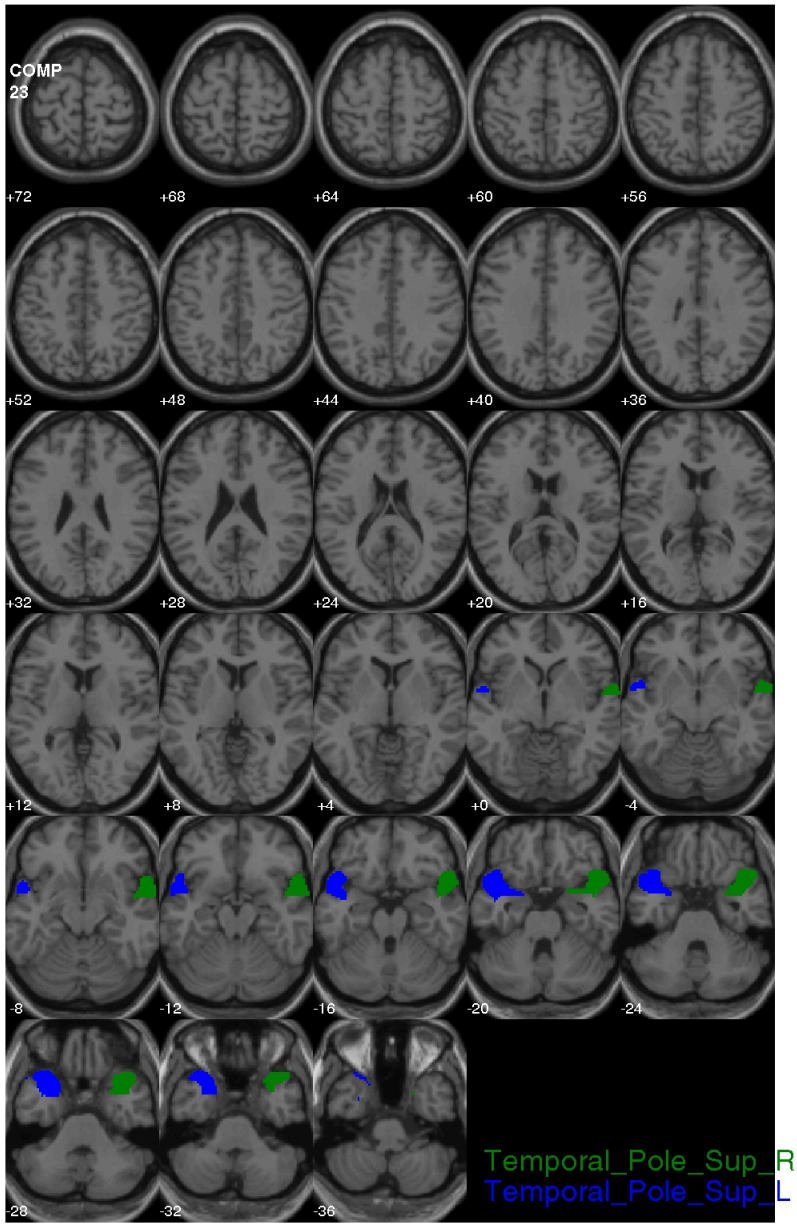

Figure C.23: The AAL regions of the 23. community detected by IRM, shown in the brain.

