

École polytechnique de Louvain

Biomarker identification for vertigo through human connectome analysis

Authors: **Laura ANIZET, Astrid DOYEN**

Supervisors: **Benoit MACQ, Laurence DRICOT**

Readers: **Frédéric CREVECOEUR, Naïma DEGGOUJ, Nicolas DELINTE**

Academic year 2021–2022

Master [120] in Biomedical Engineering

Abstract

Thanks to advances in magnetic resonance imaging (MRI), interest in the investigation of neurological disorders, such as some types of vestibular dysfunction, has increased. Vestibular system dysfunction can lead to vertigo, described as an abnormal sensation of motion. The objective of this work is to analyse diffusion and functional MRI images to identify connections between regions of the brain allowing to discriminate people suffering from vertigo from control participants. To achieve this goal, structural connectivity matrices, reflecting the fibers density between brain areas, and functional connectivity matrices, assessing the simultaneous activity between regions of the brain at rest, have been constructed. Statistical tests have been performed and two models commonly used in machine learning algorithms, the Support Vector Machine and the Random Forest, have been used. Anatomical connections between regions involved in cognitive, proprioceptive, and affective processing appear to be significant in discriminating control participants from vertigo participants. For both structural and functional features, regions located in the frontal lobe (such as the precentral gyrus and orbital gyrus), in the limbic lobe (like the cingulate gyrus) or the occipital lobe are likely to play a role in the vestibular system and its performances. To a lesser extent, regions in the temporal lobe (such as the fusiform, middle and superior temporal gyrus) and in the parietal lobe (such as the precuneus) also seem to contribute to vestibular processing. This should encourage researchers to study functional and structural connections of the vestibular system to strengthen the results of this work. Combined with a supplementary biological input from doctors, these findings could help for the diagnosis and treatment of vertigo patients.

Acknowledgements

We would like to thank Professor Benoit Macq for giving us the opportunity to work on this topic.

We would also like to thank Professor Laurence Dricot, Professor Naima Deggouj and Dr. Anais Gregoire for their help, advice and support. We would like to send a special thank you to Anais Gregoire for the time she spent preparing our data.

We would also like to sincerely thank our assistants Nicolas Delinte and Quentin Dessain for their support, their answers to our questions, the time they spent with us, their reactivity, their advice and their comments on this work.

Finally, we would like to thank Dany Rimez for his collaboration and all the people in our academic and personal spheres who have supported and encouraged us during our studies.

Technical resources have also made this master thesis possible. Computational resources have been provided by the supercomputing facilities of the Université catholique de Louvain (CISM/UCL) and the Consortium des Équipements de Calcul Intensif en Fédération Wallonie Bruxelles (CÉCI) funded by the Fond de la Recherche Scientifique de Belgique (F.R.S.-FNRS) under convention 2.5020.11 and by the Walloon Region.

Contents

Abstract	i
Acknowledgements	ii
List of Figures	viii
List of Tables	xi
List of Abbreviations	xiii
Introduction	1
I Theoretical Background	3
1 Magnetic Resonance Imaging	4
1.1 Principles of Magnetic Resonance Imaging	4
1.1.1 Magnetic properties	4
1.1.2 Relaxation	7
1.1.3 Imaging sequences	9
1.1.4 Artefacts	12
1.2 Diffusion Weighted Imaging	13
1.2.1 Diffusion tensor imaging	15
1.2.2 Tractography	15
1.3 Functional Imaging	17
2 Graph Theory and Machine Learning	19
2.1 Graph theory	19
2.2 Machine learning	21
3 Vestibular System	24
3.1 Vestibular system anatomy and physiology	24
3.1.1 Peripheral vestibular system	24
3.1.2 Central vestibular system	26
3.2 Vestibular system dysfunction and tests	27
3.2.1 Vestibular system dysfunction	27
3.2.2 Vestibular tests	28
3.3 Connectome analysis for vestibular system	29

II	Analysis of connectomes for vertigo	35
4	Methods	36
4.1	Data	36
4.2	Structural analysis	37
4.2.1	dMRI preprocessing	37
4.2.2	dMRI registration	39
4.2.3	Structural connectivity	40
4.3	Functional analysis	43
4.3.1	fMRI preprocessing	43
4.3.2	fMRI registration	45
4.3.3	Functional connectivity	46
4.4	Statistical analysis of connectivity measures	46
4.5	Machine learning and biomarker identification	48
4.5.1	Filter feature selection methods	49
4.5.2	Wrapper feature selection method	50
4.5.3	Embedded feature selection method	53
4.6	Statistical cerebellum analysis	54
5	Results	55
5.1	Connectome	55
5.1.1	Structural	55
5.1.2	Functional	56
5.2	Network measures	57
5.2.1	Structural	57
5.2.2	Functional	60
5.3	Machine learning analysis	61
5.3.1	Structural	61
5.3.2	Functional	68
5.4	Statistical cerebellum analysis	77
6	Discussion	79
6.1	Connectomes	79
6.2	Accuracy and stability of the feature selection methods	80
6.3	Selected regions of interest	82
6.4	Cerebellum	95
6.5	Limitations	96
6.6	Future work	96
7	Conclusion	98
	Bibliography	99
	Appendices	108
A	Labels of the atlas	109

B	Visualisation of the atlas	112
C	Non registered regions in the diffusion space	113
D	Non registered regions in the functional space	114
E	Example of Volbrain report	115
F	Structural features selected from correlation	119
G	Structural features selected from mutual information	122
H	Functional features selected from correlation	125
I	Functional features selected from mutual information	127
J	Complete histogram for functional features selected with mutual information method and RF model	129
K	Statistical tests on cerebellum	130
L	Analysis of features selected from machine learning methods	133
L.1	Structural features selected from correlation pre-filter method and RF model	133
L.2	Functional features selected from correlation pre-filter method and RF model	134
L.3	Functional features selected from correlation pre-filter method and SVM model (control VS vertigo + compensated vertigo)	135
L.4	Functional features selected from correlation pre-filter method and SVM model (control + compensated vertigo VS vertigo)	138
L.5	Functional features selected from MI pre-filter method and RF model	139

List of Figures

1.1	Representation of magnetic properties	5
1.2	Spiral-like motion of the magnetisation when a RF-pulse is applied	6
1.3	Effect on the magnetisation $\vec{M}$ of the application of RF-pulses . . .	6
1.4	Behaviour of the T_1 relaxation after application of a 90° RF-pulse .	7
1.5	Behaviour of the T_2 relaxation after application of a 90° RF-pulse .	8
1.6	Consequence on the magnetisation of the application of a 180° RF-pulse	9
1.7	Schematic representation of the TE and TR	10
1.8	Summary of the imaging contrasts (T1, T2 and proton density). . .	11
1.9	Pulse gradient spin echo sequence.	13
1.10	Schematic representation of a water molecule moving in the brain .	14
1.11	Example of a whole brain tractography.	16
1.12	Comparison between probabilistic and deterministic tractography. .	17
1.13	Time course of functional imaging and BOLD signal.	18
2.1	Examples of weighted/unweighted and directed/undirected networks with the corresponding connectivity matrix.	20
2.2	Schematic representation of Support Vector Machine and Random Forest models	23
3.1	Peripheral vestibular system and representation of hair cells in the peripheral vestibular system	25
3.2	Representation of the central vestibular system	27
3.3	Modules selected to belong to the multimodal vestibular cortex, each colour representing a module	32
4.1	Visualisation of the result of the preprocessing step	38
4.2	Visualisation of the result of the registration step	40
4.3	Probabilistic tractogram with and without cleaning of small fibers .	42
4.4	Graphs of residual motion for the functional images of one participant	43
4.5	High pass filter applied on one functional image	44
4.6	Methodology used for the statistical tests of the network measures .	47
4.7	Summary scheme of the methodology.	53
5.1	Visualisation of the difference between deterministic and probabilistic tractograms	56
5.2	Structural and functional connectivity matrices	56

5.3	Accuracy and stability of the structural features selected from the correlation pre-filter method and the SVM model	62
5.4	Accuracy and stability of the structural features selected from the correlation pre-filter method and the RF model	63
5.5	Structural features selected from the correlation pre-filter method and the RF model	64
5.6	Difference of fibers density for the first feature selected by the RF model with the correlation pre-filtering	64
5.7	Visualisation of second and third features selected by the RF model with the correlation pre-filtering	65
5.8	Accuracy and stability of the structural features selected from mutual information pre-filter method and the SVM model	65
5.9	Accuracy and stability of the structural features selected from mutual information pre-filter method and the RF model	66
5.10	Structural features selected from mutual information pre-filter method and the RF model	66
5.11	Visualisation of the difference in fibers density for the two first structural features selected by the RF model with mutual information pre-filtering	67
5.12	Visualisation of the difference in fibers density for the third and fourth structural features selected by the RF model with mutual information pre-filtering	67
5.13	Accuracy and stability of the functional features selected from correlation pre-filter method and the SVM model	68
5.14	Functional features selected from correlation pre-filter method for the SVM model (vertigo who have compensated or not against control)	69
5.15	Visualisation of functional features selected from the correlation pre-filter method and SVM model (vertigo participants who have compensated or not against control participants)	69
5.16	Visualisation of functional features selected with the correlation pre-filter method and SVM model (vertigo participants who have compensated or not against control participants)	70
5.17	Functional features selected from the correlation pre-filter method and the SVM model (vertigo participants against control participants and compensated vertigo participants)	70
5.18	Visualisation of functional features selected with the correlation pre-filter method and SVM model (vertigo against control and compensated vertigo)	71
5.19	Accuracy and stability of the functional features selected from correlation pre-filter method and the RF model	71
5.20	Functional features selected from correlation pre-filter method and the RF model	72
5.21	Visualisation of features selected from correlation pre-filter method and the RF model	72

5.22	Accuracy and stability of the functional features selected from mutual information pre-filter method and the SVM model	73
5.23	Functional features selected from the mutual information pre-filter method and the SVM model (vertigo participants who have compensated or not against control participants)	74
5.24	Visualisation of features selected from the mutual information pre-filter method and the SVM model (vertigo participants who have compensated or not against control participants)	74
5.25	Functional features selected from the mutual information pre-filter method and the SVM model (vertigo participants against control participants and vertigo participants who have compensated)	75
5.26	Visualisation of features selected from the mutual information pre-filter method and the SVM model (vertigo participants against control participants and vertigo participants who have compensated)	75
5.27	Accuracy and stability of the functional features selected from the mutual information pre-filter method and the RF model	76
5.28	Functional features selected from the mutual information pre-filter method and the RF model	76
5.29	Visualisation of features selected from the mutual information pre-filter method and the RF model	77
5.30	Mean and standard deviation of the cortical thickness asymmetry of lobule V and lobule IV	78
5.31	Lobule IV and lobule V of the cerebellum on a T1-weighted image of a patient	78
B.1	Visualisation of the atlas used	112
E.1	Example of a report from Volbrain	115
J.1	Functional features selected from the mutual information pre-filtering method and the RF model	129

List of Tables

3.1	Analysis of connectivity and network measures in the vestibular network.	30
3.2	Analysis of modularity in the vestibular network.	31
3.3	Analysis of the PIC and PIVC regions of the vestibular system. . .	32
3.4	Analysis of the lateralisation of the vestibular system.	33
4.1	Mean age of participants for the three classes (control participants, vertigo participants, vertigo participants who have compensated). .	50
5.1	Statistical values for the betweenness centrality in the regions located in the left hemisphere (structural features).	57
5.2	Statistical values for the betweenness centrality in the regions located in the right hemisphere (structural features).	58
5.3	Statistical values for the degree centrality in the regions located in the left hemisphere (structural features).	58
5.4	Statistical values for the degree centrality in the regions located in the right hemisphere (structural features).	59
5.5	Statistical values for the betweenness centrality in the regions located in the left hemisphere (functional features).	60
5.6	Statistical values for the betweenness centrality in the regions located in the right hemisphere (functional features).	60
5.7	Statistical values for the degree centrality in the regions located in the left hemisphere (functional features).	61
5.8	Statistical values for the degree centrality in the regions located in the right hemisphere (functional features).	61
5.9	Statistical values for the cortical thickness asymmetry of all lobules of the cerebellum.	77
5.10	Statistical values for the cortical thickness asymmetry of lobules V and IV between Control - Vertigo participants (C-V), Control - Compensated Vertigo participants (C-VC) and Vertigo - Compensated Vertigo participants (V-VC).	78
6.1	Balanced Non-balanced accuracies of the SVM and RF models for correlation and mutual information (MI) methods with the percentage of features selected chosen to analyse data. The best performances for structural and functional data are in green.	81

6.2	Analysis of structural features selected with the RF model and MI method.	84
6.3	Analysis of the third and fourth structural features selected with the RF model and MI method.	85
6.4	Analysis of the first and third functional features selected from SVM model and correlation pre-filtering method to discriminate control from vertigo participants who have compensated or not.	87
6.5	Analysis of the fourth and fifth functional features selected from SVM model and correlation pre-filtering method to discriminate control from vertigo participants who have compensated or not.	88
6.6	Analysis of functional features selected from SVM and correlation to discriminate vertigo participants from control and compensated vertigo participants.	90
6.7	Analysis of functional features selected from SVM and correlation to discriminate vertigo participants from control and compensated vertigo participants.	91
A.1	Region's names of the left hemisphere with their corresponding label.	109
A.2	Region's names of the left hemisphere with their corresponding label.	110
A.3	Region's names of the right hemisphere with their corresponding label.	110
A.4	Region's names of the right hemisphere with their corresponding label.	111
A.5	Region's names of the vermis (in the cerebellum) with their corresponding label.	111
C.1	Region's names and labels of non registered regions after the registration from the MNI space to the native space of structural images	113
D.1	Region's names and labels of non registered regions after the registration from the MNI space to the native space of functional images	114
F.1	Structural features kept after the correlation pre-filtering method for intrahemispheric connections of the left hemisphere.	119
F.2	Structural features kept after the correlation pre-filtering method for intrahemispheric connections of the right hemisphere.	120
F.3	Structural features kept after the correlation pre-filtering method for interhemispheric connections.	121
G.1	Structural features kept after the mutual information pre-filtering method for intrahemispheric connections of the left hemisphere.	122
G.2	Structural features kept after the mutual information pre-filtering method for intrahemispheric connections of the right hemisphere.	123
G.3	Structural features kept after the mutual information pre-filtering method for interhemispheric connections.	124

H.1	Functional features kept after the correlation pre-filtering method for intrahemispheric connections in the left hemisphere (left) and in the right hemisphere (right).	125
H.2	Functional features kept after the correlation pre-filtering method for interhemispheric connections.	126
I.1	Functional features kept after the mutual information pre-filtering method for intrahemispheric connections in the left hemisphere (left) and in the right hemisphere (right).	127
I.2	Functional features kept after the mutual information pre-filtering method for interhemispheric connections.	128
K.1	Statistical values for the volume asymmetry of all lobules of the cerebellum.	130
K.2	Statistical values for the grey matter asymmetry of all lobules of the cerebellum.	130
K.3	Statistical values for the grey matter volume (in cm ³) of all lobules of the cerebellum.	131
K.4	Statistical values for the mean cortical thickness of all lobules of the cerebellum.	131
K.5	Statistical values for the total volume (in cm ³) of all lobules of the cerebellum.	132
L.1	Analysis of structural features selected from RF and correlation methods.	133
L.2	Analysis of functional features selected with the RF model after the correlation prefiltering.	134
L.3	Analysis of functional features selected with SVM and MI methods to discriminate control from vertigo participants who have compensated or not.	135
L.4	Analysis of functional features selected with the SVM model after the mutual information method to discriminate control from vertigo participants who have compensated or not.	137
L.5	Analysis of functional features selected from RF and MI methods.	139

List of Abbreviations

ADC Apparent Diffusion Coefficient.

BOLD Blood Oxygen Level Dependent.

dMRI diffusion MRI.

DTI Diffusion Tensor Imaging.

DWI Diffusion Weighted Imaging.

EPI Echo Planar Imaging.

FA Fractional Anisotropy .

FID Free Induction Decay.

fMRI functional MRI.

GFA Generalised Fractional Anisotropy.

MI Mutual Information.

MNI Montreal Neurological Institute.

MR Magnetic Resonance.

MRI Magnetic Resonance Imaging.

ODF Orientation Diffusion Function.

PCVD Post-Concussive Vestibular Dysfunction.

PIC Posterior Insular Cortex.

PIVC Parieto-Insular Vestibular Cortex.

RF Random Forest.

ROI Region Of Interest.

SCC SemiCircular Canal.

SVM Support Vector Machine.

TE Echo Time.

TR Repetition Time.

VEMP Vestibular Evoked Myogenic Potential.

VN Vestibular Nuclei.

VNG Videonystagmography.

Introduction

Advances in magnetic resonance imaging (MRI) have opened new doors to the exploration of the brain and have motivated many scientists to investigate neurological disorders, some types of vestibular dysfunction being among them. Vestibular system dysfunction can lead to vertigo, described as an abnormal sensation of motion. Only recognised in modern times, the vestibular system is a highly interconnected system playing an important role in postural control, reflexes and eyes movements. Although many studies have investigated brain areas involved in the vestibular system, its structural and functional connectivity patterns are still a matter of debate. Furthermore, only very few studies have been dedicated to the search of brain areas involved in disorders causing vertigo.

Contrary to other sensory modalities, vestibular dysfunction is frequently transient. Indeed, thanks to the plasticity of the brain, vertigo symptoms rarely or partially persist in patients suffering from lesions or disorders of the vestibular system. It may even happen that patients completely compensate for vertigo, meaning they do not suffer from vestibular dysfunction anymore. It is therefore interesting to look for the neurological cause of vertigo, i.e., which areas of the brain are involved, to try to better understand these vertigo disorders.

The goal of this thesis is therefore to identify functional or structural connections between regions of the brain allowing to discriminate people suffering from vertigo from control participants.

To achieve this goal, graph theory is used in combination with machine learning. Indeed, developments in the quantitative analysis of complex networks, based on graph theory, have expanded to various scientific fields, including neuroscience. Interest in the study of the organisation of brain networks has increased and can be divided into two parts: the study of the structural network, representing the neural interaction between brain areas; and the study of the functional network, giving insight into the physiological basis of information processing in the brain. Studies of brain connectivity have revealed fundamental principles of brain architecture and function, and hold great promise for understanding certain brain disorders. On the other hand, advances in technology have allowed neuroscientists to collect more data, enabling a larger use of machine learning for brain imaging to better understand neurological diseases and dysfunctions.

The thesis is composed of two main parts: the first one gathers all the theoretical background needed to understand the methods used and the results obtained in the second part. The theoretical background covers three main topics. First, theory about MRI imaging and its application in diffusion and functional imaging is addressed. Then, two concepts used in the methodology, being graph theory and machine learning, are developed. Finally a review of the vestibular system anatomy and physiology is performed as well as a review of the regions identified in literature as being part of the vestibular system.

The second part is about the application of graph theory and machine learning on MRI images acquired on participants suffering from vertigo or not. The first part describes the methodology used, the second one enumerates the results obtained and the last one is a discussion of those results. In the methods part, the entire pipeline used to find the different regions discriminating between control and vertigo participants is described. This pipeline includes preprocessing of the diffusion and functional images, the method used for constructing tractograms based on the structural images and the method used to build the connectomes based on the structural and functional data. A description of the different machine learning algorithms used is also given. Then, the results section principally gathers a summary of the performances of machine learning algorithms, the regions selected by those algorithms and the visualisation of the regions concerned in brain images. The discussion addresses the significance of the results by comparing the different regions obtained between them and with what is known from the literature. Finally, limitations of the work performed related to some choices made in the methodology and to the available data are outlined. And lastly, to give perspectives, some ideas for future work are suggested.

Part I

Theoretical Background

Chapter 1

Magnetic Resonance Imaging

In this first chapter, necessary notions needed to understand the principles of Magnetic Resonance Imaging (MRI) are covered. Diffusion weighted imaging and functional imaging are also developed.

1.1 Principles of Magnetic Resonance Imaging

As its name indicates, MRI is an imaging technique which relies on magnetism. This non invasive technique allows to obtain different types of images (anatomical, functional, ...) in a safe manner as it does not imply the use of any radiation.

1.1.1 Magnetic properties

The nucleus of atoms is made up of a certain number of protons and neutrons. These particles describe a complex movement which includes a rotation of the particle on itself. This rotation is called spin and is represented in Figure 1.1 a. Only the particles having a nucleus with either an uneven atomic mass or uneven atomic number are characterised by a nonzero spin number. For MRI, the nucleus of main interest is the one of hydrogen because it represents two thirds of the total amount of atoms in the organism. This hydrogen atom is composed of one proton, having thus a nonzero spin number, and it also has a high intrinsic magnetic moment. This intrinsic magnetic moment is represented by the vector μ in Figure 1.1 a and is parallel to the axis of rotation. Due to this magnetic moment, the proton can be assimilated to a small magnet that orients itself in a certain direction when placed in a magnetic field $\vec{B}_0$. When no external magnetic field is applied around the body, these magnetic moments are randomly distributed and their sum is thus null. But when a magnetic field $\vec{B}_0$ is applied, protons align themselves with it and are divided in two different populations [1].

The first population is parallel to $\vec{B}_0$ (up) and the second one is antiparallel to $\vec{B}_0$ (down). However, the amount of protons up is higher than the amount of protons down (for an external field of 0.5 Tesla applied on a sample of 2 million protons at 37°, approximately 4 protons are in excess in parallel position). Therefore, this gives a resultant which is parallel to the magnetic field $\vec{B}_0$ [2, 3].

Actually, due to a torque induced by the magnetic moment in the presence of a magnetic field, protons are not perfectly aligned with the magnetic field but rotate around $\vec{B}_0$ with a given angle and a certain angular frequency ω_0 , also called the Larmor frequency, defined as

$$\omega_0 = \gamma B_0, \quad (1.1)$$

where γ is the gyromagnetic ratio specific to the nucleus. This motion is called precession.

In a body, there is a huge amount of protons all precessing around the magnetic field but with randomly distributed phases. Therefore, the resulting magnetisation $\vec{M}$ is only longitudinal (along $\vec{B}_0$) and not transverse, as illustrated in Figure 1.1 b.

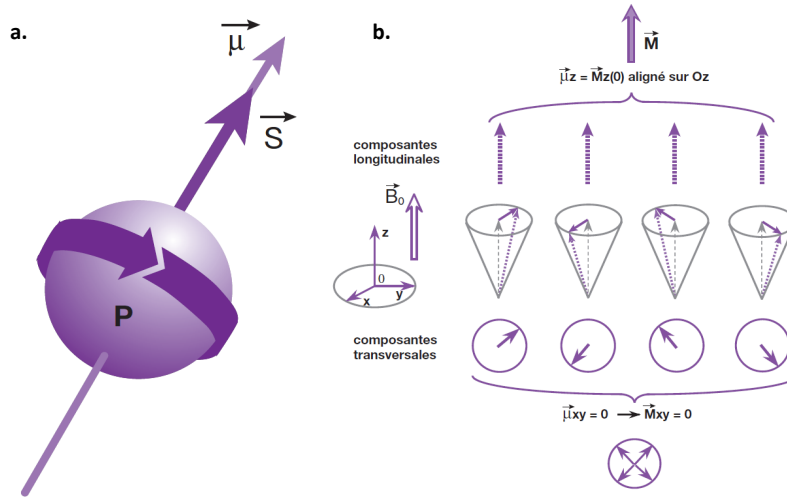


Figure 1.1: (a) Representation of the spin S and the magnetic moment μ , (b) Resulting magnetisation $\vec{M}$ when a magnetic field $\vec{B}_0$ is applied. From [1].

However, the magnetic field used for an MRI often ranges from 1.5 to 9 Tesla. Compared to this powerful magnetisation, the magnetic longitudinal resultant $\vec{M}$ of protons in the body is too small to be measured. A magnetisation not oriented along $\vec{B}_0$ should therefore be created [2].

This is done by using a Radio Frequency-pulse (RF-pulse) $\vec{B}_1$ which is applied along the xy-plane (i.e. perpendicular to $\vec{B}_0$) and for which the angular frequency ω_{RF} must be the same as ω_0 to allow the transfer of energy to the system in the equilibrium state. When the condition ($\omega_{RF} = \omega_0$) is satisfied, the two systems are said to be in resonance. The magnetisation vector $\vec{M}$, while continuing to precess around $\vec{B}_0$ at the angular frequency of $\omega_0 = \gamma B_0$, will therefore also start to precess around $\vec{B}_1$ at the angular frequency of $\omega_1 = \gamma B_1$. As a result, $\vec{M}$ goes down in a spiral-like motion as shown in Figure 1.2 [2].

When studying this spiral-like motion, it is customary to introduce a rotating frame in which spins rotating around $\vec{B}_0$ appear stationary. The motion then becomes a

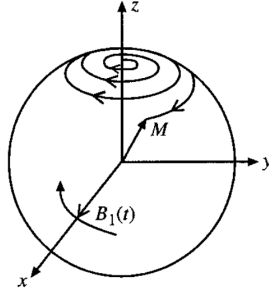


Figure 1.2: Spiral-like motion of $\vec{M}$ when a RF-pulse is applied. From [3].

simple tilting movement of $\vec{M}$ around $\vec{B}_1$ (Ox) at the angular frequency ω_1 . When submitted to a RF-pulse of finite duration T_{RF} , the magnetisation is tilted over an angle

$$\alpha = \gamma \int_{T_{RF}} B_1(t) dt,$$

which becomes $\alpha = \gamma B_1 T_{RF}$ as B_1 is constant [3]. A magnetisation tilted over an angle α is said to have been submitted to a α° RF-pulse.

In MRI, there are 2 common RF-pulses: 90° and 180° RF-pulses, which can be seen in Figure 1.3 a and c, respectively. It can be observed in Figure 1.3 c that a 180° RF-pulse reverses the magnetisation. On the other hand, it can be seen in Figure 1.3 a and b that when applying a 90° RF-pulse, the longitudinal component of $\vec{M}$ ($\vec{M}_z$) decreases while its transverse component ($\vec{M}_{xy}$) increases. The magnetisation $\vec{M}$ can therefore be decomposed in two components [3]:

$$\vec{M} = M_z \vec{1}_z + M_{xy} \vec{1}_{xy}.$$

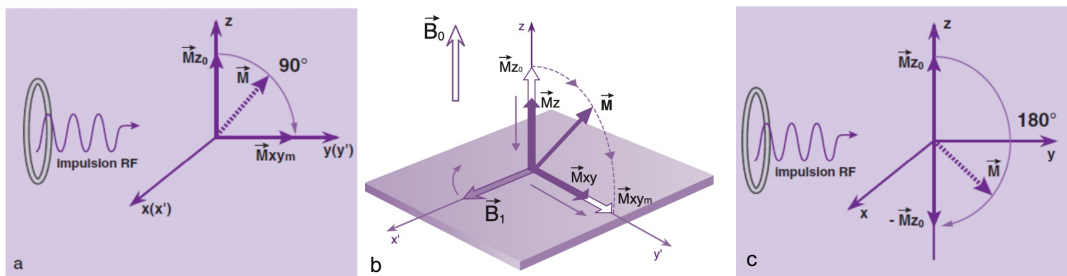


Figure 1.3: Effect on the magnetisation $\vec{M}$ of the application of (a) a 90° RF-pulse, (c) a 180° RF-pulse. (b) Effect in the rotating frame on the magnetisation $\vec{M}$ of the application of an external field $\vec{B}_0$ and a 90° RF-pulse $\vec{B}_1$. The magnetisation $\vec{M}$ is decomposed into a longitudinal component $\vec{M}_z$ and a lateral component $\vec{M}_{xy}$. From [1].

1.1.2 Relaxation

When a spin system is excited in such a way that the magnetisation $\vec{M}$ is decomposed in longitudinal and transverse components, M_z evolves back to M_0 as soon as the 90° RF-pulse is switched off, while M_{xy} disappears since a system always tends to go back to its equilibrium state. This return to the equilibrium state, also called relaxation, is recorded and gives the signal of the MRI.

The relaxation time is the time needed to return to the equilibrium state. Three main types of relaxations can be identified [2].

T_1 relaxation

The T_1 relaxation represents the exchange of energy between the excited spin and the environment. It therefore represents the transfer of protons between the parallel and antiparallel states. In other words, it describes the time needed to recover 63% of the equilibrium magnetisation, i.e., how fast the M_z component of magnetisation reappears, as illustrated in Figure 1.4.

In practice, each tissue has its own T_1 relaxation time. Spins that are very mobile and can move freely have a long T_1 relaxation time while for example large fat molecules are less mobile and have a very short T_1 relaxation time. It also depends on the type of external field. A tissue submitted to an external field $\vec{B}_0$ of 1.5T and 3T will have different T_1 relaxation times. Actually when the collision frequency of the molecules is close to the Larmor frequency ω_0 , which depends on the external field $\vec{B}_0$ applied (Eq. 1.1), the energy transfer between the spin and the environment is facilitated and T_1 relaxation time is thus short. On the contrary, the greater the difference between these two frequencies, the longer the T_1 relaxation time [1].

To give an idea about the order of magnitude, the T_1 relaxation times of white matter, grey matter and cerebral spinal fluid are approximately 600 milliseconds, 1 second and 2 seconds, respectively [2].

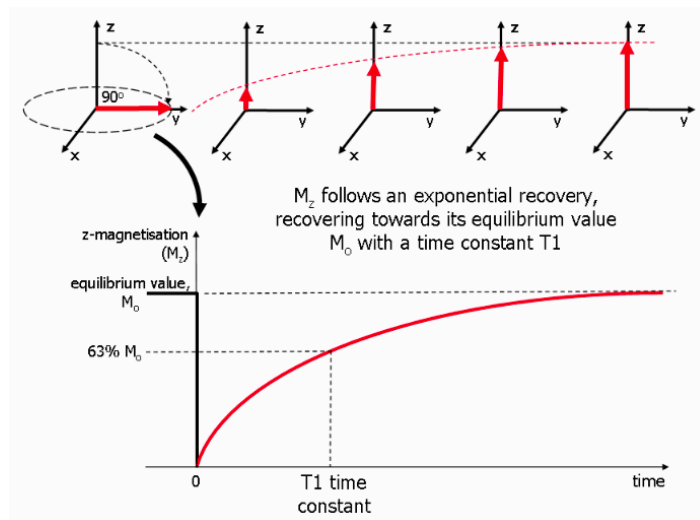


Figure 1.4: Behaviour of the T_1 relaxation after application of a 90° RF-pulse. From [4].

T_2 relaxation

The second relaxation type is the T_2 relaxation, also called spin-spin relaxation. This relaxation is caused by field fluctuations and static inhomogeneities at microscopic level. Protons evolve in different molecular environments, where small local magnetic fields will be superimposed (added or subtracted) to the main magnetic field B_0 . These phenomena are field inhomogeneities of molecular origin [1]. Due to these fluctuations, spins lose their phase coherence as represented in the upper part of Figure 1.5. Parts of spins precess faster and slower than the Larmor frequency. Spins fan out and the resultant magnetisation decreases. The disappearance of M_{xy} is thus due to the dephasing of spins. The T_2 relaxation time is thus defined as the time needed in such a way that 63% of the initial value of M_{xy} is lost, as shown in the lower part of Figure 1.5.

The T_2 relaxation time also depends on the type of tissue and is always a little bit shorter than the T_1 relaxation time. The order of magnitude for the T_2 relaxation time of biological tissues is about 50 ms to 100 ms [2].

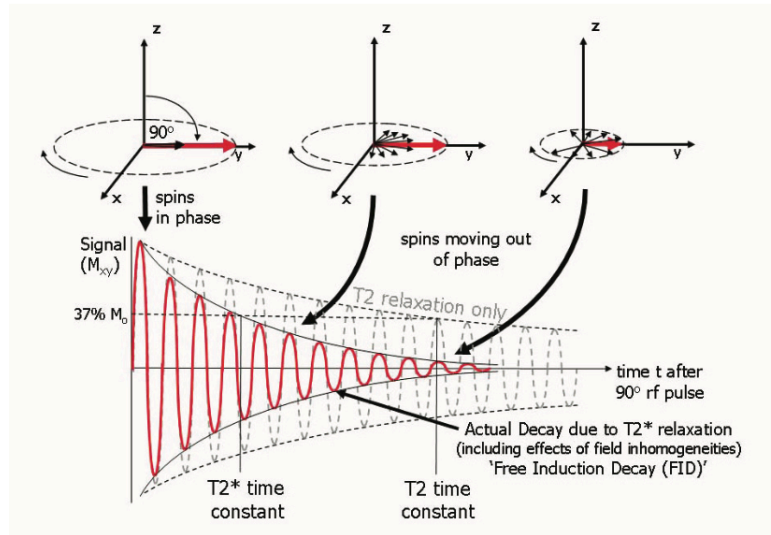


Figure 1.5: Behaviour of the T_2 relaxation after application of a 90° RF-pulse. From [4].

T_2^* relaxation

The last type of relaxation, T_2^* , represents the dephasing of spins due to spatial nonuniformities in the external field $\vec{B}_0$. The effective relaxation time T_2^* is defined as [3]:

$$\frac{1}{T_2^*} = \frac{1}{T_2} + \gamma \Delta B_0.$$

As can be seen in the two previous graphs, the behaviour of these three relaxation times follows an exponential distribution. The exponential equation of the transverse magnetisation M_{xy} when a 90° RF-pulse is applied to the system is the following:

$$M_{xy}(t) = M_0 \exp\left(\frac{-1}{T_2^*}\right) \exp(i(-\omega_0 t + \phi)).$$

This oscillating behaviour (see Figure 1.5) induces a signal known as the Free Induction Decay (FID). This signal is measured thanks to a receiver antenna placed in the xy-plane (perpendicular to the axis of the magnetic field $\vec{B}_0$ applied) and is the one allowing to obtain the final image of the MRI [3].

1.1.3 Imaging sequences

Spin Echo

As explained just before, the measured signal (FID) takes into account the fluctuations at microscopic level (T_2) but also the homogeneities linked to the external field B_0 (T_2^*). However, the signal of interest is the T_2 signal, linked to the interaction of spins together and not T_2^* linked to the quality of the imaging system (i.e. the homogeneity of B_0). The imaging sequence Spin Echo (SE) allows to get rid of these T_2^* effects and only measures the signal of interest. This sequence relies on the application of a 90° RF-pulse followed by a 180° RF-pulse which aims to cancel the dephasing induced by the B_0 field by rephasing the spins [1]. This can be observed in Figure 1.6. Part (a) represents the situation when the spins are excited by a 90° RF-pulse. The longitudinal magnetisation disappears while the transverse magnetisation becomes maximal. It can be seen on part (b) that just after the 90° RF-pulse, spins start to dephase due to the microscopic homogeneities and the homogeneities due to B_0 . Some spins (like number 4 or 5) are faster while other ones are slower (number 1). Just after the application of the 180° RF-pulse (part c), the direction of rotation stays the same but the slowest spins are now in front and the fastest behind. This 180° RF-pulse is applied at a time $t = \frac{TE}{2}$ where TE is the echo time and defined in the following paragraph. After a time TE (part d), spins are again in phase. The transverse magnetisation is again maximal, generating a spin echo signal that is measured [3].

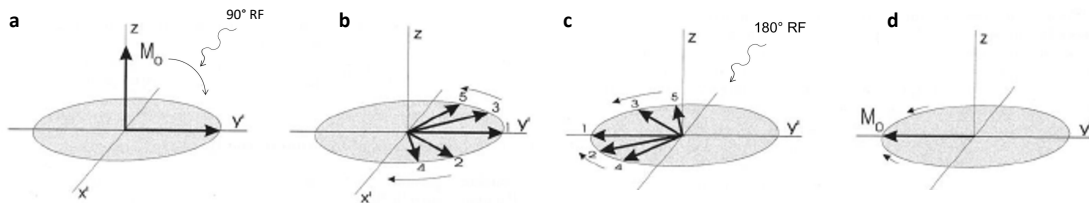


Figure 1.6: Consequence of the application of a 180° RF-pulse. From [3].

Actually, the transverse magnetisation after the 180° RF-pulse is a bit smaller than the original one because the T_2 -relaxation is not corrected by the RF-pulse. After the 180° RF-pulse, the spins are not perfectly rephased. The signal observed presents a T_2 decay. Such a sequence allows to obtain one line of the Fourier plane which is used to get the final image. It is thus repeated in order to obtain the other lines, and thus the entire image [1].

There are two main parameters in such a sequence. The first one is the echo time (TE) which corresponds to the time between the first 90° RF-pulse and the echo. The second parameter is the repetition time (TR) which is the time separating two 90° RF-pulses [1].

Three different contrasts can be obtained by playing with the parameters of the sequence. The TE parameter determines the time at which the signal is measured. The longer the TE, the lower the measurable transverse magnetisation M_{xy} . It thus determines the remaining signal level for the measurement, i.e., the "share" of T_2 in the signal, as can be seen in Figure 1.7. On the other hand, the TR represents the time left to the longitudinal magnetisation M_z to reappear and thus determines the amount of signal available. The longer TR is, the higher the longitudinal magnetisation M_z (i.e., the M_{xy} available for signal measurement after the following RF-pulse). It thus determines the initial level of the signal available for measurement, i.e., the "share" of T_1 in the signal, as can be observed in Figure 1.7 [2].

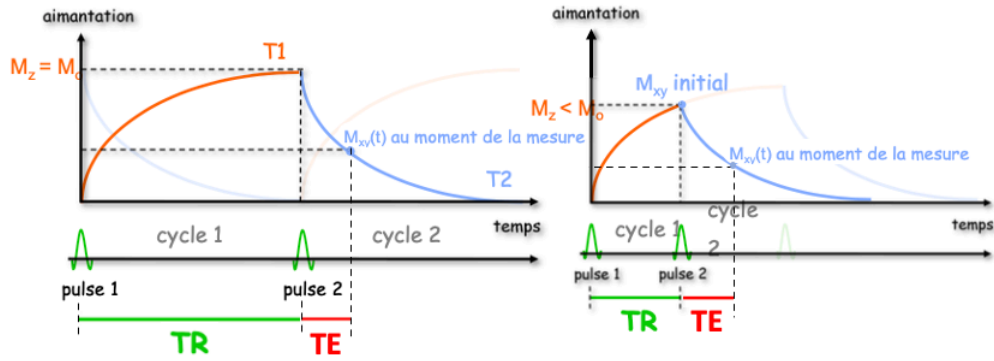


Figure 1.7: Schematic representation of the TE and TR. Adapted from [2]

T_1 contrast - During each TR interval, the longitudinal magnetisation reappears at different rates depending on the T_1 of the tissue. Therefore, if the TR is short, the magnetisation of a tissue with a small T_1 is higher than for a tissue with a longer T_1 , as can be seen in the upper right of Figure 1.8. The measured signal is therefore higher for the first tissue and a contrast can be seen on the image. This contrast is called T_1 contrast because the tissues are distinguishable based on their T_1 . The higher the TR the smaller the difference between different tissues due to T_1 . In order to have a T_1 contrast, it is therefore needed to select a small TR (400 to 600 ms) and a small TE (15 ms) to minimise the T_2 contrast (see just below) [2].

T_2 contrast - During each TE interval, the transverse magnetisation decreases due to the dephasing of spins. This decrease depends on the T_2 relaxation time. If the TE is short (lower than 20-30ms), the difference between the relaxation velocities of different tissues has not enough time to appear as can be observed in the lower right of Figure 1.8. On the contrary, if the TE is long (higher than 80-100 ms), the signal of a tissue with a longer T_2 will be higher because the dephasing is slower, resulting in a higher transverse magnetisation. In order to have a T_2 contrast, it is therefore needed to select a long TE (120 ms) and a long TR (2000 ms) to minimise the T_1 contrast [2].

Proton density contrast - To get a proton density contrast, one needs to choose a long TR (2000 ms) to minimise the T_1 contrast and a short TE (15 to 20 ms) to minimise the T_2 contrast. In that case, the longitudinal magnetisation reappears for all the tissues and is equal to the initial equilibrium value. If there is no difference in proton density between the tissues, no difference is observed in the signal. However, differences in proton density lead to differences in transverse magnetisation and therefore, to differences in signal.

An MRI image is decomposed in multiple voxels of a few millimeters where each voxel corresponds to a piece of a particular tissue. Therefore, depending on the TR and TE parameters chosen for the acquisition of the images, each voxel has a different signal intensity: a small intensity appears black while a high intensity appears white. This is what allows to obtain an image with different contrasts delimiting different tissues of the brain. An example of T_1 , T_2 and proton density contrasts of the brain can be seen here below in Figure 1.8 Left.

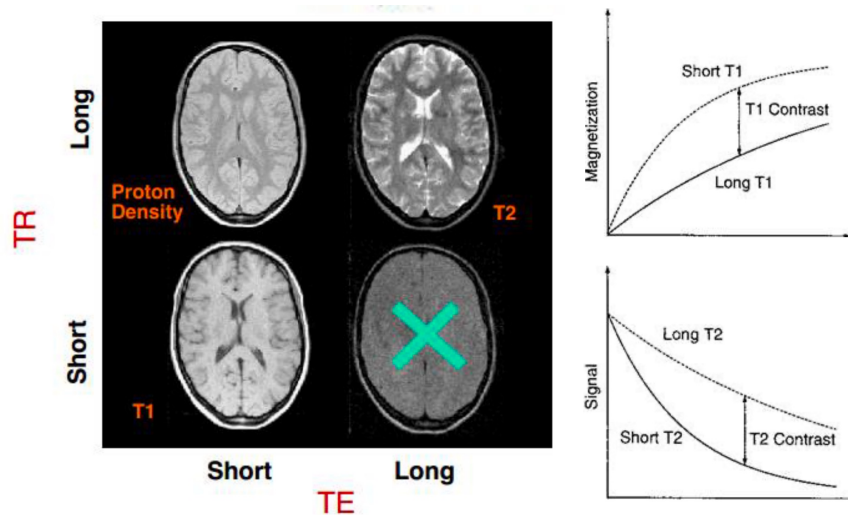


Figure 1.8: **Left:** T_1 , T_2 and proton density contrasts of brain images. From [2]. **Upper right:** Effect on the signal of different T_1 relaxation times. From [2]. **Lower right:** Effect on signal of different T_2 relaxation times. From [2].

1.1.4 Artefacts

MRI images present some artefacts as any other kind of imaging technique. It is important to know it in order to eliminate or minimise these artefacts.

Susceptibility artefacts - Substances interact differently with a magnetic field depending on their susceptibility. Diamagnetic substances have a small negative susceptibility and do not pose any problem. Paramagnetic substances have a small positive susceptibility and are weakly attracted by the external field. Ferromagnetic substances have a large positive susceptibility and are strongly attracted by the external magnetic field. Susceptibility artefacts appear at interfaces between substances of different susceptibility. The difference in susceptibility causes the distortion of the local field. This leads to a dephasing of spins (resulting in a loss of signal) and also to mismapping artefacts. Ferromagnetic substances like iron, cobalt or nickel used for some metallic implants lead to substantial artefacts [3].

Signal processing artefact - A first type of signal processing artefact is aliasing. Aliasing is an artefact characterised by data lying outside of the field of view wrapping back into the image. This can be solved by doing oversampling, for example, doubling the field of view and the number of pixels.

A second type of signal processing artefact is the Gibbs ringing. This is the apparition of the spurious ringing around the sharp transition in an image because these sharp edges contain very high spatial frequencies. A solution to reduce Gibbs ringing is to filter the data. Indeed, this allows to reduce the spatial frequency [3].

