



Universidad de Valladolid

Escuela Técnica Superior de Ingenieros de Telecomunicación

Departamento de Teoría de la Señal y Comunicaciones e Ingeniería
Telemática

TESIS DOCTORAL:

**Applications of Blind Source Separation to the
Magnetoencephalogram Background Activity in
Alzheimer's Disease**

Presentada por Javier Escudero Rodríguez para optar al grado de
doctor por la Universidad de Valladolid con Mención Europea

Dirigida por:
Roberto Hornero Sánchez
Daniel E. Abásolo Baz

Director durante la estancia europea en *Cardiff University*:
Saeid Sanei

Año académico 2009/10



Universidad de Valladolid

Escuela Técnica Superior de Ingenieros de Telecomunicación

Departamento de Teoría de la Señal y Comunicaciones e Ingeniería
Telemática

TESIS DOCTORAL:
**Applications of Blind Source Separation to the
Magnetoencephalogram Background Activity in
Alzheimer's Disease**

DOCTORANDO: JAVIER ESCUDERO RODRÍGUEZ

DIRECTORES: ROBERTO HORNERO SÁNCHEZ Y DANIEL E. ABÁSULO BAZ

DIRECTOR DE LA ESTANCIA EN CARDIFF UNIVERSITY: SAEID SANEI

TRIBUNAL

PRESIDENTE:

VOCALES:

SECRETARIO:

SUPLENTE:

FECHA DE LECTURA:

CALIFICACIÓN:



Universidad de Valladolid

Higher Technical School of Telecommunications Engineering

Department of Signal Theory and Communications and Telematics
Engineering

DOCTORAL THESIS:

**Applications of Blind Source Separation to the
Magnetoencephalogram Background Activity in
Alzheimer's Disease**

Submitted by Javier Escudero Rodríguez for the PhD degree with
the “Doctor Europeus” award at the University of Valladolid

Supervisors:

Roberto Hornero Sánchez

Daniel E. Abásolo Baz

Supervisor during the research stay at Cardiff University:

Saeid Sanei

Academic year 2009/10

A Iria y mis padres.

*“Las ideas no duran mucho.
Hay que hacer algo con ellas.”*
Santiago Ramón y Cajal.

Acknowledgements

First of all, I would like to express my gratitude to my supervisors, Dr. Hornero and Dr. Abásolo, for their helpful suggestions, advice and encouragement during the course of this work.

I also wish to state my appreciation to the members of the *Grupo de Ingeniería Biomédica* for their valuable support and discussions.

Sincere thanks are extended to the *Asociación de Familiares de Enfermos de Alzheimer* (Madrid). Without their collaboration, this Doctoral Thesis could not have been carried out. Additionally, I am grateful to *Centro de Magnetoencefalografía Dr. Pérez-Modrego* at Computense University of Madrid and, specially, to Dr. Fernández, for providing us with the magnetoencephalographic recordings analysed in this Doctoral Thesis.

Keen thanks are also due to Dr. Sanei and the members of the Centre of Digital Signal Processing for their warm welcome and help during my research stay at Cardiff University.

I am grateful to all my friends and my brother, Rubén, for their continued moral support and the good moments we have spent together.

I am particularly indebted to my parents, María del Pilar and Bernardo, for their invaluable support and devotion.

Finally, I would like to express very special thanks to Iria for her wholehearted help, encouragement, understanding and love.

This research project was partially supported by the Ministry of Education (Spain) within the grant program *Formación de Profesorado Universitario* (FPU). The author acknowledges the stipend from the FPU grant no. AP2007-03303.

Abstract

In this Doctoral Thesis, Magnetoencephalogram (MEG) background activity from 36 patients with a diagnosis of probable Alzheimer's Disease (AD) and 26 healthy elderly control subjects has been analysed with Blind Source Separation (BSS) methods. Our aim was to apply BSS techniques to help in the analysis and interpretation of this kind of brain activity, paying special attention to AD.

The MEG is the non-invasive recording of the tiny magnetic fields generated by the neurons. This neurophysiological technique measures the brain cortex activity directly, without interpreting the information on the basis of vascular or metabolic changes. Its temporal resolution is high and the magnetic recordings neither depend on any reference point nor are affected by the resistive properties of extra-cerebral tissues. On the other hand, the MEG apparatus needs superconductive materials and magnetically shielded rooms to properly acquire the brain signals. This has prevented any widespread use of this technique to record the brain activity. Despite the fact that difficulties are faced in the analysis of these signals, the MEG could provide relevant information about diverse brain states and diseases, such as AD.

AD is a progressive neurodegenerative disorder. It causes memory loss and other cognitive and behavioural symptoms that impair the activities of daily living. AD is the most common dementia in the Western World as it accounts for 50% to 60% of all cases. It shows an almost exponential increase with age. As a result, its prevalence in people over 85 years is between 24% and 33%. In clinical practice, AD must be differentiated from other dementias, though a definite diagnosis can only be made by necropsy. The criteria for AD diagnosis largely depend on the exclusion of other disorders. Of note

is that the accuracy in the clinical diagnosis is limited, with sensitivity of around 80% and specificity of 70%. Thus, it is important to develop new approaches that might help in AD detection.

The term BSS denotes a set of techniques useful to decompose multi-channel recordings into their constituent underlying components. The BSS defines a generative model for the measurements and it tries to estimate the inner components (or sources) by making a few general assumptions about the data. The most important hypothesis is that the BSS components are mutually independent or, alternatively, that they are mutually decorrelated over time. BSS extracts the sources by exploiting this assumption. Thanks to this ability, the application of BSS techniques to MEGs can help to inspect and analyse these biomedical signals from new perspectives. Thus, this could provide us with both novel methodologies to deal with problems encountered in the processing of these recordings and relevant information about brain activity.

For these reasons, in this Doctoral Thesis, MEG data were processed with BSS techniques in the context of four different applications:

- The decompositions of real MEGs computed with five common BSS algorithms were compared to assess their degree of similarity. The results showed that the most consistent (i.e., similar) pair of algorithms was AMUSE–SOBI, followed by JADE–FastICA. Additionally, the overall level of similarity increased as longer signals were decomposed.
- The ability of several combinations of BSS algorithm, epoch length and artefact detection metric to automatically reduce the cardiac, ocular and power line artefacts in the MEGs was assessed. The results indicated that a Constrained Blind Source Separation (cBSS) approach was suitable to remove the cardiac activity. Additionally, a combination of artefact detection metrics based on entropy or power criteria with AMUSE or SOBI could help to reduce the ocular contamination. Finally, the electrical noise could be reduced by means of a spectral metric and AMUSE.

- The ability of a BSS preprocessing to improve the separation between AD patients and controls' spectral and non-linear features from MEGs was measured. Ordering criteria were defined to straightforwardly compare the BSS components of different subjects. The comparisons between the classifications derived from the unprocessed and the BSS preprocessed MEG signals revealed rises in the areas under the ROC curves associated with the features between 0.023 and 0.227 and accuracy increases of up to 22.6% in the best cases. These corresponded to preprocessings developed using AMUSE and SOBI with a spectral ordering of the components.
- An adaptive framework to extract rhythmical brain activity from diverse scalp regions was introduced. This employed the local adaptiveness of an Empirical Mode Decomposition to estimate oscillating brain activity in several bands and a cBSS to extract the actual brain rhythms from each region. The statistical analysis suggested that AD might affect the Spectral Coherence ($\text{Coh}(f)$) between regions, although the results were not significant. However, a leave-one-out classification analysis based on the $\text{Coh}(f)$ computed from spectral bands classified the AD patients versus the elderly control subjects with an accuracy equal to 96.8%.

In summary, the findings of this Doctoral Thesis suggest utility of BSS to help in the processing of MEG background activity and in the identification and characterisation of AD. Therefore, BSS may be an important tool to analyse this kind of biomedical recordings. Nevertheless, further investigations are needed to confirm our results.

Índice de contenidos en Español

1	Introducción	1
1.1	Contexto: Procesado de señales biomédicas	2
1.2	Magnetoencefalografía	6
1.3	Enfermedad de Alzheimer	8
1.4	Separación ciega de fuentes	11
1.5	Hipótesis	13
1.6	Objetivos de la Tesis Doctoral	16
1.7	Plan de trabajo	19
1.8	Estructura de la Tesis Doctoral	21
2	Magnetoencefalografía	25
2.1	Introducción	26
2.2	Generación de los campos magnéticos cerebrales	27
2.2.1	Sistema nervioso	27
2.2.2	Cerebro	28
2.2.3	Neuronas y actividad neuronal	30
2.2.4	Ritmos y señales electromagnéticas cerebrales	33
2.2.5	Biomagnetismo	35
2.3	Sistemas de registro MEG	36
2.3.1	SQUIDS y transformadores de flujo	38
2.3.2	Contenedor criogénico	40
2.3.3	Habitación magnéticamente aislada	41
2.3.4	Eliminación del ruido magnético	42
2.4	Artefactos	44
2.4.1	Cardíaco	45
2.4.2	Ocular	46
2.4.3	Red eléctrica	46
2.4.4	Otros artefactos	47
2.5	Comparación con otros métodos de neuroimagen	47
2.5.1	SPECT y PET	48
2.5.2	fMRI y MRS	50

2.5.3 EEG	51
3 Enfermedad de Alzheimer	55
3.1 Introducción	56
3.2 Importancia social de la enfermedad	57
3.3 Principales características	58
3.3.1 Evolución de la enfermedad	58
3.3.2 Fisiología patológica	61
3.3.3 Factores de riesgo	63
3.3.4 Diagnóstico clínico	66
3.3.5 Posibles biomarcadores	68
3.4 Tratamiento	70
3.4.1 Tratamientos sintomáticos	71
3.4.2 Candidatos a modificar el curso de la enfermedad	73
3.4.3 Estudios epidemiológicos	74
3.4.4 Orientación a los cuidadores	76
4 Separación ciega de fuentes	77
4.1 Introducción	79
4.2 Técnicas clásicas de proyección	81
4.2.1 Análisis de componentes principales	81
4.2.2 Análisis factorial	84
4.2.3 Diferencias entre PCA, FA y BSS	85
4.3 Fundamentos	86
4.3.1 Modelo y validez de las hipótesis	87
4.3.2 Limitaciones	91
4.4 Métodos basados en estadísticas de segundo orden	95
4.4.1 Fundamentos	96
4.4.2 AMUSE	97
4.4.3 SOBI	98
4.5 Métodos basados en estadísticas de orden superior	99
4.5.1 Fundamentos	100
4.5.2 FastICA	101
4.5.3 InfoMax	104
4.5.4 JADE	106
4.6 Separación ciega de fuentes restringida	107
4.7 Relaciones entre técnicas de BSS	109
4.7.1 Consistencia de los métodos para señales reales	112
4.8 Aplicaciones al análisis del EEG y el MEG	115
4.8.1 Eliminación de artefactos	115
4.8.2 Realce de características	121

4.8.3 Extracción de ritmos cerebrales	126
5 Registro de MEG	129
5.1 Datos sociodemográficos de los sujetos	130
5.2 Registro y preprocesado de las señales MEG	131
5.3 Selección de épocas	135
6 Métodos	137
6.1 Introducción	139
6.2 Preprocesado	140
6.2.1 Selección del número de componentes	141
6.2.2 Datos sintéticos	142
6.3 Métodos de separación ciega de fuentes	145
6.3.1 SOS-BSS	146
6.3.2 HOS-BSS	146
6.3.3 cBSS	148
6.4 Evaluación de la consistencia de diversos métodos de BSS	148
6.4.1 Métrica	149
6.5 Eliminación de artefactos con BSS	150
6.5.1 Métricas de detección de artefactos	151
6.5.2 Evaluación de la eliminación de artefactos	160
6.6 Realce de características con BSS	163
6.6.1 Descripción	163
6.6.2 Métodos de caracterización de señales	165
6.7 Extracción adaptativa de ritmos cerebrales con BSS	171
6.7.1 Preprocesado	171
6.7.2 Descomposición empírica en modos	172
6.7.3 Caracterización espectral de las IMFs	173
6.7.4 Extracción regional de ritmos cerebrales	174
6.7.5 Coherencia	176
6.8 Herramientas de análisis estadístico	177
6.8.1 Herramientas de visualización	177
6.8.2 Pruebas de significación estadística	177
6.8.3 Clasificación	178
7 Resultados	181
7.1 Preprocesado	182
7.1.1 Datos sintéticos	182
7.1.2 Señales reales de MEG	185
7.2 Evaluación de la consistencia	187
7.3 Eliminación de artefactos	193

7.3.1 Artefactos en las señales originales	193
7.3.2 Separación ciega de fuentes	198
7.3.3 Detección de artefactos	198
7.3.4 Evaluación de la eliminación de artefactos	202
7.4 Realce de características	221
7.4.1 Descripción preliminar del análisis	221
7.4.2 Descomposición y ordenación de componentes	222
7.4.3 Selección de las componentes más significativas	233
7.4.4 Evaluación de la mejora en la clasificación	235
7.5 Extracción adaptativa de ritmos cerebrales	241
7.5.1 Estimación de las referencias	241
7.5.2 Análisis estadístico descriptivo	244
7.5.3 Clasificación	247
8 Discusión	249
8.1 Introducción	250
8.2 Preprocesado	252
8.3 Evaluación de la consistencia	254
8.4 Eliminación de artefactos	256
8.5 Realce de características	263
8.6 Extracción adaptativa de ritmos cerebrales	272
8.7 Limitaciones del estudio	277
9 Conclusiones y futuras líneas de investigación	281
9.1 Introducción	282
9.2 Conclusiones de la Tesis Doctoral	284
9.3 Contribuciones originales al estado del arte	287
9.4 Futuras líneas de investigación	289
A Publicaciones	293
A.1 Artículos con revisión por pares en revistas del SCI-JCR	294
A.2 Capítulos de libro internacionales	297
A.3 Contribuciones a congresos internacionales	297
A.4 Contribuciones a congresos nacionales	301
B Resúmenes en español	305
B.1 Capítulo 1: Introducción	306
B.2 Capítulo 2: Magnetoencefalografía	307
B.3 Capítulo 3: Enfermedad de Alzheimer	309
B.4 Capítulo 4: Separación ciega de fuentes	310
B.5 Capítulo 5: Registro de MEG	314

Índice de contenidos en Español	XV
B.6 Capítulo 6: Métodos	316
B.7 Capítulo 7: Resultados	318
B.8 Capítulo 8: Discusión	323
C Conclusiones en español	329
C.1 Introducción	330
C.2 Conclusiones de la Tesis Doctoral	332
C.3 Contribuciones originales al estado del arte	336
C.4 Futuras líneas de investigación	338
Acrónimos	341
Bibliografía	345

Contents

1	Introduction	1
1.1	Context: Biomedical Signal Processing	2
1.2	Magnetoencephalography	6
1.3	Alzheimer’s Disease	8
1.4	Blind Source Separation	11
1.5	Hypothesis	13
1.6	Aims of the Doctoral Thesis	16
1.7	Schedule of Activities	19
1.8	Structure of this Dissertation	21
2	Magnetoencephalography	25
2.1	Introduction	26
2.2	Generation of the Brain Magnetic Fields	27
2.2.1	Nervous System	27
2.2.2	Brain	28
2.2.3	Neurons and Neural Activity	30
2.2.4	Electromagnetic Brain Signals and Rhythms	33
2.2.5	Biomagnetism	35
2.3	MEG Recording Systems	36
2.3.1	SQUIDS and Flux Transformers	38
2.3.2	Cryogenic Container	40
2.3.3	Magnetically Shielded Room	41
2.3.4	Removal of the Magnetic Noise	42
2.4	Artefacts	44
2.4.1	Cardiac	45
2.4.2	Ocular	46
2.4.3	Power Line Noise	46
2.4.4	Other Artefacts	47
2.5	Comparison With Other Neuroimaging Methods	47
2.5.1	SPECT and PET	48
2.5.2	fMRI and MRS	50

2.5.3	EEG	51
3	Alzheimer's Disease	55
3.1	Introduction	56
3.2	Social Importance of AD	57
3.3	Main Characteristics of AD	58
3.3.1	Evolution of the Disease	58
3.3.2	Pathophysiology	61
3.3.3	Risk Factors	63
3.3.4	Clinical Diagnosis	66
3.3.5	Possible Biomarkers	68
3.4	Current Treatment	70
3.4.1	Symptomatic Treatments	71
3.4.2	Modifying-Disease Drug Candidates	73
3.4.3	Possible Drugs Based on Epidemiological Studies	74
3.4.4	Orientation to caregivers	76
4	Blind Source Separation	77
4.1	Introduction	79
4.2	Classical Projection Techniques	81
4.2.1	Principal Component Analysis	81
4.2.2	Factor Analysis	84
4.2.3	Differences Among PCA, FA and BSS	85
4.3	BSS Fundamentals	86
4.3.1	Model and Validity of the Assumptions	87
4.3.2	Limitations of BSS	91
4.4	Second Order Statistics Methods for BSS	95
4.4.1	Fundamentals	96
4.4.2	AMUSE	97
4.4.3	SOBI	98
4.5	Higher Order Statistics Methods for BSS	99
4.5.1	Fundamentals	100
4.5.2	FastICA	101
4.5.3	InfoMax	104
4.5.4	JADE	106
4.6	Constrained BSS Algorithms	107
4.7	Relationships Among BSS Algorithms	109
4.7.1	Consistency of BSS Algorithms for Real Signals	112
4.8	BSS Applications to EEG and MEG Activity	114
4.8.1	Artefact Removal	115
4.8.2	Feature Enhancement With BSS	121

4.8.3	Extraction of Brain Rhythms With BSS	124
5	MEG Recordings	129
5.1	Subjects' Sociodemographic Data	130
5.2	MEG Recording and Preprocessing	131
5.3	Epoch Selection	135
6	Methods	137
6.1	Introduction	139
6.2	Preprocessing	140
6.2.1	Selection of the Number of Components	141
6.2.2	Synthetic Data	142
6.3	Blind Source Separation Methods	145
6.3.1	SOS-BSS	146
6.3.2	HOS-BSS	146
6.3.3	cBSS	148
6.4	Consistency Evaluation of BSS Methods	148
6.4.1	Metric	149
6.5	Artefact Removal With BSS	150
6.5.1	Artefact Detection Metrics	151
6.5.2	Evaluation of the Artefact Rejection	160
6.6	Feature Enhancement With BSS	163
6.6.1	Description	163
6.6.2	Signal Features	165
6.7	Extraction of Brain Rhythms With cBSS	171
6.7.1	Preprocessing	171
6.7.2	Empirical Mode Decomposition	172
6.7.3	Frequency Characterisation of the IMFs	173
6.7.4	Regional Extraction of Brain Rhythms	174
6.7.5	Coherence	176
6.8	Statistical Analysis Tools	177
6.8.1	Visual Information Tools	177
6.8.2	Statistical Significance Tests	177
6.8.3	Classification Analysis	178
7	Results	181
7.1	Preprocessing	182
7.1.1	Synthetic Data	182
7.1.2	Real MEG Recordings	185
7.2	Evaluation of the Consistency	187
7.3	Artefact Removal	193

7.3.1	Artefacts in the Original Recordings	193
7.3.2	Blind Source Separation	198
7.3.3	Artefact Detection	198
7.3.4	Evaluation of the Artefact Removal	202
7.4	Feature Enhancement	221
7.4.1	Preliminary Description of the Analysis	221
7.4.2	Decompositions and Ordering Criteria	222
7.4.3	Selection of the Most Significant Components	233
7.4.4	Evaluation of the Classification Improvement	235
7.5	Adaptive Extraction of Brain Rhythms	241
7.5.1	Estimation of References	241
7.5.2	Descriptive Statistical Analysis	244
7.5.3	Classification Analysis	247
8	Discussion	249
8.1	Introduction	250
8.2	Preprocessing	252
8.3	Consistency Evaluation	254
8.4	Artefact Removal	256
8.5	Feature Enhancement	263
8.6	Adaptive Extraction of Brain Rhythms	272
8.7	Limitations of the Study	277
9	Conclusions and Future Research Lines	281
9.1	Introduction	282
9.2	Conclusions of the Doctoral Thesis	284
9.3	Original Contributions to the State of the Art	287
9.4	Future Research Lines	289
A	Publications	293
A.1	Peer-Reviewed International Articles in Journals Indexed in the SCI–JCR	294
A.2	International Book Chapters	297
A.3	Contributions to International Conferences	297
A.4	Contributions to National Conferences	301
B	Resúmenes en Español	305
B.1	Capítulo 1: Introducción	306
B.2	Capítulo 2: Magnetoencefalografía	307
B.3	Capítulo 3: Enfermedad de Alzheimer	309
B.4	Capítulo 4: Separación ciega de fuentes	310

Contents

XXI

B.5	Capítulo 5: Registro de MEG	314
B.6	Capítulo 6: Métodos	316
B.7	Capítulo 7: Resultados	318
B.8	Capítulo 8: Discusión	323
C	Conclusiones en Español	329
C.1	Introducción	330
C.2	Conclusiones de la Tesis Doctoral	332
C.3	Contribuciones originales al estado del arte	336
C.4	Futuras líneas de investigación	338
	Acronyms	341
	Bibliography	345

List of Figures

2.1	View of the brain showing some of its main parts. Adapted from (National Institute on Aging, 2008).	28
2.2	Example of action potential in a neuron.	32
2.3	Orthogonality between EEG and MEG signals as illustrated by Vrba and Robinson (2001) .	36
2.4	Origin of the MEG. (a) Section of the human brain. (b) The sulci and gyri of the cortex produce either tangential or radial currents. (c) Tangential currents produce magnetic fields observable outside the head. (d) Radial currents do not produce magnetic fields outside the head. (e) Magnetic fields due to cortical sources will exit and re-enter the scalp. Taken from (Vrba and Robinson, 2001).	36
2.5	Block diagram of a MEG system. Adapted from (Sternickel and Braginski, 2006).	38
2.6	Power of undesired components (external noise and artefacts) and brain signals as a function of frequency. Adapted from (Hämäläinen and Hari, 2004).	39
2.7	Cyogenic container as part of a MEG system.	41
2.8	Example of magnetically shielded room.	42
2.9	Diagram comparing diverse techniques for functional brain analysis in terms of their temporal and spatial resolution and their degree of invasiveness.	49
3.1	View on the evolution from healthy ageing to AD by National Institute on Aging (2008) .	60
3.2	Plaques and tangles in the cortex as shown by Blennow et al. (2006) .	61
3.3	Mechanism of neural death in AD.	64

3.4	FDG-PET scans in an elderly woman at baseline (1989) and over 9 years, while the patient declined to MCI and AD. Progressive reductions in glucose metabolism are illustrated by Blennow et al. (2006)	69
3.5	MRI scans in an elderly man at baseline (1993) and over 10 years, while he declined to MCI and AD. The reduction in size of the hippocampus (red) and entorhinal cortex (yellow) and the increase of the ventricle (green) with disease progression are illustrated by Blennow et al. (2006)	70
4.1	Example of the differences between PCA and ICA. The scale of the plots is arbitrary. (a) Mixed random variables with uniform distribution. (b) PCA. (c) ICA.	86
4.2	The functions G_1 and G_2 compared with the fourth power, as used in kurtosis.	104
6.1	One example of synthetic sources. S1–S4: simulated undesired signals, S5–S11: simulated ‘useful’ signals.	144
6.2	Average autocorrelation function of the controls’ MEGs for epoch lengths of 10 s, 20 s, 40 s, 60 s and 90 s.	147
6.3	Example of the cardiac activity that appears over the scalp and the reference extracted from it.	157
6.4	Location of the 13 MEG channels considered in the P_{EYES} metric.	159
6.5	Subsets of channels used in the spectral evaluation of the OA removal.	162
6.6	Distributions of the MEG sensors into central (yellow), anterior (red), lateral (blue) and posterior (green) regions for left (white text) and right (black text) hemispheres. The midline sensors are marked in white.	175
7.1	Estimated n values for each of the 1000 synthetic mixtures according to the tested criteria.	184
7.2	Average E_{norm} for diverse average SNR values in the synthetic mixtures.	185
7.3	Boxplots showing the number of components (n) estimated for the epoch lengths considered in this study.	186
7.4	Normalised external noise power in real MEG signals as modelled by the preprocessing.	186
7.5	Inner power retained for BSS by selecting the optimum number of components (n) estimated by the MDL.	187

7.6	Average F values for each pair of BSS algorithms (A: AMUSE, S: SOBI, J: JADE, eI: extended Infomax, F: FastICA) and epoch length.	189
7.7	ANOVA with n as a covariate for the F metric applied to epochs of 10 s (A: AMUSE, S: SOBI, J: JADE, e: extended Infomax, F: FastICA).	190
7.8	ANOVA with n as a covariate for the F metric applied to epochs of 20 s (A: AMUSE, S: SOBI, J: JADE, e: extended Infomax, F: FastICA).	190
7.9	ANOVA with n as a covariate for the F metric applied to epochs of 40 s (A: AMUSE, S: SOBI, J: JADE, e: extended Infomax, F: FastICA).	191
7.10	ANOVA with n as a covariate for the F metric applied to epochs of 60 s (A: AMUSE, S: SOBI, J: JADE, e: extended Infomax, F: FastICA).	192
7.11	ANOVA with n as a covariate for the F metric applied to epochs of 90 s (A: AMUSE, S: SOBI, J: JADE, e: extended Infomax, F: FastICA).	193
7.12	Average QRS complex that appears in the original recordings for epoch length equal to 10 s.	195
7.13	Median of the difference between the PSDs for the two subsets of MEG channels used to compare the impact of OAs in the recordings.	196
7.14	Average spectra centred at the PLN frequency in the original signals computed from epochs of 60 s.	197
7.15	Average QRS complex in the MEG recordings after diverse BSS-based CA removals.	209
7.16	Average QRS complex in the MEG recordings after the cBSS-based CA removal for epochs of 10 s.	210
7.17	Median of the difference between the PSDs for the two subsets of MEG channels used to compare the impact of OAs in the recordings after diverse BSS-based OA removals.	216
7.18	Average spectra centred at the PLN frequency after the artefact removal with three BSS procedures.	220
7.19	Average values of the signal features for every AMUSE component in AD patients and control subjects of the training set.	224
7.20	p -values of a one-way ANOVA with age as a covariate computed for each AMUSE component in the training set for (a) MF and SpecEn and (b) LZC and SampEn.	224

7.21	Average values of the signal features for the SOBI components orderer by their MF in AD patients and control subjects of the training set.	226
7.22	p -values of a one-way ANOVA with age as a covariate computed for SOBI components ordered by their MF in the training set for (a) MF and SpecEn and (b) LZC and SampEn. . .	226
7.23	Average values of the signal features for the SOBI components orderer by their Kurtosis in AD patients and control subjects of the training set.	228
7.24	p -values of a one-way ANOVA with age as a covariate computed for SOBI components ordered by their Kurtosis in the training set for (a) MF and SpecEn and (b) LZC and SampEn.	228
7.25	Average values of the signal features for the Extended InfoMax components orderer by their MF in AD patients and control subjects of the training set.	229
7.26	p -values of a one-way ANOVA with age as a covariate computed for the Extended InfoMax components ordered by their MF in the training set for (a) MF and SpecEn and (b) LZC and SampEn.	230
7.27	Average values of the signal features for the Extended InfoMax components orderer by their Kurtosis in AD patients and control subjects of the training set.	231
7.28	p -values of a one-way ANOVA with age as a covariate computed for the Extended InfoMax components ordered by their Kurtosis in the training set for (a) MF and SpecEn and (b) LZC and SampEn.	231
7.29	Average values of the signal features for the FastICA components orderer by their MF in AD patients and control subjects of the training set.	232
7.30	p -values of a one-way ANOVA with age as a covariate computed for the FastICA components ordered by their MF in the training set for (a) MF and SpecEn and (b) LZC and SampEn.	232
7.31	Average values of the signal features for the FastICA components orderer by their Kurtosis in AD patients and control subjects of the training set.	233
7.32	p -values of a one-way ANOVA with age as a covariate computed for the FastICA components ordered by their Kurtosis in the training set for (a) MF and SpecEn and (b) LZC and SampEn.	234
7.33	MEG channel from a control subject.	242

7.34	Example of the IMFs obtained from the EMD of a MEG channel acquired from a control subject. The IMFs are plotted following their order of extraction.	243
7.35	Example of the reference selection for the cBSS with the k -means clustering procedure.	244
7.36	Estimated marginal means plotted with the 95% confidence interval of the $\text{Coh}(f)$ for AD patients and control subjects in each spectral band.	245
7.37	Example of the estimated marginal means and the 95% confidence interval for the $\text{Coh}(f)$ of the pairs of regions in the δ band. Pairs over the same hemisphere are plotted in blue, whereas red is used to indicate pairs composed of the same regions located over different hemispheres. The black colour represents pairs composed of totally different scalp regions and hemispheres.	246
B.1	Promedio de la métrica F para cada par de algoritmos BSS (A: AMUSE, S: SOBI, J: JADE, eI: extended Infomax, F: FastICA) y longitud de época.	320

List of Tables

4.1	Summary of artefact detection studies in MEG with BSS. . . .	122
4.2	Summary of studies about the preprocessing of EEG and MEG data with spatial filtering techniques to increase the separation between subject groups.	125
5.1	AD patients' sociodemographic data.	132
5.2	Control subjects' sociodemographic data.	133
5.3	Demographic and clinical features for all participants, and for the training and test sets. Data are given as mean \pm SD. . . .	134
7.1	Average values of A_{PtP} and RMS indicating the importance of the CA in the original MEG recordings for the epoch lengths considered in this study. Data are shown as mean \pm SD. . . .	194
7.2	Number of peaks in the signals above ± 2 pT and the difference between the PSDs computed at 18 MEG channels close to the eyes and at the rest of MEG sensors indicating the importance of the OA in the original recordings for the epoch lengths considered in this study. Data are shown as median \pm SD. . . .	197
7.3	Average values of total and normalised spectral power in the band from 49 Hz to 51 Hz indicating the importance of the PLN in the original MEG recordings for the epoch lengths considered in this study. Data are shown as mean \pm SD. . . .	197
7.4	Average number of components marked for rejection for the epoch lengths and algorithms considered in this study with the metric Skew.	199
7.5	Average number of components marked for rejection for the epoch lengths and algorithms considered in this study with the metric VarSc.	199
7.6	Average number of components marked for rejection for the epoch lengths and algorithms considered in this study with the metric KrE.	201

7.7	Average number of components marked for rejection for the epoch lengths and algorithms considered in this study with the metric H_{Sh}	201
7.8	Average number of components marked for rejection for the epoch lengths and algorithms considered in this study with the metric $H_{Ré}$	201
7.9	Average number of components marked for rejection for the epoch lengths and algorithms considered in this study with the metric ApEn.	201
7.10	Average number of components marked for rejection for the epoch lengths and algorithms considered in this study with the metric P_{EYES}	202
7.11	Average number of components marked for rejection for the epoch lengths and algorithms considered in this study with the metric P_{LF}	202
7.12	Average number of components marked for rejection for the epoch lengths and algorithms considered in this study with the metric P_{50Hz}	202
7.13	Average values of A_{PtP} for the epoch lengths and BSS algorithms considered in this study after the CA removal based on the Skew. Data are shown as mean \pm SD.	203
7.14	Average values of RMS for the epoch lengths and BSS algorithms considered in this study after the CA removal based on the Skew. Data are shown as mean \pm SD.	204
7.15	Average values of A_{PtP} for the epoch lengths and BSS algorithms considered in this study after the CA removal based on the VarSc. Data are shown as mean \pm SD.	204
7.16	Average values of RMS for the epoch lengths and BSS algorithms considered in this study after the CA removal based on the VarSc. Data are shown as mean \pm SD.	204
7.17	Average values of A_{PtP} for the epoch lengths and BSS algorithms considered in this study after the CA removal based on the KrE. Data are shown as mean \pm SD.	205
7.18	Average values of RMS for the epoch lengths and BSS algorithms considered in this study after the CA removal based on the KrE. Data are shown as mean \pm SD.	205
7.19	Average values of A_{PtP} for the epoch lengths considered in this study after the CA removal based on the metric H_{Sh} . Data are shown as mean \pm SD.	206

7.20	Average values of RMS for the epoch lengths considered in this study after the CA removal based on the metric H_{Sh} . Data are shown as mean \pm SD.	206
7.21	Average values of A_{PtP} for the epoch lengths and BSS algorithms considered in this study after the CA removal based on the $H_{Ré}$. Data are shown as mean \pm SD.	206
7.22	Average values of RMS for the epoch lengths and BSS algorithms considered in this study after the CA removal based on the $H_{Ré}$. Data are shown as mean \pm SD.	207
7.23	Average values of A_{PtP} for the epoch lengths and BSS algorithms considered in this study after the CA removal based on the ApEn. Data are shown as mean \pm SD.	207
7.24	Average values of RMS for the epoch lengths and BSS algorithms considered in this study after the CA removal based on the ApEn. Data are shown as mean \pm SD.	207
7.25	Average values of A_{PtP} and RMS for the epoch lengths considered in this study after the CA rejection based on a cBSS procedure. Data are shown as mean \pm SD.	208
7.26	Number of peaks in the MEG signals above ± 2 pT for the epoch lengths and BSS algorithms considered in this study after the OA removal based on the P_{EYES} . Data are shown as median \pm SD.	210
7.27	Difference between the PSDs computed at 18 MEG channels near the eyes and at the rest of MEG sensors for the epoch lengths and BSS algorithms considered in this study after the OA removal based on the P_{EYES} . Data are shown as median \pm SD.	211
7.28	Number of peaks in the signals above ± 2 pT for the epoch lengths and BSS algorithms considered in this study after the OA removal based on the P_{LF} . Data are shown as median \pm SD.	211
7.29	Difference between the PSDs computed at 18 MEG channels close to the eyes and at the rest of MEG sensors for the epoch lengths and BSS algorithms considered in this study after the OA removal based on the P_{LF} . Data are shown as median \pm SD.	212
7.30	Number of peaks in the signals above ± 2 pT for the epoch lengths and BSS algorithms considered in this study after the OA removal based on the KrE. Data are shown as median \pm SD.	212

7.31	Difference between the PSDs computed at 18 MEG channels close to the eyes and at the rest of MEG sensors for the epoch lengths and BSS algorithms considered in this study after the OA removal based on the KrE. Data are shown as median \pm SD.	212
7.32	Number of peaks in the signals above ± 2 pT for the epoch lengths and BSS algorithms considered in this study after the OA removal based on the H_{Sh} . Data are shown as median \pm SD.	213
7.33	Difference between the PSDs computed at 18 MEG channels close to the eyes and at the rest of MEG sensors for the epoch lengths and BSS algorithms considered in this study after the OA removal based on the H_{Sh} . Data are shown as median \pm SD.	213
7.34	Number of peaks in the signals above ± 2 pT for the epoch lengths and BSS algorithms considered in this study after the OA removal based on the H_{Re} . Data are shown as median \pm SD.	213
7.35	Difference between the PSDs computed at 18 MEG channels close to the eyes and at the rest of MEG sensors for the epoch lengths and BSS algorithms considered in this study after the OA removal based on the H_{Re} . Data are shown as median \pm SD.	214
7.36	Number of peaks in the signals above ± 2 pT for the epoch lengths and BSS algorithms considered in this study after the OA removal based on the ApEn. Data are shown as median \pm SD.	214
7.37	Difference between the PSDs computed at 18 MEG channels close to the eyes and at the rest of MEG sensors for the epoch lengths and BSS algorithms considered in this study after the OA removal based on the ApEn. Data are shown as median \pm SD.	214
7.38	Average values of total spectral power in the band from 49 Hz to 51 Hz multiplied by 10^3 for the epoch lengths and BSS algorithms considered in this study after the PLN removal based on the $P_{50\text{Hz}}$. Data are shown as mean \pm SD.	217
7.39	Average values of normalised spectral power in the band from 49 Hz to 51 Hz multiplied by 10^2 for the epoch lengths and BSS algorithms considered in this study after the PLN removal based on the $P_{50\text{Hz}}$. Data are shown as mean \pm SD.	217

7.40	Average values of total spectral power in the band from 49 Hz to 51 Hz multiplied by 10^3 for the epoch lengths and BSS algorithms considered in this study after the PLN removal based on the KrE. Data are shown as mean \pm SD.	217
7.41	Average values of normalised spectral power in the band from 49 Hz to 51 Hz multiplied by 10^2 for the epoch lengths and BSS algorithms considered in this study after the PLN removal based on the KrE. Data are shown as mean \pm SD.	218
7.42	Average values of total spectral power in the band from 49 Hz to 51 Hz multiplied by 10^3 for the epoch lengths and BSS algorithms considered in this study after the PLN removal based on the H_{Sh} . Data are shown as mean \pm SD.	218
7.43	Average values of normalised spectral power in the band from 49 Hz to 51 Hz multiplied by 10^2 for the epoch lengths and BSS algorithms considered in this study after the PLN removal based on the H_{Sh} . Data are shown as mean \pm SD.	219
7.44	Average values of total spectral power in the band from 49 Hz to 51 Hz multiplied by 10^3 for the epoch lengths and BSS algorithms considered in this study after the PLN removal based on the $H_{R\acute{e}}$. Data are shown as mean \pm SD.	219
7.45	Average values of normalised spectral power in the band from 49 Hz to 51 Hz multiplied by 10^2 for the epoch lengths and BSS algorithms considered in this study after the PLN removal based on the $H_{R\acute{e}}$. Data are shown as mean \pm SD.	219
7.46	Average values of total spectral power in the band from 49 Hz to 51 Hz multiplied by 10^3 for the epoch lengths and BSS algorithms considered in this study after the PLN removal based on the ApEn. Data are shown as mean \pm SD.	221
7.47	Average values of normalised spectral power in the band from 49 Hz to 51 Hz multiplied by 10^2 for the epoch lengths and BSS algorithms considered in this study after the PLN removal based on the ApEn. Data are shown as mean \pm SD.	221
7.48	Ranges of components (estimated from the training set) selected to partially reconstruct the MEG signals for each feature and type of preprocessing (BSS algorithm and ordering criterion).	235
7.49	Classification results and AUC obtained from the MEG recordings without any preprocessing in the test set using the decision rules developed with the training set.	236

7.50	Classification results and AUC from the relative power in the α (8 Hz–13 Hz) and β_1 (13 Hz–19 Hz) bands in the original MEG recordings for the test set using the decision rules developed with the training set.	237
7.51	Classification results and AUC obtained from the α -filtered MEG signals in the test set using the decision rules developed with the training set.	237
7.52	Classification results and AUC obtained from the β_1 -filtered MEG signals in the test set using the decision rules developed with the training set.	237
7.53	Classification results and AUC obtained from the partially reconstructed MEG signals with the AMUSE preprocessing in the test set using the algorithms developed with the training set (Acc: Accuracy, Sen: Sensitivity, Spe: Specificity).	238
7.54	Classification results and AUC obtained from the partially reconstructed MEG signals with the SOBI preprocessing ordered following the MF in the test set using the algorithms developed with the training set (Acc: Accuracy, Sen: Sensitivity, Spe: Specificity).	239
7.55	Classification results and AUC obtained from the partially reconstructed MEG signals with the Extended InfoMax preprocessing ordered following the MF in the test set using the algorithms developed with the training set (Acc: Accuracy, Sen: Sensitivity, Spe: Specificity).	239
7.56	Classification results and AUC obtained from the partially reconstructed MEG signals with the FastICA preprocessing ordered following the MF in the test set using the algorithms developed with the training set (Acc: Accuracy, Sen: Sensitivity, Spe: Specificity).	240
7.57	Classification results and AUC obtained from the partially reconstructed MEG signals with the SOBI preprocessing ordered following the Kurtosis in the test set using the algorithms developed with the training set (Acc: Accuracy, Sen: Sensitivity, Spe: Specificity).	240
7.58	Classification results and AUC obtained from the partially reconstructed MEG signals with the Extended InfoMax preprocessing ordered following the Kurtosis in the test set using the algorithms developed with the training set (Acc: Accuracy, Sen: Sensitivity, Spe: Specificity).	240

7.59	Classification results and AUC obtained from the partially reconstructed MEG signals with the FastICA preprocessing ordered following the Kurtosis in the test set using the algorithms developed with the training set (Acc: Accuracy, Sen: Sensitivity, Spe: Specificity).	241
7.60	Classification results of the coherence measures computed from the adaptively extracted brain rhythms. Accuracy, sensitivity and specificity were calculated with a stepwise LDA. The number of variables automatically selected by this procedure and the related AUC are also shown.	247
8.1	Ratios of the number of data samples for each epoch length divided by the median value of the number of elements in A	255
8.2	Summary of the changes (expressed as ratios) in the metrics to assess the artefact removal. Lower ratios suggest better artefact removal.	262
8.3	Summary of the attenuation of artefactual activity in MEG recordings with BSS-based procedures.	264
8.4	Comparison of the accuracy and AUC values reached with various preprocessing techniques with the values obtained from the MEGs without any preprocessing.	269
8.5	Summary of the results reported by studies on the improvement of subject classification by means of a spatial filtering.	271
8.6	Summary of the classification results obtained from the adaptively extracted brain rhythms.	277
B.1	Principales datos sociodemográficos y clínicos de todo los participantes y la divisiones en grupos de entrenamiento y validación. Los datos se expresan como $\text{media} \pm \text{SD}$	315
B.2	Resultados de la clasificación con las medidas de coherencia calculadas a partir del procedimiento adaptativo de extracción de ritmos cerebrales.	322
B.3	Resumen de los cambios (expresados como ratios) en las métricas de evaluación del rechazo de artefactos. Ratios menores indican mejores rechazos de artefactos.	325
B.4	Comparación de los valores de precisión y AUC alcanzados mediante diversos preprocesados con los valores obtenidos a partir de las señales de MEG sin preprocesar.	327

Chapter 1

Introduction

Contents

1.1	Context: Biomedical Signal Processing	2
1.2	Magnetoencephalography	6
1.3	Alzheimer's Disease	8
1.4	Blind Source Separation	11
1.5	Hypothesis	13
1.6	Aims of the Doctoral Thesis	16
1.7	Schedule of Activities	19
1.8	Structure of this Dissertation	21

This Chapter introduces the rest of this document by presenting the concept of biomedical signal processing and its importance. Then, the basic ideas that lie at the core of this Doctoral Thesis are described. These include the magnetoencephalogram recordings, the Alzheimer's Disease and the blind source separation framework. Afterwards, the hypotheses beneath this Doctoral Thesis are listed and the aims pursued in this work, enumerated. In order to achieve those aims, a schedule of activities is proposed. Finally, the structure of the rest of this PhD dissertation is described.

1.1 Context: Biomedical Signal Processing

The field of this Doctoral Thesis is Biomedical Engineering and, to be more precise, biomedical signal processing. While the general term 'Bioengineering' comprises the research-oriented activity closely related to biotechnology and genetic engineering, the concept of 'Biomedical Engineering' is specifically associated with the development of new medical technology (Bronzino, 2006). This important multidisciplinary field aims at designing, developing and applying methods, materials and devices for clinical research and practice (Bronzino, 2006). Hence, biomedical engineers apply engineering principles and methodologies to understand, model and solve problems associated with medicine and biology. Despite the fact that its origins were mainly associated with the development of medical instruments, the range of activity of Biomedical Engineering is nowadays extremely varied. Among others, it includes (Bronzino, 2006):

- The detection, acquisition and monitoring of physiological signals.
- The diagnostic interpretation of biomedical recordings using signal processing techniques.
- The application of engineering system analysis (including modelling, simulation and control) to biologic and physiological problems.
- The computational analysis of data to help in clinical decision making and screening.

- The development of procedures and devices for rehabilitation and replacement or augmentation of bodily functions.
- The graphical display of anatomic or physiological information.
- The growth of artificial tissues and organs.

Thus, it is clear that Biomedical Engineering encompasses such different branches of expertise that close collaboration and interplay between diverse specialists, such as engineers, biologists and physicians, are needed (Bronzino, 2006).

As it was previously mentioned, this Doctoral Thesis falls into the field of *biomedical signal processing*. This scientific area deals with the acquisition and mathematical processing of informative signals derived from living systems or biological structures (Onaral, 2006).

Biomedical signals may originate from a variety of sources. Considering their origin, they can be classified into different groups such as bioelectric signals — such as Electroencephalogram (EEG), Electrocardiogram (ECG), Electrooculogram (EOG) and electromyogram —; biomagnetic signals — for instance, Magnetoencephalogram (MEG) and magnetocardiogram —; bioacoustic signals — for example, phonocardiogram and ecography—; biomechanical signals — e.g., mechanomyogram and aerial flux pressure—; biochemical signals — for instance, partial pressures of oxygen and carbon dioxide—; biooptical signals — for example, blood oxygen saturation—; etc. In addition to the former categorization, biomedical signals can also be divided according to their field of application (such as, neurology and cardiology) and their characteristics (like continuous and discrete, stationary and non-stationary signals; Cohen, 2006).

Most often in biomedical applications, the mere signal acquisition is not enough to obtain the information enclosed in it. Hence, specific actions must be applied to extract its relevant features. These procedures may be as simple as visual inspection of the signal or as complex as applying advanced signal processing techniques (Cohen, 2006). Hence, one of the major objectives of biomedical signal processing is to compute and extract signal characteristics.

Diverse signal processing techniques can help to characterise and understand physiological recordings. They are also essential to uncover signal components that may be very difficult, if not impossible, to observe by the naked eye (Laguna and Sörnmo, 2009). Biomedical signal processing also aims at reducing the subjectivity of the manual measurements. It must be noted that manual measurements made on biomedical recordings often lack accuracy, are time-consuming and depend completely on the specialist. This can reduce the accuracy of the diagnosis. By automating the measurement of the signal characteristics, subjectivity is reduced and reliability and reproducibility are increased. Finally, another major goal of biomedical signal processing is to reduce the noise present in the signals (Laguna and Sörnmo, 2009).

Generally, the recorded measurements are contaminated with undesired signals, which can interfere with the relevant physiological activity and bias the analyses. This contamination comes from both errors in the measurements produced by sensor imperfections and recorded activities from other biological systems that interfere with the desirable signal. The latter are commonly referred as artefacts (Escudero et al., 2007b; Jung et al., 2000). Hence, it is important to filter out undesired signal components of technical or physiological origin so that the Signal-to-Noise Ratio (SNR) is increased and further analysis are facilitated (Laguna and Sörnmo, 2009).

Taking into account that today's recording equipment and computers are powerful and relatively affordable, it is expected that the role of signal processing in those and other tasks will further increase in the future (Laguna and Sörnmo, 2009).

Generally speaking, the development of a typical biomedical measurement and processing system involves the following steps (Najarian and Splinter, 2006):

1. Identification of the relevant physical properties in biomedical system that will be measured by sensors.
2. Recording of a biomedical signal with a set of suitable sensors.

3. Filtering and processing of the biomedical signal to reduce the amount of noise and artefacts that contaminate it.
4. Extraction of features that represent or describe the status and conditions of the biomedical system. These features can be either based on the medical characteristics of the signal under study or defined using signal processing techniques.
5. Classification of the extracted features to obtain information about the system.
6. Finally, in some applications, other steps can be needed, such as a suitable visualisation of the information.

The biomedical signals that will be analysed in this Doctoral Thesis are known as Magnetoencephalogram (MEG) recordings. These discrete signals reflect the magnetic fields generated by the neural activity in the brain. Their main fields of application are neurology, psychiatry and psychology. Several Blind Source Separation (BSS) techniques will be applied to MEG signals acquired from patients with a diagnosis of probable Alzheimer's Disease (AD) and healthy elderly subjects with the following main objectives:

- To compare the similarity of the decompositions provided by several BSS algorithms for real MEG signals.
- To automatically and objectively assess the BSS-based removal of artefactual cardiac and ocular activities, as well as the power electrical noise, which contaminate the MEG recordings.
- To improve the classification of MEG signals acquired from AD patients against those recorded from healthy elderly subjects by means of a BSS preprocessing.
- To adaptively extract, with BSS methods, rhythmic brain activity suitable to be analysed with connectivity measurements from the MEG recordings.