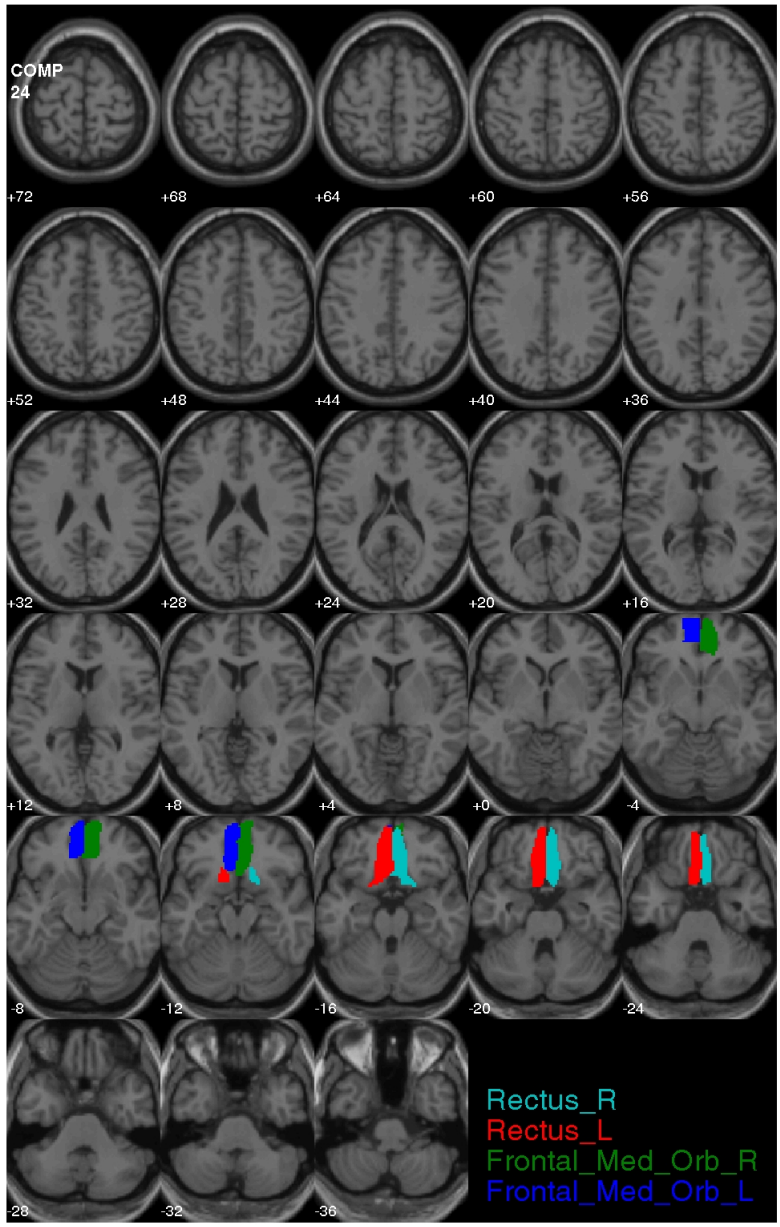

Figure C.24: The AAL regions of the 24. community detected by IRM, shown in the brain.

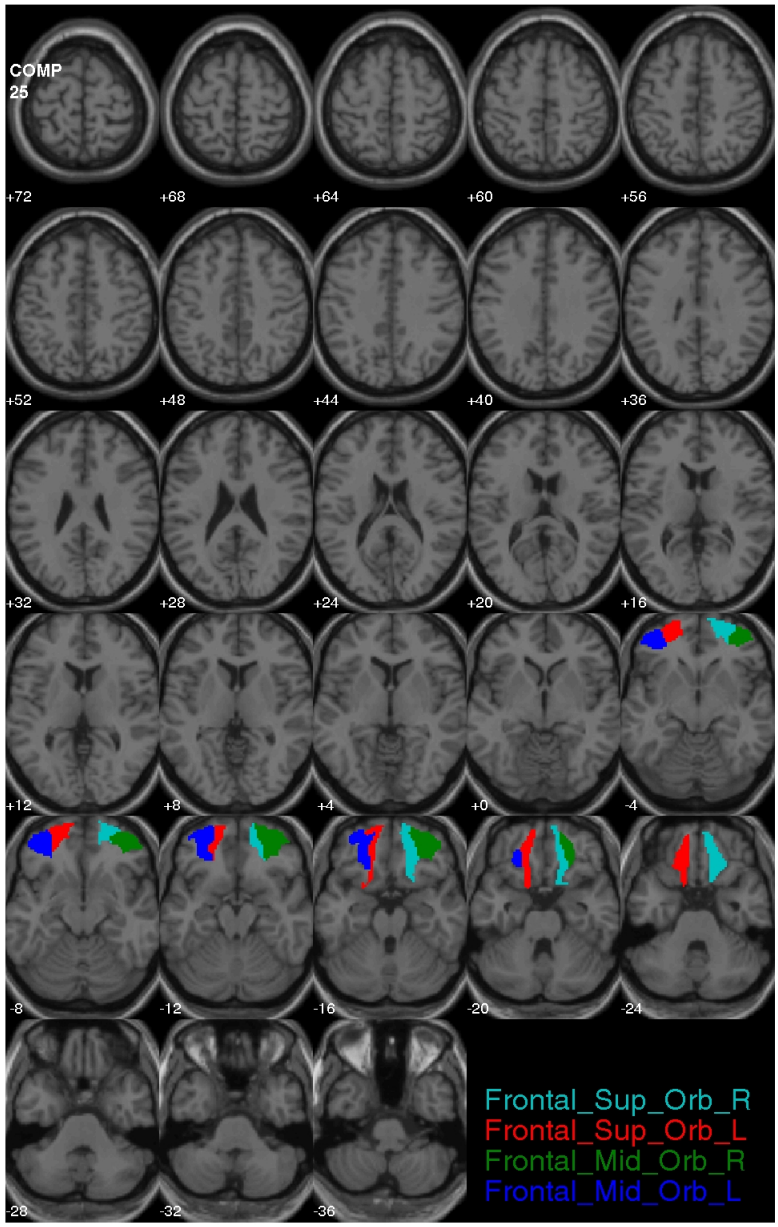

Figure C.25: The AAL regions of the 25. community detected by IRM, shown in the brain.