Chemical shift - Tissues of different composition (for example fat and water) in the human body have different resonance frequency. This can lead to misregistration because frequency encoding is used for spatial location in MRI. This means that the difference in frequency codes for one given position and not for a difference in tissues. This leads to bright and black band in the image. To reduce this artefact, it is possible to increase the magnetic gradient or to apply fat saturation pulses which are 90° RF-pulses with the resonance frequency of the fat. This allows to create a transverse magnetisation due to the fat spins and then, to apply a strong gradient to destroy transverse magnetisation of fat [3].

Motion - This kind of artefact is the most encountered one in MRI due to motion of the patient during the sequence that can last several minutes. Due to this motion, blurring can be observed along the read out direction and ghosting along the phase encoding direction. Ghosts are partial copies of the imaged object. If the motion is periodic, gating can be used to reduce the motion artefact [3].

RF and gradient artefacts & Eddy currents - Eddy currents, or Foucault currents, are caused by rapid change of gradient magnetic fields. This induces electrical current in surrounding conductors. These small electric currents distort the gradient profile by generating their own magnetic fields and cause artefacts in the image by distortion of the external magnetic field. These artefacts are important to correct especially for diffusion imaging [3, 5].

1.2 Diffusion Weighted Imaging

Diffusion imaging is an application of MRI which relies on Brownian motions of water molecules in tissues. The more water is free in a medium, the easier the motion of molecules and the greater the molecular agitation which leads to higher diffusion [3].

The imaging sequence used to see the microscopic molecular motion is a spin echo (SE) sequence with supplementary gradients applied on both sides of the 180° RF-pulse. This sequence is called a pulse gradient spin echo sequence and can be seen in Figure 1.9. To accelerate the scan time, an echo planar spin echo (EPI-SE) can be used at the cost of a lower spatial resolution.

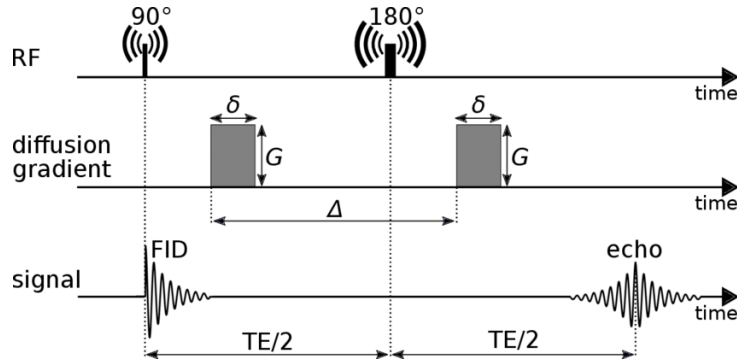


Figure 1.9: Pulse gradient spin echo sequence. From [6].

For immobile protons, the dephasing caused by the first gradient is perfectly compensated by the second gradient. The signal for immobile water molecules is therefore not attenuated. On the contrary, mobile protons also dephase due to the application of the first gradient but this is not compensated by the second one. The signal appears attenuated. The attenuation of the signal is more important for rapid molecular movements. It can be concluded that diffusion images show a hyper-signal in the regions where the diffusion is limited, such as in the grey matter, so that the signal appears light while the signal is weaker in the regions where the diffusion is high, such as in the cerebrospinal fluid, so that the signal appears darker [3]. This can be seen in Figure 1.10.

The effect of diffusion depends on the gradients used in the sequence. This effect is characterised by a gradient factor called the b-factor. The b-factor is used to describe the sensitivity to diffusion. The higher the b-factor, the greater the contrast for regions of low diffusivity. The b-factor can be expressed by the following equation:

$$b = (\gamma G \delta)^2 \left(\Delta - \frac{\delta}{3} \right) \quad [s/mm^2],$$

where γ is the gyromagnetic ratio, G is the gradient strength, δ is the gradient duration and Δ is the time period between the two gradients. Diffusion imaging can therefore be performed by adjusting the gradients [3].

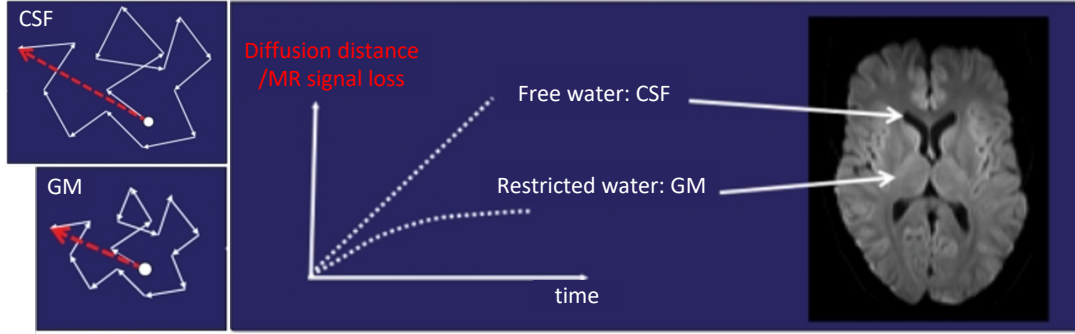


Figure 1.10: Schematic representation of a water molecule moving in the brain. The distance travelled in the grey matter (GM) is smaller than the one in cerebrospinal fluid (CSF). The graph shows that the signal loss is higher for CSF than GM. Figure adapted from [7].

When diffusion imaging is performed, one usually executes a series of T2-weighted EPI slices with a gradient factor $b = 0$ (no diffusion gradients), followed by a series of slices with a b-factor of typically 500 to 3000 $\frac{s}{mm^2}$. This allows to obtain the signal attenuation by comparing both images. The resulting signal intensity is mapped as a three-dimensional image (Diffusion Weighted Imaging (DWI) map) [7]. By increasing the b-factor, the diffusion weighting is increased, but at the expense of the signal to noise ratio which decreases. Using this DWI map, the apparent diffusion coefficient (ADC) can be calculated. This is a measure of the magnitude of diffusion within tissues [8].

The probability of molecular diffusion is equal in all directions of space for the grey matter or the cerebrospinal fluid. This diffusion is isotropic and can be represented by a sphere. However, it is not the case in the white matter. In this tissue, diffusion is facilitated along the axonal fibers and reduced perpendicularly to these fibers. The diffusion is anisotropic and can be graphically represented as an ellipsoid.

The directionality of the diffusion is obtained thanks to the application of the diffusion encoding gradients along different axes. If the principal direction of a white matter tract matches the direction of the diffusion encoding gradient, the signal will be attenuated because there is a lot of diffusion following this direction. On the contrary, if the direction of the white matter tract is perpendicular to the direction of the diffusion gradients, the signal remains intact or is less attenuated. To obtain the final image, a diffusion gradient is applied along the three principal orthogonal directions and an average is made [7].

1.2.1 Diffusion tensor imaging

Diffusion tensor imaging (DTI) allows to analyse the three-dimensional shape of the diffusion. A diffusion tensor is a three-dimensional structure composed of three diffusivities (which are the eigenvalues) associated with three mutually perpendicular directions (the eigenvectors). This tensor is represented by a 3 by 3 symmetric matrix.

$$\begin{pmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{xy} & D_{yy} & D_{yz} \\ D_{xz} & D_{yz} & D_{zz} \end{pmatrix}$$

This tensor is obtained by repeating a diffusion-weighted sequence along a minimum of six different directions corresponding to the six coefficients of the matrix.

Based on the DWI and ADC values, the fractional anisotropy (FA) map can be calculated. FA is the most common scalar metric used in diffusion imaging and takes values between 0 and 1. It indicates the presence or absence of a preferential direction of diffusion. If the FA has a value of 1, the diffusion is maximally anisotropic as for a long cylinder of minimal diameter. In this area, there is therefore a dense packing of myelinated and coherently oriented fibers. These axons act as strong constraints on diffusion. If the FA is equal to 0, the diffusion is maximally isotropic like for a perfect sphere. Areas with high degree of anisotropy are bright while areas with low degree of anisotropy are dark [7, 9].

1.2.2 Tractography

Magnetic resonance diffusion tractography is a method which relies on diffusion tensor data and aims to identify white matter pathways in the brain. It is based on information about the direction of diffusion in the brain. A model, such as the Constant Solid Angle (CSA) Orientation Distribution Function (ODF) model, can be used in order to get these directions from the diffusion dataset. This model estimates the Orientation Distribution Function in each voxel. This ODF represents the distribution of water diffusion as a function of the direction. The peaks of this function represent estimates of the orientation of the tracts [10]. Once the orientation of fibers is known in every voxel, the directions are joined up to reconstruct an approximation of white matter tracts. The simplest form of tractography consists of starting at a seed point, defined in a particular voxel of interest, and generating a track along the preferred direction, which is continuously locally updated, from one voxel to another. The track is stopped as soon as it reaches a voxel which it is unlikely to belong to.

To reconstruct these pathways, prior anatomical knowledge on the trajectory of pathways is indispensable. It is indeed needed to restrict the tracking to areas where the information about directionality is accurate. This can be done by using the areas for which the orientation function returns a significant restricted diffusion. In practice, this is done by thresholding the generalised fractional anisotropy (GFA). To generate anatomically plausible tracking, several other criteria are used such as a curvature threshold, to avoid sharp bending of a tract, and a fractional anisotropy (FA) threshold, to begin tracking in white matter regions where FA is

high. Additional masking can also be applied to limit tracking to specific regions [9, 11].

An example of tractogram of the brain is shown in Figure 5.1.1. On tractograms, principal tracts such as the corpus callosum, which connects the two hemispheres, the cingulum (in green, C-shaped) or the corticospinal tract (in blue in the central part of the image) are observed.

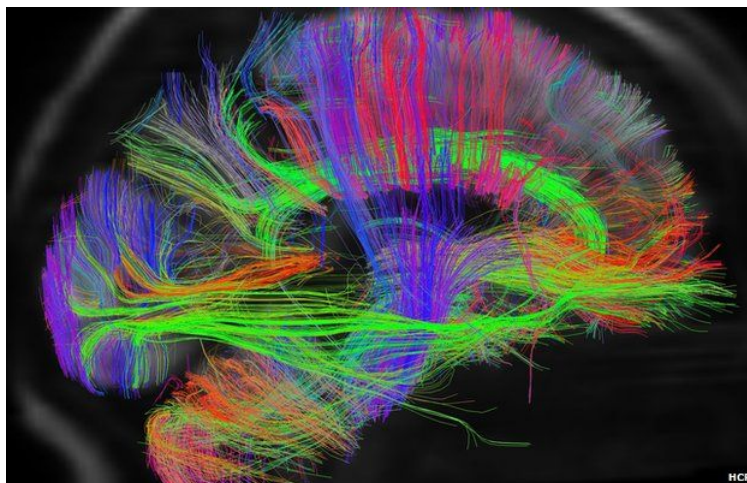


Figure 1.11: Example of a whole brain tractography. From [12].

There are multiple tractography algorithms but the two most known families are the deterministic and probabilistic algorithms.

Deterministic algorithms, on the one hand, only rely on discrete locally optimum decisions to build the tractogram, making it simple to interpret but vulnerable to local noise of the data. For deterministic algorithms, the orientation of the tracts in each voxel can be chosen based on the local maximum of the ODF calculated in each voxel [9, 11]. The model used for probabilistic algorithms includes uncertainty. This means that the distribution of tracking directions at each point is represented by a probability density function (PDF) [13]. Tracks are thus propagated even in regions where it does not seem straightforward and not necessarily in the direction which has the maximum probability.

The difference between both algorithms is illustrated with a 2D example in Figure 1.12. Each arrow represents the primary direction of the pixel concerned, which can be derived, for example, from the direction of the maxima of the ODF. Let's construct a tractogram starting at point A and ending at point Z. With the deterministic approach, only one track is constructed by following the principal direction from voxel to voxel. As seen in Figure 1.12 (i), the deterministic track (in yellow) is along the diagonal. On the other hand, with the probabilistic approach, all possible tracks are assigned to a weight. To move from one voxel to another, many possible directions exist with different probabilities (represented by multiple

arrows of different shades of red in Figure 1.12 (ii)). Therefore, even though the most probable track is still along the diagonal, as with the deterministic approach, other tracks are also possible but with a lower probability, as can be seen in Figure 1.12 (ii) [9].

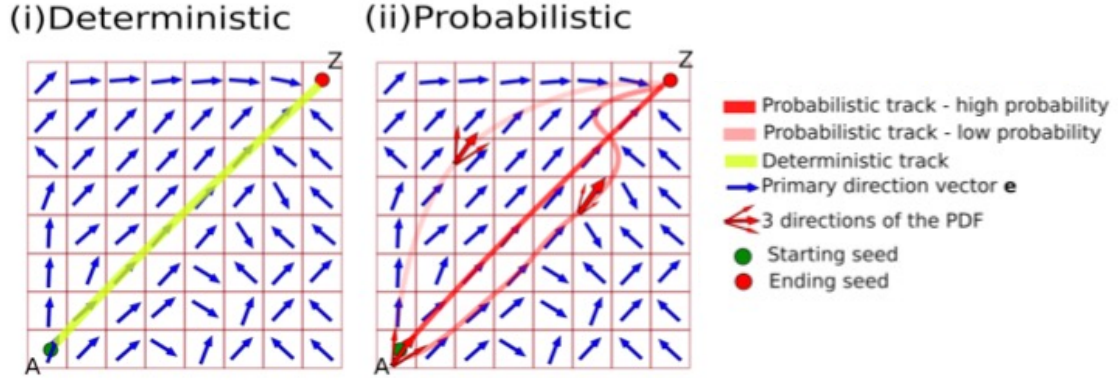


Figure 1.12: 2D example showing the difference between deterministic (i) and probabilistic (ii) tractography. (i) Deterministic approach for which the result is the single yellow line. (ii) Probabilistic approach for which the possible tracks are represented in red. Modified from [9].

1.3 Functional Imaging

Unlike structural imaging, which allows the study of brain structure through anatomical images, functional imaging allows the study of affective and cognitive processes. It allows, for example, the identification of regions activated when a patient performs a task. The task can for example be a hand movement, the pronunciation of certain words or simply doing nothing and letting the mind wander. This latter task is called resting state. There are different techniques which allow to perform analysis of functional imaging, such as the functional MRI (fMRI), the electroencephalography, the magnetoencephalography, etc. In this thesis, the fMRI data of the patients are studied [2].

The fMRI is commonly used in the field of research thanks to its non-invasive method and high spatial resolution compared to other techniques. When performing an fMRI, the result is a sequence of brain images acquired while the subject is performing a task. This can last several minutes so that the activation linked to the task can be visualised. The multiple volumes obtained are decomposed in voxels. Each voxel corresponds to a particular spatial localisation and is associated to a signal intensity. The changes in intensity between these multiple volumes are used to make inferences about brain activity during the task. This variation in intensity along the acquisition of the images is called the time course and is illustrated in Figure 1.13 (a) [2].

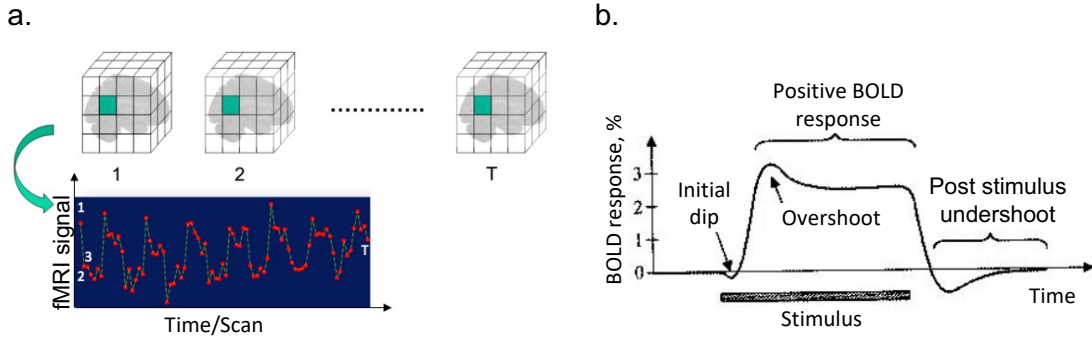


Figure 1.13: (a) Upper part: Sequence of images acquired during the task with decomposition in voxels. Lower part: Variation in intensity of green voxel along the multiple images. From [2]. (b) Variation of the BOLD response depending on the time. From [2].

The signal of the fMRI is related to changes in the blood flow and especially in the blood oxygenation. Indeed, the neuronal activity causes an increase in oxygen consumption inside the capillaries. The blood becomes locally deoxygenated which disrupts the signal since oxygenated and deoxygenated bloods have different susceptibilities. The appearance of locally deoxygenated blood in the regions activated by the task causes an increase in susceptibility. It is as if there were local inhomogeneities of the external field but now, linked to the phenomenon of blood deoxygenation. The dephasing of the spins, determined by T_2^* , will thus go much faster with a decrease in the signal (M_{xy}). This is illustrated in the first part of the curve in Figure 1.13 (b) (initial dip). When there is locally deoxygenated blood, a massive influx of oxygenated blood appears to chase away this deoxygenated blood because of neurovascular coupling. Due to a decreased susceptibility of the oxygenated blood, activated brain regions will be characterised by a longer T_2^* than non-activated brain regions. A small resumption of the signal will be therefore observed for the activated brain tissue, as seen in the second part of the curve in Figure 1.13 (b) (overshoot and positive BOLD response). In fMRI, heavily T_2^* -weighted sequences are thus used so a difference in the signal intensity is observed when performing a task.

The response obtained from an fMRI is called BOLD (Blood Oxygen Level Dependent) contrast response. The higher the field strengths, the greater the differences in susceptibility and thus, the greater the BOLD response. It is thus common to use higher fields for fMRI [2, 3].

In practice, the sequences used in fMRI are often the EPI sequences (described in Section 1.2) as the whole volume can be covered in a short scan time of 10-100ms. The time needed to observe the BOLD response of an activated tissue is approximately 12 to 13 seconds but slightly varies depending on the patient and on the region activated [2, 3].

Chapter 2

Graph Theory and Machine Learning

This chapter explains the notion of connectivity and how it can be applied to the human brain. Then, concepts about machine learning are developed.

2.1 Graph theory

Graph theory aims at understanding systems of interacting elements. A graph is used to model a system as a network of nodes linked by edges. The brain is a highly interconnected organ in which different brain regions are mutually dependent on each other. Studies on brain connectivity therefore hold great promise for shedding a new light on brain disorders. Brain connectivity is used for this reason in this thesis.

In the context of brain networks, nodes represent neural elements, such as brain regions, while edges represent connections between nodes, such as anatomical synapses or pathways [14]. Graphs can be binary, also called unweighted, or weighted. For binary graphs, all edges have the same weight equal to one. On the other hand, for weighted graphs, all edges have different weights. These weighted graphs retain more biological information and are more appropriate to data emerging from tract-tracing [15]. Graphs can also be directed, meaning they contain only directed edges, or undirected if they contain only undirected edges.

Connectivity

The connection matrix is one of the most elementary representations of a graph: nodes are represented by rows and columns of the matrix while edges are represented by matrix entries [14]. Simple examples of connection matrices obtained from a graph can be seen in Figure 2.1.

In neuroscience, a complete description of the structural connections between elements of the nervous system is called a connectome [16]. It was first defined as a matrix representing all possible pairwise anatomical connections between neural elements of the brain. After several years of research, the concept has expanded and a connectome is now defined as a matrix of anatomical connections between

large-scale brain areas but also between individual neurons or a matrix of functional interactions [17].

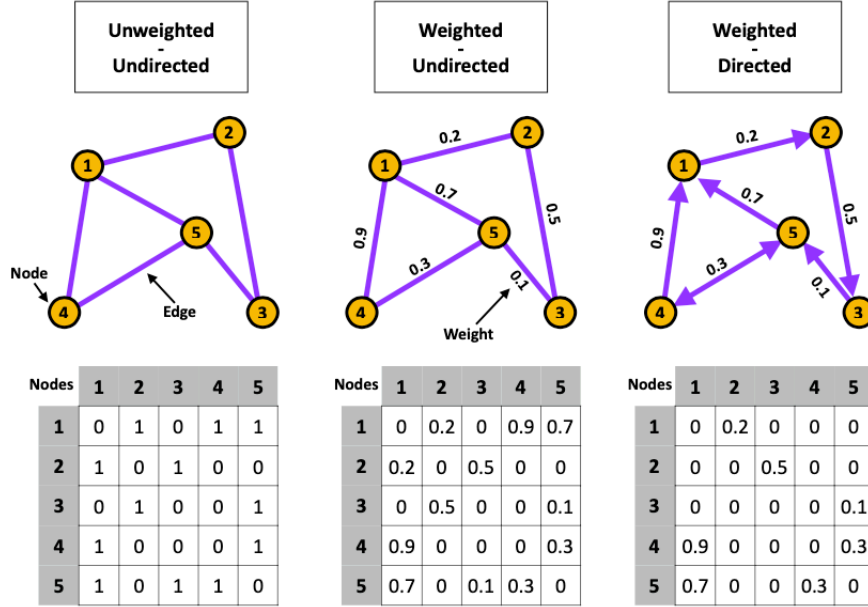


Figure 2.1: Examples of weighted/unweighted and directed/undirected networks with the corresponding connectivity matrix. Nodes are shown in orange and edges in purple. The weights of the weighted networks are overlayed on the edges. The value in the matrices is equal to the weight of the connection between the two corresponding nodes. The connectivity matrix of the directed graph (third one) is not symmetric. Figure from [18].

Structural connectivity

Tractography can be used to construct a weighted, undirected connectivity matrix. Each element of this matrix quantifies the density of axonal projections between each pair of nodes which represent regions of the brain. One needs to select appropriate initial points for the tracking algorithm. When the regions to map are situated in the grey matter, fiber tracking is more difficult because diffusion is poorly constrained in these regions. One therefore needs to place the initial points at the interface between white matter and grey matter. However, the geometry of fibres near the interface of gray and white matter is complex which can cause problems for diffusion tracking algorithms [17].

One limitation of deterministic diffusion tractography is its difficulty to reconstruct long range pathways. For long range pathways, there is indeed a higher probability that the pathway intersects other fibers. The tractography streamlines are propagated in discrete steps and each step is an opportunity for the streamlines to veer off course. The long-range pathways are thus challenging to reconstruct.

The tractogram allows to know which regions are connected but does not give any indication about the source and the target. The connectivity matrix obtained from the tractogram is therefore undirected [17].

Functional connectivity

Functional connectivity refers to symmetrical statistical dependency between elements of the brain. It is commonly measured based on resting state fMRI. The recorded signal during resting state is supposed to reflect spontaneous neural dynamics. Functional connectivity is commonly measured using the Pearson correlation coefficient between the time courses of the signal recorded in different regions. In a functional connectome, each value of the matrix represents therefore a statistical measure of synchronised activity. This does not necessarily imply an anatomical connection [16, 17].

Common network measures

A wide variety of measures can be used to quantitatively describe the topology of graphs. For example, a common measure is the degree of nodes. In undirected graphs, the degree of a node is given by the number of edges connected to that node. It is also called **degree centrality**. In directed graphs, the degree of a node is decomposed into indegree, given by the number of incoming edges, and outdegree, given by the number of outgoing edges [14].

One can also look at **betweenness centrality**. This measure allows to detect the amount of influence a node has over the flow of information in the network. This is used to find bridges from one part of a graph to another [19].

2.2 Machine learning

Machine learning is a branch of computational science that designs algorithms which learn from the surrounding environment. The algorithms use input data to achieve a desired task without being literally programmed to achieve a particular outcome. Those algorithms automatically adapt their architecture through repetitions so that they become better as the number of iterations increases. Algorithms based on machine learning have been applied in diverse fields including biomedical applications [20].

In general, machine learning algorithms are used to perform predictions on unknown data. The model built by the algorithms is first trained on data for which the outcome can be known or not. These algorithms are respectively called supervised or unsupervised. They can then be used to perform regression or classification. The predicted value of regression models is continuous, while the one of classification models is discrete [21]. In machine learning, a feature is a characteristic of the problem studied that can have more or less influence on the prediction.

Over-fitting

The data used to train and test the model must be independent to avoid over-fitting. Over-fitting happens when a model is trained on a set of samples and predicts almost perfectly the target on this set but does not give good results at all when performing it on new data. As the point in machine learning is not to predict data

for which the outcome is already known but to be able to predict data for which the outcome is not known yet, avoiding over-fitting is important [22].

A way of detecting over-fitting is to perform cross validation. This is a way of testing the performances of the model. This can be done only by using data that were not used during the training of the model. It is therefore necessary to remove some of the data from the training dataset beforehand. This data will not be used to train the model, but later to validate the model. One common type of cross-validation is the k-fold cross validation. For example, if $k = 5$, the original dataset will be separated in 5 folds. Then an iterative process takes place. The model is trained on the 4 first folds and evaluated on the fifth one, then the model is trained on folds 1, 3, 4 and 5 and evaluated on the second one, *etc.* [22].

Feature selection

When dealing with high dimensional problems (high number of features), machine learning is subjected to the curse of dimensionality, being the difficulty of working with high dimensional space. First, models built in high dimensional data show lower accuracy compared to model built in low dimensional data. Furthermore, it is complicated to visualise data in a high-dimensional space. Finally, the training time of the model increases exponentially with the number of features.

One common method used to reduce the dimensional space is feature selection. Different methods can be used for that purpose: filter methods, embedded methods and wrapper methods. Filter methods select features based on a performance measure without using a model. The modelling algorithms is used afterwards when the features are selected. Wrapper methods, on the other hand, assess the relevance of features based on the performance of a model. The model is built on different subsets of features and the best subset is chosen based on its performance. Embedded methods perform feature selection during the modelling algorithm's execution. These methods are thus embedded in the algorithm [23].

Models

Typical models used for machine learning on biological data are the Support Vector Machine (SVM) and the Random Forest (RF). The support vector machine is a supervised training method that allows to perform regression or classification. For classification, functions called kernels are used to map the data in another space in which the data are more easily separable in classes [24]. In this space, the goal is to linearly separate the data by a hyper-plane. Maximum-margin hyperplanes are then created as can be seen in Figure 2.2 a. This can be written as an optimisation problem that takes the following form :

$$\min_{w,b,\zeta} \frac{1}{2} w^T w + C \sum_{i=1}^n \zeta_i, \quad (2.1)$$

subjected to

$$y_i((w^T \cdot x_i) + b) \geq 1 - \zeta_i, \quad \zeta_i \geq 0, i = 1, \dots, n,$$

where ζ_i are the slack variables allowing some instances to fall within the margin, but penalizing them, w is the vector orthogonal to the hyperplane, C is a coefficient allowing a trade-off between the margin width and the number of misclassifications, b is a parameter such that $\frac{b}{\|w\|}$ characterises the distance between the hyperplane and the origin of the features space, y is the target (i.e., discrete values depending on the class) and x is the set of data.

The weights, i.e., the values of the w vector, are calculated in order to separate the data with a margin as large as possible with respect to the hyper-plane. Those weights allow to class the features by importance. One advantage of that model is its effectiveness in high dimensional spaces, even in cases where the number of dimensions is greater than the number of samples, which is often the case in the biomedical world. It uses a subset of training points in the decision function. These points are called the support vectors and because the number of these support vectors is small compared to the size of the dataset, it is also memory efficient [25].

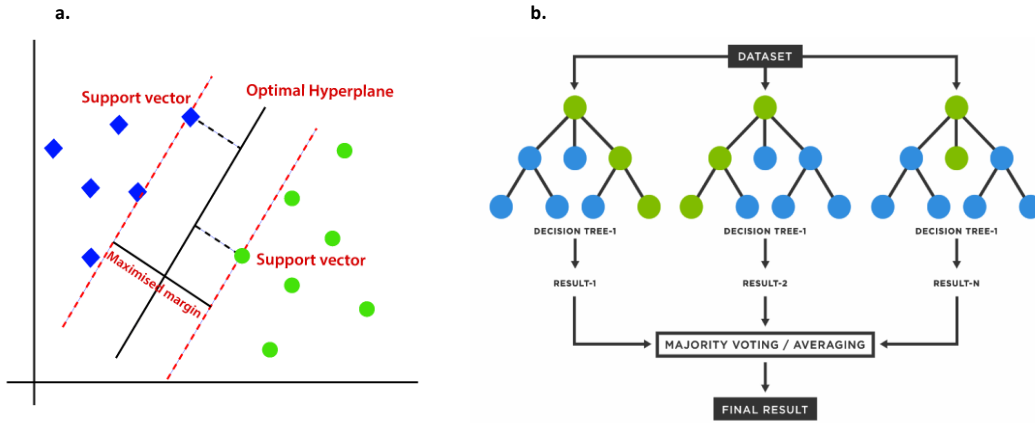


Figure 2.2: (a) Schematic representation of a SVM classifier in a two dimensional feature space. From [26]. (b) Schematic representation of a random forest classifier. From [27].

On the other hand, a random forest is an ensemble of decision trees. A decision tree is a hierarchical tree built using the features of a data set. Each node of the decision tree is split according to a criterion associated with a subset of the features and which divide the data in classes [28]. Each of these trees produces a prediction and the final prediction is chosen by taking into account the prediction of each tree. A schematic RF can be seen in Figure 2.2 b. RF has the advantage of being easily interpretable. However, computations may become far more complex compared to other algorithms and the algorithm is sensitive to small changes in input data [29].

Chapter 3

Vestibular System

This third chapter covers the functioning of the vestibular system, the disorders resulting from its dysfunction and a literature review of the characterisation of the vestibulocortical network on the brain microstructure using diffusion MRI (dMRI) and functional MRI (fMRI).

3.1 Vestibular system anatomy and physiology

The vestibular system, also often referred to as the silent sixth sense, plays an important role in postural control as well as reflexes and eye movements. It has two main tasks:

- It allows the perception of movements and position of the head and body at the level of the somaesthetic cortical areas.
- It continuously informs the integrating centers of the brainstem and cerebellum about the movements and the position of the head and body.

The involvement of the vestibular system is at two levels: the peripheral vestibular system and the central vestibular system [30].

3.1.1 Peripheral vestibular system

The main peripheral component of the vestibular system is the labyrinth, a complex set of interconnected channels. It is a part of the inner ear and can be seen in Figure 3.1 a. A membranous labyrinth is included in the bony labyrinth. This membranous labyrinth is filled by endolymph, circled by perilymph and consists in three different parts:

- the two otolith organs, **the utricle** and **the saccule**, responsible for detecting linear and angular accelerations (3 degrees of translation and rotation) and gravity,
- the **semicircular canals** responsible for detecting angular accelerations (three degrees of rotation).

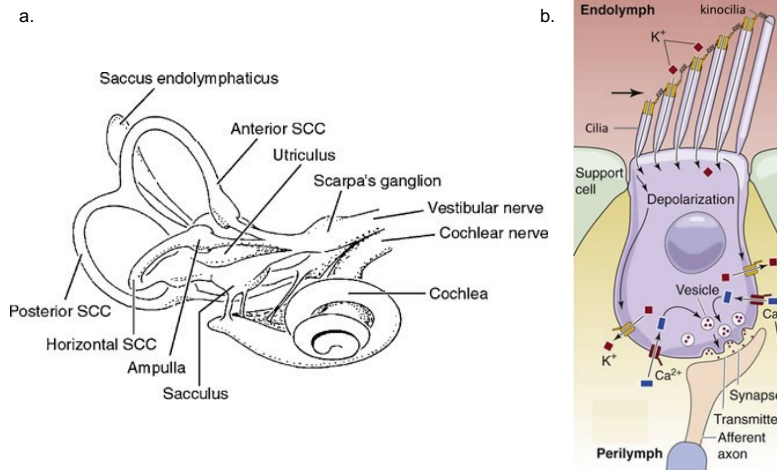


Figure 3.1: (a) Labyrinth with the semicircular canals (SCC), the saccule and the utricle. From [31]. (b) Hair cells in the labyrinth. From [32].

In order to maintain spatial orientation and to drive reflexes that stabilise vision and balance, the vestibular system transforms forces associated with head acceleration and gravity into biological signals. These signals go through the cortex to facilitate orientation and to the motor centers for postural and ocular stability [33]. Thanks to the hair cells, found in all vestibular organs, the transduction of movements into nerve impulse can be made [30].

Hair cells - Hair cells are the basis of vestibular function. In the utricle and saccule, hair cells lie in a sensory epithelium called the macula while in the semicircular canals (SCC), they are in the ampullae, which are located at the base of the semicircular canals, near the utricle as can be seen in Figure 3.1 a. On the apical end of each hair cell lies a bundle of stereocilia for which the height increases stepwise from one side to the other. A longer and thicker cilium, called kinocilium, is located next to the tallest stereocilium, as represented in Figure 3.1 b. Apical links connect the cilia to each other. Therefore, since the cilia of hair cells are organised from longest to shortest, deformation of one cilium causes a pull or a push on the apical link, leading to a cascade of deformation of the adjacent cilium causing the opening or closing of mechanosensory channels.

Hair cells systems have the particularity to be continuously activated due to a small percentage of mechanosensory channels remaining open at rest. This results in a constant depolarisation of the resting membrane potential, triggering the release of neurotransmitters on the afferent vestibular nerve fibers which are thus activated. A deflection of the cilia toward the kinocilium opens more mechanosensory channels leading to an increase of the spontaneous activity. On the contrary, a deflection of the cilia away from the kinocilium closes mechanosensory channels leading to a decrease of the spontaneous activity [30, 33].

The receptors of the inner ear all work on the same basic principle being the activation of the hair cells by the application of an external force. However, in each vestibular organ, stereocilia have a specific orientation. In semicircular canals,

stereocilia are all oriented in the same direction while in otolith organs, a region called the striola separates hair cells into two populations of opposite orientation.

Semicircular canals - There are three semicircular canals: the anterior, posterior and lateral SCC. These canals are all orthogonal to each other and each of them is filled with a fluid called endolymph and encircled by perilymph. The movements of these two liquids caused by head motion stimulate the hair cells of the vestibular receptors. Hair cells in the ampullae are embedded in a gel-like mass called the cupula. When the head is rotated in the plane of the canal, there is a relative movement of the endolymph with respect to the canal wall, which pushes on the cupula, causing the displacement of stereocilia. Depending on the semicircular canal affected by the head movement, the rotation of the head can be determined. For example, the horizontal SCC detects rotation of the head around a vertical axis, while the anterior and posterior SCCs detect rotations of the head in the sagittal and the frontal planes. Information about the strength of the rotation is also perceived to know how much and how quickly the liquid is moving.

Note that linear accelerations of the head exert equal forces on both sides of the ampullae, thus causing no displacement of the stereocilia: translational movements do not provoke any response in the semicircular canals. It is the utricle and saccule which allow to detect them [30, 34].

Otolith organs - As explained above, the saccule and utricle have a sensory epithelium, the macula, where hair cells lie. Above the hair cells is a gelatinous layer, called the otolithic membrane, in which calcium carbonate crystals are embedded.

As already mentioned, the hair cell populations are oriented with respect to the striola, which divides hair cells into two populations of opposite orientation. As a consequence, a head movement in the axis of the striola will excite the hair cells on one side and inhibit those on the other side. The saccular macula is oriented vertically and thus responds to vertical movements while the utricular macula is oriented horizontally, responding to lateral movements [30, 33]. They also respond to angular movements but are the only ones responsible for detecting linear movements.

3.1.2 Central vestibular system

The information collected from the peripheral system follows the auditory or vestibular cochlear nerve (CNVIII) to reach the brainstem. The major components of the first level of the central vestibular system are the vestibular nuclei (VN) composed of four parts: the superior, lateral, medial and inferior VN. The VN are a major source of connections including ascending pathways with the abducens, trochlear and oculomotor nuclei, all three involved in the vestibulo-ocular reflexes (VOR), and descending pathways to the spinal cord implicated in vestibulospinal reflexes. A lot of projections with the cerebellum also participate in vestibulocerebellar reflexes and with the forebrain through the thalamus, hippocampus and other limbic structures.

The information coming from the peripheral vestibular system is combined with information from other sensory systems and converges to the VN such that a central estimate or a vectorial representation of head and body position and motion in space is computed. The VN are the first relay of the central vestibular system which is largely connected to different networks involved in the perception and control of the body in space. It is thus at the interface between the sensory and motor processing.

The firing of many central nervous system neurons is modified according to the activity originating within the vestibular system [35, 36, 37].

A representation of the VN can be seen in Figure 3.2. Ascending as well as descending pathways in multiple brain areas are shown.

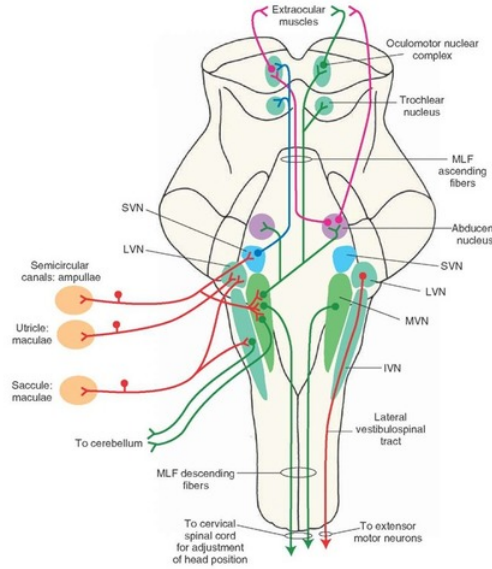


Figure 3.2: Representation of the vestibular nuclei. From [38].

3.2 Vestibular system dysfunction and tests

3.2.1 Vestibular system dysfunction

When the vestibular system is functioning normally, there is usually no conscious awareness of its activity.

However, in the presence of abnormal motion environments or pathology of the peripheral organs or their central pathways, the vestibular sensations can become obtrusive and are referred to as vertigo, dizziness or a sense of imbalance. As a result, sensations of discomfort and nausea can occur leading to motion sickness [36]. However, contrary to other sensory modalities, vestibular dysfunction is frequently transient. Indeed, vertigo symptoms rarely or partially persist in patients suffering from lesions or disorders of the peripheral or central vestibular regions. When

peripheral vestibular disorders occur, the somatosensory information of the body in its environment is unknown, resulting in a mismatch between this somatosensory information and what is known by the central vestibular system. This mismatch results in a strong sensation of imbalance, dizziness or vertigo. However, thanks to its plasticity, the central vestibular system is able, within few minutes or hours, to reorganise itself to compensate the disorder, decreasing the symptoms. This suggests that the vestibular cortical network organisation might be highly robust, redundant and plastic [39].

When the vestibular system is affected, disorders differentiate into peripheral and central causes. It is important to identify from which type of vestibular dysfunction a patient suffers, to determine the most suitable therapeutic approach. For this purpose, various vestibular tests were set up to find the dysfunction from which the patient suffers.

3.2.2 Vestibular tests

Different tests can be performed in order to assess the proper functioning of the vestibular system. Below are described the tests performed on the patients whose data were used in this master thesis.

Video Head Impulse Test (vHIT) - This test allows to analyse the integrity of the six semicircular canals. During the examination, the patient stares at a point straight ahead while the head is suddenly turned in the plane of the canal being tested by the examiner. To compensate this movement, the eyes should normally do a movement in the opposite direction in order to keep the point fixed stable on the retina. This is the vestibulo-ocular reflex. Using a camera, infrared light and software, the vestibulo-ocular reflex is recorded in order to detect correction movements in the eyes. There is an anomaly if the speed of the eyes and of the head is not the same. The eyes should remain fixed on the point in front and not move with the head movement and then return to fix the point [40].

Videonystagmography (VNG) - A nystagmus is an involuntary eye movement which makes the eye moves from side to side or up and down. This happens when conflicting information from the balance system in the inner ear and the eyes comes to the brain. This movement is considered as normal when the head is moved or when people look at specific kinds of patterns. However, it is not considered normal if the head is not moved or if it lasts a long time. During the VNG test, the patient is in a dark room and wears special goggles with a camera recording eye movements. The test is divided in three main steps. The first one is the ocular testing. During this step, the patient follows a moving or non-moving dot on a light bar. The second step is the positional testing. During this part, the head and the body are moved in different positions and one looks if a nystagmus appears or not. The last step is the caloric testing. During this step, warm and cool water or air is put in the ear. This is supposed to cause nystagmus. If hot water or air is put in the ear, the eyes should move away from the ear and then rapidly come back. This

is an ipsilateral nystagmus. If it is cold water or air, the eyes should move slowly toward the ear and rapidly come back. This is a contralateral nystagmus. If it is not the case, this may mean that there is damage to the inner ear or the nerves of the inner ear. If damaged, the response is weaker or there is no response at all [41]. Another paradigm of VNG is the rotational test. For this test, the patient is placed in a chair which rotates at constant speed after an initial acceleration. The induced eye movements are measured during rotation and when rotation is stopped. This test measures the gain of the vestibulo-ocular reflex and especially detects asymmetry between the semicircular canals of the two inner ears and also the central vestibular compensation when a persistent peripheral vestibular lesion remains stable [42].

Vestibular Evoked Myogenic Potential (VEMP) - This test allows to check the functioning of the saccule and the utricle. It relies on the fact that these two structures respond to sound stimuli. The test is composed of two parts: the cVEMP which is the cervical test for the saccule and the oVEMP which is the ocular test for the utricle. For the cVEMP test, electrodes are attached to the skin near the sternocleidomastoid muscle ipsilateral to the ear that is tested. The patient is lying on the back and raises the head in such a way that the muscle is contracted. A sound is sent and the software analyses the relaxation of the muscle. For the oVEMP test, electrodes are glued on the skin over the inferior eye muscle on the opposite side of the ear to be tested in order to collect its contraction [40].

Postural assessment - Postural control is the result of the integration of diverse sensory systems such as visual, proprioceptive and vestibular. Balance disorders are thus frequent in vestibular pathology. There are several techniques to quantify these balance disorders. Fixed or mobile platforms can be used with static or dynamic posturography methods.

For example, the Fukuda test consists in asking to the subject to walk on the spot with his index fingers pointing forward. In case of vestibular asymmetry, the patient shows a postural deviation to the left or to the right.

Another example is the equitest, a device allowing to test the static and dynamic balance of the patient. It includes two main types of tests: a sensory organisation test and a motor test. The first one assesses the effect of six different sensory environments on postural control. To this end, the equitest consists in a force platform and a visual screen located in front of him, both either fixed or controlled by the subject's center of gravity. For the motor test, the latencies of the postural responses in response to anteroposterior or angular displacements of the platform are observed.

The equitest is very helpful for the quantification of balance disorders [42].

3.3 Connectome analysis for vestibular system

In this section are presented articles addressing vestibular system and networks, making the link between the macro- and microstructural information obtained with dMRI and fMRI, the measures of connectivity and the dysfunction of the vestibular system. The regions involved in the vestibular system and studied in the literature

are varied. This section aims to summarise some of the literature by topics in order to highlight the regions and results and to compare them.

Connectivity and network measures have been used in several articles to investigate the vestibular system. The outcomes of these articles can be seen in Table 3.1.

Table 3.1: Analysis of connectivity and network measures in the vestibular network.