1.2 Magnetoencephalography

The Magnetoencephalogram (MEG) is the *non-invasive* recording of the weak magnetic fields generated by the brain neural activity. In addition to the Electroencephalogram (EEG), it is the only neurophysiological technique able to *directly measure* the activity of the brain cortex without having to interpret the information on the basis of vascular or metabolic changes (Hari, 2004).

The recording of MEG signals is relatively new. Whereas the first EEG signal was acquired 80 years ago (Sanei and Chambers, 2007), the first measurement of MEG activity dates from just 1968 (Vrba and Robinson, 2001). It consisted of α -related activity recorded using a one-million-turn coil (Cohen, 1968). In fact, this measurement was carried out before any kind of superconductive sensor was available, thus resulting in poor-quality signals. However, superconductive sensors were soon available (Zimmerman et al., 1970) and the quality of the following MEG recordings was significantly improved (Cohen, 1972). The first MEG measurements were performed with a one-channel system. Hence, the equipment had to be moved to measure activity over different parts of the scalp. Eventually, novel instrumentation equipped with gradiometers was launched. These systems were able to simultaneously record the activity of one brain hemisphere with several channels (Hämäläinen et al., 1993). Of note is that current MEG systems are made of hundreds of channels (Stam, 2010; Vrba and Robinson, 2001).

Considering that EEG and MEG record the electromagnetic activity generated by the same primary currents in the brain, similarities between their waveforms are to be expected. The distributions of the electric potential and the magnetic field generated by a current dipole in a spherical volume conductor are dipolar, but they are rotated by 90° with respect to each other (Hari, 2004). Hence, there are differences in the sensitivity of both techniques to the orientation of the currents. MEG is mostly sensitive to tangential currents (Hämäläinen et al., 1993), whereas some current sources (very deep and radial) are more reliably picked up by EEG than MEG (Hari, 2004).

Both EEG and MEG can record the brain oscillations simultaneously all

over the scalp. Moreover, their temporal resolution is high, up to the millisecond time-scale (Hämäläinen et al., 1993). For MEG, this high temporal resolution can be accompanied of good spatial resolution in locating cortical events (Hari, 2004; Hari and Forss, 1999). Although both techniques are complementary to study the brain activity, the MEG recording offers some important advantages over the EEG. These include the facts that the magnetic recordings do not depend on any reference point (Hämäläinen et al., 1993; Hari et al., 2000; Najarian and Splinter, 2006) and that the resistive properties of the extra-cerebral tissues do not affect the magnetic fields (Hämäläinen et al., 1993; Hari et al., 2000).

On the other hand, the MEG recording is subjected to some considerable limitations, including the need for superconductive materials and magnetically shielded rooms to properly record the signals (Hari, 2004; Hari et al., 2000). Hence, the investment cost in the MEG system and shielded room is high, which has prevented any widespread use of this technique to acquire the brain activity (Sternickel and Braginski, 2006).

The MEG sensors are made of superconductive materials — called Superconductive QUantum Interference Devices (SQUIDS) — based on the Josephson Effect (Zimmerman et al., 1970). Moreover, the extremely weak amplitude of the brain magnetic fields needs the application of additional procedures to attenuate the noisy signals produced by extra-cerebral sources that affect these signals (Hari et al., 2000). This poses strong restrictions on the recording process and makes the MEG non-portable (Hari, 2004).

The most important noise affecting the MEG signals is that produced by outside disturbances caused by moving magnetic objects or electrical equipment (Hari et al., 2000; Vrba and Robinson, 2001). Bodily organs and movements also introduce *artefacts* in MEG signals. This includes the eye movements and blinks and the cardiac activity (Hämäläinen et al., 1993).

Normal brain activity generates magnetic fields whose amplitude ranges from 10 fT to 500 fT for evoked responses to a few pT (pico-Teslas) for epileptic spikes. However, the heart produces a significantly stronger field (Jousmäki and Hari, 1996). Obviously, the magnetic shielding does not help to reduce the Cardiac Artefact (CA) since the subject is within the shielded

room.

Furthermore, significant contamination can be caused by eye blinks and movements (Antervo et al., 1985). Both types of Ocular Artefacts (OAs) may be time-locked to the stimuli, specially if they are strong or infrequent (Hämäläinen and Hari, 2004). It must be noticed that the amplitude of the artefacts caused by eye movements can be about 3 pT to 4 pT above the lateral aspects of the orbits (Antervo et al., 1985).

Apart from the biological artefacts, the Power Line Noise (PLN) may also be a strong source of contamination in brain recordings (Escudero et al., 2007b; Iriarte et al., 2003; Jung et al., 2000). Moreover, removal of the PLN and its harmonics with notch filters may not be a satisfactory solution if the frequency band of the disturbances overlaps that of the signals (Hämäläinen et al., 1993; Sternickel and Braginski, 2006). In any case, prevention is always the best way to deal with artefacts.

It must be taken into account that MEG is not the only technique to study the brain activity. Other methods are available, such as the Positron Emission Tomography (PET), the Single Photon Emission Computed Tomography (SPECT), the Functional Magnetic Resonance Imaging (fMRI), the Magnetic Resonance Spectroscopy (MRS) and the EEG. However, EEG and MEG avoid not only the need for opening the skull to examine the subject's brain but also to be exposed to X rays, radioactive tracers or time-varying and strong static magnetic fields (Hämäläinen et al., 1993).

1.3 Alzheimer's Disease

Alzheimer's Disease (AD) is a slowly progressive and fatal neurodegenerative disorder with insidious onset. In addition to the well-known memory loss, AD also results in other cognitive and behavioural symptoms that progressively impair the activities of daily living. The cognitive deficit manifest itself as disorientation and problems with reasoning and thinking, whereas the behavioural symptoms include anxiety, delusions, depression, insomnia and wandering (Blennow et al., 2006; Cummings, 2004; Lahiri et al., 2002).

AD is the *most common type of dementia* in the Western World. It ac-

counts for 50% to 60% of all cases (Blennow et al., 2006). AD is one of the most disabling and burdensome health conditions worldwide. Its prevalence is below 1% in subjects aged 60 to 64 years, but it shows an almost exponential increase with age. The prevalence roughly doubles with every five years of age (Cummings, 2004). As a result, the prevalence of AD in people aged 85 years or older is between 24% and 33% (Blennow et al., 2006). In 2001, more than 24 million people had dementia and it is expected that this figure will double every 20 years up to 81 million in 2040 (Blennow et al., 2006). Moreover, it is estimated that 4.6 million new cases of dementia appear each year (Ferri et al., 2006).

It is clear that AD is a very important economic and social problem. It has been recently shown that dementia contributed more than 10% of years lived with disability in people over 60 (Ferri et al., 2006). This accounts for more cases than stroke (9.5%), musculoskeletal disorders (8.9%), cardiovascular disease (5.0%), and all forms of cancer (2.4%) (Ferri et al., 2006). Furthermore, the disability weight for dementia was higher than for any other health condition except for spinal-cord injury and terminal cancer (Ferri et al., 2006). It is expected that diagnostic testing for this dementia will contribute only marginally to this cost. Hence, the development of *methods to help in the diagnosis* may likely improve the care of patients and allow more accurately targeted future therapies. Eventually, this may lead to reduce the costs associated with this disorder (Lahiri et al., 2002).

AD patients usually die within 7 to 10 years after diagnosis (Masters et al., 2006). However, it is estimated that the neurodegeneration begins between 20 and 30 years before clinical onset (Blennow et al., 2006). During this preclinical phase, the number of microscopical lesions increases and, at a certain threshold, the first symptoms appear and gradually worsen (Blennow et al., 2006; Nestor et al., 2004). In the last stages of the dementia, AD patients are utterly dependent on their caregivers and cannot perform almost any task without assistance (Jeong, 2004; National Institute on Aging, 2008).

At the microscopic level, the hallmark lesions of AD are senile plaques (also known as neuritic plaques) and neurofibrillary tangles (Mattson, 2004). Additionally, the neurons and synapses degenerate (Blennow et al., 2006).

This process produces the characteristic cognitive deficit in AD. There is increasing consensus that the production and accumulation of Amyloid β peptide ($A\beta$) is central to the pathogenesis of this dementia (Cummings, 2004; Masters et al., 2006). The generation and deposition of $A\beta$ has been related to the formation of neurofibrillary tangles, oxidation and lipid peroxidation, glutamatergic excitotoxicity, inflammation and cell death (Cummings, 2004).

A *definite diagnosis* of AD can only be made by necropsy (the microscopic analysis of the patient's brain tissue after death; Blennow et al., 2006). In clinical practice, AD must be differentiated from other pathologies though. In fact, the criteria for detection of probable AD largely depend on the exclusion of other dementias. The AD diagnosis is usually based on the criteria developed by the National Institute of Neurological and Communicative Disorders and Stroke – Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) according to which the diagnosis is classified as definite, probable or possible (McKhann et al., 1984).

Nevertheless, the accuracy in the clinical diagnosis is *limited*. Even in patients who have been followed up clinically for several years at expert research centres, the diagnostic accuracy is relatively low, with sensitivity of around 80% and specificity of 70% (Blennow et al., 2006). Some authors even suggested lower typical sensitivity and specificity values for the diagnosis of probable AD (Cummings, 2004) and that these sensitivity and specificity figures may be significantly lower in primary care settings and in patients with mild AD (Blennow et al., 2006; Knopman et al., 2001).

The medical history together with the clinical, neurological, and psychiatric examination serves as the basis in the diagnostic work-up. Laboratory studies are necessary to identify secondary causes of dementia and coexisting disorders that are common in elderly people (Blennow et al., 2006; Cummings, 2004). Neuroimaging, including Computerised Axial Tomography (CAT) and Magnetic Resonance Imaging (MRI), is useful to exclude alternative causes of dementia (Cummings, 2004; Knopman et al., 2001). Additionally, tests are used to screen for cognitive impairment (Folstein et al., 1975). Moreover, new biomarkers would be of great value as diagnostic tools for the clinical diagnosis of AD (Blennow et al., 2006).

1.4 Blind Source Separation

The application of Blind Source Separation (BSS) in the field of biomedical signal processing has become popular in the last years (James and Hesse, 2005; Onton et al., 2006; Vigário and Oja, 2008). These techniques are helpful to separate multi-channel biomedical recordings into their constituent underlying components (James and Hesse, 2005; Vigário and Oja, 2008). Their use in biomedical signal processing frequently involves the separation or extraction of statistically independent sources underlying the multivariate recorded biomedical signals (James and Hesse, 2005; Vigário and Oja, 2008).

As it was describe in Section 1.1, biomedical signal processing deals with the recording and mathematical processing of informative signals derived from living systems (Onaral, 2006). Although biomedical signals may originate from a wide variety of sources, we focus now on electromagnetic measurements. Most often, the recorded signals are at least temporally or spatially correlated. Additionally, the informative components of the signal may be contaminated with noise and artefacts making them difficult to observe by the naked eye (Laguna and Sörnmo, 2009). That is, the recorded data contain measurements of a finite set of separate but overlapping in space and time activities. As a result, the Signal-to-Noise Ratio (SNR) of the interesting signal is generally poor (James and Hesse, 2005). While the clinicians may be able to visually extract the useful information from the signals, it would be extremely helpful to automate the analysis and isolate the constituent inner components of interest in the signals (James and Hesse, 2005).

BSS denotes a set of data-driven approaches that aim at extracting a set of underlying sources or components from a number of measurements or random variables. Very frequently, the term Independent Component Analysis (ICA) is also used to refer to the BSS techniques that are strictly based on Higher-Order Statistics (HOS) and attempt to minimise the statistical dependence of the data (Vigário and Oja, 2008). The BSS techniques define a generative model for the observed measurements and try to estimate the inner components from a large set of observed multivariate recordings (James and Hesse, 2005).

A few strong general assumptions must be made about the data at hand before applying the BSS algorithms and proceeding to properly interpret the results (James and Hesse, 2005). The most important hypothesis is the assumption that the sources are mutually *independent* or, alternatively, that they are mutually *decorrelated at any time delay* (James and Hesse, 2005; Onton et al., 2006). BSS extracts the sources, or components, by exploiting their independence or spatio-temporal decorrelation (Cichocki and Amari, 2002; Hyvärinen et al., 2001).

BSS is, indeed, a more powerful technique than classical decomposition methods such as Principal Component Analysis (PCA) and Factor Analysis (FA) (James and Hesse, 2005). For instance, the aim of PCA is to find temporally orthogonal directions that each successively explain as much of the remaining data variance as possible, where BSS tries to find directions in the joint data whose activities are as distinct from one another as possible. For this reason, PCA will *lump* whilst BSS will *split* the source activities across the resultant signal components (Onton et al., 2006).

The most well-known BSS algorithms were developed about 10 to 15 years ago (Cichocki and Amari, 2002; Hyvärinen et al., 2001). Soon after their development, these techniques were applied to diverse fields, including the study of cardiac and neurophysiologic recordings (James and Hesse, 2005). One of the most fruitful applications of BSS algorithms is the artefact rejection in EEG and MEG recordings. This is based on the hypothesis that the artefacts are independent from the brain activity (James and Hesse, 2005; Onton et al., 2006). Thus, the artefactual activity can be isolated into a few BSS components (Barbati et al., 2004; Cichocki and Amari, 2002; Delorme et al., 2007; Escudero et al., 2007b; Ting et al., 2006; Vorobyov and Cichocki, 2002). Nonetheless, this is not the only application of BSS to brain activity. These techniques have also been used to extract specific brain activity (Hung et al., 2007; Jin et al., 2002; Kobayashi et al., 1999). Despite these and other successful applications, it must be noticed that certain ambiguities exist in the BSS model and that the results must be interpreted carefully (James and Hesse, 2005).

1.5 Hypothesis

The different studies performed in this Doctoral Thesis are based on the corresponding hypotheses, which are introduced and discussed in the following lines.

1. Firstly, although the MEG measures the synchronous oscillations of the cortex directly and non-invasively, some problems arise when analysing these recordings (Hari, 2004). BSS algorithms are useful tools to overcome these some of these difficulties (Hyvärinen et al., 2001; James and Hesse, 2005; Onton et al., 2006). These techniques are able to decompose the observations assuming the mutual independence of the sources (James and Hesse, 2005; Onton et al., 2006; Vigário and Oja, 2008). Therefore, they have been widely applied to brain signals in various applications (James and Hesse, 2005; Onton et al., 2006; Vigário and Oja, 2008).

However, not all BSS algorithms are based on the same principles. On the one hand, this wide variety of techniques provides the researchers with very diverse approaches to perform the signal decomposition. On the other, it may make difficult to a priori select the most appropriate algorithm for a particular application (Fitzgibbon et al., 2007; Zavala Fernández et al., 2006).

A few studies have compared some BSS algorithms (see Klemm et al., 2009, and references therein). However, most of these analyses were based on artificially created signals. As a consequence, a detailed study on the similarity of the decompositions for real brain recordings computed with different algorithms is lacking (Onton et al., 2006). Thus, it can be hypothesised that the comparison of the outcomes provided by diverse BSS algorithms applied to the same real MEG signals can yield important information to further understand the relationships and similarities between BSS techniques.

2. Secondly, the MEG captures the neural activity with high temporal and spatial resolution by measuring the brain magnetic fields. This

technique is not invasive, it does not depend on any reference point and the recordings are less distorted than the EEG by the extra-cerebral tissues (Hämäläinen et al., 1993; Hari, 2004). However, non-cerebral sources (i.e., artefacts) always appear mixed with the desired brain signals in the MEG recordings (Hämäläinen et al., 1993; Hari, 2004).

The artefacts can bias the analyses, since their power may be larger than that of the brain sources (James and Hesse, 2005). For instance, the CA is very noticeable in MEG data (Jousmäki and Hari, 1996). OAs can also be visible in these signals (Antervo et al., 1985). Additionally, MEG data may have strong PLN contamination (Hari, 2004). Several methods have been used to remove these artefacts from MEG data, including: epoch rejection, spectral filtering, regression techniques and PCA. However, all these methods pose important problems (Jung et al., 2000).

An alternative is BSS, which has been used recently in the artefact rejection problem (Escudero et al., 2007b; James and Hesse, 2005; Vigário and Oja, 2008). Furthermore, the assumptions made about the data by BSS seem to be suitable for MEG recordings (James and Hesse, 2005; Onton et al., 2006; Vigário and Oja, 2008). Thus, it can be hypothesised that the application of BSS, together with some specific features designed to detect artefacts, can help to reduce the CA, OA and PLN in MEG background activity.

3. Thirdly, AD produces progressive and irreversible impairment of mental functions (Blennow et al., 2006; Cummings, 2004). This dementia is characterised by neural loss. Moreover, neuritic plaques and neurofibrillary tangles also appear in the patient's brain (Blennow et al., 2006; Mattson, 2004). The only definitive method for AD diagnosis is the direct pathological examination of brain tissue (Blennow et al., 2006). Nevertheless, a probable diagnosis based on neuroimaging, medical history studies and several mental tests is attempted (Blennow et al., 2006; Cummings, 2004).

Due to its relevance and difficult diagnosis, the utility of the EEG and

MEG brain activity in the detection of AD has been widely researched in the last decades (Hornero et al., 2009b; Jeong, 2004; Stam, 2005, 2010). Nevertheless, there is room for further improvement (Cichocki et al., 2005; Henderson et al., 2006). A technique that may improve the subject classification based on features extracted from MEG data is BSS (Cichocki et al., 2005; Escudero et al., 2008d, 2009b).

Considering the intrinsic complexity of the brain recordings, some BSS components may have certain features that could make them more sensitive to particular brain states, such as AD (Cichocki et al., 2005; Escudero et al., 2008d, 2009b). Hence, the most relevant components may be selected and the brain recordings may be partially reconstructed using only these components (Cichocki et al., 2005; Escudero et al., 2008d, 2009b). Thus, it can be hypothesised that a preprocessing based on BSS may help to differentiate AD patients' MEG background activity from that of control subjects when features such as Median Frequency (MF), Spectral Entropy (SpecEn), Lempel-Ziv Complexity (LZC) and Sample Entropy (SampEn) are used in the analysis.

4. Finally, it is important to note that AD not only modifies the signal spectrum and the non-linear properties of EEG and MEG signals (Hornero et al., 2009b; Jeong, 2004; Stam, 2005), but also is characterised by a 'disconnection syndrome' (Delbeuck et al., 2003; Jeong, 2004). Although several connectivity measures exist (Pereda et al., 2005), one of the most commonly applied measures is coherence, which quantifies the correlation between two signals as a function of frequency (Jeong, 2004). It has been found that AD patients show decreased coherence levels in α and β bands (Jeong, 2004).

Nevertheless, the domain where the brain recordings are analysed may be, at least, as important as the actual connectivity measure applied to the data. In addition to study the channels of the EEG and MEG recordings, the equivalent current dipoles can be analysed (Rossini et al., 2007). However, other approaches are possible. For example, the activity of interest within a particular spectral band located on a

particular scalp region may be extracted with Constrained Blind Source Separation (cBSS) techniques (James and Hesse, 2005).

In order to provide a suitable reference for the extraction, the Intrinsic Mode Functions (IMFs) derived from the Empirical Mode Decomposition (EMD) could be considered (Huang et al., 1998). The EMD decomposes complex signals into a reduced number of IMFs. Furthermore, this decomposition method is adaptive and, therefore, highly efficient (Huang et al., 1998). Thus, it can be hypothesised that the IMFs derived from a EMD can be suitable references to extract rhythmic activity from AD patients and healthy elderly subjects' MEGs with a cBSS. It can also be hypothesised that this extracted activity can be characterised by connectivity techniques.

1.6 Aims of the Doctoral Thesis

The MEG is the non-invasive technique that records the tiny magnetic fields produced by the brain neural activity. Moreover, it does not require the interpretation of the brain activity on the basis of vascular or metabolic changes (Hari, 2004). This recording is closely related to the EEG since both techniques acquire the electromagnetic activity generated by the same primary currents in the brain (Hari, 2004). The EEG and MEG are complementary to study the brain activity. However, the MEG recording has some important advantages over the EEG. For example, it does not depend on any reference point and it is less affected by extra-cerebral tissues (Hämäläinen et al., 1993; Hari, 2004; Hari et al., 2000).

AD is a slowly progressive and fatal neurodegenerative disorder. It is characterised by memory loss and cognitive and behavioural symptoms that progressively impair the activities of daily living (Blennow et al., 2006; Cummings, 2004; Lahiri et al., 2002). In the last stages of the dementia, the patients are utterly dependent on their caregivers and cannot perform almost any task without assistance (Jeong, 2004). AD is the most common type of dementia in the Western World (Blennow et al., 2006). Moreover, a

definite diagnosis of AD can only be made by necropsy and the accuracy in the clinical diagnosis is limited (Blennow et al., 2006). Hence, it is important to improve the diagnostic testing for this dementia. The development of methods able to help in the the diagnosis may likely improve the care of patients and allow more accurately targeted future therapies (Lahiri et al., 2002).

The processing of MEG recordings may be helpful in the detection of AD (Hornero et al., 2009b; Stam, 2005, 2010). However, some difficulties are encountered in the analysis of these signals. BSS techniques may be very useful tools to solve these problems (James and Hesse, 2005; Vigário and Oja, 2008). For instance, BSS could remove the artefact that contaminate brain recordings (Escudero et al., 2006c,d, 2007b, 2008c). These methods could also be applied as a preprocessing step to increase the accuracy in the classification of AD patients versus healthy elderly subjects (Escudero et al., 2006e, 2007c,d, 2008d, 2009b,c). Nevertheless, there is a wide variety of BSS techniques and it may be difficult to select a priori the most appropriate technique for a particular analysis. Hence, it is important to compare their outcomes when applied to the same real data (Escudero et al., 2009d). Finally, it can be expected that the extraction of rhythmic activity by means of cBSS techniques (Huang and Mi, 2007; Lu and Rajapakse, 2005) could offer relevant information for the connectivity analysis of the brain in this dementia.

Therefore, the main objective of this Doctoral Thesis is to **apply BSS techniques to help in the analysis and interpretation of MEG background activity from AD patients and elderly control subjects**. The basic assumption beneath this Doctoral Thesis is that some characteristics of the MEG signals may be more easily identified using the underlying components than in the actual recorded MEG channels. For example, the artefactual activity that contaminates the MEG channels and is overlapped in frequency with the brain signals may be isolated into single components by BSS. Similarly, some specific features of AD may be enhanced when a BSS preprocessing is applied to the MEG signals or when specific brain rhythms are extracted with BSS. In order to achieve these major objectives, the fol-

lowing partial objectives must be fulfilled:

1. To create a database containing MEG background activity recordings from AD patients and healthy elderly subjects.
2. To research into the state of the art of BSS techniques.
3. To select signal processing techniques suitable to decompose biomedical signals and, particularly, MEG background activity.
4. To process the MEG recordings with the selected BSS methods to:
 - Compare the outcome of several BSS algorithms when applied to real MEG background signals.
 - Automatically and objectively assess the removal of the CA, OAs and PLN that contaminate the magnetic brain activity.
 - Improve the discrimination between features extracted from control subjects and AD patients' MEG background activity recordings.
 - Extract rhythmic activity from the MEG recordings over diverse scalp regions to study the coherence between them.
5. To study using appropriate statistical tools and objective measurements:
 - The similarity between the decomposition provided by diverse BSS techniques.
 - The performance of the artefact removal in MEG recordings.
 - The improvement in the discrimination of AD patients and control subjects.
 - The differences in the spectral coherence patterns of both subject groups.
6. To publish the results of this Doctoral Thesis, if possible, in international scientific journals with high impact factors indexed in the Science Citation Index – Journal Citation Reports® (SCI–JCR).

1.7 Schedule of Activities

In order to achieve the objectives pursued in this Doctoral Thesis, the following schedule of activities is defined:

1. **Selection of patients with probable AD and healthy elderly control subjects.** The AD patients came from the “Asociación de Familiares Enfermos de Alzheimer” (Madrid) while the control subjects were voluntary recruited to participate in this research among the elderly subjects attending the “Unidad de Geriatria del Hospital Clínico Universitario San Carlos” (Madrid). The subject groups were carefully matched for age in order to minimise the influence of this parameter in the analyses.
2. **Recording of MEG signals from AD patients and control subjects.** Recordings are acquired with a 148-channels whole-head magnetometer (MAGNES 2500 WH de 4D Neuroimaging) placed in a magnetically shielded room of the “Centro de Magnetoencefalografía Dr. Pérez-Modrego” at Complutense University of Madrid. The signals are recorded while the subjects were lying on a patient bed, with eyes closed and in a relaxed state. Moreover, the participants were asked to avoid movements to minimise the number of artefacts. For each subject, five minutes of MEG background activity were recorded at a sampling rate of 678.17 Hz using a hardware filter with cut-off frequencies at 0.1 Hz and 200 Hz. Afterwards, data were down-sampled by a factor of four, resulting in an effective sampling frequency of 169.54 Hz.
3. **Development of a database with the MEG signals.** All MEG recordings, together with sociodemographic data (age and sex) and neuropsychological tests — Global Deterioration Scale / Functional Assessment Staging (GDA/FAST) and Mini-Mental State Examination (MMSE) — were stored in a database.
4. **Bibliographic research about methodologies based on BSS.** The aim is to compare the outcome of diverse BSS algorithms, remove

artefactual contamination of MEG signals, improve the classification of AD patients and healthy elderly subjects and extract rhythmic activity from MEG signals. It must be borne in mind that the algorithms will be applied to noisy biomedical signals composed of a large number of channels, such as the MEG recordings. Diverse techniques and analysis methods will be applied and compared at each stage in order to provide a comprehensive perspective of the analyses and their performance.

5. **Selection of the most suitable approaches on the basis of the previous bibliographic research** to assess their utility in the comparison of algorithms, artefact removal, classification of demented patients and study of connectivity in MEG signals.
6. **Processing of the MEG background activity recordings** from healthy elderly control subjects and patients with probable AD with the selected methods. It is expected that the application of BSS techniques will allow to:
 - Provide a comparison based on real brain signals of the similarity between the decompositions of diverse BSS algorithms.
 - Remove or minimise the amount of CA, OAs and PLN contaminating the MEG activity.
 - Improve the classification between signal features extracted from AD patients and control subjects' MEG signals.
 - Extract rhythmic activity with useful information about the coherence of diverse scalp regions from the MEG signals¹.
7. **Statistical analysis of the performance provided by the methods** included in this Doctoral Thesis. The main tools applied in this step will be ANalysis Of VAriance (ANOVA), Receiver Operating Characteristic (ROC) curves, box-plots and Linear Discriminant Analysis (LDA).

¹This study about the adaptive extraction of brain rhythms with EMD and cBSS and their connectivity was carried out during the PhD candidate's research stay at Cardiff University from 15 July 2009 to 31 October 2009 under the supervision of Dr. Sanei.

8. **Drawing of conclusions** from the obtained results and comparison with previously published research.
9. **Publication of the results** derived from these studies in national and international conferences and in international journals with high impact factor indexed in the SCI–JCR.

1.8 Structure of this Dissertation

Once the context of this Doctoral Thesis, the hypotheses beneath it and its main aims have been introduced, the structure of the rest of this dissertation is outlined. It comprises eight more chapters and three appendices. Their contents, in addition to the present Chapter 1, is described in the following lines:

- Chapter 2: Magnetoencephalography. The second Chapter begins explaining the basic principles beneath the generation of the brain magnetic field. Then, the particular details of the MEG recording equipment are explained and justified. Additionally, the nature of the main artefacts contaminating this kind of signals and its effects are presented. This Chapter finishes by comparing the MEG with other commonly used neuroimaging techniques.
- Chapter 3: Alzheimer’s Disease. This part of the dissertation summarises the concepts associated with AD. First of all, the social importance of this dementia is highlighted. Then, the main characteristics of AD are presented. This includes information about its evolution, pathophysiology, risk factors and diagnostic techniques. Finally, current and possible future treatments for AD are discussed.
- Chapter 4: Blind Source Separation. The fundamentals of the BSS are detailed in this Chapter. After an introduction about this topic, the classical projection techniques of PCA and FA are described for the sake of comparison with BSS. Afterwards, the principles of this framework are explained, considering both the Second-Order Statistics

(SOS)- and HOS-based approaches, as well as the cBSS. Next, the relationships and similarities among algorithms are briefly reviewed. Finally, the state of the art relevant to the applications considered in this Doctoral Thesis is detailed.

- Chapter 5: MEG Recordings. This Chapter describes the sociodemographic data of the 36 AD patients and 26 healthy elderly subjects included in this study and the diagnosis criteria. Additionally, the MEG recording process and basic preprocessing is explained.
- Chapter 6: Methods. Account of the methods and techniques applied in this Doctoral Thesis is given in this Chapter. It starts providing details about the preprocessing and dimension reduction techniques applied to the MEG data when needed. Additionally, the selection of the input parameters employed in the BSS algorithms is justified. Afterwards, the methodology used to compare the outcomes of the BSS is described. Finally, the metrics, features and signal processing techniques used in the other applications and analyses carried out in this Doctoral Thesis are explained.
- Chapter 7: Results. This Chapter presents the most relevant results obtained from the studies and methodology described in the previous Chapter.
- Chapter 8: Discussion. On the basis of the previously reported results, this Chapter discussed the main finding of this Doctoral Thesis considering the state of the art in the application of BSS techniques to MEG recordings and AD. It also acknowledges the main limitations of the studies carried out in this Doctoral Thesis
- Chapter 9: Conclusions and Future Research Lines. In this final Chapter, the main conclusions of this Doctoral Thesis are drawn. Additionally, the main contributions of this Thesis to the state of the art are pointed out and future research lines are suggested.

- Appendix A: Publications. This Appendix details the publications in journals, books and conferences in which the PhD candidate has collaborated thanks to his work in this Doctoral Thesis.
- Appendix B: Resúmenes en Español. This Appendix contains a summary, written in Spanish, of the first eight chapters of this dissertation.
- Appendix C: Conclusiones en Español. This Appendix is a translation into Spanish of the Chapter 9: Conclusions and Future Research Lines.

Chapter 2

Magnetoencephalography

Contents

2.1	Introduction	26
2.2	Generation of the Brain Magnetic Fields	27
2.2.1	Nervous System	27
2.2.2	Brain	28
2.2.3	Neurons and Neural Activity	30
2.2.4	Electromagnetic Brain Signals and Rhythms	33
2.2.5	Biomagnetism	35
2.3	MEG Recording Systems	36
2.3.1	SQUIDs and Flux Transformers	38
2.3.2	Cryogenic Container	40
2.3.3	Magnetically Shielded Room	41
2.3.4	Removal of the Magnetic Noise	42
2.4	Artefacts	44
2.4.1	Cardiac	45
2.4.2	Ocular	46
2.4.3	Power Line Noise	46
2.4.4	Other Artefacts	47
2.5	Comparison With Other Neuroimaging Methods	47
2.5.1	SPECT and PET	48
2.5.2	fMRI and MRS	50
2.5.3	EEG	51

This Chapter starts with a historical introduction about MEG. Then, the fundamentals of the brain magnetic fields generation are described. Additionally, the basic techniques supporting the recording of this signal are explained and the nature of the main artefacts contaminating the recordings is pinpointed. Finally, a comparison between MEG and other neuroimaging techniques is offered.

2.1 Introduction

The MEG is the non-invasive recording of the tiny magnetic fields generated by the brain neural activity. In addition to the EEG, it is the only neurophysiologic technique able to *directly measure* the activity of the brain cortex, without having to interpret the information on the basis of vascular or metabolic changes (Hari, 2004).

The recording of MEG signals is relatively new. Whereas the first EEG signal was acquired 80 years ago (Sanei and Chambers, 2007), the first measurement of MEG activity dates from just 1968 (Cohen, 1968). It consisted of α -related activity recorded using a one-million-turn coil (Cohen, 1968). In fact, this measurement was carried out before any kind of superconductive sensor was available, thus resulting in poor-quality signals. However, superconductive sensors were soon available (Zimmerman et al., 1970) and the quality of the following MEG recordings was significantly improved (Cohen, 1972).

The first MEG measurements were performed with a one-channel system. Hence, the equipment had to be moved to measure activity over different parts of the scalp. Eventually, novel instrumentation equipped with gradiometers was launched. These systems were able to simultaneously record the activity of one brain hemisphere with several channels (Hämäläinen et al., 1993). Current MEG systems contain hundreds of channels in a helmet-shaped array and they can operate while the subject is either seated or supine (Stam, 2010; Vrba and Robinson, 2001). Moreover, MEG is now being accepted in routine clinical practice throughout the world since it can be useful in pre-surgical mapping of epilepsy and patients with brain tu-

mours. Additionally, MEG can also help to reveal how the brain processes information (Stufflebeam et al., 2009).

2.2 Generation of the Brain Magnetic Fields

MEG signals are recorded on the surface of the scalp. Although they reflect the currents generated in the functioning part of the brain, most of the signal comes from the *cerebral cortex*. This part of the brain contains well-aligned pyramidal cells, which consist of dendrites, the cell body and an axon. A set of axons composes the nerve fibres which interconnect diverse regions of the brain. The information is transmitted from the dendrites to the cell body through the synapses. Of note is that there are approximately 10^5 to 10^6 cells in an area of 10 mm^2 of cortex. In the whole brain, there are approximately 10^{11} cells and about 10^{14} synaptic connections (National Institute on Aging, 2008; Vrba and Robinson, 2001). The number of synapses per neuron increases with age, whereas the number of neurons decreases (Sanei and Chambers, 2007).

2.2.1 Nervous System

The nervous system is a complex network that communicates diverse nervous tissues whose function is to transmit information in order to control the bodily activities and functions. It enables the body to detect and react to changes in the environment and to interpret the resulting nervous impulses. Attending to anatomical and physiological criteria, the nervous system can be divided into two systems (Carr and Brown, 2001; National Institute on Aging, 2008):

- Central Nervous System (CNS). It is the largest part of the nervous system. It is contained in the skull and spinal cavity, which protect the brain and the spinal cord, respectively.
- Peripheral Nervous System (PNS). It is composed of the nervous tissue that extend from or arrive at the skull and spinal cavity. The PNS

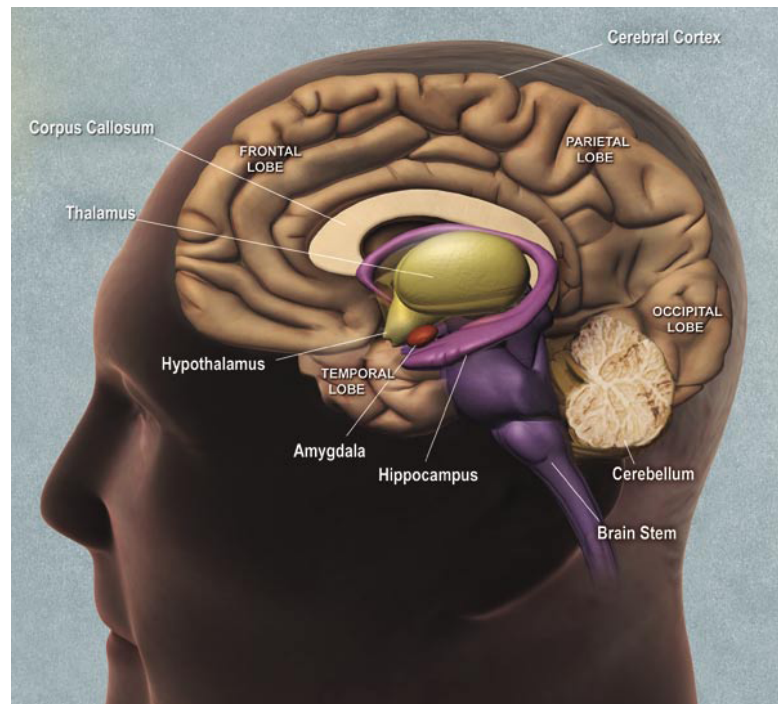


Figure 2.1: View of the brain showing some of its main parts. Adapted from (National Institute on Aging, 2008).

extends through the limbs and the flesh of the torso. It can be divided into:

- Somatic nervous system: responsible for coordinating voluntary body movements.
- Autonomic nervous system: responsible for coordinating involuntary functions.

2.2.2 Brain

The brain is the largest part of the CNS. As shown in Fig. 2.1, it is composed of several structures, which will be introduced in the following lines.

- The brain stem sits at the base of the brain. It connects the spinal cord with the rest of the brain, transmitting the nervous impulses between

them (National Institute on Aging, 2008). It controls involuntary functions such as respiration, heart regulation, biorhythms, hormone secretions and the sleep cycle (National Institute on Aging, 2008; Sanei and Chambers, 2007).

- The cerebellum lays above the brain stem (National Institute on Aging, 2008). It coordinates voluntary movements of muscles and maintains balance (Sanei and Chambers, 2007). The cerebellum receives information about the body's movements and position to process it and send instructions to the body through the rest of the CNS. It is also involved in motor learning (National Institute on Aging, 2008).
- The cerebrum is the most evolved part of the brain. It is divided into two cerebral hemispheres connected by nerves at the corpus callosum (National Institute on Aging, 2008). The cerebrum includes the regions for movement initiation, conscious awareness of sensation, complex analysis and expression of emotions and behaviour (Sanei and Chambers, 2007). The outer layer of the hemispheres is the *cortex*, which plays a key role in memory, attention, thought, language and other cognitive functions (National Institute on Aging, 2008). Attending at the most prominent fissures of the cortex, each hemisphere can be divided into four lobes, each of which has different roles (National Institute on Aging, 2008):
 - The frontal lobe, which lays at the front of the brain, controls executive functions like thinking, planning, and problem solving, as well as memory, attention and movement.
 - The parietal lobe, which sits behind the frontal lobe, deals with the perception and integration of stimuli from the senses.
 - The temporal lobe, which runs along the side of the brain under the frontal and parietal lobes, deals with the senses of smell, taste and sound and the formation and storage of memories.
 - The occipital lobe, which is at the back of the brain, is concerned with vision.

Other important parts of the brain lie deep inside the cerebral hemispheres forming a network of structures called the limbic system, which connects the brain stem with the cognitive elements of the cerebral cortex. The limbic system plays a key role in developing and carrying out instinctive behaviours and emotions. The limbic system includes ([National Institute on Aging, 2008](#)):

- The amygdala, involved in processing and remembering strong emotions such as fear.
- The hippocampus, which is important for learning and short-term memory as it is hypothesised that this part of the brain converts short-term into long-term memories for storage in other brain areas.
- The thalamus, which receives and processes sensory and limbic information to send it to the cerebral cortex.
- The hypothalamus, which monitors activities such as body temperature and food intake and controls the body's internal clock.

After briefly explaining the basic structure of the most important parts of the CNS, Section 2.2.3 will provide information about its basic unit, the neuron, and the electrical activity associated with it.

2.2.3 Neurons and Neural Activity

Generally speaking, the CNS is composed of glial and nerve cells ([Sanei and Chambers, 2007](#)). The glial cells are located between neurons, holding them in place and supporting them with nutrients. They also remove damaged cells and other debris and provide insulation to the neurons. However, the transmission and processing of the information in the nervous system is only carried out by the neurons ([National Institute on Aging, 2008](#)).

Typically, each neuron consists of axon, cell body (or soma) and dendrites ([National Institute on Aging, 2008](#)):

- The neuron's cell body has a single nucleus and contains most of the cell metabolism.

- The dendrites are branched projections of the cell body that transmits the stimulation between other neurons cells and the soma ([National Institute on Aging, 2008](#)). The dendrites are connected to either the axon or dendrites of other cells. In the human brain, each nerve is connected to approximately 10000 other neurons, mostly through dendritic connections ([Sanei and Chambers, 2007](#)).
- The axon is a long cylinder, which transmits an electrical impulse. Its length may vary from a fraction of a millimetre to more than a metre ([Sanei and Chambers, 2007](#)). Axons are covered with an insulating layer called myelin, which is produced by glial cells. Myelin is essential to the proper functioning of the nervous system as it increases the speed of nerve signal transmissions through the axon ([National Institute on Aging, 2008](#)).

Of note is that the activity recorded in the EEG and MEG is the ultimate result of the *transmission of the nerve impulse* or action potential. Action potentials are temporary changes in the neuron membrane potential which are transmitted along the axon. They are produced as a result of an exchange of ions across the neuron membrane and the stimulus for their creation is the chemical activity of neurotransmitters at the synapses. The conduction velocity of action potentials ranges from 1 m/s to 100 m/s ([Sanei and Chambers, 2007](#)).

A schematic of an action potential is shown in Fig. 2.2. After the peak of the previous nerve impulse, the membrane repolarises. That is, it becomes more negative than the resting potential. After approximately two milliseconds, the potential returns to normal and another stimulus may be presented ([Sanei and Chambers, 2007](#)). This is called refractory period. Then, in the presence of a stimulus from an excitatory synapse, an Excitatory Post-Synaptic Potential (EPSP) occurs and the membrane potential may depolarise (become more positive), producing a spike ([Sanei and Chambers, 2007](#)).

It must be noted that the stimulus must be above a certain threshold level to produce an action potential. Weak stimuli can cause a small local

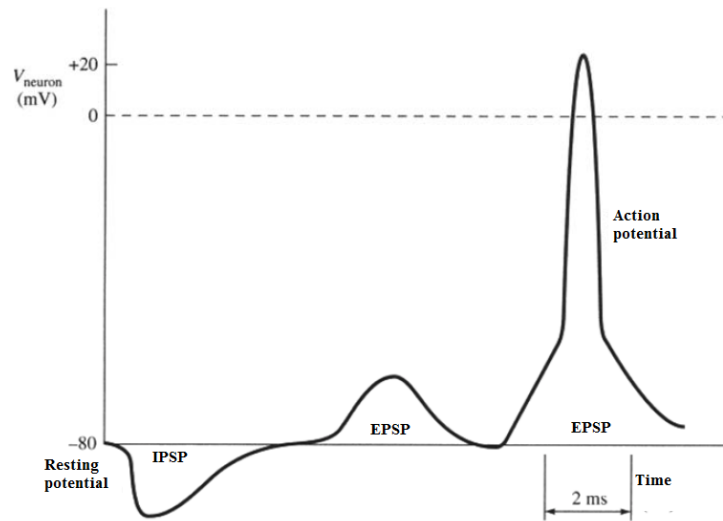


Figure 2.2: Example of action potential in a neuron.

electrical disturbance, but this does not produce a transmitted impulse. As soon as the stimulus strength is big enough, an action potential appears and travels down the nerve. On the contrary, if the fibre ends in an inhibitory synapse, hyperpolarisation will occur, indicating an Inhibitory Post-Synaptic Potential (IPSP). The action potentials of most nerves last between 5 and 10 milliseconds (Sanei and Chambers, 2007).

Changes in the membrane permeability produce the opening or closing of ionic channels in the cell membrane. This alters its resting state, where the membrane potential is about -80 mV. This potential is supported by a non-uniform distribution of ions sodium (Na^+) and chlorine (Cl^-) outside the neuron and ions potassium (K^+) and certain proteins inside it (Carr and Brown, 2001). The Na pump produces gradients of the ions Na^+ and K^+ . These ions are used in the mechanism that creates the action potential. Neurons have special Na^+ and K^+ voltage-gated channels. When the ionic Na^+ channel opens, the positively charged Na^+ ions rush into the cell. This makes the membrane potential positive (depolarisation), producing the spike (Sanei and Chambers, 2007). As a consequence, the polarity reverses and this provokes the opening of the K^+ channel. Then, the ions K^+ move out of the cell in order to equalise the concentration of charges and reach again the

resting state. However, as these K^+ channels are much slower to open, the depolarisation has time to be completed (Sanei and Chambers, 2007).

The repolarisation typically overshoots the resting potential. This hyperpolarisation is important in the transmission of the information as raises the neural threshold for any new stimulus and prevents any stimulus already sent up an axon from triggering an impulse in the opposite direction (Sanei and Chambers, 2007). This is the reason why the pulse propagates in only one direction: from the dendrites to the end of the axon. There, the neurotransmitters are segregated which cause the opening of the ionic channels in the post-synaptic cell after moving across a tiny gap (or synapse) to specific receptors on the receiving, or post-synaptic, end of dendrites of nearby neurons (National Institute on Aging, 2008). This stimulus produces another action potential in the next neuron (Carr and Brown, 2001).

2.2.4 Electromagnetic Brain Signals and Rhythms

When neurons are activated, the synaptic currents are produced within the dendrites and propagated through the soma and axon. This current generates a magnetic field measurable by MEG equipment and a secondary electrical field over the scalp measurable by EEG equipment. These fields usually lie below 100 Hz (Sanei and Chambers, 2007). However, only when large populations of active neurons are simultaneously active, a sufficiently large signal is generated to be recorded on the scalp (Sanei and Chambers, 2007).

The cortex can generate electromagnetic activity spontaneously. This activity is composed of *superimposed rhythms*. It must be noticed that, the higher the synchronisation of nearby neurons, the larger the amplitude and lower the frequency of the resulting activity. In healthy adults, the amplitudes and frequencies of such signals change from one state of a human to another, such as wakefulness or sleep (Sanei and Chambers, 2007). The characteristics of the waves also change with age. Usually, five major brain waves are distinguished by their frequency range in the brain activity: δ (0.5 Hz – 4 Hz), θ (4 Hz – 8 Hz), α (8 Hz – 13 Hz), β (13 Hz – 30 Hz) and γ (over 30 Hz) (Sanei and Chambers, 2007).

The δ activity ranges from about 0.5 Hz to 4 Hz. It is mainly associated with deep sleep but it may also be present in the waking state. Sometimes, this activity may be confused with artefact signals caused by large muscles of the neck and jaw (Sanei and Chambers, 2007).

The θ waves lie within the range of 4 Hz to 8 Hz. They have been linked with access to unconscious material, creative inspiration and deep meditation. The θ wave is often accompanied by other frequencies and seems to be related to the level of arousal. It plays an important role in infancy and childhood. Large amounts of θ activity in the waking adult are abnormal and it may indicate pathological problems (Sanei and Chambers, 2007). It is also remarkable that changes in the θ rhythms are analysed for maturational and emotional studies (Sanei and Chambers, 2007).

The α waves appear in the posterior half of the head. Moreover, they are dominant over the occipital region of the brain. Their frequency lies within the range of 8 Hz to 13 Hz. These waves usually appear as a round or sinusoidal shaped signal (Sanei and Chambers, 2007). This rhythm has been related to both a relaxed awareness without any attention and concentration. Most subjects produce α waves with their eyes closed. For this reason, it has been argued that it actually is a waiting or scanning pattern produced by the visual regions of the brain. α activity is greatly reduced by opening the eyes, hearing unfamiliar sounds, anxiety or attention (Sanei and Chambers, 2007).

The β rhythm denotes the brain electromagnetic activity varying between 13 Hz and 30 Hz. It is the usual waking rhythm of the brain associated with active thinking, active attention or solving problems (Sanei and Chambers, 2007). β activity is found in normal adults, mainly over the frontal and central regions. It may also be enhanced around tumoural regions (Sanei and Chambers, 2007).

The γ range corresponds to frequencies over 30 Hz. The amplitude of this rhythm is very low and its occurrence is rare. However, detection of this activity can be used for confirmation of certain brain diseases (Sanei and Chambers, 2007).

2.2.5 Biomagnetism

The generation of MEG signals has some close similarities to that of EEG activity. The exchange of ions through the cell membrane produces an electric field with its corresponding magnetic field. Due to the ionic exchange between the cell and its environment, the equilibrium between diffusion processes and electrical forces is reached at a negative potentials of about -70 mV within the cell (Vrba and Robinson, 2001). As previously explained, cell stimulation can alter this potential and produce cell depolarization (or hyperpolarization). Due to the conductivity of the cell, this generates a current flow within the cell, which is known as impressed or intracellular current, and a return current outside the cell, called volume or extracellular current (Vrba and Robinson, 2001).

The extra- and intra-cellular currents flow towards the region with less electrical resistance. The former flow from the dendrites to the soma, and then they are radially distributed, which decreases their intensity. On the other hand, intracellular currents flow from the soma to the dendrites to close the loop. Hence, their intensity is kept constant. As a consequence, the intra-cellular currents are considered the ultimate cause of the magnetic field acquired with the MEG. This is illustrated in Fig. 2.3, where the recorded magnetic field is mainly due to the intra-cellular current, whereas the EEG measures the volume currents and, to a lesser extent, the intracellular ones (Maestú et al., 1999; Rampp and Stefan, 2007; Vrba and Robinson, 2001).