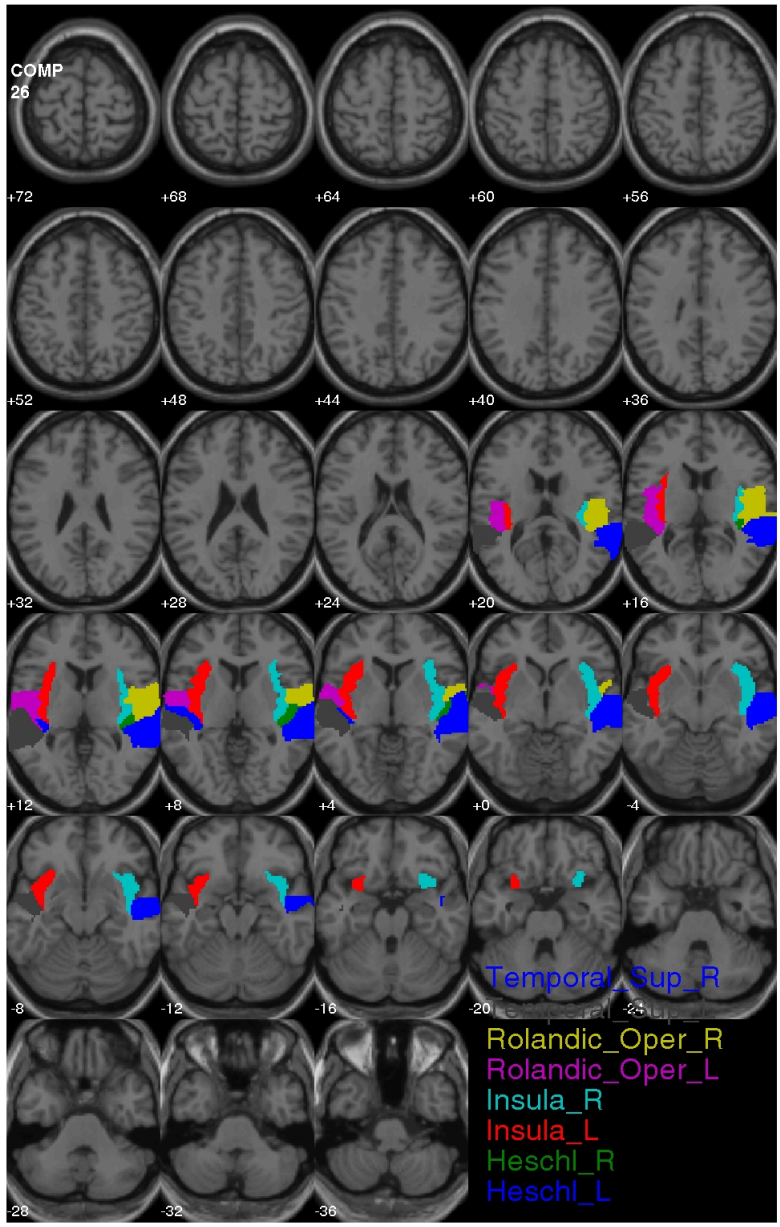

Figure C.26: The AAL regions of the 26. community detected by IRM, shown in the brain.

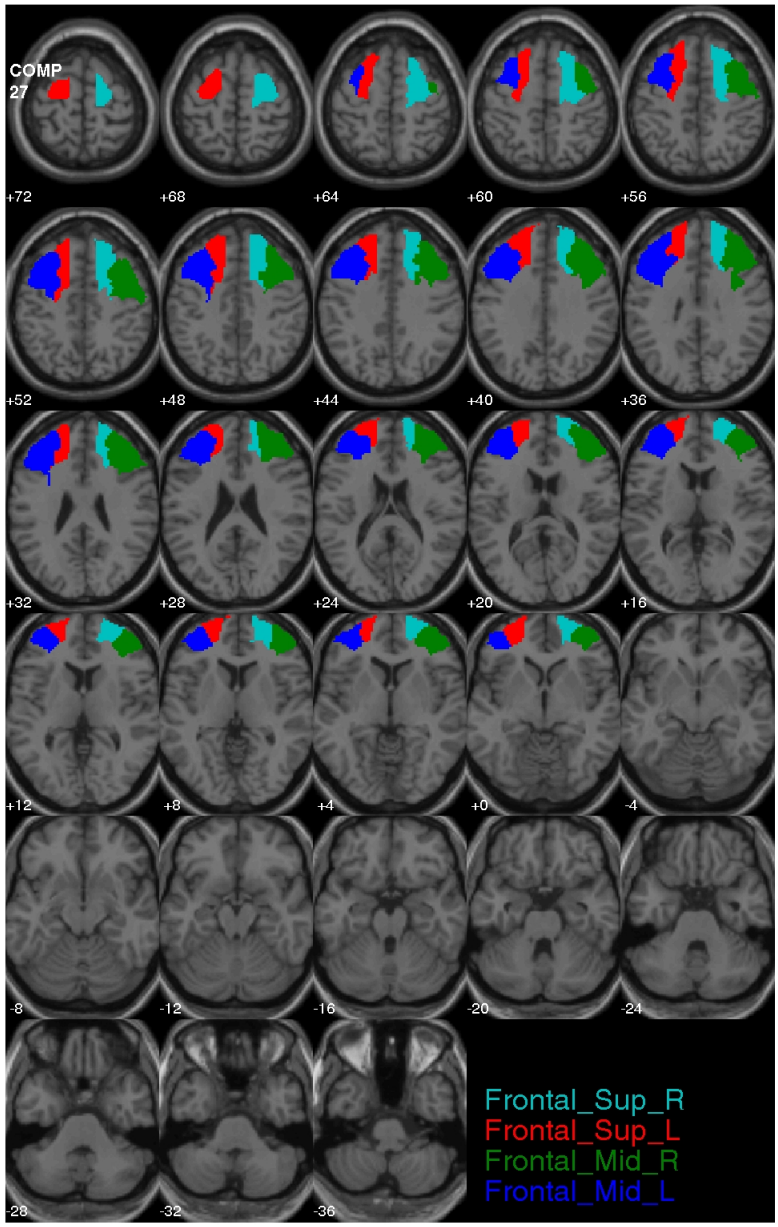

Figure C.27: The AAL regions of the 27. community detected by IRM, shown in the brain.