“The “vestibular neuromatrix”: A proposed, expanded vestibular network from graph theory in post-concussive vestibular dysfunction”, Jeremy L. Smith et al. [43]	
Population	27 healthy control subjects and 18 patients with subacute post-concussive vestibular dysfunction (PCVD)
Methods	Construction of functional connectome (computed from Pearson correlation) and structural connectomes (based on probabilistic tractography) between 82 regions of interest (ROIs). Then, analysis of the node centrality and community structure (= the ability of nodes of the network to be grouped into (potentially overlapping) sets of nodes such that each set of nodes is densely connected internally).
Results	- Consensus network not only including vestibular areas but also regions subserving visuospatial processing and attention as well as motor and affective processes. - Consensus network common to both healthy and PCVD patients: proposed to represent the necessary substrate for higher vestibular function. - From nodal community structure analysis, vestibular and cognitive-affective ROIs are closely associated and group-wise differences are observed in the integration of some areas such as cingulate, frontal eye field or posterior hippocampal areas in the functional networks. ⇒ Cognitive, affective, and proprioceptive substrates may be associated to vestibular dysfunction, as observed in clinical practice, and therefore highlight the need to consider these areas during clinical diagnosis.
“The human corticocortical vestibular network”, Raiser et al. [39]	
Population	29 right-handed healthy subjects.
Methods	Construction of structural and functional connectomes between 10 regions of interest of the vestibular system. Then, analysis of common network measures such as degree and betweenness centrality.
Results	- Robust intrahemispheric structural connectivity. - Strong functional connection between homotopic nodes of the left and right hemispheres, despite the absence of structural connections.

“Alterations in Resting-State Functional Brain Connectivity and Correlations with Vestibular/Ocular-Motor Screening Measures in Postconcussion Vestibular Dysfunction”, Anna Trofimova et al. [44]	
Population	12 patients with subacute PCVD and 10 healthy subjects.
Methods	Analysis of resting-state fMRI brain connectivity alterations. The correlations between the resting-state fMRI connectivity and the symptoms provoked during vestibular/ocular-motor screening assessment were also studied.
Results	Significant group differences between right posterior hippocampus and right anterior insula and between the vestibular-sensorimotor network and the right prefrontal cortex highlighted from statistical tests on graph theory measures (degree centrality) of fMRI brain connectivity.

Analyses of the vestibular networks have been performed to try to identify modules and hubs, i.e., cluster of highly connected areas, within the vestibular system. One article addressing this topic can be seen in Table 3.2.

Table 3.2: Analysis of modularity in the vestibular network.

“Structural connectome and connectivity lateralization of the multimodal vestibular cortical network”, Iole Indovina et al. [45]	
Population	974 healthy subjects (794 right handed, 530 females)
Methods	Investigation of the structural connectivity properties of the vestibular system and identification of module. Construction of connectivity matrices based on tractogram and grey matter parcellation. Modularity of the network estimated by using the Newman-Girvan algorithm (each module is extracted as a group of non-overlapping nodes by maximising the number of links inside a module and minimising the number of links outside the module). Louvain-like algorithm then used to maximise modularity.
Results	Three modules found in the right hemisphere and four in the left one: - <i>parietal-insular prefrontal module</i> (light green regions for right hemisphere and light red + dark green regions for the left one in Figure 3.3) - <i>visual module</i> (purple regions for the right hemisphere and dark green for the left one) - <i>limbic-subcortical module</i> (in yellow for right hemisphere and in blue and dark red for the left one) including the hippocampal formation, the cingulate cortex, parieto-occipital regions, the thalamus, dorsolateral caudal region of BA 6 and ventral insula. This one is split in two for the left hemisphere

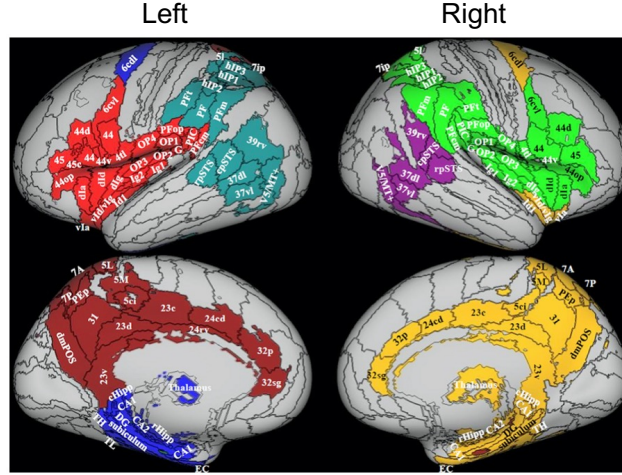


Figure 3.3: Modules selected to belong to the multimodal vestibular cortex, each colour representing a module. From [45]

Two key regions of the cortical vestibular network have been investigated by many researchers: the parieto-insular vestibular cortex (PIVC) and the posterior insular cortex (PIC). Different articles analysing those regions are presented in Table 3.3.

Table 3.3: Analysis of the PIC and PIVC regions of the vestibular system.

“Structural connectome and connectivity lateralization of the multimodal vestibular cortical network”, Iole Indovina et al. [45]	
Population	974 healthy subjects (794 right handed, 530 females)
Methods	Structural connectivity matrices made from multi-shell dMRI and definition of network hubs of the vestibular system by analysing node strength, local efficiency and betweenness centrality.
Results	- Highest hubness scores for regions corresponding to PIC and Parietal Operculum (OP2). - PIC plays a key role in visuo-vestibular processing while OP2 has a structural connectivity pattern consistent with integration of mainly vestibular somato-sensory and motor information.
“White Matter Connectivity of the Visual–Vestibular Cortex Examined by Diffusion-Weighted Imaging”, Anna M. Wirth et al. [46]	
Population	15 right-handed healthy participants
Methods	Analysis of probabilistic fiber tracking based on dMRI to study the ipsilateral connectivity of PIVC and PIC.
Results	Both PIVC and PIC share ipsilateral connections to the insular/lateral sulcus, superior temporal cortex, and inferior frontal gyrus and are connected with lateral nuclei of the thalamus and the basal ganglia. PIVC is linked with posterior parietal and inferior frontal cortex, whereas PIC is linked with superior temporal and inferior parietal cortex.

“The parieto-insular vestibular cortex in humans: More than a single area?”, Sebastian M. Frank and Mark W. Greenlee. [47]	
Population	1 average MRI human brain
Methods	Analysis of functional brain images acquired while performing a galvanic vestibular stimulation
Results	- PIVC coincides with the Parietal Operculum (OP2-3), situated in the mid-posterior Sylvian fissure while PIC is located at a posterior site in the Sylvian fissure, in the retroinsular cortex, posterior to the PIVC. - Both areas are functionally separated: PIVC is a vestibulo-somatosensory and motor integration area, while PIC is a visual-vestibular area
“Structural and functional connectivity mapping of the vestibular circuitry from human brainstem to cortex”, V. Kirsch et al. [48]	
Population	24 healthy right-handed volunteers
Methods	Analysis of vestibular connectivity in the living human brain between the VN and the PIVC by combined structural and functional connectivity mapping.
Results	- Functional and structural link between the VN and the ipsi- and contralateral PIVC. - 5 separate vestibular pathways. 3 run ipsilaterally with 2 running through the posterolateral or paramedian thalamic subnuclei, and a 3rd bypassing the thalamus to reach the inferior part of the insular cortex. The 2 other pathways cross either in the pons or the midbrain, running through the posterolateral thalamus to the vestibular cortex. - Transcallosal interconnection of PIVC region through the antero-caudal splenium, suggesting a vestibular circuitry extending from the brainstem to the cortex with 3 crossings in the brainstem (VN, pons and midbrain), and 1 cortical crossing through the splenium.

Finally, in many articles, the lateralisation of the vestibular system has been proven, as can be seen in Table 3.4.

Table 3.4: Analysis of the lateralisation of the vestibular system.

“Alterations in Resting-State Functional Brain Connectivity and Correlations with Vestibular/Ocular-Motor Screening Measures in Postconcussion Vestibular Dysfunction”, Anna Trofimova et al. [44]	
Population	12 subacute PCVD patients & 10 healthy subjects.
Methods	Analysis of correlations between the resting-state fMRI connectivity and the symptoms provoked during vestibular/ocular-motor screening assessment.
Results	Predominance of right cerebral hemisphere foci.

“White Matter Connectivity of the Visual–Vestibular Cortex Examined by Diffusion-Weighted Imaging”, Anna M. Wirth et al. [46]	
Population	15 right-handed healthy subjects
Methods	Analysis of probabilistic fiber tracking based on dMRI to study the ipsilateral connectivity of PIVC and PIC.
Results	Right hemispherical dominance except for PIC connectivity in the caudal middle frontal cortex.
“The human corticocortical vestibular network”, Raiser et al. [39]	
Population	29 right-handed healthy subjects.
Methods	Construction of structural and functional connectomes between 10 ROIs of the vestibular system. Then, analysis of common network measures (degree and betweenness centrality) and of the vulnerability of each node, giving an indication about the importance of the node on the global efficiency (the more distant the nodes, the less efficient for exchanging information) of the network.
Results	- Right-sided dominance with larger nodes and stronger connectivity values both structurally and functionally. - Higher functional vulnerability in the right cortex. This lateralisation might explain the association of higher prevalence of vestibular symptoms with modified right hemispheric connectivity.
“Structural connectome and connectivity lateralization of the multimodal vestibular cortical network”, Iole Indovina et al. [45]	
Population	974 healthy subjects (794 right handed, 530 females)
Methods	Investigation of the structural connectivity properties of the vestibular system and identification of module. Construction of connectivity matrices based on tractogram and grey matter parcellation.
Results	Significant level of lateralisation of structural connectivity for several regions. In general, localisation to the left side. Structural connectivity lateralisation did not depend on the handedness nor on gender. Left lateralisation of the PIC and OP2 selected as hubs of the vestibular network.

As can be seen from this literature review addressing vestibular system, MRI and networks, debates are still open concerning the structural and functional connectivity patterns of the human vestibular system. Further investigations on connectomics are therefore needed to better understand this complex system and the effects of clinical vestibular syndromes.

Part II

Analysis of connectomes for vertigo

Chapter 4

Methods

4.1 Data

The data used in this thesis were acquired in Saint-Luc hospital on a 3T Siemens MRI scanner in the framework of the PhD thesis of Dr. Anais Gregoire. The images were acquired on 74 deaf participants. The participants are distributed as follows:

- 7 control participants (who have never experienced vertigo).
- 13 participants who have experienced vertigo but have compensated.
- 54 participants who still experience vertigo.

Participants were classified in these three categories based on the results obtained for the vestibular tests described in the section *Vestibular Tests*.

Three modalities of images are used: anatomical images (T1-weighted of resolution 256x256x224, TE = 4.92 [ms], TR = 21 [ms]), resting state functional images (resolution 64x64x35x180, TE = 30 [ms], TR = 2000 [ms]), DTI images (resolution 128x128x56x65, TE = 74 [ms], TR = 9000 [ms]). For the DTI images, one image is taken with a b-value equal to 0 [s/mm^2] and the 64 others with a b-value equal to 800 [s/mm^2]. Because only one b-value is used for the 64 images, the images are called single-shell.

The main objective of this thesis is to use these data to build structural and functional connectomes serving as input to machine learning algorithms in order to identify connections between regions of the brain which allow to differentiate between control participants, participants suffering from vertigo and participants that have compensated for vertigo.

Besides, as the lower part of the brain was cut on the dMRI and fMRI data, the cerebellum was not complete and a specific analysis of anatomical data is thus also used to find regions of interest in the cerebellum as explained in the *Statistical cerebellum analysis* section.

To analyse the connectivity, one needs to define specific regions of the brain as regions of interest. Two atlases are used for that purpose. The first one is the 'Brainnetome' atlas consisting in 246 regions and which has been designed to "help researchers to more accurately describe the locations of the activation or connectivity in the brain" [49]. It has indeed already been used in studies such as the one made by Iole Indovina et al. to analyse the structural connectome and the connectivity lateralisation of the multimodal vestibular cortical network [45]. This atlas covers cortical regions and subcortical nuclei but not the cerebellum. The latter playing an important role in the vestibular system, a second atlas is used to cover the cerebellum. The atlas is 'Suit' and contains 34 regions [50]. These two atlases are put together and labelled in such a way that the regions of the left hemisphere (including the left part of the cerebellum) are labelled from 1 to 136, the regions of the right hemisphere (including the right part of the cerebellum) are labelled from 137 to 272 and the regions of the vermis in the cerebellum are labelled from 273 to 280. The correspondence of the labels with the regions name is available in Appendix A. More details on each region's name can be found on the site [49]. The visualisation of the atlas on a brain image can be seen in Appendix B. The complete atlas was used for further analyses in order to avoid making any assumptions and to keep open all possibilities of regions allowing to differentiate between vertigo and non-vertigo participants.

Data were received in the DICOM format and were converted to Nifti with the MRICron tool [51]. Then, before beginning the preprocessing step, data have been inspected and it was found that the diffusion images of one vertigo participant were not correctly acquired and were thus unusable. Moreover, as there were no functional images available for this participant, he had to be removed from the study.

4.2 Structural analysis

Structural data are used to build the structural connectome. For this purpose, tractography must be used. However, data must first be preprocessed in order to remove some artefacts described in the section *Artefacts* and to obtain some specific information from these data, such as the FA.

4.2.1 dMRI preprocessing

The ElikoPy library [52] is used to preprocess the data. This allows to correct and enhance the quality of raw images in an automatic way and allows to deal with several participants. The library is used on the CECI-Clusters, to have access to powerful computing equipment [53]. This preprocessing pipeline includes brain extraction, reslicing, denoising, suppression of the Gibbs ringing artefacts, estimation and correction of the susceptibility induced and Eddy current induced distortions, modelling and correction of participant movements and finally bias field correction. Some of these steps were applied on the data depending on the

parameters given to the *preproc* function of the ElikoPy library. The steps used are:

- the brain extraction step which allows to extract the brain from the skull and other tissues surrounding the brain. This step was mandatory and applied with the default values of the parameters [52]. This also allows to get a brain mask which is a binary mask where the voxels corresponding to the brain take value 1 while the other tissues and the background take value 0.
- the denoising step which allows to reduce the Rician noise typically found in MR images due to the coil system [54].
- the susceptibility field estimation which allows to estimate the susceptibility artefacts. The estimation of this susceptibility requires images acquired with a reverse phase encoding. In this case, images acquired with reverse encoding are not available. The susceptibility is therefore estimated based on the T1 structural image.
- the effect on the brain of Eddy current, motion and susceptibility artefacts (estimated in the previous step) are the same and are therefore corrected together by using the eddy parameter. The images are acquired on adult people and the motion is therefore not expected to be large. Inter-volume motion correction might thus be sufficient.
- biasfield which allows to compute the variability of the signal in tissues. It is used with the default values for the parameters.

The result of this preprocessing can be seen in Figure 4.1.

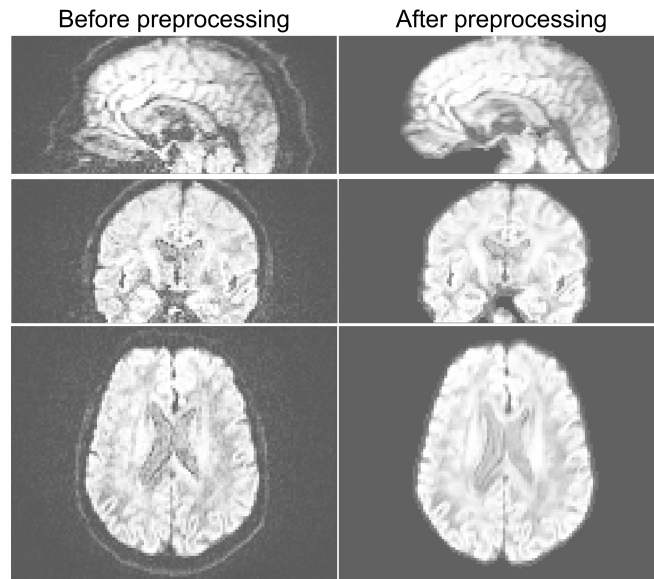


Figure 4.1: Comparison of the DTI data **Left:** before and **Right:** after preprocessing with ElikoPy.

Then, the *dti* function from ElikoPy was finally used to get the Fractional Anisotropy (FA) map that is used for the registration.

4.2.2 dMRI registration

The regions of interest listed in Appendix A must then be registered on the brain of each participant. Image registration is the task of finding a meaningful spatial transformation or mapping from one image to another taken in different times or from different viewpoints or by different imaging systems as in multi-modality imaging [55]. This is an iterative process which allows to find a deformation field to apply to the image to be registered. This process includes affine transformations such as translation, rotation, scale and shear. The proximity between the images to be registered is evaluated based on the Mutual Information.

The combination of the two atlases used is in the Montreal Neurological Institute (MNI) space, which is a standard space defined by averaging several human brains [56]. The images of each participant are not in this standard space but in the native space. One needs to register the participant's images and the atlas in order to identify the regions of interest correctly on each participant's brain. This is done using Dipy [57]. The resolution of the atlas is higher than the one of the DTI image. However, the registration is made from the MNI space to the native space to keep only the regions present on the dMRI images but also, to allow performing tractogram from the dMRI images. The deformation field between a FA map in the MNI space (downloaded from [58]) and the FA map in the native space of each participant is computed. This deformation field is applied to the atlas in the MNI space to get the atlas registered in the space of each participant. The difference between the non-registered and registered atlas on the FA map can be seen in Figure 4.2. It can be seen that the registration is not perfect and that some regions spill over in some places, like at the lower part of the brain in the lower right image. However, because the main purpose is to find the connections between regions inside the brain, the outline of the brain is less important and the registration has not been refined.

During this process, some regions were not found on the images in the native space. On the one hand, most of the regions of the cerebellum were not registered because for all the participants, the cerebellum is cut on the DTI images. On the other hand, for three participants (one vertigo participant and two participants who have compensated for vertigo), images were also cut in the upper part because they were placed too high in the MR scanner during the acquisition. These three participants were therefore removed from the participants' set of structural analysis. When a region was not registered for one participant, it was also removed for all other participants. A list containing non-registered regions that were removed for the structural images can be seen in Appendix C.

For further analyses on structural data, 249 regions are thus used out of the 280 and the dataset is composed of 7 control participants, 11 participants who have

experienced vertigo but have compensated and 52 participants who still experience vertigo.

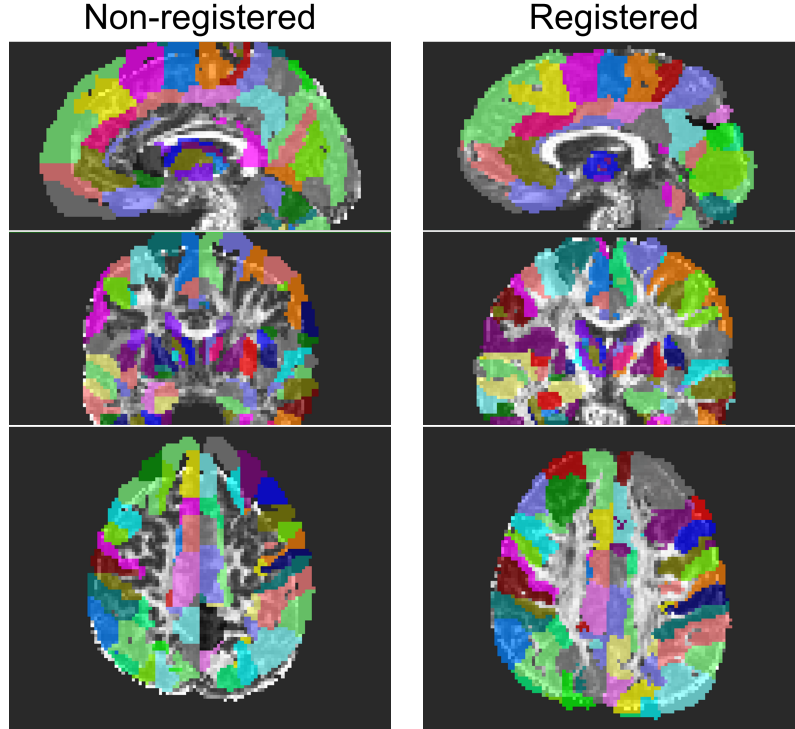


Figure 4.2: Comparison between a FA map on which **Left:** the brain atlas is not registered or **Right:** is registered.

4.2.3 Structural connectivity

Tractogram

As mentioned in the section [Connectivity](#), tractography is needed in order to build the structural connectomes. Two kinds of tractography algorithms can be used: deterministic or probabilistic algorithms. The goal here is to build a whole-brain tractogram to find as many connections as possible between all the regions of interest in the brain. As explained in the section [Tractography](#), using probabilistic tractography has the advantage of allowing to track the fibers through regions of high uncertainty. The deterministic algorithms would stop in these regions [59]. However, it has been shown that probabilistic algorithms yield inherently noisy connection matrices, at least at the single participant level, and hence likely contain numerous false positives [45]. Both methods are therefore used to evaluate which one is the best.

To use tractography algorithms, one needs to provide diffusion data, seed points and a mask in which the tracks are computed. The Dipy tutorials were used to implement these algorithms [13, 60].

First, the deterministic algorithm is used with a white matter mask. The seed points are taken in this white matter mask. Because the regions of interest are located in the cortex, the white matter mask was dilated in order to have an overlap between the grey and white matter and to have seed points located in the regions of interest. The density of seed points, i.e., the number of tracts starting from one voxel, can be tuned. This value was put to 2, meaning 8 tracts start from each voxel. Other parameters of the model are tuned. First, the FA threshold, being the value of the *relative_peak_threshold* on the FA map such that only peaks greater than $\text{relative_peak_threshold} \cdot m$ are used, where m is the largest peak. To obtain anatomically consistent tractograms, this parameter can not be changed too much and was put to 0.8. In the same way, the maximum angle between two tractography steps can be tuned in order to avoid too sharp angles in the tractogram. A value of 45° was used. The step size for tracking was put to half a voxel, i.e., 1mm. The parameters used remain pretty strict.

For probabilistic tractograms, the brain mask (obtained from ElikoPy as explained in Section [dMRI preprocessing](#)) and a higher density of seed points were used. Indeed, a value of density of 3 was chosen, meaning 27 tracts starting from one voxel. Since the goal was to generate as many fibers as possible between the regions of interest, the least stringent parameters possible were favoured. Probabilistic algorithms allow to generate tractograms that are still anatomically correct while using less strict parameters than with the deterministic algorithms. Therefore, the FA threshold was put to 0.4 and the Generalised Fractional Anisotropy (GFA) threshold to 0.25. The GFA threshold is a threshold allowing the tract to propagate when the GFA is higher than the threshold. It demonstrates a net enhancement, in terms of tracking, relative to FA when multiple white matter fiber bundle orientations are present in a voxel [61]. The maximum angle between two tractography steps was put to 20° . The step size for tracking was also put to 1mm. The probabilistic algorithm was therefore used with parameters being less strict than for the deterministic one but the obtained tractograms were still anatomically relevant. The obtained tractogram was further cleaned to remove all the streamlines smaller than 10mm. Indeed, as the brain mask was used instead of the white matter mask, small non-desired streamlines were created in the surrounding of the brain and needed to be removed. To determine the threshold below which streamlines had to be removed, the TrackVis tool was used [62]. It allows to download tractograms and the *Length Threshold* parameter of the *Track* tab allows to see streamlines having lengths greater or lower than a predefined threshold. 10 mm was the threshold used because it allowed to remove the streamlines surrounding the skull, not coming from the white matter as can be seen in Figure 4.3.

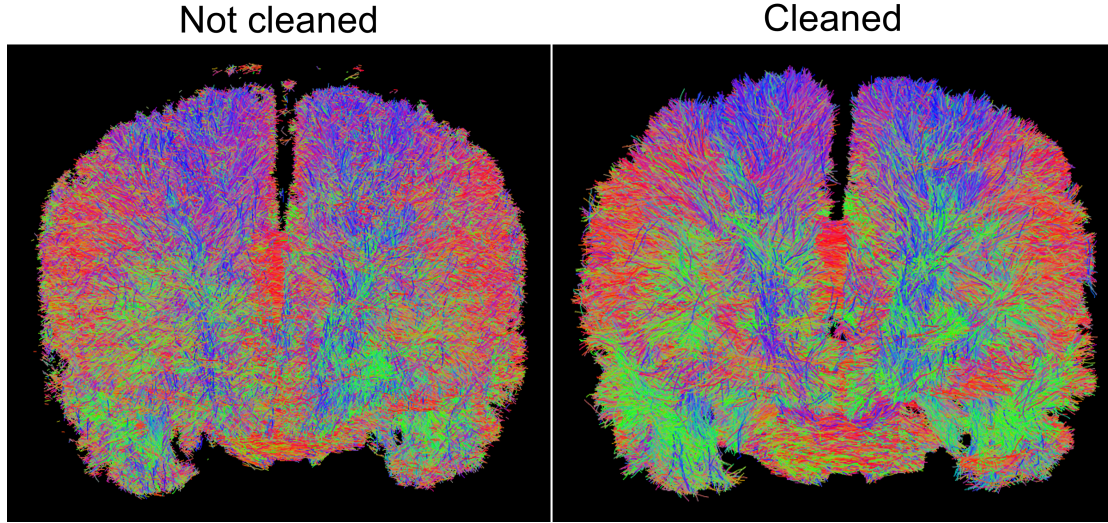


Figure 4.3: Probabilistic tractograms obtained **Left:** without removal of fibers smaller than 10mm, **Right:** with removal of fibers smaller than 10mm.

Structural connectomes

The tractogram obtained is used to build the structural connectome. The values in this connectome represent the number of fibers between two regions of interest. However, the regions in the atlas are not equivalent in size. A large number of fibers start from big regions compared to smaller ones. This difference in size is taken into account by dividing each entry of the matrix by the sum of the volume of the two regions, as done in the articles written by Z. Gu *et al.* [63] and L. Cammoun *et al.* [64]. Moreover, a linear bias towards long fibers is introduced by the tractography algorithm. To correct for this bias, the values are also normalised by the average length of the streamlines between each pair of regions [65].

To build the structural connectivity matrix, the `utils.connectivity_matrix` function of Dipy is used [66]. This function has the argument `"inclusive"`. When passed to true, the connectivity matrix is built taking into account, for each region, streamlines starting from the seed points in this region but also all the streamlines passing through this region. However, the computational time was considerably increased and it was therefore decided to only look at streamlines starting from a region and not passing through it. The hypothesis was made that the impact on the results is limited because the regions of interest are located in the grey matter while the tracts are propagating in the white matter. There are probably few regions crossed by a tract, except perhaps for the subcortical nuclei. It is thus important to keep in mind that only streamlines starting from one region and ending in another are considered for the measure of density.

4.3 Functional analysis

As for the structural data, functional data first have to go through a data preprocessing pipeline before being analysed.

4.3.1 fMRI preprocessing

Five steps are usually performed to preprocess functional data: motion correction, slice-time correction, temporal filtering, spatial filtering and co-registration [67, 68].

Motion correction

Acquiring fMRI images takes several seconds. It is therefore very likely that the participant moves during this time and because of this motion, a voxel of two acquired images may not correspond to the same location. To overcome this issue, motion correction is performed to line-up each functional image with the one before it. The *MCFLIRT* function of the FSL tool is used for this purpose. MCFLIRT is a motion correction tool designed for use on fMRI time series and based on optimisation and registration techniques [68, 69].

As a result of this preprocessing step, six parameters of movement (3 translation and 3 rotation parameters) are obtained. A residual motion less than 0.5-0.75 mm in translation and 0.5° in rotation is considered acceptable [70]. The graphs of the motion of the functional images of one participant can be seen in Figure 4.4. It can be observed that the residual motion after motion correction is in the acceptable range. When this is not the case, the motion correction procedure can be applied several times until residual motion becomes acceptable.

Also note that the motion correction needs a volume of reference on which the correction of all the other volumes is based. The first volume is often taken because an anatomical image is performed before acquiring functional images. The first volume is thus the closest one to the anatomical image on which data are co-registered (see section *co-registration*).

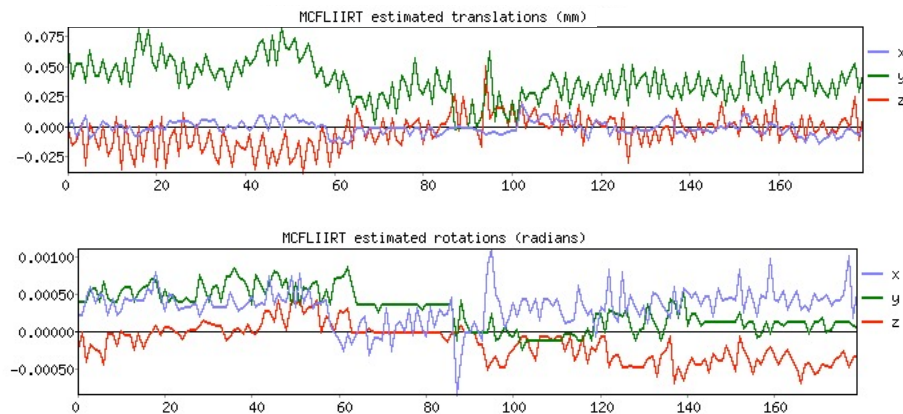


Figure 4.4: Graphs of residual motion for the functional images of one participant. Upper graph: translation motion; Lower graph: rotation motion.

Slice time correction

To obtain fMRI data, 2D slices are acquired and stacked together to create a 3D image. However, the acquisition of the slices is not instantaneous: it results in an accumulation of delays between the first slice and the remaining ones. A slice time correction is therefore needed to correct the fact that the slices are not taken simultaneously. It consists in temporally interpolating the signal in such a way that it appears as though all of the slices were acquired at the same time [71]. The *slice time* function of the FSL tool is used for the slice time correction [72].

Temporal filtering

Due to the heating of the magnet and the gradients MR scanner, the recorded signal is slightly increased. This generates signals of low frequency which constitute noise. The removal of these low-frequency drifts is important as they reduce the power of statistical data analysis and invalidate event-related averaging [73]. A high pass filter is therefore necessary to eliminate this noise. It consists in removing the low frequencies of the signal (those that represent scanner-related noise) while letting the high frequencies pass [2]. The *fslmaths* function of the FSL tool is used for temporal filtering. This is a simple but powerful program to allow mathematical manipulation of images including temporal filtering [74]. The image of a participant before and after applying the high-pass filter can be seen in Figure 4.5.

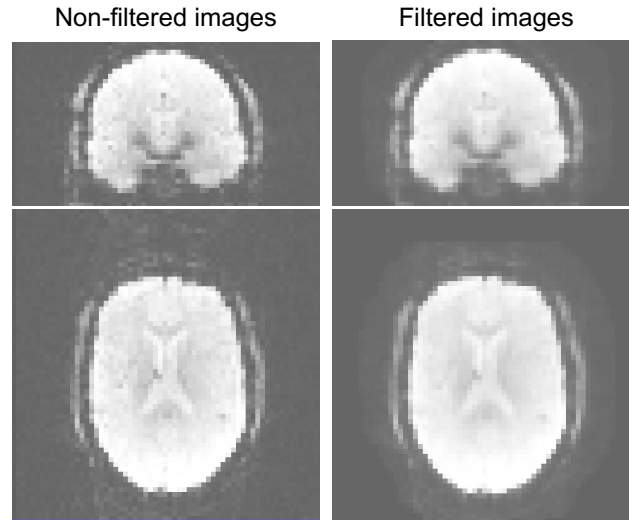


Figure 4.5: High pass filter applied on one functional image. **Left:** Non-filtered image; **Right:** Filtered image.

Spatial smoothing

Some software, such as FSL, usually performs spatial smoothing. To apply spatial smoothing on images, a FWHM Gaussian filter can be used. However, it was decided not to use the spatial smoothing corrected images for further analysis. Indeed, smoothing induces a loss of spatial specificity by spreading out each participant's

signal in space. The signal is therefore no more specific to one particular location [75]. Because the atlas used is made of a lot of regions of interest with some of them being smaller, it seems important to keep spatial specificity. Moreover, it has been proven that the spatial smoothing of fMRI data may overestimate activation spots in the correlation functional map, i.e., a functional connectome computed with the Pearson correlation. The goal here is to build such a functional map and an overestimation of the activity is to be avoided [76].

Co-registration

Finally, as the quality of anatomical images is better than the one of functional images, the latter are often co-registered on the anatomical images, if these are available [2]. To do that, as explained in the section about *dmRI registration*, the deformation field between the moving image, being the first volume of functional images, and the static image, being the anatomical image, has to be computed. As all volumes are corrected based on the first one, the deformation field is the same between all volumes and the anatomical image. Therefore, only one computation of the deformation field is necessary. The deformation field is then applied on each volume of the functional image to obtain co-registered functional images.

4.3.2 fMRI registration

As for the dmRI registration, one needs to register the atlas on the functional images to delimit the regions of interest. To do so, the deformation field between a T1 image in the MNI space, downloaded from the site [77], and the T1 image of the participants is computed. This deformation field is then applied to the atlas so that the registration of the atlas on the T1 images in the participant's space is obtained. Then, as the functional images have been co-registered on the T1 images, an overlap between the co-registered functional images and the atlas registered on the T1 images has been performed. Contrary to the T1 images, the cerebellum was cut on some fMRI. Therefore, the regions of the atlas registered on the T1 images but not present on the fMRI were removed.

As for the diffusion images, if one region was not registered on the fMRI of one participant, this region was removed for further analysis on all the participants. A list containing non-registered regions for the functional images can be seen in Appendix D.

Moreover, the images of one vertigo participant who has compensated were cut in the upper part because the participant has been placed too high in the MRI scanner. This participant was therefore removed from the participants' set of functional analysis.

For further analyses on functional data, 256 regions are thus used out of the 280 and the dataset is composed of 7 control participants, 12 participants who have experienced vertigo but have compensated and 53 participants who still experience vertigo.

4.3.3 Functional connectivity

As stated in the [Connectivity](#) section, functional connectivity measures the dependence in activity over time between brain regions. To do so, several measures can be used. However, because of the poor temporal resolution of fMRI, the range of possible metrics is constrained and the most commonly used is the Pearson correlation which measures the linear dependency between time series. Based on many articles which used the Pearson correlation to build the functional connectome, such as the article written by *Leonardo Gutiérrez-Gómez et al.* [65] who tried to identify a biomarker in schizophrenia by analysing the human functional connectome, this metric has been chosen for this thesis although other measures, such as the mutual information, the coherence or the partial correlation are also used in the literature [18]. The generation of the functional matrices was performed using the Nilearn tools [78] which propose statistical and machine learning methods for neuroimaging data.

Finally, as the metric chosen is the Pearson correlation, the entries of the functional matrix vary between -1 and 1. Based on the work of *Leonardo Gutiérrez-Gómez et al.* and other researchers, the values are taken in absolute value [65].

4.4 Statistical analysis of connectivity measures

The goal of this section is to statistically analyse the connectomes obtained previously based on network measures: degree centrality and betweenness centrality.

As it is commonly done with functional connectivity matrices computed with the Pearson's coefficient in literature [79, 80, 81], the functional matrix of each participant was normalised with the Fisher z-transform. The Fisher z-transform is a way to transform the sampling distribution of Pearson's correlation coefficient so that it becomes normally distributed. Fisher z' is used to test the significance of the difference between two correlation coefficients from independent samples. Using this z-score seems therefore appropriate as statistical tests are performed to assess the significance of the difference between measures computed based on the correlation coefficients of participants belonging to different categories [82].

Two network measures are calculated and compared between the three classes. This is done for structural and functional networks. The centrality degree and the centrality betweenness are calculated for each node of the network of each participant. For the centrality degree, the formula used is:

$$c_{degree}(\nu) = \frac{n_{c,\nu}}{n - 1}, \quad (4.1)$$

where $n_{c,\nu}$ is the number of edges connected to the node ν and n is the total number of edges in the network. It thus represents the fraction of nodes to which node ν is connected [83]. The centrality betweenness is computed using the following

formula:

$$c_{betweenness}(\nu) = \sum_{s,t \in V} \frac{\sigma(s,t|\nu)}{\sigma(s,t)}, \quad (4.2)$$

where V is the set of nodes, $\sigma(s,t)$ is the number of shortest (s,t)-paths, and $\sigma(s,t|\nu)$ is the number of those paths passing through some node ν other than (s,t). It thus represents the sum of the fraction of all-pairs shortest paths that pass through node ν [84]. Then, statistical tests are performed to observe whether some nodes present significant differences between the three classes.

To perform statistical tests between three classes, one needs to use ANOVA tests. ANOVA uses F-tests to statistically test for equality of means. To apply the ANOVA test, several hypotheses must be verified. The first one is the independence of observations. Each participant must belong to only one group and repeated measurements for the same participants are not allowed. This hypothesis is verified. The second hypothesis is the normality. Each data should be approximately normally distributed. This is checked by a Shapiro test. This test is particularly powerful for small groups which is the case in this work. It allows to get a p-value. If this p-value is lower than a defined threshold, here 0.05, the normality condition is not verified [85]. The last hypothesis is the homogeneity of variances. The variance of each variable must be equal [86]. This condition is checked with the Bartlett test. This test compares the variances of two or more groups. Data must be normally distributed. This test also allows to compute p-values. As for the Shapiro test, if this p-value is lower than a defined threshold, here 0.05, the homogeneity of variances condition is not verified [87]. If these conditions are not met, Kruskal-Wallis tests are applied. This is a non-parametric alternative. A summary of the methodology of statistical tests can be seen in Figure 4.6.

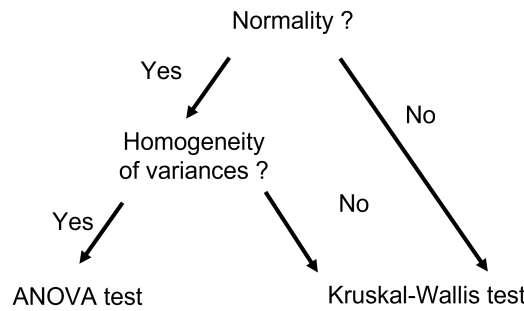


Figure 4.6: Methodology used for the statistical tests of the network measures. This methodology aims at verifying the hypothesis of the ANOVA test to know which test can be applied.

Because multiple statistical tests are performed, some fraction of the resulted p-values can be false positives. This false discovery rate is controlled by the Benjamini–Hochberg procedure or by the Bonferroni correction.

The Bonferroni correction consists in defining a new significance level based on

the number of tests performed. The classical value of the significance threshold is $\alpha = 0.05$. The new significance threshold with the Bonferroni correction becomes

$$\alpha_{Bonferroni} = \frac{0.05}{\text{Number of tests performed}}. \quad (4.3)$$

The values lower than this threshold are considered as significant.

The Benjamini-Hochberg technique consists in ranking the p-values in increasing order and attribute a number to each p-value. The smallest p-value taking the value $i=1$, the second, the value $i=2$, *etc.*. Then, the Benjamini-Hochberg critical value is computed by using the formula:

$$\alpha_{Benjamini} = \frac{i}{m}Q, \quad (4.4)$$

where i is the rank of the p-value, m is the number of tests performed and Q is the false discovery rate. This value is put to 0.1. Indeed, as written by John H. McDonald, the false discovery rate Q should take a fairly high value (0.1 or 0.2) when the cost of having a false negative, i.e., missing potentially important discovery, is high and the cost of performing additional experiments is low. The cost of performing additional experiments in this case can not be considered low as acquiring MRI images for a patient in a hospital is quite expensive. However, it was decided to perform the statistical test with a Q value equal to 0.05 only if features were significant with the Q value of 0.1 since 0.05 is more restrictive than 0.1. The largest p-value P such that $P < \frac{i}{m}Q$ is significant and all the other p-values that are smaller are also significant even if they are higher than their Benjamini-Hochberg critical values [88].

4.5 Machine learning and biomarker identification

The next analysis is the application of a feature selection procedure on brain connectivity data to identify biomarkers that are relevant to discriminate between control and vertigo participants. A biomarker in this case represents a structural or functional link between two regions of interest in the brain.

Once the functional and structural connectivity matrices are built, the table with all the features has to be created. One feature represents the corrected streamlines density value, in the structural case, or the Pearson correlation of neuronal activity, in the functional case, between two regions. So, for example, an entry in the structural connectivity matrix between the *Left A8m* region and the *Right A8m* region constitutes a feature called "corrected density value of Left_A8mToRight_A8m", shortened to "*Left_A8mToRight_A8m*", and its value is the value of the entry in the matrix. As the connectivity matrices are symmetrical (the density of fibers between regions A and B is the same than the one between regions B and A, and the same happens for neuronal activity), only half of the matrix is necessary. Moreover, the diagonal representing the entries of a region with itself is not used.

The size of the structural connectivity matrix being (249x249), 30876 structural features are thus available. On the other hand, the functional connectivity matrix has a size of (256x256) creating 32640 functional features. As each patient has its own structural and functional connectivity matrices, the final table of structural features is a (70, 30876) table while the one of functional features has a size of (72, 32640).

As can be observed from the size of the features tables, the number of available samples (i.e., the number of participants in this case) is much smaller than the dimension of the features space, increasing the risk of unexpected and poor behaviours of the machine learning algorithms, as explained in the [Machine Learning](#) section. The dimension of features must thus be reduced as much as possible to maximise the chance of having good results with machine learning methods. Feature selection can therefore be used. But before applying feature selection methods, it has first been decided to remove each feature that has a value of 0 for more than 60% of the participants in the three categories. Indeed, it has been shown that probabilistic algorithms produce noisy connection matrices and are thus likely to contain false positives [45]. The more a feature contains null values, the greater the probability that the non-null values obtained for this feature are false positives. Therefore, if more than 60% (set arbitrarily) of the values of a feature are null for the three classes, this feature is removed.

4.5.1 Filter feature selection methods

Two common filter methods used in machine learning for feature selection are the correlation and the mutual information (MI). The correlation only measures linear relation between two features while the mutual information measures any relation between two features. Features having a high correlation or mutual information with the target are expected to be more relevant to discriminate between the classes. These two filter methods have been applied on the data set.

First, the correlation has been computed between each feature and the target. Here, the target corresponds to the class to which the participant belongs. Numbers have been attributed to each class: 0 for control participants, 1 for vertigo participants and 2 for vertigo participants who have compensated. The correlation takes values between -1 and 1, 1 (resp. -1) meaning that the two features are positively (resp. negatively) correlated, i.e. if values of one feature increase, values of the other one also increase (resp. decrease). To reduce the features tables in such a way that the dimension and the computational time become acceptable, features having a correlation above 0.3 with the target have been kept.

Moreover, the correlation between the features has also been checked to remove possible redundant features, i.e. features that give the same information. If two features have an absolute correlation higher than 0.8, the one having the lowest absolute correlation with the target is removed.

A correlation equal to 0 means that the variables are not linearly dependent. They can nevertheless be non-linearly correlated and therefore, the mutual information can be useful to detect non-linear dependency between variables. Contrary to the correlation, if the mutual information between a feature and the target is 0, it is sure that this feature does not contribute to the target.

Unlike the correlation, the values of mutual information are not bounded and it is thus difficult to determine a threshold below which features are removed. Therefore, it has been decided to keep the same amount of features than the one kept after having performed the correlation filter and the redundancy check. The features with the higher mutual information are kept.

After performing the filter methods, the age of the participants has been added to the table of features kept. Indeed, as can be seen in Table 4.1, there is a difference between the means of the participants' age for the three classes. However, researches have shown that ageing has an impact on the vestibular processing, older patients exhibiting lower activation of visual and somatosensory regions and a poorer postural control [89]. Therefore, differences in brain's regions or activation between participants could be due to difference in age leading to a wrong interpretation of some features selected by the machine learning algorithm. To avoid having features selected due to the influence of age, it has been added as a feature such that features selected by the algorithm are not affected by age [90].