It must also be considered that, although the flow of the dendritic current due to cell depolarization (or hyperpolarization) is approximately perpendicular to the cortex, this is convoluted with multiple sulci and gyri. Hence, the current flow may be either radial or tangential to the surface depending on where it occurs, as it is exemplified by Fig. 2.4 (Vrba and Robinson, 2001).

The magnetic field generated by the current flow of a single cell is too small to be measured outside the brain. Only if, approximately, 10^5 to 10^6 cells are simultaneously activated, the field can be detected (Maestú et al., 1999; Vrba and Robinson, 2001).

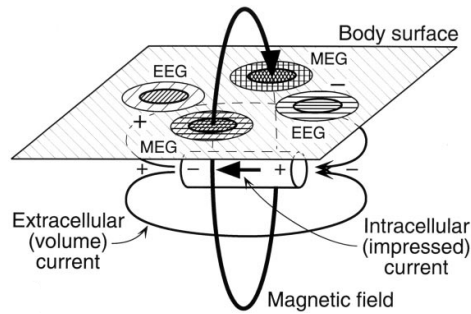


Figure 2.3: Orthogonality between EEG and MEG signals as illustrated by Vrba and Robinson (2001).

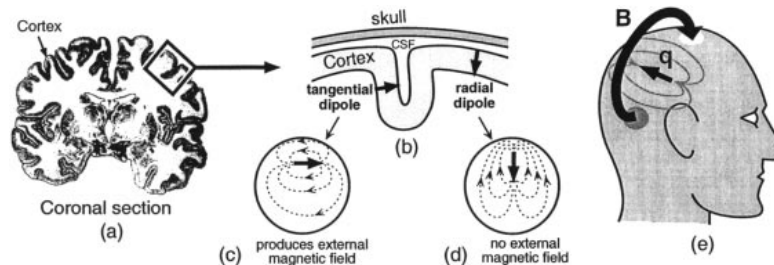


Figure 2.4: Origin of the MEG. (a) Section of the human brain. (b) The sulci and gyri of the cortex produce either tangential or radial currents. (c) Tangential currents produce magnetic fields observable outside the head. (d) Radial currents do not produce magnetic fields outside the head. (e) Magnetic fields due to cortical sources will exit and re-enter the scalp. Taken from (Vrba and Robinson, 2001).

2.3 MEG Recording Systems

When recording the subject's brain magnetic activity, the position of the head must be known exactly. In order to do so, an instrument called Head-Position Indicator (HPI) is used. The HPI provides information about the exact measurement sites and the orientations of the sensors with respect to the head. These data are obtained by placing three or more small coils on known sites on the scalp. Then, the field produced by currents led through these coils is measured by the MEG equipment. As an alternative, the HPI may be composed of a transmitter connected to magnetic sensors, which are fixed on the dewar, and three receivers are placed on the subject's head. In either case, the HPI provides information about the position of the MEG

equipment with respect to the head with 2- to 3-mm accuracy (Hari, 2004).

Once the position of the head is known, the helmet that contains the array of sensors is placed over the subjects' scalp while they are typically sitting on a chair or lying on a patient bed (Hari, 2004). The main sensory organs of the head remain uncovered by the helmet so that the magnetic brain responses to certain external stimuli can be acquired in addition to the recording of the brain background activity. Nevertheless, it is of utter importance to keep the head immobile during the recording process. Hence, MEG activity cannot be measured during major motor seizures. Similarly, tasks involving head movements and long-term monitoring are not feasible. Problems can also appear in studies with uncooperative subjects who either cannot keep still during the recording or are unwilling to perform the tasks (Hari, 2004; Hari et al., 2000).

As it was previously mentioned, the MEG sensors are made of *superconductive materials* — called SQUIDs — based on the Josephson Effect (Zimmerman et al., 1970). These devices are extremely sensitive to tiny magnetic fields such as those produced by the neurons. However, their configuration is not best suited for the direct detection of the brain magnetic fields. Thus, superconductive flux transformers are needed to couple the brain fields into the SQUIDs sensors, thus increasing the system sensitivity (Vrba and Robinson, 2001).

The temperature of the sensors must be kept close to the absolute zero in order to ensure their proper performance. Therefore, these devices are immersed in liquid helium at -269°C (Hämäläinen and Hari, 2004; Vrba and Robinson, 2001). The thermally insulated container (known as dewar) must be refilled periodically to replace the evaporating helium, approximately once every week (Hari, 2004). Obviously, the dewar must be electromagnetically transparent so that the signals can reach the superconductive equipment. Finally, the signals recorded by the SQUIDs are amplified and processed. The whole set-up can also include a video-recorder and an intercom for observation and communication with the subject in the magnetically shielded room (Vrba and Robinson, 2001). These elements are represented in Fig. 2.5.

In addition to requiring superconductive materials in the MEG sensors,

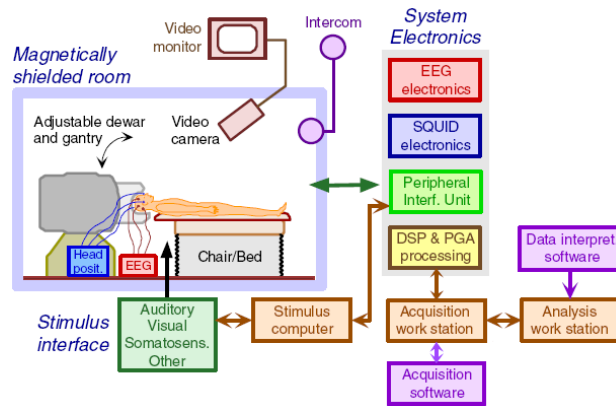


Figure 2.5: Block diagram of a MEG system. Adapted from (Sternickel and Braginski, 2006).

the extremely weak amplitude of the brain magnetic fields needs the application of additional procedures to attenuate the noisy signals produced by extra-cerebral sources that can affect these signals. It should be noticed that the MEG signals are usually in the fT (10^{-15} T) range whereas, for example, the steady magnetic field of the Earth is about 5×10^{-5} T (Hari et al., 2000), as it is shown in Fig. 2.6. This poses strong restrictions on the recording process and makes the MEG non-portable. In addition, the installation and maintenance of an MEG equipment is costly (Hari, 2004).

2.3.1 SQUIDS and Flux Transformers

The SQUID sensor are at the core of a MEG recording system. They are essential to provide the *high sensitivity to detect the tiny magnetic fields* generated by the brain. The most popular types of SQUIDS are the dc and the rf SQUIDS. These names come from the methods of their biasing, which stand for direct current and radio frequency, respectively (Vrba and Robinson, 2001).

A dc SQUID is a ring of superconductive material interrupted by two resistively shunted Josephson junctions, which are based on the Josephson effect (Vrba and Robinson, 2001). This effect relies on the properties that exhibit two pieces of superconductive material separated by a thin film of

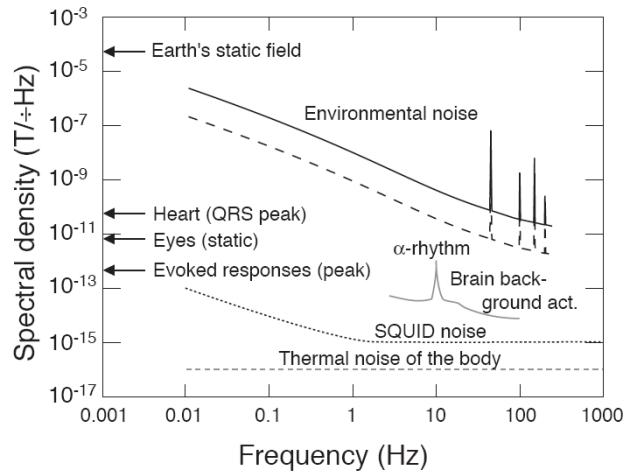


Figure 2.6: Power of undesired components (external noise and artefacts) and brain signals as a function of frequency. Adapted from (Hämäläinen and Hari, 2004).

isolating material. This film acts as a barrier which opposes to the current. Hence, the ring operates as a resistive component. However, if the current is small enough, the ring behaves as a superconductive material. Moreover, the level of this current is modified by magnetic fields. Thus, the SQUIDS are able to detect changes in the magnetic activity (Maestú et al., 1999; Vrba and Robinson, 2001).

On the other hand, the rf SQUIDS were popular in the early days of magnetometry. This was due to the fact that they only required one Josephson junction instead of two, like the dc SQUIDS. Nevertheless, rf SQUIDS have been displaced by dc SQUIDS in modern MEG equipment since the sensibility of the latter to the magnetic field is 10 times higher (Vrba and Robinson, 2001). Recently, there has been some interest in rf SQUIDS as a result of research in high temperature superconductivity (Vrba and Robinson, 2001).

Furthermore, flux transformers are necessary to help coupling SQUID sensors to the measured signals. In this way, the overall magnetic field sensitivity is increased. Flux transformers are superconductive devices made of one or more pick-up coils exposed to the measured fields and a coupling coil which inductively couples the flux transformer to the SQUID ring. Thanks to the fact that the flux transformers are superconductive, they do not add

noise to the system (Vrba and Robinson, 2001).

The flux transformer pick-up coils may have diverse configurations, including radial and tangential magnetometers, and diverse types of gradiometers, which are composed of two magnetometers connected in series but wired in opposition. This configuration is sensitive only to changes in the magnetic field across the device dimension. These devices are called first-order gradiometers, which can be combined with opposing polarity to form second-order gradiometers, and so on (Hämäläinen et al., 1993; Vrba and Robinson, 2001).

An important function of flux transformers is to help reduce environmental noise. For this reason, the most commonly used flux transformers are the gradiometers instead of magnetometers, since the former help to attenuate the signals produced by distant noise sources (Vrba and Robinson, 2001).

2.3.2 Cryogenic Container

The MEG sensors and their associated equipment are built with superconductive materials. Nowadays, the commercial equipment using this type of devices needs to keep the temperature of these elements close to the absolute zero. These temperatures can be achieved either with cryocoolers or with a cryogenic bath in contact with the superconductive components.

Although the cryocoolers may be an attractive solution because they eliminate the need for periodic refilling of the cryogenic container (Vrba and Robinson, 2001), they contribute to electromagnetic interference. Hence, they are not suitable for MEG systems. Therefore, current commercial systems rely on cooling the devices by placing them in a bath of liquid He in a dewar. The dewar itself is a complex device that is composed of several layers, both thermally insulating and heat-conductive (Vrba and Robinson, 2001).

Due to the fact that the thermal difference between the environment and the He liquid is about 300°C, tens of layers are needed to isolate the container of the dewar from the room temperature. The cold gases resulting from the evaporating He transport energy that is captured in the dewar neck and



Figure 2.7: Cryogenic container as part of a MEG system.

conducted back into the dewar vacuum space to reduce the thermal gradient between inside and outside the dewar. Current designs are highly efficient and MEG systems only consume about 10 litres of liquid He per day (Vrba and Robinson, 2001). A cryogenic container as part of a MEG recording system is illustrated in Fig 2.7.

2.3.3 Magnetically Shielded Room

The most straightforward and reliable way to reduce the importance of the external noise is to enclose the MEG equipment in a magnetically shielded room (Hämäläinen et al., 1993; Vrba and Robinson, 2001). One example of this kind of room is shown in Fig 2.8. The walls of these rooms are composed of several sheets of μ -metal (an alloy of nickel and iron with a high magnetic permeability) and high conductivity aluminium (Sternickel and Braginski, 2006; Vrba and Robinson, 2001). These rooms are able to attenuate the external noise by factor of at least 100 at 0.1 Hz and well over 10^5 for frequencies higher than 100 Hz (Sternickel and Braginski, 2006).



Figure 2.8: Example of magnetically shielded room.

2.3.4 Removal of the Magnetic Noise

As a consequence of the extremely weak amplitude of the brain magnetic fields and the different sources of noise and artefacts that can contaminate the signals (illustrated in Fig. 2.6), diverse procedures are needed to attenuate this undesirable contamination. There are mainly three types of approaches to reduce the influence of these noise sources in the MEG signals (Sternickel and Braginski, 2006; Vrba and Robinson, 2001):

- Electromagnetic shielding. As it was mentioned in Section 2.3.3, the most straightforward and reliable way of reducing the importance of external magnetic disturbances is to perform the measurements in magnetically shielded rooms (Hämäläinen et al., 1993). As a complement to the passive magnetic shielding, active shielding can be used to increase the attenuation at low frequencies. Typically, this consists of reference sensors used to measure the residual noise. Their output is employed to cancel this noise with opposing fields of large orthogonal coils via suitable negative feedback circuitry (Hämäläinen and Hari, 2004; Sternickel and Braginski, 2006).

- Gradiometer systems. As explained in Section 2.3.1, these devices try to reject the external magnetic disturbances (Hämäläinen and Hari, 2004). For instance, an axial first-order gradiometer consists of a pick-up (lower) coil and a compensation (upper) coil. They are identical in area and are connected in series, but wound in opposition. The pick-up coil is closer to the desired, useful source to be measured (i.e., the brain) whereas the other is at least 3 cm to 4 cm away. Both sensors record a measurement composed of the signal of interest plus some background noise. While the amplitude of signal of interest is much stronger at the coil close to the source than at the remote one, the environmental noise can be considered similar at both recording places (Hämäläinen and Hari, 2004; Hämäläinen et al., 1993). The difference of both measurements allows to reject the uniform noise by a factor of 10^2 to 10^4 , thus resulting into the desired magnetic activity. Hence, this system of two coils is insensitive to spatially uniform changes in the background field, but it responds to inhomogeneous changes. However, noise gradients shall not be subtracted (Sternickel and Braginski, 2006).
- Signal processing methods. A straightforward solution to reduce the effect of the environmental noise is to apply a hardware or software band-pass filter. When studying the brain response to external stimuli or the background activity, low-pass filters with a cut-off frequency about 40 Hz or 50 Hz can be used (Hämäläinen et al., 1993). However, it is important to realise that the use of non-adaptive filtering techniques may distort the signals or lead to information loss. For instance, the notch filters commonly used to remove the PLN at 50 Hz or 60 Hz distort the signal (Sternickel and Braginski, 2006).

Another alternative is to compute ensemble averages of the MEG signals to reduce the uncorrelated noise and improve the SNR. Nevertheless, this procedure can only be applied when recording responses to stimuli and it is not suitable for background activity (Hämäläinen et al., 1993; Sternickel and Braginski, 2006). Moreover, it removes the subtle individual variations that may appear in the individual responses

(Vigário and Oja, 2008).

In addition to the aforementioned approaches, there are mathematical and signal processing techniques to decompose the recordings into their inner components so that the artefactual activity can be isolated and removed from the rest of the signal. Examples of these methods are the PCA (Sadasivan and Dutt, 1996), the Wavelet Transform (WT) (Sanei and Chambers, 2007) and, specially, the BSS techniques, including the so-called ICA algorithms (Sternickel and Braginski, 2006).

2.4 Artefacts

The most important noise affecting the MEG signals is that produced by outside disturbances. This can be caused by moving magnetic objects (i.e., moving vehicles, people and trains) or electrical equipment, such as computers (Hari et al., 2000; Vrba and Robinson, 2001). Medical equipment also generates strong noise interferences, as shown in Fig. 2.6. For instance, the superconductive magnets used in MRI produce magnetic fields which are 14 or 15 orders of magnitude stronger than the brain activity (Hämäläinen et al., 1993). Finally, bodily organs and movements also introduce *artefacts* in the MEG signals. This includes the eye movements and blinks and the cardiac activity (Hämäläinen et al., 1993).

In any case, prevention is always the best way to deal with artefacts. It often works for external magnetic noise, stimulation-related artefacts, etc. It is also possible to reject all traces coinciding with some biological artefacts, such as OA. However, this procedure can bias the results in some experiments because certain states of the subject are not included to the analysis (Hämäläinen and Hari, 2004). Moreover, it may imply a significant data loss and it does not work for artefacts that are always present in the MEG signals, such as the CA (Castellanos and Makarov, 2006; Fatourechi et al., 2007; Jung et al., 2000).

2.4.1 Cardiac

Normal brain activity generates magnetic fields whose amplitude ranges from 10 fT to 500 fT for evoked responses to a few pT (pico-Teslas) for epileptic spikes. However, the heart produces a significantly stronger field. Its amplitude reaches a few hundred pT over the chest and might, therefore, be of considerable magnitude around the head (Jousmäki and Hari, 1996).

Obviously, the magnetic shielding does not help to reduce the CA since the subject is within the shielded room. The use of gradiometers is partially helpful since, unlike distant environmental noise contributions, the heart is not very far from the sensors (Samonas et al., 1997).

The CA is produced by the electric activity of the heart (Jousmäki and Hari, 1996). The number of significantly contaminated channels tends to be higher over the left than the right hemisphere and depends on the position of the head with respect to the heart (Hämäläinen and Hari, 2004). These artefacts may have a considerable effect on recordings of spontaneous brain activity, and thus should be considered in the analysis (Jousmäki and Hari, 1996). The magnitude of the CA varies from subject to subject and it may be stronger in children due to the shorter heart-to-brain distance (Hämäläinen and Hari, 2004; Jousmäki and Hari, 1996).

It has been shown that the P wave, which is related to the atrial depolarization, is more prominent in the Electrocardiogram (ECG) than in MEG. This might be due to the fact that the current distribution is more complex during atrial than during ventricular depolarization. Hence, the magnetic field of such a complex pattern may attenuate more rapidly with distance than that of a dipolar current (Jousmäki and Hari, 1996).

Other artefacts related to the cardiac cycle are the blood-flow-related susceptibility artefacts and the ballistocardiogram artefacts (Hämäläinen and Hari, 2004; Jousmäki and Hari, 1996). It has been shown that the former are negligible in healthy subjects (Jousmäki and Hari, 1996). The latter are caused by body movements related to heart beats and their main peak lags the QRS complex by several hundred milliseconds. They may exist in MEG recordings if the subject's lung, skin, or clothes contain magnetic contami-

nants (Jousmäki and Hari, 1996). Hence, all magnetic materials on the subject must be removed before entering into the recording room (Hämäläinen et al., 1993).

2.4.2 Ocular

Furthermore, significant contamination can be caused by eye blinks and movements (Antervo et al., 1985). Both types of OAs may be time-locked to the stimuli, specially if they are strong or infrequent (Hämäläinen and Hari, 2004). The origin of the OAs can be explained considering a simplified model where each eye is represented by a small electrical dipole oriented from the negatively charged retina to the positively charged cornea (Croft and Barry, 2000). Obviously, any eye movement modifies the orientation of such dipole, thus altering the associated electromagnetic fields near the eyes. As for the artefacts caused by blinks, these are due to a change in the intensity of the dipole created by the movement of the eyelid over the eyeball (Croft and Barry, 2000; Hämäläinen and Hari, 2004).

It must be noticed that the amplitude of the artefacts caused by eye movements can be about 3 pT to 4 pT above the lateral aspects of the orbits and that the polarity of the magnetic fields is opposite over both hemispheres (Antervo et al., 1985). The amplitude of these artefacts increases with light and decreases with darkness in a similar way for eye blinks and movements (Antervo et al., 1985).

2.4.3 Power Line Noise

Apart from the biological artefacts, the PLN may also be a strong source of contamination in brain recordings (Escudero et al., 2007b; Jung et al., 2000). Moreover, removal of the PLN and its harmonics with notch filters may not be a satisfactory solution if the frequency band of the disturbances overlaps that of the signals (Hämäläinen et al., 1993; Iriarte et al., 2003; Sternickel and Braginski, 2006).

2.4.4 Other Artefacts

Muscle contractions in the neck and face areas can also produce artefacts in MEG signals, although they seem to be weaker than in EEG measurements. This may be due to the fact that the distance to the muscular sources is significantly larger for MEG sensors than for EEG electrodes (Hämäläinen and Hari, 2004). As it was previously mentioned, magnetic lung-contamination can cause respiration-related slow shifts (Hari, 2004).

2.5 Comparison With Other Neuroimaging Methods

It must be taken into account that the MEG is not the only technique to study the brain activity. Other methods are available, such as the PET, SPECT, fMRI, MRS and EEG. However, EEG and MEG avoid not only the need for opening the skull to examine the subject's brain but also for being exposed to X rays, radioactive tracers or strong time-varying magnetic fields (Hämäläinen et al., 1993).

These techniques have dissimilar characteristics. However, it is important to regard them as complementary techniques rather than competitors. For instance, changes in synchrony of the spontaneous firing or in time- and phase-locking of the signals with respect to external stimuli are seen with MEG and EEG but may be invisible for fMRI or PET. On the other hand, MEG and EEG can be blind to the slowly rising or sustained neuronal activity that is related to with metabolic changes (Hari et al., 2000). Moreover, all of them share the important feature of not needing any kind of surgery (Hämäläinen et al., 1993).

The structural information provided by MRI can be merged with that offered by MEG to perform a non-invasive cartographic study of the brain function (Hari, 2004). Additionally, MEG information can also be combined with other neuroimaging techniques, such as CAT. The advantage of combining these techniques is that, whereas MEG provides high temporal resolution, MRI and CAT are able to offer accurate images of the brain anatomy with

millimetre resolution (Hämäläinen et al., 1993). This combination of MEG and neuroimaging techniques can be helpful to estimate the current sources by imposing constraints derived from the brain anatomy (Hari, 2004). Finally, the study of the brain can be complemented with SPECT or PET, which measure the functional activity of the brain (Hämäläinen et al., 1993; Hari and Forss, 1999).

This Section briefly describes the main characteristics, advantages and limitations of the most commonly used techniques in the functional study of the brain. These techniques can be classified as follows (Maestú et al., 1999):

- Techniques that measure the radiation emitted by particles and radioisotopes: SPECT and PET.

- Techniques that acquire signals produced by electromagnetic sources.
 - Methods that apply external electromagnetic sources to produce a image of the brain: fMRI and MRS.
 - Methods that record the electromagnetic activity generated by the brain: EEG and MEG.

Fig. 2.9 shows a classification of these techniques in terms of their temporal and spatial resolution as well as their degree of invasiveness (Maestú et al., 1999).

2.5.1 SPECT and PET

Both SPECT and PET are nuclear medicine tomographic imaging techniques that provide diagnostic information about the brain by analysing the tridimensional distribution of the electromagnetic radiation produced by a radioisotope (Knoll, 1983). They differ from CAT scanning in the sense that the source of the electromagnetic radiation, instead of being an external X-ray tube, is a radioisotope distributed within the body (Knoll, 1983). Hence, while CAT displays absorption properties related to anatomy, SPECT and PET reveal the distribution of radioactive tracers that can often indicate various aspects of physiological function (Knoll, 1983). This information is

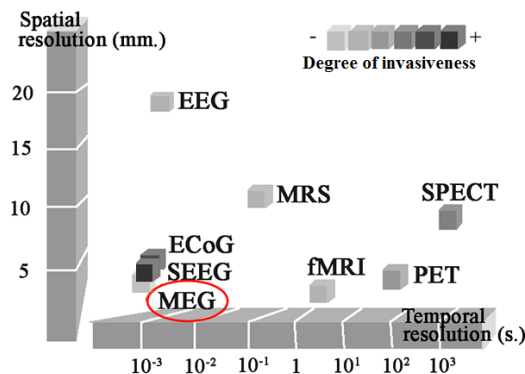


Figure 2.9: Diagram comparing diverse techniques for functional brain analysis in terms of their temporal and spatial resolution and their degree of invasiveness.

provided as a set of thin two-dimensional sections (or slices) along a particular axis (Knoll, 1983).

In CAT, the source of X-rays is a well-focussed spot on the anode of an X-ray tube. With this information, the detection point uniquely defines a single ray of interest. In SPECT and PET, the situation is more complicated as the position of the source point is not known a priori. Actually, the knowledge of this position is the general goal of the measurement. SPECT and PET differ in the methods used to define the ray direction (Knoll, 1983).

PET is limited to those radioisotopes that decay by emitting a positron. This positron travels a short distance (a few millimetres at most) before annihilating with a normal electron from the subject's tissue. The resulting annihilation radiation consists of two electromagnetic photons emitted in time coincidence and are oppositely directed along exactly the same line. By placing two detectors on opposite sides of the subject, it is possible to define the ray (Knoll, 1983). This technique, because of the short half lives of the isotopes involved, requires that a expensive cyclotron production facility be located within the hospital (Knoll, 1983). PET is able to measure the tissue metabolic activity in terms of regional glucose uptake. It is particularly useful to study the functional activity of the central nervous system when a high temporal resolution is not needed (Maestú et al., 1999).

SPECT differs from PET in that any radioisotope that emits γ rays may

be used. These γ rays are emitted as single individual photons. Thus, SPECT allows a much broader range of applications. The agents usually have relatively long half lives and are available at low cost (Knoll, 1983). The maps offered by SPECT represent the regional perfusion, the density of neuro-receptors or the metabolic activity of a brain region depending on which radiopharmaceutical was taken (Knoll, 1983). Hence, SPECT is a helpful technique to measure the brain blood flow with a relatively low cost. Moreover, it can be used to carry out semi-functional studies, where diverse activated brain areas can be recorded for a limited period of time (Maestú et al., 1999). However, its spatial resolution is lower than that of PET (Maestú et al., 1999). Similarly to the PET, its use is limited by the fact that it does involve exposure to ionizing radiation (Maestú et al., 1999).

2.5.2 fMRI and MRS

fMRI is a procedure that employs MRI images to measure small metabolic changes due to the activation of small cerebral regions. It has very high spatial resolution, but it cannot follow the changes in activity in real time. Moreover, it is necessary to wait for a short period of time before acquiring the measurement again. It is extremely useful to perform functional studies of the brain, largely innocuous, relatively fast and comfortable for the subject. Nevertheless, similarly to the previously mentioned techniques, it does not measure the neural activity directly, but the fMRI must interpret it in terms of vascular and metabolic changes. Furthermore, some types of mental activities, brain disorders and malfunctions of the brain cannot be recorded using fMRI since their effect on the level of oxygenated blood is low (Sanei and Chambers, 2007).

On the other hand, MRS is a technique based on similar principles to those of the fMRI. The main difference is the way in which the data are shown. Instead of the anatomic image, a spectrum related to several biochemical compounds is shown. MRS can analyse volumes of up to 10 cm^3 and its temporal resolution is of about tens of seconds. It is used to obtain biochemical information about the tissues of the human body in a non-invasive

way (without the need for a biopsy; [Maestú et al., 1999](#)).

Both fMRI and MRS have the important advantage over the SPECT and PET of not requiring ionising radiation to perform the recordings. Additionally, there is no accumulative effect due to the repetitive performance of measurements. Hence, fMRI and MRS studies can be repeated as many times as necessary. However, in fMRI and MRS the subject is exposed to very strong magnetic fields (between 20 and 80000 times the Earth magnetic field, depending of the equipment) whose effects on the human body are not completely clear. Thus, its use is not recommended in certain cases, such as pregnancy ([Maestú et al., 1999](#)).

2.5.3 EEG

EEG is a physiological techniques that records the electrical brain activity originated by the neurons with a set of electrodes located on the scalp ([Sanei and Chambers, 2007](#)). It must be noticed that the EEG varies notably depending on the location of the electrodes. Hence, it is necessary to standardise the electrode placement. The most frequent setting is the International 10–20 System, recommended by the International Federation of Societies for Electroencephalography and Clinical Neurophysiology ([Sanei and Chambers, 2007](#)).

EEG recordings are able to reflect the changes in the electrical brain activity due to several physiological and pathological causes. Hence, they are commonly employed in clinical practice, including monitoring of comma and brain death, studying brain lesions and tumours and assessing anaesthesia depth ([Sanei and Chambers, 2007](#)). Moreover, their utility in diverse brain disorders, including AD has been studied ([Jeong, 2004](#)).

It is important to note that EEG and MEG record the electromagnetic activity generated by the same primary currents in the brain ([Hari, 2004](#)). Hence, similarities between their waveforms are to be expected. The distributions of the electric potential and the magnetic field generated by a current dipole in a spherical volume conductor are dipolar (i.e., they display two extrema of opposite polarities) but they are rotated by 90° with respect

to each other (Hari, 2004). Moreover, both signals provide complementary information (Rampp and Stefan, 2007) and this notion of complementarity can be mathematically defined (Dassios et al., 2007). In a spherical head model, the part of the neuronal current that generates the electric potential is complementarily orthogonal to the part of the current creating the magnetic induction field. Thus, for a continuously distributed neuronal current, information missing in EEG is available in the MEG data, and vice versa (Dassios et al., 2007).

There are differences, though, in the sensitivity of both techniques to the orientation of the currents. MEG is mostly sensitive to tangential currents (Hämäläinen et al., 1993). Actually, this sensitivity to tangential currents in the presence of several simultaneous sources is an important advantage in practical work, whereas some current sources (very deep and radial) are more reliably picked up by EEG than MEG (Hari, 2004).

When comparing these techniques, it should be kept in mind that both signals can record the brain oscillations simultaneously all over the scalp. Moreover, their temporal resolution is high. It is possible to follow the dynamics of the brain oscillations on a millisecond time-scale (Hämäläinen et al., 1993). For MEG, this high temporal resolution can be accompanied of good spatial resolution in locating cortical events if the dipoles are just a few millimetres away of each other and their temporal waveforms differ (Hari, 2004; Hari and Forss, 1999). Although both techniques are complementary to study the brain activity, the MEG recording offers some important advantages over the EEG:

1. The magnetic recordings do not depend on any reference point due to the fact that MEG is not a differential measure between two different channels. On the other hand, reference-free EEG signals can be obtained by computing the surface Laplacians. However, even in this case, it is difficult to make a reasonable guess of the source locations and the Laplacians cannot be computed for the outermost electrodes (Hämäläinen et al., 1993; Hari et al., 2000; Najarian and Splinter, 2006; Rampp and Stefan, 2007).

2. The resistive properties of the skull and other extra-cerebral tissues do not affect the magnetic fields in contrast to the significant distortion and smearing that they produce in the electric potentials ([Hämäläinen et al., 1993](#); [Hari et al., 2000](#)).
3. Measurements from a very large numbers of sensors are more easily done with MEG than with EEG, thanks to the fact that the time-consuming application of electrodes on the scalp is not needed ([Rampp and Stefan, 2007](#); [Stam, 2010](#)).

On the other hand, the MEG recording is subjected to some considerable limitations ([Hari, 2004](#); [Hari et al., 2000](#); [Rampp and Stefan, 2007](#)):

1. The only technology able to produce good-quality measurements of the extremely weak brain magnetic fields is based on superconductive materials.
2. The recording process must take place in a magnetically shielded room to minimise ambient magnetic field variations.

Hence, the investment cost in the MEG system and shielded room is high, which has prevented any widespread use of this technique to acquire the brain activity ([Sternickel and Braginski, 2006](#)).

Chapter 3

Alzheimer's Disease

Contents

3.1	Introduction	56
3.2	Social Importance of AD	57
3.3	Main Characteristics of AD	58
3.3.1	Evolution of the Disease	58
3.3.2	Pathophysiology	61
3.3.3	Risk Factors	63
3.3.4	Clinical Diagnosis	66
3.3.5	Possible Biomarkers	68
3.4	Current Treatment	70
3.4.1	Symptomatic Treatments	71
3.4.2	Modifying-Disease Drug Candidates	73
3.4.3	Possible Drugs Based on Epidemiological Studies	74
3.4.4	Orientation to caregivers	76

This Chapter of the PhD dissertation presents the main concepts related to AD. It begins with a brief introduction about this dementia and pointing out its social importance in the Western World. Then, the main characteristics of AD are introduced. This includes information about its evolution, pathophysiology, risk factors and diagnostic techniques. Finally, current and possible future treatments for AD are discussed.

3.1 Introduction

AD is a slowly progressive and fatal neurodegenerative disorder with insidious onset. The first case, a 51 year-old woman named Auguste D. with a 5-year history of progressive cognitive impairment, was first described in 1906 by Dr. Alois Alzheimer during a congress in Tübingen, Germany (Blennow et al., 2006; Ferri et al., 2006; Samanta et al., 2006). In this first case, Dr. Alzheimer introduced the typical clinical characteristics of the dementia that now bears his name, including the memory disturbances and instrumental signs. He also identified the plaques and tangles that are known today as the hallmarks of the disease (Blennow et al., 2006).

In addition to the very well-known memory loss, AD also results in other cognitive and behavioural symptoms that progressively impair the activities of daily living. The cognitive deficit manifest itself as disorientation, confusion and problems with reasoning and thinking, whereas the behavioural symptoms include agitation, anxiety, delusions, depression, hallucinations, insomnia and wandering (Blennow et al., 2006; Cummings, 2004; Lahiri et al., 2002).

Ageing is the most obvious risk factor for AD. Additionally, several epidemiological studies have pinpointed possible associations of AD with a decreased reserve brain capacity, which is determined by the number of neurons and their synaptic and dendritic density (Blennow et al., 2006). Other risk factors are related to vascular disease, including, among others, hypertension, atherosclerosis, smoking and obesity (Blennow et al., 2006; Lahiri et al., 2002) although these associations have recently been questioned by Purnell et al. (2009). Dietary habits, such as the intake of vitamin B12 or antioxidants,

have also been linked with a reduced risk of AD, but they are not conclusive enough to make any dietary recommendations (Blennow et al., 2006).

Moreover, AD has a significant genetic background (Blennow et al., 2006) in its both forms: familiar and sporadic (Blennow et al., 2006). The familiar form has been related to mutations in the Amyloid Precursor Protein (APP), Presenilin 1 (PSEN1) and Presenilin 2 (PSEN2) genes (Blennow et al., 2006; Reitz and Mayeux, 2009; van Es and van den Berg, 2009). On the other hand, the much more common sporadic form is associated with the Apolipoprotein E (APOE) $\epsilon 4$ allele (Blennow et al., 2006; Reitz and Mayeux, 2009). Additionally, it has been shown that the SORL1-Related receptor (SORL1), Clusterin (CLU), Phosphatidylinositol-binding Clathrin Assembly Protein (PICALM) and Complement Component (3b/4b) Receptor 1 (CR1) genes are also associated with AD (Harold et al., 2009; Lambert et al., 2009; National Institute on Aging, 2008).

3.2 Social Importance of AD

AD is the most common type of dementia in the Western World. It accounts for 50% to 60% of all cases (Blennow et al., 2006). It is, indeed, one of the most disabling and burdensome disorders worldwide. It is 1.5 times more common than stroke or epilepsy and as common as congestive heart failure (Samanta et al., 2006). Its prevalence is below 1% in subjects aged 60 to 64 years, but it shows an almost exponential increase with age. Approximately, the prevalence doubles with every five years of age (Cummings, 2004). As a result, the prevalence of AD in people aged 85 years or older is between 24% and 33% (Blennow et al., 2006). In 2001, more than 24 million people had dementia. Due to the anticipated increase in life expectancy, it is expected that this figure will double every 20 years up to 81 million in 2040 (Blennow et al., 2006). Moreover, it is estimated that about 4.6 million new cases of dementia appear each year (Ferri et al., 2006). Hence, it is clear that AD is a very important economic and social problem.

It has been recently shown that dementia contributed more than 10% of years lived with disability in people over 60 (Ferri et al., 2006). This ac-

counts for more cases than stroke (9.5%), musculoskeletal disorders (8.9%), cardiovascular disease (5.0%) and all forms of cancer (2.4%) (Ferri et al., 2006). Furthermore, the disability weight for dementia was higher than for any other health condition except for spinal-cord injury and terminal cancer (Ferri et al., 2006).

People with dementia cause high costs as a result of their use of the health system and the community and residential care (Ferri et al., 2006). For instance, it has been estimated that, in 2005, nearly half of the elderly people with cognitive impairment in the UK live in institutions at a cost of £4.6 billion per year. This represented the 0.6% of the UK gross domestic product (Ferri et al., 2006). Furthermore, the care and attention needed by the patients is exceedingly intensive. The national direct and indirect costs of caring for people with AD are estimated to be more than \$100 billion per year in USA (National Institute on Aging, 2008).

It is expected that diagnostic testing for this dementia will contribute only marginally to this cost. Hence, the development of methods able to help in the the diagnosis may likely improve the care of patients and allow more accurately targeted future therapies. Eventually, this may lead to reduce the costs associated with this disorder (Lahiri et al., 2002).

3.3 Main Characteristics of AD

3.3.1 Evolution of the Disease

AD is a slowly progressive and fatal neurodegenerative brain disorder. It is characterised by memory disturbances and other cognitive and behavioural symptoms that progressively impair the activities of daily living (Blennow et al., 2006). AD patients usually die within 7 to 10 years after diagnosis (Masters et al., 2006). However, it is estimated that the neurodegeneration begins between 20 and 30 years before clinical onset (Blennow et al., 2006). AD starts near the hippocampus and eventually affects it. As it was indicated in Section 2.2.2, the hippocampus plays a major role in learning and converting short-term memories to long-term memories (National Institute

on Aging, 2008). During this preclinical phase, the number of microscopical lesions increases and at a certain threshold the first symptoms appear and gradually worsen (Blennow et al., 2006; Nestor et al., 2004). Once the cognitive deficits become global and severe enough to interfere with normal living, established criteria for a clinical diagnosis of AD are met (Nestor et al., 2004).

Subjects in the pre-dementia stage of AD are considered to be cognitively impaired but do not meet criteria for dementia. This stage is commonly known as Mild Cognitive Impairment (MCI). This is due to the fact that the subjects' cognitive deficits are only limited to memory and/or their everyday life activities are preserved (Nestor et al., 2004). MCI usually precedes the formal diagnosis of AD (Masters et al., 2006) and it is defined on the basis of reports of memory loss verified by objective tests (Blennow et al., 2006).

The concept of MCI represents a transitional period before the development of full-blown AD (National Institute on Aging, 2008). About 40% to 60% of carefully characterised subjects with MCI will progress to meet criteria for AD within 3 or 4 years (Masters et al., 2006). This rate of conversion to AD is about ten times higher than that for the general population (Nestor et al., 2004). The current view on evolution from healthy ageing to MCI and, eventually, AD is shown in Fig. 3.1. However, the narrow definition of MCI as an amnesic syndrome may fail to capture the heterogeneity of clinical AD presentations (Nestor et al., 2004).

Once the patient has reached the state of early dementia, the most obvious symptom is memory loss, such as difficulty in remembering recently learned facts. However, episodic, semantic and implicit memory are less affected than new facts or memories. Additionally, the language deteriorates due to a decreased vocabulary and word fluency, but the AD patient can still communicate basic ideas. A mild apraxia may also appear at this stage (Gwyther, 2001; National Institute on Aging, 2008). Eventually, AD patients may need assistance or supervision with the most cognitively demanding activities (Gwyther, 2001; National Institute on Aging, 2008).

At some point, the progressive deterioration of mental functions hinders independence. Then, the patients are unable to perform most common activities of daily living (Gwyther, 2001; National Institute on Aging, 2008).

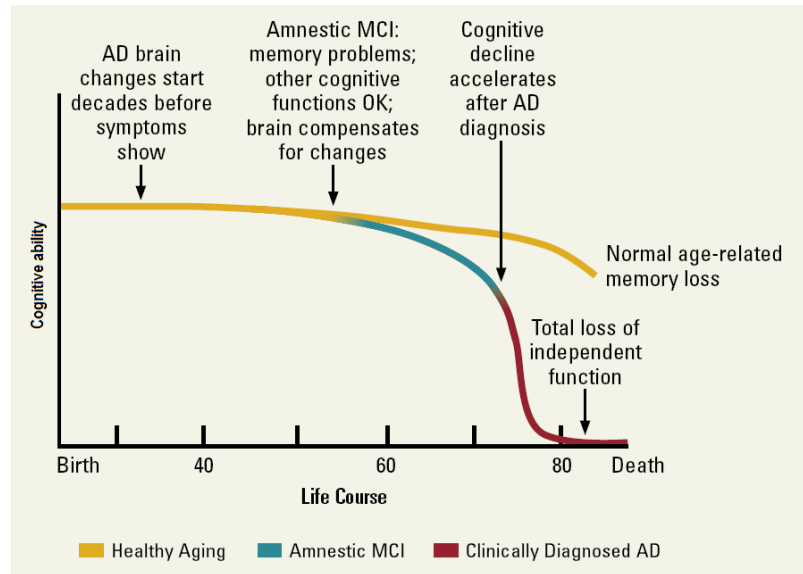


Figure 3.1: View on the evolution from healthy ageing to AD by [National Institute on Aging \(2008\)](#).

Speech difficulties become more evident and complex motor sequences become less coordinated. At this phase of moderate dementia, memory problems worsen (e.g., the person cannot recognise close relatives) and long-term memory begins to be affected by the disease ([Gwyther, 2001](#); [National Institute on Aging, 2008](#)). Moreover, behavioural and neuropsychiatric changes become more prevalent, including wandering, irritability or crying. In some cases, delusional symptoms may also appear ([Gwyther, 2001](#)).

Finally, as the disease advances, the symptoms include confusion, irritability and a more generalised loss of language and long-term memory ([Cummings, 2004](#); [Gwyther, 2001](#); [National Institute on Aging, 2008](#)). However, patients can often understand and return emotional signals. In the last stages, the previous aggressiveness tends to be replaced by apathy. The patients are utterly dependent on their caregivers and cannot perform almost any task without assistance ([Jeong, 2004](#)). Eventually, an external factor produces the death of the AD patient ([Gwyther, 2001](#)).

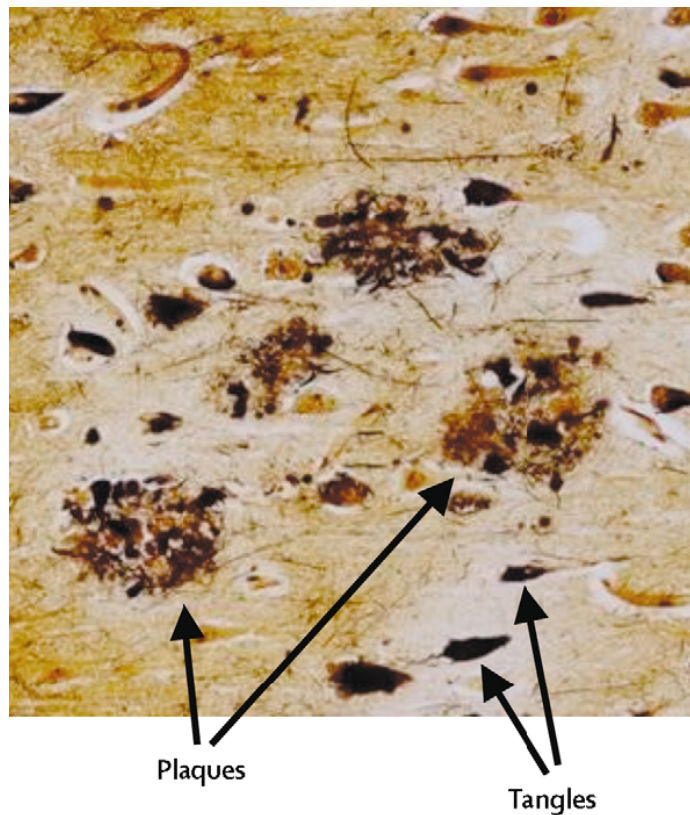


Figure 3.2: Plaques and tangles in the cortex as shown by [Blennow et al. \(2006\)](#).

3.3.2 Pathophysiology

At the microscopic level, the hallmark lesions in AD are senile plaques (also known as neuritic plaques) and neurofibrillary tangles, which are shown in a histological sample in Fig. 3.2 ([Blennow et al., 2006](#); [Mattson, 2004](#)). However, the mechanism whereby the formation of these lesions leads to neurodegeneration is poorly understood ([van Es and van den Berg, 2009](#)).

In addition to the appearance of plaques and tangles, AD is also characterised by degeneration of neurons and synapses ([Blennow et al., 2006](#)). This process produces the typical cognitive deficit in AD. There is increasing consensus that the production and accumulation of $A\beta$ is central to the pathogenesis of this dementia ([Cummings, 2004](#); [Masters et al., 2006](#)). The generation and deposition of $A\beta$ has been related to the formation of

neurofibrillary tangles, oxidation and lipid peroxidation, glutamatergic excitotoxicity, inflammation, and activation of the cascade of apoptotic cell death (Cummings, 2004).

The senile plaques are extracellular deposits of $A\beta$ surrounded by dystrophic neurites, reactive astrocytes, and activated microglia (Blennow et al., 2006; Masters et al., 2006). $A\beta$ is derived from the APP gene on chromosome 21 (Masters et al., 2006). Additionally, there is an association of the APOE $\epsilon 4$ allele with AD (Blennow et al., 2006) as it is a risk factor for the disease. APOE acts as a cholesterol transporter in the brain with APOE $\epsilon 4$ being less efficient than the other variants ($\epsilon 2$ and $\epsilon 3$) in neural repair and reuse of membrane lipids (Blennow et al., 2006). APOE $\epsilon 4$ enhances $A\beta$ aggregation and reduces $A\beta$ clearance (Mattson, 2004). It is also essential for $A\beta$ deposition, promoting $A\beta$ fibrillation and plaque formation. The APOE $\epsilon 4$ allele has been estimated to account for most of the genetic risk in sporadic AD (Blennow et al., 2006; Masters et al., 2006).

The other important histo-pathological hallmark of AD is the neurofibrillary tangles. They are intracellular bundles of a hyper-phosphorylated form of the microtubule-associated protein tau (Blennow et al., 2006). Tau is a normal axonal protein found in the neurons. It binds to microtubules through its microtubule-binding domains. This promotes microtubule assembly and stability, which helps in the axonal transport (Blennow et al., 2006). In AD, tau is hyper-phosphorylated inside the neurons. This leads to sequestration of normal tau and other microtubule-associated proteins, causing disassembly of microtubules. As a consequence, the axonal transport is impaired and this compromises the neuronal and synaptic function (Blennow et al., 2006). Tau also becomes prone to aggregation into insoluble fibrils in tangles, further compromising neuronal function. Tau pathology starts early in the AD process in neurons in the transentorhinal region, spreads to the hippocampus and amygdala, and later to the neo-cortical association areas (Blennow et al., 2006).

The finding of a correlation between plaque counts and dementia severity suggests that the plaques are involved in the pathogenesis of the disease (Blennow et al., 2006). However, whether tau hyper-phosphorylation and

tangle formation are a cause or consequence of AD is unknown (Blennow et al., 2006). Additionally, brain regions involved in learning and memory processes, including the temporal and frontal lobes, are reduced in size in AD patients as the result of degeneration of synapses and neural death (Mattson, 2004).

To sum up, the central hypothesis for the cause of AD is the amyloid cascade hypothesis (Blennow et al., 2006; Masters et al., 2006). This indicates that an imbalance between the production and clearance of $A\beta$ (a product of APP processing) in the brain is the initiating event, eventually leading to neuronal degeneration and dementia (Blennow et al., 2006; Masters et al., 2006; van Es and van den Berg, 2009).

This hypothesis is supported by the facts that the genes relevant to the familial form of AD are related to $A\beta$ generation (Blennow et al., 2006). Soluble $A\beta$ is thought to suffer a conformational change to high β -sheet content, thus becoming prone to aggregate into soluble oligomers and larger insoluble fibrils in plaques (Blennow et al., 2006). In this process, the fibrillogenic $A\beta_{42}$ isoform provokes the misfolding of other $A\beta$ species.

Furthermore, tau hyper-phosphorylation in AD starts intracellularly. It causes sequestration of normal tau and other microtubule-associated proteins. This produces disassembly of microtubules and thus impaired axonal transport, compromising neuronal and synaptic function. Additionally, insoluble tangles composed of abnormally hyper-phosphorylated tau protein are developed, further compromising neuronal function (Blennow et al., 2006; Masters et al., 2006). The mechanism of neural death is graphically summarised in Fig. 3.3 (Samanta et al., 2006).