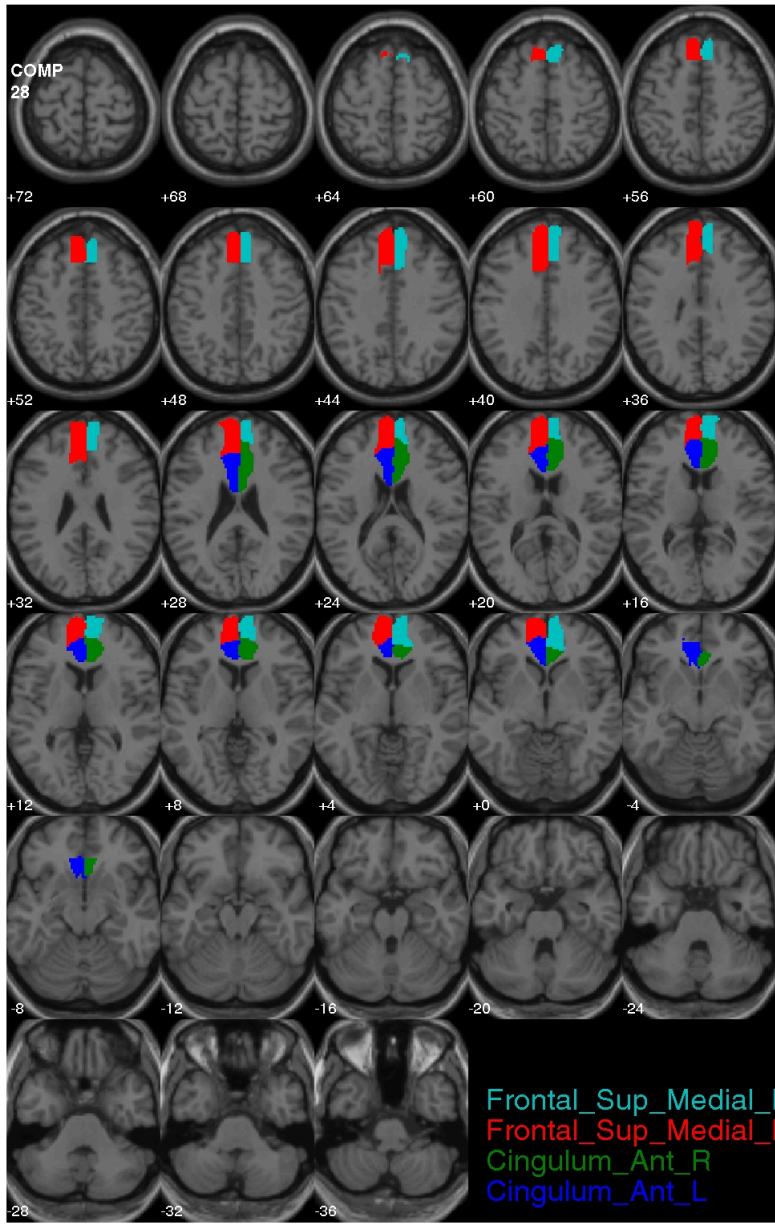

Figure C.28: The AAL regions of the 28. community detected by IRM, shown in the brain.