Table 4.1: Mean age of participants for the three classes (control participants, vertigo participants, vertigo participants who have compensated).

	Control	Vertigo	Compensated vertigo
Mean age (years)	45 $\pm$ 18	61 $\pm$ 16	54 $\pm$ 16

4.5.2 Wrapper feature selection method

Contrary to the filter methods, the wrapper methods use the model to determine whether the subset of features selected is a subset that allows to discriminate between the different classes. In this thesis, a forward sequential feature selection has been performed. It starts with an empty subset, S , and consists in adding one feature at each step. To choose which feature to add, all possible subsets of $S \cup \{x_i\}$ are evaluated, where x_i is a feature of the original set of features that is not in S [91]. The assessment of the best subset is based on a cross-validation score. This means that the model is built on each feature or subset of features to determine which one gives the best result. Therefore, the smaller the dimension of input features, the smaller the computational time. The features at the output of the filter methods (correlation and mutual information) are thus used as input to the sequential feature selection.

The two models used in this wrapper feature selection are the support vector machine (SVM) with a linear kernel and the random forest (RF), described in the *Machine Learning* section.

As a reminder, the goal of SVM is to linearly separate the data by a hyper-plane. The weights are calculated in order to separate the data with a larger margin with respect to this hyper-plane. As there are three classes in this thesis (class 1 = control participants, class 2 = vertigo participants and class 3 = vertigo participants who have compensated), three different weights are assigned to each feature by the classifier. The first weight assigned to a given feature gives the importance of that feature to differentiate class 1 from class 2 and 3 together; the second weight gives the importance of the feature to differentiate class 2 from class 1 and 3 together, while the third weight is the one giving the importance of the feature to differentiate class 3 from class 1 and 2 together. In this thesis, there is no need to evaluate how features can differentiate vertigo participants who have compensated from control and vertigo participants together as the goal is to be able to differentiate vertigo from control participants. Therefore, the third weight is useless for the analyses. The higher the weight of a feature, the more it is supposed to be relevant to discriminate between classes.

As explained in the [Machine Learning](#) section, the random forest model is the implementation of multiple decision trees. A decision tree splits data based on the information gain which gives the information whether the data still need to be split or not. In opposite to the SVM, the random forest model assigns only one weight to each feature and therefore, does not allow to know whether the feature best discriminates between class 1 and classes 2 and 3, class 2 and classes 1 and 3 or class 3 and classes 1 and 2.

Before applying the wrapper method, the hyper parameters of each model giving the best performances should first be found. The C parameter of the SVM, used in equation 2.1, is an important parameter as it allows to find a compromise between maximising the margin width with respect to the hyper-plane while minimising the number of misclassifications. It has thus been tested. The hyper-parameter tested for the random forest is the number of trees. To determine the best hyper-parameter, a grid search is performed. During this process, several values of hyper-parameters are tested. For each set of parameters, a cross-validation is performed and at the end of the process, the set of parameters giving the highest score is returned. This score is based on the balanced accuracy which takes into account the fact that the data set is imbalanced (there are much more participants in the vertigo class than in the two others).

Once the best hyper-parameters are found, the sequential feature selection can be performed. To do this, a cross-validation of 5-folds is also done in such a way that the subset of features returned at the end of the process is the one giving the best performances on most of the validation sets. This procedure is repeated 35 times allowing to keep a reasonable computational time, in such a way that the stability can be computed. Indeed, the subset of features returned by the method depends on the training/testing split of the data. If the subset returned among the 35 iterations is always the same, the features contained in the subsets are more

likely to be consistent. On the other hand, if subsets are different between each iteration, features are more likely to have been returned by chance and are not relevant to discriminate between classes. One challenge in feature selection is thus to try to have reproducible results. The stability index gives an indication on how similar subsets of features chosen from the sequential feature selection model are and is defined as

$$I = \frac{2}{k(k-1)} \sum_{i=1}^{k-1} \sum_{j=i+1}^k KI(f_i, f_j), \quad (4.5)$$

where k is the number of repetitions of the sequential feature selection method, i.e., the number of subsets available, f_i and f_j , two subsets of features and KI , the consistency index of the two subsets [91]. This consistency index is defined as

$$KI(f_i, f_j) = \frac{rN - s^2}{s(N - s)}, \quad (4.6)$$

where r represents the number of features present in both feature sets f_i and f_j , N is the total number of features given as input to the method and s is the size of both feature sets, $|f_i| = |f_j| = s$ [91]. It satisfies the following properties:

- The larger the intersection between the two subsets, the higher the value of the consistency index.
- It is bounded between -1 and 1 and reaches a maximum (resp. minimum) value when the two subsets are identical (resp. perfectly disjoint).
- It corrects for chance: consistency value is high only if it exceeds the consistency by chance.

The total index of stability I is thus the average of all pairwise consistency indices. If this index is high, the way of determining the importance of features to discriminate between classes is based on the mean of the weights attributed to the features by the model (SVM or RF) across the k iterations of the sequential feature selection. On the other hand, if the stability is low, the rank is based on the number of times the feature has been returned by the method [91].

Finally, the last step consists in computing the accuracy of the models on the different subsets selected by the sequential feature selection. To avoid over-fitting as much as possible, a cross-validation is performed with stratified 5-folds. It is a variation of k -fold that returns folds preserving the percentage of samples of each class in the split. Ideally, this accuracy should have been computed on a set of data different from the one used to determine the best hyper-parameters. However, since a cross-validation is performed in both steps and because of the small number of samples available, especially for the control class, splitting the initial set in two and then performing different folds would have resulted in too small subsets with the risk, during the cross-validation, of having no sample left to represent the control class.

A summary scheme of the methodology used for the analysis of structural and functional data, from raw data to features selected by statistical tests and machine learning algorithms, can be seen in Figure 4.7.

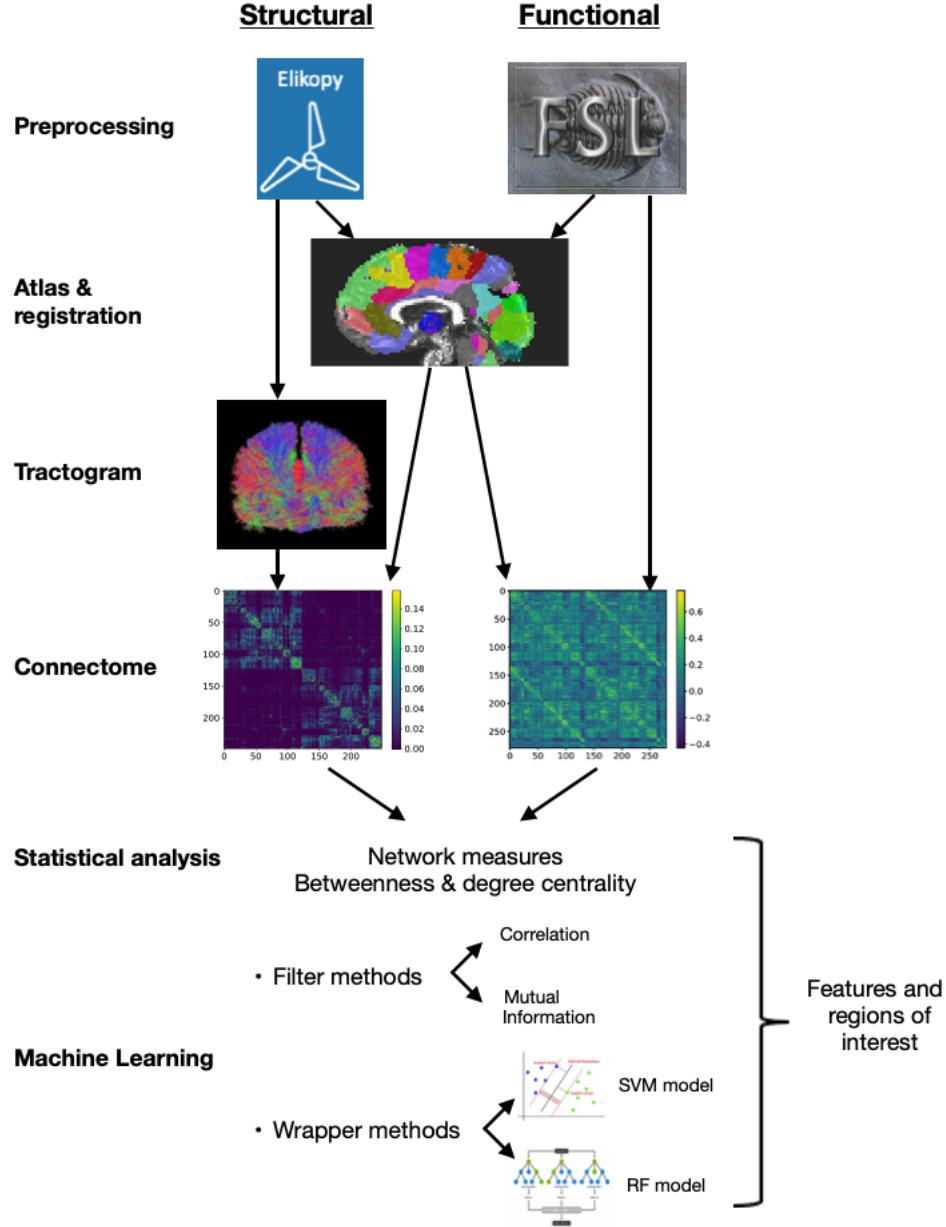


Figure 4.7: Summary scheme of the methodology.

4.5.3 Embedded feature selection method

Based on what was done in the article [65], embedded method has been applied to the data set. More specifically, recursive feature elimination has been applied with an SVM and an RF model. This method is an iterative algorithm using a ranking criterion for eliminating the features backwards. For example, the SVM

model is built on the training set and the features are ranked according to the weights assigned by the algorithm. The least important features are removed and the remaining ones are used to train a new model. This process is repeated until the desired number of features is reached. Different numbers of features kept were tested as well as different step ranges, the step range being the number of features removed at each step. However, this method has been abandoned because over-fitting was detected.

4.6 Statistical cerebellum analysis

As explained previously, the cerebellum plays an important role in the vestibular network. This section aims at looking at some characteristics of the cerebellum and comparing them between the three classes. For this part, the T1-weighted images are used. These are put on the VolBrain platform, which is an automated MRI brain volumetry system [92]. The VolBrain pipeline first performs some preprocessing steps to set the input images in the same geometric and intensity space as the one of the VolBrain library. Then, estimation of volumes at different scales is performed. The algorithm chosen is the CERES which allows to segment the cerebellum. The platform provides a report containing volumetric information for each participant. The platform also returns *pdf* and *csv* files summarising the volume, cortical thickness and asymmetry ratios estimated from the data. The asymmetry is defined as the difference between the right and left volumes divided by their mean (in percent). An example of report and data returned by VolBrain can be seen in Appendix E.

The cerebellum was divided in 12 parts: Lobule I-II, Lobule III, Lobule IV, Lobule V, Lobule VI, Lobule Crus I, Lobule Crus II, Lobule VIIIB, Lobule VIIIA, Lobule VIIIB, Lobule IX, Lobule X. For each of these regions, 6 variables are analysed: the total volume, the volume asymmetry, the mean cortical thickness, the cortical thickness asymmetry, the grey matter volume and the grey matter volume asymmetry.

Statistical tests are used to discriminate between the three classes based on these six variables. The methodology for the tests is the same as the one described in the section *Statistical analysis of connectivity measures*.

The tests are realised separately for each variable on the 12 lobules and the correction for multiple tests used is the same as in section *Statistical analysis of connectivity measures*.

The variables corresponding to the significant p-values allow to statistically discriminate between control participants, vertigo participants and vertigo participants who have compensated. However, it is not known whether the variable is significant thanks to a difference between control and vertigo participants, between control and vertigo participants who have compensated or between vertigo participants who have and have not compensated. A second statistical test has therefore been performed between each pair of classes to see which classes are the most different.

Chapter 5

Results

This fifth chapter covers the results obtained after applying the methodology on the data, including the functional and structural connectomes, tables containing the results of the statistical tests on the connectomes for the betweenness and degree centrality and on the cerebellum, as well as graphs illustrating the performances and the significant features obtained from the machine learning methods.

5.1 Connectome

Connectomes obtained for structural and functional data are displayed.

5.1.1 Structural

For structural connectomes, tractograms had first to be built.

Tractogram

The tractograms, generated from (a) the deterministic algorithm and (b) the probabilistic algorithm, between two regions of the right hemisphere can be seen in Figure 5.1. In orange is the medial part of area 6 of the superior frontal gyrus (A6m) while in purple is the dorsolateral putamen of the basal ganglia (dlPu). As expected, more fibers are created with the probabilistic algorithm (933 fibers) compared to the deterministic one (22 fibers).

Connectivity matrix

The structural connectivity matrix obtained for one participant who has compensated for vertigo is shown in the left of Figure 5.2. As a reminder, the upper left square represents the connections between regions located in the left hemisphere, the lower right square represents the connections between regions located in the right hemisphere and the last two squares, i.e., the upper right and lower left squares, represent inter-hemispheric connections. The matrix represents the structural connectome corrected for the mean length of the tracts and the volume of the regions.

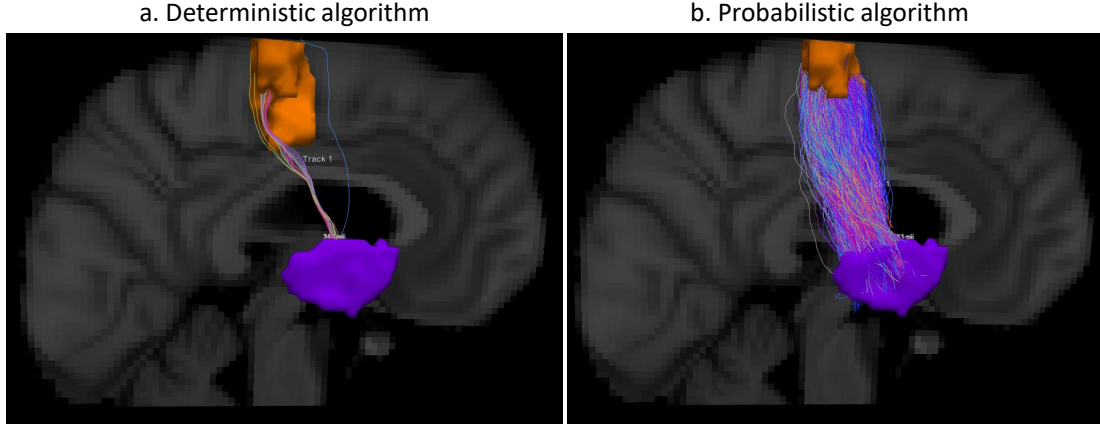


Figure 5.1: Fibers between two regions (Right A6m in orange and Right dlPu in purple) in tractograms generated from (a) the deterministic algorithm and (b) the probabilistic algorithm.

5.1.2 Functional

The functional connectivity matrix obtained for a control participant can be seen in the right of Figure 5.2. As for the structural connectivity matrix, the upper left square represents the connections between regions located in the left hemisphere, the lower right square represents the connections between regions located in the right hemisphere and the last two squares, i.e., the upper right and lower left squares, represent inter-hemispheric connections.

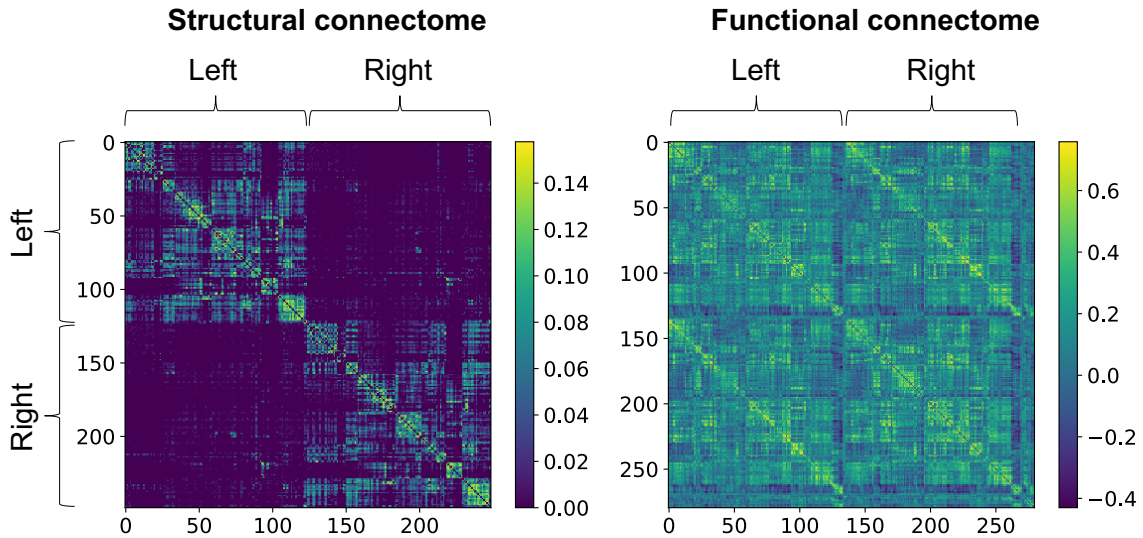


Figure 5.2: **Left:** Structural connectivity matrix obtained for a vertigo participant who has compensated after correction by the mean length of the tracts and the volumes of the regions. **Right:** Functional connectivity matrix obtained for a control participant.

5.2 Network measures

5.2.1 Structural

Based on the corrected structural connectivity matrices obtained, network measures are calculated and statistical analysis is performed on these measures. The statistical analysis concerning the betweenness centrality (Eq. 4.2) for left and right hemispheres can respectively be found in Tables 5.1 and 5.2. It can be seen that none of these regions show a significant p-value when comparing with the threshold corrected with the Benjamini (Eq. 4.4) or the Bonferroni (Eq. 4.3) method. The displayed values are therefore the ones which are significant with respect to the non-corrected threshold, i.e., 0.05.

Table 5.1: Statistical values for the betweenness centrality in the regions located in the left hemisphere (structural features).

	p-value	Benjamini correction	Bonferroni correction
Left_cpSTS	0.0108	0.0008	0.0004
Left_A41/42	0.0163	0.0016	
Left_A9/46v	0.0190	0.0024	
Left_A20rv	0.023	0.0032	
Left_A6cvl	0.0373	0.004	
Left_vId/vIg	0.0378	0.0048	
Left_TE1.0/TE1.2	0.04	0.0056	
Left_A39rv	0.0438	0.0065	
Left_A6dl	0.0479	0.0073	

For the left hemisphere regions listed in Table 5.1, four regions belong to the temporal lobe: *Left_cpSTS* in the posterior superior temporal sulcus, *Left_A41/42* and *Left_TE1.0/TE1.2* in the superior temporal gyrus and *Left_A20rv* in the fusiform gyrus. Three regions are located in the frontal lobe: *Left_A9/46v* in the middle frontal gyrus, *Left_A6cvl* in the precentral gyrus and *Left_A6dl* in the superior frontal gyrus. *Left_vId/vIg* is in the insular gyrus of the insular lobe. *Left_A39rv* is in the parietal lobe, more precisely in the inferior parietal lobule.

For the right hemisphere regions listed in Table 5.2, three regions belong to the subcortical nuclei: *Right_rHipp* is in the hippocampus, *Right_Stha* is in the thalamus and *Right_lAmyg* is in the amygdala. Two regions are in the frontal lobe: *Right_A44op* is in the inferior frontal gyrus and *Right_A12/47l* is in the orbital gyrus. *Right_I_IV* represents the four first lobules of the cerebellum. *Right_A5l* is in the parietal lobe, in the superior parietal lobule. *Right_A23d* is in the cingulate gyrus of the limbic lobe.

Table 5.2: Statistical values for the betweenness centrality in the regions located in the right hemisphere (structural features).

	p-value	Benjamini correction	Bonferroni correction
Right_A5l	0.015	0.0008	0.0004
Right_rHipp	0.029	0.0016	
Right_A23d	0.0325	0.0024	
Right_A44op	0.04	0.0032	
Right_I_IV	0.0436	0.004	
Right_Stha	0.0469	0.0048	
Right_lAmyg	0.0488	0.0056	
Right_A12/47l	0.049	0.0064	

The statistical analysis concerning the degree centrality (Eq. 4.1) for the nodes located in the left and in the right hemispheres can respectively be seen in Tables 5.3 and 5.4.

Table 5.3: Statistical values for the degree centrality in the regions located in the left hemisphere (structural features).

	p-value	Benjamini correction	Bonferroni correction
Left_A20rv	0.0172	0.0008	0.0004
Left_A21c	0.0247	0.0016	
Left_msOccG	0.0259	0.0024	
Left_A9/46d	0.0263	0.0032	
Left_A9/46v	0.0292	0.0040	
Left_A5m	0.0317	0.0048	
Left_A21r	0.0327	0.0056	
Left_A38m	0.034	0.0065	
Left_A20cl	0.037	0.0073	
Left_rCunG	0.0375	0.0081	
Left_A20il	0.0379	0.0089	
Left_A45c	0.039	0.0097	
Left_vIa	0.0419	0.0105	
Left_cTtha	0.0432	0.0113	
Left_A7pc	0.0455	0.0121	
Left_A37vl	0.0468	0.0129	
Left_A1/2/3tru	0.0477	0.0137	
Left_dmPOS	0.049	0.0145	

For the left hemisphere regions listed in Table 5.3, seven regions are in the temporal lobe: *Left_A20rv* is in the fusiform gyrus, *Left_A21c* and *Left_A21r* are in

the middle temporal gyrus, *Left_A38m* is in the superior temporal gyrus and *Left_A20cl*, *Left_A20il* and *Left_A37vl* are in the inferior temporal gyrus. Four regions are in the parietal lobe: *Left_A5m* and *Left_dmPOS* are in the precuneus, *Left_A7pc* is in the superior parietal lobe and *Left_A1/2/3tru* is in the postcentral gyrus. Two regions, *Left_A9/46d* and *Left_A9/46v* are in the middle frontal gyrus of the frontal lobe and one, *Left_A45c*, is in the inferior frontal gyrus. Two regions are in the occipital lobe: *Left_msOccG* is in the lateral occipital cortex and *Left_rCunG* is in the MedioVentral occipital lobe. The region *Left_vIa* is in the insular gyrus of the insular lobe. The regions *Left_cTha* belongs to the subcortical nuclei, more precisely to the thalamus.

Table 5.4: Statistical values for the degree centrality in the regions located in the right hemisphere (structural features).

	p-value	Benjamini correction	Bonferroni correction
Right_A40rd	0.0086	0.0008	0.0004
Right_A14m	0.0125	0.0016	
Right_A37elv	0.0203	0.0024	
Right_A24rv	0.0231	0.0032	
Right_A31	0.029	0.0040	
Right_A5l	0.0295	0.0048	
Right_A7pc	0.031	0.0056	
Right_A5m	0.0323	0.0064	
Right_A8vl	0.0337	0.0072	
Right_Otha	0.0395	0.0080	
Right_A9l	0.0427	0.0088	
Right_A44d	0.0433	0.0096	
Right_vCa	0.0434	0.0104	
Right_NAC	0.0446	0.0112	
Right_A7ip	0.0465	0.0120	
Right_A32sg	0.047	0.0128	
Right_A10m	0.048	0.0136	

For the right hemisphere regions listed in Table 5.4, six regions are in the parietal lobe: *Right_A40rd* is in the inferior parietal lobule, *Right_A31* and *Right_A5m* are in the precuneus and *Right_A5l*, *Right_A7pc* and *Right_A7ip* are in the superior parietal lobule. Five regions are in the frontal lobe: *Right_A9l* and *Right_A10m* are in the superior frontal gyrus, *Right_A14m* is in the orbital gyrus, *Right_A8vl* is in the middle frontal gyrus and *Right_A44d* is in the inferior frontal gyrus. Three regions belong to the subcortical nuclei: *Right_vCa* and *Right_NAC* are part of the basal ganglia and *Right_Otha* is part of the thalamus. Two regions, *Right_A24rv* and *Right_A32sg* belong to the cingulate gyrus of the limbic lobe. The region *Right_A37elv* is located in the temporal lobe, more precisely in the inferior temporal gyrus.

5.2.2 Functional

In a similar way to what is done for structural data, network measures have been calculated based on the functional connectivity matrix. The statistical tests corresponding to the degree betweenness (Eq. 4.2) for the left and the right hemispheres can respectively be found in Tables 5.5 and 5.6. As for the structural part, any of these regions are significant with the corrected thresholds and the values displayed are therefore the ones that are significant without correction.

Table 5.5: Statistical values for the betweenness centrality in the regions located in the left hemisphere (functional features).

	p-value	Benjamini correction	Bonferroni correction
Left_A23d	0.0168	0.0008	0.00039
Left_A37dl	0.0175	0.0016	
Left_A7ip	0.0277	0.0023	
Left_A20rv	0.0293	0.0031	
Left_A23c	0.0359	0.0039	
Left_A13	0.0381	0.0047	
Left_A2	0.0463	0.0055	

For the left hemisphere, two regions are in the limbic lobe: *Left_A23d* and *Left_A23c*. Both are located in the cingulate gyrus. Two regions are in the temporal lobe: *Left_A37dl* in the middle temporal gyrus and *Left_A20rv* in the fusiform gyrus. Two regions are located in the parietal lobe: *Left_A7ip* in the superior parietal lobule and *Left_A2* in the post-central gyrus. *Left_A13* is in the frontal lobe, in the orbital gyrus.

Table 5.6: Statistical values for the betweenness centrality in the regions located in the right hemisphere (functional features).

	p-value	Benjamini correction	Bonferroni correction
Right_A5m	0.0139	0.0008	0.000394
Right_A21r	0.0144	0.0016	
Right_rHipp	0.0227	0.0024	
Right_A7ip	0.0274	0.0032	
Right_A4ul	0.0329	0.0039	
Right_A20rv	0.0358	0.0047	
Right_V5/MT+	0.0405	0.0055	
Right_A23c	0.0492	0.0062	

For the right hemisphere, two regions are located in the parietal lobe. *Right_A5m* is in the precuneus while *Right_A7ip* is in the superior parietal lobule. Two regions

are located in the temporal lobe. *Right_A21r* is in the middle temporal gyrus and *Right_A20rv* is in the fusiform gyrus. *Right_rHipp* is in the subcortical nuclei, more precisely in the hippocampus. *Right_A4ul* is in the frontal lobe, in the precentral gyrus. *Right_V5/MT+* is in the occipital lobe, in the lateral occipital cortex. *Right_A23c* is in the limbic lobe, in the cingulate gyrus.

The values of the statistical tests for the degree centrality (Eq. 4.1) for the left and right hemispheres can respectively be found in Tables 5.7 and 5.8.

Table 5.7: Statistical values for the degree centrality in the regions located in the left hemisphere (functional features).

	p-value	Benjamini correction	Bonferroni correction
Left_V	0.0274	0.0008	0.00039

The *Left_V* region represents the fifth lobule of the cerebellum.

Table 5.8: Statistical values for the degree centrality in the regions located in the right hemisphere (functional features).

	p-value	Benjamini correction	Bonferroni correction
Right_TI	0.0127	0.0008	0.000394

The *Right_TI* region is located in the temporal lobe, in the parahippocampal gyrus.

5.3 Machine learning analysis

The number of structural and functional features (including the age) kept after applying the filter methods are 117 and 53, respectively. As a reminder, the number of features kept is the same after performing the correlation or the mutual information. The structural features kept from the correlation method can be seen in Appendix F, and the ones from the mutual information, in Appendix G. The functional features kept after applying the correlation filter method are listed in Appendix H and the ones kept after the mutual information can be seen in Appendix I.

5.3.1 Structural

Pre-filtered with correlation

Once the correlation applied, the wrapper method is performed both with the Support Vector Machine (SVM) and the Random Forest (RF) models. As a reminder, the wrapper method consists in a forward sequential feature selection.

SVM

The wrapper method has first been applied with the SVM model with a C value of 0.01 (best hyper-parameter returned by the grid search).

The two left graphs in Figure 5.3 represent the balanced and non-balanced accuracies obtained for the SVM model on the train (in orange) and test (in green) sets for different percentages of structural features selected by the wrapper method with, as input, the features chosen by the pre-filter method of correlation. For each value of percentage of features kept, 35 iterations of sequential feature selection are performed and 35 subsets of features are therefore obtained. For each subset of features chosen by the sequential feature selection procedure, 5 iterations of a stratified 5-folds are performed to compute the accuracy. At each iteration of the 5-folds, a SVM model is built on the train set and the model then predicts the target for the train and test sets. The accuracy is obtained by comparing this prediction with the known value of the target on each set. One value of accuracy is therefore computed for each iteration of the stratified 5-folds which is itself performed 5 times for each subset of features. The accuracy plotted in Figure 5.3 is therefore the average of the accuracy calculated across all these iterations.

The right graph in Figure 5.3 shows the stability of the 35 subsets, corresponding to 35 iterations of the sequential feature selection procedure, generated by the wrapper method for different percentages of features selected based on the SVM model with, as input, the features chosen by the pre-filter method of correlation. One point in this graph therefore represents the stability index between 35 subsets of features. For each point, the size of the 35 subsets is different depending on the percentage of selected features.

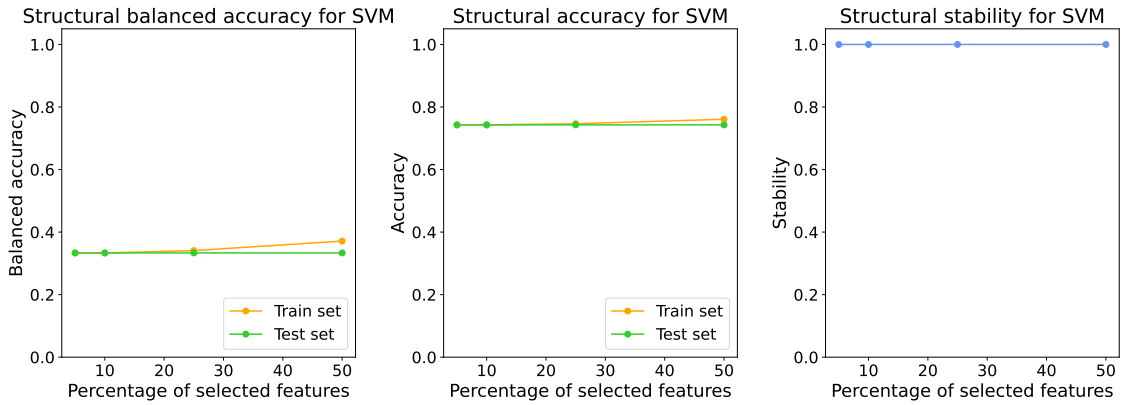


Figure 5.3: Accuracy and stability of the structural features selected from the pre-filter method of correlation and the wrapper method with the SVM model.

The results obtained by this model are not satisfactory. Indeed, about 75% of the participants in the dataset are vertigo participants. This means that when the classes are non-balanced, vertigo participants represent about 75% of the data and when the classes are balanced, they represent about 33% of the data because there

are three classes. The performances of the model are of 75% when measured with the non-balanced accuracy and 33% when measured with the balanced accuracy. A model assigning all participants to the vertigo class would show the same performance.

RF

The forward sequential feature selection has then been applied with the RF model using 10 decision trees (best hyper-parameter returned by the grid search).

The graphs in Figure 5.4 are similar to the ones in Figure 5.3 except for the model used being the RF model in this case, instead of the SVM model.

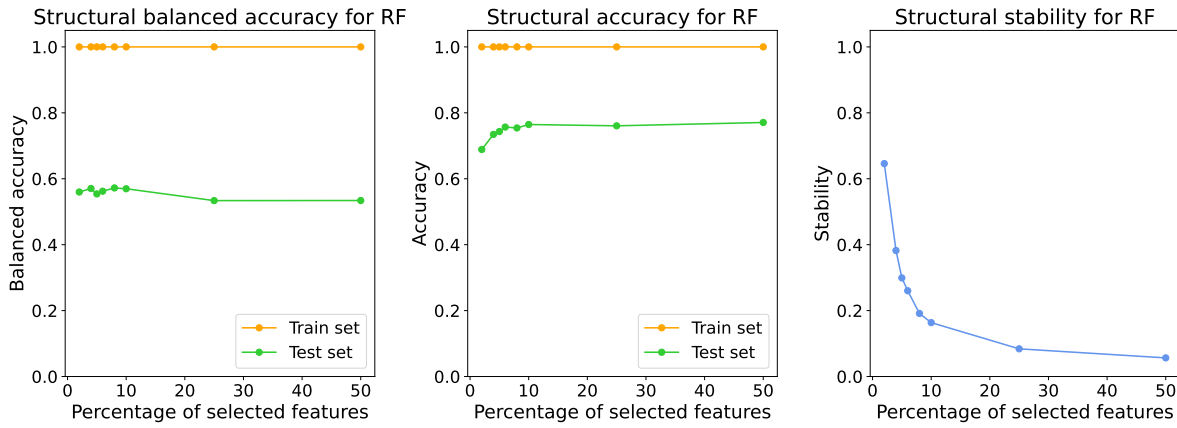


Figure 5.4: Accuracy and stability of the structural features selected from the pre-filter method of correlation and the wrapper method with the RF model.

A histogram showing the number of times a feature has been selected over the 35 iterations of the wrapper method based on the RF model when 2% of structural features are selected (i.e., 2 features at each iteration) is illustrated in Figure 5.5. The percentage of features selected analysed (2% in this case) is chosen based on the results of the accuracies and the stability. Since having reproducible results is as important as having good performances, the percentage chosen is the one showing the best trade-off between high performances and high stability. The black line represents the mean of the number of times a feature is selected. A feature above this line is a feature having a number of selections higher than the overall average. The line is present on all the following histograms and has the same meaning for the RF models.

The features for which the bar is in red are features showing a decrease of mean value for vertigo participants compared to control participants. Values are in this case the entries of the structural connectivity matrix, i.e., the weighted density of fibers between two regions. On the contrary, green features are features for which values are higher for vertigo participants compared to control participants.

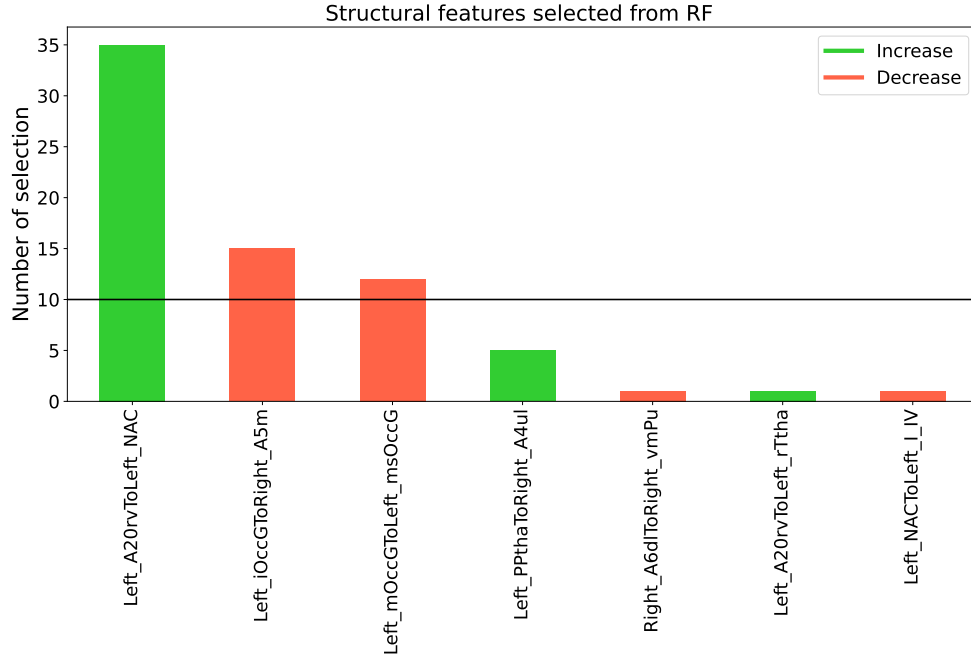


Figure 5.5: Structural features selected from the pre-filter method of correlation and the wrapper method with the RF model.

The difference in fibers density between one control participant and one vertigo participant can be seen in Figure 5.6 for the first feature selected. An augmentation in fibers density can be observed between the two regions for the vertigo participant, which confirms what is shown in the histogram of Figure 5.5.

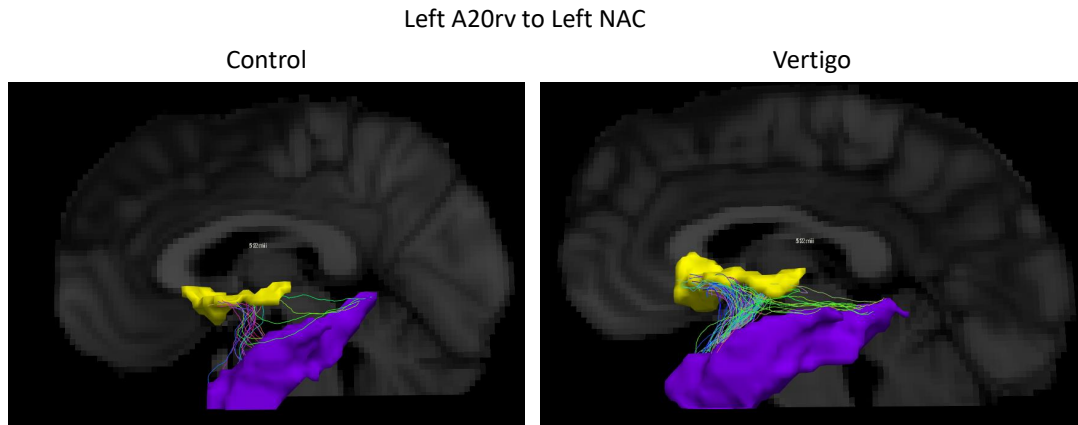


Figure 5.6: Difference of fibers density for the first feature between **Left:** one control participant and **Right:** one vertigo participant.

The two other features which have a number of selections higher than the mean are visualised on a brain in Figure 5.7 with the legend of colours in the caption.

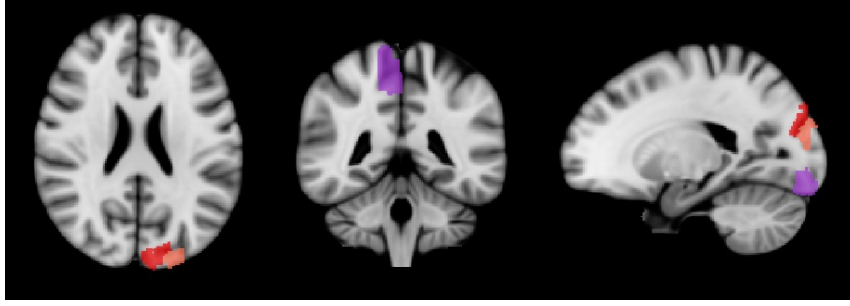


Figure 5.7: Visualisation of features selected. In purple: Left iOccG to Right A5m. In red: Left msOccG (dark) to Left mOccG (light).

Pre-filtered with mutual information

The forward sequential feature selection has then been performed on the features kept from the mutual information.

SVM

First, the wrapper method is applied using the SVM model with a C value of 0.01. The graphs obtained in Figure 5.8 are similar to those obtained in Figure 5.3, except that in this case, the features have been pre-selected with the mutual information.

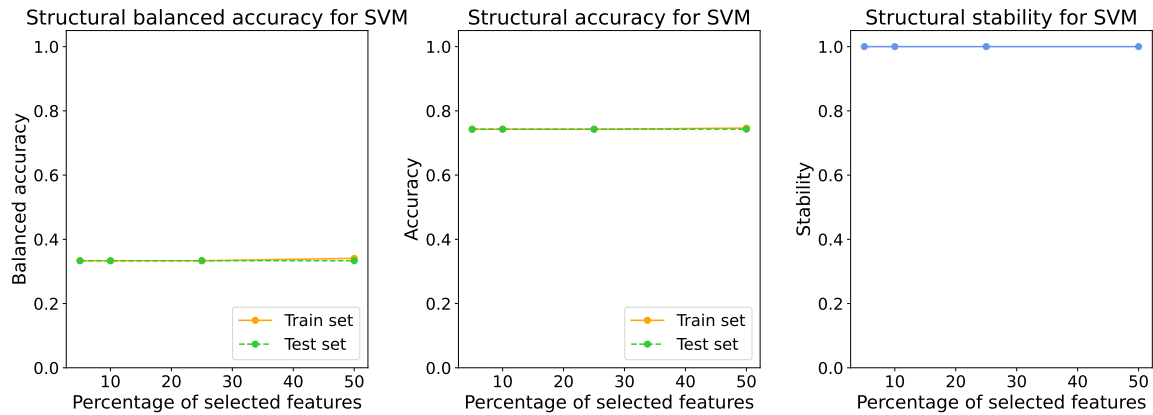


Figure 5.8: Accuracy and stability of the structural features selected from the pre-filter method of mutual information and the wrapper method with the SVM model.

As for the results of the features kept from the correlation, the results of balanced and non-balanced accuracies are equivalent to a model attributing each participant to the vertigo class. The features selected by this method are therefore not analysed.

RF

Then, the RF model with 150 decision trees has been used for the forward sequential feature selection.

Figure 5.9 is similar to the previous Figure 5.8, except that the model used for the selection of structural features in the wrapper method is, in this case, the RF model instead of the SVM model.

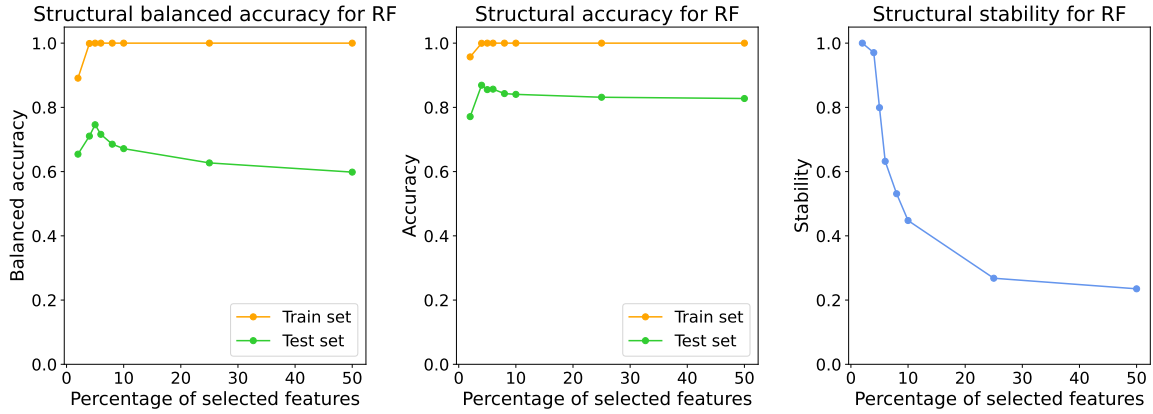


Figure 5.9: Accuracy and stability of the structural features selected from the pre-filter method of mutual information and the wrapper method with the RF model.

A histogram showing the number of selections of each feature selected by the wrapper method when 4% of structural features are selected (i.e., 4 features at each iteration) is illustrated in Figure 5.10. The input of the wrapper method are the features selected by the pre-filter method of mutual information.