3.3.3 Risk Factors

In addition to ageing, which is the most obvious risk factor for AD, epidemiological studies have pinpointed several possible associations of diverse factors with AD. Some of them are reduced brain size, low educational and occupational attainment, low mental ability during early life, and reduced mental and physical activity in late life (Blennow et al., 2006; National In-

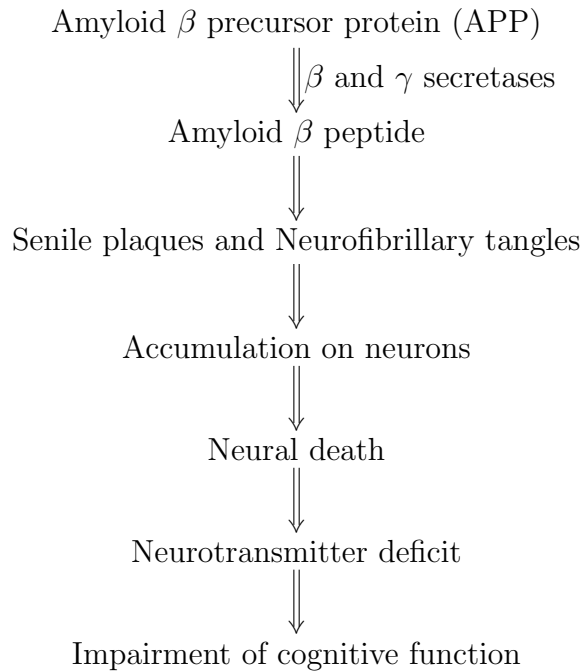


Figure 3.3: Mechanism of neural death in AD.

stitute on Aging, 2008; Reitz and Mayeux, 2009). All these factors have to do with a decreased reserve brain capacity, which is determined by the number of neurons and their synaptic and dendritic density. A low reserve capacity has been associated with early onset of the disease (Blennow et al., 2006; Reitz and Mayeux, 2009). Additionally, head injury can be another risk factor (Lahiri et al., 2002). However, it is unclear whether this may lead to creation of plaques and tangles or whether it just reduces the brain reserve capacity (Blennow et al., 2006).

Cardiovascular factors have also been related to AD. They include, among others, hypertension, atherosclerosis, atrial fibrillation, smoking and obesity (Blennow et al., 2006; Lahiri et al., 2002). Nevertheless, it has not been established whether they lead to plaque and tangle formation or whether they produce cerebrovascular pathology which adds to a pre-existing dementia (Blennow et al., 2006). What is more, a recent study suggests that there is limited evidence supporting the fact that single cardiovascular risk factors

affect AD risk and that the strength of the measured association is affected by the interactions between single factors (Purnell et al., 2009).

Finally, other dietary habits, such as the intake of vitamin B12, antioxidants or unsaturated fatty acids, have been linked with a reduced risk of AD (National Institute on Aging, 2008), but they are not conclusive enough to make any dietary recommendations (Blennow et al., 2006).

Nevertheless, AD has a significant genetic background (Blennow et al., 2006). This dementia is a heterogeneous disorder with both familial and sporadic forms (Blennow et al., 2006). The onset of the familiar type occurs before age 65 years. The first mutation producing this form of AD was identified in the APP gene on chromosome 21 and several additional mutations of this gene have been found. However, most of the familiar cases are explained by mutations in the PSEN1 and PSEN2 genes (Blennow et al., 2006; Reitz and Mayeux, 2009; van Es and van den Berg, 2009). These genes are thought to increase $A\beta$ production (van Es and van den Berg, 2009). In any case, it must be noted that the prevalence of the familiar form of the disease with autosomal dominant inheritance is very rare, about 1% (Blennow et al., 2006; van Es and van den Berg, 2009).

On the other hand, it is well-known that there exists an association between the APOE $\epsilon 4$ allele and the sporadic form of AD (Blennow et al., 2006; Reitz and Mayeux, 2009). This link was discovered in 1993, and subsequent studies have shown that APOE may account for up to half the genetic risk for AD (van Es and van den Berg, 2009). Additionally, the gene SORL1 is another risk-factor for AD. SORL1 has to do with the recycling of APP and underexpression of SORL1 increases $A\beta$ levels (National Institute on Aging, 2008; Reitz and Mayeux, 2009). Recently, other additional risk factors for AD have been discovered (Carrasquillo et al., 2009; Harold et al., 2009; Lambert et al., 2009). For instance, a gender-linked susceptibility gene for late-onset AD has been found in the PCDH11X gene of the X chromosome (Carrasquillo et al., 2009). Women with two copies of this gene are at considerably greater risk of developing AD (Carrasquillo et al., 2009). Other three additional risk factors for AD have been discovered even more recently (Harold et al., 2009; Lambert et al., 2009). Of note is that these genes — CLU, PICALM and

CR1 — may be involved in pathways regulating $A\beta$ clearance from the brain for the following reasons (Harold et al., 2009; Lambert et al., 2009; van Es and van den Berg, 2009):

- CLU has some similarities in common with APOE. They are related to major brain apolipoproteins and amyloid plaques. Moreover, they interact with $A\beta$ in its conversion into insoluble forms. Finally, CLU and APOE seem to cooperate in reducing $A\beta$ deposition and in modifying $A\beta$ clearance (Harold et al., 2009; Lambert et al., 2009).
- There is evidence that CR1 has a protective role for the brain via the generation and binding of C3b, which may contribute to $A\beta$ clearance (Lambert et al., 2009).
- PICALM has to do with the intracellular trafficking of proteins, lipids, growth factors and neurotransmitters. It has been shown that PICALM is involved in retrieving APP from the cell surface and that an aberrant processing leads to altered $A\beta$ levels (Harold et al., 2009).

3.3.4 Clinical Diagnosis

A definite diagnosis of AD can only be made by necropsy (the microscopical analysis of the patient's brain tissue after death; Blennow et al., 2006). In clinical practice, AD must be differentiated from other dementias though. In fact, the criteria for detection of probable AD largely depend on the exclusion of other dementias. The diagnosis of AD is usually based on the criteria developed by the NINCDS-ADRDA according to which it is classified as (McKhann et al., 1984):

- Definite AD:
 - Clinical criteria of probable AD.
 - Histologic confirmation via autopsy or biopsy.
- Probable AD: typical clinical syndrome without histologic confirmation. This requires the following:

- Diagnosis of dementia established with a clinical examination and supported by a mental state test.
 - Diagnosis of dementia confirmed with a neuropsychological test.
 - Cognitive impairment in two or more areas of cognition.
 - Progressive impairment of memory and other cognitive functions.
 - Absence of alterations in the conscience.
 - Onset of the deficits between the ages of 40 and 90 years.
 - Absence of other diseases capable of causing dementia.
- Possible AD: atypical clinical features but no alternative diagnosis apparent and no histologic confirmation. This refers to:
 - Presence of a disease or mental disorder capable of producing dementia without being able to confirm it as the cause of the cognitive impairment.
 - Progressive impairment of only one cognitive function without any other known cause.

The accuracy of clinical diagnosis is limited. Even in patients who have been followed up clinically for several years at expert research centres, the diagnostic accuracy is relatively low, with sensitivity of around 80% and specificity of about 70% (Blennow et al., 2006). Some authors even suggest lower typical sensitivity and specificity values for the diagnosis of probable AD (Cummings, 2004) and that these sensitivity and specificity figures may be significantly lower in primary care settings and in patients with mild AD (Blennow et al., 2006; Knopman et al., 2001).

The medical history together with clinical, neurological, and psychiatric examinations serves as the basis in the diagnostic work-up. In very early cases, neuropsychological testing can help to obtain objective signs of memory disturbances. Laboratory studies, such as thyroid-function and serum vitamin B12 tests, are necessary to identify secondary causes of dementia and coexisting disorders that are common in elderly people (Blennow et al., 2006; Cummings, 2004). Neuroimaging, including CAT and MRI, is useful

to exclude alternative causes of dementia, such as brain tumour and subdural haematoma (Cummings, 2004; Knopman et al., 2001). Cerebral atrophy, visualised as enlarged ventricles and cortical sulci, is also identified by those neuroimaging techniques, but the overlap with normal ageing and other dementias is too large to confirm their diagnostic value (Blennow et al., 2006). However, neuroimaging is valuable to detect cerebrovascular disease, such as cerebral infarcts and white-matter lesions, which is important to identify Vascular Dementia (VaD) or mixed dementia (AD/VaD; Blennow et al., 2006).

Additionally, tests are used to screen for cognitive impairment. The most commonly used test is the MMSE (Folstein et al., 1975). It includes 30 short questions (35 in its Spanish version; Lobo et al., 1979) that sample some cognitive functions including arithmetic, memory and orientation (Folstein et al., 1975). Moreover, new biomarkers would be of great value as diagnostic tools, both for the clinical diagnosis of AD and the prediction of incipient AD in MCI cases (Blennow et al., 2006). It must be borne in mind that the low rates of recognition of dementia by family members and physicians constitute a major barrier to appropriate care for many patients with AD (Cummings, 2004).

3.3.5 Possible Biomarkers

As it was previously mentioned, new biomarkers would be of great value as diagnostic tools for the clinical diagnosis of AD (Blennow et al., 2006). An ideal biomarker for this dementia should detect a fundamental characteristic of the neuropathology and be validated in neuropathologically confirmed cases, with sensitivity and specificity no lower than 80% (Blennow et al., 2006).

The utility of some methods in this setting has been studied. For instance, the levels of $A\beta$ and tau proteins in the CerebroSpinal Fluid (CSF) have been widely researched (Blennow et al., 2006). The brain imaging techniques PET and MRI can also complement other biomarkers. This is due to the fact that PET images may be useful to reveal the progressive reduction

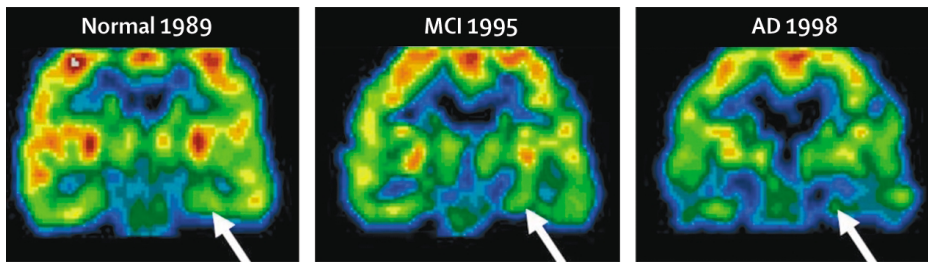


Figure 3.4: FDG-PET scans in an elderly woman at baseline (1989) and over 9 years, while the patient declined to MCI and AD. Progressive reductions in glucose metabolism are illustrated by [Blennow et al. \(2006\)](#).

in glucose metabolism associated with AD, as Fig. 3.4 shows. Additionally, MRI images similar to that plotted in Fig. 3.5 can illustrate that the hippocampus and entorhinal cortex typically get smaller, while the ventricle increases in size with disease progression. However, few studies have addressed the differentiation of AD from other dementias, and there are few autopsy confirmation data available ([Blennow et al., 2006](#)).

In spite of the possible useful information that the aforementioned techniques may provide, their use in clinical practice is not recommended yet. This is due to the fact that more exhaustive clinical trials are needed to confirm their utility and performance in AD diagnosis ([Blennow et al., 2006](#)).

The analysis of MEG recordings is not included in the clinical diagnosis of AD. This is due to the facts that the availability of such systems is still reduced and to their limited diagnostic value. However, in recent years, the availability of whole-head MEG equipment has increased ([Stam, 2010](#)). This technique can record brain activity directly, and has several advantages over the conventional EEG (see Section 2.5.3; [Hari, 2004](#)). Therefore, MEG may offer a more accurate image of ongoing neural activity. Furthermore, significant advances have been made in neuroscience regarding the understanding of oscillatory and synchronized brain activities and abnormal patterns associated with various brain disorders, including AD ([Stam, 2010](#)). Hence, it is possible that these changes, which reflect abnormalities in specific networks and neurotransmitter systems, and could be useful in future AD diagnosis and monitoring ([Stam, 2010](#)).

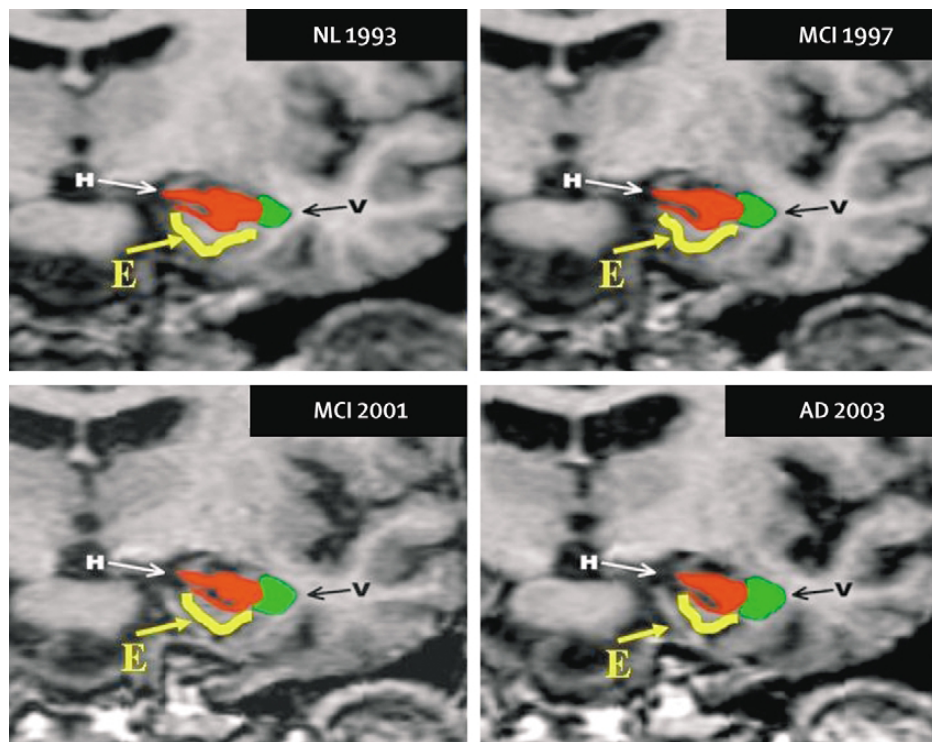


Figure 3.5: MRI scans in an elderly man at baseline (1993) and over 10 years, while he declined to MCI and AD. The reduction in size of the hippocampus (red) and entorhinal cortex (yellow) and the increase of the ventricle (green) with disease progression are illustrated by [Blennow et al. \(2006\)](#).

3.4 Current Treatment

Current therapies for AD are able to deal with the symptoms, temporally improving the quality of life for the patients and their caregivers. Nowadays, four different approaches are being considered to treat AD: symptomatic treatments, drug candidates with potential disease-modifying effects, treatments based on epidemiological studies and orientation to the caregiver ([Blennow et al., 2006](#); [Cummings, 2004](#); [National Institute on Aging, 2008](#)). Each of them has, or might have, its own effect on the neuropsychiatric symptoms, subject's cognitive functions or evolution of the disease ([Cummings, 2004](#)).

3.4.1 Symptomatic Treatments

Thanks to the knowledge about the neurotransmitter disturbances in AD, several drugs with symptomatic effects have been developed (Blennow et al., 2006). They treat the cognitive loss temporarily to improve the quality of life. However, none cure AD, nor do they halt disease progression (Lahiri et al., 2002). This section describes some of these treatments (Blennow et al., 2006).

Acetylcholinesterase Inhibitors

The cholinergic hypothesis suggests that AD is caused by degeneration of cholinergic neurons in the basal forebrain nuclei. This affects the pre-synaptic cholinergic terminals in the hippocampus and neo-cortex, which are important for cognitive processes such as memory (Blennow et al., 2006). Hence, it is assumed that the symptoms of AD can be temporarily mitigated by enhancing cholinergic neurotransmission. In order to do so, one tries to increase the availability of acetylcholine by inhibiting acetylcholinesterase, the enzyme that degrades acetylcholine in the synaptic cleft (Blennow et al., 2006).

The acetylcholinesterase inhibitors donepezil, rivastigmine, and galantamine are approved for clinical use in AD following this approach (Blennow et al., 2006; Doody et al., 2001; National Institute on Aging, 2008). Donepezil and galantamine are selective acetylcholinesterase inhibitors, while rivastigmine can also inhibit acetylcholinesterase and butyrylcholinesterase. Finally, galantamine can modulate pre-synaptic nicotinic receptors as well (Blennow et al., 2006; Doody et al., 2001).

Clinical trials have shown that these drugs produce modest positive effects on cognitive symptoms. They are also beneficial for functional and behavioural symptoms, without conclusive differences between their efficiency (Blennow et al., 2006). Moreover, their side-effects are usually limited to gastrointestinal symptoms such as nausea, vomiting and diarrhoea, which can often be reduced by taking these drugs with food and starting their prescription with low doses (Blennow et al., 2006).

Memantine

Memantine is a non-competitive *N*-methyl-D-aspartate (NMDA) receptor antagonist (Cummings, 2004). These receptors are related to memory and other cognitive functions. Memantine decreases the excitotoxicity produced by the glutamate, one of the major excitatory neurotransmitters in the brain (Blennow et al., 2006). Under healthy conditions, glutamate and the NMDA receptors are key for cognitive processes. Nevertheless, in AD, an increased glutamatergic activity can lower the activation of NMDA receptors, thus impairing the neuronal function (Blennow et al., 2006).

Memantine is slightly helpful for cognitive and behavioural symptoms in AD and improves the ability to carry out daily living activities in patients with moderate to severe AD (Cummings, 2004; National Institute on Aging, 2008). Moreover, it has been shown that the combination on memantine with donepezil or memantine produces more positive effects on the symptoms than the latter alone (Cummings, 2004). However, there is no conclusive evidence indicating that memantine is beneficial in mild disease. Of note is that this drug is generally well tolerated in general, with few side-effects (Blennow et al., 2006).

Treatment of Behavioural Signs

AD patients very frequently suffer from behavioural symptoms such as aggression, psychomotor agitation, hallucinations and delusions. These signs of AD are common in the late stages of the disease (Blennow et al., 2006). They not only worsen the patients' quality of life of life, but also increase the burden of their caregivers. Hence, it is important to manage these behavioural signs. Atypical antipsychotic drugs can be used to deal with psychosis or agitation (Blennow et al., 2006; Doody et al., 2001). Alternative treatments include anticonvulsants and short-acting benzodiazepines. Furthermore, the treatment with acetylcholinesterase inhibitors can also help to improve behavioural symptoms, which can be partially due to the cholinergic deficit (Blennow et al., 2006).

3.4.2 Modifying-Disease Drug Candidates

Due to the key role that $A\beta$ plays in AD important efforts have been devoted to inhibiting its production in the brain and increasing its clearance (Blennow et al., 2006). This has led to several research lines, which are listed in the following lines (Blennow et al., 2006).

Secretase Modulators

Secretases are enzymes that snip pieces off the APP. α -secretase enhances the non-amyloidogenic pathway of the APP. Hence, one possible approach is trying to stimulate this pathway in order to reduce $A\beta$ production (Blennow et al., 2006). On the other hand, β -secretase inhibitors have been developed to reduce brain $A\beta$ concentrations in AD transgenic mice (Blennow et al., 2006; Cummings, 2004). Finally, the inhibition of the γ -secretase might be another useful approach to decrease the amyloidogenic pathway of the APP. Moreover, γ -secretase can be more easily modulated than β -secretase. Nevertheless, there may be adverse effects when inhibiting γ -secretase (Blennow et al., 2006).

$A\beta$ Immunotherapy

The principle of $A\beta$ immunotherapy is trying to provoke an immunoresponse against the peptide deposits responsible for AD. This effect might be achieved by antibodies for $A\beta$ that bind to $A\beta$ plaques and induce $A\beta$ clearance by microglia or bind soluble $A\beta$ in the periphery, thus driving an $A\beta$ efflux away from the brain (Blennow et al., 2006). Clinical trials with active immunisation started a few years ago, but they had to be interrupted because 6% of cases developed encephalitis (Blennow et al., 2006; Cummings, 2004). Current research attempts at developing a second generation of immunotherapy with reduced side-effects (Blennow et al., 2006).

A β Fibrillisation Inhibitors

This research line study small peptides that interfere with A β -A β or A β -APOE interactions to prevent A β fibrillisation. It has been found that two peptides are able to reduce A β fibrillisation in vitro and brain A β load in AD transgenic mice, without inducing an immune response (Blennow et al., 2006). Some of these compounds are subject to clinical trials (Blennow et al., 2006).

Inhibitors of the Hyperphosphorylation of Tau Protein

Drug candidates that reduce tau phosphorylation by inhibiting tau kinases are in the preclinical phase (Blennow et al., 2006). However, since tau phosphorylation is regulated by the balance between multiple kinases and phosphates, inhibition of a single kinase might be insufficient to normalise tau phosphorylation (Blennow et al., 2006).

3.4.3 Possible Drugs Based on Epidemiological Studies

Epidemiological studies have provided the theoretical foundations for several treatment approaches. Observational studies have suggested a protective effect of different types of drugs or supplements. Nevertheless, when tested in randomised well-controlled clinical trials, beneficial effects have been difficult to prove (Blennow et al., 2006). Some of these approaches are listed below.

Anti-Inflammatory Drugs

Inflammatory characteristics accompany the plaques in AD. Nevertheless, it is unclear whether this inflammation has neurotoxic effects or only constitutes a secondary reaction to A β deposition (Blennow et al., 2006). Epidemiological studies indicate that the risk of this dementia is reduced in patients treated with non-steroidal anti-inflammatory drugs, which, in some cases, reduce brain A β burden in AD transgenic mice. Nevertheless, clinical trials on anti-inflammatory drugs showed no effects on cognition in AD (Blennow et al., 2006; Cummings, 2004). This can be due to the fact that these drugs

might be protective only if given during mid-life, but will not reverse the degenerative process in patients with established pathology (Blennow et al., 2006).

Oestrogens

Epidemiological studies pointed out a link between reduced risk of dementia and postmenopausal oestrogen supplementation (Doody et al., 2001; Reitz and Mayeux, 2009). Additionally, animal studies suggested that oestrogens could have beneficial effects on neuronal function (Blennow et al., 2006).

However, large well-controlled clinical trials have not shown a reduced risk of the disease (Blennow et al., 2006). Moreover, it has been suggested that therapy with a combination of oestrogens and progesterone may increase the risk of AD (Samanta et al., 2006).

Antioxidants

Observational studies suggest that dietary intake of antioxidants (e.g., vitamin E) could reduce the risk of the AD. It was found that vitamin E supplementation showed only marginal effects on time to institutionalisation and need of care in AD patients (Blennow et al., 2006; Doody et al., 2001). However, another study found no link between this supplementation and the rate of progression from MCI to AD (Blennow et al., 2006).

Cholesterol-Lowering Drugs

Several studies have been devoted to analysing the link between cholesterol and AD. However, this topic is still open to question, as both reduced and increased brain A β load has been reported when feeding AD transgenic mice a high-cholesterol diet (Blennow et al., 2006). In the past, retrospective studies suggesting that treatment with cholesterol-lowering drugs reduce the incidence of the dementia and the brain A β load received much attention. However, more recent studies have shown no association between the use of such drugs and reduced risk of AD. Moreover, a recent study suggests that

the relationship between cholesterol as a risk factor and AD may be more complex than expected, if not even inconsistent (Purnell et al., 2009).

3.4.4 Orientation to caregivers

In the early stages of AD, the clinician should encourage health maintenance activities, including exercise, the control of hypertension and other medical conditions (Doody et al., 2001). In later phases of the dementia, it is important to address basic requirements such as nutrition, hydration, and skin care (Cummings, 2004). Moreover, an alliance between the clinician and the caregiver is essential in treating patients with AD (Cummings, 2004).

Caregivers are responsible for supervising patients, administering medication, implementing non-pharmacologic treatment, and promoting the patient's general health, well-being and meaningful quality of life. Studies have shown that caregivers of patients with this dementia rated their own health as relatively poor (Cummings, 2004). Furthermore, they endured a greater number of illnesses, had more somatic symptoms, depression and anxiety, used more health care, and engaged in fewer preventive-health activities than people who were not AD caregivers. Interventions such as self-help groups, skill training, counselling and psychotherapy may help caregivers and reduce their psychological distress, yet they have failed to reduce the caregiver's burden (Cummings, 2004; Doody et al., 2001; National Institute on Aging, 2008).

Chapter 4

Blind Source Separation

Contents

4.1	Introduction	79
4.2	Classical Projection Techniques	81
4.2.1	Principal Component Analysis	81
4.2.2	Factor Analysis	84
4.2.3	Differences Among PCA, FA and BSS	85
4.3	BSS Fundamentals	86
4.3.1	Model and Validity of the Assumptions	87
4.3.2	Limitations of BSS	91
4.4	Second Order Statistics Methods for BSS	95
4.4.1	Fundamentals	96
4.4.2	AMUSE	97
4.4.3	SOBI	98
4.5	Higher Order Statistics Methods for BSS	99
4.5.1	Fundamentals	100
4.5.2	FastICA	101
4.5.3	InfoMax	104
4.5.4	JADE	106
4.6	Constrained BSS Algorithms	107
4.7	Relationships Among BSS Algorithms	109
4.7.1	Consistency of BSS Algorithms for Real Signals	112
4.8	BSS Applications to EEG and MEG Activity	114

4.8.1	Artefact Removal	115
4.8.2	Feature Enhancement With BSS	121
4.8.3	Extraction of Brain Rhythms With BSS	124

This Chapter details the fundamentals of BSS techniques. After an introduction about this topic and its utility, the classical techniques of PCA and FA are explained for the sake of comparison with BSS. Then, the basic ideas beneath this set of techniques are presented, attending at the SOS or HOS nature of the methods. The cBSS methodology is also considered. Afterwards, the relationships among diverse algorithms are briefly reviewed. Finally, the state of the art in the applications studied in this Doctoral Thesis is detailed.

4.1 Introduction

The application of BSS in the field of biomedical signal processing has become popular in the last years (James and Hesse, 2005; Onton et al., 2006; Vigário and Oja, 2008). These techniques are helpful to separate multi-channel biomedical recordings into their constituent underlying components (James and Hesse, 2005; Vigário and Oja, 2008). Their use in biomedical signal processing frequently involves the separation or extraction of statistically independent sources underlying the multivariate recorded biomedical signals (James and Hesse, 2005; Vigário and Oja, 2008).

As it was describe in Section 1.1, biomedical signal processing deals with the recording and mathematical processing of informative signals derived from living systems (Onaral, 2006). Although biomedical signals may originate from a wide variety of sources, we will focus on electromagnetic measurements. Most often, the recorded signals are at least temporally or spatially correlated (James and Hesse, 2005).

Additionally, the informative components of the signal may be contaminated with noise and artefacts making them difficult to observe by the naked eye (Laguna and Sörnmo, 2009). That is, the recorded data contain measurements of a finite set of separate but overlapping in space and time activities. As a result, the SNR of the interesting signal is generally poor (James and Hesse, 2005). While the clinicians may be able to visually extract the useful information from the signals, it would be extremely helpful to automate the analysis and isolate the constituent inner components of interest in the

signals (James and Hesse, 2005).

BSS denotes a set of data-driven approaches whose aim is to extract a set of underlying sources or components from a number of measurements or random variables. Very frequently, the term ICA is also used to refer to the BSS techniques that are strictly based on HOS and try to minimise the statistical dependence of the data (Vigário and Oja, 2008). BSS techniques define a generative model for the observed measurements and attempt to estimate the inner components from a large set of observed, multivariate recordings (James and Hesse, 2005).

A few strong general assumptions must be made about the data at hand before applying the BSS algorithms and interpreting the results (James and Hesse, 2005). The most important hypothesis is that the sources are mutually independent or, alternatively, that they are mutually decorrelated at any time delay. BSS extracts the sources, or components, by exploiting this independence, or spatio-temporal decorrelation (Hyvärinen et al., 2001; James and Hesse, 2005; Onton et al., 2006).

BSS is, indeed, a more powerful technique than classical decomposition methods such as PCA and FA (Hyvärinen et al., 2001; James and Hesse, 2005). For instance, the aim of PCA is to find temporally orthogonal directions that each successively explain as much of the remaining data variance as possible, where BSS tries to find directions in the joint data whose activities are as distinct from one another as possible. For this reason, PCA will *lump* whilst BSS will *split* the source activities across the resultant signal components (Onton et al., 2006).

The most well-known BSS algorithms were developed about 10 to 15 years ago (Cichocki and Amari, 2002; Comon, 1994; Hyvärinen et al., 2001). Soon after their development, these techniques were applied to diverse fields, including the study of cardiac and neurophysiologic recordings (James and Hesse, 2005).

One of the most fruitful applications of BSS algorithms is the artefact rejection in EEG and MEG recordings. This is based on the hypothesis that the artefacts are independent from the brain activity (James and Hesse, 2005; Onton et al., 2006). Thus, the artefactual activity can be isolated into a few

BSS components (Barbati et al., 2004; Cichocki and Amari, 2002; Delorme et al., 2007; Escudero et al., 2007b; Ting et al., 2006; Vorobyov and Cichocki, 2002). Nonetheless, this is not the only application of BSS to brain activity. These techniques have also been used to extract specific brain activity (Hung et al., 2007; Jarchi et al., 2009; Jin et al., 2002; Kobayashi et al., 1999).

In spite of those and other successful applications, certain ambiguities exist in the BSS model and that the results must be interpreted carefully (James and Hesse, 2005). Hence, it is important to know the fundamentals of BSS techniques to be able to properly apply these techniques and interpret the results.

4.2 Classical Projection Techniques

BSS techniques tend to produce more reliable and meaningful decompositions of multi-channel signals than classical decomposition methods, such as PCA and FA (James and Hesse, 2005). This is due to the principle of independence (or time-spatial decorrelation) beneath the former type of methods (James and Hesse, 2005). Nevertheless, the latter techniques are commonly applied as a preprocessing stage before the actual BSS algorithms (James and Hesse, 2005). Furthermore, some ideas and techniques are shared by PCA, FA and BSS (Cichocki and Amari, 2002; Hyvärinen et al., 2001). Hence, the classical methods of PCA and FA are briefly introduced in this Section.

4.2.1 Principal Component Analysis

PCA is one of the oldest and most well-known techniques in multivariate analysis and data-mining (Cichocki and Amari, 2002). Its objective is to find orthogonal directions that successively explain as much data variance as possible (Onton et al., 2006). This means that it tries to derive a relatively small number of decorrelated linear combinations (i.e., principal components) whilst retaining the maximum of the original variance (i.e., it attempts to reduce the data redundancy; Cichocki and Amari, 2002; Hyvärinen et al., 2001). It is used in data decomposition, classification, filtering and whitening,

which is a frequent preprocessing step in connection with BSS (Hyvärinen et al., 2001; Sanei and Chambers, 2007).

It is important to note that PCA, as well as FA and BSS, produce a transformation which depends on the data themselves. PCA is characterised by being efficient in terms of *energy compaction* (Hyvärinen et al., 2001; Sanei and Chambers, 2007). Usually, the principal components with the largest variance are supposed to have important signal information, whereas those with the smallest variance tend to be associated with noise (Cichocki and Amari, 2002).

The solution of PCA is based on the eigenstructure of the data covariance matrix by computing its eigenvectors and eigenvalues (Cichocki and Amari, 2002). Let $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_m(t)]^T$ denote a multidimensional signal composed of m simultaneous measurements at a particular time instant t . Here T denotes transposition. For the sake of simplicity, $\mathbf{x}(t)$ is supposed to have zero mean but no other explicit assumptions on the probability distributions of the data are made in PCA (Hyvärinen et al., 2001). PCA is based on computing the eigenvectors and eigenvalues of the estimated covariance matrix, $\mathbf{C}_{\mathbf{xx}}^0$, following (Cichocki and Amari, 2002):

$$\mathbf{C}_{\mathbf{xx}}^0 = E \{ \mathbf{x}(t) \mathbf{x}^T(t) \} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T \in \Re^{m \times m}, \quad (4.1)$$

where $\mathbf{\Lambda} = \text{diag} \{ \lambda_1, \lambda_2, \dots, \lambda_m \}$ is a diagonal matrix with the m eigenvalues sorted in descending order and $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m] \in \Re^{m \times m}$ is the corresponding unitary matrix containing the principal eigenvectors. It is important to note that all random variables are assumed to have zero mean (Cichocki and Amari, 2002).

The n first principal components of $\mathbf{x}(t)$, which are denoted as $\mathbf{y}_P(t) = [y_1(t), y_2(t), \dots, y_n(t)]^T$, are computed using the following linear transformation (Cichocki and Amari, 2002):

$$\mathbf{y}_P(t) = \mathbf{V}_n^T \mathbf{x}(t), \quad (4.2)$$

where $\mathbf{V}_n = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n] \in \Re^{m \times n}$ is the subset of the first n eigenvectors.

The principal components y_i are mutually uncorrelated and that their variances are the corresponding eigenvalues of the decomposition (Cichocki and Amari, 2002; Hyvärinen et al., 2001).

A central issue in PCA is the *selection of the number of principal components* to be retained (Cichocki and Amari, 2002; Hyvärinen et al., 2001). Since the principal components are characterised by retaining the maximum variance of the data, the usual approaches set a threshold in the sequence of eigenvalues that allows to divide them into the ‘signal’ and ‘noise’ subspaces of the data (Cichocki and Amari, 2002). However, this criterion may pose some problems when applied to brain recordings as a preprocessing step previous to a BSS, as it will be shown in Section 4.3.2 (Escudero et al., 2007b; James and Hesse, 2005).

It is remarkable that the estimation of the inner components in the BSS (or ICA) problem is often greatly simplified if the observed mixture vectors are first whitened (or sphered). A zero-mean random vector is ‘white’ if its elements z_i are uncorrelated and have unit variances:

$$\mathbb{E} \{z_i z_j\} = \delta_{ij}, \quad (4.3)$$

which is equivalent to

$$\mathbb{E} \{\mathbf{z}\mathbf{z}^T\} = \mathbf{I}, \quad (4.4)$$

where \mathbf{I} is the unitary matrix (Hyvärinen et al., 2001).

PCA is helpful to whiten the data, which can be achieved, for instance, by means of:

$$\mathbf{z} = \mathbf{Q}\mathbf{x}, \quad (4.5)$$

with

$$\mathbf{Q} = \mathbf{\Lambda}^{-1/2}\mathbf{V}^T, \quad (4.6)$$

where $\mathbf{\Lambda}^{-1/2} = \text{diag} \{ \lambda_1^{-1/2}, \lambda_2^{-1/2}, \dots, \lambda_m^{-1/2} \}$ and λ_i and \mathbf{V} were defined in Eq. (4.1) (Hyvärinen et al., 2001).

4.2.2 Factor Analysis

Although PCA does not assume any particular statistical model for the data, it can also be derived from the following generative latent model of FA (Hyvärinen et al., 2001):

$$\mathbf{x} = \mathbf{A}\mathbf{z} + \mathbf{n}, \quad (4.7)$$

where \mathbf{z} is gaussian, zero-mean and white (i.e, $E\{\mathbf{z}\mathbf{z}^T\} = \mathbf{I}$) and \mathbf{n} represents zero-mean gaussian white noise.

Assuming this model, if \mathbf{z} is gaussian, so is \mathbf{x} . In this case, the likelihood function can be formulated and the rows of \mathbf{A} can be estimated from the maximum likelihood solution when the noise power tends to zero (Hyvärinen et al., 2001). This approach, which is called principal factor analysis, is one of the methods for the statistical technique FA. Nevertheless, the goal in FA is usually different from that of PCA. FA generally tries to find *relevant and meaningful factors* able to explain the observations \mathbf{x} using the model (4.7) instead of to reduce the redundancy in the data. In the FA case, the elements of \mathbf{A} are the factor loadings whereas \mathbf{n} is usually interpreted as vector of specific factors instead of noise (Hyvärinen et al., 2001).

Let us assume that the variances of \mathbf{z} can be incorporated into \mathbf{A} and that the elements of \mathbf{n} are uncorrelated with each other and with the factors \mathbf{z} . Hence, $\mathbf{\Psi} = E\{\mathbf{n}\mathbf{n}^T\}$ is a diagonal matrix without restrictions on the values of the elements in the diagonal: ψ_{ii} . The covariance matrix of the observations \mathbf{x} is given by (Hyvärinen et al., 2001):

$$E\{\mathbf{x}\mathbf{x}^T\} = \mathbf{C}_{\mathbf{xx}}^0 = \mathbf{A}\mathbf{A}^T + \mathbf{\Psi}. \quad (4.8)$$

Unfortunately, there is no close analytic solution for \mathbf{A} and $\mathbf{\Psi}$ (Hyvärinen et al., 2001).

Nevertheless, if an estimation of $\mathbf{\Psi}$ is available, one might try to find \mathbf{A} from:

$$\mathbf{A}\mathbf{A}^T = (\mathbf{C}_{\mathbf{xx}}^0 - \mathbf{\Psi}). \quad (4.9)$$

However, Eq. (4.9) cannot be solved exactly either due to the fact that the

number of factors (n) is constrained to be smaller than the number of original dimensions in the data (m). Therefore, a least-square approach can be used. Nevertheless, this solution is not unique since any orthogonal transformation or rotation $\mathbf{A} \rightarrow \mathbf{AT}$, with $\mathbf{TT}^T = \mathbf{I}$, will also produce $(\mathbf{C}_{\mathbf{xx}}^0 - \Psi)$ (Hyvärinen et al., 2001). Thus, additional constraints, typically in the form of sparsity in the mixing matrix, are taken (Hyvärinen et al., 2001).

4.2.3 Differences Among PCA, FA and BSS

At this point, it may be convenient to clarify some key differences between these classical projection techniques, PCA and FA, and BSS.

First of all, *PCA does not assume any generative model* of the data, although it can be derived from one. Additionally, only if all the principal components are retained, the PCA is a completely invertible linear transformation of the data (Hyvärinen et al., 2001).

Additionally, FA does assume a generative latent variable model, but the values of the factors cannot be directly calculated from the observations due to the term of *specific factors*, or noise, \mathbf{n} (Hyvärinen et al., 2001). The rows of \mathbf{A} are not directly given by the eigenvectors of $\mathbf{C}_{\mathbf{xx}}^0$ (Hyvärinen et al., 2001).

Finally, both FA and PCA rely on the covariance of the observed variables at time lag $\tau = 0$ to estimate the components or factors (Hyvärinen et al., 2001). On the other hand, BSS methods have an underlying generative model similar to that of FA but they remove the rotational redundancy of the FA model by imposing the *independence of non-gaussian variables* or using the *temporal structure of the data* (Hyvärinen et al., 2001; James and Hesse, 2005).

To sum up, BSS techniques can be seen as an extension of PCA and FA which is capable of finding the inner factors in situations where those techniques fail (Hyvärinen et al., 2001). The differences in the outputs of this variables are illustrated in Fig. 4.1, which shows two mixed random variables with uniform distributions and the outcome of a PCA and an ICA.

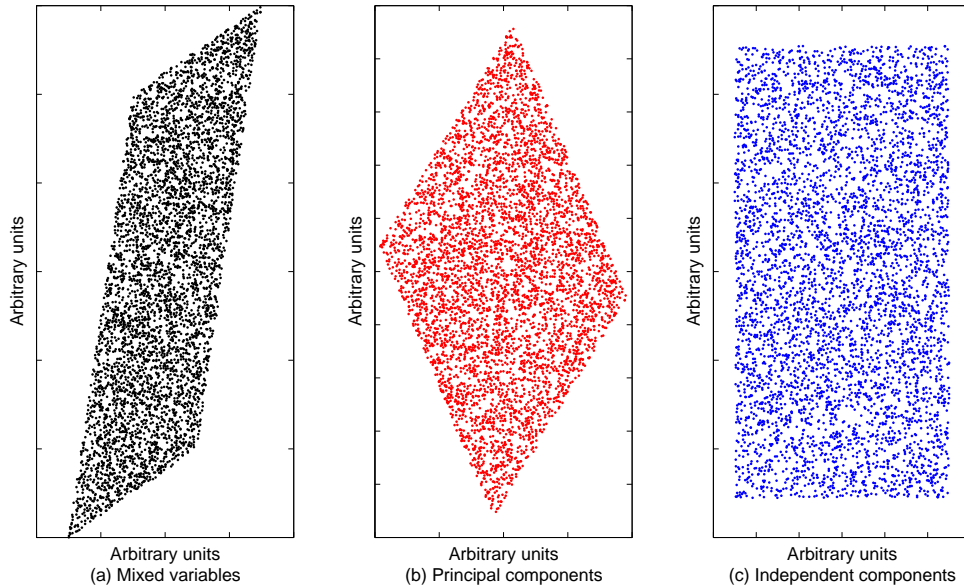


Figure 4.1: Example of the differences between PCA and ICA. The scale of the plots is arbitrary. (a) Mixed random variables with uniform distribution. (b) PCA. (c) ICA.

4.3 BSS Fundamentals

The concept of BSS refers to a set of statistical and computational techniques for revealing hidden components that underlie sets of random variables, measurements or signals (Hyvärinen et al., 2001). Here, *source* refers to an original signal while *blind* indicates that very little, if any, information is known about the mixing process of those sources (Hyvärinen et al., 2001).

BSS defines a generative model of the observed data and tries to estimate the constituent inner components assuming that they are independent of each other (Hyvärinen et al., 2001). A set of variables is strictly independent from a statistical point of view if its joint probability distribution, $p_{\mathbf{x}}(\mathbf{x})$ is the product of the marginal distributions, $p_x(x_i)$ (Comon, 1994; Sanei and Chambers, 2007):

$$p_{\mathbf{x}}(\mathbf{x}) = \prod_i p_x(x_i). \quad (4.10)$$

However, the actual probability distributions are exceedingly difficult to estimate. Thus, alternative approaches must be taken (Hyvärinen et al., 2001).

4.3.1 Model and Validity of the Assumptions

The problem that BSS techniques (Cichocki and Amari, 2002; Hyvärinen et al., 2001) attempt to solve is stated in the following lines. Likewise Section 4.2.1, let the m -dimensional vector $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_m(t)]^T$ denote a set of m data samples at a particular time instant t . $\mathbf{x}(t)$ is supposed to be a combination of n underlying, inner components, or sources, $\mathbf{s}(t) = [s_1(t), s_2(t), \dots, s_n(t)]^T$. In the simplest case, the mixing process is assumed to be linear and the external noise is disregarded, which leads to:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t), \quad (4.11)$$

where \mathbf{A} is a full rank $m \times n$ mixing matrix. Additionally, it is assumed that $m \geq n$ (Hyvärinen et al., 2001; James and Hesse, 2005; Vigário and Oja, 2008).

BSS algorithms compute an estimation, $\hat{\mathbf{s}}(t)$, where $\hat{\cdot}$ denotes an estimated variable, of the real underlying components $\mathbf{s}(t)$ from the measurements $\mathbf{x}(t)$ by means of:

$$\hat{\mathbf{s}}(t) = \mathbf{B}\mathbf{x}(t), \quad (4.12)$$

where \mathbf{B} is the estimated separating, or de-mixing, matrix, which is the pseudo-inverse of \mathbf{A} (Cichocki and Amari, 2002; Hyvärinen et al., 2001).

The model presented in Eq. (4.11) is actually a simplification made for the ease of implementation (Cichocki and Amari, 2002; James and Hesse, 2005). A more general mixing model would include non-linearity in the system and external noise, resulting in:

$$\mathbf{x}(t) = \mathbf{f}[\mathbf{s}(t)] + \mathbf{n}(t), \quad (4.13)$$

where $\mathbf{f}[\cdot]$ denotes a function and $\mathbf{n}(t)$ is the additive external noise (James and Hesse, 2005).

Some *assumptions* are needed to tackle the estimation of the inner components if no information is available about $\mathbf{f}[\cdot]$, $\mathbf{s}(t)$ or $\mathbf{n}(t)$ (Cichocki and Amari, 2002; Hyvärinen et al., 2001). Hence, the following hypotheses, which are feasible in most biomedical signal processing applications, are assumed to

make the problem more tractable (James and Hesse, 2005; Sanei and Chambers, 2007).

Instantaneous Linear Mixing

The assumption of linear mixing is essential to simplify the model. This transform Eq. (4.13) into:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t). \quad (4.14)$$

The linear and instantaneous mixing hypothesis implies that the recorded measurements are a superposition of attenuated sources at each channel (Cichocki and Amari, 2002; James and Hesse, 2005). This assumption holds for electromagnetic brain signals, where the signals have narrow bandwidths and the sampling frequency is relatively low (Sanei and Chambers, 2007). Nevertheless, it may pose problems for other kind of signals, such as sound signals measured via microphones, which may involve convolutive mixtures (James and Hesse, 2005; Sanei and Chambers, 2007).

Stationary Mixing

This is another common hypothesis. It consists of assuming that the mixing matrix \mathbf{A} does not change with time (Cichocki and Amari, 2002; Hyvärinen et al., 2001). This implies that the physical properties of the mixing process remain stable. For instance, this hypothesis may not hold when ECG is recorded on chest electrodes which move over time due to breathing (James and Hesse, 2005). Fortunately, the situation for electromagnetic brain recordings is simpler. For these signals, the assumption of a stationary mixing matrix can be interpreted as the biophysical structure of the brain itself is fixed while the sources that are activated within this structure may change their intensity over time (James and Hesse, 2005; Onton et al., 2006).

Independence and Temporal Decorrelation

The most important assumption made in BSS is that the sources are *mutually statistically independent* or, alternatively, that they are *decorrelated at any time delay* (Cichocki and Amari, 2002). This is a stronger assumption than mere uncorrelatedness at time lag zero. While statistically independent sources are necessarily uncorrelated, the converse does not hold (Hyvärinen et al., 2001).

As is was shown in Eq. (4.10), the general condition for two random variables to be statistically independent is that their joint distribution equals the product of their marginal probability functions (Comon, 1994; Hyvärinen et al., 2001; James and Hesse, 2005). This implies that independent variables are not only uncorrelated at time lag zero but also that they have no higher order correlations and that they are uncorrelated for any time lag (Cichocki and Amari, 2002). This hypothesis is quite plausible in some situations (James and Hesse, 2005). For instance, artefacts in EEG and MEG signals can be considered independent from the brain activity (Vigário and Oja, 2008).

It is important to notice that the algorithms based on the statistical independence require all components but one to have non-gaussian probability distributions. This need for non-gaussian probability distributions can be intuitively justified considering the Central Limit Theorem (Hyvärinen et al., 2001). Briefly, this theorem states that the distribution of a sum of independent random variables tends towards a gaussian distribution under certain conditions. Thus, it can be hypothesised that a sum of two or more independent random variables is often more gaussian than any of the original variables and that, when looking for the components of this sum, the least gaussian estimated component is likely to be an actual original source (Hyvärinen et al., 2001).

On the other hand, in order for the methods based on the temporal structure of the data to work properly, the sources must have different frequency spectra so that their temporal structure capture by a set of covariances matrices differs (Hyvärinen et al., 2001).

Justification of the Previous Assumptions

The main assumptions of linear, instantaneous and stationary mixing process have been validated for EEG and MEG recordings (Cichocki and Amari, 2002; Hyvärinen et al., 2001; James and Hesse, 2005; Vigário and Oja, 2000, 2008).

The hypothesis that mixing process is linear and instantaneous is supported by the Maxwell Equations (Hyvärinen et al., 2001; Vigário and Oja, 2000, 2008). This is due to the fact that the quasi-static approximation is valid as the spectral content of the signals lies below 1 kHz (Hämäläinen et al., 1993).

The stationarity of the mixture can be justified since the populations of neurons do not move in relation to the sensors (Vigário and Oja, 2008). In some cases, there exists rapid cortical signal propagation. This violates the stationarity assumption (Vigário and Oja, 2008). Nevertheless, a set of multiple consecutive stationary components may be considered instead (Vigário and Oja, 2008).

It is clear that the different nature and origin of artefacts and brain signals justifies the important assumption of independence in this case (Vigário and Oja, 2008). Within brain activity, the independence assumption is consistent with the fact that cortex is organized into compact regions of specialized function (Onton et al., 2006). Consider the fact that the density of long-range cortical connections is low compared with the density of local connections. Thus, it can be hypothesised that, over sufficient time, locally synchronous activities within particular regions of the cortex are, indeed, nearly temporally independent of each other. Hence, they may be viewed as single, distinct, temporally independent sources of activity (Onton et al., 2006). On the other hand, locally synchronized activities in a pair of cortical regions that are densely connected to each other may form a single effective source (Onton et al., 2006). Therefore, the recorded brain activity may be modelled as a sum of distinct, nearly-independent, signals (Onton et al., 2006).