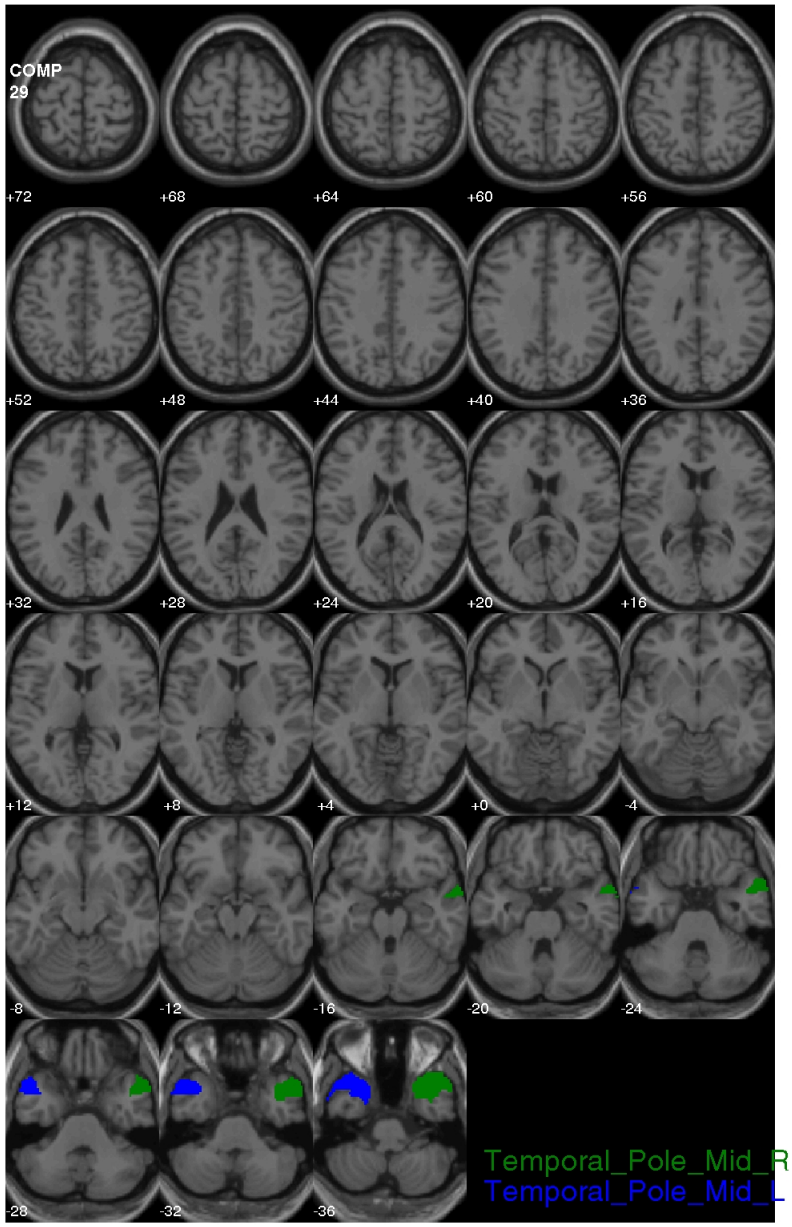

Figure C.29: The AAL regions of the 29. community detected by IRM, shown in the brain.

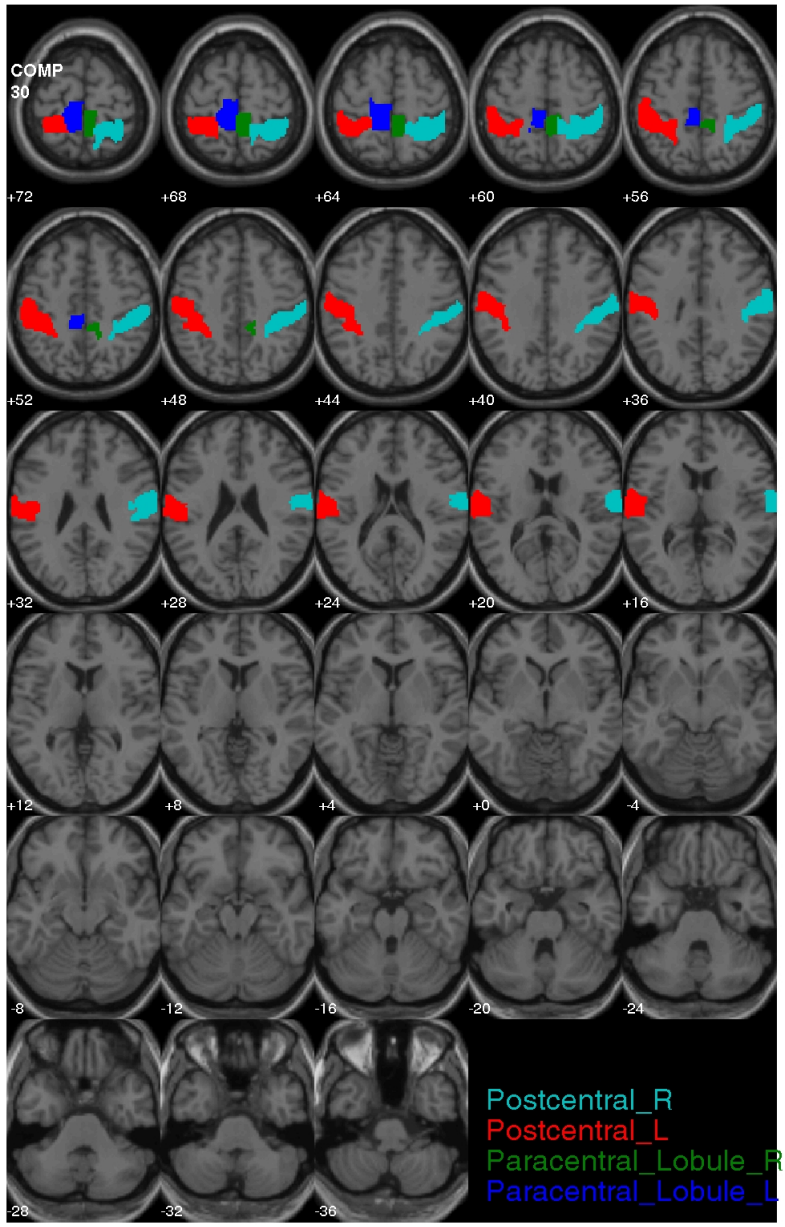

Figure C.30: The AAL regions of the 30. community detected by IRM, shown in the brain.