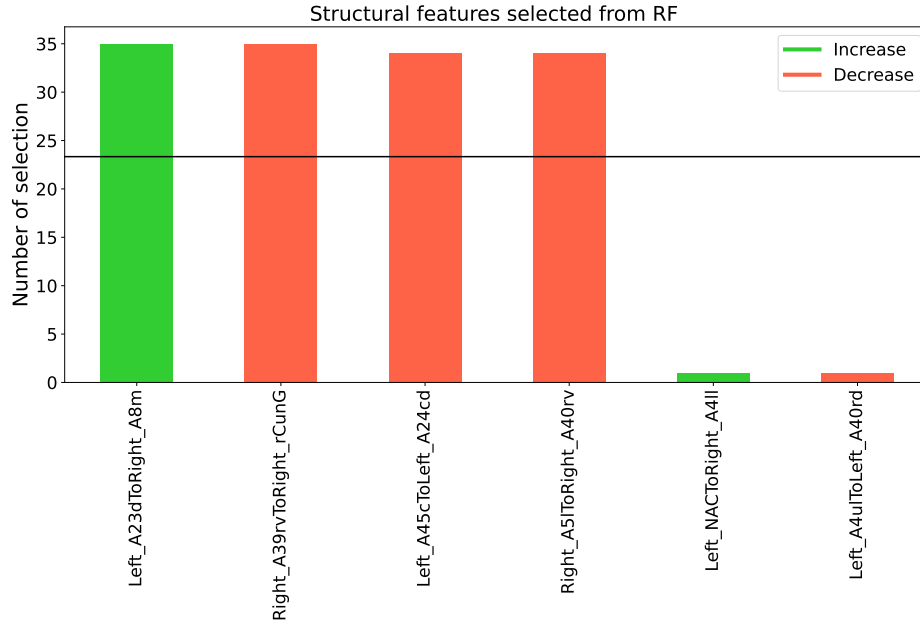


Figure 5.10: Structural features selected from the pre-filter method of mutual information and the wrapper method with the RF model.

The first four features, which have a number of selections higher than the mean, are represented in Figures 5.11 and 5.12. The left part of Figure 5.11 shows an increase in fibers density while all other images show a decrease in fibers density for vertigo participants compared to control ones.

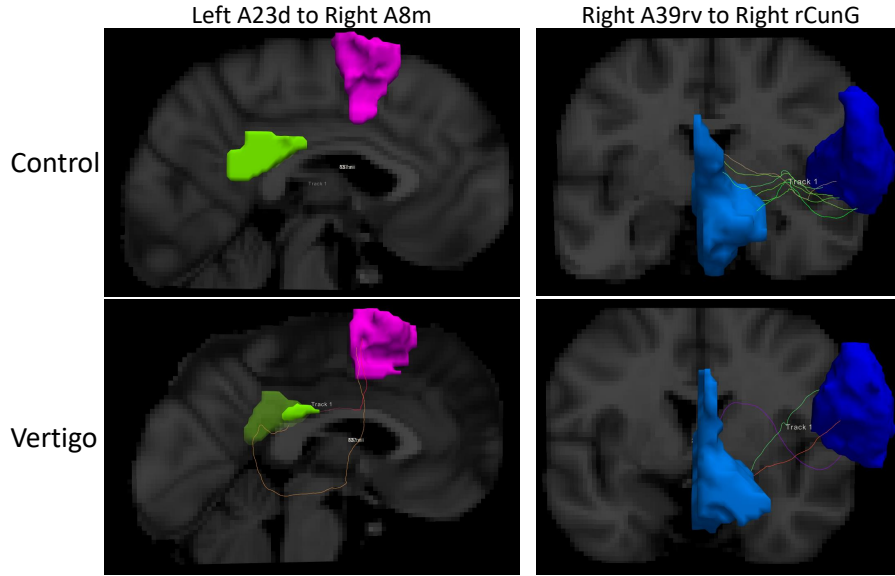


Figure 5.11: Difference of fibers density for the two first features selected by the sequential feature selection method. **Left**, first feature, **Right**, second feature. The upper image is from a control participant and the lower image is from a vertigo participant.

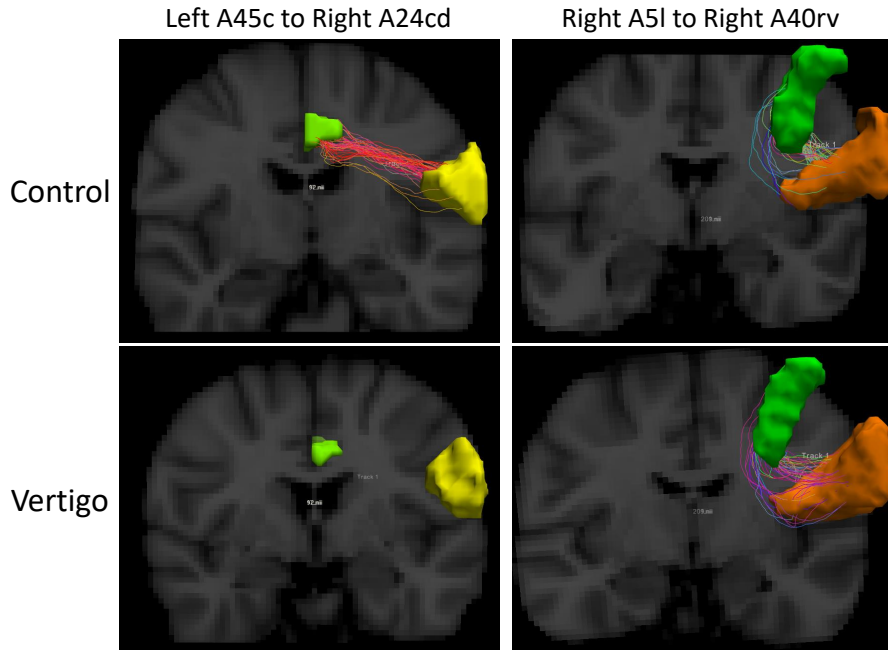


Figure 5.12: Difference of fibers density for the third and fourth features selected by the sequential feature selection method. **Left**, third feature, **Right**, fourth feature. The upper image is from a control participant and the lower image is from a vertigo participant.

5.3.2 Functional

Pre-filtered with correlation

Based on the features generated from the functional connectivity matrices, the same procedure than the one used on structural features has been applied. The forward sequential feature selection has thus first been performed with both SVM and RF models on the functional features kept from the correlation.

SVM

The wrapper method has first been applied with the SVM model with a C value of 10.

The three graphs obtained in Figure 5.13 are similar to the ones obtained in Figure 5.3 except that features selected from the wrapper method are, in this case, entries of the functional connectivity matrix.

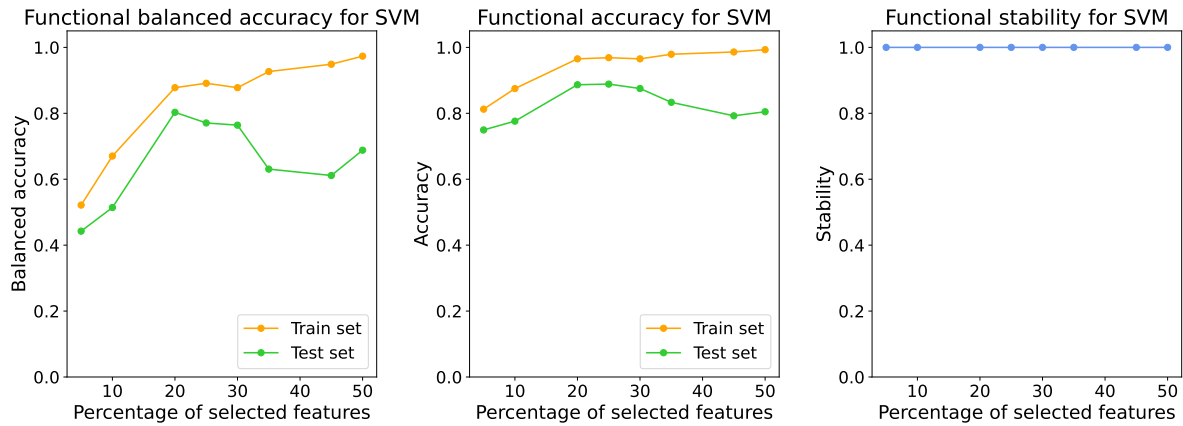


Figure 5.13: Accuracy and stability of the functional features selected from the pre-filter method of correlation and the wrapper method with the SVM model.

The histogram in Figure 5.14 shows the mean of the weights attributed to each functional feature when 20% of features are selected from the wrapper method with, as input, the features chosen by the pre-filter method of correlation. The percentage of features selected chosen (20% in this case) depends only on the accuracies obtained since the stability is always the same with the SVM model for each percentage. The weight is the sum of all weights attributed to the feature divided by the number of iterations (35) of the wrapper method. It represents the importance of the feature to discriminate control participants from vertigo participants who have compensated or not. Red (resp. green) bars correspond to functional features showing a decrease (resp. an increase) of mean value for vertigo participants who have compensated or not compared to the control participants. Mean values are in this case the entries of the functional connectivity matrix, i.e., the simultaneous activity of the brain.

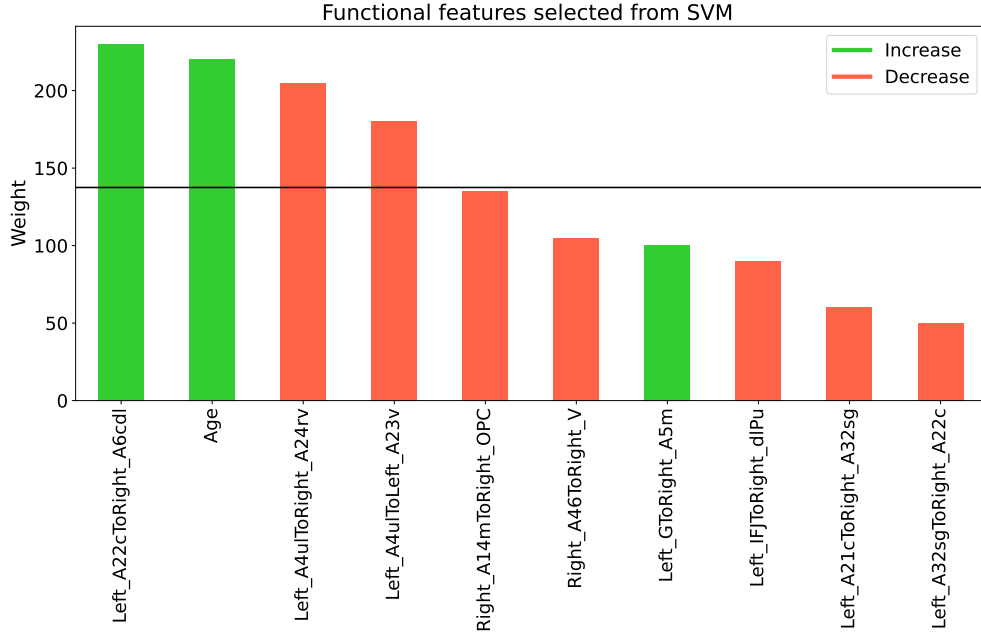


Figure 5.14: Functional features selected from the pre-filter method of correlation and the wrapper method with the SVM model to differentiate control participants from vertigo participants who have compensated or not.

The regions of the functional features present in Figure 5.14 can be visualised on a brain in Figures 5.15 and 5.16. The legend of the colours is in the caption of the figure. Two regions of same colour represent a feature selected from the method.

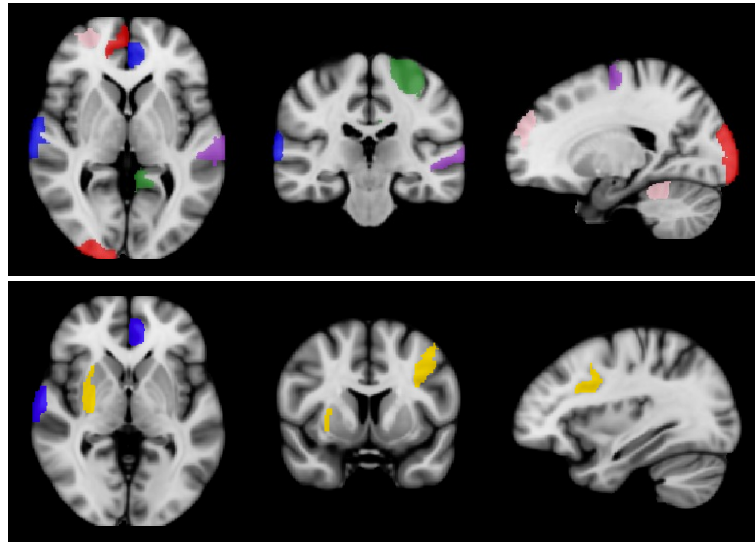


Figure 5.15: Visualisation of features selected. In purple: Left A22c to Right A6cdl. In green: Left Aul to Left A23v. In red: Right A14m to Right OPC. In pink: Right A46 to Right V. In yellow: Left IFJ to Right dIPu. In blue: Left A32sg to Right A22c.

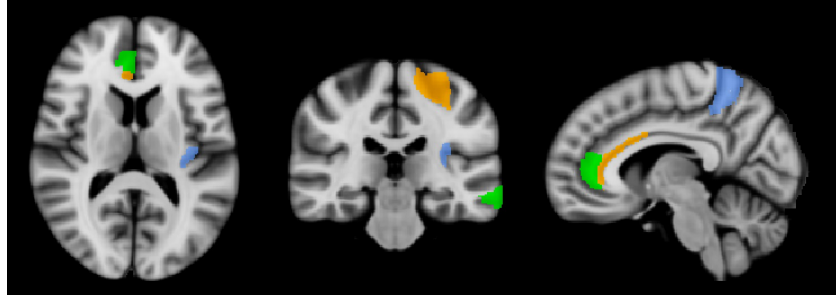


Figure 5.16: Visualisation of features selected. In orange: Left A4ul to Right A24rv. In green: Left A21c to Right A32sg. In blue: Left G to Right A5m.

The histogram in Figure 5.17 is the same as the one in Figure 5.14 except that in this case, the weight represents the importance of the feature to discriminate vertigo participants from control participants and vertigo participants who have compensated. Therefore, a red (resp. green) bar indicates that mean entries of the functional matrix for vertigo participants are lower (resp. higher) than the ones for control and compensated vertigo participants.

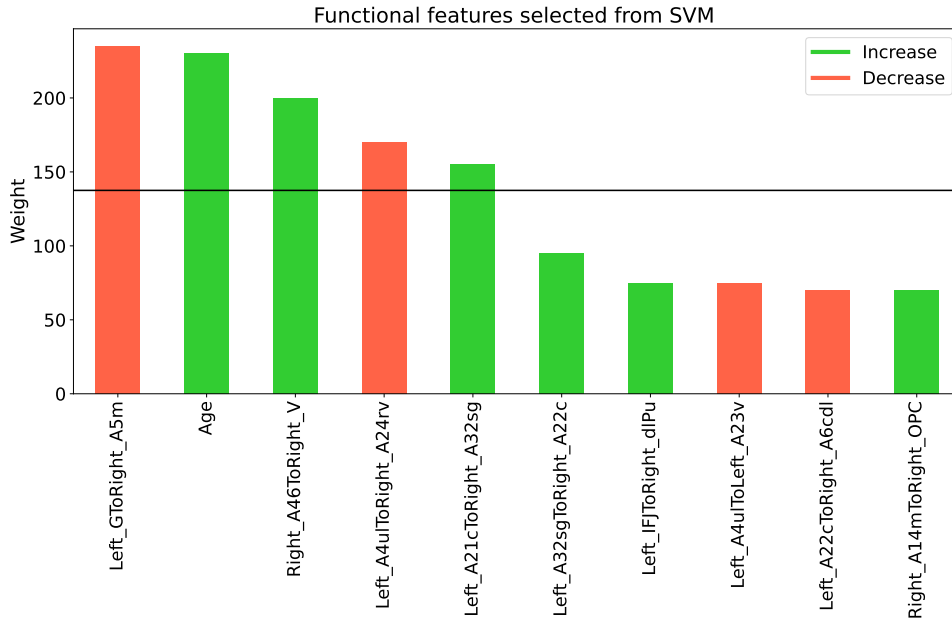


Figure 5.17: Functional features selected from the pre-filter method of correlation and the wrapper method with the SVM model to differentiate vertigo participants from control participants and vertigo participants who have compensated.

The regions of the functional features present in the histogram of Figure 5.17 can be visualised on a brain in Figure 5.18 (with the legend of the colours in the caption). Two regions of same colour represent a feature selected from the method.

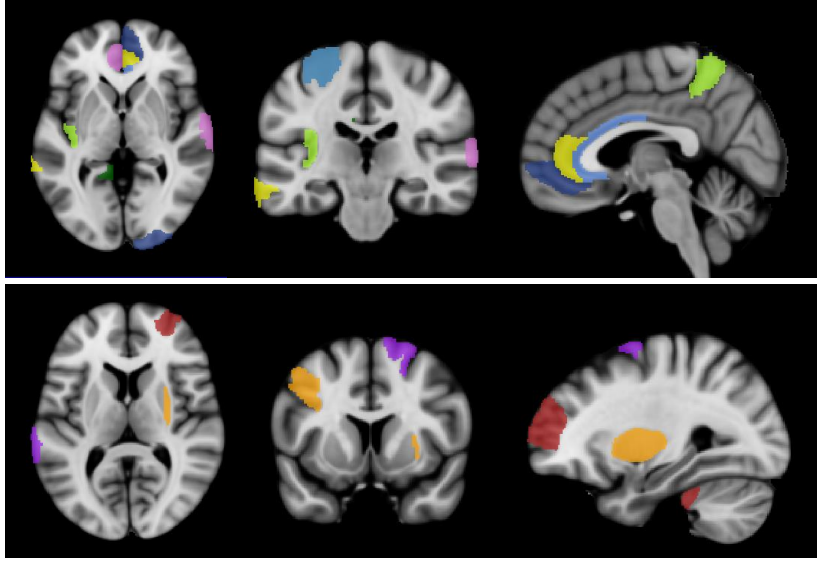


Figure 5.18: Visualisation of the features selected. In light green: Left G to Right A5m. In red: Right A46 to Right V. In light blue: Left A4ul to Right A24rv. In yellow: Left A21c to Right A32sg. In pink: Left A32sg to Right A22c. In orange: Left IFJ to Right dlPu. In dark green: Left A4ul to Left A23v (Left A4ul is in light blue). In purple: Left A22c to Right A6dl. In dark blue: Right A14m to Right OPC.

RF

The forward sequential feature selection has then been applied with the RF model using 150 decision trees. The three graphs obtained in Figure 5.19 are similar to the ones obtained in Figure 5.4 except that features selected from the wrapper method are, in this case, entries of the functional connectivity matrix.

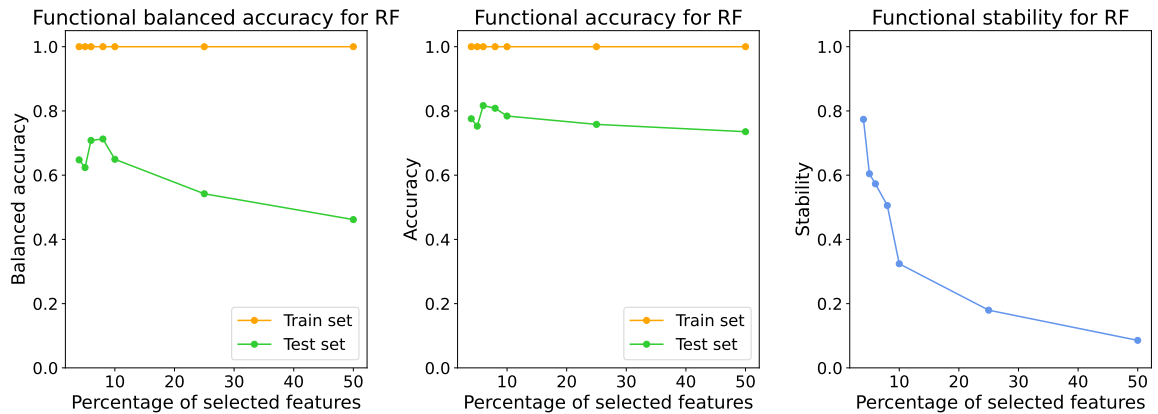


Figure 5.19: Accuracy and stability of the functional features selected from the pre-filter method of correlation and the wrapper method with the RF model.

Figure 5.20 illustrates the histogram showing the number of times a feature has been selected over the 35 iterations from the wrapper method based on the RF model when 4% of functional features (i.e., 2 features at each iteration) are selected. The meaning of colours bars is the same than the one for Figure 5.5.

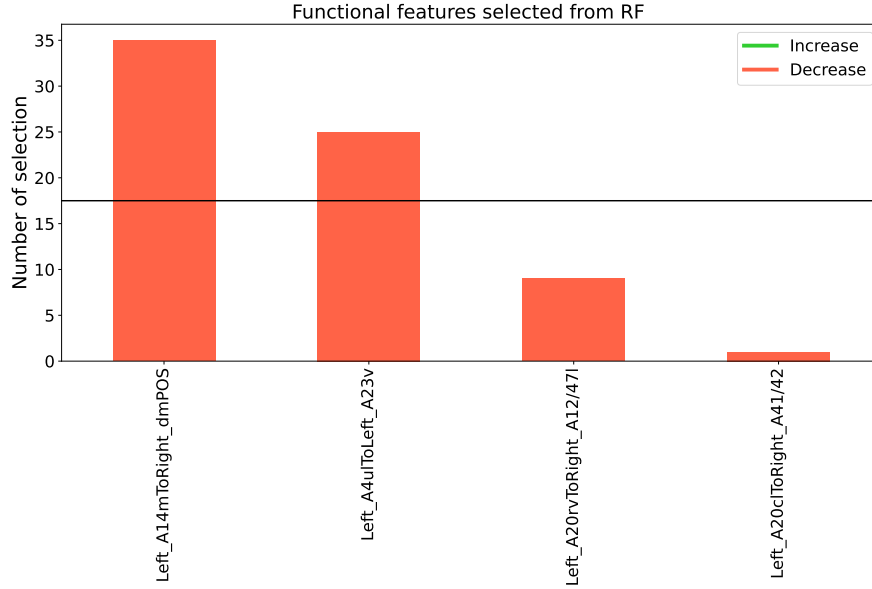


Figure 5.20: Functional features selected from the pre-filter method of correlation and the wrapper method with the RF model.

The regions of the functional features present in the histogram of Figure 5.20 can be visualised on a brain in Figure 5.21. The legend of the colours is in the caption of the figure. Two regions of same colour represent a feature selected from the method.

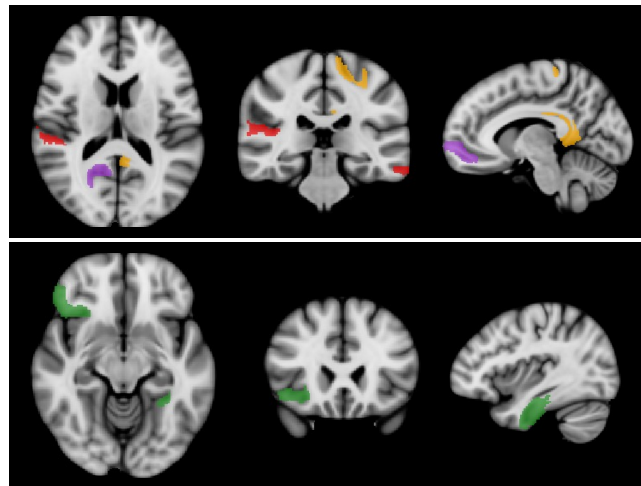


Figure 5.21: Visualisation of features selected. In purple: Left A14m to Right dmPOS. In orange: Left A4ul to Left A23v. In red: Left A20cl to Right A41/42. In green: Left A20rv to Right A12/47l.

Pre-filtered with mutual information

The wrapper method has then been applied on the features kept from the mutual information.

SVM

First, the forward sequential feature selection has been applied using the SVM model with a C value of 10.

The graphs obtained in Figure 5.22 are similar to those obtained from Figure 5.13 except that the features have been pre-selected with the mutual information method, instead of the correlation method.

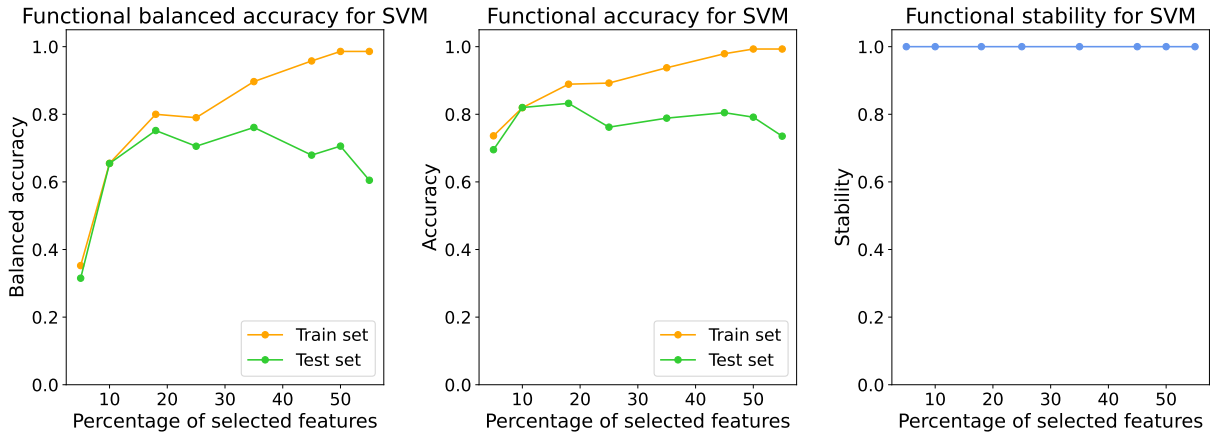


Figure 5.22: Accuracy and stability of the functional features selected from the pre-filter method of mutual information and the wrapper method with the SVM model.

The histogram in Figure 5.23 illustrates the mean of the weights attributed to each functional feature selected from the wrapper method for 20% of features kept. The weight indicates the importance of the feature to discriminate vertigo participants who have compensated or not from control participants.

The regions of the functional features present in the histogram of Figure 5.23 can be visualised on a brain in Figure 5.24. The legend of the colours is in the caption of the figure. Two regions of same colour represent a feature selected from the method.

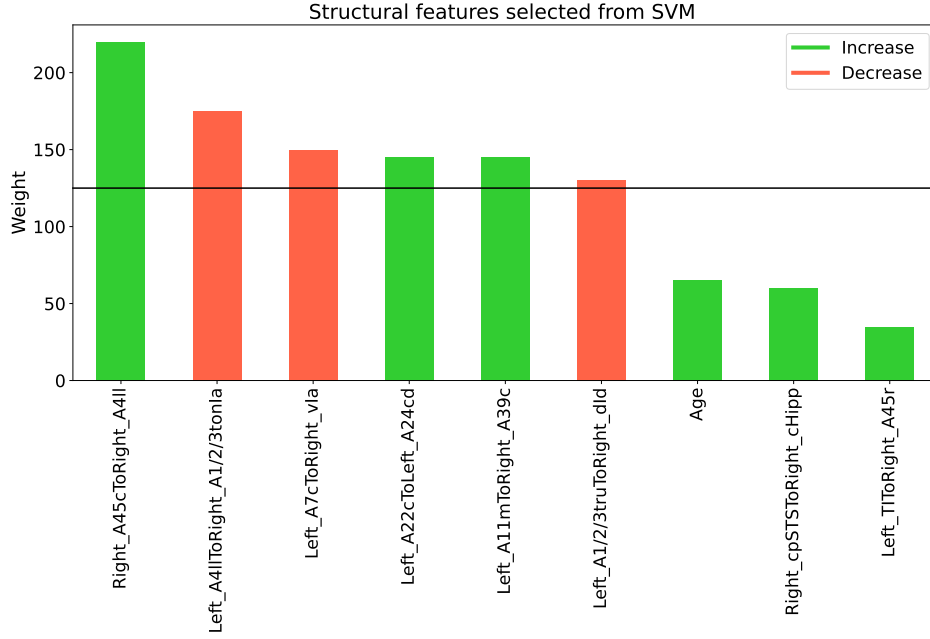


Figure 5.23: Functional features selected from the pre-filter method of mutual information and the wrapper method with the SVM model to differentiate control participants from vertigo participants who have compensated or not.

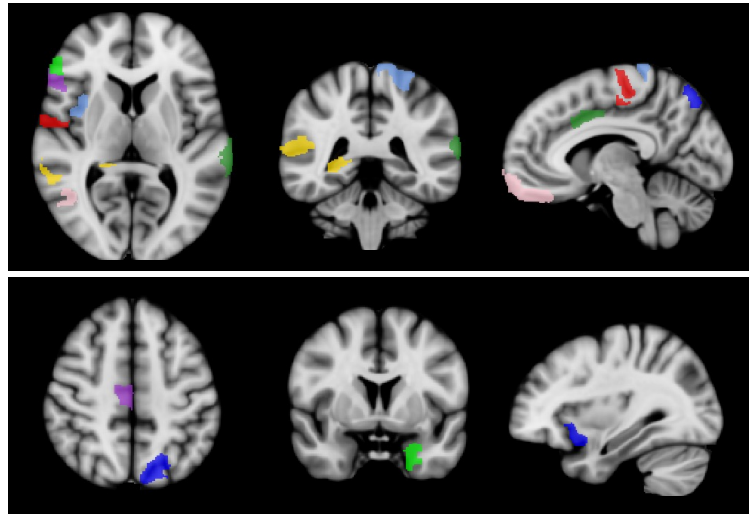


Figure 5.24: Visualisation of features selected. In purple: Right A45c to Right A4ll. In red: Left A4ll to Right A1/2/3tonIa. In royal blue: Left A7c to Right vIa. In dark green: Left A22c to Left A24cd. In pink: Left A11m to Right A39c. In light blue: Left A1/2/3tru to Right dId. In yellow: Right cpSTS to Right cHipp. In light green: Left TI to Right A45r.

The histogram in Figure 5.25 is the same as the previous one in Figure 5.23 except that, in this case, the weights of the features selected are an indication on how the feature allows to discriminate vertigo participants from control participants and vertigo participants who have compensated.

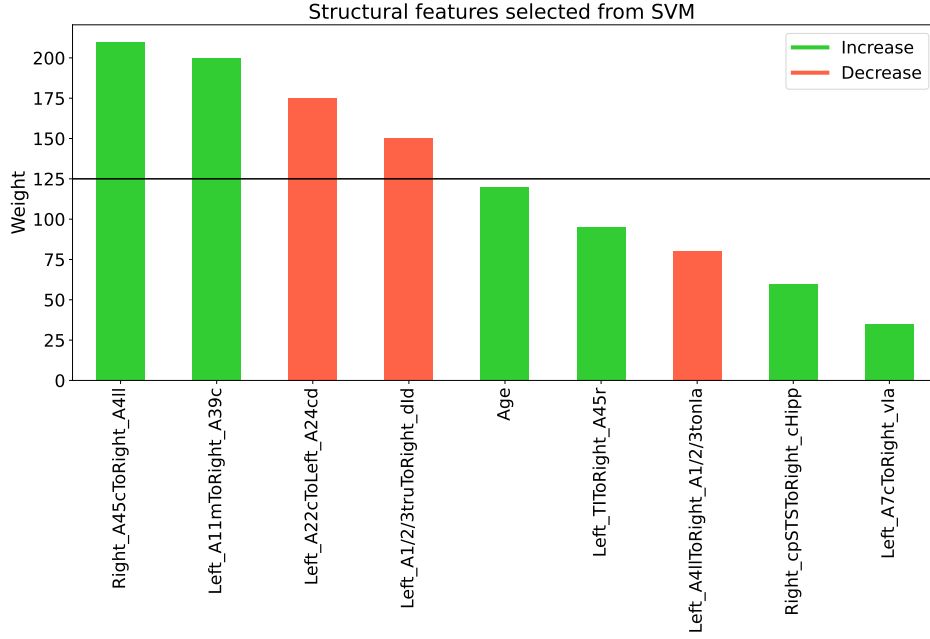


Figure 5.25: Functional features selected from the pre-filter method of Mutual Information and the wrapper method with the SVM model to differentiate control participants and vertigo participants who have compensated from vertigo ones.

The regions of the functional features present in the histogram of Figure 5.25 can be visualised on a brain in Figure 5.26 (the legend of the colours is in the caption). Two regions of same colour represent a feature selected from the method.

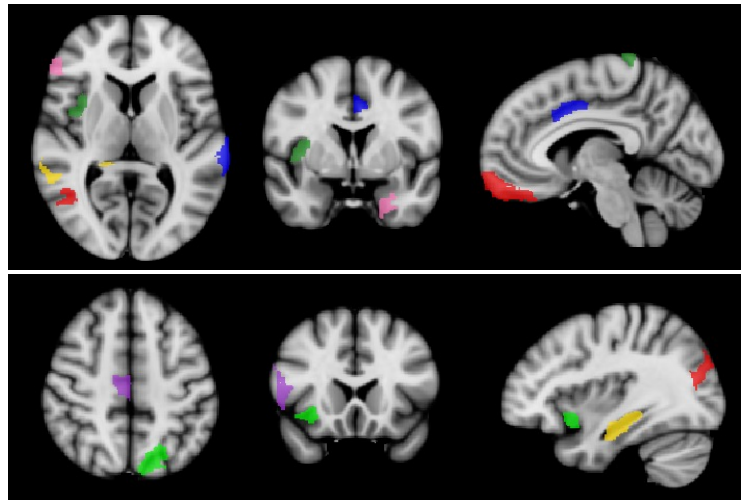


Figure 5.26: Visualisation of features selected. In purple: Right A45c to Right A4ll. In red: Left A11m to Right A39c. In royal blue: Left A22c to Left A24cd. In dark green: Left A1/2/3tru to Right dld. In pink: Left Tl to Right A45r. In light blue: Left A4ll to Right A1/2/3tonla. In yellow: Right cpSTS to Right cHipp. In light green: Left A7c to Right vla.

RF

The wrapper method has then been performed using the RF model with 10 decision trees.

The graphs obtained in Figure 5.27 are similar to those obtained from Figure 5.19 except that the features have been pre-selected with the mutual information method, instead of the correlation method.

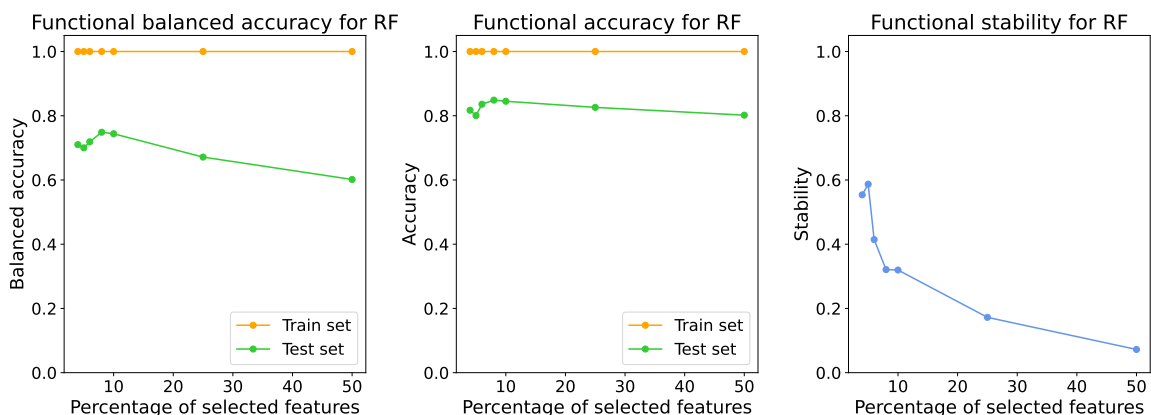


Figure 5.27: Accuracy and stability of the functional features selected from the pre-filter method of mutual information and the wrapper method with the RF model.

The features having a number of selections higher than the mean when 5% of features (i.e., 2 at each iteration) are selected can be seen in Figure 5.28. The complete histogram can be found in Appendix J.



Figure 5.28: Functional features (having a number of selection higher than the mean) selected from the pre-filter method of mutual information and the wrapper method with the RF model.

The regions of the functional features present in the histogram of Figure 5.28 can be visualised on a brain in Figure 5.29. The legend of the colours is in the caption of the figure. Two regions of same colour represent a feature selected from the method.

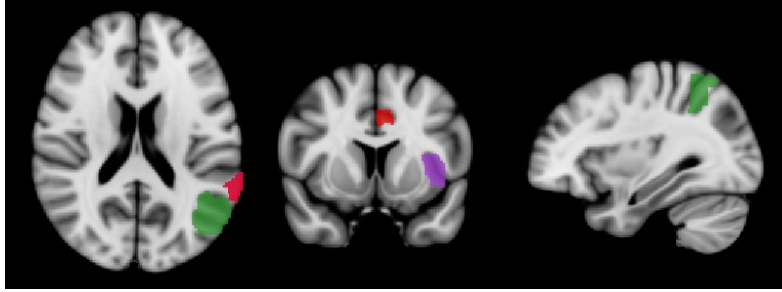


Figure 5.29: Visualisation of features selected. In red: Left A22c to Left A24cd. In purple: Left A22c (in red in the first image) to Left dId. In green: Left A39rv to Right A7ip.

5.4 Statistical cerebellum analysis

The results obtained for the statistical tests on the cortical thickness asymmetry can be seen in Table 5.9. It can be seen that two lobules (*lobule V* and *lobule IV*) present significant differences between the three classes. The tests for the 5 other variables can be found in Appendix K.

Table 5.9: Statistical values for the cortical thickness asymmetry of all lobules of the cerebellum.

	p-value	Benjamini correction, Q = 0.1	Benjamini correction, Q = 0.05	Bonferroni correction
Lobule V	0.00245	0.00833	0.00417	0.004167
Lobule IV	0.00354	0.01667	0.00833	
Lobule VIIIA	0.0497	0.025	0.0125	
Lobule IX	0.08	0.033	0.0168	
Lobule I-II	0.1488	0.0417	0.0208	
Lobule X	0.2003	0.05	0.025	
Lobule VIIIB	0.2505	0.0583	0.0292	
Lobule VIIB	0.6264	0.0667	0.033	
Lobule Crus II	0.6666	0.075	0.0375	
Lobule III	0.7518	0.083	0.0417	
Lobule Crus I	0.7983	0.0916	0.0458	
Lobule VI	0.9076	0.1	0.05	

Pairwise statistical tests are realised to identify from which classes the difference is the more significant. The results can be seen in Table 5.10.

Table 5.10: Statistical values for the cortical thickness asymmetry of lobules V and IV between Control - Vertigo participants (C-V), Control - Compensated Vertigo participants (C-VC) and Vertigo - Compensated Vertigo participants (V-VC).

	C-V		C-VC		V-VC	
	p-value	Benjamini	p-value	Benjamini	p-value	Benjamini
Lobule V	0.0001	0.0167	0.0115	0.0667	0.6287	0.0833
Lobule IV	0.0004	0.0333	0.0024	0.05	0.8429	0.1

The mean and standard deviation of the cortical thickness asymmetry for lobules V and IV can be seen in the left and right graphs, respectively, of Figure 5.30 for each class (control, vertigo and vertigo compensated classes). The blue squares represent the mean of the cortical thickness asymmetry for the concerned class while the black lines represent the standard deviation. The dashed red line indicates the 0 value to better visualise the positive and negative means.

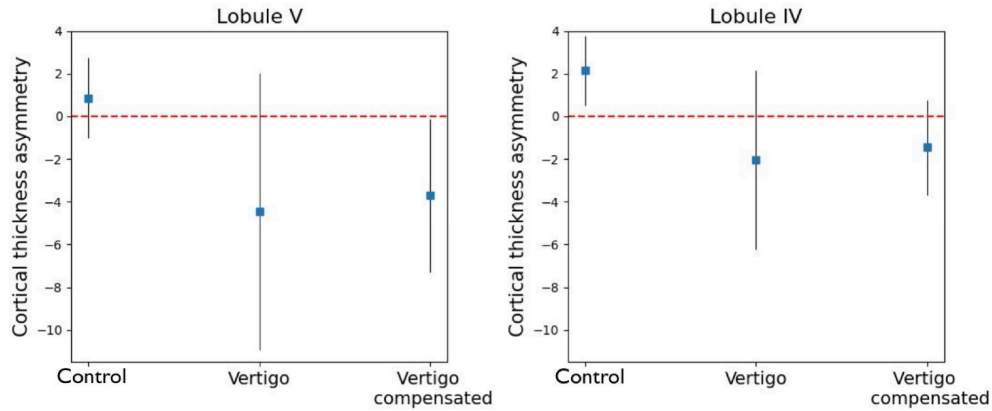


Figure 5.30: Mean (blue rectangle) and standard deviation (black line) of the cortical thickness asymmetry of the lobule V (left) and the lobule IV (right) for the control participants, vertigo participants and vertigo participants who have compensated.

The cerebellum can be visualised in green on the T1-weighted image in the MNI space in Figure 5.31. Lobule IV is represented in orange while lobule V is coloured in purple.



Figure 5.31: Lobule IV (orange) and Lobule V (purple) of the cerebellum (green) on a T1 image in the MNI space.

Chapter 6

Discussion

The sixth chapter includes the discussion of the results shown in the fifth chapter. First, the structural and functional connectomes are discussed followed by a discussion about features selection methods. Then, features or regions selected as significant from the statistical tests and the machine learning methods are analysed.

6.1 Connectomes

Structural

The weighted structural matrix obtained from the dMRI images for a vertigo participant who has compensated is illustrated in Figure 5.2 (Left). As can be seen in the figure, intrahemispheric structural connections are favoured. Indeed, two bright squares can be distinguished in the upper left and lower right of the matrix, corresponding to connectivity between regions of the same hemisphere. On the other hand, there are almost no anatomical connections between regions of left and right hemispheres. This is coherent with the anatomical structure of the brain as both hemispheres are only connected through some fibers forming a bundle of fibers called corpus callosum. These findings are consistent with the article written by *Raiser et al.* described in the [Connectome analysis for vestibular system](#) section [39].

Functional

The functional matrix obtained from the fMRI images for a control participant is illustrated in Figure 5.2 (Right). In contrast to the structural matrix, the functional matrix shows a correlated activity between the right and left hemispheres of the brain. The yellow squares correspond to highly correlated functional groups in the brain. Three bright diagonals can be observed in the matrix. The one at the center shows clusters of correlated regions close to each other. The two other ones, on the other hand, show that the homotopic nodes of the left and right hemispheres have a strong correlated activity. These findings are also consistent with the article written by *Raiser et al.* described in the [Connectome analysis for vestibular system](#)

section [39]. The dark blue cross in the middle of the matrix corresponds to the non-registered regions of the cerebellum that were subsequently removed. The less marked bands at the bottom of the matrix correspond to the vermis and were also non-registered for several patients. They were therefore removed for the analyses of features.

6.2 Accuracy and stability of the feature selection methods

Different methods were used to select features in this work. There are two possibilities for the pre-filtering method: the correlation which is a linear metric and the mutual information which is a non-linear metric. Then, two machine learning algorithms are used being the SVM which is a linear algorithm or the RF which is a non-linear algorithm.