Further Assumptions: Noiseless and Square Mixing

Some other assumptions can be made in order to further simplify the problem. However, these are more problematic in the context of electromagnetic brain signal analysis (James and Hesse, 2005):

- Noiseless mixing. A commonly made assumption is that the observations $\mathbf{x}(t)$ are noiseless or, at least, that the power of the external noise vector $\mathbf{n}(t)$ is negligible (Hyvärinen et al., 2001; James and Hesse, 2005; Vigário and Oja, 2008). Thus, Eq. (4.14) becomes Eq. (4.11). This hypothesis is not realistic from the point of view that biomedical recordings are always measured with some additive noise. Nevertheless, this assumption still allows the techniques to decompose the signals even if the actual separate components themselves are contaminated by some measurement noise (James and Hesse, 2005).
- Square mixing. The general model for BSS assumes that $m \geq n$. Thus, \mathbf{A} may be non-square (Cichocki and Amari, 2002; James and Hesse, 2005). Actually, from a physiological point of view, it is likely that the number of significant inner *meaningful* components (n) is less than the number of measured channels, m (James and Hesse, 2005). Nevertheless, many BSS algorithms assume that the number of components is equal to the number of recorded signals ($m = n$). This simplifies the decomposition problem (Hyvärinen et al., 2001; James and Hesse, 2005). Nevertheless, high-density measurements are sometimes made during short time periods. In these cases, this assumption is less likely (James and Hesse, 2005). Due to this reason, dimensionality-reduction techniques are usually applied as a preprocessing before the BSS techniques (James and Hesse, 2005; Vigário and Oja, 2008).

4.3.2 Limitations of BSS

Despite the increasing popularity of BSS techniques in the biomedical signal processing field, it is important to recall that these techniques make a few strong, but plausible, assumptions about the data. Hence, careful attention

must be paid to the fulfilment of these hypotheses and to the interpretation of the results (James and Hesse, 2005). It is also necessary to consider whether the data at hand are suitable to be fed into a particular kind of BSS algorithm or whether some preprocessing is needed (Hyvärinen et al., 2001). To sum up, the application of BSS techniques as ‘black-box’ method may degrade the quality of the results (James and Hesse, 2005). Therefore, some limitations foreseen in the application of BSS techniques to biomedical recordings are presented in this Section.

Ambiguities

It must also be noticed that, due to the fact that neither \mathbf{A} nor $\mathbf{s}(t)$ are known, some ambiguities appear in the BSS decomposition (Comon, 1994; James and Hesse, 2005). These indeterminacies are not insurmountable, though. They do not pose serious limitations on the accuracy and applicability of the BSS analysis (James and Hesse, 2005; Sanei and Chambers, 2007). They are the following:

- *Neither the energies nor signs of the components can be computed.* The reason is that, both $\mathbf{s}(t)$ and \mathbf{A} being unknown, any scalar multiplier in the one of the sources, s_i can always be cancelled by dividing the corresponding column \mathbf{a}_i of \mathbf{A} by the same scalar (Comon, 1994; Hyvärinen et al., 2001). To solve this ambiguity, either each component computed in $\hat{\mathbf{s}}(t)$ or every column of the estimated mixing matrix $\hat{\mathbf{A}}$ may be normalised to, for example, unit variance and the remaining factor may be incorporated into the corresponding term (James and Hesse, 2005). Thus, neither the sign of the scalp maps nor the sign of the activations are meaningful in themselves, but only their product, which determines the sign of the signal acquired at each scalp channel (Onton et al., 2006).
- *There is no fixed ordering between the sources a priori.* The reason is the following. Again, both $\mathbf{s}(t)$ and \mathbf{A} are unknown. Hence, it is possible to change the order of the terms in Eq. (4.11) using a permutation matrix \mathbf{P} to give $\mathbf{x} = (\mathbf{A}\mathbf{P}^{-1})(\mathbf{P}\mathbf{s})$ (Comon, 1994; Hyvärinen

et al., 2001). However, the components may be ranked a posteriori following any particular suitable criterion, such as their power (James and Hesse, 2005), temporal structure (Vorobyov and Cichocki, 2002) or spectral content (Escudero et al., 2008d, 2009b).

Interpretation

Despite the fact that the previous ambiguities do not pose insoluble problems in the analysis of the BSS results, caution should always be taken in their interpretation (James and Hesse, 2005).

First of all, it should be noted that the BSS components are not necessarily equivalent current dipoles or fMRI activated zones. Instead, they are concurrent, spatially and temporally decorrelated electromagnetic activities which were added over the scalp to originate the measured recordings (Cichocki and Amari, 2002; Vorobyov and Cichocki, 2002).

Secondly, the real components underlying the measured recordings are almost never known in biomedical signal processing applications. Therefore, the interpretation of the results may sometimes be based on subjective criteria related to the expected outcomes of the analysis (James and Hesse, 2005).

Finally, HOS-BSS techniques are regarded as very helpful in finding artefacts in EEG and MEG signals, but Hyvärinen et al. (2010) have suggested that they are less so in finding components related to brain activity. This might be due to the fact that ICA looks for components with maximum non-gaussian distributions, while the ‘interesting’ sources in EEG and MEG are sometimes not very far from gaussian (Hyvärinen et al., 2010). Another suggested problem may be the implicit assumption that each oscillatory source is observed in all the channels at the same phase (or with π radian phase shift). This is true for a single current dipole source but it may not be the case for more distributed sources (Hyvärinen et al., 2010).

Model Order Selection

As for the restriction $m \geq n$, it is obvious that the total number of possible cortical activity generators far exceeds the number of channels recorded in EEG or MEG. Yet, brain activity can only be measured over the scalp if post-synaptic activity from millions of synapses is synchronously active. As a consequence, the number of *effective* sources is considerably small in comparison with the number of channels in the EEG and, specially, MEG equipment (Hesse, 2008; Vigário and Oja, 2008). Thus, the selection of model order (n) may be a challenging task in the case of MEG recordings, where hundreds of channels are simultaneously recorded (James and Hesse, 2005). Despite the attention that BSS has received in recent years, the correct estimation of this parameter has been somewhat disregarded considering that an accurate estimation of n can have major impact on the quality of the decomposition (James and Hesse, 2005).

Therefore, it can be useful to apply a dimensionality reduction before BSS. Firstly, when high-density recording equipment is used, the number of macroscopic inner components may be less than the number of available channels for analysis (Ikeda and Toyama, 2000; James and Hesse, 2005).

Secondly, a dimensionality reduction can sometimes be needed to avoid ‘overfitting’ or ‘overlearning’ (Hyvärinen et al., 2001; Vigário and Oja, 2008), which can lead to the extraction of meaningless components (James and Hesse, 2005). ‘Overfitting’ may occur if a too high value is assigned to n . This would imply that the number of free parameters in the model is too high in comparison with the available sample size (Hyvärinen et al., 2001; James and Hesse, 2005; Vigário and Oja, 2008). ‘Overfit’ components of HOS techniques have a single spike or bump and are practically zero everywhere while for SOS algorithms they are characterised by various periodic solutions, mostly sinusoidal (Vigário and Oja, 2008).

Finally, it must be considered that computing the BSS with fewer components than the optimal value will result in some or all of the extracted components summing activity from more than one underlying source. Nevertheless, even in this case, the algorithms should efficiently produce mixtures

with minimal mutual information (Onton et al., 2006). Additionally, the dimensionality reduction helps to diminish the importance of the outer noise if this is considered in the model (Ikeda and Toyama, 2000).

Usually, the dimensionality reduction is performed by standard PCA (Hyvärinen et al., 2001; James and Hesse, 2005). The number of sources is supposed to be equal to the number of dominant eigenvalues, which is defined as the number of them that account for a certain fraction of the total observed variance (for example, 95% or 99%) or the number of eigenvalues with a contribution larger than a threshold (for instance, 1%) to the total variance (Hesse, 2008; Hyvärinen et al., 2001; James and Hesse, 2005; Ting et al., 2006). However, this approach has some major drawbacks. First of all, this criterion involves some arbitrariness and there is no theoretical guidelines supporting one choice or another (Hyvärinen et al., 2001; James and Hesse, 2005). Moreover, it assumes that the power of the additive is close to zero (Hyvärinen et al., 2001), something that may not be true in un-averaged MEG data (Cao et al., 2003; Hesse, 2008).

To overcome this problem, the eigenvalue spectrum may be split into ‘signal’ and ‘noise’ subspaces by a power threshold (Cichocki and Amari, 2002; Hyvärinen et al., 2001). Then, the external noise power is estimated from the noise subspace and it is subtracted from the signal subspace (Barbati et al., 2004; Hesse, 2008; Ting et al., 2006). However, this method assumes that all MEG channels have the same noise power, and it retains the aforementioned subjectivity. Moreover, there is no a priori reason to suppose that the sources of interest are contained in the signal subspace spanned by the dominant principal components (James and Hesse, 2005).

4.4 Second Order Statistics Methods for BSS

There are different approaches to perform the BSS of multi-dimensional time series (Cichocki and Amari, 2002; Hyvärinen et al., 2001). Whereas some methods are based on HOS of the source waveform distribution (commonly known as ICA; Vigário and Oja, 2008), other consider the temporal (i.e., spectral) structure of the signals and are known as SOS-based techniques (James

and Hesse, 2005).

The assumption that the components are independent implies that they have *no spatial temporal or spatial time-frequency correlations* (James and Hesse, 2005). Hence, it is possible to capture the temporal dependency of the measurements using a stack of square matrices to, afterwards, estimate the de-mixing matrix of the BSS model as the approximate joint diagonaliser of the stack (James and Hesse, 2005). Actually, these techniques are able to estimate the model where methods based on HOS fail since, for example, the sources are correlated over time but gaussian (Hyvärinen et al., 2001).

4.4.1 Fundamentals

The simple zero-lagged covariance does not contain enough information to estimate the independent sources. Actually, diagonalising this covariance produce white, but not independent, signals (Hyvärinen et al., 2001). However, the matrices \mathbf{A} in Eq. (4.11) and \mathbf{B} in Eq. (4.12) can be estimated if the temporal structure is taken into account in the form of *a set of lagged covariances* (Hyvärinen et al., 2001).

The simplest type of time structure is given by autocovariances. Let $\mathbf{C}_{\mathbf{xx}}^\tau$ be the τ -th covariance matrix of the measurements $\mathbf{x}(t)$, that is, $\mathbf{C}_{\mathbf{xx}}^\tau = E \left\{ \mathbf{x}(t) \mathbf{x}(t - \tau)^\top \right\}$. Likewise, $\mathbf{C}_{\mathbf{ss}}^\tau$ is the τ -th covariance matrix for the components $\mathbf{s}(t)$. In this formulation, τ indicates a matrix in the stack and represents the time lags ($\tau = 0, 1, 2, \dots, L$) at which the cross-covariances were computed (James and Hesse, 2005). All random variables and signals are assumed to have zero mean.

They are related to each other via:

$$\mathbf{C}_{\mathbf{xx}}^\tau = \mathbf{A} \mathbf{C}_{\mathbf{ss}}^\tau \mathbf{A}^\top, \quad (4.15)$$

where \mathbf{A} represents the mixing matrix as in Eq. (4.11) (James and Hesse, 2005). Therefore, the covariance of the components is obtained from $\mathbf{C}_{\mathbf{xx}}^\tau$ with the inversion:

$$\mathbf{C}_{\mathbf{ss}}^\tau = \mathbf{B} \mathbf{C}_{\mathbf{xx}}^\tau \mathbf{B}^\top, \quad (4.16)$$

where \mathbf{B} is the de-mixing matrix (James and Hesse, 2005).

The most common approach is to estimate the de-mixing matrix (\mathbf{B}) first. This is based on the idea that \mathbf{C}_{ss}^τ must be diagonal. Thus, \mathbf{B} can be estimated so that the matrix $\mathbf{B}\mathbf{C}_{xx}^\tau\mathbf{B}^T$ is as diagonal as possible (James and Hesse, 2005), where the matrix ‘diagonality’ may be quantified, for example, as the weight of the off-diagonal elements (James and Hesse, 2005). These methods require the components to have unique power spectra (Cichocki and Amari, 2002). On the other hand, they can deal with gaussian sources and they are robust in the presence of white noise (James and Hesse, 2005; Sanei and Chambers, 2007).

Actually, this approach is conceptually simple. It was originally proposed for only two time lags on the basis of two consecutive eigenvalue decompositions, resulting in the Algorithm for Multiple Unknown Signals Extraction (AMUSE) which is described in Section 4.4.2 (Tong et al., 1991). This methodology can be extended to several time lags, although in this cases iterative techniques must be used, such as in the Second-Order Blind Identification (SOBI) algorithm introduced in Section 4.4.3 (Belouchrani et al., 1997; Ziehe and Muller, 1998).

4.4.2 AMUSE

In the simplest case, only two time lags are considered. AMUSE is a very simple algorithm which cancels the instantaneous covariances at those time lags (Hyvärinen et al., 2001; Tong et al., 1991). Since only two time delays are considered, its computational complexity is low (Cichocki and Amari, 2002). Furthermore, AMUSE always offers the same separation when it is applied to the same input data set and it orders the components by decreasing linear predictability (Cichocki et al., 2005).

First of all, AMUSE decorrelates the signals at $\tau_0 = 0$ by applying a PCA to the input data, $\mathbf{x}(t)$. The result is a whitened signal (Cichocki and Amari, 2002; Hyvärinen et al., 2001; Tong et al., 1991). This whitening process can be done by:

$$\mathbf{z}(t) = \mathbf{Q}\mathbf{x}(t), \quad (4.17)$$

where $\mathbf{Q} = [\mathbf{C}_{\mathbf{xx}}^0]^{-1/2}$ (Escudero et al., 2008d, 2009b).

Afterwards, the signals are decorrelated at another time delay, τ_1 (usually $\tau_1 = 1$ sample; Cichocki and Amari, 2002; Escudero et al., 2009b; Hyvärinen et al., 2001; Ting et al., 2006). A time-delayed covariance matrix is computed as $\mathbf{C}_{\mathbf{xx}}^{\tau_1} = E \left\{ \mathbf{x}(t) \mathbf{x}(t - \tau_1)^T \right\}$ and the eigenvalue decomposition of:

$$\bar{\mathbf{C}}_{\mathbf{xx}}^{\tau_1} = \frac{\mathbf{C}_{\mathbf{xx}}^{\tau_1} + \mathbf{C}_{\mathbf{xx}}^{\tau_1 T}}{2} \quad (4.18)$$

is calculated (Cichocki and Amari, 2002; Escudero et al., 2008d, 2009b).

If \mathbf{V}_{τ_1} denotes the eigenmatrix computed from this decomposition, the demixing matrix \mathbf{B} is estimated as (Tong et al., 1991):

$$\mathbf{B} = \mathbf{V}_{\tau_1}^T \mathbf{Q}. \quad (4.19)$$

As it has been described, AMUSE consists of a whitening process followed by a singular value decomposition. Thus, the computational complexity of this algorithm is low (Cichocki et al., 2005; Hyvärinen et al., 2001). However, this method requires the eigenvectors of $\bar{\mathbf{C}}_{\mathbf{xx}}^{\tau_1}$ to be uniquely defined, which might restrict the applicability of this method (Hyvärinen et al., 2001).

4.4.3 SOBI

An extension of the AMUSE technique that improves the quality of the separation is to consider more than two several time lags: $\tau = 0, 1, 2, \dots, L$. In this way, it is enough that the components' covariances for one of those time lags are different to estimate the components (Hyvärinen et al., 2001). However, the simultaneous diagonalisation of the corresponding covariance matrices cannot be computed exactly (Belouchrani et al., 1997; Hyvärinen et al., 2001; James and Hesse, 2005).

The matrices $\mathbf{C}_{\mathbf{ss}}^{\tau}$ must be diagonal. Therefore, \mathbf{B} is estimated so that those matrices are as diagonal as possible (Belouchrani et al., 1997; James and Hesse, 2005). The 'diagonality' of a matrix \mathbf{M} can be measured with the

operator (Belouchrani et al., 1997; Hyvärinen et al., 2001):

$$\text{off}(\mathbf{M}) = \sum_{i \neq j} |m_{ij}|^2. \quad (4.20)$$

Hence, the sources are estimated by minimising the sum of the off-diagonal elements of several lagged covariances of $\mathbf{y} = \mathbf{W}\mathbf{z}$, where \mathbf{z} represents the already whitened signals. This leads to the following objective function, \mathcal{J} , where S denotes the set of considered time lags, τ (Belouchrani et al., 1997; Hyvärinen et al., 2001):

$$\mathcal{J}(\mathbf{W}) = \sum_{\tau \in S} \text{off}(\mathbf{W}\bar{\mathbf{C}}_{\mathbf{z}\mathbf{z}}^{\tau}\mathbf{W}^T), \quad (4.21)$$

with $\bar{\mathbf{C}}_{\mathbf{z}\mathbf{z}}^{\tau}$ computed as in Eq. (4.18) (Hyvärinen et al., 2001).

The estimation method consists of minimising \mathcal{J} under the constraint that \mathbf{W} is orthogonal. This is the principle of SOBI (Belouchrani et al., 1997).

It must be noted that, if the components have identical autocovariances (i.e., identical power spectra), they cannot be estimated by either AMUSE or SOBI (Hyvärinen et al., 2001). This contrasts to HOS methods, which allow the components to have identical power spectra or even equal distributions as long as they are non-gaussian (Hyvärinen et al., 2001).

4.5 Higher Order Statistics Methods for BSS

Other algorithms employ HOS to solve the BSS problem, an approach usually known as ICA. This methodology assumes that the recordings $\mathbf{x}(t)$ are observations of random variables linearly mixed through \mathbf{A} , where the temporal structure is irrelevant (Cichocki and Amari, 2002; Hyvärinen et al., 2001; James and Hesse, 2005). Thus, these are the only methods that can be applied to purely random variables without any temporal structure (Hyvärinen et al., 2001).

The probability distributions of the sources must be non-gaussian. This

procedure relies on the *Central Limit Theorem*: sums of non-gaussian random variables are closer to gaussian than the original ones (Hyvärinen et al., 2001). Hence, searching for non-gaussian sources is equivalent to searching for more independent sources (Hyvärinen et al., 2001; James and Hesse, 2005).

It must be borne in mind that these methods will provide successful decomposition only if all independent sources but one have non-gaussian distributions (Hyvärinen et al., 2001; James and Hesse, 2005). Despite this limitation, this approach is exceedingly popular in the biomedical signal processing field and some of the most commonly employed BSS techniques implement it: Hyvärinen–Oja’s FastICA Algorithm (FastICA), Bell & Sejnowski’s Information Maximisation Algorithm (InfoMax) and Joint Approximate Diagonalization of Eigenmatrices (JADE).

4.5.1 Fundamentals

These techniques assess the independence of the estimated components relying on statistics such as kurtosis and negentropy (Hyvärinen et al., 2001). Therefore, these metrics will be introduced in the following lines.

The *kurtosis*, $\text{kurt}(\cdot)$, is the fourth-order cumulant of a random variable and can be computed as (Hyvärinen et al., 2001):

$$\text{kurt}(x) = E\{x^4\} - 3(E\{x^2\})^2. \quad (4.22)$$

Kurtosis can be either positive or negative. The random variables associated with positive and negative kurtoses are called supergaussian and subgaussian, respectively (Hyvärinen et al., 2001). Supergaussian random variables typically have a ‘spiky’ probability distribution with heavy tails, whereas subgaussian ones tend to be ‘flat’ with small tails (Hyvärinen et al., 2001). Thanks to the fact that kurtosis is identically zero for gaussian distributed signals, maximising the magnitude of this cumulant is equivalent to make the inner components as non-gaussian (i.e., as independent) as possible (James and Hesse, 2005).

Negentropy is a statistic based on the concept of differential entropy (Hyvärinen et al., 2001). Negentropy, $J(\mathbf{x})$, is zero only for a gaussian vari-

able and always non-negative. It is defined as (Comon, 1994; Hyvärinen et al., 2001):

$$J(\mathbf{x}) = H(\mathbf{x}_{gauss}) - H(\mathbf{x}), \quad (4.23)$$

where $H(\cdot)$ represents the entropy of a variable and \mathbf{x}_{gauss} is a gaussian random vector with the same covariance matrix, $\mathbf{C}_{\mathbf{xx}}^0$, as \mathbf{x} . It is important to note that negentropy is invariant for invertible linear transformations and that, for a standardised random variable, the following approximation holds (Hyvärinen et al., 2001, p. 115):

$$J(x) \approx \frac{1}{12} E\{x^3\}^2 + \frac{1}{48} \text{kurt}(x)^2, \quad (4.24)$$

which gives an idea of its utility as a non-gaussianity measure.

4.5.2 FastICA

FastICA (Hyvärinen, 1999; Hyvärinen and Oja, 1997) tries to decompose the signals on the basis of their non-gaussianity (Hyvärinen et al., 2001). It is a fast fixed-point iterative algorithm which searches projections that maximize the non-gaussianity of components by their kurtosis (Hyvärinen and Oja, 1997).

The simplest cost function for an algorithm based on the kurtosis, $\mathcal{J}(\mathbf{W})$, is defined as (Sanei and Chambers, 2007):

$$\mathcal{J}(\mathbf{W}) = -\frac{1}{4} |\text{kurt}(\mathbf{y})| = -\frac{\beta}{4} \text{kurt}(\mathbf{y}), \quad (4.25)$$

where \mathbf{y} is an estimated component and β is the sign of the kurtosis (Sanei and Chambers, 2007).

This cost function can be minimised using the standard gradient descent approach, which leads to:

$$\mathbf{W}(k+1) = \mathbf{W}(k) - \mu \left. \frac{\partial \mathcal{J}(\mathbf{W})}{\partial \mathbf{W}} \right|_{\mathbf{w}(k)} = \mathbf{W}(k) + \mu(k) \varphi(\mathbf{y}(k)) \mathbf{x}(k), \quad (4.26)$$

where $\mu(k)$ is the learning rate and $\phi(\mathbf{y}(k))$ depends on the second, third and

fourth order moments of \mathbf{y} (Sanei and Chambers, 2007). Since this algorithm extracts one source at a time, a deflationary process must be followed to exclude the extracted source from the remaining mixtures (Hyvärinen, 1999).

This kind of gradient descent approach enables a fast adaptation in a non-stationary environment. However, its convergence is slow and depends on a good choice of the learning rate sequence. Therefore, other approaches are needed to solve these problems. A possible alternative is the fixed-point iteration algorithm (Hyvärinen, 1999; Hyvärinen and Oja, 1997; Hyvärinen et al., 2001).

In order to derive a more efficient fixed-point iteration, it is important to note that, at a *stable point* of the gradient algorithm, the gradient must point in the same direction of \mathbf{w} . Then, adding the gradient to \mathbf{w} does not change its direction, and we can have convergence. This means that:

$$\mathbf{w} \propto \left[E \left\{ \mathbf{z} (\mathbf{w}^T \mathbf{z})^3 \right\} - 3 \|\mathbf{w}\|^2 \mathbf{w} \right], \quad (4.27)$$

which leads to the following algorithm for estimating \mathbf{w} (Hyvärinen and Oja, 1997; Hyvärinen et al., 2001):

$$\mathbf{w} \leftarrow E \left\{ \mathbf{z} (\mathbf{w}^T \mathbf{z})^3 \right\} - 3\mathbf{w}. \quad (4.28)$$

The vector \mathbf{w} must be divided by its norm after each iteration. The final value of \mathbf{w} provides one of the independent components via $\mathbf{w}^T \mathbf{z}$ (Hyvärinen and Oja, 1997; Hyvärinen et al., 2001).

This is the basic principle beneath the FastICA algorithm. It must be noticed that the convergence of this algorithm is cubic and it does not depend on adjusting any learning rate (Hyvärinen, 1999; Hyvärinen et al., 2001).

Nonetheless, kurtosis is *not robust* against outliers in the data since its value may depend on only very few observations on the tails of the distributions, which may be irrelevant or even erroneous (Hyvärinen, 1999; Hyvärinen et al., 2001). Hence, other metrics are often used instead of the kurtosis in the fixed-point iteration. These metrics are derived as approximations to negentropy. Actually, they may not be very accurate approximations, but

still they can be used to build a measure of non-gaussianity consistent in the sense that it is always non-negative and zero only for gaussian distributions (Hyvärinen et al., 2001). Under certain circumstances, negentropy, J , may be approximated using a non-quadratic function, $G(y)$, according to:

$$J(y) \propto [E\{G(y)\} - E\{G(\nu)\}]^2, \quad (4.29)$$

where ν denotes a standardised gaussian random variable (Hyvärinen, 1999; Hyvärinen et al., 2001).

Taking $G(y) = y^4$, the previous version of the algorithm is retrieved. However, by selecting $G(y)$ functions that do not grow too fast, it is possible to obtain more robust algorithms. The following choices have proven to be adequate:

$$G_1(y) = \frac{1}{a_1} \log [\cosh(a_1 y)], \quad (4.30)$$

$$G_2(y) = -\exp(-y^2/2), \quad (4.31)$$

where $1 \leq a_1 \leq 2$ is a constant (Hyvärinen, 1999; Hyvärinen et al., 2001). A illustration of these functions is shown in Fig. 4.2. It must be noticed that G_1 and G_2 grow less fast than the kurtosis-based function.

These functions have several advantages. G_1 is a good general-purpose function and G_2 is appropriate for highly supergaussian components or when robustness is very important. On the other hand, kurtosis is justified only for subgaussian components without outliers (Hyvärinen, 1999; Hyvärinen et al., 2001). Moreover, these functions are conceptually simple and fast to compute. Hence, they are used in the FastICA algorithm, which actually includes the derivative of the non-quadratic function $G(\cdot)$.

The algorithm including the aforementioned approximations can be derived by means of an approximate Newton method (Hyvärinen et al., 2001). Nevertheless, this algorithm only estimates one independent component. Thus, some decorrelation scheme must be used to estimate more components since the vectors \mathbf{w}_i are orthogonal in the whitened space (Hyvärinen, 1999; Hyvärinen et al., 2001).

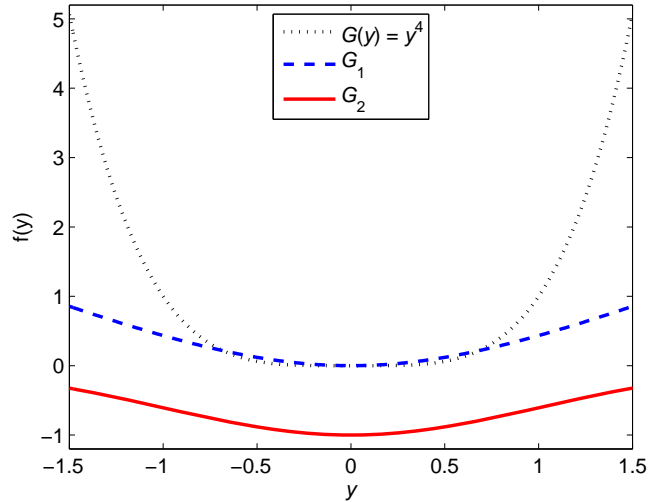


Figure 4.2: The functions G_1 and G_2 compared with the fourth power, as used in kurtosis.

4.5.3 InfoMax

The InfoMax approach implements a *neural network gradient-based algorithm* where the learning rule relies on the information maximization principle (Bell and Sejnowski, 1995). InfoMax maximises the output entropy of a neural network with non-linear outputs. It has been shown that this methodology is equivalent to the maximization of the non-gaussianity of the sources (Hyvärinen et al., 2001; James and Hesse, 2005). Actually, this principle is closely related to the maximum likelihood of the BSS model. From Eq. (4.11), the density $p_{\mathbf{x}}$ of the mixture vector can be formulated as (Hyvärinen et al., 2001):

$$p_{\mathbf{x}}(\mathbf{x}) = |\det \mathbf{B}| p_{\mathbf{s}}(\mathbf{s}) = |\det \mathbf{B}| \prod_i p_i(s_i), \quad (4.32)$$

where $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n]^T$ is the separating matrix.

Denoting by T the total number of samples, the likelihood can be obtained as the product of the density given by Eq. (4.32) evaluated at the T points (Hyvärinen et al., 2001):

$$L(\mathbf{B}) = \prod_{t=1}^T \prod_{i=1}^n p_i(\mathbf{b}_i^T \mathbf{x}(t)) |\det \mathbf{B}|, \quad (4.33)$$

which, after some operations and taking logarithms, leads to (Hyvärinen et al., 2001):

$$\frac{1}{T} \log [L(\mathbf{B})] = E \left\{ \sum_{i=1}^n \log p_i(\mathbf{b}_i^T \mathbf{x}) \right\} + \log |\det \mathbf{B}|. \quad (4.34)$$

However, this model does not only need to estimate \mathbf{B} , but also the densities of the components. This makes the estimation much more complicated, because the estimation of the densities is a non-parametric problem. This difficulty may be avoided by, for example, approximating them by a *family of densities* specified by a limited number of parameters. Actually, only two approximations of the density of an independent component are enough since it is only necessary to decide which of the two approximations is better for each component. This choice is based on the information whether the independent components are super-gaussian or sub-gaussian (Hyvärinen et al., 2001).

The simplest algorithms for maximising likelihood are obtained by gradient methods. The stochastic gradient of the log-likelihood in Eq. (4.34) is (Hyvärinen et al., 2001):

$$\frac{1}{T} \frac{\partial \log L}{\partial \mathbf{B}} = [\mathbf{B}^T]^{-1} + E \{ \mathbf{g}(\mathbf{B}\mathbf{x}) \mathbf{x}^T \}, \quad (4.35)$$

where $\mathbf{g}(y)$ is a component-wise vector function that is composed of the score functions g_i which are defined on the basis of the two families of densities approximating the space of distributions. As a result, the following algorithm is obtained (Bell and Sejnowski, 1995; Hyvärinen et al., 2001):

$$\Delta \mathbf{B} \propto [\mathbf{B}^T]^{-1} + E \{ \mathbf{g}(\mathbf{B}\mathbf{x}) \mathbf{x}^T \}, \quad (4.36)$$

which constitutes the algorithm first derived by Bell and Sejnowski (1995), although the seminal derivation was based on the InfoMax principle. This is based on maximising the output entropy, or information flow, of a neural network with non-linear outputs (Bell and Sejnowski, 1995). If the nonlinearities used in the neural network are chosen as the cumulative distribu-

tion functions corresponding to the densities p_i , the output entropy is equal to the likelihood. This means that the InfoMax approach is equivalent to maximum likelihood estimation (Hyvärinen et al., 2001).

The original InfoMax can only extract super-gaussian sources. For this reason, it was soon modified to include an ‘*extended*’ mode able to estimate sub-gaussian sources (e.g., sinusoids). The resulting technique is the Lee & Sejnowski’s Extended Information Maximisation Algorithm (eInfoMax) by Lee et al. (1999).

4.5.4 JADE

Finally, another usual approach to HOS-BSS is the JADE algorithm (Cardoso and Souloumiac, 1993, 1996), which is sometimes referred as ICA by tensorial methods using higher-order cumulant tensors (Hyvärinen et al., 2001; James and Hesse, 2005). Whereas the covariance matrix is the second-order cumulant tensor, the fourth-order tensor is defined by the *fourth-order cumulants*: $\text{cum}(x_i, x_j, x_k, x_l)$ (Cardoso and Souloumiac, 1993; Hyvärinen et al., 2001; James and Hesse, 2005).

Whereas an eigenvalue decomposition of the second-order cumulant tensor is helpful to decorrelate the variables, making the fourth-order cumulants zero or as close to zero as possible implies statistical independence of the sources (Cardoso and Souloumiac, 1993; James and Hesse, 2005). It is important to note that, if the x_i are independent, all the cumulants with at least two different indices are zero (Hyvärinen et al., 2001). These are the basic principles of the JADE algorithm, which jointly diagonalises a set of matrices (Cardoso and Souloumiac, 1993, 1996).

The JADE algorithm uses certain matrices $\mathbf{Q}_z(\mathbf{M})$ formed by the inner product of the fourth-order cumulant tensor of the outputs with the arbitrary matrix \mathbf{M} (Cardoso and Souloumiac, 1993; Hyvärinen et al., 2001; Sanei and Chambers, 2007):

$$\{\mathbf{Q}_z(\mathbf{M})\}_{ij} = \sum_k \sum_l \text{Cum}(z_i, z_j^*, z_k, z_l^*) m_{lk}, \quad (4.37)$$

where m_{lk} denotes an element of the \mathbf{M} matrix and $*$ represents complex conjugate. $\mathbf{Q}_z(\mathbf{M})$ is diagonalised by the correct rotation matrix \mathbf{U} , which implies that $\mathbf{U}^H \mathbf{Q} \mathbf{U} = \mathbf{\Lambda}_M$. Here, $\mathbf{\Lambda}_M$ is a diagonal matrix whose diagonal elements depend on \mathbf{M} and \mathbf{z} (Sanei and Chambers, 2007). Therefore, applying Eq. (4.37) for a set of different matrices \mathbf{M} , a set of $\mathbf{Q}_z(\mathbf{M})$ can be computed. The rotation matrix \mathbf{U} that approximately jointly diagonalises those cumulants provides the solution to the decomposition of \mathbf{z} (Cardoso and Souloumiac, 1993; Hyvärinen et al., 2001; Sanei and Chambers, 2007).

The problem of this joint diagonalisation can be stated as the minimisation of $\mathcal{J}(\mathbf{u})$ given by Eq. (4.38) (Hyvärinen et al., 2001; Sanei and Chambers, 2007):

$$\mathcal{J}(\mathbf{u}) = \sum_i \sum_j \text{off}(\mathbf{u}^H \mathbf{Q}_{ij} \mathbf{u}), \quad (4.38)$$

where $\text{off}(\cdot)$ is a measure of the importance of the off-diagonal terms as in Eq. (4.20).

A natural choice for this set of matrices is to take the eigenmatrices of the cumulant tensor. In this way, this set of n matrices provide all the relevant information on the cumulants since they span the same subspace as the cumulant tensor (Cardoso and Souloumiac, 1993; Hyvärinen et al., 2001). Finally, it must be noticed that this method may have *computational problems* when applied to high dimensional problems since the memory requirements grow quite rapidly with n (Hyvärinen et al., 2001).

4.6 Constrained BSS Algorithms

The algorithms introduced in Section 4.4 and Section 4.5 are able to decompose the measurements into a set of sources (Cichocki and Amari, 2002; Hyvärinen et al., 2001). FastICA is even able to extract the components one by one and each estimated component is characterised, at least in theory, by having the least gaussian distribution among the set of all possible remaining sources (Hyvärinen, 1999; James and Hesse, 2005). Nevertheless, the algorithms may not be guaranteed to converge to the *global optimum* due to their random initialization and other computational factors (James and Gibson,

2003; Lu and Rajapakse, 2005).

In some particular cases, one may desire to *extract only one particular component* which does not correspond with that of maximum negentropy (James and Hesse, 2005). What is more, if the BSS algorithm extracts the components one at a time errors committed at each step are propagated through to the following estimates (Vigário and Oja, 2008). It must be noticed that reducing the dimension of the data with PCA is not helpful in this situation since the source of interest may be lost (James and Gibson, 2003).

To solve these problems, the source separation may be guided using cBSS techniques so that the first components are close to a particular source (or reference) of interest (Vigário and Oja, 2008). Obviously, this requires to have access to a fair *estimate* of the activity of interest, but there are some cases in neurophysiological analysis where this a priori information does exist (James and Gibson, 2003; Vigário and Oja, 2008). However, if this is the case, the application of a cBSS algorithm removes the need for manual identification and labelling of the components, which may be highly laborious and subjective (James and Gibson, 2003).

Diverse implementations of cBSS have been proposed (Huang and Mi, 2007; Lu and Rajapakse, 2005, 2006). These cBSS algorithms convert the constrained problem to an unconstrained one by introducing a regularisation parameter (e.g., an augmented Lagrange function; Hyvärinen et al., 2001), in the decomposition algorithm (Huang and Mi, 2007; Lu and Rajapakse, 2005; Sanei and Chambers, 2007). In this way, it is possible to obtain a component that is statistically independent of other sources and closest to some reference signal $\mathbf{r}(t)$ (James and Hesse, 2005). This constrain reference signal does not need to be perfectly identical to the source of interest, but it should be similar enough to drive the algorithm into the direction of the desired component (James and Hesse, 2005).

The closeness constrain between the reference, $\mathbf{r}(t)$, and the estimated component, $\mathbf{y}_e(t)$ may be written as (Huang and Mi, 2007; James and Gibson, 2003):

$$g(\mathbf{y}_e, \mathbf{r}) = \epsilon(\mathbf{y}_e, \mathbf{r}) - \xi \leq 0, \quad (4.39)$$

where ξ is some convergence threshold and $\epsilon(\mathbf{y}_e, \mathbf{r})$ measures the closeness between the reference and the current output, which is computed as $\mathbf{y}_e(t) = \mathbf{w}^T \mathbf{x}$ with \mathbf{w} being the de-mixing vector (Huang and Mi, 2007; James and Gibson, 2003; Lu and Rajapakse, 2006). This closeness function may be formulated in several ways, including (Huang and Mi, 2007):

$$\epsilon(\mathbf{y}_e, \mathbf{r}) = E \{ (\mathbf{y}_e - \mathbf{r})^2 \}, \quad (4.40)$$

$$\epsilon(\mathbf{y}_e, \mathbf{r}) = 1 / (E \{ \mathbf{y}_e \mathbf{r} \})^2. \quad (4.41)$$

Thanks to this modelling of the constraint, the cBSS problem can be formulated in the context of Lagrange multipliers (Hyvärinen et al., 2001). This means that the optimisation of any of the cost functions, \mathcal{J} , derived for ICA algorithms — see, for example, Eq. (4.25) — is subject to (James and Gibson, 2003; Lu and Rajapakse, 2005):

$$g(\mathbf{y}_e, \mathbf{r}) \leq 0, \quad (4.42)$$

$$E \{ \mathbf{y}_e^2 \} - 1 = 0 \text{ and} \quad (4.43)$$

$$E \{ \mathbf{r}^2 \} - 1 = 0. \quad (4.44)$$

These conditions are necessary to fix the amplitude of the output and assure its similarity to the reference (James and Gibson, 2003).

4.7 Relationships Among BSS Algorithms

As it has been previously explained, there is a relatively wide variety of BSS algorithms with different theoretical foundations and practical implementations (Cichocki and Amari, 2002; Hyvärinen et al., 2001). This makes important to think about the different suitability of these techniques to analyse the data at hand. The selection of the BSS algorithm to be applied must consider a few aspects.

First of all, one has to decide between SOS and HOS methods. As it has been explained in Section 4.4 and Section 4.5 the assumptions beneath these two types of algorithms differ substantially (Cichocki and Amari, 2002).

Hence, the choice of the criterion should depend on the type of data to be studied, including whether the variables have temporal structure or not and if the inner components are assumed non-gaussian (Hyvärinen et al., 2001).

Afterwards, one has to opt for one particular algorithm. This decision is influenced by certain issues, such as the need to estimate all components in parallel or one by one, the type of metric on which the algorithm is based, and other parameters (Hyvärinen et al., 2001). Obviously, the knowledge of reference signal can be helpful to constrain the extraction (James and Gibson, 2003).

In the following lines, the theoretical relationships among different BSS methods will be briefly reviewed.

Mutual information, $MI(\cdot)$, provides a unifying perspective to compare the fundamentals of diverse HOS techniques. It is a concept derived from entropy which measures both linear and non-linear dependences in data by evaluating the information that members of a set of random variables have on the other random variables in the set (Comon, 1994; Escudero et al., 2008b, 2009a; Hyvärinen et al., 2001). For a set of variables y_i that are uncorrelated and have unit variance, the following straightforward relation between mutual information and negentropy, $J(\cdot)$, holds (Hyvärinen et al., 2001):

$$MI(y_1, y_2, \dots, y_n) = \text{constant} - \sum_i J(y_i). \quad (4.45)$$

Hence, if the variables y_i are the estimated components by an ICA algorithm via $\mathbf{y} = \mathbf{B}\mathbf{x}$, the transformation \mathbf{B} that minimises the mutual information between the outputs also provides the directions in which the negentropy is maximised (Hyvärinen et al., 2001). Since negentropy is a measurement of non-gaussianity, Eq. (4.45) shows that maximising the sum of non-gaussianities of the estimated components is equivalent to minimising their mutual information (Hyvärinen et al., 2001). Moreover, from Eq. (4.34), if the p_i were equal to the real probability density functions of $\mathbf{b}_i^T \mathbf{x}$, the likelihood would be equal, up to an additive constant given by the total entropy \mathbf{x} , to the negative of the mutual information (Hyvärinen et al., 2001).

In summary, mutual information relates to the principles of estimating the

independent components both by means of maximising the non-gaussianity and the likelihood. Hence, it becomes clear that the fundamental ideas beneath FastICA and InfoMax are mostly equivalent (Hyvärinen et al., 2001).

The connection of these principles to cumulant-based criteria is revealed taking into account the approximation of negentropy by cumulants shown by Eq. (4.24), where the first term may be omitted for symmetrical probability distributions (Hyvärinen et al., 2001). Furthermore, considering Eq. (4.24) and Eq. (4.45), the following approximating of $MI(\cdot)$ may be derived:

$$MI(\mathbf{y}) \approx \text{constant} - \text{constant} \sum_i \text{kurt}(y_i)^2, \quad (4.46)$$

which suggests that the maximisation of the cumulants produces the minimisation of mutual information (Hyvärinen et al., 2001). Despite these similarities, it must be borne in mind that there are several differences among the algorithms, including the actual form of the objective functions of the techniques and the need to estimate the sub- or super-gaussian nature of the components (Hyvärinen et al., 2001).

On the other hand, both AMUSE and SOBI use the temporal information of the signal to find the inner components. They diagonalise a set of covariance matrices to minimise the time dependencies among the sources. However, an important difference between those techniques is the number of time lags, τ , considered in the analysis. Whereas only two time delays are considered in AMUSE (usually $\tau = 0$ and $\tau = 1$ sample), SOBI employs several time delays (up to a few tens in some cases; Tang et al., 2005). This makes the algorithm more robust against noise but increases its complexity (Cichocki and Amari, 2002). Actually, the principle of joint diagonalisation (of cumulants or variance matrices) is shared between the SOBI and JADE algorithms (Belouchrani et al., 1997; Cardoso and Souloumiac, 1993).

The theoretical principles of HOS techniques, which relies on concepts of Information Theory, and SOS methods, which employs the temporal structure of the signals can be unified under the perspective of the Kolmogorov complexity (Hyvärinen et al., 2001). The Kolmogorov complexity allows to quantify the structure of the data by measuring the length of the shortest

possible code for it (Hyvärinen et al., 2001). Therefore, this generalisation implies that the BSS can be seen as transformations that *minimise the sum of the coding lengths* of the signals (Hyvärinen et al., 2001). This leads to a principle closely related to the Minimum Description Length (MDL). If the signals have no time structure, their Kolmogorov complexities are given by their entropies. This concept is also related to the time structure of the signals. However, this measure is a rather theoretical and its optimisation is difficult to perform in practice (Hyvärinen et al., 2001).

4.7.1 Consistency of BSS Algorithms for Real Signals

As it was mentioned in Section 2.5.3, both EEG and MEG are the only techniques that measure the synchronous oscillations of the cortex directly and non-invasively (Hari, 2004). Although these recordings have slightly different characteristics, the primary generating currents of the EEG and MEG are the same (Hari, 2004) and similar problems are faced when analysing both recordings. BSS is useful to overcome some of these difficulties (James and Hesse, 2005; Onton et al., 2006; Vigário and Oja, 2008). Hence, these techniques have been widely applied to EEG and MEG data in very diverse applications (James and Hesse, 2005; Onton et al., 2006; Vigário and Oja, 2008).

There is a wide variety of BSS techniques available and not all algorithms are based on the same principles (for a review see, for example, Cichocki and Amari, 2002; Hyvärinen et al., 2001; James and Hesse, 2005). Some theoretical relationships exist among the metrics used to measure the independence of the signals but it might be difficult to a priori select the most appropriate algorithm for a particular application (Fitzgibbon et al., 2007; Zavala Fernández et al., 2006).

A few studies have compared some BSS algorithms (see Klemm et al., 2009, and references therein). However, most of these analyses were based on artificially created signals. For instance, FastICA, JADE and InfoMax were compared by Li et al. (2001). However, some of the basic hypotheses in ICA were violated in the design of the experiments, which can limit the

reliability of the results. Moreover, the analysis focused on acoustic signals, making the extension of the results to brain electromagnetic recordings not straightforward at all (Li et al., 2001).

Computational and statistical comparisons among diverse HOS methods were performed in (Hyvärinen et al., 2001). The results were based again on synthetic signals, which turned out to be super-gaussian. The main conclusions of this study support the robustness of HOS techniques under slight violations of the assumptions and suggest that different techniques may reveal different components when applied to real signals (Hyvärinen et al., 2001).

Other studies have compared diverse BSS algorithms bearing in mind their application to the rejection of artefacts in EEG recordings (Crespo-Garcia et al., 2008; Fitzgibbon et al., 2007; Romero et al., 2008, 2009). The independence of the extracted components was checked by Krishnaveni et al. (2005, 2006) in the removal of OA although the most commonly used algorithms were left out of the analysis. The evaluation was done in terms of mutual information. This fact might bias the analysis in favour of those algorithms directly based on this parameter. Moreover, the significance of the differences among algorithms were not tested (Krishnaveni et al., 2006).

Other analyses have evaluated the performance of diverse BSS algorithms on the terms of the quality of the artefact removal achieved with them (Delorme et al., 2007). Recently, an extensive study oriented to the analysis of EEG data has been published (Klemm et al., 2009). Nevertheless, it was entirely based on synthetic data (Klemm et al., 2009). Even more recently, the outputs of FastICA, JADE and SOBI were compared against a new approach to BSS based on the short-time Fourier transform of the signals (Hyvärinen et al., 2010). This study suggests that, in the case of spontaneous activity, HOS methods tend to focus on the extraction of artefacts whereas SOBI failed since it tended to extract components with very similar spectra (Hyvärinen et al., 2010). However, this analysis was mainly carried out in specific the framework of the study of the phase differences between components with data from only one subject (Hyvärinen et al., 2010).

To sum up, most comparisons among different BSS algorithms were car-

ried out with only simulated signals or in very particular contexts, such as artefact removal. Those studies tried to assess the quality of the brain recordings after a BSS-based artefact removal procedure or the quality of the artefactual signals isolated by BSS (Fitzgibbon et al., 2007; Romero et al., 2008, 2009; Zavala Fernández et al., 2006). This may limit the applicability of the results to other settings and a detailed *study on the similarity* of the decompositions for real brain recordings computed with different algorithms is *lacking* (Onton et al., 2006).

It must be noticed that these methods are data-driven and, by their own nature, exploratory (Vigário and Oja, 2008). If the generating process is known, an objective evaluation is possible. However, this is not the case in almost all cases, hence the term *blind* (Vigário and Oja, 2008).

Therefore, it is important to study the consistency (i.e., similarity) of the separations estimated from real electromagnetic recordings so that the relationships among algorithms applied to the same real brain data can be further understood. This could help to decide a priori which particular algorithms provide more similar results, thus limiting the search for appropriate techniques for each case (Escudero et al., 2009d). To achieve this goal, in this Doctoral Thesis MEG background activity will be blindly decomposed using five of the most common used BSS algorithms in the analysis of electromagnetic brain background activity: AMUSE, SOBI, JADE, eInfoMax and FastICA.

4.8 BSS Applications to EEG and MEG Activity

Although EEG and MEG have slightly different characteristics, similar problems are faced when analysing both recordings. BSS is useful to overcome some of these difficulties (James and Hesse, 2005; Onton et al., 2006; Vigário and Oja, 2008). Thus, BSS has been widely applied to EEG and MEG data in very diverse applications (James and Hesse, 2005; Onton et al., 2006; Vigário and Oja, 2008).