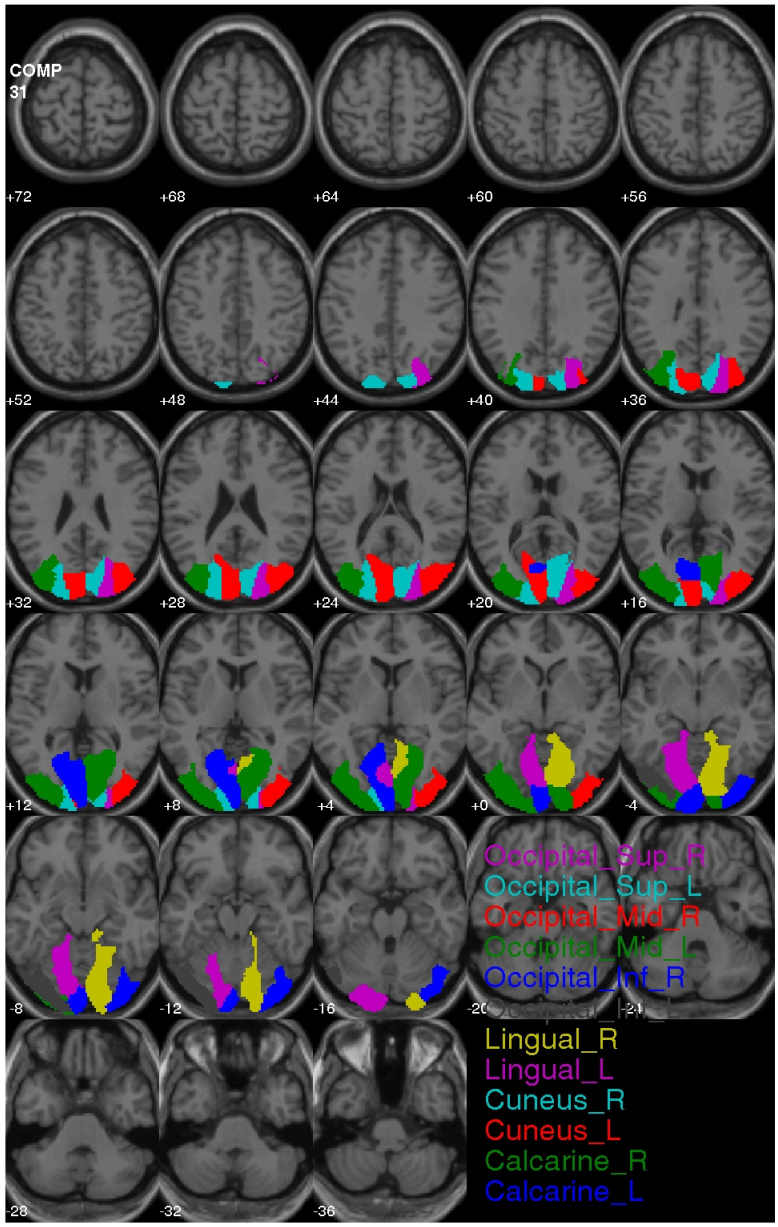

Figure C.31: The AAL regions of the 31. community detected by IRM, shown in the brain.

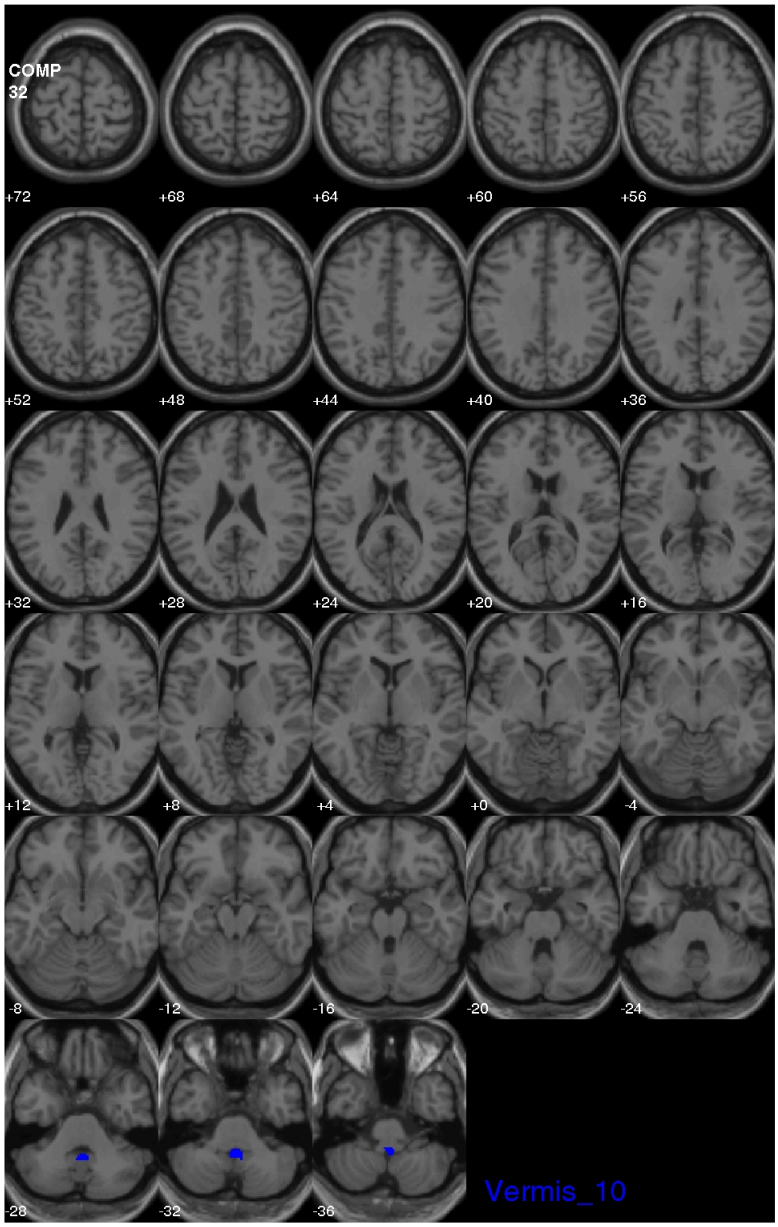

Figure C.32: The AAL regions of the 32. community detected by IRM, shown in the brain.