First, it can be seen in Figure 5.3 that the SVM model did not give any satisfactory results with structural features. Indeed, as a reminder, the performances based on the balanced and non-balanced accuracies are respectively equal to 0.33 and 0.75 which are the performances equivalent to the ones of a model assigning all participants to the vertigo class. The performances of the SVM model on structural data are thus not good. On the other hand, Figure 5.4 shows that the RF model gives better results as the balanced accuracy is around 0.56 and the non-balanced accuracy around 0.78. This is a first sign that the link between the structural features and the target is rather non-linear as a non-linear model is able to give better results than the random model. Moreover, if one compares Figures 5.3 and 5.8, it can be seen that the performances of the RF model when the input features are selected by the mutual information are better than the ones when the input features are selected by the correlation. This is another indication that the link between structural features and the target is rather non-linear.

On the other hand, the performances of the SVM on the functional features is better than the ones of the RF as can be seen in Figures 5.13 and 5.19. The performances are also a bit better when the pre-filtering is performed with correlation compared to the mutual information, as can be seen when comparing Figure 5.13 and 5.22. This shows that the link between the functional features and the target is rather linear.

In Table 6.1 are summarised the balanced and non-balanced accuracies obtained for each method with the percentage of selected features chosen to construct the histograms for the (a) structural and (b) functional features. As can be seen, the non-linear pre-filter method with non-linear model in the wrapped method gives the best performances (highlighted in green) for the structural features while the linear filter method with linear model gives the best performances (highlighted in green) for the functional features.

Table 6.1: *Balanced / Non-balanced accuracies of the SVM and RF models for correlation and mutual information (MI) methods with the percentage of features selected chosen to analyse data. The best performances for structural and functional data are in green.*

<i>a. Structural features</i>			<i>b. Functional features</i>		
	SVM linear	RF non-linear		SVM linear	RF non-linear
Correlation linear	0.33 0.75	0.56 0.69	Correlation linear	0.80 0.89	0.65 0.78
MI non-linear	0.33 0.74	0.71 0.87	MI non-linear	0.75 0.83	0.70 0.80

When analysing the trend of orange and green curves on the graphs showing the balanced and non-balanced accuracies for each method, the phenomenon of over-fitting can be observed. As a reminder, over-fitting occurs when a model is trained on a set of samples (i.e., train set) for which it predicts almost perfectly the target but does not give good results at all when performing it on new data (i.e., test set). It can be observed on the graphs that the difference between the orange and the green curves is sometimes big, showing a very good (sometimes almost perfect) performance on the train set (orange curve) but a less good performance on the test set (green curve). It can also be observed that the higher the dimensionality of features space, the greater the chance to have over-fitting. This can be seen in Figure 5.22, where the accuracies on the train (orange) and test (green) sets are almost the same when only 5% or 10% of the features are selected. However, the higher the percentage of features selected, the greater the difference between the orange and green curves. The green curve starts to decrease while the orange curve tends towards a perfect accuracy (1).

When analysing the stability of the different methods, it can be observed that the SVM model always has a stability of 1, whatever the pre-filtering method and the percentage of features selected. Indeed, the same features were always selected at each run of the sequential feature selection. Even though the stability of 1 is surprising, it has been demonstrated by I. Buciuc et al. that SVM models are stable classifiers [93].

6.3 Selected regions of interest

To analyse features selected by the machine learning methods, a table is constructed comprising only the features for which the number of iterations (with the RF model) or their mean weight (with the SVM model) is above the mean (dark line in the histograms). For the RF model, the number of features above the mean is equal to or greater than the number of features that had to be selected by the sequential feature selection method. For this model, all the features are analysed by looking at the mean number of fibers (for structural data) or mean functional activity (for functional data).

However, for the SVM model, the number of features selected by the wrapper method is higher than the number of features above the mean. As the number of features selected by the wrapper method in this case is equal to 9 or 10, this makes a features space in high dimension (9 or 10 dimensions). Interpreting the results in such a high dimension space becomes tricky. It was thus decided to only analyse features above the mean. However, it is important to keep in mind that it is the combination of all the features selected by the wrapper method that allows to best predict the class the participant belongs to. Looking at the specific mean number of fibers or mean functional activity of each feature is therefore less useful in understanding why the feature was selected by the SVM model. It was thus decided to look rather at the trend in the difference between classes (increase or decrease of mean fibers density or mean functional activity).

For both RF and SVM models, tables are created and include information such as the location of regions in the brain, the functions it is known to play, the number of iterations the feature has been selected (for RF model) or its mean weight (for SVM model), whether regions have obtained a p-value smaller than 0.05 for the statistical tests of betweenness and degree centrality, upon which classes a difference is observed and finally, whether the region has already been reported in the literature developed in section [Connectome analysis for vestibular system](#). In an article developed in this section, a hubness score was attributed for some regions encountered in the tables below. This score can be equal to 1, 2 or 3. This means that the corresponding region belongs to the 30% of areas having a higher value in 1, 2 or 3 of the following metrics: nodal strength, betweenness centrality or local efficiency. It should also be noted that the reference to the section [Connectome analysis for vestibular system](#) allows to assess the individual consistency of each region from each feature but, to our best knowledge, none of the features identified in this work are found back in the literature as no articles performing machine learning on the basis of connectomes using the Brainnetome and Suit atlases were found in the literature to analyse dysfunctions of the vestibular system. However, information on the structural or functional connections between brain areas involved in certain processes, such as visual-spatial or cognitive integration is found in certain articles (see [Connectome analysis for vestibular system](#) section) and the results obtained can thus be compared with these findings.

To facilitate the visibility of the discussion of the results, only tables of the features selected from methods having the best performances are shown, i.e., the features selected from the MI pre-filtering method and RF model for structural data and the features selected from the correlation pre-filtering method and SVM model for functional data. The features selected by all the other methods can be found in Appendix L.

Finally, in the following tables, the abbreviations C, V and VC are used. They respectively refer to Control participants, Vertigo participants and Vertigo participants who have Compensated.

Structural

As a reminder, the performances of the SVM model with the structural data are not good at all and the structural features selected by this model are thus not analysed. Only the structural features selected from the RF models are therefore discussed.

Correlation pre-filter method

The first regions analysed are the ones obtained with the correlation pre-filtering and with the RF model. The most important features are predominantly (5 features over 7 in Figure 5.5) features representing intra-hemispheric connections. This makes sense since, as explained in section [Connectomes](#), most of the structural connections are intrahemispheric connections. Three features are above the mean selection number and thus analysed in the table of Appendix L.1. In the tables comprising the structural features, the mean number of fibers refers to the weighted values.

A feature selected is the anatomical connection between regions *A20 rv* and *NAC* of the left hemisphere. This feature is increased for participants suffering from vertigo, and still more for vertigo participants who have compensated, compared to control participants. Although the nucleus accumbens of the basal ganglia is not identified in literature developed in section [Connectome analysis for vestibular system](#), it has been shown that PIVC and PIC share ipsilateral connections to the lateral nuclei of the basal ganglia.

A second feature selected is the anatomical connection between the *iOccG* region of the left hemisphere and the *A5m* region of the right hemisphere. In opposite to the first feature, the difference between the three classes for the means of weighted fibers density between the *iOccG* and *A5m* regions is very small. Actually, only 14% of control participants have non-null fibers density values, 3.8% for vertigo participants and 45% for vertigo participants who have compensated. The number of null values in the connectivity matrix for the second feature is thus high (88.6%). However, the feature has not been removed because the number of zeros for the participants that suffered from vertigo but have compensated is lower than the arbitrary threshold of 60% fixed in the methodology part. The feature was thus selected by the machine learning algorithm because the difference between vertigo

participants who have compensated on the one hand and the control and vertigo participants on the other hand is high. However, in view of the number of zeros in the connectivity matrix, it is clearly not the best feature to discriminate between vertigo and control participants. The threshold arbitrarily put to 60% of null values in each class may not be the best choice, and other thresholds or other ways of eliminating some features before the pre-filtering can be considered and compared with the threshold of 60%.

The third feature selected is the structural connection between the *mOccG* and *msOccG* regions of the left hemisphere. However, the mean fibers density being much higher for the vertigo participants who have compensated compared to the one of control and vertigo participants, this feature seems to be used to discriminate control and vertigo participants on the one hand and vertigo participants who have compensated on the other hand. Indeed, as the model must be able to separate the participants in three classes, not all the features used show significant difference between control and vertigo participants.

Mutual information pre-filter method

Then, the regions obtained with the mutual information pre-filtering and with the RF model can be analysed. The combination of these two methods give the highest performances for structural data, as can be seen in Table 6.1. As with the correlation pre-filtering, the features selected are predominantly (4 over 6 features in Figure 5.10) intrahemispheric connections. Only the four features with a number of selection higher than the mean are analysed in Tables 6.2 and 6.3.

Table 6.2: Analysis of structural features selected with the RF model and MI method.

Features selected	Left A23d to Right A8m	Right A39rv to Right rCunG
Number of iterations	35/35	35/35
Location	- <i>A23d</i>: dorsal area 23 in the cingulate gyrus of the limbic lobe. - <i>A8m</i>: medial area 8 in the superior frontal gyrus of the frontal lobe.	- <i>A39rv</i>: rostroventral area 39 in the inferior parietal lobule. - <i>rCunG</i>: rostral cuneus gyrus in the MedioVentral occipital cortex.
Functions	- <i>A23d</i>: the cingulate gyrus coordinates sensory inputs (e.g. role in visuo-spatial orientation) [94, 95]. - <i>A8m</i>: involved in impulse control and modulates inhibitory control and motor urgency [96].	- <i>A39rv</i>: plays a role in various cognitive functions such as attention, action processing and language [97]. - <i>rCunG</i>: plays a role in the visuo-spatial attention tasks [98].

Features selected	Left A23d to Right A8m	Right A39rv to Right rCunG
Difference between classes	Mean fibers density: 0 for C, 1.4 for V, 0.5 for VC.	Mean fibers density: 3.3 for C, 1.7 for V, 2.9 for VC.
p-value < 0.05 for between-ness	- <i>A23d</i> : right hemisphere for the structural connectome, left hemisphere for the functional connectome. - <i>A8m</i> : No.	- <i>A39rv</i> : left hemisphere for the structural connectome. - <i>rCunG</i> : No.
p-value < 0.05 for degree	- <i>A23d</i> : No. - <i>A8m</i> : No.	- <i>A39rv</i> : No. - <i>rCunG</i> : left hemisphere for the structural connectome.
Literature	From <i>Indovina et Al.</i> study, region <i>A23rd</i> belongs to the multimodal vestibular network and has a hubness score equal to 3.	From <i>Indovina et Al.</i> study, region <i>A39rv</i> belongs to the multimodal vestibular network.

Table 6.3: Analysis of the third and fourth structural features selected with the RF model and MI method.

Features selected	Left A45c to Left A24cd	Right A5l to Right A40rv
Number of iterations	34/35	34/35
Location	- <i>A45c</i> : caudal area 45 of the inferior frontal gyrus. - <i>A24cd</i> : caudodorsal area 24 of the cingulate gyrus of the limbic lobe.	- <i>A5l</i> : lateral area 5 in the superior parietal lobule. - <i>A40rv</i> : rostroventral area 40 in the inferior parietal lobe.
Functions	- <i>A45c</i> : responsible for speech production [96]. - <i>A24cd</i> : the cingulate gyrus coordinates sensory inputs (e.g. role in visuo-spatial orientation) [94, 95].	- <i>A5l</i> : superior parietal lobule is closely linked with the occipital lobe and is involved in aspects of attention and visuospatial perception [99]. - <i>A40rv</i> : implicated in diverse mental processes, such as attention, language and social cognition [100].

Features selected	Left A45c to Left A24cd	Right A5l to Right A40rv
Difference between classes	Mean fibers density: 167 for C, 3 for V, 11 for VC.	Mean fibers density: 84 for C, 56 for V, 148 for VC.
p-value < 0.05 for between-ness	- <i>A45c</i> : No. - <i>A24cd</i> : No.	- <i>A5l</i> : right hemisphere for the structural connectome. - <i>A40rv</i> : No.
p-value < 0.05 for degree	- <i>A45c</i> : left hemisphere for the structural connectome. - <i>A24cd</i> : No.	- <i>A5l</i> : right hemisphere for the structural connectome. - <i>A40rv</i> : No.
Literature	From <i>Indovina et al.</i> both regions <i>A45c</i> and <i>A24cd</i> belong to the multimodal vestibular network and have a hubness score of 1.	From <i>Indovina et al.</i> regions <i>A5l</i> is belonging to the multimodal vestibular network and have a hubness score of 3.

For the feature *Left A23d to Right A8m*, none of the control participants shows a connection between those regions, only seven of the vertigo participants have a value in the connectivity matrix and 6 over 11 participants present a connection between those regions. As previously, the number of zeros in the connectivity matrix for this connection is probably too high to consider this feature as being significant to discriminate between control, vertigo and vertigo participants who have compensated. Other thresholds than 60% of null values should be tested.

For feature *Left A45c to Left A24cd*, the means show that this feature allows to discriminate between control participants on the one hand and vertigo participants who have compensated or not on the other hand.

Unlike the previous feature, the last feature *Right A5l to Right A40rv* seems to be used to discriminate between control and vertigo participants on the one hand and vertigo participants who have compensated on the other hand.

By retaining only the most consistent regions, anatomical connections between regions involved in cognitive (*A39rv* region), proprioceptive (*rCunG*, *A45c*, *A24cd*, *A20rv* regions) and affective (*NAC*) processing appear to be significant in discriminating control participants from vertigo participants who have compensated or not. In the article written by Jeremy L. Smith et al., described in the [Connectome analysis for vestibular system](#) section, findings on the structural and functional connectomes indicated that cognitive, affective and proprioceptive substrates contribute to vestibular processing and performance.

Functional

The SVM models have better performances for functional data compared to structural data, as can be seen in Table 6.1. The features selected by the SVM models are thus also analysed, in this section, contrary to structural data where only features selected by the RF models were discussed.

In the case of functional connectomes, very few entries of the matrix are null. Therefore, all the features chosen by the wrapper method show less than 60% of null values for the three classes. The individual features obtained can thus be analysed without taking into account the number of zeros present in each class. However, as previously noticed, it can not be forgotten that it is the combination of features which allows to get the best performance and not each feature individually.

Correlation pre-filter method with SVM model

Features selected from the correlation pre-filtering method and the SVM model can first be analysed. The combination of these two methods give the highest performances for functional data. The analysis can be split in two parts: first the regions allowing to discriminate vertigo participants who have compensated or not from control participants, and secondly regions allowing to discriminate vertigo participants from control and compensated vertigo participants.

In the first case, five features above the mean in Figure 5.14 are returned by the methods including the age, being the second feature with the highest weight, meaning it has a great impact on the functional connectivity. The four other features are analysed in Tables 6.4 and 6.5.

Table 6.4: Analysis of the first and third functional features selected from SVM model and correlation pre-filtering method to discriminate control from vertigo participants who have compensated or not.

Features selected	Left A22c to Right A6cdl	Left A4ul To Right A24rv
Mean weights	230	205
Location	- A22c: caudal area 22 in the superior temporal gyrus. - A6cdl: caudal dorsolateral area 6 in the precentral gyrus.	- A4ul: Area 4 of the precentral gyrus in the frontal lobe. - A24rv: rostroventral area 24 of the cingulate gyrus in the limbic lobe.

Features selected	Left A22c to Right A6cdl	Left A4ul To Right A24rv
Functions	- <i>A22c</i>: superior temporal gyrus is the site of auditory association cortex (and a site of multisensory integration) [101]. - <i>A6cdl</i>: precentral gyrus contains the primary motor cortex which is responsible for controlling the voluntary movements of specific body parts [102].	- <i>A4ul</i>: precentral gyrus contains the primary motor cortex which is responsible for controlling the voluntary movements of specific body parts [102]. - <i>A24rv</i>: the cingulate gyrus coordinates sensory inputs (e.g. role in visuo-spatial orientation) [94, 95].
Difference between classes	Simultaneous activity higher for vertigo participants who have compensated or not compared to control ones.	Simultaneous activity smaller for vertigo participants who have compensated or not compared to control ones.
p-value < 0.05 for between-ness	- <i>A22c</i>: No. - <i>A6cdl</i>: No.	- <i>A4ul</i>: right hemisphere for the functional connectome. - <i>A24rv</i>: No.
p-value < 0.05 for degree	- <i>A22c</i>: No. - <i>A6cdl</i>: No.	- <i>A4ul</i>: No. - <i>A24rv</i>: right hemisphere for the structural connectome.
Literature	From <i>Indovina et al.</i> , region <i>A6cdl</i> belongs to the multimodal vestibular network.	From <i>Indovina et al.</i> , region <i>A24rv</i> belongs to the multimodal vestibular network and has a hubness score equal to 1.

Table 6.5: Analysis of the fourth and fifth functional features selected from SVM model and correlation pre-filtering method to discriminate control from vertigo participants who have compensated or not.

Features selected	Left A4ul to Right A23v	Right A14m To Right OPC
Mean weights	180	135
Location	- <i>A4ul</i>: Area 4 of the precentral gyrus in the frontal lobe. - <i>A23v</i>: ventral area 23 of cingulate gyrus in the limbic lobe.	- <i>A14m</i>: medial area 14 of the orbital gyrus in the frontal lobe. - <i>OPC</i>: occipital polar cortex in lateral occipital lobe.

Features selected	Left A4ul to Right A23v	Right A14m To Right OPC
Functions	- <i>A4ul</i>: precentral gyrus contains the primary motor cortex which is responsible for controlling the voluntary movements of specific body parts [102]. - <i>A23v</i>: the cingulate gyrus coordinates sensory inputs (e.g. role in visuo-spatial orientation) [94, 95].	- <i>A14m</i>: orbital gyrus is implicated in intellectual and emotional expression [103]. - <i>OPC</i>: occipital lobe (both right and left) mainly handle visual perception [104].
Difference between classes	Simultaneous activity smaller for vertigo participants who have compensated or not compared to control ones.	Simultaneous activity smaller for vertigo participants who have compensated or not compared to control ones.
p-value < 0.05 for between-ness	- <i>A4ul</i>: right hemisphere for the functional connectome. - <i>A23v</i>: No.	- <i>A14m</i>: No. - <i>OPC</i>: No.
p-value < 0.05 for degree	- <i>A4ul</i>: No. - <i>A23v</i>: No.	- <i>A14m</i>: right hemisphere for structural connectome. - <i>OPC</i>: No.
Literature	From <i>Indovina et al.</i> , region <i>A23v</i> belongs to the multimodal vestibular network and has a hubness score equal to two.	/

Looking at these results, it can be observed that many features selected are regions located in the frontal lobe (precentral and orbital gyrus) functionally connected to regions playing a role in visuo-spatial orientation (cingulate gyrus) or visual perception (occipital lobe) with a decreasing trend for vertigo participants compared to control ones. In the article written by Jeremy L. Smith et al., described in the *Connectome analysis for vestibular system* section, the connections with the cingulate cortex were found to be highly associated with the post-concussive vestibular dysfunction of patients. Visual, supplementary motor, and frontal and cingulate eye field nodes were also exhibiting significant differences for patients suffering from PCVD but with an increased centrality trend.

In the second case where features allow to discriminate vertigo participants from control and compensated vertigo participants together, the analysis of the four features above the mean (see Figure 5.17) can be seen in Tables 6.6 and 6.7. Again, the age is the second feature with the highest weight and has thus an impact on the functional activity of brain.

Table 6.6: Analysis of functional features selected from SVM and correlation to discriminate vertigo participants from control and compensated vertigo participants.

Features selected	Left G To Right A5m	Right A46 To Right V
Mean weights	235	200
Location	- <i>G</i>: hypergranular insular in the insular gyrus. - <i>A5m</i>: medial area 5 in the precuneus (parietal lobe).	- <i>A46</i>: area 46 of the middle frontal gyrus. - <i>V</i>: Lobule V of the cerebellum.
Functions	- <i>G</i>: the insular gyrus is involved in sensory preprocessing and hypergranular insula has significant connections with the auditory cortex [105]. - <i>A5m</i>: the precuneus is implicated in the sense of self, autobiographical memory, spatial function, and navigation [106].	- <i>A46</i>: a site of convergence of the dorsal and ventral attention networks [107]. - <i>V</i>: the cerebellum modulates the motor pathways that control balance and gaze [108].
Difference between classes	Simultaneous activity smaller for vertigo participants compared to control and compensated vertigo ones.	Simultaneous activity higher for vertigo participants compared to control and compensated vertigo ones.
p-value < 0.05 for between-ness	- <i>G</i>: No. - <i>A5m</i>: right hemisphere for the functional connectome.	- <i>A46</i>: No. - <i>V</i>: No.
p-value < 0.05 for degree	- <i>G</i>: No. - <i>A5m</i>: left hemisphere for the structural connectome.	- <i>A46</i>: No. - <i>V</i>: left hemisphere for the functional connectome.
Literature	From <i>Indovina et al.</i> , region <i>A5m</i> belongs to the multimodal vestibular network and has a hubness score equal to 1 for the right hemisphere. Region <i>G</i> belongs to the multimodal vestibular network.	/

Table 6.7: Analysis of functional features selected from SVM and correlation to discriminate vertigo participants from control and compensated vertigo participants.

Features selected	Left A4ul To Right A24rv	Left A21c To Right A32sg
Mean weights	170	155
Location	- <i>A4ul</i>: area 4 (upper limb region) in the precentral gyrus of the frontal lobe. - <i>A24rv</i>: rostroventral area 24 in the cingulate gyrus of the limbic lobe.	- <i>A21c</i>: caudal area 21 in the middle temporal gyrus. - <i>A32sg</i>: subgenual area 32 in the cingulate gyrus of the limbic lobe.
Functions	- <i>A4ul</i>: precentral gyrus contains the primary motor cortex which is responsible for controlling the voluntary movements of specific body parts [102]. - <i>A24rv</i>: the cingulate gyrus coordinates sensory inputs (e.g. role in visuo-spatial orientation) [94, 95].	- <i>A21c</i>: middle temporal gyrus subserves language and semantic memory processing, visual perception, and multimodal sensory integration [109]. - <i>A32sg</i>: the cingulate gyrus coordinates sensory inputs (e.g. role in visuo-spatial orientation) [94, 95].
Difference between classes	Simultaneous activity smaller for vertigo participants compared to control and compensated vertigo ones.	Simultaneous activity higher for vertigo participants compared to control and compensated vertigo ones.
p-value < 0.05 for between-ness	- <i>A4ul</i>: right hemisphere for the functional connectome. - <i>A24rv</i>: No.	- <i>A21c</i>: No. - <i>A32sg</i>: No.
p-value < 0.05 for degree	- <i>A4ul</i>: No. - <i>A24rv</i>: right hemisphere for the structural connectome.	- <i>A21c</i>: left hemisphere for the structural connectome. - <i>A32sg</i>: right hemisphere for the structural connectome.
Literature	From <i>Indovina et al.</i> , region <i>A24rv</i> belongs to the multimodal vestibular network and has a hubness score equal to 1.	From <i>Indovina et al.</i> , region <i>A32sg</i> belongs to the multimodal vestibular network.

The significant features in discriminating vertigo participants from control and compensated vertigo participants include an increased functional connection between the insular gyrus (*G region*) and precuneus (*A5m region*), and between the temporal gyrus (*A21c region*) and cingulate gyrus (*A32sg region*) while a decreased functional connection between the middle frontal gyrus (*A46 region*) and lobule V of the cerebellum. The last feature (*Left A4ul To Right A24rv*) is the same as the previous case. Actually, when comparing the histograms in Figure 5.14 (allowing to discriminate vertigo participants who have compensated or not from control participants), and in Figure 5.17 (allowing to discriminate vertigo participants from control and compensated vertigo participants), it can be observed that the 9 features selected in the first case are also selected in the second one (which makes sense as the features output from the same model). However, the weights attributed to each feature are different. Indeed, three of the four first features selected with the highest weight in each case are different, meaning that in each case, three features receive relatively different weights. This may indicate that the class of the vertigo participants who have compensated influences the discrimination of vertigo participants from control participants because, depending on the class they are put with, the most important features are different.

For the following methods, as the performances are lower than the SVM model with the correlation pre-filter method, the tables comprising the analysis of features selected are shown in Appendices. These features are not discussed individually in this section but features that are redundant for different methods will be more analysed in the section *Redundancy*.

Correlation pre-filter method with RF model

The two features obtained from the correlation pre-filtering method and with the RF model above the mean are analysed in the table of Appendix L.2.

Mutual information pre-filter method with SVM model

The features obtained from the mutual information pre-filtering method and with the SVM model are analysed in the tables of Appendix L.3. As with the correlation pre-filtering, the analysis can be split in two: on the one hand, features allowing to discriminate vertigo participants who have compensated or not from control participants, and on the other hand features allowing to discriminate vertigo participants from control participants and vertigo participants who have compensated.

There are six features that can be identified in the first case and four in the second case. As those four features are included in the 6 previous ones, all features can be found in the same table of Appendix L.3. The name of these four features as well as their trend between classes for this second case can be seen in Appendix L.4. Contrary to what was obtained for the SVM model with the correlation pre-filtering, being that the most important features are mainly different for both parts of the analysis, in this case, the features having the higher weights for the second part

of the analysis are involved in the ones obtained for the first part of the analysis. This may indicate that the participants who have compensated do not influence the classification based on these features because the same features are chosen regardless of the class the compensated vertigo participants are placed with. The age is also a feature selected in both cases (histograms in Figures 5.23 and 5.25) but contrary to the correlation pre-filtering method, it is not a feature having a weight above the mean.

Mutual information pre-filter method with SVM model

Finally, the three features obtained from the mutual information pre-filtering method and with the RF model being selected a number of times above the mean, are analysed in Appendix L.5.

Redundancy

As can be seen from the review of the results obtained, 30 features are finally selected among the initial ones to discriminate between the three classes. 36 different regions are included among these 30 features. In general, regions located in the frontal lobe (such as the precentral gyrus, involved in motor control, and orbital gyrus implicated in intellectual and emotional expression), in the limbic lobe (like the cingulate gyrus which coordinates sensory inputs, playing a role in visuo-spatial orientation) or the occipital lobe (principally involved in vision) are likely to play a role in the vestibular system, specifically to discriminate participants suffering from vertigo from control participants. To a lesser extent, regions in the temporal lobe (such as the fusiform and middle temporal gyrus, playing a role in vision, and the superior temporal gyrus, involved in the processing of auditory information) and in the parietal lobe (such as the precuneus involved in the sense of self, autobiographical memory, spatial function, and navigation) also seem to contribute to vestibular processing and performances.

It can be observed that some regions are repeatedly present, particularly for functional features (which makes sense since the results obtained with the SVM model for the structural features were not analysed because of their low performances).

Two regions are selected from three different methods: regions *Left A22c* and *Left A24cd*. However, as can be seen in Table 6.4, the region *A22c* is involved in the auditory system. All the participants in this work are deaf but participants may have different auditory residuals in the low frequencies. One therefore needs to ensure that all the participants from different classes are comparable from an auditory point of view to assess the role of this region in the discrimination between vertigo and control participants.

The region *A24cd* is the anterior cingulate cortex which is involved in intention, inhibition, conflict or error handling [110]. The cingulate gyrus also plays an important role in visuo-spatial orientation by coordinating sensory inputs [94, 95]. Its

structural connection with the *Left A45c* region has been revealed to be significant as seen in Table 6.3. This feature shows a higher mean of fibers density for control participants (167) compared to vertigo participants who have compensated or not (3 and 11, respectively), suggesting an alteration of the structural connection between these two regions when one suffers from vertigo. Its functional connection with the *Left A22c* region has also been retained from both the SVM and RF models after the mutual information pre-filtering. However, as explained, the *A22c* region being involved in the auditory system, the comparison of patients' auditory abilities should be performed.

Then, several regions are selected by two different methods.

First, region *Right A5m* is selected one time in a structural feature and a second time in a functional feature. This region is in the precuneus which is a brain region involved in a variety of complex functions including recollection and memory, integration of information relating to perception of the environment, cue reactivity, mental imagery strategies, episodic memory retrieval, and affective responses to pain [111]. In the structural feature, it is connected to the *Left iOccG* region which plays a role in vision and perception, but as explained, this feature contains numerous null values and rather allows to discriminate vertigo participants who have compensated from control and vertigo participants. In the functional feature, it is connected to region *Left G* in the insular gyrus, showing a smaller simultaneous activity for vertigo participants as compared to control and compensated vertigo participants.

Region *A39rv* is also selected once in a structural feature in the right hemisphere and once in a functional feature in the left hemisphere. Region *A39rv* is in the inferior parietal lobule, involved in sensorimotor integration. For the structural feature, it is connected to the *Right rCunG* region which is in the occipital lobe and involved in visual perception. The mean fibers density of this feature is observed to be higher for control participants compared to vertigo ones, who contain numerous null values. For the functional feature, it is connected to the *Right A7ip* which is in the superior parietal lobe, closely linked with the occipital lobe. The connection shows a higher simultaneous activity for control participants compared to vertigo who have compensated or not.

Another region selected twice is the *A45c*, also once for the left hemisphere in a structural feature and once for the right hemisphere in a functional feature. This region belongs to the frontal lobe, involved in the planning of movement. More precisely, the right inferior frontal gyrus is involved in the unrelated-speech movements while the left one is responsible for the processing of related-speech gesture [112]. The structural connection of the latter with the *A24cd* region has already been mentioned above. The functional connection of the *Left A45c* region with the *Right A4ll*, involved in motor and sensory function related to the lower limb, has been selected from the SVM model. The results suggest that participants suffering from vertigo show a higher simultaneous functional activity compared to control ones.

Another region of the cingulate gyrus, the *A23v*, has also been selected by two methods. The SVM model with the correlation prefiltering method has selected the feature *Right A23v to Left A4ul*, the latter playing a role in the control of voluntary movements. The simultaneous functional activity of these regions are smaller for participants suffering from vertigo compared to control ones. On the other hand, the RF model with the correlation prefiltering has selected the *Left A23v to Left A4ul*, also showing a smaller simultaneous activity for participants suffering from vertigo compared to control ones.

Region *Right A14m*, which is in the orbital frontal gyrus, is selected two times in a functional feature. The first connection is with the region *Right OPC* in the occipital polar cortex which is involved in visual perception and the second one is with the region *Right dmPOS* in the precuneus, involved in spatial function and navigation. Both features show a smaller activity for vertigo participants who have compensated or not as compared to control ones.

Finally, region *dId* is selected two times in functional features. However, as can be seen in table L.4, it has significant connections with the auditory cortex and the comparability of patients from an auditory point of view must be checked as for region *A22c*.

Even if several regions show coherence with the involvement in vestibular system, biological and medical advice is still needed to better filter these data, to highlight those that are most relevant from a medical point of view and to confirm the consistence of the connections between those different regions.

6.4 Cerebellum

The significant regions identified by the statistical tests in section *Statistical cerebellum analysis* are *lobules IV* and *V* for the cortical thickness asymmetry. The right *lobule IV* has also a p-value smaller than 0.05 for the statistical test performed with the structural connectome on the betweenness centrality. The left *lobule V*, on the other hand, has a p-value smaller than 0.05 for the statistical test performed with the functional connectome on the degree centrality. Moreover, the functional connection between the right *lobule V* and the *Right A46* region has been revealed to be a significant feature to discriminate between classes. *Lobules IV* and *V* correspond to what is called the culmen. The culmen takes part to the spinocerebellum which is responsible for controlling the muscular activity of the posture by influencing the muscle tone. The spinocerebellum receives major inputs from the spinocerebellar tract and its output projects notably to the vestibulospinal tracts. It is involved in the integration of sensory input with motor commands to produce adaptive motor coordination [113].

The results observed for mean cortical thickness asymmetry of *lobules IV* and *V* (Figure 5.30) show that values for vertigo participants who have compensated are more closer to the values of vertigo participants rather than control ones.

6.5 Limitations

The work done in this thesis obviously suffers from some limitations. First, as previously explained, the threshold on null values to select the features before the pre-filtering method of correlation or mutual information was put arbitrarily to 60%. This may not give the best results and it also has an important influence on the results obtained.

Moreover, when using machine learning algorithms, the more data available the better, as seen in the *Machine learning* section. However, the number of available data for this thesis is small (72 participants) compared to the number of features, making it difficult for the machine learning model to achieve very good performances. Moreover, classes are not evenly distributed among the data and a difference between the mean age of each class is observed, as shown in Table 4.1. Although corrections were made to avoid influencing the results as much as possible, they may not be perfect.

Furthermore, as seen in the section *Connectome analysis for vestibular system*, a lateralisation of the vestibular system has already been proven by some researchers. The handedness varies between participants of this study: 5 participants are left-handed, 2 ambidextrous, 13 for which the information is not known and the rest of the participants are right-handed. It might therefore be possible that this difference in handedness alters the results obtained.

Finally, based on previous works, the absolute value of the functional connectivity entries was taken. Therefore, if two regions activate simultaneously positively for vertigo participants but negatively for control participants, this feature is not selected by the machine learning method to discriminate between classes as the regions become positively activated for all classes when taking the absolute value.

6.6 Future work

Based on what is done in this thesis and based on the work of Benjamin Chiem [18], it would be interesting to characterise vertigo based on multimodal connectivity data. The method would remain the same but input features would be composed of both structural and functional data. This would make it possible to know if vertigo disorders are best explained with structural data, functional data or a combination of both.

Moreover, to address the inequality between classes and the difficulty of accurately defining when patients have compensated for vertigo or not, the machine learning part could be applied with the results to the different vestibular tests as target and not the three classes 'control participants', 'vertigo participants' and 'vertigo participants who have compensated'.

Furthermore, the work can be extended to participants who are not deaf to avoid including the effect of possible different auditory abilities on the results.

Finally, two filter methods (the correlation and the mutual information) as well as two models of machine learning (the support vector machine and the random forest) are tested in this thesis. However, these are not the only ones and further work should test alternative classification and feature selection methods.

Chapter 7

Conclusion

Finding regions involved in the vestibular system, and in particular, finding features that differentiate people suffering from vertigo from control participants is a real challenge. The combination of tools such as magnetic resonance imaging, graph theory and machine learning is one way of identifying some regions among the wide range of possibilities.

In this work, a literature review of the methods employed and the regions of the vestibular system already identified has been performed. To follow, a description of the methodology used to identify brain regions has been given. Connectivity matrices, statistical tests and machine learning models such as Support Vector Machine and Random Forest were applied to magnetic resonance images of control participants, vertigo participants and vertigo participants who have compensated.

The goal of this thesis was to identify the features which best discriminate between the participants of the different classes. Although the methodology used presents limitations, some regions and features were identified as relevant. Anatomical connections between regions involved in cognitive, proprioceptive, and affective processing appear to be significant in discriminating control participants from vertigo participants. For both structural and functional features, regions located in the frontal lobe (such as the precentral gyrus, involved in motor control, and orbital gyrus, implicated in intellectual and emotional expression), in the limbic lobe (like the cingulate gyrus which coordinates sensory inputs, playing a role in visuo-spatial orientation) or the occipital lobe (principally involved in vision) are likely to play a role in the vestibular system and its performances. To a lesser extent, regions in the temporal lobe (such as the fusiform and middle temporal gyrus, playing a role in vision, and the superior temporal gyrus, involved in the processing of auditory information) and in the parietal lobe (such as the precuneus involved in the sense of self, autobiographical memory, spatial function, and navigation) also seem to contribute to vestibular processing.

Most of these findings are consistent with what can be found in the literature and once again highlight the involvement of various parts of the brain in the vestibular system. This should encourage further research in the study of functional and structural connections of the vestibular system so as to strengthen the results of this work. Combined with a supplementary biological input from doctors, these findings could help for the diagnosis and treatment of vertigo patients.

Bibliography

- [1] B Kastler et al. *Comprendre l'IRM*. 2003. ISBN: 2913495095.
- [2] Dricot Laurence Ir. “LGBIO1115: Introduction aux Neurosciences THEMATIQUE NEUROIMAGERIE”. In: *Course given at UCLouvain* (2019).
- [3] Frank Peeters. “LGBIO2050: Medical Imaging, PART2: MRI”. In: *Course given at UCLouvain* (2020).
- [4] John P. Ridgway. “Cardiovascular magnetic resonance physics for clinicians: Part i”. In: *Journal of Cardiovascular Magnetic Resonance* 12.1 (2010). ISSN: 10976647. DOI: [10.1186/1532-429X-12-71](https://doi.org/10.1186/1532-429X-12-71).
- [5] Jeremy Jones and Sarah Li. “Eddy currents”. In: *Radiopaedia.org* (Aug. 2017). DOI: [10.53347/RID-55242](https://doi.org/10.53347/RID-55242). URL: <http://radiopaedia.org/articles/55242>.
- [6] Rutger Fick. “Advanced dMRI Signal Modeling for Tissue Microstructure Characterization”. PhD thesis. 2017.
- [7] T. A.G.M. Huisman. “Diffusion-weighted and diffusion tensor imaging of the brain, made easy”. In: *Cancer Imaging* 10.1A (2010), S163. ISSN: 14707330. DOI: [10.1102/1470-7330.2010.9023](https://doi.org/10.1102/1470-7330.2010.9023). URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2967146/>.
- [8] Yahya Baba and Mohammad Niknejad. “Apparent diffusion coefficient”. In: *Radiopaedia.org* (Feb. 2013). DOI: [10.53347/RID-21759](https://doi.org/10.53347/RID-21759).
- [9] Eleftherios Garyfallidis. “Towards an Accurate Brain Tractography”. In: May (2012), pp. 1–181.
- [10] *DIPY : Docs 1.5.0 - Introduction to Basic Tracking*. Last accessed: 2022-05-11. URL: https://dipy.org/documentation/1.5.0/examples_built/tracking_introduction_eudx/#example-tracking-introduction-eudx.
- [11] Derek K. Jones and Alexander Leemans. *Diffusion tensor imaging*. Vol. 711. 2010, pp. 127–144. ISBN: 9781617379918. DOI: [10.1007/978-1-61737-992-5_6](https://doi.org/10.1007/978-1-61737-992-5_6).
- [12] Christopher Bergland. *What Is the Human Connectome Project? Why Should You Care? / Psychology Today*. Psychology Today. 2013. URL: <https://www.psychologytoday.com/us/blog/the-athletes-way/201311/what-is-the-human-connectome-project-why-should-you-care>.
- [13] *DIPY : Docs 1.5.0 - An introduction to the Probabilistic Direction Getter*. Last accessed: 2022-05-07. URL: https://dipy.org/documentation/1.5.0/examples_built/tracking_probabilistic/#example-tracking-probabilistic.
- [14] Olaf Sporns. *Networks of the brain*. 2011. ISBN: 9780262014694.

-
- [15] Danielle S. Bassett and Edward T. Bullmore. “Small-World Brain Networks Revisited”. In: *The Neuroscientist : a review journal bringing neurobiology, neurology and psychiatry* 23.5 (Oct. 2017), pp. 499–516. ISSN: 1089-4098. DOI: [10.1177/1073858416667720](https://doi.org/10.1177/1073858416667720). URL: <https://pubmed.ncbi.nlm.nih.gov/27655008/>.
- [16] Ed Bullmore and Olaf Sporns. “Complex brain networks: Graph theoretical analysis of structural and functional systems”. In: *Nature Reviews Neuroscience* 10.3 (2009), pp. 186–198. ISSN: 1471003X. DOI: [10.1038/nrn2575](https://doi.org/10.1038/nrn2575).
- [17] Alex Fornito, Andrew Zalesky, and Edward Bullmore. *Fundamentals of Brain Network Analysis*. 2016.
- [18] Benjamin Chiêm. “Thesis: Interweaving Structure and Function in Human Brain Networks”. PhD thesis.
- [19] *Betweenness Centrality - Neo4j Graph Data Science*. Last accessed: 2022-04-20. URL: <https://neo4j.com/docs/graph-data-science/current/algorithms/betweenness-centrality/>.
- [20] Issam El Naqa and Martin J. Murphy. “What Is Machine Learning?” In: *Machine Learning in Radiation Oncology* (2015), pp. 3–11. DOI: [10.1007/978-3-319-18305-3_1](https://doi.org/10.1007/978-3-319-18305-3_1). URL: https://link.springer.com/chapter/10.1007/978-3-319-18305-3_1.
- [21] *Tous les modèles de Machine Learning expliqués en 8 minutes*. Last accessed: 2022-05-11. URL: <https://moncoachdata.com/blog/modeles-de-machine-learning-expliques/>.
- [22] Professor Michel Verleysen Professor John Lee. “LELEC2870 Machine learning : regression, deep networks and dimensionality reduction”. In: *Course given at UCLouvain* (2021).
- [23] A. Jović, K. Brkić, and N. Bogunović. “A review of feature selection methods with applications”. In: *2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2015 - Proceedings* (July 2015), pp. 1200–1205. DOI: [10.1109/MIPRO.2015.7160458](https://doi.org/10.1109/MIPRO.2015.7160458).
- [24] V Kecman. “Support Vector Machines – An Introduction BT - Support Vector Machines: Theory and Applications”. In: *Support Vector Machines: Theory and Applications* 47 (2005), pp. 1–47. URL: http://dx.doi.org/10.1007/10984697_1.
- [25] *1.4. Support Vector Machines — scikit-learn 1.1.0 documentation*. Last accessed: 2022-05-12. URL: <https://scikit-learn.org/stable/modules/svm.html>.
- [26] *Support Vector Machine (SVM) Algorithm - Javatpoint*. Last accessed: 2022-05-12. URL: <https://www.javatpoint.com/machine-learning-support-vector-machine-algorithm>.

- [27] *Tree-Based Algorithms — Decision Tree and Random Forest* / by İbrahim Halil Kaplan / *Python in Plain English*. URL: <https://python.plainenglish.io/tree-based-algorithms-decision-tree-and-random-forest-cb5c5ccaf43b>.
- [28] S. Suthaharan. “A Cognitive Random Forest: An Intra- and Intercognitive Computing for Big Data Classification Under Cune Condition”. In: *Handbook of Statistics* 35 (Jan. 2016), pp. 207–227. ISSN: 0169-7161. DOI: [10.1016/BS.HOST.2016.07.006](https://doi.org/10.1016/BS.HOST.2016.07.006).
- [29] *Random Forest Algorithm Advantages and Disadvantages - Dhiraj K - Medium*. Last accessed: 2022-05-12. URL: <https://dhirajkumarblog.medium.com/random-forest-algorithm-advantages-and-disadvantages-1ed22650c84f>.
- [30] Professor Deggouj Naima et Al. “WMDS2215: Secteur tête et cou (ophtalmologie, oto-rhino-laryngologie, stomatologie)”. In: *Course given at UCLouvain* (2019).
- [31] Werner M Graf. “Evolution of the Vestibular System”. In: *Encyclopedia of Neuroscience*. Ed. by Marc D Binder, Nobutaka Hirokawa, and Uwe Windhorst. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 1440–1448. ISBN: 978-3-540-29678-2. DOI: [10.1007/978-3-540-29678-2_3175](https://doi.org/10.1007/978-3-540-29678-2_3175). URL: http://link.springer.com/10.1007/978-3-540-29678-2_3175.
- [32] Walter F. Boron Emile L. Boulpaep. *Vestibular and Auditory Transduction: Hair Cells - Sensory Transduction - The Nervous System - Medical Physiology, 3rd Edition*. Last accessed: 2022-04-07. 2016. URL: <https://doctorlib.info/physiology/medical/83.html>.
- [33] Martha C. Johnson and Mary L. Cutler. *Anatomy and physiology of the breast*. 2010. 2016, pp. 1–39. ISBN: 9783319463568. DOI: [10.1007/978-3-319-46356-8_1](https://doi.org/10.1007/978-3-319-46356-8_1).
- [34] Emilie Lacroix. “Thesis: A neuropsychological approach to better understand cognition in adult and child patients with vestibular loss”. In: (2018).
- [35] *INSTANT NEURO - Vestibular Pathways*. Last accessed: 2022-04-08. URL: <https://www.youtube.com/watch?v=cXhe20THed8>.
- [36] Donald W. Pfaff and Nora D. Volkow. *Neuroscience in the 21st century: From basic to clinical, second edition*. 2016, pp. 1–4155. ISBN: 9781493934744. DOI: [10.1007/978-1-4939-3474-4](https://doi.org/10.1007/978-1-4939-3474-4).
- [37] Arthur N. Popper Stephen M. Highstein Richard R. Fay. *The vestibular system*. 2003. ISBN: 0387983147.
- [38] *Auditory and Vestibular Systems (sensory System) Part 3*. Last accessed: 2022-04-08. URL: <http://what-when-how.com/neuroscience/auditory-and-vestibular-systems-sensory-system-part-3/>.
- [39] T. M. Raiser et al. “The human corticocortical vestibular network”. In: *NeuroImage* 223 (Dec. 2020), p. 117362. ISSN: 1053-8119. DOI: [10.1016/J.NEUROIMAGE.2020.117362](https://doi.org/10.1016/J.NEUROIMAGE.2020.117362).