For instance, the signals acquired at a particular sensor are a weighted linear mixture of the underlying brain activity (Onton et al., 2006). Therefore, the isolation and analysis of the electromagnetic activity generated by a specific source of interest is a complex task in which the BSS methods may be helpful (Jarchi et al., 2009; Jin et al., 2002; Klemm et al., 2009; Onton et al., 2006). Moreover, the brain activity is usually recorded together with undesired signals (i.e., artefacts) of physiological or environmental origin (James and Hesse, 2005; Onton et al., 2006). BSS methods have been used to detect and remove artefactual activity from brain signals (Delorme et al., 2007; Escudero et al., 2007b; Fitzgibbon et al., 2007). BSS can also help in the identification of diverse brain diseases (Cichocki et al., 2005; Escudero et al., 2008d, 2009b).

Thus, this Section reviews the state of the art in the applications of BSS techniques to EEG and MEG signals that are related to the studies carried out in this Doctoral Thesis.

4.8.1 Artefact Removal

MEG captures the neural activity with high spatial resolution by measuring the brain magnetic fields. This technique is not invasive and it does not depend on any reference point (Hari, 2004). Moreover, magnetic fields are less distorted than electric ones by the skull and the scalp (Hari, 2004). However, MEG data must be recorded in magnetically shielded rooms with SQUID sensors to reduce external noise (Hari, 2004).

As it was explained in Section 2.4, external noise is not the only undesired signal in MEG data. In these recordings, non-cerebral sources (i.e., artefacts) always appear mixed with brain signals (Hyvärinen et al., 2001). The artefacts could bias the analyses, since their power may be larger than that of the brain sources (James and Hesse, 2005). For instance, the CA is usually noticeable in MEG data (Jousmäki and Hari, 1996). OAs can also be visible in these signals (Antervo et al., 1985). Although they can be partially controlled by the subject in short data epochs, these artefacts are likelier to appear in long recordings or when the MEG is recorded from non-

collaborative subjects. In addition to these artefacts, MEG data may have strong PLN (Hari, 2004). Furthermore, some authors have even claimed that any inner source without time structure should be removed, since it provides no information about the brain activity (Vorobyov and Cichocki, 2002).

Several methods have been used to remove artefacts from EEG and MEG data: epoch rejection, regression techniques (Croft and Barry, 2000), PCA (Sadasivan and Dutt, 1996) or BSS (Cichocki and Amari, 2002; James and Hesse, 2005; Vigário and Oja, 2008). The main characteristics of each of these methodologies will be introduced in the following lines.

Epoch Rejection

The simplest method to avoid artefacts in EEG and MEG data is epoch rejection, which discards raw data epochs highly contaminated by artefacts. Usually, the assessment is performed individually for each signal epoch, which may be very time-consuming (Fatourehchi et al., 2007). Nevertheless, automatic techniques may be used. These are often based on detecting fragments of the signal whose amplitude is higher than a certain threshold. This approach is useful to perform a preliminary screening of the data (Osipova et al., 2003, 2005, 2006a,b; Pekkonen et al., 1999, 2005).

However, epoch rejection has some major drawbacks. First of all, it may produce significant *data loss* (Castellanos and Makarov, 2006; Fatourehchi et al., 2007; Jung et al., 2000; Vigário and Oja, 2000). Additionally, both automatic and manual evaluations of the epochs are intrinsically *subjective* (Castellanos and Makarov, 2006; Fatourehchi et al., 2007). Finally, in some applications, it may cause problems associated with the stationarity of the recordings as this technique produces *discontinuous segments* of data (Castellanos and Makarov, 2006) or it may bias the analyses by removing activity related to the rejected artefacts (Fatourehchi et al., 2007; Vigário, 1997).

Regression Techniques

Other method to remove artefacts from MEG and, specially, EEG signals is removing the projection of a reference signal from the brain recordings

(Fatourehchi et al., 2007; Jung et al., 2000; Samonas et al., 1997; Vigário and Oja, 2000). This technique has been widely applied to remove the OA from EEG recordings (Croft and Barry, 2000; Schlögl et al., 2007). The projection of a reference signal — e.g., Electrooculogram (EOG) — on the signals of interest is computed either in the time or frequency domain. Then, this projection is subtracted from the brain recordings to obtain artefact-free signals (Croft and Barry, 2000; Schlögl et al., 2007). Similarly, adaptive filtering may be used to remove the PLN (Wesson et al., 2009).

Nevertheless, this technique may introduce new unexpected artefacts in the brain recordings (Jung et al., 2000; Vigário and Oja, 2000). This is due to the fact that the reference signal, which is most often the EOG, may be ‘contaminated’ with brain activity (i.e., *bidirectional contamination*), which would be removed from the EEG or MEG by this procedure (Fatourehchi et al., 2007; Vigário, 1997). Moreover, this technique needs to record the reference signal and the brain data simultaneously, which may not always be possible (Fatourehchi et al., 2007; Jung et al., 2000; Li et al., 2006). Hence, there is some evidence that the artefact rejection based on BSS is superior to regression techniques (Hoffmann and Falkenstein, 2008; Romero et al., 2008, 2009), although some studies have claimed the opposite (Schloegl et al., 2009; Wallstrom et al., 2004).

Principal Component Analysis

In contrast, PCA can be applied to reject any kind of artefact without reference signals (Fatourehchi et al., 2007; Sadasivan and Dutt, 1996). This technique finds orthogonal directions of greatest variance in data. Principal components are uncorrelated with each other, but not really statistically independent (Hyvärinen et al., 2001).

PCA can detach artefacts from brain signals completely only when they are orthogonal to each other, their amplitudes are dissimilar, and the additive noise power is low enough. These *hypotheses are difficult to fulfil* for EEG and MEG data (Fatourehchi et al., 2007; Jung et al., 2000; Shoker et al., 2005b; Vorobyov and Cichocki, 2002). Nevertheless, other authors have suggested

that PCA might reject artefacts in EEG signals successfully and that, in some cases, it might even be a better approach than some BSS techniques (Kierkels et al., 2006; Wallstrom et al., 2004).

Blind Source Separation

BSS (Cichocki and Amari, 2002) has been used in the artefact rejection problem (Delorme et al., 2007; Escudero et al., 2007b; Jung et al., 2000; Sander et al., 2002; Ting et al., 2006; Vigário and Oja, 2008; Vorobyov and Cichocki, 2002). BSS needs neither previous information nor orthogonality between artefacts and brain signals (James and Hesse, 2005; Vigário and Oja, 2008).

The assumptions made about the data by BSS (or ICA) seem to be suitable for MEG recordings (James and Hesse, 2005; Vigário and Oja, 2008). This kind of procedure needs neither an accurate model of the process that generated the artefacts nor indicating signal intervals that mainly contain artefacts nor additional inputs (Hyvärinen et al., 2001). However, the use of reference signals or any other a priori information may help to extract the artefacts from the brain recordings by means of a cBSS (James and Gibson, 2003; Nazarpour et al., 2008; Shoker et al., 2005b).

A major problem in this artefact rejection method is the *recognition* of the undesired activity. Despite the fact that BSS can isolate artefactual activity into one or a few components (James and Hesse, 2005), the manual identification of the contaminating signals is a very time-consuming process. This was the approach taken in the first artefact removal studies based on BSS (Iriarte et al., 2003; Jung et al., 2000; Vigário, 1997; Vigário and Oja, 2000).

Afterwards, several studies have introduced *automatic or semi-automatic criteria* to mark diverse artefacts in order to simplify this process. For instance, the use of statistical moments and metrics (such as skewness, kurtosis and entropy) has been proposed to identify artefacts in EEG and MEG recordings (Barbati et al., 2004; Dammers et al., 2008; Delorme et al., 2007; Escudero et al., 2006c,d, 2007b, 2008c; LeVan et al., 2006; Mammone and

Morabito, 2008; Mantini et al., 2008; Rong and Contreras-Vidal, 2006; Shoker et al., 2005a).

In addition to being used in a cBSS setting (James and Gibson, 2003; Shoker et al., 2005b), reference signals like the EOG can help to identify the components corresponding to artefacts. The most common approach is to mark as artefacts the components whose correlation with the reference is above a particular threshold (Barbati et al., 2004; Dammers et al., 2008; Flexer et al., 2005; Joyce et al., 2004; Shoker et al., 2005a) although approaches based on mutual information have also been used (Nicolaou and Nasuto, 2007).

Other approaches have been suggested. For instance, the BSS components can be sorted depending on their time structure (Vorobyov and Cichocki, 2002). Additionally, the scalp topography was used to remove OAs from data since most of the ocular activity is located near the eyes (Campos Viola et al., 2009; Escudero et al., 2007b; Li et al., 2006; Okada et al., 2007; Romero et al., 2008; Rong and Contreras-Vidal, 2006). Amplitude thresholds have also been applied to the components to decide which sources are contaminated by artefacts such as the ocular blinks (Castellanos and Makarov, 2006; Delorme et al., 2007; Ting et al., 2006). In some cases, such as OAs and PLN, spectral features of the components may also be helpful in the identification of the artefacts (Joyce et al., 2004; LeVan et al., 2006; Nicolaou and Nasuto, 2007; Romero et al., 2008).

Finally, it is important to note that some of the aforementioned criteria have been combined to detect various types of artefacts (Barbati et al., 2004; Escudero et al., 2007b; Joyce et al., 2004; LeVan et al., 2006; Romero et al., 2008; Shao et al., 2008).

Of note is that the BSS-based artefact rejection cleans the data in such a way that the classification of different pathological groups is less affected by those undesired sources, this being improved (Melissant et al., 2005; Vialatte et al., 2009). However, in these studies, the selection of the rejected components was done manually (Melissant et al., 2005; Vialatte et al., 2009). Moreover, it has been recently shown that this type of BSS artefact removal does not denature the electromagnetic brain recordings. By reducing the

amount of artefacts in the data, the subjects' data included in the analysis became closer to the corresponding group's centres (Vialatte et al., 2009). A more consistent topographical distribution of the EEG power was also found after the artefact rejection (Vialatte et al., 2009). All these findings suggest that the *artefact rejection based on BSS does not denature the signals*. Instead, it improves their quality (Vialatte et al., 2009).

Another open issue is how to select the appropriate number of BSS components (n) when a large number of channels (m) are available. Traditionally, few statistical criteria have been used to estimate this parameter (Cao et al., 2003; Ikeda and Toyama, 2000). Some studies have set $m = n$ (Castellanos and Makarov, 2006; Flexer et al., 2005; Joyce et al., 2004; Li et al., 2006; Romero et al., 2008; Sander et al., 2002). Nevertheless, as previously explained, a dimensionality reduction can be needed to avoid 'overfitting', which can lead to the extraction of components without physiological meaning (James and Hesse, 2005; Vigário and Oja, 2008). Hence, some authors have set a threshold on the spectrum of eigenvalues of the data covariance matrix (Barbati et al., 2004; Ting et al., 2006). Nevertheless, this approach is subjective and statistical criteria may be more suitable (Escudero et al., 2007b; Hesse, 2007, 2008; Ikeda and Toyama, 2000).

Despite the relatively large number of metrics proposed to detect artefactual activity in EEG and MEG signals (Barbati et al., 2004; Castellanos and Makarov, 2006; Dammers et al., 2008; Delorme et al., 2007; Escudero et al., 2006c,d, 2007b, 2008c; Joyce et al., 2004; LeVan et al., 2006; Li et al., 2006; Mantini et al., 2008; Okada et al., 2007; Romero et al., 2008; Shao et al., 2008; Ting et al., 2006), few comparative studies are available in the literature. The evaluation of the artefact rejection has quite often been made on subjective bases and few quantitative measurements have been used to objectively quantify the amount of artefacts in the recordings before and after the methodology. Furthermore, most of these efforts have been devoted to the ocular artefact in EEG signals (Castellanos and Makarov, 2006; Fitzgibbon et al., 2007; Joyce et al., 2004; Romero et al., 2008, 2009; Shao et al., 2008; Ting et al., 2006), with the MEG recordings receiving less attention (Dammers et al., 2008; Escudero et al., 2007b; Mantini et al., 2008). In order to clarify

the state of the art in the artefact removal from MEGs with BSS, Table 4.1 shows the most relevant studies on this topic.

Therefore, this Doctoral Thesis seeks to apply diverse straightforward criteria for the detection of CA, OAs and PLN and to compare their performance in the artefact rejection for MEG background activity. Additionally, this Doctoral Thesis also aims at developing a robust dimensionality reduction scheme able to estimate the number of components to be extracted.

4.8.2 Feature Enhancement With BSS

EEG and MEG have been analysed with several signal processing techniques to gain insight into AD (Hornero et al., 2009b; Jeong, 2004; Rossini et al., 2007; Stam, 2005, 2010). For instance, spectral features have been used to quantify the abnormalities in the frequency distribution of AD patients' EEG and MEG signals (Hornero et al., 2008; Jeong, 2004; Stam, 2010). Additionally, non-linear analysis methods can offer useful information about the brain dynamics in this dementia (Escudero et al., 2006a; Hornero et al., 2009b; Stam, 2005). Nevertheless, it is desirable to develop *novel strategies to help in AD detection* from the analysis of the electromagnetic brain activity (Cichocki et al., 2005; Henderson et al., 2006; Woon et al., 2007).

Techniques based on spatial filtering can help in the recognition of particular brain states, such as dementia. This is due to the fact that these algorithms offer additional perspectives to examine EEG and MEG signals (Cichocki et al., 2005; Escudero et al., 2008d; Woon et al., 2007). For instance, Common Spatial Patterns (CSP) have been applied to enhance characteristics of EEG recordings in MCI patients who eventually developed AD (Woon et al., 2007). CSP finds spatial filters which maximise the difference in signal power between two classes to be discriminated (e.g., patients and controls) (Woon et al., 2007).

Obviously, another type of spatial filtering techniques is BSS (Cichocki and Amari, 2002; Hyvärinen et al., 2001; James and Hesse, 2005). BSS methods estimate the underlying components of the EEG and MEG signals without a priori information about those components (Cichocki and Amari,

Table 4.1: Summary of artefact detection studies in MEG with BSS.

Study	Database	Artefact	Approach	Type of evaluation
Sander et al. (2002)	1200 epochs from 1 subject.	CA.	Study of the CA components extracted with a SOS-BSS technique in a visual stimulation task.	Study on the preservation of the evoked signal with geometrical and temporal measures.
James and Gibson (2003)	3 epochs from 3 subjects.	CA, OA and PLN.	cBSS scheme.	Visualisation of single examples.
Barbati et al. (2004)	3 epochs from 3 subjects.	CA, OA and PLN.	HOS-BSS technique with metrics based on kurtosis, entropy and correlation.	Visual inspection of the difference between the original and the processed MEGs.
Rong and Contreras-Vidal (2006)	Up to 123 trials from 5 subjects.	Muscular, CA and OA.	Categorisation of InfoMax sources with a clustering and an artificial neural network on the basis of HOS and comparison with templates.	Comparison of the variance at different subsets of channels before and after the procedure.
Escudero et al. (2007b)	66 epochs from 6 subjects.	CA, OA and PLN.	HOS-BSS technique with Skew, KrE, P _{EVES} , P _{LF} and P _{50Hz} .	Comparison of the artefacts' power in the whole dataset.
Okada et al. (2007)	2100 trials from 7 subjects.	OA.	Identification of ocular InfoMax components by their scalp topography.	Visual examples.
Mantini et al. (2008)	12 subjects.	Muscular, CA and OA.	Detection of artefacts in FastICA components with ApEn.	Power of the QRS complex, number of peaks, noise in the prestimulus interval and performance of a dipole localisation.
Dammers et al. (2008)	10 experiments from 6 subjects.	Muscular, CA and OA.	Detection of artefacts in InfoMax components with HOS, correlation and phase statistics.	Power in the QRS complex and around eye blinks.

2002; James and Hesse, 2005; Vigário and Oja, 2008). Since these techniques isolate specific physiological activities into different components, they have been used to reject artefacts, as it was explained in Section 4.8.1 (Escudero et al., 2007b; James and Hesse, 2005; Vigário and Oja, 2008).

EEG and MEG data can also be processed with BSS methods to help in the recognition of neurological disorders and states. For example, BSS can help to separate specific brain activity related to epilepsy or Creutzfeldt-Jakob Disease (CJD) as shown by Jarchi et al. (2009); Kobayashi et al. (1999) and Hung et al. (2007), respectively. Considering these research studies and the intrinsic complexity of brain recordings, it can be hypothesised that the application of BSS, together with features extracted from electromagnetic brain activity recordings, may *enhance features associated with particular brain states* such as AD (Cichocki et al., 2005) or audio-visual stimulation (Jin et al., 2002). This is due to the fact that some BSS components of the EEG and MEG signals may be more sensitive to those brain states than others (Cichocki et al., 2005; Escudero et al., 2008d, 2009b; Vialatte et al., 2005).

Therefore, the most relevant components may be selected and the brain recordings may be partially reconstructed using only those components (Cichocki et al., 2005). As a consequence, diverse brain states (e.g., AD) might be better differentiated in comparison to the situation where BSS is not used.

Nevertheless, despite the advantages that this approach may provide, few studies have implemented it (Escudero et al., 2008d, 2009b). For instance, a subset of five AMUSE components extracted from 21 EEG channels were retained to improve the accuracy of the classification of 22 MCI patients who later proceeded to AD against 38 control subjects up to of 10% when six spectral features of the EEG were used in the classification (Cichocki et al., 2005). A later study used the partially reconstructed EEGs obtained by Cichocki et al. (2005) to further improve the classification of the MCI patients (Vialatte et al., 2005). This subsequent study achieved an additional rise of about 13% over the accuracy reported by Cichocki et al. (2005) using a ‘bump modelling’ of the WT computed from the EEGs and a neural network classifier (Vialatte et al., 2005). A summary of relevant studies on this

application of spatial filtering techniques is displayed in Table 4.2.

It must be borne in mind that the approach taken in this part of the Doctoral Thesis is different from the artefact rejection framework. Now, we did *not* aim to remove artefacts from MEG recordings with *minimal* brain activity distortion. Instead, we attempted to *emphasise the differences* between AD patients and control subjects' MEG recordings by retaining only the components which account for the most relevant differences between groups. Of course, this will produce some distortion in the signals, but the resulting data should increase the separability between the subject groups. That is, the partially reconstructed MEG recordings do not reflect the brain activity accurately as in [Melissant et al. \(2005\)](#) and [Vialatte et al. \(2009\)](#). Instead, they should provide a more efficient separation between both subject groups than the MEG data this BSS procedure ([Escudero et al., 2006e, 2007c,d, 2008d, 2009b,c](#)).

In this Doctoral Thesis, we applied diverse BSS algorithms (AMUSE, SOBI, eInfoMax and FastICA) to background MEG recordings from AD patients and control subjects. Two spectral, MF and SpecEn, and two non-linear variables, LZC and SampEn ([Escudero et al., 2008a](#)), were used to characterise every MEG channel and BSS component. The components of both groups will be analysed to assess how those four variables varied within them and to decide which provided the most relevant information to classify the subjects. In addition, the separation between AD patients and elderly control subjects achieved with and without the BSS preprocessing will be compared. Additionally, we also aim at determining whether the subsets of BSS components with significant differences between demented patients and controls notably differ for the four variables considered.

4.8.3 Extraction of Brain Rhythms With BSS

It is well known that AD causes the electromagnetic brain activity to slow down. This dementia is related to an increase of power in the low-frequency range (δ and θ) and a decrease of power in the high-frequency range (α and β) ([Jeong, 2004](#); [Stam, 2010](#)). Abnormalities in the non-linear dynamics of

Table 4.2: Summary of studies about the preprocessing of EEG and MEG data with spatial filtering techniques to increase the separation between subject groups.

Study	Subjects	Signal	Approach
Cichocki et al. (2005)	22 MCI patients who later proceeded to AD and 38 controls.	EEG	Selection of 5 AMUSE components and classification with 6 spectral band powers.
Vialatte et al. (2005)	22 MCI patients who later proceeded to AD and 38 controls.	EEG	Selection of 5 AMUSE components and classification with WT features and a neural network.
Woon et al. (2007)	22 MCI patients who later proceeded to AD and 38 controls.	EEG	CSP as a preprocessing with spectral powers and SampEn.
Escudero et al. (2008d)	21 AD patients and 21 controls.	MEG	Selection of AMUSE components and classification with MF and SpecEn.
Escudero et al. (2009b)	36 AD patients and 26 controls.	MEG	Selection of AMUSE components and classification with MF and SpecEn (spectral features) and LZC and SampEn (non-linear features).
Escudero et al. (2009c)	18 AD patients and 13 controls.	MEG	Selection of AMUSE, SOBI and eInfoMax components and classification with MF.

the recordings also appear (Hornero et al., 2009b; Stam, 2005). Additionally, the brain activity of AD patients tend to be less coherent than in control patients (Jeong, 2004). Thus, several neuropathological, electrophysiological and neuroimaging studies have suggested that AD may also be characterised by a ‘disconnection syndrome’ (Delbeuck et al., 2003; Jeong, 2004). However, these effects are not always easily detectable (Dauwels et al., 2010). Furthermore, there is some consensus that a *lower synchronization* level in α and β bands is associated with AD but contradictory results have been found in the other frequency bands (Jeong, 2004).

To assess the ‘disconnection syndrome’, different connectivity measures have been applied to the EEG and MEG signals. Nevertheless, it is important to note that the MEG recordings may be more suitable than the EEG signals to evaluate the brain connectivity for the following reasons (Gómez et al., 2009c; Hämäläinen et al., 1993; Hari, 2004):

- EEG signals are strongly influenced by methodological issues such as sensor placement.
- MEG has higher spatial resolution than the EEG.
- MEG provides reference-free recordings, which are not distorted by the extra-cerebral tissues.

Bearing this in mind, it should be noticed that several approaches are available to study the connectivity of the brain recordings (Pereda et al., 2005). Spectral Coherence ($\text{Coh}(f)$) is the most widely used measure to assess connectivity. It is a normalised linear measure of the correlation between two signals as a function of frequency. It is computed as the square of the cross-spectrum of the two signals divided by the product of the power spectra of each individual recording (Jeong, 2004; Nunez et al., 1997, 1999; Pereda et al., 2005). Diverse studies have found a decrease in the $\text{Coh}(f)$ of AD patients’ brain activity in α and β bands (Jeong, 2004).

Due to the fact that $\text{Coh}(f)$ is a linear measure, it may fail to detect non-linear interdependencies in the data. To overcome this limitation, several other non-linear connectivity measures have been proposed (Pereda et al.,

2005). However, it is not clear that all of them are superior to $\text{Coh}(f)$ to distinguish AD or MCI (Dauwels et al., 2010; Gómez et al., 2009c).

Nevertheless, it is important to note that the ‘*domain where the brain activity is analysed*’ may be, at least, as important as the actual connectivity measure applied to the data. The most straightforward approach is to analyse the signals acquired at each particular EEG or MEG channel (Dauwels et al., 2010; Gómez et al., 2009c; Osipova et al., 2003). In other cases, the equivalent current dipoles are localised and analysed (Rossini et al., 2007; Stam, 2010). Nevertheless, other approaches are possible. Instead of analysing all the numerous channels recorded or trying to localise the equivalent current dipoles, one may think of characterising the rhythmic brain activities from diverse scalp regions out of the the whole set of EEG or MEG channels (Stam et al., 2006). This approach allows to study the connectivity relationships between different sets of channels, or regions (Stam et al., 2006). In order to do so, a channel-wise procedure may be undertaken. In this way, the analysis is carried out channel by channel, and the the results from different sets of channels are grouped together (Stam et al., 2006). Instead, a useful alternative may be to apply BSS techniques to *extract the activity that characterises a particular scalp region and spectral band*.

To be more precise, cBSS can extract the physiological activity of interest without having to compute the whole decomposition of the brain recordings (James and Gibson, 2003; James and Hesse, 2005). Thus, it could be used to extract the typical brain rhythms (δ , θ , α and β activity) from the set channels located together over a specific region.

It must be noted that a suitable reference is needed to perform the cBSS (Huang and Mi, 2007; James and Gibson, 2003; Lu and Rajapakse, 2005). If the objective is to extract activity related to the well-known δ , θ , α and β brain rhythms, one possibility is to constrain the extraction to be similar to noise filtered in the band of interest (James and Hesse, 2005). However, this approach is not adaptive as it needs to fix bandpass cut-off frequencies *a priori*. Hence, other approaches should be employed instead. One alternative is the EMD recently introduced by Huang et al. (1998).

The EMD allows to decompose any complicated signal into a finite, and

often small, number of IMFs that can be well analysed by means of the Hilbert spectrum (Huang et al., 1998). It is important to note that, in contrast to the BSS techniques (Cichocki and Amari, 2002; Hyvärinen et al., 2001), the EMD provides a scale decomposition of one uni-dimensional signal. In this sense, it is relatively similar to, for example, the WT (Huang et al., 1998). The main difference between them is that the EMD only uses the information available in the data, thus being completely adaptive. On the other hand, the WT employs a set of pre-fixed filters based on the selection of the mother wavelet (Andrade et al., 2006). The EMD is highly adaptive and, therefore, efficient. Since it is based on the local characteristic time-scale of the data, it is applicable to complex signals, such as biomedical recordings (Andrade et al., 2006; Blanco-Velasco et al., 2008; Wu and Hu, 2006; Ye et al., 2007; Zhang et al., 2008). Moreover, the IMFs provide accurate information about the frequency content of the signal (Huang et al., 1998) and the phase-locking between them can help to evaluate brain synchronisation (Sweeney-Reed and Nasuto, 2009).

Therefore, the rhythmic activity obtained from the adaptive EMD (Huang et al., 1998) of the MEGs could be used to constrain a BSS applied to the recordings themselves. Hence, the cBSS could be able to extract brain rhythmic activity belonging to the spectral bands of interest (δ , θ , α and β) from diverse scalp regions. Afterwards, the activity extracted thanks to the cBSS can be studied using connectivity measures, such as $\text{Coh}(f)$, to study the connectivity between scalp regions for each spectral band in AD.

Chapter 5

MEG Recordings

Contents

5.1	Subjects' Sociodemographic Data	130
5.2	MEG Recording and Preprocessing	131
5.3	Epoch Selection	135

This Chapter describes the sociodemographic data of the 36 patients with a diagnosis of probable AD and 26 healthy elderly control subjects who participated in the studies of this Doctoral Thesis. The diagnostic criteria are also defined. Additionally, the MEG recording process is described and the basic preprocessing of the recordings is explained.

5.1 Subjects' Sociodemographic Data

MEG background activity signals recorded from 62 subjects were analysed in this Doctoral Thesis. The AD group consisted of 36 patients with probable AD — 12 male and 24 female; age = 74.06 ± 6.95 years; mean \pm Standard Deviation (SD). All of them fulfilled the criteria of probable AD according to the guidelines provided by the NINCDS-ADRDA (McKhann et al., 1984) and were recruited from the “Asociación de Familiares de Enfermos de Alzheimer” in Madrid. To diagnose the dementia, brain scans (SPECT and MRI) and thorough medical, physical, neurological, psychiatric and neurophysiologic examinations were performed. Finally, MMSE and GDA/FAST tests (Folstein et al., 1975; Reisberg, 1988) were used to evaluate the cognitive impairment in this group.

The Spanish version of the MMSE (Lobo et al., 1979) was applied. The MMSE is a simple test to assess the patient's global cognitive state (Folstein et al., 1975). On the other hand, the patients' functional impairment was evaluated with the GDA/FAST (Reisberg, 1988). The GDA/FAST consists of a scale of seven levels that allows to evaluate the progressive impairment of mental functions, from normal to severe deterioration associated with AD. The average scores for these tests were 18.06 ± 3.36 and 4.17 ± 0.45 for MMSE and GDA/FAST, respectively. None of them was receiving medication that could affect the MEG activity.

The diagnosis of the AD was complemented with:

- Neuropsychological evaluation. The Spanish versions of the following tests were applied: Wechsler Memory Scale 3rd Edition, Boston Naming Test, Stroop Test, Wisconsin Card Sorting Test and Silhouettes

Test of the Visual Object and Space Battery.

- Neuroimaging techniques: SPECT and MRI.
- Blood tests. These include thyroid-function, serum vitamin B12, folic acid, creatine, glucose, phosphorus and potassium tests. Proteino-graphic, bilirubin, alkaline phosphatase, blood plasma transaminase and serum tumour marker studies were also carried out.

Additionally, 26 healthy elderly subjects formed the control group (9 males and 17 females; age = 71.77 ± 6.38 years; MMSE = 28.88 ± 1.18 ; GDA/FAST = 1.73 ± 0.45). There was no significant difference between the mean age of both subject groups ($p = 0.1911 > 0.05$). Informed consent was obtained from all control subjects and AD patients' caregivers for participation in the study, which was approved by the local ethics committee. Table 5.1 and Table 5.2 show the sociodemographic data for the AD patients and control subjects included in this Doctoral Thesis, respectively.

In order to objectively evaluate the improvement in the classification of AD patients versus control subjects' MEG data achieved with the BSS preprocessing (see Section 6.6), the population was divided randomly into a training set (18 AD patients and 13 control subjects) and a test set (formed by other 18 demented patients and 13 controls). The training set was used to develop the BSS preprocessing and to find the classification rules for each case. Then, these algorithms were applied, without further modification, to the test set to independently assess the improvement in the separation between AD patients and control subjects due to the BSS preprocessing. The demographic data and clinical features of training and test sets are also summarized in Table 5.3.

5.2 MEG Recording and Preprocessing

The MEG signals studied in this Doctoral Thesis were recorded using a 148-channel whole-head magnetometer (MAGNES 2500WH, 4D Neuroimaging) placed in a magnetically shielded room of the "Centro de Magnetoencefalo-

Table 5.1: AD patients' sociodemographic data.

Identifier	Gender	Age	MMSE	FAST
Alz-01	Female	71	15	4
Alz-02	Male	67	12	4
Alz-03	Female	56	14	4
Alz-04	Female	64	15	4
Alz-05	Female	59	20	4
Alz-06	Male	60	16	4
Alz-07	Female	72	15	4
Alz-08	Female	71	15	4
Alz-09	Female	75	22	4
Alz-10	Female	82	21	4
Alz-11	Female	72	17	4
Alz-12	Male	80	24	4
Alz-13	Male	83	10	5
Alz-14	Female	77	21	4
Alz-15	Male	82	19	4
Alz-16	Female	83	20	4
Alz-17	Female	73	23	4
Alz-18	Male	79	19	4
Alz-19	Male	83	16	4
Alz-20	Female	72	23	3
Alz-21	Female	69	16	5
Alz-22	Male	77	21	4
Alz-23	Female	74	16	5
Alz-24	Female	81	21	5
Alz-25	Female	81	17	4
Alz-26	Female	78	15	5
Alz-27	Male	68	21	4
Alz-28	Female	78	15	5
Alz-29	Female	72	22	4
Alz-30	Male	79	15	5
Alz-31	Female	78	18	4
Alz-32	Male	71	20	4
Alz-33	Male	78	18	4
Alz-34	Female	75	16	4
Alz-35	Female	78	21	4
Alz-36	Female	68	21	4

Table 5.2: Control subjects' sociodemographic data.

Identifier	Gender	Age	MMSE	FAST
Con-01	Female	68	30	2
Con-02	Female	61	29	2
Con-03	Female	70	30	2
Con-04	Female	64	30	1
Con-05	Male	60	30	1
Con-06	Female	63	30	1
Con-07	Male	73	29	1
Con-08	Female	69	29	1
Con-09	Male	79	29	2
Con-10	Male	79	30	2
Con-11	Female	75	29	2
Con-12	Male	67	29	2
Con-13	Female	68	29	2
Con-14	Male	84	29	2
Con-15	Female	68	27	2
Con-16	Male	73	30	2
Con-17	Female	71	29	1
Con-18	Male	74	30	2
Con-19	Male	78	27	2
Con-20	Female	76	29	2
Con-21	Female	83	26	2
Con-22	Female	68	28	2
Con-23	Female	68	30	1
Con-24	Female	72	27	2
Con-25	Female	77	29	2
Con-26	Female	78	27	2

grafía Dr. Pérez-Modrego” at the Complutense University of Madrid. In order to minimise the amount of artefacts occurring in the MEG background activity, the recordings were acquired while the subjects were lying on a patient bed, awake and in a relaxed state with their eyes closed.

The sampling frequency of the recording process was 678.17 Hz. A hardware band-pass filter with cut-off frequencies at 0.1 Hz and 200 Hz was applied. Additionally, a notch filter at 50 Hz was used to reduce the power level of the electrical power interference. Using this setting, five minutes of MEG background activity were acquired for each subject. Afterwards, the record-

Table 5.3: Demographic and clinical features for all participants, and for the training and test sets. Data are given as mean \pm SD.

All subjects		
	AD patients	Control Subjects
Number of subjects	36	26
Number of females	24	17
Age	74.06 \pm 6.95	71.77 \pm 6.38
MMSE	18.06 \pm 3.36	28.88 \pm 1.18
FAST	4.17 \pm 0.45	1.73 \pm 0.45
Training set		
	AD patients	Control Subjects
Number of subjects	18	13
Number of females	12	9
Age	74.11 \pm 7.38	71.38 \pm 4.84
MMSE	17.72 \pm 3.63	28.92 \pm 1.04
FAST	4.17 \pm 0.38	1.85 \pm 0.38
Test set		
	AD patients	Control Subjects
Number of subjects	18	13
Number of females	12	8
Age	74.00 \pm 6.70	72.15 \pm 7.82
MMSE	18.39 \pm 3.15	28.85 \pm 1.34
FAST	4.17 \pm 0.51	1.62 \pm 0.51

ings were down-sampled by a factor of four to reduce their length. Finally, the MEG signals were stored in American Standard Code for Information Interchange (ASCII) for further analysis.

Finally, the Filter Design & Analysis Tool included in Matlab was used to design band-pass Finite Impulse Response (FIR) filters with a Hamming window. These filters were used to limit the bandwidth of the MEG signals in the different applications. For instance, the cut-off frequencies of this filter were set at 0.5 Hz and 60 Hz for the comparison of diverse BSS techniques and the artefact removal. This setting allowed us to study the broad-band spectral content of the signals. On the other hand, in the study of the BSS

techniques as a preprocessing to improve the classification of AD patients and controls, the cut-off frequencies were set at 1.5 Hz and 40 Hz. The lower cut-off frequency allows to reduce the importance of the OAs while the upper one minimises the number of muscular artefacts and removes the PLN. It is important to note that linear filtering is allowed before the application of BSS and ICA techniques because it does not change the generative model. However, it may be possible that a low-pass filtering reduce the information in the data which might lead to a reduction of independence. On the other hand, high-pass filtering may be useful in certain cases to increase the independence of the components by removing slow fluctuations from the signals (Hyvärinen et al., 2001).

5.3 Epoch Selection

The presence of artefacts in the MEG recordings may bias the results of the analysis. One possible application of BSS techniques is the artefact removal in brain recordings (James and Hesse, 2005; Vigário and Oja, 2008). However, this approach was not applied in the study of the classification improvement due to BSS (Section 6.6) or in the extraction of rhythmic activity with BSS (Section 6.7). Instead, previously to the analyses, the MEG epochs with the most clear artefacts were rejected. The epoch rejection method was applied in order to avoid surplus complexity in these analyses and to assess the performance of these approaches without any other kind of preprocessing.

When needed, the artefactual epochs were rejected on the basis of a semi-automatic procedure which combined a visual inspection of the recordings with an amplitude-thresholding method to detect spurious activity. Several specialists, without previously knowing the subject's diagnosis, visually discarded the signal epochs contaminated with noise and strong artefacts. Additionally, the specialists were helped by the amplitude-thresholding procedure to pinpoint noisy signal epochs. This kind of process has been previously used in MEG studies on dementia (Osipova et al., 2003, 2005, 2006a,b; Pekkonen et al., 1999, 2005). In the case of this Doctoral Thesis, this consisted of setting a variable threshold for each channel, defined as the median of the

maximum values computed for windows of 100 data samples. Empirically, it was determined that setting the threshold to four times the median of the maximum values automatically marked a large fraction of the artefactual epochs.

Chapter 6

Methods

Contents

6.1	Introduction	139
6.2	Preprocessing	140
6.2.1	Selection of the Number of Components	141
6.2.2	Synthetic Data	142
6.3	Blind Source Separation Methods	145
6.3.1	SOS-BSS	146
6.3.2	HOS-BSS	146
6.3.3	cBSS	148
6.4	Consistency Evaluation of BSS Methods	148
6.4.1	Metric	149
6.5	Artefact Removal With BSS	150
6.5.1	Artefact Detection Metrics	151
6.5.2	Evaluation of the Artefact Rejection	160
6.6	Feature Enhancement With BSS	163
6.6.1	Description	163
6.6.2	Signal Features	165
6.7	Extraction of Brain Rhythms With cBSS	171
6.7.1	Preprocessing	171
6.7.2	Empirical Mode Decomposition	172
6.7.3	Frequency Characterisation of the IMFs	173
6.7.4	Regional Extraction of Brain Rhythms	174

6.7.5	Coherence	176
6.8	Statistical Analysis Tools	177
6.8.1	Visual Information Tools	177
6.8.2	Statistical Significance Tests	177
6.8.3	Classification Analysis	178

This Chapter details the methods and techniques applied in each of the studies carried out in this Doctoral Thesis. Firstly, the preprocessing used to determine the number of components to be extracted from the recordings is explained. Next, the selection of the input parameters employed in the BSS algorithms is justified. Afterwards, it is described how the outcomes of diverse BSS methods are compared to reveal their similarities and differences. The framework used to assess the artefact removal with BSS techniques is also detailed. Then, the use of BSS as a preprocessing to improve the classification of features computed from AD patients and control subjects' MEGs is described. Finally, the methodology developed to adaptively extract brain rhythms in certain frequency bands and scalp regions is presented.

6.1 Introduction

For the sake of clarity, let us recall the basic ideas behind the BSS techniques. The m MEG channels, $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_m(t)]^T$ are considered a linear mixture of n underlying, inner components, or sources, $\mathbf{s}(t) = [s_1(t), s_2(t), \dots, s_n(t)]^T$. Here T denotes transposition and all the variables are assumed to have zero mean. As it was already shown in Eq. (4.11), the simplest model is:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t), \quad (6.1)$$

where \mathbf{A} is a full rank $m \times n$ mixing matrix with $m \geq n$ (Hyvärinen et al., 2001; James and Hesse, 2005; Vigário and Oja, 2008).

This simplified model can be suitable if additive sensor noise is low enough (James and Hesse, 2005). However, a more realistic model may be used to consider external noise, which is modelled as an m -dimensional vector of additive spatially uncorrelated gaussian noise, $\mathbf{n}(t)$, (Barbati et al., 2004; Escudero et al., 2007b; Ting et al., 2006; Vorobyov and Cichocki, 2002). Thus, the model becomes:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t). \quad (6.2)$$

In this setting, only the observations, $\mathbf{x}(t)$, are available. Thus, $\mathbf{x}(t)$ and \mathbf{A} , have to be estimated blindly from $\mathbf{x}(t)$. Using the assumptions described in Section 4.3.1, the estimation of the components, $\hat{\mathbf{s}}(t)$, is done as:

$$\hat{\mathbf{s}}(t) = \mathbf{B}\mathbf{x}(t) = \mathbf{B}[\mathbf{A}\mathbf{s}(t) + \mathbf{n}(t)] = \mathbf{s}(t) + \mathbf{B}\mathbf{n}(t), \quad (6.3)$$

where \mathbf{B} is the BSS demixing matrix which must fulfil $\mathbf{B} = \hat{\mathbf{A}}^+$ (Cichocki and Amari, 2002). Here, $\hat{}$ and $^+$ denote a pseudo-inverse matrix and an estimated variable, respectively.

6.2 Preprocessing

In some cases, it may be helpful to preprocess the data to reduce the problem dimensionality and to estimate m (Escudero et al., 2007b; Hesse, 2008; Hyvärinen et al., 2001). Ideally, this preprocessing should be robust to external noise (Escudero et al., 2007b; Hesse, 2007, 2008).

It can be useful to apply a dimensionality reduction before BSS for several reasons. Firstly, when high-density recording equipment is used, the number of inner meaningful components may be less than the number of available channels for analysis (Ikeda and Toyama, 2000; James and Hesse, 2005). Secondly, a dimensionality reduction can sometimes be needed to avoid ‘overfitting’ (Vigário and Oja, 2000, 2008). Finally, the dimensionality reduction helps to reduce the importance of the external noise (Ikeda and Toyama, 2000).

Ordinarily, the preprocessing is performed by standard PCA (James and Hesse, 2005). However, this approach has some drawbacks, including some arbitrariness and the assumption that the power of $\mathbf{n}(t)$ is negligible, which may not be true in un-averaged MEG data (Cao et al., 2003). To overcome this problem, the eigenvalue spectrum may be split into signal and noise subspaces by a power threshold so that the external noise power is estimated from the noise subspace (Barbati et al., 2004; Cichocki and Amari, 2002; Ting et al., 2006). However, this method assumes that all MEG channels have the same noise power, and it retains the aforementioned subjectivity (Escudero

et al., 2007b).

In contrast to those techniques, we used the preprocessing with noise reduction proposed in (Cao et al., 2003): the unweighted least squares method of FA. This method takes into account the *diagonal elements in the covariance of the external noise*: $\Psi = E \left\{ \mathbf{n}(t) \mathbf{n}(t)^T \right\}$ (Cao et al., 2003).

Let $\hat{\mathbf{A}}_{\text{pr}}$ be the estimated preprocessing mixing matrix that relates the preprocessed data, $\mathbf{z}(t)$, to $\mathbf{x}(t)$:

$$\mathbf{x}(t) = \hat{\mathbf{A}}_{\text{pr}} \mathbf{z}(t) + \mathbf{n}(t) \quad (6.4)$$

and let $\mathbf{C}_{\text{xx}}^0 = E \left\{ \mathbf{x}(t) \mathbf{x}^T(t) \right\}$. $\hat{\mathbf{A}}_{\text{pr}}$ is iteratively computed from the eigenvalue decomposition of $(\mathbf{C}_{\text{xx}}^0 - \Psi)$, and Ψ is also iteratively estimated as the diagonal elements of $(\mathbf{C}_{\text{xx}}^0 - \hat{\mathbf{A}}_{\text{pr}} \hat{\mathbf{A}}_{\text{pr}}^T)$. The full process is detailed in (Cao et al., 2003).

Once $\hat{\mathbf{A}}_{\text{pr}}$ and $\hat{\Psi}$ have reached stable values, we calculate the whitening matrix, \mathbf{Q} , as the pseudo-inverse of $\hat{\mathbf{A}}_{\text{pr}}$. Since pseudo-inverse matrices are not unique, we used (Cao et al., 2003; Ikeda and Toyama, 2000):

$$\mathbf{Q} = \left(\hat{\mathbf{A}}_{\text{pr}}^T \hat{\Psi}^{-1} \hat{\mathbf{A}}_{\text{pr}} \right)^{-1} \hat{\mathbf{A}}_{\text{pr}}^T \hat{\Psi}^{-1}, \quad (6.5)$$

which considers $\hat{\Psi}$.

Using \mathbf{Q} , the preprocessed data are obtained by means of (Cao et al., 2003; Escudero et al., 2007b):

$$\hat{\mathbf{z}}(t) = \mathbf{Q} \mathbf{x}(t). \quad (6.6)$$

6.2.1 Selection of the Number of Components

The preprocessing previously detailed assumes that n is known. However, n is a parameter that must also be estimated blindly from the data. Considering $n \leq m$ and the number of free parameters, a bound for the number of components (n_{max}) in this model can be found as (Cao et al., 2003; Ikeda

and Toyama, 2000):

$$n_{max} = \left\lfloor \frac{1}{2} \left(2m + 1 - \sqrt{8m + 1} \right) \right\rfloor, \quad (6.7)$$

where $\lfloor a \rfloor$ denotes the integer part of a .

In order to determine the optimum value of n , we used a method derived from FA based on *statistical model selection* with information criteria. We estimated $\hat{\mathbf{A}}_{pr}$ and $\hat{\Psi}$ for each $1 \leq n \leq n_{max}$. Then, the MDL was computed for each n value as follows (Ikeda and Toyama, 2000):

$$\begin{aligned} \text{MDL}(n) = & \frac{1}{2} \text{tr} \left[\mathbf{C}_{xx}^0 \left(\hat{\Psi} + \hat{\mathbf{A}}_{pr} \hat{\mathbf{A}}_{pr}^T \right)^{-1} \right] + \\ & \frac{1}{2} \log \left[\det \left(\hat{\Psi} + \hat{\mathbf{A}}_{pr} \hat{\mathbf{A}}_{pr}^T \right) \right] + \\ & \frac{m}{2} \log(2\pi) + \frac{\log N}{N} \left[m(n+1) - \frac{n(n-1)}{2} \right], \end{aligned} \quad (6.8)$$

where N is the number of signal samples, and $\text{tr}(\cdot)$ denotes the trace of a matrix. Finally, the data were preprocessed with the matrix \mathbf{Q} corresponding to the value of n that minimises the MDL (Escudero et al., 2007b; Ikeda and Toyama, 2000).

This preprocessing and component selection procedure was applied in the analyses carried out in Section 6.4 and Section 6.5. On the other hand, by the own nature of the studies performed in Section 6.6 and Section 6.7, this kind of procedure was not used.

6.2.2 Synthetic Data

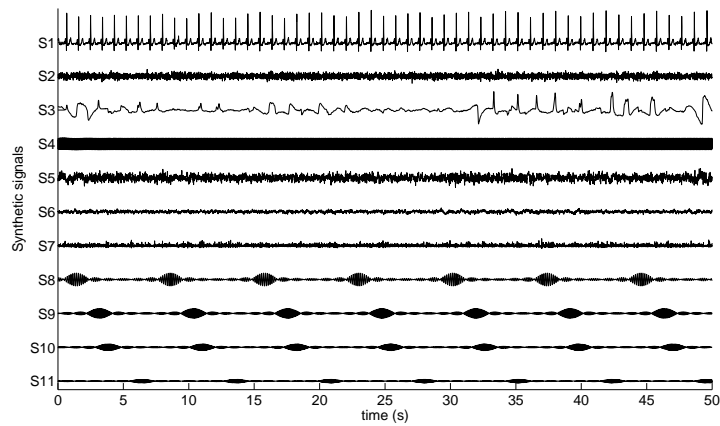
We used simulated data to *test the performance of the preprocessing* (Escudero et al., 2007b). These data were composed by 11 sources. Four of them (S1 to S4) were considered artefactual components, whereas the other seven (S5 to S11) simulated ‘useful’ signals. These signals had the same sample frequency (169.54 Hz) as the real MEG signals analysed in this Doctoral Thesis. Their length was set to 50 s, a signal epoch length previously used in artefact removal studies (Escudero et al., 2006c,d, 2007b, 2008c).

Fig. 6.1 depicts one example of each synthetic source. Their time plot, Normalised Power Spectral Density (PSD_n) and histogram are shown. Their description follows (Escudero et al., 2007b):

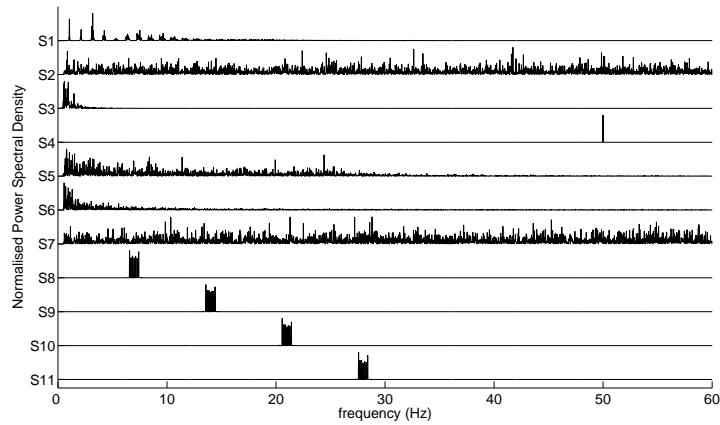
1. S1 corresponds to a real ECG signal.
2. S2 is an inner white gaussian noise source.
3. S3 is a real EOG.
4. S4 simulated PLN (a sine wave at 50 Hz).
5. S5 is a real MEG epoch recorded at a central position of the head. It was selected to have minimal ocular and cardiac activity. The PLN was reduced using a Q -notch digital filter. Moreover, in order to remove any possible remainder of cardiac activity, the projection of a simultaneous ECG recording onto the epoch was computed. Then, this projection was removed from S5 (Escudero et al., 2007b; Samonas et al., 1997).
6. S6 is a $1/f$ noise source. Similarly to the EOG, this noise has most of its power in low frequencies.
7. Likewise the ECG, S7 is a skewed and supergaussian signal. It is a white exponential noise source with the parameter $\lambda = 1$.
8. Finally, S8 to S11 represented rhythmic activity. Their main frequencies are 7 Hz, 14 Hz, 21 Hz and 28 Hz, respectively, and their bandwidth is 1 Hz.