Bibliography

[1] The illuminated brain. `http://theilluminatedbrain.com/2010/09/`, January. Accessed: 28/01/2013.

[2] K.W. Andersen, M. Mørup, H. Siebner, K.H. Madsen, and L.K. Hansen. Identifying modular relations in complex brain networks. Technical report, DTU Informatics, Technical University of Denmark and Danish Research Centre for Magnetic Resonance, Cophenhagen University Hospital Hvidovre, 2012.

[3] J. Ashburner, G. Barnes, and et al. `http://www.fil.ion.ucl.ac.uk/spm/doc/manual.pdf`. Accessed: 28/01/2013.

[4] C.F. Beckmann, M. DeLuca, J.T. Devlin, and S.M. Smith. Investigations into resting-state connectivity using independent component analysis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1457):1001–1013, 2005.

[5] B. Biswal, F. Zerrin Yetkin, V.M. Haughton, and J.S. Hyde. Functional connectivity in the motor cortex of resting human brain using echo-planar mri. *Magnetic resonance in medicine*, 34(4):537–541, 1995.

[6] S. Bonavita, A. Gallo, R. Sacco, M. Della Corte, A. Bisecco, R. Docimo, L. Lavorgna, D. Corbo, A. Di Costanzo, F. Tortora, et al. Distributed changes in default-mode resting-state connectivity in multiple sclerosis. *Multiple Sclerosis Journal*, 17(4):411–422, 2011.

[7] R.L. Buckner, J.R. Andrews-Hanna, and D.L. Schacter. The brain's default network. *Annals of the New York Academy of Sciences*, 1124(1):1–38, 2008.

[8] Bishop. C.M. *Pattern Recognition and Machine Learning*. Springer, 2006.

[9] A. Compston and A. Coles. Multiple sclerosis. *Lancet*, 359:1221–1231, 2002.

[10] D. Cordes, V.M. Haughton, K. Arfanakis, J.D. Carew, P.A. Turski, C.H. Moritz, M.A. Quigley, and M.E. Meyerand. Frequencies contributing to functional connectivity in the cerebral cortex in "resting-state" data. *American Journal of Neuroradiology*, 22(7):1326–1333, 2001.

[11] J.S. Damoiseaux, S. Rombouts, F. Barkhof, P. Scheltens, C.J. Stam, S.M. Smith, and C.F. Beckmann. Consistent resting-state networks across healthy subjects. *Proceedings of the national academy of sciences*, 103(37):13848–13853, 2006.

[12] M. De Luca, C.F. Beckmann, N. De Stefano, P.M. Matthews, and S.M. Smith. fmri resting state networks define distinct modes of long-distance interactions in the human brain. *Neuroimage*, 29(4):1359–1367, 2006.

[13] A-M. Dogonowski. *Resting-state Functional Connectivity of the Motor System in Multiple Sclerosis*. PhD thesis, Faculty Of Health and Medical Science, 2012.

[14] J. Eising. *Lineær ALGEBRA*. Institut for Matematik Danmarks Tekniske Universitet, 2006.

[15] T. Fawcett. An introduction to roc analysis. *Pattern recognition letters*, 27(8):861–874, 2006.

[16] M. Filippi and M.A. Rocca. Mri evidence for multiple sclerosis as a diffuse disease of the central nervous system. *Journal of neurology*, 252:16–24, 2005.

[17] S. Fortunato. Community detection in graphs. *Physics Reports*, 486(3):75–174, 2010.

[18] M.D. Fox and M.E. Raichle. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature Reviews Neuroscience*, 8(9):700–711, 2007.

[19] M.D. Fox, A.Z. Snyder, J.L. Vincent, M. Corbetta, D.C. Van Essen, and M.E. Raichle. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences of the United States of America*, 102(27):9673–9678, 2005.

[20] K. Friston. `http:_/_/www.fil.ion.ucl.ac.uk_/spm_/doc_/intro_/`. Accessed: 28/01/2013.

[21] A. Goldenberg, A.X. Zheng, S.E. Fienberg, and E.M. Airoldi. A survey of statistical network models. *Foundations and Trends® in Machine Learning*, 2(2):129–233, 2010.

[22] M.D. Greicius, B. Krasnow, A.L. Reiss, and V. Menon. Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences*, 100(1):253–258, 2003.

[23] M.D. Greicius, G. Srivastava, A.L. Reiss, and V. Menon. Default-mode network activity distinguishes alzheimer's disease from healthy aging: evidence from functional mri. *Proceedings of the National Academy of Sciences of the United States of America*, 101(13):4637–4642, 2004.

[24] J.M. Hofman and C.H. Wiggins. Bayesian approach to network modularity. *Physical review letters*, 100(25):258701, 2008.