- [40] *vHIT (Video Head Impulse Test)* / Polyclinique Centre-Ville. Last accessed: 2022-04-06. URL: <https://pcvmontreal.com/tests-en-clinique/examens/equilibre-et-vertige/vhit-video-head-impulse-test>.
- [41] *Videonystagmography (VNG): MedlinePlus Medical Test*. Last accessed: 2022-04-06. URL: <https://medlineplus.gov/lab-tests/videonystagmography-vng/>.
- [42] *Les différents examens vestibulaires pour le diagnostic des vertiges*. Last accessed: 2022-04-28. URL: <https://www.lesvertiges.com/oreille-interne/diagnostic-vestibulaire.html>.
- [43] Jeremy L. Smith et al. “The “vestibular neuromatrix”: A proposed, expanded vestibular network from graph theory in post-concussive vestibular dysfunction”. In: *Human Brain Mapping* 43.5 (2022), pp. 1501–1518. ISSN: 10970193. DOI: [10.1002/hbm.25737](https://doi.org/10.1002/hbm.25737).
- [44] Anna Trofimova et al. “Alterations in Resting-State Functional Brain Connectivity and Correlations with Vestibular/Ocular-Motor Screening Measures in Postconcussion Vestibular Dysfunction”. In: *Journal of neuroimaging : official journal of the American Society of Neuroimaging* 31.2 (Mar. 2021), pp. 277–286. ISSN: 1552-6569. DOI: [10.1111/JON.12834](https://doi.org/10.1111/JON.12834). URL: <https://pubmed.ncbi.nlm.nih.gov/33476477/>.
- [45] Iole Indovina et al. “Structural connectome and connectivity lateralization of the multimodal vestibular cortical network”. In: *NeuroImage* 222 (Nov. 2020), p. 117247. ISSN: 1053-8119. DOI: [10.1016/j.neuroimage.2020.117247](https://doi.org/10.1016/j.neuroimage.2020.117247).
- [46] Anna M. Wirth et al. “White Matter Connectivity of the Visual-Vestibular Cortex Examined by Diffusion-Weighted Imaging”. In: 8.4 (May 2018), pp. 235–244. ISSN: 21580022. DOI: [10.1089/BRAIN.2017.0544](https://doi.org/10.1089/BRAIN.2017.0544). URL: <https://www.liebertpub.com/doi/abs/10.1089/brain.2017.0544>.
- [47] Sebastian M. Frank and Mark W. Greenlee. “The parieto-insular vestibular cortex in humans: More than a single area?” In: *Journal of Neurophysiology* 120.3 (2018), pp. 1438–1450. ISSN: 15221598. DOI: [10.1152/jn.00907.2017](https://doi.org/10.1152/jn.00907.2017).
- [48] V. Kirsch et al. “Structural and functional connectivity mapping of the vestibular circuitry from human brainstem to cortex”. In: *Brain Structure and Function* 221.3 (Apr. 2016), pp. 1291–1308. ISSN: 18632661. DOI: [10.1007/S00429-014-0971-X](https://doi.org/10.1007/S00429-014-0971-X). URL: <https://link.springer.com/article/10.1007/s00429-014-0971-x>.
- [49] *Brainnetome Atlas / Home*. Last accessed: 2022-05-06. URL: <http://atlas.brainnetome.org/>.
- [50] J. Diedrichsen. *A spatially unbiased atlas template of the human cerebellum*. Neuroimage. 2006. URL: https://www.diedrichsenlab.org/imaging/suit_download.htm.
- [51] *NITRC: MRICron: Tool/Resource Info*. Last accessed: 2022-05-15. URL: <https://www.nitrc.org/projects/mricron/>.

- [52] *Introduction to ElikoPy — ElikoPy documentation*. Last accessed: 2022-04-29. URL: <https://elikopy.readthedocs.io/en/latest/introduction.html>.
- [53] *CECI*. Last accessed: 2022-05-06. URL: <https://www.cec-hpc.be/clusters.html>.
- [54] Jian Yang et al. “Brain MR image denoising for Rician noise using pre-smooth non-local means filter”. In: *BioMedical Engineering Online* 14.1 (Jan. 2015), pp. 1–20. ISSN: 1475925X. DOI: [10.1186/1475-925X-14-2](https://doi.org/10.1186/1475-925X-14-2). URL: <https://biomedical-engineering-online.biomedcentral.com/articles/10.1186/1475-925X-14-2>.
- [55] Noppadol Chumchob. “A robust affine image registration method”. In: *International Journal of Numerical Analysis and Modeling* 6 (2009), pp. 311–334.
- [56] *Registration (aka spatial normalization) — Introduction to MRI*. Last accessed: 2022-05-07. URL: http://jpeelle.net/mri/image_processing/registration.html.
- [57] *DIPY : Docs 1.5.0 - Symmetric Diffeomorphic Registration in 3D*. Last accessed: 2022-05-15. URL: https://dipy.org/documentation/1.5.0/examples_built/syn_registration_3d/#example-syn-registration-3d.
- [58] *FSL_HCP1065_FA_1mm.nii.gz*. Last accessed: 2022-05-15. URL: <https://balsa.wustl.edu/file/mDB4X>.
- [59] Ben Jeurissen et al. “Diffusion MRI fiber tractography of the brain”. In: *NMR in biomedicine* 32.4 (Apr. 2019). ISSN: 1099-1492. DOI: [10.1002/NBM.3785](https://doi.org/10.1002/NBM.3785). URL: <https://pubmed.ncbi.nlm.nih.gov/28945294/>.
- [60] *DIPY : Docs 1.5.0 - An introduction to the Deterministic Maximum Direction Getter*. Last accessed: 2022-05-07. URL: https://dipy.org/documentation/1.5.0/examples_built/tracking_deterministic/#example-tracking-deterministic.
- [61] G. Russell Glenn et al. “Quantitative Assessment of Diffusional Kurtosis Anisotropy”. In: *NMR in biomedicine* 28.4 (Apr. 2015), p. 448. ISSN: 10991492. DOI: [10.1002/NBM.3271](https://doi.org/10.1002/NBM.3271). URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4378654/>.
- [62] *TrackVis*. Last accessed: 2022-05-16. URL: <http://trackvis.org/>.
- [63] Zijin Gu et al. “Heritability and interindividual variability of regional structure-function coupling”. In: *Nature Communications* 2021 12:1 12.1 (Aug. 2021), pp. 1–12. ISSN: 2041-1723. DOI: [10.1038/s41467-021-25184-4](https://doi.org/10.1038/s41467-021-25184-4). URL: <https://www.nature.com/articles/s41467-021-25184-4>.
- [64] Leila Cammoun et al. “Mapping the human connectome at multiple scales with diffusion spectrum MRI”. In: *Journal of Neuroscience Methods* 203.2 (Jan. 2012), pp. 386–397. ISSN: 01650270. DOI: [10.1016/J.JNEUMETH.2011.09.031](https://doi.org/10.1016/J.JNEUMETH.2011.09.031).

- [65] Leonardo Gutiérrez-Gómez et al. “Stable biomarker identification for predicting schizophrenia in the human connectome”. In: *NeuroImage: Clinical* 27.June (2020), p. 102316. ISSN: 22131582. DOI: [10.1016/j.nicl.2020.102316](https://doi.org/10.1016/j.nicl.2020.102316). URL: <https://doi.org/10.1016/j.nicl.2020.102316>.
- [66] *DIPY : Docs 1.5.0 - Connectivity Matrices, ROI Intersections and Density Maps*. Last accessed: 2022-05-15. URL: https://dipy.org/documentation/1.5.0/examples_built/streamline_tools/#example-streamline-tools.
- [67] Mark Jenkinson. “FSL Pre-Processing Pipeline FSL Pre-Processing Pipeline”. In: *OHBM 2016 Educational Course Presentation* (2016). URL: https://www.humanbrainmapping.org/files/2015/Ed%20Materials/FSL_PreProcessing_Pipeline_OHBM15_Jenkinson.pdf.
- [68] Mark Jenkinson et al. “Improved optimization for the robust and accurate linear registration and motion correction of brain images”. In: *NeuroImage* 17.2 (2002), pp. 825–841. ISSN: 1053-8119. DOI: [10.1016/S1053-8119\(02\)91132-8](https://doi.org/10.1016/S1053-8119(02)91132-8). URL: <https://pubmed.ncbi.nlm.nih.gov/12377157/>.
- [69] *MCFLIRT - FslWiki*. Last accessed: 2022-05-08. URL: <https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/MCFLIRT>.
- [70] “fMRI data preprocessing and Functional connectivity”. PhD thesis. URL: https://www.da.ugent.be/school2014/pages/en/Presentations/tutorial_preprocessing_fMRI_Roma.pdf.
- [71] David B. Parker and Qolamreza R. Razlighi. “The benefit of slice timing correction in common fMRI preprocessing pipelines”. In: *Frontiers in Neuroscience* 13.JUL (2019), p. 821. ISSN: 1662453X. DOI: [10.3389/FNINS.2019.00821](https://doi.org/10.3389/FNINS.2019.00821).
- [72] *Slicetimer*. Last accessed: 2022-05-08. URL: <https://poc.vl-e.nl/distribution/manual/fsl-3.2/slicetimer/index.html>.
- [73] *BrainVoyager User’s Guide: Temporal High-Pass Filtering*. Last accessed: 2022-04-29. URL: <https://www.brainvoyager.com/bv/doc/UsersGuide/Preprocessing/TemporalHighPassFiltering.html>.
- [74] *Fslutils - FslWiki*. Last accessed: 2022-05-09. URL: <https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Fslutils>.
- [75] *Smoothing — Introduction to MRI*. Last accessed: 2022-05-15. URL: http://jpeelle.net/mri/image_processing/smoothing.html.
- [76] Peng Liu, Vince Calhoun, and Zikuan Chen. “Functional overestimation due to spatial smoothing of fMRI data”. In: *Journal of Neuroscience Methods* 291 (Nov. 2017), pp. 1–12. ISSN: 1872678X. DOI: [10.1016/J.JNEUMETH.2017.08.003](https://doi.org/10.1016/J.JNEUMETH.2017.08.003).
- [77] *MNI space average T1-weighted MRI of the brain with anatomical labels*. Last accessed: 2022-05-02. URL: https://figshare.com/articles/dataset/MNI_space_average_T1_weighted_MRI_of_the_brain_with_anatomical_labels/2066037.

- [78] *Nilearn: Statistical Analysis and Machine learning for NeuroImaging in Python*. Last accessed: 2022-05-07. URL: <https://nilearn.github.io/modules/generated/nilearn.connectome.ConnectivityMeasure.html#nilearn.connectome.ConnectivityMeasure>.
- [79] Jiao Li et al. “Exploring the functional connectome in white matter”. In: *Human Brain Mapping* 40.15 (Oct. 2019), pp. 4331–4344. ISSN: 1097-0193. DOI: [10.1002/hbm.24705](https://doi.org/10.1002/hbm.24705). URL: <https://onlinelibrary.wiley.com/doi/full/10.1002/hbm.24705>.
- [80] Xi Nian Zuo et al. “Network Centrality in the Human Functional Connectome”. In: *Cerebral Cortex* 22.8 (Aug. 2012), pp. 1862–1875. ISSN: 1047-3211. DOI: [10.1093/CERCOR/BHR269](https://doi.org/10.1093/CERCOR/BHR269). URL: <https://academic.oup.com/cercor/article/22/8/1862/321860>.
- [81] Corey Horien et al. “The individual functional connectome is unique and stable over months to years”. In: *NeuroImage* 189 (Apr. 2019), pp. 676–687. ISSN: 1095-9572. DOI: [10.1016/J.NEUROIMAGE.2019.02.002](https://doi.org/10.1016/J.NEUROIMAGE.2019.02.002). URL: <https://pubmed.ncbi.nlm.nih.gov/30721751/>.
- [82] *Fisher Z-Transformation - Statistics How To*. Last accessed: 2022-05-15. URL: <https://www.statisticshowto.com/fisher-z/>.
- [83] *degree_centrality — NetworkX 2.8 documentation*. Last accessed: 2022-05-14. URL: https://networkx.org/documentation/stable/reference/algorithms/generated/networkx.algorithms centrality.degree_centrality.html.
- [84] *betweenness_centrality — NetworkX 2.8 documentation*. Last accessed: 2022-05-14. URL: https://networkx.org/documentation/stable/reference/algorithms/generated/networkx.algorithms centrality.betweenness_centrality.html#networkx.algorithms centrality.betweenness_centrality.
- [85] Ricco Rakotomalala. “Tests de normalité: Techniques empiriques et tests statistiques”. 2011.
- [86] *ANOVA dans R: Excellente Référence - Datanovia*. Last accessed: 2022-05-07. URL: <https://www.datanovia.com/en/fr/lessons/anova-dans-r/>.
- [87] *Le test des variances de Bartlett – lemakistatheux*. Last accessed: 2022-05-16. URL: <https://lemakistatheux.wordpress.com/2013/05/31/le-test-de-bartlett/>.
- [88] *Multiple comparisons - Handbook of Biological Statistics*. Last accessed: 2022-05-08. URL: <http://www.biostathandbook.com/multiplecomparisons.html>.
- [89] Fatemeh Noohi et al. “Deactivation of somatosensory and visual cortices during vestibular stimulation is associated with older age and poorer balance”. In: *PLoS ONE* 14.9 (2019), pp. 1–18. ISSN: 19326203. DOI: [10.1371/journal.pone.0221954](https://doi.org/10.1371/journal.pone.0221954).

- [90] Farshad Falahati et al. “The Effect of Age Correction on Multivariate Classification in Alzheimer’s Disease, with a Focus on the Characteristics of Incorrectly and Correctly Classified Subjects”. In: *Brain Topography* 29.2 (Mar. 2016), pp. 296–307. ISSN: 15736792. DOI: [10.1007/S10548-015-0455-1](https://doi.org/10.1007/S10548-015-0455-1). URL: <https://link.springer.com/article/10.1007/s10548-015-0455-1>.
- [91] Ludmila I Kuncheva. *A STABILITY INDEX FOR FEATURE SELECTION*. Last accessed: 2022-05-15. URL: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.101.6458&rep=rep1&type=pdf>.
- [92] José V. Manjón and Pierrick Coupé. “Volbrain: An online MRI brain volumetry system”. In: *Frontiers in Neuroinformatics* 10.JUL (July 2016). ISSN: 16625196. DOI: [10.3389/FNINF.2016.00030](https://doi.org/10.3389/FNINF.2016.00030). URL: <https://www.volbrain.upv.es/index.php>.
- [93] I. Buciu, C. Kotropoulos, and I. Pitas. “Demonstrating the stability of support vector machines for classification”. In: *Signal Processing* 86.9 (2006), pp. 2364–2380. ISSN: 01651684. DOI: [10.1016/j.sigpro.2005.11.005](https://doi.org/10.1016/j.sigpro.2005.11.005).
- [94] *Le gyrus cingulaire : ses différentes parties et fonctions - Nos Pensées*. URL: <https://nospensees.fr/le-gyrus-cingulaire-ses-differentes-parties-et-fonctions/>.
- [95] Rachel H. Tan et al. “The pathogenesis of cingulate atrophy in behavioral variant frontotemporal dementia and Alzheimer’s disease”. In: *Acta Neuropathologica Communications* 2.1 (Jan. 2014). ISSN: 20515960. DOI: [10.1186/2051-5960-1-30](https://doi.org/10.1186/2051-5960-1-30).
- [96] Rami M. El-Baba and Mark P. Schury. “Neuroanatomy, Frontal Cortex”. In: *StatPearls* (June 2021). URL: <https://www.ncbi.nlm.nih.gov/books/NBK554483/>.
- [97] Svenja Caspers et al. “Organization of the Human Inferior Parietal Lobule Based on Receptor Architectonics”. In: *Cerebral Cortex* 23.3 (Mar. 2013), pp. 615–628. ISSN: 1047-3211. DOI: [10.1093/CERCOR/BHS048](https://doi.org/10.1093/CERCOR/BHS048). URL: <https://academic.oup.com/cercor/article/23/3/615/315396>.
- [98] Marianne Dieterich et al. “Dominance for Vestibular Cortical Function in the Non-dominant Hemisphere”. In: *Cerebral Cortex* 13.9 (Sept. 2003), pp. 994–1007. ISSN: 1047-3211. DOI: [10.1093/CERCOR/13.9.994](https://doi.org/10.1093/CERCOR/13.9.994). URL: <https://academic.oup.com/cercor/article/13/9/994/342646>.
- [99] Paul Johns. “Functional neuroanatomy”. In: *Clinical Neuroscience* (2014), pp. 27–47. DOI: [10.1016/B978-0-443-10321-6.00003-5](https://doi.org/10.1016/B978-0-443-10321-6.00003-5).
- [100] Ole Numssen, Danilo Bzdok, and Gesa Hartwigsen. “Functional specialization within the inferior parietal lobes across cognitive domains”. In: *eLife* 10 (Mar. 2021). ISSN: 2050084X. DOI: [10.7554/ELIFE.63591](https://doi.org/10.7554/ELIFE.63591).
- [101] J. Zevin. “Word Recognition”. In: *Encyclopedia of Neuroscience* (2009), pp. 517–522. DOI: [10.1016/B978-008045046-9.01881-7](https://doi.org/10.1016/B978-008045046-9.01881-7).

- [102] *Neuroanatomy Online: Lab 1 - Overview of the Nervous System - Telencephalon - A Closer Look*. URL: https://nba.uth.tmc.edu/neuroanatomy/L1/Lab01p10_index.html.
- [103] *Neuroanatomy Online: Lab 1 - Overview of the Nervous System - Central Lobes & Frontal Lobes - Inferior View*. URL: https://nba.uth.tmc.edu/neuroanatomy/L1/Lab01p11_index.html.
- [104] Brian D. Loftus, Sudhir S. Athni, and Igor M. Chermes. “Clinical Neuroanatomy”. In: *Neurology Secrets* (2010), pp. 18–54. DOI: [10.1016/B978-0-323-05712-7.00002-7](https://doi.org/10.1016/B978-0-323-05712-7.00002-7).
- [105] Yang Zhang et al. “The Roles of Subdivisions of Human Insula in Emotion Perception and Auditory Processing”. In: *Cerebral Cortex* 29.2 (Feb. 2019), pp. 517–528. ISSN: 1047-3211. DOI: [10.1093/CERCOR/BHX334](https://doi.org/10.1093/CERCOR/BHX334). URL: <https://academic.oup.com/cercor/article/29/2/517/4802217>.
- [106] Edmund T. Rolls. “The neuroscience of emotional disorders”. In: *Handbook of Clinical Neurology* 183 (Jan. 2021), pp. 1–26. ISSN: 22124152. DOI: [10.1016/B978-0-12-822290-4.00002-5](https://doi.org/10.1016/B978-0-12-822290-4.00002-5).
- [107] Shruti Japee et al. “A role of right middle frontal gyrus in reorienting of attention: A case study”. In: *Frontiers in Systems Neuroscience* 9.MAR (Mar. 2015). ISSN: 16625137. DOI: [10.3389/FNSYS.2015.00023](https://doi.org/10.3389/FNSYS.2015.00023). URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4347607/>.
- [108] Jay M. Goldberg et al. “The Vestibular System: A Sixth Sense”. In: *The Vestibular System: A Sixth Sense* (May 2012), pp. 1–560. DOI: [10.1093/ACPROF:OS0/9780195167085.001.0001](https://doi.org/10.1093/ACPROF:OS0/9780195167085.001.0001).
- [109] Toshiaki Onitsuka et al. “Middle and Inferior Temporal Gyrus Gray Matter Volume Abnormalities in Chronic Schizophrenia: An MRI Study”. In: *The American journal of psychiatry* 161.9 (2004), p. 1603. ISSN: 0002953X. DOI: [10.1176/APPI.AJP.161.9.1603](https://doi.org/10.1176/APPI.AJP.161.9.1603). URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2793337/>.
- [110] *Le cortex préfrontal et la dopamine striatale dans l'apprentissage guidé par la récompense*. Last accessed: 2022-05-29. URL: <https://tel.archives-ouvertes.fr/tel-00160868/document>.
- [111] David Borsook, Nasim Maleki, and Rami Burstein. “Migraine”. In: *Neurobiology of Brain Disorders: Biological Basis of Neurological and Psychiatric Disorders* (2015), pp. 693–708. DOI: [10.1016/B978-0-12-398270-4.00042-2](https://doi.org/10.1016/B978-0-12-398270-4.00042-2).
- [112] Miriam Steines et al. “The role of the left and right inferior frontal gyrus in processing metaphoric and unrelated co-speech gestures”. In: *NeuroImage* 237 (Aug. 2021). ISSN: 1095-9572. DOI: [10.1016/J.NEUROIMAGE.2021.118182](https://doi.org/10.1016/J.NEUROIMAGE.2021.118182). URL: <https://pubmed.ncbi.nlm.nih.gov/34020020/>.

- [113] *Cerebellum (Section 3, Chapter 5) Neuroscience Online: An Electronic Textbook for the Neurosciences / Department of Neurobiology and Anatomy - The University of Texas Medical School at Houston*. Last accessed: 2022-05-21. URL: <https://nba.uth.tmc.edu/neuroscience/m/s3/chapter05.html>.
- [114] *Morphologie - Cortex - Aires de Brodman*. Last accessed: 2022-05-28. URL: <http://www.chups.jussieu.fr/ext/neuranat/morphologie/cortex/cortex03.html>.
- [115] David J. Scott et al. “Individual Differences in Reward Responding Explain Placebo-Induced Expectations and Effects”. In: *Neuron* 55.2 (2007), pp. 325–336. ISSN: 08966273. DOI: [10.1016/j.neuron.2007.06.028](https://doi.org/10.1016/j.neuron.2007.06.028).
- [116] Isabel Schwienbacher et al. “Temporary inactivation of the nucleus accumbens disrupts acquisition and expression of fear-potentiated startle in rats”. In: *Brain Research* 1027.1-2 (2004), pp. 87–93. ISSN: 00068993. DOI: [10.1016/j.brainres.2004.08.037](https://doi.org/10.1016/j.brainres.2004.08.037).
- [117] Andrea E. Cavanna and Michael R. Trimble. “The precuneus: a review of its functional anatomy and behavioural correlates”. In: *Brain : a journal of neurology* 129.Pt 3 (2006), pp. 564–583. ISSN: 1460-2156. DOI: [10.1093/BRAIN/AWL004](https://doi.org/10.1093/BRAIN/AWL004). URL: <https://pubmed.ncbi.nlm.nih.gov/16399806/>.
- [118] Grace Carpenter and Dayu Gai. “Paracentral lobule”. In: *Radiopaedia.org* (Aug. 2015). DOI: [10.53347/RID-39075](https://doi.org/10.53347/RID-39075).
- [119] Krishnagopal Dharani. “Functional Anatomy of the Brain”. In: *The Biology of Thought* (2015), pp. 3–29. DOI: [10.1016/B978-0-12-800900-0.00001-4](https://doi.org/10.1016/B978-0-12-800900-0.00001-4).
- [120] Svenja Caspers et al. “Organization of the Human Inferior Parietal Lobule Based on Receptor Architectonics”. In: *Cerebral Cortex* 23.3 (Mar. 2013), pp. 615–628. ISSN: 1047-3211. DOI: [10.1093/CERCOR/BHS048](https://doi.org/10.1093/CERCOR/BHS048). URL: <https://academic.oup.com/cercor/article/23/3/615/315396>.

Appendix A

Labels of the atlas

Table A.1: Region's names of the left hemisphere with their corresponding label.

Region's name	Label	Region's name	Label	Region's name	Label
Left_A8m	1	Left_A22c	38	Left_A5m	75
Left_A8dl	2	Left_A38l	39	Left_dmPOS	76
Left_A9l	3	Left_A22r	40	Left_A31	77
Left_A6dl	4	Left_A21c	41	Left_A1/2/3ulhf	78
Left_A6m	5	Left_A21r	42	Left_A1/2/3tonIa	79
Left_A9m	6	Left_A37dl	43	Left_A2	80
Left_A10m	7	Left_aSTS	44	Left_A1/2/3tru	81
Left_A9/46d	8	Left_A20iv	45	Left_G	82
Left_IFJ	9	Left_A37elv	46	Left_vIa	83
Left_A46	10	Left_A20r	47	Left_dIa	84
Left_A9/46v	11	Left_A20il	48	Left_vId/vIg	85
Left_A8vl	12	Left_A37vl	49	Left_dIg	86
Left_A6vl	13	Left_A20cl	50	Left_dId	87
Left_A10l	14	Left_A20cv	51	Left_A23d	88
Left_A44d	15	Left_A20rv	52	Left_A24rv	89
Left_IFS	16	Left_A37mv	53	Left_A32p	90
Left_A45c	17	Left_A37lv	54	Left_A23v	91
Left_A45r	18	Left_A35/36r	55	Left_A24cd	92
Left_A44op	19	Left_A35/36c	56	Left_A23c	93
Left_A44v	20	Left_TL	57	Left_A32sg	94
Left_A14m	21	Left_A28/34	58	Left_cLinG	95
Left_A12/47o	22	Left_TI	59	Left_rCunG	96
Left_A11l	23	Left_TH	60	Left_cCunG	97
Left_A11m	24	Left_rpSTS	61	Left_rLinG	98
Left_A13	25	Left_cpSTS	62	Left_vmPOS	99
Left_A12/47l	26	Left_A7r	63	Left_mOccG	100
Left_A4hf	27	Left_A7c	64	Left_V5/MT+	101
Left_A6cdl	28	Left_A5l	65	Left_OPC	102
Left_A4ul	29	Left_A7pc	66	Left_iOccG	103
Left_A4t	30	Left_A7ip	67	Left_msOccG	104
Left_A4tl	31	Left_A39c	68	Left_lsOccG	105
Left_A6cvl	32	Left_A39rd	69	Left_mAmyg	106
Left_A1/2/3ll	33	Left_A40rd	70	Left_lAmyg	107
Left_A4ll	34	Left_A40c	71	Left_rHipp	108
Left_A38m	35	Left_A39rv	72	Left_cHipp	109
Left_A41/42	36	Left_A40rv	73	Left_vCa	110
Left_TE1.0/TE1.2	37	Left_A7m	74	Left_GP	111

APPENDIX A. LABELS OF THE ATLAS

Table A.2: Region's names of the left hemisphere with their corresponding label.

Region's name	Label	Region's name	Label	Region's name	Label
Left_NAC	112	Left_Otha	121	Left_VIIIa	130
Left_vmPu	113	Left_cTtha	122	Left_VIIIb	131
Left_dCa	114	Left_lPFtha	123	Left_IX	132
Left_dIPu	115	Left_I_IV	124	Left_X	133
Left_mPFtha	116	Left_V	125	Left_Dentate	134
Left_mPMtha	117	Left_VI	126	Left_Interposed	135
Left_Stha	118	Left_CrusI	127	Left_Fastigial	136
Left_rTtha	119	Left_CrusII	128		
Left_PPtha	120	Left_VIIb	129		

Table A.3: Region's names of the right hemisphere with their corresponding label.

Region's name	Label	Region's name	Label	Region's name	Label
Right_A8m	137	Right_A6cdl	164	Right_A35/36r	191
Right_A8dl	138	Right_A4ul	165	Right_A35/36c	192
Right_A9l	139	Right_A4t	166	Right_TL	193
Right_A6dl	140	Right_A4tl	167	Right_A28/34	194
Right_A6m	141	Right_A6cvl	168	Right_TI	195
Right_A9m	142	Right_A1/2/3ll	169	Right_TH	196
Right_A10m	143	Right_A4ll	170	Right_rpSTS	197
Right_A9/46d	144	Right_A38m	171	Right_cpSTS	198
Right_IFJ	145	Right_A41/42	172	Right_A7r	199
Right_A46	146	Right_TE1.0/TE1.2	173	Right_A7c	200
Right_A9/46v	147	Right_A22c	174	Right_A5l	201
Right_A8vl	148	Right_A38l	175	Right_A7pc	202
Right_A6vl	149	Right_A22r	176	Right_A7ip	203
Right_A10l	150	Right_A21c	177	Right_A39c	204
Right_A44d	151	Right_A21r	178	Right_A39rd	205
Right_IFS	152	Right_A37dl	179	Right_A40rd	206
Right_A45c	153	Right_aSTS	180	Right_A40c	207
Right_A45r	154	Right_A20iv	181	Right_A39rv	208
Right_A44op	155	Right_A37elv	182	Right_A40rv	209
Right_A44v	156	Right_A20r	183	Right_A7m	210
Right_A14m	157	Right_A20il	184	Right_A5m	211
Right_A12/47o	158	Right_A37vl	185	Right_dmPOS	212
Right_A11l	159	Right_A20cl	186	Right_A31	213
Right_A11m	160	Right_A20cv	187	Right_A1/2/3ulhf	214
Right_A13	161	Right_A20rv	188	Right_A1/2/3tonIa	215
Right_A12/47l	162	Right_A37mv	189	Right_A2	216
Right_A4hf	163	Right_A37lv	190	Right_A1/2/3tru	217

APPENDIX A. LABELS OF THE ATLAS

Table A.4: Region's names of the right hemisphere with their corresponding label.

Region's name	Label	Region's name	Label	Region's name	Label
Right_G	218	Right_V5/MT+	237	Right_PPtha	256
Right_vIa	219	Right_OPC	238	Right_Otha	257
Right_dIa	220	Right_iOccG	239	Right_cTtha	258
Right_vId/vIg	221	Right_msOccG	240	Right_lPFtha	259
Right_dIg	222	Right_lsOccG	241	Right_I_IV	260
Right_dId	223	Right_mAmyg	242	Right_V	261
Right_A23d	224	Right_lAmyg	243	Right_VI	262
Right_A24rv	225	Right_rHipp	244	Right_CrusI	263
Right_A32p	226	Right_cHipp	245	Right_CrusII	264
Right_A23v	227	Right_vCa	246	Right_VIIb	265
Right_A24cd	228	Right_GP	247	Right_VIIIa	266
Right_A23c	229	Right_NAC	248	Right_VIIIb	267
Right_A32sg	230	Right_vmPu	249	Right_IX	268
Right_cLinG	231	Right_dCa	250	Right_X	269
Right_rCunG	232	Right_dIPu	251	Right_Dentate	270
Right_cCunG	233	Right_mPFtha	252	Right_Interposed	271
Right_rLinG	234	Right_mPMtha	253	Right_Fastigial	272
Right_vmPOS	235	Right_Stha	254		
Right_mOccG	236	Right_rTtha	255		

Table A.5: Region's names of the vermis (in the cerebellum) with their corresponding label.

Region's name	Label	Region's name	Label	Region's name	Label
Vermis_VI	273	Vermis_VIIb	276	Vermis_IX	279
Vermis_CrusI	274	Vermis_VIIIa	277	Vermis_X	280
Vermis_CrusII	275	Vermis_VIIIb	278		

Appendix B

Visualisation of the atlas

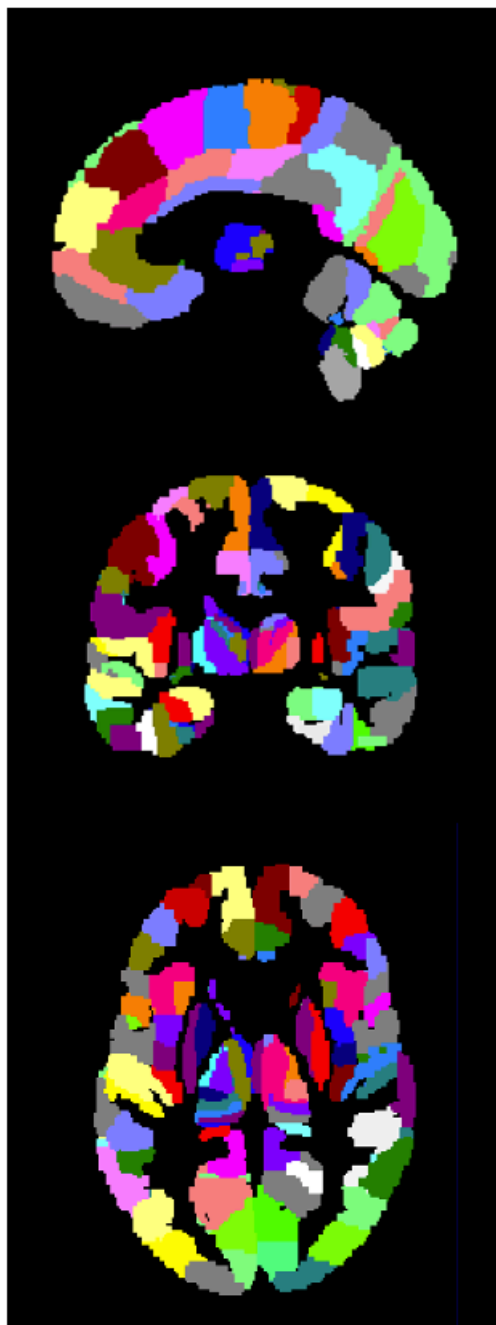


Figure B.1: Visualisation of the atlas used, each colour represents a region.

Appendix C

Non registered regions in the diffusion space

Table C.1: Region's names and labels of non registered regions after the registration from the MNI space to the native space of structural images. The C letter refers to regions belonging to the cerebellum.

Region's name	Label
Left_A20r	47
Left_VI (C)	126
Left_CrusI (C)	127
Left_CrusII (C)	128
Left_VIIb (C)	129
Left_VIIIa (C)	130
Left_VIIIb (C)	131
Left_IX (C)	132
Left_X (C)	133
Left_Dentate (C)	134
Left_Interposed (C)	135
Left_Fastigial (C)	136
Right_VI (C)	262
Right_CrusI (C)	263
Right_CrusII (C)	264
Right_VIIb (C)	265
Right_VIIIa (C)	266
Right_VIIIb (C)	267
Right_IX(C)	268
Right_X (C)	269
Right_Dentate (C)	270
Right_Interposed (C)	271
Right_Fastigial (C)	272
Vermis_VI (C)	273
Vermis_CrusI (C)	274
Vermis_CrusII (C)	275
Vermis_VIIb (C)	276
Vermis_VIIIa (C)	277
Vermis_VIIIb (C)	278
Vermis_IX (C)	279
Vermis_X (C)	280

Appendix D

Non registered regions in the functional space

Table D.1: Region's names and labels of non registered regions after the registration from the MNI space to the native space of functional images. The C letter refers to regions belonging to the cerebellum.

Region's name	Label
Left_CrusII (C)	128
Left_VIIb (C)	129
Left_VIIIa (C)	130
Left_VIIIb (C)	131
Left_IX (C)	132
Left_X (C)	133
Left_Dentate (C)	134
Left_Fastigial (C)	136
Right_CrusI (C)	263
Right_CrusII (C)	264
Right_VIIb (C)	265
Right_VIIIa (C)	266
Right_VIIIb (C)	267
Right_IX(C)	268
Right_X (C)	269
Right_Dentate (C)	270
Right_Fastigial (C)	272
Vermis_CrusI (C)	274
Vermis_CrusII (C)	275
Vermis_VIIb (C)	276
Vermis_VIIIa (C)	277
Vermis_VIIIb (C)	278
Vermis_IX (C)	279
Vermis_X (C)	280

Appendix E

Example of Volbrain report

CERES Volumetry Report. version 1.0 release 11-11-2021

Patient ID	Sex	Age	Report Date	
job379861	Male	73	14-Mar-2022	
Image Information				
Orientation		neurological		
Scale factor		0.79		
SNR		30.40		
Total intracranial volume (cm ³)		1417.68		
Volumes	Total (cm ³ /%)	Right (cm ³ /%)	Left (cm ³ /%)	Asym.(%)
Cerebellum	122.74 (8.6575) [7.2275, 9.7910]	61.05 (4.3061) [3.6370, 4.9291]	61.69 (4.3514) [3.5798, 4.8726]	-1.0471 [-2.3492, 5.0177]
Lobule I-II	0.09 (0.0062) [0.0000, 0.0229]	0.05 (0.0038) [0.0006, 0.0107]	0.03 (0.0024) [0.0000, 0.0126]	45.0450 [-43.7789, 43.3922]
Lobule III	1.06 (0.0744) [0.0675, 0.1508]	0.58 (0.0412) [0.0320, 0.0757]	0.47 (0.0333) [0.0332, 0.0773]	21.3213 [-25.6851, 21.4993]
Lobule IV	5.17 (0.3644) [0.2005, 0.3708]	2.70 (0.1903) [0.0873, 0.1848]	2.47 (0.1741) [0.1042, 0.1950]	8.8957 [-35.1154, 16.5216]
Lobule V	7.64 (0.5392) [0.2005, 0.3708]	4.36 (0.3076) [0.0873, 0.1848]	3.28 (0.2316) [0.1042, 0.1950]	28.1746 [-35.1154, 16.5216]
Lobule VI	15.54 (1.0964) [0.8721, 1.4701]	7.84 (0.5533) [0.4250, 0.7467]	7.70 (0.5431) [0.4329, 0.7375]	1.8657 [-15.5658, 15.2140]
Lobule Crus I	20.88 (1.4727) [1.2801, 2.2035]	8.86 (0.6250) [0.6503, 1.1247]	12.02 (0.8477) [0.6090, 1.0996]	-30.2391 [-11.3431, 18.9711]
Lobule Crus II	16.39 (1.1559) [0.7014, 1.3716]	8.92 (0.6294) [0.3438, 0.7131]	7.46 (0.5266) [0.3384, 0.6777]	17.7933 [-16.7390, 24.0860]
Lobule VIIIB	9.14 (0.6447) [0.4341, 0.8127]	4.48 (0.3158) [0.2098, 0.4253]	4.66 (0.3289) [0.2037, 0.4081]	-4.0576 [-24.6842, 31.1924]
Lobule VIIIA	14.83 (1.0457) [0.6006, 1.0037]	7.44 (0.5248) [0.2796, 0.5133]	7.39 (0.5210) [0.2969, 0.5145]	0.7269 [-26.0363, 21.4929]
Lobule VIIIB	7.79 (0.5495) [0.3949, 0.7333]	3.65 (0.2575) [0.1857, 0.3788]	4.14 (0.2920) [0.1875, 0.3762]	-12.5305 [-27.9369, 28.1987]
Lobule IX	6.89 (0.4863) [0.3193, 0.6622]	3.47 (0.2448) [0.1580, 0.3337]	3.42 (0.2415) [0.1581, 0.3317]	1.3793 [-12.3218, 13.4280]
Lobule X	1.42 (0.1000) [0.3193, 0.6622]	0.71 (0.0499) [0.1580, 0.3337]	0.71 (0.0501) [0.1581, 0.3317]	-0.3354 [-12.3218, 13.4280]

*All the volumes are presented in absolute value (measured in cm³) and in relative value (measured in relation to the ICV).

*The Asymmetry Index is calculated as the difference between right and left volumes divided by their mean (in percent).

*Cortical thickness is given in absolute value (mm) and also normalized in relation to the cube root of the intracranial volume (adimensional).

*Result images located in the MNI space (neurological orientation).

Figure E.1: Example of a report from Volbrain.

APPENDIX E. EXAMPLE OF VOLBRAIN REPORT

Grey matter vol.	Total (cm ³ /%)	Right (cm ³ /%)	Left (cm ³ /%)	Asym.(%)
<i>Cerebellum</i>	85.74 (6.0477) [5.3643, 7.4925]	42.99 (3.0328) [2.6585, 3.7618]	42.74 (3.0149) [2.6961, 3.7404]	0.5896 [-4.5605, 3.9767]
<i>Lobule I-II</i>	0.07 (0.0048) [0.0008, 0.0111]	0.04 (0.0030) [0.0006, 0.0051]	0.03 (0.0018) [0.0000, 0.0063]	58.7002 [-71.8993, 57.4197]
<i>Lobule III</i>	0.87 (0.0614) [0.0367, 0.0939]	0.50 (0.0350) [0.0171, 0.0466]	0.37 (0.0263) [0.0180, 0.0489]	35.8616 [-45.0944, 31.6145]
<i>Lobule IV</i>	4.57 (0.3223) [0.1723, 0.3205]	2.41 (0.1698) [0.0765, 0.1582]	2.16 (0.1525) [0.0876, 0.1706]	13.5705 [-48.5485, 22.2689]
<i>Lobule V</i>	6.80 (0.4794) [0.1723, 0.3205]	3.85 (0.2719) [0.0765, 0.1582]	2.94 (0.2075) [0.0876, 0.1706]	33.8980 [-48.5485, 22.2689]
<i>Lobule VI</i>	13.62 (0.9604) [0.7472, 1.2954]	6.78 (0.4785) [0.3581, 0.6568]	6.83 (0.4819) [0.3757, 0.6520]	-0.8813 [-24.6024, 19.5537]
<i>Lobule Crus I</i>	14.42 (1.0170) [1.0493, 1.8759]	5.73 (0.4039) [0.5195, 0.9529]	8.69 (0.6131) [0.5116, 0.9411]	-51.9359 [-19.7265, 23.1505]
<i>Lobule Crus II</i>	12.33 (0.8697) [0.6151, 1.2088]	6.67 (0.4704) [0.2982, 0.6234]	5.66 (0.3992) [0.2986, 0.6036]	20.6665 [-26.0156, 31.0976]
<i>Lobule VIIIB</i>	7.60 (0.5358) [0.3905, 0.7310]	3.83 (0.2702) [0.1893, 0.3845]	3.77 (0.2656) [0.1823, 0.3653]	2.1853 [-32.7646, 42.8349]
<i>Lobule VIIIA</i>	11.89 (0.8388) [0.5360, 0.8944]	6.08 (0.4292) [0.2477, 0.4545]	5.81 (0.4096) [0.2665, 0.4616]	5.8864 [-37.2050, 26.8665]
<i>Lobule VIIIB</i>	6.52 (0.4602) [0.3425, 0.6425]	3.26 (0.2297) [0.1589, 0.3281]	3.27 (0.2305) [0.1637, 0.3342]	-0.4599 [-41.3892, 36.0838]
<i>Lobule IX</i>	5.30 (0.3741) [0.2342, 0.5261]	2.96 (0.2086) [0.1116, 0.2642]	2.35 (0.1654) [0.1191, 0.2653]	29.1519 [-23.9161, 18.4565]
<i>Lobule X</i>	1.36 (0.0960) [0.2342, 0.5261]	0.67 (0.0472) [0.1116, 0.2642]	0.69 (0.0487) [0.1191, 0.2653]	-3.9692 [-23.9161, 18.4565]

*All the volumes are presented in absolute value (measured in cm³) and in relative value (measured in relation to the ICV).