From these 11 source components, 52 mixed signals were created using mixing matrices randomly generated by a gaussian process with zero mean and SD equal to 1 (Barbati et al., 2004; Escudero et al., 2007b).

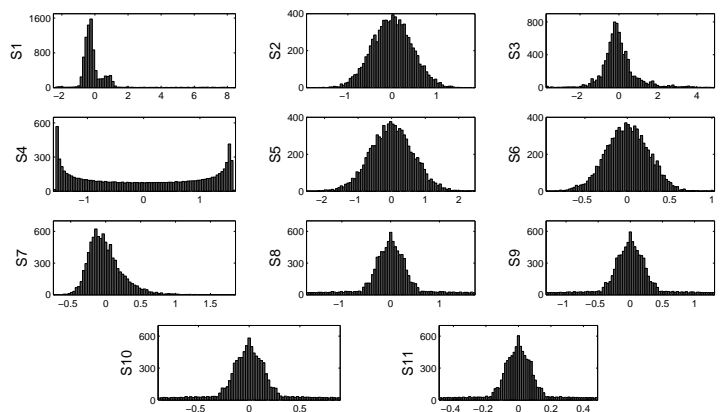
If the simulated artefacts (S1 to S4) and the useful sources (S5 to S11) were considered inner noise and useful signal, respectively, the inner SNR would be about -5.5 dB. Moreover, additive gaussian noise with different power was added to each of the 52 mixed signals. The additive noise power was varied, and the source delay (or phase) was changed in every realisation



(a) Time plots.



(b) Normalised PSD.



(c) Normalised histograms.

Figure 6.1: One example of synthetic sources. S1–S4: simulated undesired signals, S5–S11: simulated ‘useful’ signals.

of the data set (Escudero et al., 2007b). Finally, all the mixtures were filtered using a the same band-pass FIR filter (Hamming window) with cut-off frequencies at 0.5 Hz and 60 Hz that was applied to the real MEG data in the evaluation of the BSS algorithms consistency and artefact removal (Escudero et al., 2007b, 2009d).

As it has been previously mentioned, these synthetic data were used to test the performance of preprocessing. This includes the accuracy in the estimation of the number of inner components and the external noise power (Escudero et al., 2007b). To assess the latter, the following approach was taken. Let Ψ be the actual covariance matrix of the additive external noise added to the mixtures and $\hat{\Psi}$ denote the estimation of Ψ provided by the preprocessing. The error in the estimation of the external noise power E_{norm} was defined as (Escudero et al., 2007b):

$$E_{norm} = \frac{\sum \text{abs} \left[\text{tr} \left(\hat{\Psi} \right) - \text{tr} \left(\Psi \right) \right]}{\sum \text{tr} \left(\Psi \right)}, \quad (6.9)$$

where $\text{tr}(\cdot)$ and $\text{abs}(\cdot)$ denote trace of a matrix and absolute value, respectively.

6.3 Blind Source Separation Methods

This Section describes the selection of the parameters and the particular details of the BSS algorithms used in this Doctoral Thesis. These are: AMUSE (detailed in Section 4.4.2; Tong et al., 1991), SOBI (described in Section 4.4.3; Belouchrani et al., 1997), FastICA (presented in Section 4.5.2; Hyvärinen, 1999; Hyvärinen and Oja, 1997), eInfoMax (introduced in Section 4.5.3; Lee et al., 1999), JADE (explained in Section 4.5.4; Cardoso and Souloumiac, 1993, 1996) and the cBSS algorithm by Huang and Mi (2007) whose basic ideas were outlined in Section 4.6.

The implementations of these algorithm are available at the EEGLAB (Delorme and Makeig, 2004), FastICA (Gävert et al., 2009) and ICALAB (Cichocki et al., 2009) toolboxes or upon request (Huang and Mi, 2007).

6.3.1 SOS-BSS

AMUSE and SOBI employ SOS. They are time structure based methods (Hyvärinen et al., 2001), which assume that the sources have no spatial-temporal correlations (James and Hesse, 2005). Thus, these techniques try to diagonalise a set of cross-covariance matrices computed from $\mathbf{x}(t)$ (James and Hesse, 2005).

AMUSE (Tong et al., 1991) only considers two time delays — usually $\tau = 0$ and $\tau = 1$ sample (Cichocki and Amari, 2002). It is important to note that AMUSE orders the components by decreasing linear predictability (Cichocki et al., 2005), a criterion which is closely related to the spectral content of the signals (Escudero et al., 2008d, 2009b).

On the other hand, SOBI (Belouchrani et al., 1997) uses iterative procedures to simultaneously diagonalise multiple temporal lags (Hyvärinen et al., 2001; James and Hesse, 2005). Considering the results reported by Sutherland et al. (2004) and Tang et al. (2005) about the ‘standard set’ of time delays for application of SOBI to brain activity, and similarly to (Crespo-Garcia et al., 2008), SOBI was applied with 50 consecutive time lags from $\tau = 1$ sample to $\tau \approx 300$ ms (Escudero et al., 2009d). Moreover, this choice is supported by the fact that this set of delays covers a wide time interval without extending beyond the support of the autocorrelation function (Tang et al., 2005), as it can be seen in Fig. 6.2.

6.3.2 HOS-BSS

On the other hand, JADE, eInfoMax and FastICA rely on HOS or statistical parameters like negentropy. They look for non-gaussian sources assuming that $\mathbf{x}(t)$ contains observations of random variables where the temporal order is irrelevant (Hyvärinen et al., 2001; James and Hesse, 2005).

In this Doctoral Thesis, FastICA was applied with $\tanh(\cdot)$ as the non-linearity (Escudero et al., 2009d), which corresponds to the function $G_1(y) = \frac{1}{a_1} \log[\cosh(a_1 y)]$. This function was selected for being a good general-purpose function (Hyvärinen, 1999; Hyvärinen et al., 2001). It has been used in previous artefact removal studies (Frank and Frishkoff, 2007; Halder

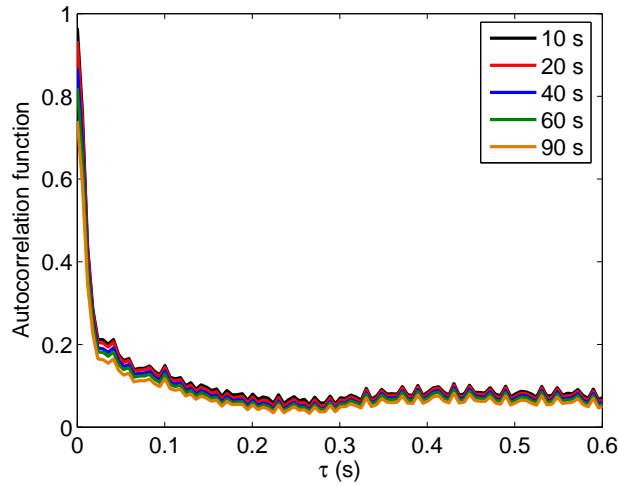


Figure 6.2: Average autocorrelation function of the controls' MEGs for epoch lengths of 10 s, 20 s, 40 s, 60 s and 90 s.

et al., 2007). The deflation approach (i.e., estimation of the components one-by-one) was used in the comparison of BSS algorithms (Section 6.4) and in the artefact removal (Section 6.5). This is due to the fact that the CA and OAs are usually characterised by extreme values of HOS (Escudero et al., 2007b). Hence, the deflation approach may be helpful to extract this kind of components more reliably (Hyvärinen et al., 2001). On the other hand, the symmetric approach was applied in the BSS preprocessing studied in Section 6.6 in order to assure that the possible errors in the estimation of the components were distributed among all components (Hyvärinen et al., 2001). Moreover, this approach is faster if all components are to be estimated (Halder et al., 2007).

The extended version of InfoMax was used in order to estimate both sub- and super-gaussian sources (Escudero et al., 2009d). This version of the algorithms has been widely applied in the artefact removal problem (Crespo-Garcia et al., 2008; Fitzgibbon et al., 2007; Hoffmann and Falkenstein, 2008; Romero et al., 2008; Rong and Contreras-Vidal, 2006). The number of each type of components was automatically determined (Lee et al., 1999).

JADE has no tunable parameters (Escudero et al., 2009d; Hyvärinen et al., 2001; Zavala Fernández et al., 2006). Additionally, it is important

to note that JADE could not be used in the BSS preprocessing described in Section 6.6. The reason was that the high dimensionality of the data ($n = 148$) produces an extremely high memory requirement.

6.3.3 cBSS

The cBSS algorithm applied in this Doctoral Thesis was developed by Huang and Mi (2007) as a modification of the method originally introduced by Lu and Rajapakse (2005, 2006). This framework incorporates a priori information from a reference signal into the negentropy contrast function as inequality and equality constrained terms by means of an augmented Lagrangian function (Huang and Mi, 2007; Lu and Rajapakse, 2006).

The algorithm by Huang and Mi (2007) improves the performance and the quality of the separation of the original one by Lu and Rajapakse (2006). The norm used to measure the similarity between the extracted component and the reference was the mean square error: $E \{[\mathbf{y}(t) - \mathbf{r}(t)]^2\}$, which has been suggested to be appropriate for brain recordings (Huang and Mi, 2007). Finally, it must be noticed that the reference signal was calculated as the sign function of the activity supposed to follow the activity of interest (Huang and Mi, 2007; James and Gibson, 2003).

6.4 Consistency Evaluation of BSS Methods

A completely accurate quantification of the performance provided by a BSS algorithm q can only be done if either the original mixing matrix, \mathbf{A} , or the set of sources, $\mathbf{s}(t)$, is known (Hyvärinen et al., 2001; Melissant et al., 2005). In this case, the evaluation can be carried out using a framework like the one proposed by Vincent et al. (2006). In the case of real EEG and MEG recordings, these data are not available. However, the consistency (i.e., similarity) of various BSS algorithms can still be precisely computed (Cichocki et al., 2009).

In order to achieve this goal, two different BSS algorithms (algorithm q and algorithm r) must be applied to the *same data* in order to estimate the

corresponding mixing matrices: \mathbf{A}^q and \mathbf{A}^r (Cichocki et al., 2009). Then, the columns of these matrices are normalized to unit length vectors and a matrix \mathbf{P}^{qr} is computed as:

$$\mathbf{P}^{qr} = (\mathbf{A}^q)^{-1} \mathbf{A}^r. \quad (6.10)$$

If the two algorithms q and r provide exactly the same separation, \mathbf{P}^{qr} will be a generalized permutation matrix. Similarly, the closer \mathbf{P}^{qr} is to a permutation matrix, the more consistent (i.e., similar) the separations of the algorithms q and r are (Cichocki et al., 2009).

6.4.1 Metric

In the aforementioned methodology, it is necessary to define a metric to assess the degree to which \mathbf{P}^{qr} is close to a permutation matrix. Hence, we introduce the metric F as:

$$F = \frac{F_1 + F_2}{2}, \quad (6.11)$$

with F_1 and F_2 computed as in (Escudero et al., 2009d; Melissant et al., 2005):

$$F_1 = \frac{1}{n} \sum_{i=1}^n \left[\frac{1}{n-1} \left(\sum_{j=1}^n \frac{|p_{ij}|}{\max_k |p_{ik}|} - 1 \right) \right], \quad (6.12)$$

and

$$F_2 = \frac{1}{n} \sum_{j=1}^n \left[\frac{1}{n-1} \left(\sum_{i=1}^n \frac{|p_{ij}|}{\max_k |p_{kj}|} - 1 \right) \right], \quad (6.13)$$

where p_{ij} denotes an element of \mathbf{P}^{qr} and n is the number of components.

F_1 measures the average coupling of other sources into one particular component, whereas F_2 accounts for the fact that two or more estimated components represent exactly the same original source (Melissant et al., 2005). It is worth noting that F_1 and F_2 are normalized so that their values do not depend on the dimensions of \mathbf{P}^{qr} . Since F_1 and F_2 are bounded between 0 (for a perfect generalized permutation matrix) and 1, F also ranges between 0 and 1. Hence, the lower the value of F for a pair of algorithms, the more consistent they are (i.e., the outcomes of both algorithms are more similar) (Escudero et al., 2009d).

6.5 Artefact Removal With BSS

BSS may be useful in the artefact removal problem, since it decomposes the recordings into components with minimal dependence between them. Then, these components can be inspected to find out which are responsible for the artefacts (Barbati et al., 2004; Escudero et al., 2007b; Ting et al., 2006; Vigário, 1997).

The marked artefactual components can be removed from the recordings to reduce the importance of the artefacts in brain signals. Criteria to determine which sources are artefacts will be described in Section 6.5.1.

Once the artefactual components have been identified, the artefact-free signals can be reconstructed by two approaches (Mantini et al., 2008; Ting et al., 2006):

1. The first one is done by re-mixing the artefact-free components:

$$\mathbf{x}_{clean}(t) = \sum_{i \in I} \mathbf{a}_i \mathbf{s}_i(t), \quad (6.14)$$

where $\mathbf{x}_{clean}(t)$ denotes the reconstructed, artefact-free, MEG signals and I represents the set of components *not being rejected*.

2. As an alternative, the identified artefact sources can be subtracted from the observed signals:

$$\mathbf{x}_{clean}(t) = \mathbf{x}(t) - \sum_{j \in J} \mathbf{a}_j \mathbf{s}_j(t), \quad (6.15)$$

where J denotes the indices of *all artefactual components*.

The second approach is adopted in this Doctoral Thesis. This is due to the fact that the noise term $\mathbf{n}(t)$ in Eq. (6.2) can be considered as a modelling error (Ting et al., 2006). Hence, the second approach tries to minimize the distortion to the *true* brain activity by the artefact removal process (Ting et al., 2006). Moreover, the second approach can be used together with a cBSS algorithm. Typically, these techniques only extract one component

with collects the artefactual activity. Hence, the first reconstruction introduced in Eq. (6.14) cannot be applied, whereas the term \mathbf{a}_j in Eq. (6.15) can be estimated as the correlation coefficients between the original recorded channels and the extracted components (Huang and Mi, 2007).

Only the MEG recordings of the subjects in the control group have been studied in this part of the Doctoral Thesis. The objective was to compare the performance of diverse artefact removal procedures avoiding the possible influences that AD could produce. In order to analyse the effect of different epoch lengths on the performance of the artefact removal, the BSS algorithms were applied to signal epochs of 10 s, 20 s, 40 s, 60 s and 90 s.

6.5.1 Artefact Detection Metrics

In this Section, we present the metrics used in this Doctoral Thesis to detect the considered artefacts: CA, OA and PLN.

Before describing each of the metrics implemented in this study, it is important to notice that a ‘segment’ approach was taken to compute the metrics based on statistical properties from the signals. This procedure consists of dividing the components estimated from the corresponding signal epochs into *non-overlapping segments* (also known as trials) of 1 s (169 data samples) (Dammers et al., 2008; Mammone and Morabito, 2008). For the sake of clarity, it is recalled that the signal epoch length considered in this study are: 10 s, 20 s, 40 s, 60 s and 90 s. Then, within each component, the parameter is computed for each segment (or trial) separately, as proposed by Barbati et al. (2004); Dammers et al. (2008); Greco et al. (2006) and Mammone and Morabito (2008). In case a certain fraction of the segments of a BSS component exceeds a predefined threshold, that component is marked for rejection (Barbati et al., 2004; Dammers et al., 2008; Greco et al., 2006; Mammone and Morabito, 2008).

All distributions derived from this ‘segment’ approach were normalized to zero mean and SD equal to 1 with respect to all components of the database extracted with the same BSS algorithm and epoch length (Barbati et al., 2004; Dammers et al., 2008; Greco et al., 2006; Mammone and Morabito,

2008). Then, a component was marked as artefact if 30% or more of its segments exceeded a threshold value of ± 2.0 (Dammers et al., 2008). This criterion is more restrictive than the commonly used consisting of marking a BSS source if 20% or more of its trials exceeded a value of ± 1.64 (Barbati et al., 2004), and it aims at *minimising the false positive artefact detection* (Dammers et al., 2008).

When the probability distribution needed to be computed, the recommendation by Delorme et al. (2007) was followed so that the number of bins of the histogram was set to the number of data samples divided by five.

Cardiac Artefact

The following metrics were used in the detection of cardiac components.

Kurtosis Excess Kurtosis has been used in several studies to recognise diverse artefacts, including cardiac and ocular ones (Barbati et al., 2004; Dammers et al., 2008; Escudero et al., 2006c, 2007b; Mammone and Morabito, 2008).

Let $m_n = E\{(x - E\{x\})^n\}$ the n th central moment of the amplitude distribution of the signal x . Kurtosis Excess (KrE) is defined as (Escudero et al., 2006c, 2007b):

$$\text{KrE} = \frac{m_4}{(m_2)^2} - 3. \quad (6.16)$$

KrE is negative for sub-gaussian amplitude distributions ('flatter' than the gaussian one). However, if the samples are highly gathered around the central values and the distribution has 'heavy tails', KrE is positive (super-gaussian distribution) (Barbati et al., 2004; Delorme et al., 2007; Escudero et al., 2007b). Thus, large KrE values are associated with super-gaussian components, which may be due to the CA and OA (Barbati et al., 2004; Escudero et al., 2007b).

This metric was applied to the components following the segment 'approach' previously introduced and it has been tested in the rejection of all kind of artefacts considered in this Doctoral Thesis.

Skewness Likewise KrE, Skewness (Skew) has been proposed to detect artefactual activity isolated in BSS components (Dammers et al., 2008; Escudero et al., 2006c, 2007b; Shao et al., 2008; Shoker et al., 2005a).

Using the same notation as in Eq. (6.16), Skew is computed as (Dammers et al., 2008; Escudero et al., 2007b; Shoker et al., 2005a):

$$\text{Skew} = \frac{m_3}{(m_2)^{\frac{3}{2}}}. \quad (6.17)$$

Skew measures the asymmetry degree of a distribution. Only if the amplitude distribution is symmetrical, Skew is zero (Escudero et al., 2007b). Thus, large abs(Skew) values, where abs(\cdot) denotes absolute value, are associated with asymmetric components, which may be due to ocular and, specially, cardiac activity (Escudero et al., 2007b; Shoker et al., 2005a). Hence, this metric, which was computed in segments of the BSS components, was applied to assess the reduction of cardiac activity in the data.

Shannon Entropy The detection of artefacts in BSS components from brain activity was also carried out using Shannon Entropy (H_{Sh}) (Barbati et al., 2004; Greco et al., 2006; Rong and Contreras-Vidal, 2006).

Here, the H_{Sh} is computed as (Dammers et al., 2008):

$$H_{Sh} = - \sum_i p_i(x) \log [p_i(x)], \quad (6.18)$$

where the probabilities were estimated from the signal histogram calculated with a number of bins equal to the number of samples divided by five (Delorme et al., 2007).

Entropy is a measure of disorder or irregularity. Higher entropy values correspond to more unstructured signals. On the other hand, small entropy values correspond to components whose amplitude distributions are contained in a few limited intervals with high probabilities, as it happens in certain types of artefacts with ‘spiky’ probability densities (Barbati et al., 2004; Dammers et al., 2008). Thus, H_{Sh} was used to mark BSS components with all types of artefacts considered in this Doctoral Thesis.

Rényi Entropy H_{Sh} can be replaced with Rényi Entropy ($H_{R\acute{e}}$) in the detection of artefactual components (Greco et al., 2006; Mammone and Morabito, 2008). The computation of $H_{R\acute{e}}$ depends on an entropic index, q . Actually, $H_{R\acute{e}}$ is linked to the distribution of the signals through q . A high q value emphasises super-gaussian distributions whereas a low q emphasises the sub-gaussian ones (Mammone and Morabito, 2008). If the mixtures belong to different kurtosis classes, a quadratic entropy ($q = 2$) can be used to put equal emphasis on all data points regardless of their probability density (Greco et al., 2006; Mammone and Morabito, 2008).

The $H_{R\acute{e}}$ is computed as follows (Greco et al., 2006; Mammone and Morabito, 2008):

$$H_{R\acute{e}} = \frac{1}{1-q} \log \sum_i p_i^q(x), \quad (6.19)$$

with $q = 2$ (Greco et al., 2006; Mammone and Morabito, 2008). This metric was tested in the detection of CA, OA and PLN.

Approximate Entropy Another entropic measure is Approximate Entropy (ApEn). The use of ApEn as a metric to detect artefactual components in MEG activity was proposed by Mantini et al. (2008). It was shown that this metric could classify the components extracted with FastICA into three groups corresponding to cerebral activations, non-cerebral biological signals and environmental disturbances (Mantini et al., 2008). ApEn is expected to be small for non-cerebral biological signals, such as cardiac and ocular activities (Mantini et al., 2008).

ApEn is a family of statistics introduced to quantify the regularity of a sequence (Pincus, 2001). It assigns a non-negative number to a time series, with larger values corresponding to more irregularity in the data. A run length m and a tolerance window r must be specified to compute it (Pincus, 2001). Given N points, $\text{ApEn}(m, r, N)$ measures the logarithmic likelihood that runs of patterns that are close (within r) for m contiguous observations remain close (within the same tolerance width r) on subsequent incremental comparisons (Pincus, 2001).

Given N data points from a time series $\{x(n)\} = \{x(1), x(2), \dots, x(N)\}$,

one should follow these steps to compute ApEn (Pincus, 2001):

1. Form $N - m + 1$ vectors $X(1), X(2), \dots, X(N - m + 1)$ defined by: $X(i) = [x(i), x(i + 1), \dots, x(i + m - 1)]$, $i = 1, \dots, N - m + 1$. These vectors represent m consecutive x values, commencing with the i th point.
2. Define the distance between $X(i)$ and $X(j)$, $d[X(i), X(j)]$, as the maximum norm:

$$d[X(i), X(j)] = \max_{k=1, \dots, m} (|x(i + k - 1) - x(j + k - 1)|). \quad (6.20)$$

3. For a given $X(i)$, count the number of j ($j = 1, \dots, N - m + 1$) so that $d[X(i), X(j)] \leq r$, denoted as $N^m(i)$. Then, for $i = 1, \dots, N - m + 1$,

$$C_r^m(i) = \frac{N^m(i)}{N - m + 1} \quad (6.21)$$

$C_r^m(i)$ measures, within a tolerance r , the frequency of patterns similar to a given one of window length m .

4. Compute the natural logarithm of each $C_r^m(i)$, and average it over i ,

$$\phi^m(r) = \frac{1}{N - m + 1} \sum_{i=1}^{N-m+1} \ln C_r^m(i). \quad (6.22)$$

5. Increase the dimension to $m + 1$. Repeat steps 1 to 4 and find $C_r^{m+1}(i)$ and $\phi^{m+1}(r)$.
6. ApEn is defined by:

$$\text{ApEn}(m, r, N) = \phi^m(r) - \phi^{m+1}(r). \quad (6.23)$$

ApEn was computed with the values of $m = 2$ and $r = 0.2$ times the SD of the analysed signal, as proposed by Mantini et al. (2008) and this parameter was applied following the ‘segment’ procedure.

In this Doctoral Thesis, ApEn has been tested in the removal of CA, OAs and PLN.

Variance of the scalp distribution The Variance of the Scalp Distribution (VarSc) is introduced as feature able to detect CAs in a BSS component since it has been suggested that the scalp distributions associated with cardiac activity gives smaller variance compared to other types of components (Shao et al., 2008).

This metric is computed as (Shao et al., 2008):

$$\text{VarSc} = \text{var} \left(\frac{\mathbf{a}_i}{\|\mathbf{a}_i\|} \right), \quad (6.24)$$

where $\text{var}(\cdot)$ represents the variance and \mathbf{a}_i is the scalp distribution coefficients in the mixing matrix corresponding to the component i (Shao et al., 2008).

Due to the facts that this metric intends to detect components whose scalp distribution has extremely low variance and that the cardiac contamination always appears in this recordings (Escudero et al., 2007b), the component with the minimum value of VarSc for each signal epoch was marked for rejection.

Constrained Blind Extraction of the Cardiac Artefact In addition to applying a BSS method and characterising the extracted components, a cBSS approach can be taken to reduce the cardiac contamination in the MEGs (James and Gibson, 2003). This is due to the fact that cBSS extracts components that are characterised by having minimal dependence of the other underlying sources and being similar to some reference signal (Huang and Mi, 2007; James and Gibson, 2003). However, this reference should be derived from the data or from some a priori knowledge (Huang and Mi, 2007; James and Gibson, 2003).

In this Doctoral Thesis, a suitable reference for the extraction of the CA has been computed taking into account that the MEG signals contain background activity. Thus, an average of all MEG channels tends to cancel

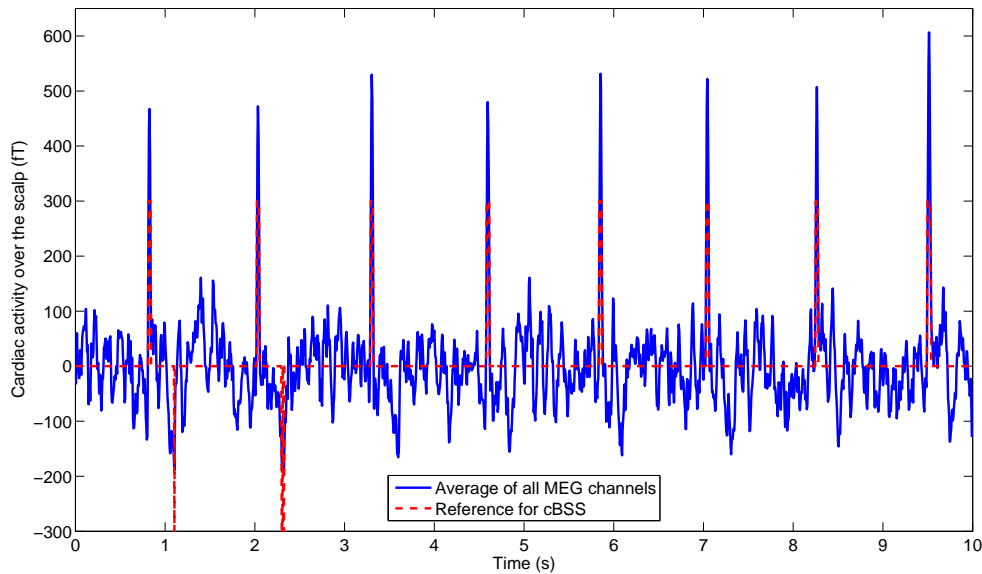


Figure 6.3: Example of the cardiac activity that appears over the scalp and the reference extracted from it.

the brain activity while the only synchronous activity (the CA) remains in the average signal (Escudero et al., 2007b). This is illustrated in Fig. 6.3, which shows the average of a 10 s MEG epoch.

It is more important for the reference to capture the temporal features of interest than follow the relative morphology of the activity to be extracted (James and Gibson, 2003). Moreover, the average of all MEG channels may not be a pure cardiac signal since some other activities might also appear in it. Hence, the actual reference introduced to the cBSS algorithm was computed as a series of positive and/or negative pulses synchronised with the R-peaks of the cardiac activity (James and Gibson, 2003). This reference is also plotted in Fig. 6.3. Finally, the extracted activity is projected to the MEG channels space and subtracted from the recorded signals (Huang and Mi, 2007; James and Gibson, 2003).

Ocular Artefact

This Doctoral Thesis also aims at reducing the importance of the OAs in the MEG recordings. The ocular components share some characteristics with

the CA (e.g., both are usually characterised by extreme values of kurtosis or entropy). Therefore, some of the metrics proposed to detect the cardiac components can be also used to mark ocular sources. These are (Barbati et al., 2004; Dammers et al., 2008; Mammone and Morabito, 2008; Mantini et al., 2008; Rong and Contreras-Vidal, 2006):

- KrE.
- H_{Sh} .
- $H_{Ré}$.
- ApEn.

Similarly to the case of the CA, these features were calculated using the ‘segment’ approach and the criterion for rejection was also that 30% or more of the segments in a component have a normalised value that exceeds ± 2.0 (Dammers et al., 2008).

Additionally, a couple of metrics specially focused on the detection of the OA were also tested in this Doctoral Thesis. They are detailed in the following lines.

Power near the eyes The power of the OA is mainly gathered near the eyes. Hence, the scalp distribution has already been used to help in the detection of this artefact (Escudero et al., 2007b; Fatourechi et al., 2007; Frank and Frishkoff, 2007; Li et al., 2006; Rong and Contreras-Vidal, 2006). Therefore, the metric Power near the Eyes (P_{EYES}) is proposed (Escudero et al., 2006d, 2007b). This is computed as the fraction of the power located on the 13 frontal peripheral channels nearer to the eyes, which are depicted in Fig. 6.4.

For a component \mathbf{s}_j , this metric can be computed as follows (Escudero et al., 2007b):

$$P_{EYES} = \sum_{k \in K_{EYES}} a_{kj}^2 / \sum_k \mathbf{a}_{kj}^2, \quad (6.25)$$

where \mathbf{a}_{kj} is an element of the mixing matrix \mathbf{A} and K_{EYES} denotes the set of channels considered.

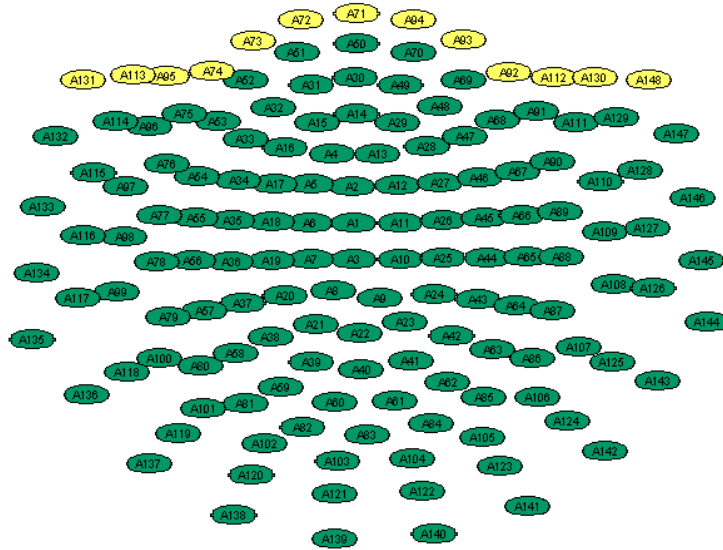


Figure 6.4: Location of the 13 MEG channels considered in the P_{EYES} metric.

This metric intends to detect components whose P_{EYES} is extremely high. Moreover, not all signal epoch are contaminated by OAs. Therefore, in order to minimise the removal of non-artefactual activity, the values of P_{EYES} were normalized to zero mean and $SD = 1$ with respect to all components derived from the same BSS method and epoch length and the detection threshold was set to $+3.5$ (Dammers et al., 2008).

Power in Low Frequencies The energy of the OAs tends to be centred in lower frequencies than in brain or cardiac signals (Escudero et al., 2007b; Fatourechi et al., 2007; Joyce et al., 2004). The low frequency content of the OAs can be helpful to mark these undesired components (Escudero et al., 2006d, 2007b). Thus, we computed the fraction of the Power Spectral Density (PSD) that each BSS component has from 0.5 Hz to 2.5 Hz: the Power in Low Frequencies (P_{LF}) (Escudero et al., 2007b).

Similarly to P_{EYES} , the values of P_{LF} were normalized to mean = 0 and $SD = 1$ with respect to all the relevant components. Then, the detection threshold was set to $+3.5$ to try to minimise the amount of brain activity lost (Dammers et al., 2008).

Power Line Noise

Finally, we introduce the metrics employed to mark components that mainly contain PLN. It has been suggested that ApEn might be able to detect this kind of contamination (Mantini et al., 2008). Additionally, the amplitude distribution of an harmonic component may be different from that of a brain signal. Hence, the kurtosis and entropic metrics will also be used. In addition, the following metric has been tested in this Doctoral Thesis.

Power at the Line Frequency If a BSS component could isolate PLN, its spectrum would be centred at the power line frequency (50 Hz in this case) (Escudero et al., 2006c, 2007b; Iriarte et al., 2003; Jung et al., 2000; LeVan et al., 2006; Nicolaou and Nasuto, 2007). Hence, we calculate a spectral metric, Power at the Line Frequency ($P_{50\text{Hz}}$), which measures the fraction of the PSD contained from 49.5 Hz to 50.5 Hz for each component (Escudero et al., 2007b).

Large values of $P_{50\text{Hz}}$ are due to BSS components that have most of their energy round 50 Hz. The same normalisation procedure and detection criterion as P_{EYES} and P_{LF} (threshold at +3.5 of the normalised values) is applied in this case.

6.5.2 Evaluation of the Artefact Rejection

Most often, the quality of the artefact detection and removal based on BSS has been evaluated using some sort of synthetic signals (Barbati et al., 2004; Delorme et al., 2007; Kierkels et al., 2006; Romero et al., 2008, 2009; Wallstrom et al., 2004) or by means of a visual inspection of the recordings (Flexer et al., 2005; Greco et al., 2006; Joyce et al., 2004; Jung et al., 2000; LeVan et al., 2006; Li et al., 2006; Mammone and Morabito, 2008; Shoker et al., 2005a; Ting et al., 2006; Vigário, 1997; Vorobyov and Cichocki, 2002). However, it is desirable to carry out objective evaluations of the artefact removal as done by Dammers et al. (2008); Escudero et al. (2007b) or Mantini et al. (2008). Thus, this Section describes the measurements used to try to *automatically and objectively assess the artefact removal* from the MEG back-

ground activity.

Cardiac Artefact

In order to assess the CA removal, we detected the QRS-complexes in the average of all 148 MEG channels before the artefact rejection method (Escudero et al., 2007b). This approach is motivated by the fact that the only synchronous activities in the MEG recordings are the CA and the PLN. Therefore, it is possible to locate the clear R-peaks in this average signal (before the artefact rejection process) and estimate an average QRS-complex for each case (Escudero et al., 2007b). The importance of the CA can be characterised with the Peak-to-Peak Amplitude (A_{PtP}) (Escudero et al., 2007b; Mantini et al., 2008) or the Root Mean Square (RMS) value (Dammers et al., 2008) of this ‘typical’ QRS complex. By comparing these values with the corresponding parameters once the BSS-based artefact removal has been applied, it is possible to evaluate the amount of cardiac activity removed.

Ocular Artefact

The evaluation of the OA removal is far more complex than the cases of the CA and PLN. This is due to the fact that the ocular activity is not always present in the MEG recordings and it is not as well localised in the frequency spectrum as the PLN. However, the effects of the BSS artefact rejection procedure on the MEG signals can still be quantified.

For instance, for every signal epoch, the average PSD can be computed at the subset of 18 MEG channels located over the anterior part of the head and close to the eyes and at the remaining set of 130 channels. These two subsets of channels are shown in Fig 6.5. It is important to note that the first subset of channels is not equal to that used in the metric P_{EYES} , in order to avoid biasing the results.

The subtraction of those two PSDs, computed over completely disjoint sets of channels, shows the difference between the MEG activity recorded over those two areas as a function of the frequency. Hence, one can compare a discrepancy measure before and after the OA removal to evaluate its effect.

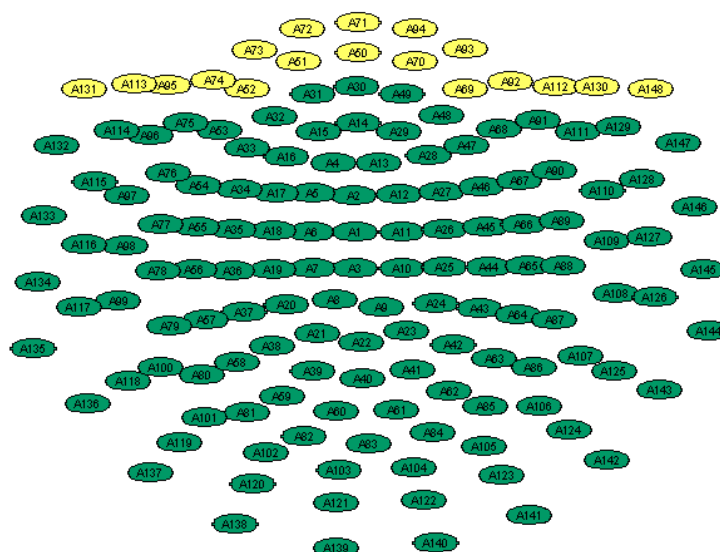


Figure 6.5: Subsets of channels used in the spectral evaluation of the OA removal.

This is done by quantifying the difference between the PSDs between 0.5 Hz and 6.5 Hz. These frequencies are selected since it has been shown that most of the OA energy is located at low frequencies (Romero et al., 2008, 2009; Schlögl et al., 2007).

Additionally, an amplitude threshold was set in the MEG recordings to count the number of peaks in the signals whose amplitude was above ± 2 pT before and after the artefact removal. This measurement was compared as an indication the suppression of peaks in the recordings (Mantini et al., 2008).

Power Line Noise

Thanks to the fact that the spectral location of the PLN is very well defined, it was possible to assess the reduction of this interference by calculating the mean of the PSD of the MEG epochs before and after the artefact rejection and inspecting the amount of power contained between 49 Hz and 51 Hz (Escudero et al., 2007b).

6.6 Feature Enhancement With BSS

This applications tries to improve the discrimination between AD patients and control subjects based on spectral and non-linear signal features by means of a BSS preprocessing.

This is based on the hypothesis that the application of BSS may enhance the separation of those subject groups since some BSS components of the EEG and MEG signals may be more sensitive to AD than others (Cichocki et al., 2005; Escudero et al., 2008d, 2009b; Vialatte et al., 2005). Hence, the most relevant components may be selected and the electromagnetic brain signals may be partially reconstructed using only these components to achieve a better discrimination between AD patients and healthy elderly subjects (Escudero et al., 2008d, 2009b).

In this Doctoral Thesis, we wanted to evaluate whether a BSS preprocessing (using the algorithms AMUSE, SOBI, FastICA and eInfoMax) might enhance the separation between AD patients and elderly control subjects based on spectral and non-linear features of MEG signals. Additionally, we aimed at determining whether the range of BSS components with significant differences between demented patients and controls differed when both kinds of features (spectral and non-linear ones) or BSS techniques were considered.

6.6.1 Description

The methodology is described in the following lines:

1. Firstly, MEG epochs with minimal artefactual contamination were selected for analysis. Then, the signals were decomposed with four BSS algorithms: AMUSE (Tong et al., 1991), SOBI (Belouchrani et al., 1997), FastICA (Hyvärinen and Oja, 1997) and eInfoMax (Lee et al., 1999).

AMUSE and SOBI were applied to the signal epochs of 10 s. In contrast, FastICA and eInfoMax were applied to the whole MEG recordings of five minutes in order to have enough data samples for a reliable

estimation of the decomposition (Cichocki and Amari, 2002). Afterwards, the same signal epochs as in the SOS case were selected for the HOS algorithms (Escudero et al., 2009d).

For simplicity, we assume that $m = n$ thanks to the fact that only the most relevant components will be retained to partially reconstruct the MEG signals. Moreover, considering $m = n$ allows us to consistently compare the same number of extracted components instead of estimating different values of n for each signal epoch (Escudero et al., 2008d, 2009b).

2. Secondly, for every algorithm, the extracted components were ordered to perform objective comparisons between different subjects (Escudero et al., 2008d, 2009b). Two kinds of ordering were used:
 - The MF. This is due to the fact that previous results have shown a direct relationship between the AMUSE ranking used in the seminal paper by Cichocki et al. (2005) and the components' MF (Escudero et al., 2008d, 2009b).
 - The kurtosis. This ordering is considered as a comparison with the previous one based on MF, which can be considered as a summary of the signal spectral content (Poza et al., 2007b). On the other hand, the kurtosis is representative of HOS (Hyvärinen et al., 2001).

Hence, by ordering the components by their MF and kurtosis, we are considering metrics related to both SOS and HOS techniques.

3. Thirdly, on the basis of the previous orderings, the features (MF, SpecEn, LZC and SpecEn) of the AD patients and control subjects' components was compared to define, for each algorithm, the range of components with the most significant differences between subject groups.
4. Afterwards, the MEG signals were partially reconstructed using only the ranges of components with the most significant differences (Ci-

chocki et al., 2005; Escudero et al., 2008d, 2009b). That is, if $\hat{\mathbf{s}}_{range}(t)$ denotes the range of the most sensitive components to AD, a *partial reconstruction of the MEG signals* — $\mathbf{x}_{partial}(t)$ — that may have enhanced features of AD is computed using the estimated mixing matrix $\hat{\mathbf{A}} = \mathbf{B}^{-1}$:

$$\mathbf{x}_{partial}(t) = \hat{\mathbf{A}}\hat{\mathbf{s}}_{range}(t). \quad (6.26)$$

5. Finally, the signal features (MF, SpecEn, LZC and SampEn) were computed from the partially reconstructed MEGs — $\mathbf{x}_{partial}(t)$ — and the original recordings without the BSS preprocessing: $\mathbf{x}(t)$. The classification results obtained in each case were compared to decide which BSS preprocessing improved the discrimination between subject groups.

In order to test these BSS preprocessing on completely unseen data, the selection of the most sensitive components to AD was performed using the training set of subjects, whereas the assessment of the improvement in the separation between AD patients and controls was carried out with the test set (Escudero et al., 2009b). It is recalled that the demographic data and clinical features of these subsets of subjects are shown in Table 5.3.

In contrast to the applications of BSS in the artefact removal from EEG and MEG signals, now we do not aim at isolating specific physiological activity (Escudero et al., 2007b). Instead, we apply BSS as a preprocessing step to enhance the differences between AD patients and controls' brain activity (Cichocki et al., 2005; Escudero et al., 2008d, 2009b). Thus, $\mathbf{x}_{partial}(t)$ does *not intend to resemble the brain activity accurately*, but it may provide a better separation between subject groups than the MEG data without the BSS preprocessing (Escudero et al., 2008d, 2009b).

6.6.2 Signal Features

In this application, every MEG channel and BSS component was characterised with two *spectral* (MF and SpecEn) and two *non-linear* features (LZC and SampEn). These features were selected on the basis of previous studies that showed their usefulness to distinguish AD patients' EEGs and MEGs

from those of healthy elderly subjects (Abásolo et al., 2006c; Escudero et al., 2008d; Gómez et al., 2006a; Hornero et al., 2008, 2009b; Poza et al., 2007b). Moreover, since two of them are spectral features (MF and SpecEn) and the other two (LZC and SampEn) are non-linear analysis methods, the usefulness of the BSS and component selection procedure could be tested with both types of techniques (Escudero et al., 2009b).

Median Frequency

MF has been used to study the electromagnetic brain activity in AD (Hornero et al., 2008; Poza et al., 2007b) since this dementia is associated with a slowing of brain frequencies (Jeong, 2004). This feature summarises the signal spectrum. It provides information about the relative power of low- and high-frequency electromagnetic oscillations produced by local synchrony of neural assemblies. In order to calculate the MF, the PSD of each signal is estimated as the Fourier Transform of its autocorrelation function (Poza et al., 2007b). Then, the MF is computed as the frequency which contains half the PSD power (Escudero et al., 2009b):

$$\frac{1}{2} \left[\sum_{f=1.5 \text{ Hz}}^{40 \text{ Hz}} \text{PSD}(f) \right] = \sum_{f=1.5 \text{ Hz}}^{\text{MF}} \text{PSD}(f). \quad (6.27)$$

It is important to note that the use of MF is *twofold* in this BSS preprocessing (Escudero et al., 2008d, 2009b):

1. It serves as criterion to order the components extracted by the BSS techniques so that comparisons can be made between AD patients and control subjects.
2. It is a signal characterising feature which can be used to classify the subjects.

Spectral Entropy

SpecEn has been applied to AD patients' EEG and MEG recordings to measure the flatness of the signal spectrum (Abásolo et al., 2006c; Poza et al.,

2007b). This is due to the fact that this dementia causes a slowing in the frequency content of the electromagnetic brain signals (Jeong, 2004) and this measure is a convenient way of quantifying these changes. A broad and flat spectrum entails high SpecEn values, whereas a predictable signal with narrow spectral content offers a low SpecEn (Sleigh et al., 2004). In order to estimate the SpecEn, the PSD is normalised — PSD_n — so that (Abásolo et al., 2006c; Escudero et al., 2008d; Poza et al., 2007b):

$$\sum_{f=1.5 \text{ Hz}}^{40 \text{ Hz}} \text{PSD}_n(f) = 1. \quad (6.28)$$

Afterwards, SpecEn is computed applying the Shannon's entropy to the PSD_n (Sleigh et al., 2004):

$$\text{SpecEn} = \frac{-1}{\log N} \sum_{f=1.5 \text{ Hz}}^{40 \text{ Hz}} \text{PSD}_n(f) \log [\text{PSD}_n(f)]. \quad (6.29)$$

Lempel-Ziv Complexity

LZC has been used to analyse various biomedical signals (Aboy et al., 2006), including AD patients' MEG activity (Gómez et al., 2006a, 2009a; Hornero et al., 2008, 2009b). This metric evaluates the signal complexity by measuring the number of distinct substrings and their rate of recurrence along the time series (Lempel and Ziv, 1976). It assigns larger values to more complex data (Lempel and Ziv, 1976). This non-linear feature can offer information about the electromagnetic brain dynamics taking into account that it mainly depends on the signal bandwidth and, to a smaller degree, on the sequence probability density function (Aboy et al., 2006; Ferenets et al., 2006). Additionally, LZC can be interpreted as a harmonic variability metric (Aboy et al., 2006). To compute this feature, the recording must be coarse-grained into a finite symbol sequence (Aboy et al., 2006). In this study, this transformation has been performed by comparing the data points with the median of the signal due to its well-known robustness to outliers (Escudero et al., 2009b; Hornero et al., 2008).

To estimate the LZC complexity, a uni-dimensional signal, $\{x(n)\}$, must be transformed into a finite symbol sequence, typically a binary one (Lempel and Ziv, 1976). By comparison with a threshold T_d , the original signal samples are converted into a 0-1 sequence $\{s(n)\} = \{s(1), s(2), \dots, s(N)\}$, with $s(i)$ defined as (Zhang et al., 2001):

$$s(i) = \begin{cases} 0 & \text{if } s(i) < T_d, \\ 1 & \text{if } s(i) \geq T_d. \end{cases} \quad (6.30)$$

The sequence $\{s(n)\}$ is then scanned from left to right and the complexity counter $c(N)$ is increased by one unit every time a new subsequence of consecutive characters is encountered (Zhang et al., 2001). In order to obtain a complexity measure which is independent of the sequence length, $c(N)$ should be normalized (Zhang et al., 2001). In general, $N/\log_\alpha(N)$ is the upper bound of $c(N)$, where the base of the logarithm α is the number of symbols (two for a binary sequence). Thus,

$$\lim_{N \rightarrow \infty} c(N) = b(N) \equiv \frac{N}{\log_\alpha(N)} \quad (6.31)$$

where \equiv denotes identity and $c(N)$ can be normalized via $b(N)$:

$$C(N) = \frac{c(N)}{b(N)} \quad (6.32)$$

$C(N)$, the normalized LZC complexity, reflects the arising rate of new patterns along with the sequence, capturing its temporal structure. Larger values correspond to more complexity (Zhang et al., 2001).