[25] S. Jain and R.M. Neal. A split-merge markov chain monte carlo procedure for the dirichlet process mixture model. *Journal of Computational and Graphical Statistics*, 13(1):158–182, 2004.

[26] B. Karrer and M.E.J. Newman. Stochastic blockmodels and community structure in networks. *Physical Review E*, 83(1):016107, 2011.

[27] C. Kemp, J.B. Tenenbaum, T.L. Griffiths, T. Yamada, and N. Ueda. Learning systems of concepts with an infinite relational model. In *Proceedings of the national conference on artificial intelligence*, volume 21, page 381. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, 2006.

[28] J.F. Kurtzke. Rating neurologic impairment in multiple sclerosis an expanded disability status scale (edss). *Neurology*, 33(11):1444–1444, 1983.

[29] A.R. Laird, S.B. Eickhoff, K. Li, D.A. Robin, D.C. Glahn, and P.T. Fox. Investigating the functional heterogeneity of the default mode network using coordinate-based meta-analytic modeling. *The Journal of Neuroscience*, 29(46):14496–14505, 2009.

[30] E. Lesage, M.A.J. Apps, A.L. Hayter, C.F. Beckmann, D. Barnes, D.W. Langdon, and N. Ramnani. Cerebellar information processing in relapsing-remitting multiple sclerosis (rrms). *Behavioural Neurology*, 23(1):39–49, 2010.

[31] Hanson. L.G. Introduction to magnetic resonance imaging techniques. `www.eprints.drcmr.dk/37/2/MRI_English_letter.pdf`, August 2009. Accessed: 28/01/2013.

[32] H.K. Mathiesen and P.S. Sørensen. Ny simplere diagnosticering af multipel sklerose. *Ugeskrift For Læger*, 174(13):862–865, 2012.

[33] P.M. Matthews and P. Jezzard. Functional magnetic resonance imaging. *Journal of Neurology, Neurosurgery & Psychiatry*, 75(1):6–12, 2004.

[34] M. Mørup, K.H. Madsen, A.M. Dogonowski, H. Siebner, and L.K. Hansen. Infinite relational modeling of functional connectivity in resting state fmri. *Neural Information Processing Systems 23*, 2010.

[35] M. Mørup and M.N. Schmidt. Bayesian community detection. *Neural Computation*, 24(9):2434–2456, 2012.

[36] T.J Nowak and A.G. Handford. *Pathophysiology.* The McGraw-Hill Companies, Inc., 2004.

[37] S. Ogawa, T.M Lee, A.R Kay, and D.W Tank. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences*, 87(24):9868–9872, 1990.

[38] C.E. Pizoli, M.N. Shah, A.Z. Snyder, J.S. Shimony, D.D. Limbrick, M.E. Raichle, B.L. Schlaggar, and M.D. Smyth. Resting-state activity in development and maintenance of normal brain function. *Proceedings of the National Academy of Sciences*, 108(28):11638–11643, 2011.

[39] M.E. Raichle. The restless brain. *Brain Connectivity*, 1(1):3–12, 2011.

[40] M.E. Raichle, A.M. MacLeod, A.Z. Snyder, W.J. Powers, D.A. Gusnard, and G.L. Shulman. A default mode of brain function. *Proceedings of the National Academy of Sciences*, 98(2):676–682, 2001.

[41] R. Salvador, J. Suckling, M.R. Coleman, J.D. Pickard, D. Menon, and E.D. Bullmore. Neurophysiological architecture of functional magnetic resonance images of human brain. *Cerebral Cortex*, 15(9):1332–1342, 2005.

[42] M. N. Schmidt and M. Mørup. Non-parametric bayesian modeling of complex networks. Technical report, Cognitive Systems, DTU Informatics, Technical University of Denmark, 2012.

[43] B.D. Trapp, J. Peterson, R.M. Ransohoff, R. Rudick, S. Mörk, and L. Bö. Axonal transection in the lesions of multiple sclerosis. *New England Journal of Medicine*, 338(5):278–285, 1998.

[44] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, M. Joliot, et al. Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain. *Neuroimage*, 15(1):273–289, 2002.

[45] M.P. van den Heuvel and H.E. Hulshoff Pol. Exploring the brain network: a review on resting-state fmri functional connectivity. *European Neuropsychopharmacology*, 20(8):519–534, 2010.

[46] M.P. Van den Heuvel, C.J. Stam, M. Boersma, P.H.E. Hulshoff, et al. Small-world and scale-free organization of voxel-based resting-state functional connectivity in the human brain. *Neuroimage*, 43(3):528, 2008.

[47] K.R.A. Van Dijk, T. Hedden, A. Venkataraman, K.C. Evans, S.W. Lazar, and R.L. Buckner. Intrinsic functional connectivity as a tool for human connectomics: theory, properties, and optimization. *Journal of neurophysiology*, 103(1):297–321, 2010.

[48] S.G. Waxman. Axonal conduction and injury in multiple sclerosis: the role of sodium channels. *Nature Reviews Neuroscience*, 7(12):932–941, 2006.

[49] W.M. Wells, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis. Multi-modal volume registration by maximization of mutual information. *Medical image analysis*, 1(1):35–51, 1996.

[50] D. Zhang and M.E. Raichle. Disease and the brain's dark energy. *Nature Reviews Neurology*, 6(1):15–28, 2010.