*The Asymmetry Index is calculated as the difference between right and left volumes divided by their mean (in percent).

*Cortical thickness is given in absolute value (mm) and also normalized in relation to the cube root of the intracranial volume (adimensional).

*Result images located in the MNI space (neurological orientation).

APPENDIX E. EXAMPLE OF VOLBRAIN REPORT

Cortical thickness	Mean (mm/norm.)	Right (mm/norm.)	Left (mm/norm.)	Asym.(%)
Cerebellum	4.34 (3.864) [4.076, 4.725]	4.32 (3.842) [4.022, 4.718]	4.37 (3.886) [4.111, 4.751]	-1.1570 [-0.0445, 0.1004]
Lobule I-II	2.71 (2.408) [0.303, 2.200]	2.77 (2.468) [0.307, 2.185]	2.60 (2.311) [0.276, 2.226]	6.5201 [-0.5170, 0.5478]
Lobule III	3.42 (3.046) [2.226, 3.838]	3.62 (3.221) [2.130, 3.846]	3.16 (2.813) [2.239, 3.892]	13.3686 [-0.2909, 0.3919]
Lobule IV	5.04 (4.488) [4.094, 5.091]	5.13 (4.567) [4.030, 5.089]	4.94 (4.402) [4.110, 5.130]	3.6797 [-0.1090, 0.1618]
Lobule V	5.05 (4.497) [4.094, 5.091]	5.05 (4.497) [4.030, 5.089]	5.05 (4.496) [4.110, 5.130]	0.0329 [-0.1090, 0.1618]
Lobule VI	4.87 (4.331) [4.110, 4.919]	4.79 (4.265) [4.048, 4.931]	4.94 (4.398) [4.134, 4.937]	-3.0814 [-0.0867, 0.1281]
Lobule Crus I	3.75 (3.335) [3.931, 4.773]	3.27 (2.909) [3.823, 4.782]	4.06 (3.616) [3.982, 4.820]	-21.1948 [-0.1019, 0.1923]
Lobule Crus II	3.73 (3.317) [4.165, 4.880]	3.66 (3.254) [4.127, 4.912]	3.81 (3.391) [4.135, 4.909]	-4.1240 [-0.1406, 0.1413]
Lobule VIIIB	4.53 (4.033) [4.226, 4.953]	4.63 (4.121) [4.177, 4.959]	4.43 (3.946) [4.224, 4.992]	4.3314 [-0.1041, 0.1384]
Lobule VIIIA	4.65 (4.144) [4.198, 4.906]	4.70 (4.181) [4.142, 4.891]	4.61 (4.105) [4.216, 4.960]	1.8292 [-0.0762, 0.1400]
Lobule VIIIB	4.69 (4.177) [4.212, 4.977]	4.78 (4.255) [4.104, 4.974]	4.61 (4.105) [4.257, 5.040]	3.5948 [-0.0905, 0.1880]
Lobule IX	3.66 (3.260) [2.690, 4.436]	3.98 (3.542) [2.616, 4.409]	3.26 (2.901) [2.717, 4.502]	19.6610 [-0.1615, 0.2703]
Lobule X	3.01 (2.675) [2.690, 4.436]	2.84 (2.527) [2.616, 4.409]	3.19 (2.836) [2.717, 4.502]	-11.5541 [-0.1615, 0.2703]

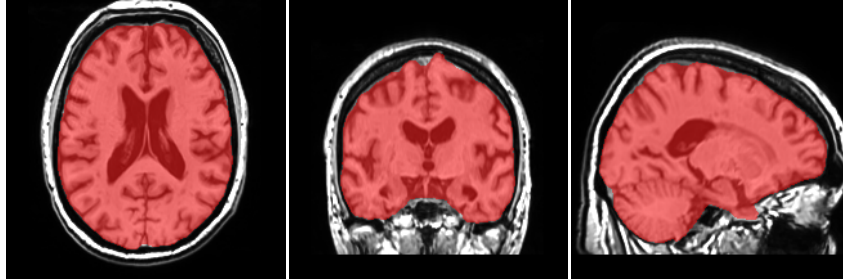
*All the volumes are presented in absolute value (measured in cm^3) and in relative value (measured in relation to the ICV).

*The Asymmetry Index is calculated as the difference between right and left volumes divided by their mean (in percent).

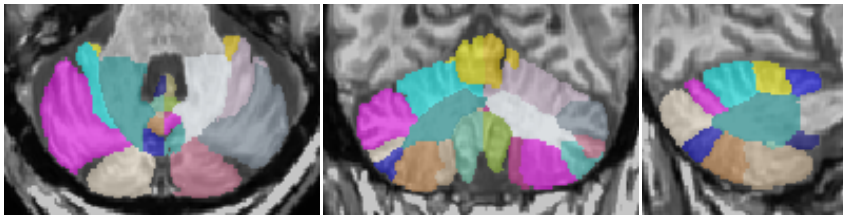
*Cortical thickness is given in absolute value (mm) and also normalized in relation to the cube root of the intracranial volume (adimensional).

*Result images located in the MNI space (neurological orientation).

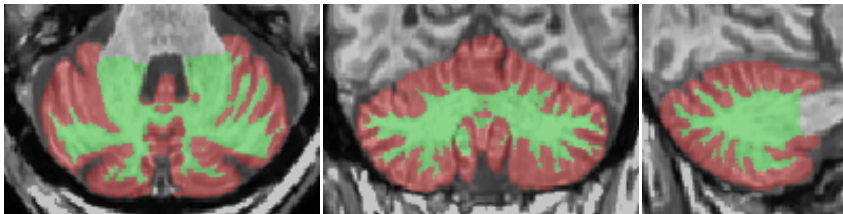
Intracranial cavity extraction



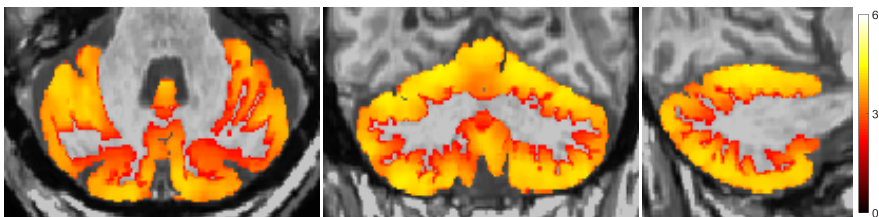
Lobules segmentation



Tissue classification



Cortical thickness



*All the volumes are presented in absolute value (measured in cm^3) and in relative value (measured in relation to the ICV).

*The Asymmetry Index is calculated as the difference between right and left volumes divided by their mean (in percent).

*Cortical thickness is given in absolute value (mm) and also normalized in relation to the cube root of the intracranial volume (adimensional).

*Result images located in the MNI space (neurological orientation).

Appendix F

Structural features selected from correlation

Table F.1: Structural features kept after the correlation pre-filtering method for intrahemispheric connections of the left hemisphere.

Left-to-left	Left-to-left
Left_A9lToLeft_A46	Left_A20rvToLeft_IPFtha
Left_A6dlToLeft_A37vl	Left_A37mvToLeft_dmPOS
Left_A6dlToLeft_dld	Left_A37mvToLeft_cCunG
Left_A9/46vToLeft_A6cvl	Left_TIToToLeft_A3l
Left_A6vlToLeft_A1/2/3ll	Left_TIToToLeft_rTtha
Left_A6vlToLeft_A37dl	Left_THToToLeft_A23v
Left_A44dToLeft_A20cl	Left_A7rToLeft_A7c
Left_A44dToLeft_A24cd	Left_A7rToLeft_A7ip
Left_A45cToLeft_A4ul	Left_A7cToLeft_A39c
Left_A44opToLeft_mPMtha	Left_A5lToLeft_A40rd
Left_A44opToLeft_cTtha	Left_A5lToLeft_A1/2/3ulhf
Left_A44vToLeft_A21c	Left_A5lToLeft_dld
Left_A14mToLeft_A23d	Left_A7ipToLeft_mPMtha
Left_A12/47lToLeft_mPMtha	Left_A40cToLeft_A23v
Left_A4hfToLeft_A37dl	Left_A7mToLeft_A24rv
Left_A4ulToLeft_I_IV	Left_A7mToLeft_lsOccG
Left_A4tToLeft_NAC	Left_A5mToLeft_A23c
Left_A4tlToLeft_A5l	Left_A31ToLeft_A24rv
Left_A38mToLeft_NAC	Left_vIaToLeft_vmPu
Left_A21rToLeft_A37vl	Left_A23dToLeft_A24rv
Left_A37dlToLeft_cpSTS	Left_A24cdToLeft_mPMtha
Left_A37dlToLeft_mOccG	Left_A24cdToLeft_IPFtha
Left_aSTSToToLeft_NAC	Left_mOccGToLeft_msOccG
Left_A20ilToLeft_G	Left_rHippToLeft_cTtha
Left_A37vlToLeft_vmPOS	Left_NACToToLeft_I_IV
Left_A20clToRight_NAC	Left_SthaToLeft_I_IV
Left_A20rvToLeft_rTtha	Left_IPFthaToLeft_I_IV

APPENDIX F. STRUCTURAL FEATURES SELECTED FROM
CORRELATION

Table F.2: Structural features kept after the correlation pre-filtering method for intrahemispheric connections of the right hemisphere.

Right-to-Right
Right_A6dlToRight_vCa
Right_A6dlToRight_vmPu
Right_A6mToRight_mAmyg
Right_A6mToRight_vmPu
Right_A9mToRight_A9/46d
Right_A9mToRight_A8vl
Right_A9mToRight_A7pc
Right_IFJToRight_vmPu
Right_A9/46vToRight_A6cvl
Right_A8vlToRight_lsOccG
Right_A44dToRight_A6cvl
Right_A45cToRight_A12/47l
Right_A44vToRight_lAmyg
Right_A44vToRight_vCa
Right_A44vToRight_NAC
Right_A11mToRight_A22r
Right_A4hfToRight_PPtha
Right_A4hfToRight_I_IV
Right_A6cdlToRight_dIg
Right_A1/2/3llToRight_- A1/2/3tru
Right_A1/2/3llToRight_A24cd
Right_A1/2/3llToRight_dIPu

Right-to-Right
Right_A38mToRight_A37vl
Right_A38mToRight_V5/MT+
Right_A38mToRight_dIPu
Right_TE1.0/TE1.2ToRight_- rHipp
Right_A22cToRight_mAmyg
Right_A38lToRight_TI
Right_A22rToRight_I_IV
Right_aSTSToRight_GP
Right_A37elvToRight_IPFtha
Right_A20clToRight_lAmyg
Right_A20cvToRight_lAmyg
Right_A35/36cToRight_TL
Right_A7rToRight_A7ip
Right_A7cToRight_A5m
Right_A5lToRight_cHipp
Right_A7pcToRight_A23c
Right_dmPOSToRight_rCunG
Right_A1/2/3ulhfToRight_Stha
Right_A1/2/3truToRight_dIPu
Right_vId/vIgToRight_Stha
Right_mPFthaToRight_Otha

APPENDIX F. STRUCTURAL FEATURES SELECTED FROM
CORRELATION

Table F.3: Structural features kept after the correlation pre-filtering method for interhemispheric connections.

Left-to-Right
Left_A38mToRight_dmPOS
Left_A37elvToRight_rCunG
Left_A20ilToRight_rCunG
Left_A37vlToRight_cLinG
Left_A20clToRight_NAC
Left_A20rvToRight_A5m
Left_A37mvToRight_cLinG
Left_A7cToRight_A5mu
Left_A23vToRight_A23d
Left_A23vToRight_cCunG
Left_A23vToRight_lAmyg
Left_vmPOSToRight_A5l
Left_V5/MT+ToRight_rCunG
Left_iOccGToRight_A5m
Left_msOccGToRight_V5/MT+
Left_NACToRight_A4ll
Left_PPthaToRight_A4ul
Left_I_IVToRight_rHipp
Left_A4ulToRight_A41/42

Appendix G

Structural features selected from mutual information

Table G.1: Structural features kept after the mutual information pre-filtering method for intrahemispheric connections of the left hemisphere.

Left-to-left	Left-to-left
Left_A41/42ToLeft_PPtha	Left_A41/42ToLeft_A20cl
Left_A9/46vToLeft_A24rv	Left_rHippToLeft_mPFtha
Left_A20rvToLeft_rCunG	Left_lAmygToLeft_Otha
Left_A7rToLeft_dIPu	Left_A45cToLeft_A24cd
Left_A1/2/3ulhfToLeft_A1/2/3tonIa	Left_A35/36rToLeft_A35/36c
Left_A4ulToLeft_A40rd	Left_aSTSToLeft_dmPOS
Left_dIPuToLeft_Stha	Left_IFJToLeft_dCa
Left_A6mToLeft_rHipp	Left_A22cToLeft_A20il
Left_A20cvToLeft_vId/vIg	Left_A22cToLeft_A31
Left_A6cdlToLeft_A1/2/3tonIa	Left_A21rToLeft_NAC
Left_A5lToLeft_A40rd	Left_A1/2/3llToLeft_A24rv
Left_A35/36rToLeft_TL	Left_A44opToLeft_A23c
Left_A6mToLeft_dIPu	Left_vId/vIgToLeft_PPtha
Left_IFSToLeft_A4ul	Left_A45cToLeft_Otha
Left_vIaToLeft_lAmyg	Left_A37dlToLeft_A39rd
Left_A22rToLeft_vmPu	Left_A9/46vToLeft_A32sg
Left_A6dlToLeft_A24cd	Left_A4hfToLeft_Otha
Left_A9/46vToLeft_A32p	Left_dmPOSToLeft_A32p
Left_A1/2/3ulhfToLeft_vmPOS	Left_A22rToLeft_vId/vIg
Left_A37dlToLeft_A7pc	Left_A37vlToLeft_dIPu
Left_A5lToLeft_lsOccG	Left_A13ToLeft_A39rv
Left_TLToLeft_TH	Left_A6cdlToLeft_vIa
Left_A10mToLeft_IPFtha	Left_A40rvToLeft_A5m
Left_A6mToLeft_Stha	Left_A45cToLeft_A4hf
Left_A6dlToLeft_dId	Left_A39cToLeft_A40rv
Left_cpSTSToLeft_dIg	Left_A20ivToLeft_vId/vIg
Left_A20ivToLeft_rCunG	Left_A46ToLeft_mPFtha
Left_A7rToLeft_I_IV	Left_A4hfToLeft_vmPu
Left_A23dToLeft_A32sg	Left_A4llToLeft_I_IV
Left_A6cvlToLeft_A1/2/3tru	Left_A20ilToLeft_IPFtha

APPENDIX G. STRUCTURAL FEATURES SELECTED FROM MUTUAL
INFORMATION

Table G.2: Structural features kept after the mutual information pre-filtering method for intrahemispheric connections of the right hemisphere.

Right-to-Right
Right_A39rvToRight_V5/MT+
Right_A7mToRight_lAmyg
Right_A9mToRight_A8vl
Right_A14mToRight_dmPOS
Right_A20rvToRight_cpSTS
Right_A8dlToRight_NAC
Right_A38mToRight_lAmyg
Right_lsOccGToRight_cTtha
Right_vId/vIgToRight_NAC
Right_A38mToRight_rHipp
Right_A10mToRight_vId/vIg
Right_A44opToRight_NAC
Right_A7ipToRight_mOccG
Right_A6vlToRight_dId
Right_A5mToRight_IPFtha
Right_dmPOSToRight_A23v
Right_A2ToRight_IPFtha
Right_A11lToRight_A38m
Right_A6cvlToRight_dId
Right_A6mToRight_Otha
Right_A22rToRight_A31

Right-to-Right
Right_A8mToRight_vmPu
Right_A6mToRight_G
Right_A32pToRight_A32sg
Right_A6dlToRight_lAmyg
Right_A5lToRight_A40rv
Right_A22cToRight_A5m
Right_A39rdToRight_cHipp
Right_A10mToRight_A38l
Right_A40rvToRight_lAmyg
Right_A1/2/3llToRight_A23d
Right_A4tlToRight_A6cvl
Right_A39rvToRight_rCunG
Right_A7ipToRight_A40rv
Right_A37elvToRight_A40rd
Right_A37dlToRight_rpSTS
Right_A37elvToRight_A7m
Right_A12/47lToRight_Otha
Right_THToRight_rHipp
Right_A20clToRight_A20rv
Right_A45cToRight_A44v
Right_A1/2/3llToRight_A4ll
Right_A6dlToRight_mPMtha

APPENDIX G. STRUCTURAL FEATURES SELECTED FROM MUTUAL
INFORMATION

Table G.3: Structural features kept after the mutual information pre-filtering method for interhemispheric connections.

Left-to-Right
Left_A5mToRight_A5m
Left_I_IVToRight_A1/2/3tru
Left_NACToRight_A4ll
Left_I_IVToRight_A6cdl
Left_A10mToRight_vCa
Left_mPFthaToRight_rTtha
Left_I_IVToRight_A4ul
Left_IFSToRight_A24rv
Left_dmPOSToRight_A5m
Left_rLinGToRight_lsOccG
Left_A4ulToRight_A8m
Left_A23dToRight_A8m
Left_A1/2/3llToRight_A4ll

Appendix H

Functional features selected from correlation

Table H.1: Functional features kept after the correlation pre-filtering method for intrahemispheric connections in the left hemisphere (left) and in the right hemisphere (right).

Left-to-left
Left_A8dlToLeft_A35/36r
Left_A46ToLeft_A7r
Left_A44dToLeft_TL
Left_A45rToLeft_cCunG
Left_A14mToLeft_OPC
Left_A14mToLeft_A4ul
Left_A14mToLeft_A37lv
Left_A14mToLeft_mOccG
Left_A14mToLeft_V5/MT+
Left_A12/47oToLeft_TI
Left_A11lToLeft_A11m
Left_A4ulToLeft_A23v
Left_A38mToLeft_A7r
Left_A22cToLeft_A24cd

Right-to-Right
Right_A46ToRight_V
Right_A14mToRight_OPC
Right_A14mToRight_V
Right_A11mToRight_A21r
Right_A32sgToRight_V
Right_A13ToRight_A1/2/3tru
Right_vId/vIgToRight_A24cd
Right_A20clToRight_A20rv

APPENDIX H. FUNCTIONAL FEATURES SELECTED FROM
CORRELATION

Table H.2: Functional features kept after the correlation pre-filtering method for interhemispheric connections.

Left-to-Right
Left_A8dlToRight_dlPu
Left_A45rToRight_A35/36c
Left_A10mToRight_dmPOS
Left_IFJToRight_dlPu
Left_A8vlToRight_Stha
Left_A45rToRight_A20r
Left_A44opToRight_A7m
Left_A4ulToRight_A14m
Left_A14mToRight_cpSTS
Left_A14mToRight_dmPOS
Left_A21cToRight_A14m
Left_A21cToRight_A32sg
Left_A11mToRight_A13
Left_A12/47lToRight_A7m
Left_A6cdlToRight_A24rv
Left_A4ulToRight_A24rv
Left_A4ulToRight_A32p
Left_A4ulToRight_A32sg
Left_A22cToRight_A6cdl
Left_A40rvToRight_A14m
Left_A37vlToRight_A8dl
Left_A20clToRight_A12/47o
Left_A20clToRight_A41/42
Left_A20rvToRight_A12/47l
Left_lAmygToRight_A8dl
Left_A32sgToRight_A22c
Left_A7rToRight_A32p
Left_A7cToRight_vIa
Left_vId/vIgToRight_A10m
Left_GToRight_A5m

Appendix I

Functional features selected from mutual information

Table I.1: Functional features kept after the mutual information pre-filtering method for intrahemispheric connections in the left hemisphere (left) and in the right hemisphere (right).

Left-to-left
Left_A22cToLeft_A24cd
Left_rCunGToLeft_dIPu
Left_rHippToLeft_I_IV
Left_A37dlToLeft_A37elv
Left_A22cToLeft_dId
Left_A39rdToLeft_A1/2/3tonIa
Left_A11lToLeft_OPC
Left_mAmygToLeft_V
Left_A37elvToLeft_A39rv
Left_A35/36rToLeft_dId
Left_A1/2/3truToLeft_cLinG
Left_rLinGToLeft_rHipp
Left_dIaToLeft_lsOccG
Left_vId/vIgToLeft_dIg

Right-to-Right
Right_A4hfToRight_lAmyg
Right_A12/47lToRight_A7m
Right_A45cToRight_A4ll
Right_cpSTSToRight_cHipp
Right_rLinGToRight_NAC
Right_A1/2/3tonIaToRight_vmPu
Right_A6dlToRight_lsOccG

APPENDIX I. FUNCTIONAL FEATURES SELECTED FROM MUTUAL
INFORMATION

Table I.2: Functional features kept after the mutual information pre-filtering method for interhemispheric connections.

Left-to-Right
Left_A38lToRight_rLinG
Left_IFSToVermis_VI
Left_TIToRight_A45r
Left_A12/47lToRight_Otha
Left_dlaToRight_vId/vIg
Left_TE1.0/TE1.2ToRight_TI
Left_A7pcToRight_A46
Left_mPMthaToRight_vmPu
Left_A7cToRight_mAmyg
Left_cHippToRight_TI
Left_A4llToRight_A1/2/3tonIa
Left_A20rvToRight_A38m
Left_THToRight_A40rv
Left_A39rvToRight_A7ip
Left_A31ToRight_A20cl
Left_IFSToRight_A6m
Left_A14mToRight_dmPOS
Left_vIaToRight_A13
Left_A20clToRight_A20rv
Left_A22cToRight_VI
Left_A7cToRight_vIa
Left_A39cToRight_rCunG
Left_rCunGToRight_A4ll
Left_A44dToRight_dCa
Left_A35/36rToRight_A38m
Left_A11mToRight_A39c
Left_A37mvToRight_A1/2/3tru
Left_A21cToRight_TI
Left_dCaToRight_rHipp
Left_A1/2/3truToRight_dId
Left_A39cToRight_mPFtha

Appendix J

Complete histogram for functional features selected with mutual information method and RF model

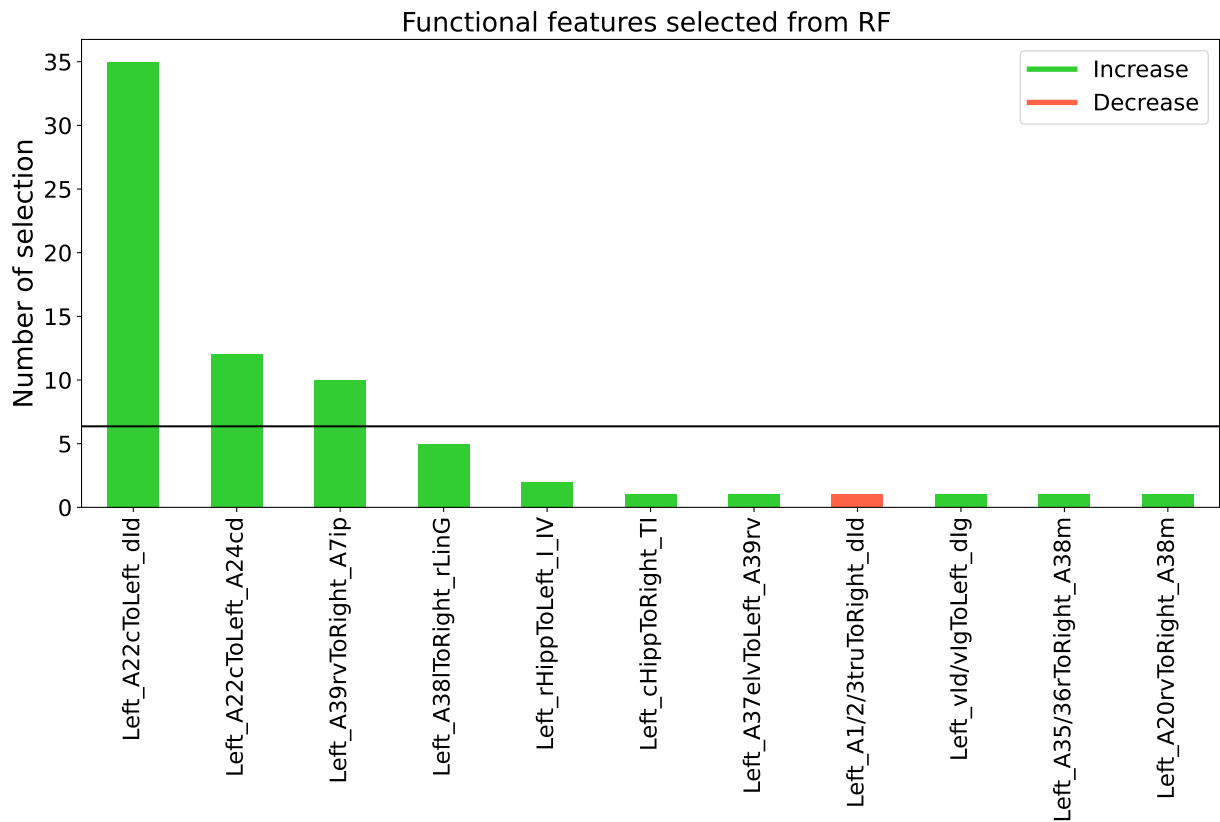


Figure J.1: Functional features selected from the pre-filtering method of mutual information and the wrapper method with the RF model.

Appendix K

Statistical tests on cerebellum

Table K.1: Statistical values for the volume asymmetry of all lobules of the cerebellum.

	p-value	Benjamini correction	Bonferroni correction
Lobule IV	0.03325	0.00833	0.004167
Lobule IX	0.07916	0.01667	
Lobule VIIIB	0.31846	0.025	
Lobule Crus I	0.37078	0.03333	
Lobule V	0.43114	0.04167	
Lobule VIIIA	0.65857	0.05	
Lobule III	0.83032	0.05833	
Lobule Crus II	0.85361	0.06667	
Lobule VIIB	0.87841	0.075	
Lobule X	0.90948	0.08333	
Lobule VI	0.92662	0.09167	
Lobule I-II	0.97091	0.1	

Table K.2: Statistical values for the grey matter asymmetry of all lobules of the cerebellum.

	p-value	Benjamini correction	Bonferroni correction
Lobule IV	0.00812	0.00833	0.004167
Lobule IX	0.02693	0.01667	
Lobule VIIIB	0.31377	0.025	
Lobule V	0.33537	0.03333	
Lobule VIIIA	0.34504	0.04167	
Lobule Crus II	0.53996	0.05	
Lobule Crus I	0.56319	0.05833	
Lobule VIIB	0.80641	0.06667	
Lobule III	0.88023	0.075	
Lobule I-II	0.95653	0.08333	
Lobule X	0.96560	0.09167	
Lobule VI	0.97945	0.1	

Table K.3: Statistical values for the grey matter volume (in cm^3) of all lobules of the cerebellum.

	p-value	Benjamini correction	Bonferroni correction
Lobule VI	0.17673	0.00833	0.004167
Lobule I-II	0.24064	0.01667	
Lobule VIIIB	0.33433	0.025	
Lobule X	0.34704	0.03333	
Lobule V	0.38735	0.04167	
Lobule VIIIA	0.417	0.05	
Lobule III	0.59864	0.05833	
Lobule VIIB	0.67136	0.06667	
Lobule IX	0.75065	0.075	
Lobule Crus II	0.88318	0.08333	
Lobule Crus I	0.91704	0.09167	
Lobule IV	0.95621	0.1	

Table K.4: Statistical values for the mean cortical thickness of all lobules of the cerebellum.

	p-value	Benjamini correction	Bonferroni correction
Lobule Crus II	0.46737	0.00833	0.004167
Lobule VI	0.48103	0.01667	
Lobule VIIIB	0.55674	0.025	
Lobule VIIIA	0.56961	0.03333	
Lobule Crus I	0.75884	0.04167	
Lobule X	0.83835	0.05	
Lobule III	0.87502	0.05833	
Lobule IX	0.88057	0.06667	
Lobule VIIB	0.89317	0.075	
Lobule IV	0.89388	0.08333	
Lobule V	0.97284	0.09167	
Lobule I-II	0.99093	0.1	

Table K.5: Statistical values for the total volume (in cm^3) of all lobules of the cerebellum.

	p-value	Benjamini correction	Bonferroni correction
Lobule VI	0.0792	0.00833	0.004167
Lobule VIIIB	0.2643	0.01667	
Lobule V	0.30412	0.025	
Lobule X	0.34593	0.03333	
Lobule VIIIA	0.36296	0.04167	
Lobule I-II	0.41612	0.05	
Lobule VIIIB	0.62865	0.05833	
Lobule III	0.66086	0.06667	
Lobule Crus II	0.83889	0.075	
Lobule IX	0.85137	0.08333	
Lobule IV	0.88939	0.09167	
Lobule Crus I	0.91343	0.1	

Appendix L

Analysis of features selected from machine learning methods

L.1 Structural features selected from correlation pre-filter method and RF model

Table L.1: Analysis of structural features selected from RF and correlation methods.

Features selected	Left A20rv To Left NAC	Left iOccG To Right A5m	Left mOccG To Left msOccG
Number of iterations	35/35	15/35	12/35
Location	- <i>A20rv</i>: rostroventral area 20 in the fusiform gyrus of the temporal lobe. - <i>NAC</i>: nucleus accumbens in the basal ganglia.	- <i>iOccG</i>: left inferior occipital gyrus. - <i>A5m</i>: medial area 5 in the precuneus (parietal lobe).	- <i>mOccG</i>: middle occipital gyrus. - <i>msOccG</i>: medial superior occipital gyrus.
Functions	- <i>A20rv</i>: implicated in the perception, language and vision [114]. - <i>NAC</i>: important role in the reward system and addiction, laughter, pleasure, fear and the placebo effect [115, 116].	- <i>iOccG</i>: implicated in vision and perception [49]. - <i>A5m</i>: precuneus implicated in the sense of self, spatial function, autobiographical memory, and navigation [117].	The occipital lobe (both right and left) mainly handles visual perception [104].
Difference between classes	Mean fibers density: 52 for C, 149 for V, 331 for VC.	Mean fibers density: 0.06 for C, 0.03 for V, 0.3 for VC.	Mean fibers density: 460 for C, 372 for V, 1093 for VC.
p-value < 0.05 for between-ness	- <i>A20rv</i>: Left hemisphere for structural and functional connectomes. - <i>NAC</i>: No.	- <i>iOccG</i>: No. - <i>A5m</i>: Right hemisphere for functional connectome.	- <i>mOccG</i>: No. - <i>msOccG</i>: No.

Features selected	Left A20rv To Left NAC	Left iOccG To Right A5m	Left mOccG To Left msOccG
p-value < 0.05 for degree	- <i>A20rv</i>: Left hemisphere for structural connectome. - <i>NAC</i>: Right hemisphere for structural connectome.	- <i>iOccG</i>: No. - <i>A5m</i>: Left hemisphere for structural connectome.	- <i>mOccG</i>: No. - <i>msOccG</i>: Left hemisphere for structural connectome.
Literature	- From <i>Indovina et al.</i> study, the <i>Left A20rv</i> region has a hubness score equal to 1.	- From <i>Indovina et al.</i> , <i>A5m</i> region belongs to the multi-modal vestibular network and has a hubness score equal to 1.	/

L.2 Functional features selected from correlation pre-filter method and RF model

Table L.2: Analysis of functional features selected with the RF model after the correlation prefiltering.

Features selected	Right A14m to Right dmPOS	Left A4ul To Left A23v
Number of iterations	35/35	25/35
Location	- <i>A14m</i>: medial area 14 in the orbital gyrus of the frontal lobe. - <i>dmPOS</i>: dorsomedial parietooccipital sulcus in the precuneus (parietal lobe).	- <i>A4ul</i>: area 4 (upper limb region) in the precentral gyrus of the frontal lobe. - <i>A23v</i>: ventral area 23 in the cingulate gyrus of the limbic lobe.
Functions	- <i>A14m</i>: the orbital gyrus is implicated in intellectual and emotional expression [103]. - <i>dmPOS</i>: the precuneus is implicated in the sense of self and agency, autobiographical memory, spatial function, and navigation [106].	- <i>A4ul</i>: the precentral gyrus contains the primary motor cortex which is responsible for controlling the voluntary movements of specific body parts [102]. - <i>A23v</i>: the cingulate gyrus coordinates sensory inputs (e.g. role in visuo-spatial orientation) [94, 95].

APPENDIX L. ANALYSIS OF FEATURES SELECTED FROM MACHINE
LEARNING METHODS

Features selected	Right A14m to Right dmPOS	Left A4ul To Left A23v
Difference between classes	Mean functional entries: 0.30 for C, 0.18 for V, 0.10 for VC.	Mean functional entries: 0.30 for C, 0.17 for V, 0.12 for VC.
p-value < 0.05 for between-ness	- <i>A14m</i> : No. - <i>dmPOS</i> : No.	- <i>A4ul</i> : Right hemisphere for the functional connectome. - <i>A23v</i> : No.
p-value < 0.05 for degree	- <i>A14m</i> : Right hemisphere for the structural connectome. - <i>dmPOS</i> : Left hemisphere for the structural connectome.	- <i>A4ul</i> : No. - <i>A23v</i> : No.
Literature	From <i>Indovina et al.</i> , region <i>dmPOS</i> belongs to the multimodal vestibular network.	From <i>Indovina et al.</i> , region <i>A23v</i> belongs to the multimodal vestibular network and has a hubness score equal to 2.

L.3 Functional features selected from correlation pre-filter method and SVM model (control VS vertigo + compensated vertigo)

Table L.3: Analysis of functional features selected with SVM and MI methods to discriminate control from vertigo participants who have compensated or not.

Features selected	Right A45c to Right A4ll	Left A4ll To Right A1/2/3tonIa	Left A7c To Right vIa
Mean weight	220 to discriminate C from VC and V. 210 to discriminate C and VC from V.	175 to discriminate C from VC and V.	150 to discriminate C from VC and V.
Location	- <i>A45c</i> : caudal area 45 in the inferior frontal gyrus. - <i>A4ll</i> : area 4 (lower limb region) in the paracentral lobule of the frontal lobe.	- <i>A4ll</i> : area 4 (lower limb region) in the paracentral lobule of the frontal lobe. - <i>A1/2/3tonIa</i> : area 1/2/3 (tongue and larynx region) in the postcentral gyrus of the parietal lobe.	- <i>A7c</i> : caudal area 7 of the superior parietal lobule. - <i>vIa</i> : ventral agranular insular in the insular gyrus of the insular lobe.

APPENDIX L. ANALYSIS OF FEATURES SELECTED FROM MACHINE
LEARNING METHODS

Features selected	Right A45c to Right A4ll	Left A4ll To Right A1/2/3tonIa	Left A7c To Right vIa
Functions	- <i>A45c</i>: responsible for speech production [96]. - <i>A4ll</i>: implicated in motor and sensory functions related to the lower limb [118].	- <i>A4ll</i>: implicated in motor and sensory functions related to the lower limb [118]. - <i>A1/2/3tonIa</i>: post-central gyrus constitutes the somatosensory cortex where the opposite half of the body's sensory inputs, such as touch, pain, temperature, etc., is registered [119].	- <i>A7c</i>: superior parietal lobule is involved in attention and visuospatial perception. [99]. - <i>vIa</i>: insular gyrus is involved in sensory preprocessing and hypergranular insula has significant connections with the auditory cortex [105].
Difference between classes	Simultaneous activity higher for vertigo participants who have compensated or not compared to control ones.	Simultaneous activity smaller for vertigo participants who have compensated or not compared to control ones.	Simultaneous activity smaller for vertigo participants who have compensated or not compared to control ones.
p-value < 0.05 for between-ness	- <i>A45c</i>: No. - <i>A4ll</i>: No.	- <i>A4ll</i>: No. - <i>A1/2/3tonIa</i>: No.	- <i>A7c</i>: No. - <i>vIa</i>: No.
p-value < 0.05 for degree	- <i>A45c</i>: Left hemisphere for the structural connectome. - <i>A4ll</i>: No.	- <i>A4ll</i>: No. - <i>A1/2/3tonIa</i>: No.	- <i>A7c</i>: No. - <i>vIa</i>: Left hemisphere for the structural connectome.
Literature	From <i>Indovina et al.</i> , region <i>A45c</i> belongs to the multimodal vestibular network and has a hubness score equal to 1.	From <i>Indovina et al.</i> , region <i>A1/2/3tonIa</i> has a hubness score equal to 2.	From <i>Indovina et al.</i> , region <i>A7c</i> belongs to the multimodal vestibular network and has a hubness score equal to 2. Region <i>vIa</i> belongs to the multimodal vestibular network.

APPENDIX L. ANALYSIS OF FEATURES SELECTED FROM MACHINE
LEARNING METHODS

Table L.4: Analysis of functional features selected with the SVM model after the mutual information method to discriminate control from vertigo participants who have compensated or not.

Features selected	Left A22c to Left A24cd	Left A11m To Right A39c	Left A1/2/3tru To Right dId
Mean weights	145 to discriminate C from VC and V. 175 to discriminate C and VC from V.	145 to discriminate C from VC and V. 200 to discriminate C and VC from V.	130 to discriminate C from VC and V. 150 to discriminate C and VC from V.
Location	- <i>A22c</i> : caudal area 22 of the superior temporal gyrus. - <i>24cd</i> : caudodorsal area 24 of the cingulate gyrus of the limbic lobe.	- <i>A11m</i> : medial area 11 in the orbital gyrus of the frontal lobe. - <i>A39c</i> : caudal area 39 in the inferior parietal lobule.	- <i>A1/2/3tru</i> : post-central gyrus of the parietal lobe. - <i>dId</i> : dorsal dysgranular insular in the insular gyrus of the insular lobe.
Functions	- <i>A22c</i> : superior temporal gyrus is the site of auditory association cortex (and a site of multisensory integration) [101]. - <i>24cd</i> : the cingulate gyrus coordinates sensory inputs (e.g. role in visuo-spatial orientation) [94, 95].	- <i>A11m</i> : orbital gyrus is implicated in intellectual and emotional expression [103]. - <i>A39c</i> : inferior parietal lobule plays a key role in attention, language and action processing [120].	- <i>A1/2/3tru</i> : post-central gyrus constitutes the somatosensory cortex where the opposite half of the body's sensory inputs, such as touch, pain, temperature, etc., is registered [119]. - <i>dId</i> : insular gyrus is involved in sensory preprocessing and hypergranular insula has significant connections with the auditory cortex [105].
Difference between classes	Simultaneous activity higher for vertigo participants who have compensated or not compared to control ones.	Simultaneous activity higher for vertigo participants who have compensated or not compared to control ones.	Simultaneous activity smaller for vertigo participants who have compensated or not compared to control ones.

Features selected	Left A22c to Left A24cd	Left A11m To Right A39c	Left A1/2/3tru To Right dId
p-value < 0.05 for between-ness	- A22c: No. - A24cd: No.	- A11m: No. - A39c: No.	- A1/2/3tru: No. - dId: No.
p-value < 0.05 for degree	- A22c: No. - A24cd: No.	- A11m: No. - A39c: No.	- A1/2/3tru: Left hemisphere for the structural connectome. - dId: No.
Literature	From <i>Indovina et al.</i> , region A24cd belongs to the multimodal vestibular network and has a hubness score equal to one.	From <i>Indovina et al.</i> , region A39c belongs to the multimodal vestibular network.	From <i>Indovina et al.</i> , region dId belongs to the multimodal vestibular network.

L.4 Functional features selected from correlation pre-filter method and SVM model (control + compensated vertigo VS vertigo)

The four features selected from the mutual information pre-filtering method and with the SVM model to discriminate vertigo participants from control participants and vertigo participants who have compensated are included in the six features selected to discriminate vertigo participants who have compensated or not from control participants. These four features are *Right_A45cToRight_A4ll*, *Left_A11mToRight_A39c*, *Left_A22cToLeft_A24cd* and *Left_A1/2/3truToRight_dId*.

The simultaneous activity for vertigo participants compared to control and compensated vertigo ones is higher with the *Right_A45cToRight_A4ll* and *Left_A11mToRight_A39c* features but smaller with the *Left_A22cToLeft_A24cd* and *Left_A1/2/3truToRight_dId* features.

L.5 Functional features selected from MI pre-filter method and RF model

Table L.5: Analysis of functional features selected from RF and MI methods.

Features selected	Left A22c to Left dId	Left A22c To Left A24cd	Left A39rv To Right A7ip
Number of iterations	35/35	12/35	10/35
Location	- <i>A22c</i>: caudal area 22 of the superior temporal gyrus. - <i>dId</i>: dorsal dysgranular insular of the insular gyrus.	- <i>A22c</i>: caudal area 22 of the superior temporal gyrus. - <i>A24cd</i>: caudodorsal area 24 in the cingulate gyrus of the limbic lobe.	- <i>A39rv</i>: rostroventral area 39 of the inferior parietal lobule. - <i>A7ip</i>: intraparietal area 7 of the superior parietal lobule.
Functions	- <i>A22c</i>: superior temporal gyrus is the site of auditory association cortex (and a site of multisensory integration). - <i>dId</i>: insular gyrus is involved in sensory preprocessing and hypergranular insula has significant connections with the auditory cortex [105].	- <i>A22c</i>: superior temporal gyrus is the site of auditory association cortex (and a site of multisensory integration) [101]. - <i>A24cd</i>: the cingulate gyrus coordinates sensory inputs (e.g. role in visuo-spatial orientation) [94, 95].	- <i>A39rv</i>: inferior parietal lobule plays a key role in attention, language and action processing [120]. - <i>A7ip</i>: superior parietal lobule is closely linked with the occipital lobe and is involved in aspects of attention and visuospatial perception [99].
Difference between classes	Mean functional entries: 0.13 for C, 0.23 for V, 0.25 for VC.	Mean functional entries: 0.11 for C, 0.21 for V, 0.27 for VC.	Mean functional entries: 0.18 for C, 0.19 for V, 0.17 for VC.
p-value < 0.05 for between-ness	- <i>A22c</i>: No. - <i>dId</i>: No.	- <i>A22c</i>: No. - <i>A24cd</i>: No.	- <i>A39rv</i>: left hemisphere for structural connectome. - <i>A7ip</i>: left hemisphere for functional connectome.

APPENDIX L. ANALYSIS OF FEATURES SELECTED FROM MACHINE
LEARNING METHODS

Features selected	Left A22c to Left dId	Left A22c To Left A24cd	Left A39rv To Right A7ip
p-value < 0.05 for degree	- <i>A22c</i> : No. - <i>dId</i> : No.	- <i>A22c</i> : No. - <i>A24cd</i> : No.	- <i>A7ip</i> : Right hemisphere for structural connectome. - <i>A39rv</i> : No.
Literature	From <i>Indovina et al.</i> , region <i>dId</i> belongs to the multimodal vestibular network.	From <i>Indovina et al.</i> , region <i>24cd</i> belongs to the multimodal vestibular network and has a hubness score of 1.	From <i>Indovina et al.</i> , region <i>A39rv</i> belongs to the multimodal vestibular network. Region <i>A7ip</i> belongs to the multimodal vestibular network and has a hubness score of 3 for the left hemisphere and 2 for the right hemisphere.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
École polytechnique de Louvain

Rue Archimède, 1 bte L6.11.01, 1348 Louvain-la-Neuve, Belgique | www.uclouvain.be/epl