Sample Entropy

In 1991, ApEn was introduced to assess the irregularity of biomedical recordings by evaluating the appearance of repetitive patterns in the data (Pincus, 2001). This statistic counts each sequence as matching itself to avoid the occurrence of $\log(0)$ in the computations. Thus, this irregularity estimation is biased (Richman and Moorman, 2000). To reduce this bias, SampEn was de-

veloped as a modification of ApEn (Richman and Moorman, 2000). Likewise ApEn, SampEn is an irregularity metric that associates higher values with more irregular signals. It can provide information about changes in the regularity of brain local synchronizations (Abásolo et al., 2006c; Hornero et al., 2009b). In addition to the number of samples in the signal, N , this metric has two input parameters: a run length m (here m is a different parameter from the number of measurements in the BSS) and a tolerance window r (Richman and Moorman, 2000).

SampEn(m, r, N) is the negative logarithm of the conditional probability that two sequences similar for m points remain similar at the next point, where self-matches are not included in the calculation. Thus, a lower value of SampEn also indicates more self-similarity in the time series. SampEn is largely independent of record length and displays relative consistency under circumstances where ApEn does not (Richman and Moorman, 2000).

Formally, given a time series $\{x(n)\} = \{x(1), x(2), \dots, x(N)\}$, to define SampEn, one should follow these steps (Richman and Moorman, 2000):

1. Form m -vectors $X_m(1), X_m(2), \dots, X_m(N - m + 1)$ following the procedure defined in the first step of the algorithm for the computation of ApEn. The distance between $X_m(i)$ and $X_m(j)$ is defined as in Eq. (6.20).
2. For a given $X_m(i)$, count the number of j ($1 \leq j \leq N - m, j \neq i$), denoted as B_i , such that the distance between $X_m(i)$ and $X_m(j)$ is less than or equal to r . Then, for $1 \leq i \leq N - m$,

$$B_i^m(r) = \frac{1}{N - m - 1} B_i \quad (6.33)$$

3. Define $B^m(r)$ as:

$$B^m(r) = \frac{1}{N - m} \sum_{i=1}^{N-m} B_i^m(r) \quad (6.34)$$

4. We increase the dimension to $m + 1$ and calculate A_i as the number of

$X_{m+1}(i)$ within r of $X_{m+1}(j)$, where j ranges from 1 to $N - m$ ($j \neq i$). We then define $A_i^m(r)$ as:

$$A_i^m(r) = \frac{1}{N - m - 1} A_i \quad (6.35)$$

5. We set $A^m(r)$ as:

$$A^m(r) = \frac{1}{N - m} \sum_{i=1}^{N-m} A_i^m(r) \quad (6.36)$$

Thus, $B^m(r)$ is the probability that two sequences will match for m points, whereas $A^m(r)$ is the probability that two sequences will match for $m + 1$ points.

6. We define SampEn by:

$$\text{SampEn}(m, r, N) = -\ln \left[\frac{A^m(r)}{B^m(r)} \right] \quad (6.37)$$

It is imperative to consider ApEn (m, r, N) and SampEn (m, r, N) as families of parameters: comparisons are intended with fixed values of m , r and N . Although m and r are critical in determining the outcome of ApEn and SampEn, no guidelines exist for optimizing their values. In principle, the accuracy and confidence of the entropy estimate improve as the number of matches of length m and $m + 1$ increases. The number of matches can be increased by choosing small m (short templates) and large r (wide tolerance) (Escudero et al., 2006a). However, there are penalties for too relaxed criteria (Pincus, 2001). For smaller r values, one usually achieves poor conditional probability estimates, while for larger r values, too much detailed system information is lost. It is convenient to set the tolerance as r times the SD of the original data sequence. In this study, SampEn was estimated with $m = 1$ and $r = 0.25$ times the SD of the original time series (Abásolo et al., 2006c; Richman and Moorman, 2000).

6.7 Extraction of Brain Rhythms With cBSS

This part of the Doctoral Thesis attempts to adaptively extract, using a cBSS, rhythmic activity related to the main frequency bands δ , θ , α and β from diverse scalp regions.

Instead of setting a priori the boundaries of these spectral bands, an EMD (Huang et al., 1998) will be used to decompose the activity recorded at each channel into rhythmical IMFs. Afterwards, all the IMFs computed from all channels in the same region are fed into a clustering procedure to automatically select a reference for every brain rhythm (δ , θ , α and β). This reference is employed with a cBSS to extract activity simultaneously from all channels of the region. Finally, the $\text{Coh}(f)$ is calculated for each pair of extracted activities to assess the level of brain connectivity in AD patients versus healthy elderly subjects.

6.7.1 Preprocessing

The following preprocessing was applied to the signals before proceeding to the EMD, cBSS and calculation of $\text{Coh}(f)$.

Firstly, as it was explained in Section 5.3, and similarly to Section 6.6, the MEG recordings were divided into epochs of 10 s and only the fragments of the signal with minimal artefactual activity were selected for further analysis. This was done in order to avoid surplus complexity in the analysis.

Additionally, it is important to note that the CA always appears in the MEG activity (Escudero et al., 2007b; Jousmäki and Hari, 1996). Furthermore, due to the ability of the EMD to isolate specific rhythms in natural data (Huang et al., 1998) — including biomedical signals (Wu and Hu, 2006; Ye et al., 2007) — the QRS complexes of the cardiac activity may spread across the mid- to high-frequency IMFs computed with the EMD (Blanco-Velasco et al., 2008; Tang et al., 2008). Hence, in order to avoid the cardiac contamination of the IMFs that must serve as references to extract the rhythmic brain activity, we reduced the CA in the MEG signals by means of a cBSS procedure as explained in Section 6.5.1.

6.7.2 Empirical Mode Decomposition

In this Doctoral Thesis, EMD was applied to obtain the rhythmic components that compose each of the MEG channels recorded.

EMD (Huang et al., 1998) is a non-linear technique to *adaptively* represent non-stationary signals as sum of their IMFs. EMD considers the oscillations in signals at a very *local* level. Each resulted IMF by the EMD method satisfies two basic conditions (Huang et al., 1998):

1. In the complete data set, the number of extrema and the number of zero crossings must be the same or differ at most by one.
2. At any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.

The EMD of the signal $x(t)$ can be computed following the next algorithm (Flandrin et al., 2004):

1. Set $g_1(t) = x(t)$.
2. Detect the extrema (both maxima and minima) of $g_1(t)$.
3. Generate the upper and lower envelopes $e_m(t)$ and $e_l(t)$, respectively, by connecting the maxima and minima separately with cubic spline interpolation.
4. Determine the local mean as:

$$m(t) = \frac{e_m(t) + e_l(t)}{2}. \quad (6.38)$$

5. The IMF should have zero local mean. Thus, subtract $m(t)$ from the original signal as: $g_1(t) = g_1(t) - m(t)$.
6. Decide whether $g_1(t)$ is an IMF or not by checking the two basic conditions described above.
7. Repeat steps 2 to 6 and end when an IMF $g_1(t)$ is obtained.

When the first IMF is derived, set $d_1(t) = g_1(t)$, which is the smallest temporal scale (higher frequencies) in $x(t)$ (Flandrin et al., 2004). In order to find the remaining IMFs, generate the residue $r_1(t)$ of the data by subtracting $d_1(t)$ from the signal as: $r_1(t) = x(t) - d_1(t)$ (Flandrin et al., 2004).

$r_1(t)$ is then considered as the new data and subjected to the same sifting process described above (Flandrin et al., 2004). The sifting process is continued until the final residue is a constant, monotonic function, or a function with one maxima and one minima from which no more IMFs can be derived (Flandrin et al., 2004).

At the end of the decomposition, the unidimensional signal $x(t)$ is represented as:

$$x(t) = \sum_{p=1}^M d_p(t) + r_M(t), \quad (6.39)$$

where M is the number of IMFs and $r_M(t)$ is the final residue.

In this Doctoral Thesis, the stopping criteria proposed by Rilling et al. (2003) was used.

6.7.3 Frequency Characterisation of the IMFs

The EMD was applied to adaptively extract the rhythms composing each MEG channel. Since the objective of this part of the Doctoral Thesis is to analyse the brain connectivity in different spectral bands, it is necessary to characterise the spectral content of each IMF.

In order to do so, the instantaneous frequency, $\omega(t)$, of each IMF was computed as (Huang et al., 1998):

$$\omega(t) = \frac{d\theta(t)}{dt}, \quad (6.40)$$

where $\theta(t)$ is the phase of the analytic signal, $z(t) = a(t) \exp[i\theta(t)] = x(t) + iy(t)$, computed from each IMF by means of the Hilbert Transform (Huang et al., 1998).

Afterwards, a weighted average value of the instantaneous frequency, $\langle \omega \rangle$, was computed from each $\omega(t)$ (and thus from each IMF) considering the

corresponding amplitude following:

$$\langle \omega \rangle = \frac{E \{a(t) \omega(t)\}}{E \{a(t)\}}. \quad (6.41)$$

$\langle \omega \rangle$ represents the average frequency of each IMF and is the feature used to characterise the spectral content of every IMF.

6.7.4 Regional Extraction of Brain Rhythms

The EMD allowed us to obtain rhythmic activities from each MEG channel. Then, the IMFs were characterised with their $\langle \omega \rangle$ values.

Nevertheless, it must be noticed that the MEG recordings are composed of a large number of channels (148 in this Doctoral Thesis). Hence, it may be helpful to group the channels in regions (Poza et al., 2008b; Stam et al., 2006). The regions considered in Doctoral Thesis are shown in Fig. 6.6. This division is similar to that made by Gómez et al. (2009b) and Poza et al. (2008b), but it also considers the difference between left and right hemispheres like Stam et al. (2006). Midline sensors were not used (Stam et al., 2006). Hence, MEG channels were grouped into: Left Central Region (LeftC), Right Central Region (RightC), Left Anterior Region (LeftA), Right Anterior Region (RightA), Left Lateral Region (LeftL), Right Lateral Region (RightL), Left Posterior Region (LeftP) and Right Posterior Region (RightP).

The IMFs are representative of the oscillating brain activity recorded at a particular channel, but not necessarily of the brain rhythms over a specific region. Therefore, a cBSS technique has been applied to extract the brain activity corresponding to a particular region (see Fig. 6.6) for a specific frequency band — δ (1 Hz to 4 Hz), θ (4 Hz to 8 Hz), α (8 Hz to 13 Hz) and β (13 Hz to 30 Hz). In order to perform the extraction, the cBSS requires a suitable reference signal (Huang and Mi, 2007; James and Gibson, 2003). This reference signal is selected by means of a k -means clustering procedure applied to the average frequency of all IMFs from a region.

The clustering procedure is as follows. Firstly, for each region, select the values of $\langle \omega \rangle$ ranging from 1 Hz to 30 Hz. Then, these $\langle \omega \rangle$ are fed

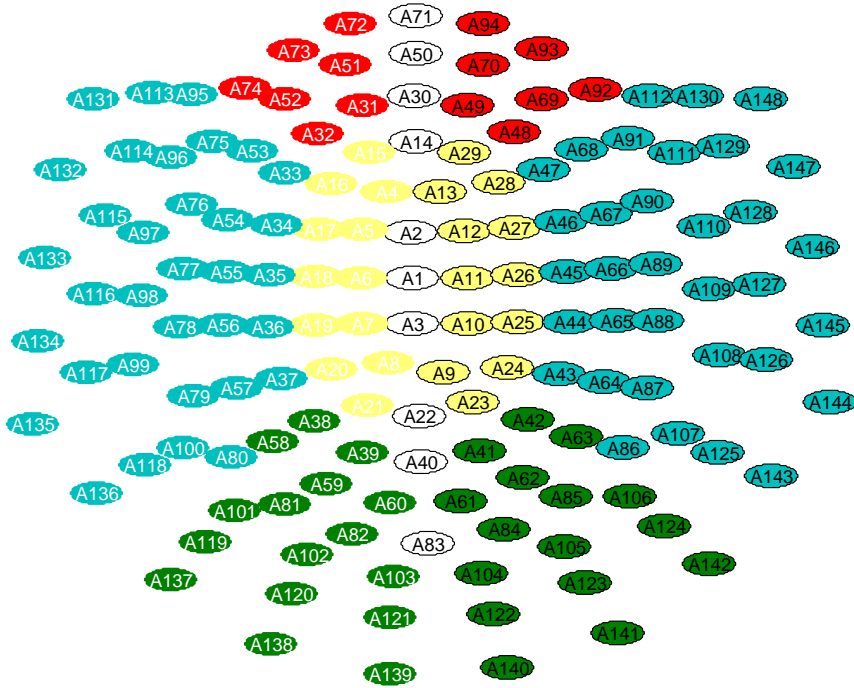


Figure 6.6: Distributions of the MEG sensors into central (yellow), anterior (red), lateral (blue) and posterior (green) regions for left (white text) and right (black text) hemispheres. The midline sensors are marked in white.

into a k -means clustering procedure which is initialised with four centroids located in the middle of the four spectral bands considered (δ , θ , α and β). The outcomes of the clustering showed that the $\langle\omega\rangle$ values are assigned to disjoint clusters, each of which can be identified with one of the four spectral bands. Finally, the IMF corresponding with the $\langle\omega\rangle$ value closest to the centroid of each cluster is selected as the reference for a cBSS applied to all MEG channels from the region in order to extract activity representative of that spectral band.

The k -means algorithm is an effective and relatively simple clustering procedure (Hartigan and Wong, 1979). Hence, it has been applied in diverse applications (Sanei and Chambers, 2007). This algorithm divides a set of features (the values of $\langle\omega\rangle$ in this case) into k clusters automatically and without the need of supervision (Hartigan and Wong, 1979; Sanei and Chambers, 2007).

The algorithm is initialised by setting k to the expected number of clusters, which in our case is four corresponding to the four spectral bands: δ , θ , α and β . The centre of every cluster has been initialised to the medium frequency in each band. After this initialisation, the k -means clustering assigns the data points to the closest cluster centre. Then, new cluster centres are calculated. This is the essence of the algorithm: once the clusters have been established (i.e., each data point has been assigned to its closest cluster centre), the geometric centre of each cluster is recomputed. These steps are repeated until convergence is reached (Hartigan and Wong, 1979; Jobson, 1991b; Sanei and Chambers, 2007).

6.7.5 Coherence

Once the cBSS had extracted the brain activity related to each band (δ , θ , α and β) from every region, the $\text{Coh}(f)$ was used to measure the synchrony. The $\text{Coh}(f)$ function quantifies linear correlations as a function of the frequency (Dauwels et al., 2010; Sanei and Chambers, 2007). This measure can detect linear synchronisation between two signals, but it does not discriminate the directionality of the coupling (Sanei and Chambers, 2007).

In order to compute $\text{Coh}(f)$, the two signals x and y are subdivided in M segments of equal length L . The $\text{Coh}(f)$ is calculated by averaging over those segments as (Dauwels et al., 2010):

$$\text{Coh}(f) = \frac{|\langle X(f) Y^*(f) \rangle|}{|\langle X(f) \rangle| |\langle Y(f) \rangle|}, \quad (6.42)$$

where $X(f)$ and $Y(f)$ denote the Fourier Transform of $x(t)$ and $y(t)$, in that order, and $*$ indicates the complex conjugate. $|\cdot|$ and $\langle \cdot \rangle$ stand for magnitude and average computed over the M segments, respectively (Dauwels et al., 2010). Experiments with EEG data suggest that this parameters is strongly correlated with other commonly used synchronisation measures (Dauwels et al., 2010).

6.8 Statistical Analysis Tools

Diverse statistical analysis tools have been used in this Doctoral Thesis to assess the results. They are briefly described in the following lines grouped in three main categories.

6.8.1 Visual Information Tools

A useful tool to obtain a visual summary of a data distribution is the boxplot. This diagram is composed of a box with three horizontal lines at the lower quartile, median and upper quartile values. Hence, the median of the distribution is represented as the line which appears, approximately, in the middle of the box. Usually, the confidence interval of the median is indicated with a couple of notches. For two or more boxplots, if there is no overlap between these notches, there are significant differences between the corresponding data distributions. The boxplot also has two whiskers, which are lines extending from each end of the boxes to show the extent of the rest of the data. Their maximum length is usually estimated as 1.5 times the interquartile range. Values beyond the end of the whiskers are considered outliers, which are marked with a '+' (Tukey, 1977). This tool has been used in the study of the results derived from the Section 6.2.1.

6.8.2 Statistical Significance Tests

A Student's t -test was used in the analysis of the results from Section 6.6 to decide which BSS components provided the most significant difference between subject groups. This statistical test evaluates the null hypothesis that the means of two independent and normally distributed samples are equal. This procedure also assumes that the variances of the two populations are equal (Jobson, 1991a). Additionally, the variance of a quantitative predictor (covariate) can be removed from the samples by a regression in order to account for some variability of the samples and increase the statistical power (Jobson, 1991a).

As a generalisation of the Student's t -test, the ANOVA allows us to test

whether the means of several groups are all equal. Moreover, this procedure offers the possibility of partitioning the observed covariance in the data into components due to diverse explanatory variables (Jobson, 1991a). Similarly to the Student's t -test, ANOVA assumes independence of cases, normality of the distributions and homogeneity of variances. ANOVA splits the total sum of squares of the data into components related to the effects included into the model (Jobson, 1991a). In Section 6.7, a one-way, repeated-measures ANOVA was done with Greenhouse-Geisser corrected degrees of freedom to correct the lack of sphericity. The variable 'Diagnosis' was taken as the inter-subject factor while the variables 'Band' and 'Pair of regions' were considered intra-subject factors. Special attention was paid to the estimated marginal means corresponding to each factor (Searle et al., 1980). A significance level of $p < 0.05$ was used.

6.8.3 Classification Analysis

In the classification analyses, results are showed in terms of sensitivity, specificity and accuracy. Specificity was defined as the percentage of healthy subjects correctly detected and sensitivity represented the proportion of all AD patients for whom the test was positive. Accuracy denoted the total fraction of subjects well recognised.

A LDA may be useful to perform a classification analysis of the observations. On the basis of information derived from a known sample distribution, one or more discriminant functions are estimated to classify unknown data of the same type. Under the assumption of multivariate normality with common covariance matrix, an observation is assigned to the group whose Mahalanobis distance from the observation is smallest. This criterion is equivalent to classify a sample as belonging to the group with the largest likelihood value (Jobson, 1991b). It is important to note that for discrete data in small samples this criterion performs as well or better than methods based on various discrete distributions. For continuous data, transformations to normality do not usually improve classification error rates and, in some cases, the error rates are even larger than for non-transformed data (Jobson, 1991b). Clas-

sification analysis using LDA have been carried out to analyse the results obtained from Section 6.6 and Section 6.7.

Additionally, it is important to decide whether all variables are necessary to obtain a good classification. For instance, the larger the number of variables included in the LDA classification rule, the larger must be the sample size to achieve the same level of precision (Jobson, 1991b). Hence, statistical tests can be employed to test the hypothesis that the Mahalanobis distance between two groups is the same for a subset of $q < p$ of the variables as it is for the full set of p variables (Jobson, 1991b). These tests can offer a criterion to determine if a particular variable should enter or exit the discriminant rule in a stepwise way. This stepwise procedure was applied to select the variables included in the classification studies of Section 6.7.

ROC curves (Fawcett, 2006) are helpful to analyse and visually assess the performance of classifiers. They are two-dimensional graphs in which true positive rate (i.e., sensitivity) is plotted against the false positive rate (computed as one minus the specificity). The ROC curves depicts the trade-off between true positives and false positives for a classifier (Fawcett, 2006). It may be useful to summary the ROC performance to a single scalar value in order to compare different classifiers. This can be done by calculating the Area Under the ROC Curve (AUC) (Fawcett, 2006). The AUC is a fraction of the area of the unit square. Hence, its value ranges between 0 and 1.0. However, due to the fact that completely random classifiers produces a diagonal with an associated AUC of 0.5, no realistic classifier should have an AUC smaller than 0.5 (Fawcett, 2006). It is important to note that the AUC of a classifier is equivalent to the probability that the classifier will rank a randomly chosen positive instance higher than a randomly chosen negative instance. This is equivalent to the Wilcoxon test of ranks (Fawcett, 2006). The AUCs of diverse classifiers are given in Section 7.4 and Section 7.5.

In the analyses carried out in Section 6.6, the subject population was divided randomly into a training set and a test set (see Table 5.3). The training set was used to develop the BSS preprocessing and to find the classification rules for each case. Then, these rules were applied, without further modification, to the test set to independently assess the improvement in the

separation between AD patients and control subjects due to the BSS preprocessing in order to avoid over-estimation of the classification rate (Escudero et al., 2009b; Jobson, 1991b).

A leave-one-out cross-validation procedure was used to avoid the appearance of over-fitting and bias in the classification analysis performed with the results of Section 6.7. The leave-one-out cross-validation classifies each single case using the decision rule derived from all remaining data (Jobson, 1991a,b). Then, this process is repeated for all cases. Although this procedure typically reduces the sensitivity, specificity and accuracy values, it increases the reproducibility of the results (Cichocki et al., 2005).

Chapter 7

Results

Contents

7.1	Preprocessing	182
7.1.1	Synthetic Data	182
7.1.2	Real MEG Recordings	185
7.2	Evaluation of the Consistency	187
7.3	Artefact Removal	193
7.3.1	Artefacts in the Original Recordings	193
7.3.2	Blind Source Separation	198
7.3.3	Artefact Detection	198
7.3.4	Evaluation of the Artefact Removal	202
7.4	Feature Enhancement	221
7.4.1	Preliminary Description of the Analysis	221
7.4.2	Decompositions and Ordering Criteria	222
7.4.3	Selection of the Most Significant Components	233
7.4.4	Evaluation of the Classification Improvement	235
7.5	Adaptive Extraction of Brain Rhythms	241
7.5.1	Estimation of References	241
7.5.2	Descriptive Statistical Analysis	244
7.5.3	Classification Analysis	247

This Chapter reports the results obtained in this Doctoral Thesis. The structure of the Chapter follows quite closely the description of the methodology presented in Chapter 6. Hence, the results obtained from the testing of the preprocessing with synthetic data are first presented. The Chapter continues with the results derived when the preprocessing was applied to the real MEG signals. Afterwards, the measurements of similarity between BSS algorithms are reported. This Chapter also contains the results obtained for the automatic evaluation of the artefact removal based on BSS. Next, the classification improvements gained from the use of BSS techniques as a preprocessing are reported. Finally, results about the extraction of brain rhythms over specific scalp regions and the connectivity analysis between them are given.

7.1 Preprocessing

This part of the dissertation details the results obtained from the application of the processing presented in Section 6.2 to the synthetic signals and the real dataset of control subjects' recordings.

7.1.1 Synthetic Data

In order to evaluate the degree to which the preprocessing (Cao et al., 2003; Escudero et al., 2007b; Ikeda and Toyama, 2000) was able to accurately estimate the number of inner components and the level of the additive noise added to the mixtures, 1000 different synthetic mixtures were created with diverse levels of external additive noise. Then, the preprocessing was applied ten times to each of these mixtures and the results were averaged for every dataset (Escudero et al., 2007b). The average SNR of the simulated mixed signals originated from the 11 inner synthetic components ranged between -14 dB and 22 dB.

First of all, the performance of the preprocessing to find out the correct number of inner components was analysed (Escudero et al., 2007b). The scheme using the MDL was compared against three PCA-based ap-

proaches (Escudero et al., 2007b):

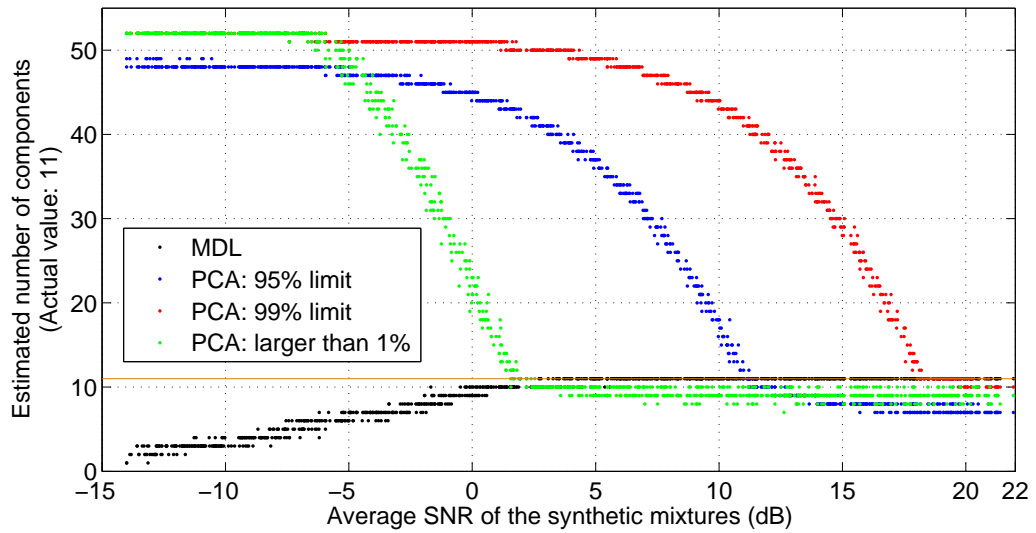
- To set n to the number of eigenvalues of the data covariance that explain a certain part of the total power (Escudero et al., 2007b; James and Hesse, 2005; Ting et al., 2006). This fraction was set to 95% (“PCA: limit 95%”) and 99% (“PCA: limit 99%”).
- To consider those components that individually account for at least 1% of the total variance of the data (Escudero et al., 2007b; James and Hesse, 2005). This procedure was denoted as “PCA: larger than 1%”.

The results of this comparative analysis appear in Fig. 7.1. As it could be expected, the criteria based on setting a threshold on the cumulative value of explained variance over-estimate n when the noise power is not negligible (Escudero et al., 2007b). For instance, the criterion “PCA: limit 99%” over-estimated n when the noise power produces average SNRs lower than 17 dB. The criterion “PCA: limit 95%” offers more accurate estimations of this parameter for SNRs larger than 11 dB, but even in this case it does not provide the correct number of components. What is more, below that value of 11 dB, this criterion behaves very similarly to “PCA: limit 99%”.

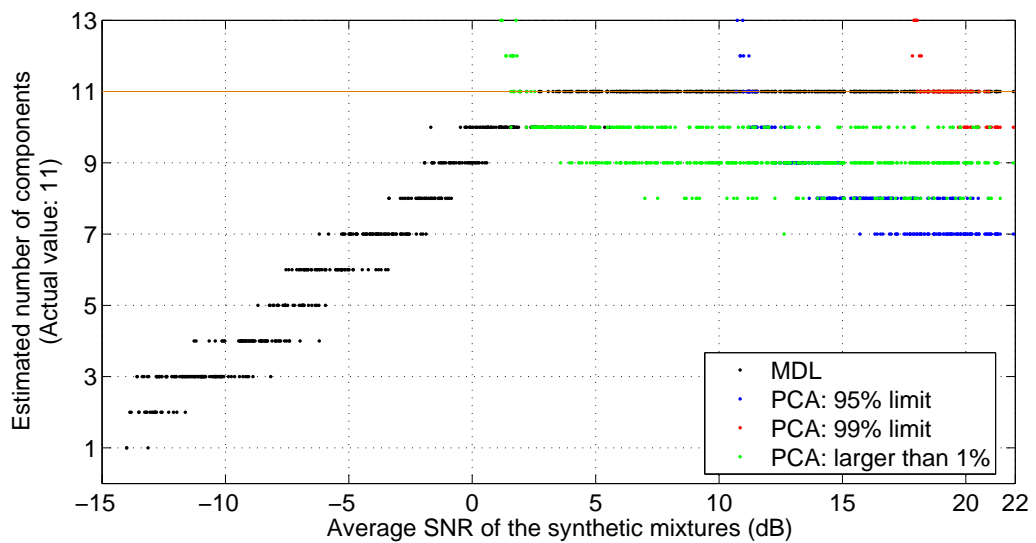
The MDL criterion gives accurate estimations of the number of components that originated the mixtures for SNR larger than 3 dB. For higher noise power, and contrary to the PCA-based criteria, MDL tends to underestimate n .

The criterion “PCA: larger than 1%” provides values of n around nine or ten (close to the real value of 11) for average SNR values larger than 3 dB. Nevertheless, for $\text{SNR} < 3$ dB, this criterion assigns extremely high values to n , similarly to the other PCA-based criteria (Escudero et al., 2007b).

Additionally, the error in the estimation of the external noise power, E_{norm} , was evaluated with Eq. (6.9) (Escudero et al., 2007b). Fig. 7.2 depicts the average value of E_{norm} computed from the ten applications of the preprocessing to each of the 1000 realisations of the synthetic mixtures. Here, the number of components (n) was set to the value that minimised the MDL (Escudero et al., 2007b; Ikeda and Toyama, 2000). It can be seen that, in most cases, the normalised error was lower than 3.5%.



(a) Estimated number of components for all criteria and SNR values considered in this study.



(b) Detail of the results obtained for MDL.

Figure 7.1: Estimated n values for each of the 1000 synthetic mixtures according to the tested criteria.

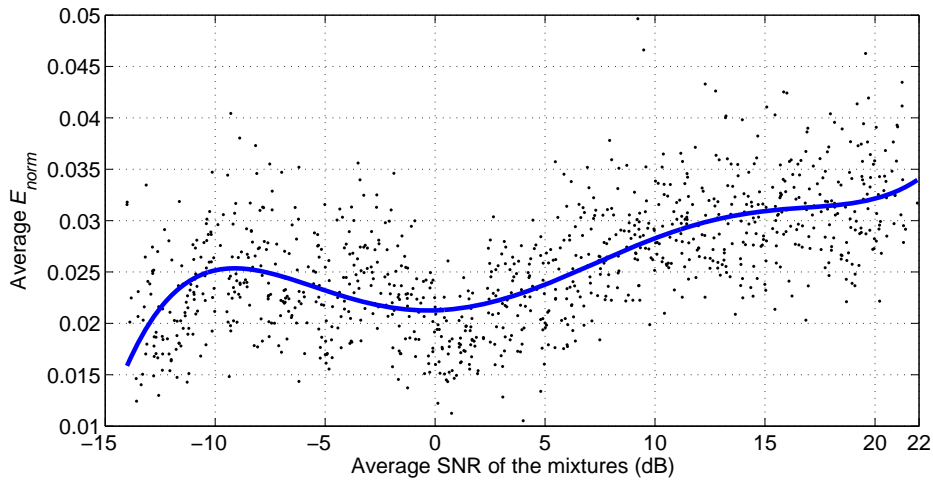


Figure 7.2: Average E_{norm} for diverse average SNR values in the synthetic mixtures.

7.1.2 Real MEG Recordings

The preprocessing was applied to the real MEG recordings acquired from the 26 control subjects. These signals were split into epochs of 10 s, 20 s, 40 s, 60 s and 90 s. In each case, the optimum number of components (n) and the external noise power were estimated. These results are presented in this Section. These preprocessed data are the starting point of the analyses and results about the comparison of BSS algorithms presented in Section 7.2 and the artefact rejection performance summarised in Section 7.3.

Firstly, the value of n was estimated for each case with the MDL (Escudero et al., 2007b; Ikeda and Toyama, 2000). Fig. 7.3 shows the boxplots representing the distributions of n for every epoch length. As it can be expected, n tended to increase with the epoch length (Escudero et al., 2009d).

For the previously computed n values, the power of the external noise was computed. The results, in terms of normalised power (power assigned to the additive noise divided by the total power in the data) are depicted in Fig. 7.4. As it can be seen, the interquartile range of the additive term power represents about 2.8% to 5.6% of the total power in the MEGs. This implies that the SNR of the recordings usually lay between 15.4 dB and 12.3 dB.

Finally, it must be noted that the preprocessing implies a dimensionality

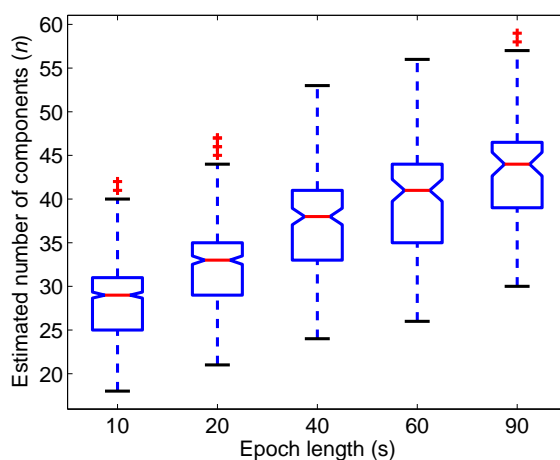


Figure 7.3: Boxplots showing the number of components (n) estimated for the epoch lengths considered in this study.

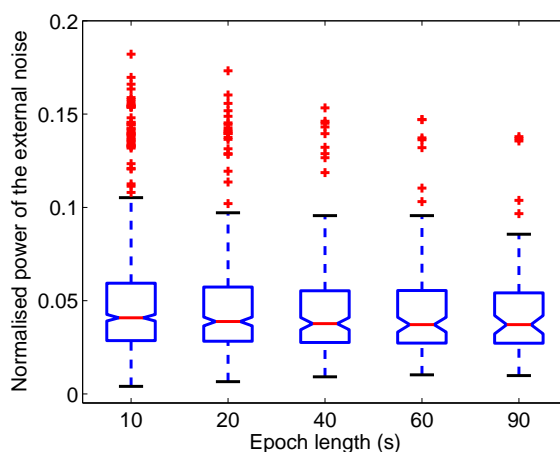


Figure 7.4: Normalised external noise power in real MEG signals as modelled by the preprocessing.

reduction caused by the selection of a value for $n < m = 148$ (Escudero et al., 2007b). The additive vector $\mathbf{n}(t)$ can be interpreted as a specific error term in the model (6.4). Hence, the objective is to retain the maximum amount of energy associated by the processing with the inner components for the BSS analysis. Hence, considering the previously estimated n values, Fig. 7.5 illustrates the fraction of energy retained. It is clear that almost all power attributed to the underlying components will be considered in the following steps.

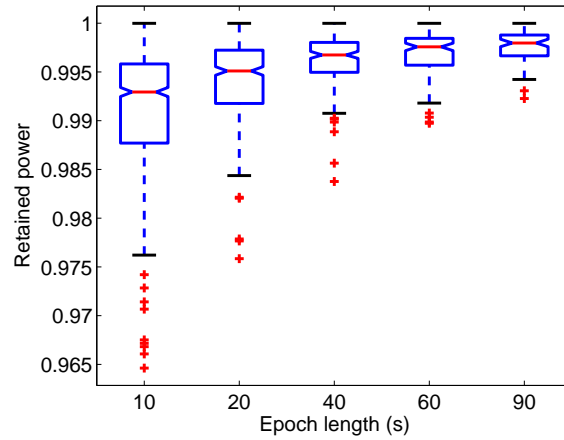


Figure 7.5: Inner power retained for BSS by selecting the optimum number of components (n) estimated by the MDL.

7.2 Evaluation of the Consistency

In this Section of the Doctoral Thesis, the preprocessed MEG signal epochs with the optimum n values were decomposed with five BSS algorithms: AMUSE, SOBI, JADE, eInfoMax and FastICA. Similarly to Section 7.1.2, epoch lengths equal to 10 s, 20 s, 40 s, 60 s and 90 s were considered.

After the BSS, the matrices \mathbf{P}^{qr} — see Eq. (6.10) — were computed for each epoch and pair of algorithms. Then, they were characterized with the metric F defined in Eq. (6.11). In order to reduce the amount of data to be analysed, only the matrix \mathbf{P}^{qr} was studied instead of both \mathbf{P}^{qr} and \mathbf{P}^{rq} (Escudero et al., 2009d). This decision was supported by the fact that the average absolute differences for the F metric between \mathbf{P}^{qr} and \mathbf{P}^{rq} were always lower than 1.2%.

For each epoch length, the F values obtained for every pair of algorithms were averaged. These results are depicted in Fig. 7.6, where all subplots are represented with the same colour range. As it was previously explained in Section 6.4.1, lower F values are related to more consistent (i.e., more similar) pairs of algorithms (Escudero et al., 2009d). For all epoch lengths, the most consistent pair of algorithms was AMUSE–SOBI (SOS-based methods), followed by the pair JADE–FastICA, which involve HOS. Moreover, Fig. 7.6 shows that the general level of consistency improved as the epoch length in-

creased. This means that the separations provided by different algorithms tended to converge as larger signals were decomposed.

For each epoch length, a one-way ANOVA analysis of variance with the Scheffé's multiple comparison procedure, 'Pair of algorithms' as the grouping factor and 'Number of estimated components' (n) as a covariate was used to statistically evaluate the differences in the F values.

For an epoch length equal to 10 s, there were significant differences in the F values as a consequence of the grouping factor 'Pair of Algorithms', the covariate 'Number of estimated components' and their interaction ($p \ll 0.0001$ in all cases). Additionally, the slopes of the regression of F against n were significantly different from 0 ($p < 0.05$) for the pairs of algorithms 'AMUSE-SOBI', 'AMUSE-eInfoMax' and 'SOBI-eInfoMax'. Whereas in the first case F slightly increased with n (indicating that a larger number of components made the outcomes of the decompositions more different), for the other two pairs, more components produced lower F . Finally, the multiple comparison procedure confirmed that the level of consistency of the 'AMUSE-SOBI' pair was significantly lower from that of 'JADE-FastICA', and that the F values for these two pairs also differed significantly from the other eight pairs. Of note is that 'AMUSE-JADE' and 'SOBI-JADE' offered the most different separations. All values of F obtained from the 10 s epochs and the regression lines are shown in Fig. 7.7.

The results obtained for epochs of 20 s are very similar to those previously reported for 10 s. There were significant differences in the F values due to the 'Pair of Algorithms', the 'Number of estimated components' and their interaction ($p \ll 0.0001$ in all cases). The regression of F against n had slopes that were significantly positive for the pairs of algorithms 'AMUSE-SOBI' and 'JADE-FastICA' and significantly negative for 'AMUSE-eInfoMax' and 'SOBI-eInfoMax'. By means of the multiple comparison procedure, it was found out that the level of consistency of 'AMUSE-SOBI' was significantly lower from that of 'JADE-FastICA', and that their F values were also significantly different from the other 8 algorithms. Again, the pairs 'AMUSE-JADE' and 'SOBI-JADE' were characterised by providing the most dissimilar decompositions. Fig. 7.8 summarises this analysis graphically.

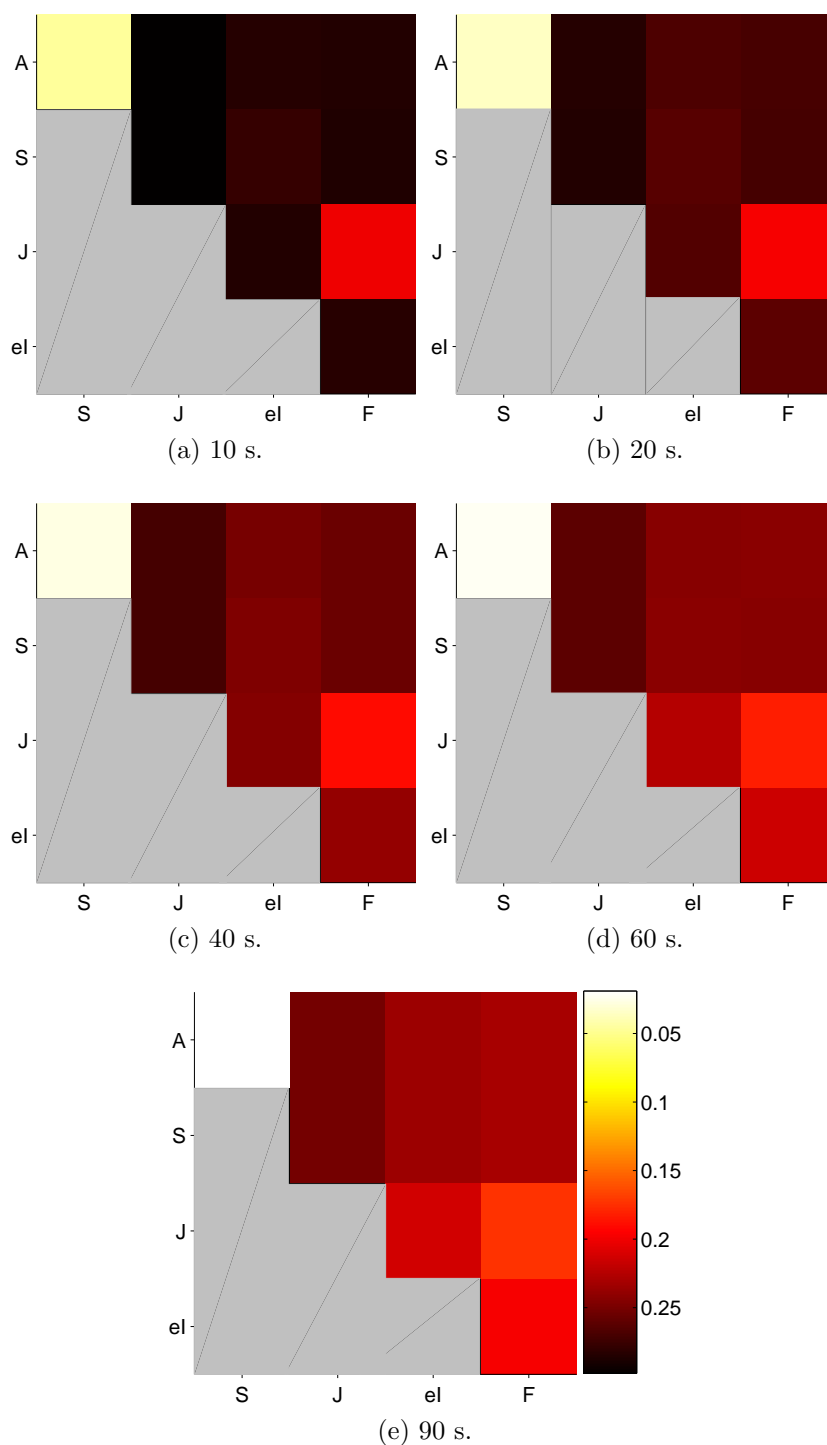


Figure 7.6: Average F values for each pair of BSS algorithms (A: AMUSE, S: SOBI, J: JADE, el: extended Infomax, F: FastICA) and epoch length.

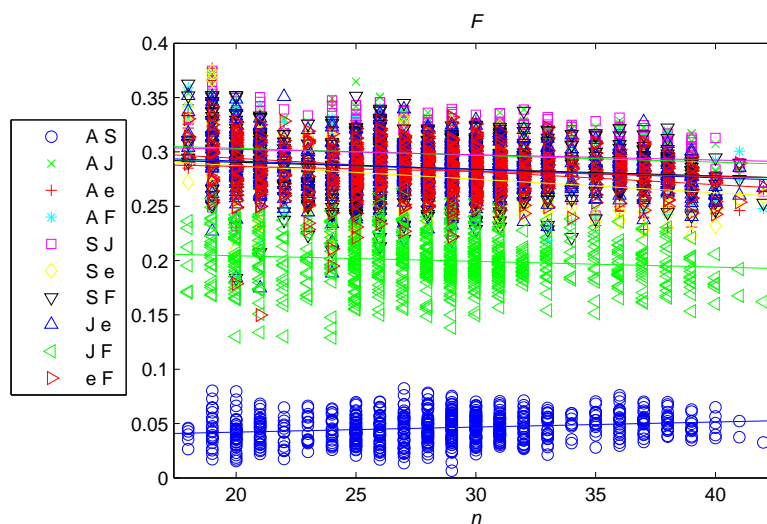


Figure 7.7: ANOVA with n as a covariate for the F metric applied to epochs of 10 s (A: AMUSE, S: SOBI, J: JADE, e: extended Infomax, F: FastICA).

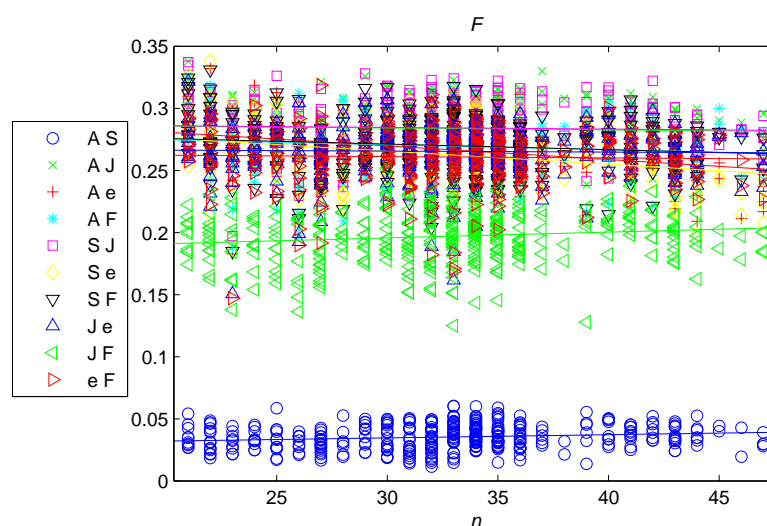


Figure 7.8: ANOVA with n as a covariate for the F metric applied to epochs of 20 s (A: AMUSE, S: SOBI, J: JADE, e: extended Infomax, F: FastICA).

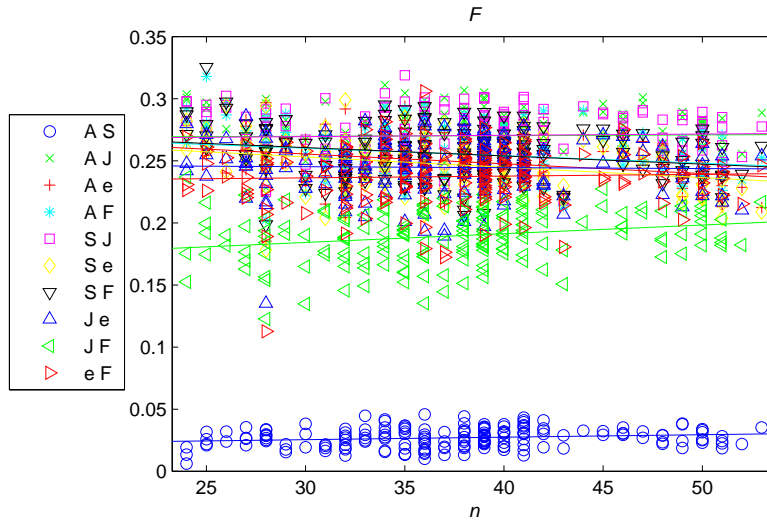


Figure 7.9: ANOVA with n as a covariate for the F metric applied to epochs of 40 s (A: AMUSE, S: SOBI, J: JADE, e: extended Infomax, F: FastICA).

The case of epoch length equal to 40 s presents slight deviations from the results introduced above. The F values also varied significant as a result of the ‘Pair of Algorithms’ ($p \ll 0.0001$), ‘Number of estimated components’ ($p = 0.0100$) and their interaction ($p \ll 0.0001$). In this occasion, the slopes of the regression for ‘AMUSE–eInfoMax’, ‘AMUSE–FastICA’, ‘SOBI–eInfoMax’ and ‘SOBI–FastICA’ decreased with the number of components, whereas the pair ‘JADE–FastICA’ offered less similar separations for larger values of n . The multiple comparison procedure reported the same outcome as in the two previous cases. Fig. 7.9 illustrates this analysis.

When the BSS of epochs of 60 s is studied, the F values only present significant differences for the grouping factor ‘Pair of Algorithms’ and the interaction of this factor with the number of components ($p \ll 0.0001$ in both cases). The slopes that are significantly different from zero are identical to those indicated in the analysis made for epochs of 40 s: the consistency of ‘AMUSE–eInfoMax’, ‘AMUSE–FastICA’, ‘SOBI–eInfoMax’ and ‘SOBI–FastICA’ increased with n , while ‘JADE–FastICA’ produced larger F values for more components. Once again, the multiple comparison procedure suggested that ‘AMUSE–SOBI’ and ‘JADE–FastICA’, in that order, are the most consistent pairs of algorithms. It also indicated that ‘AMUSE–JADE’

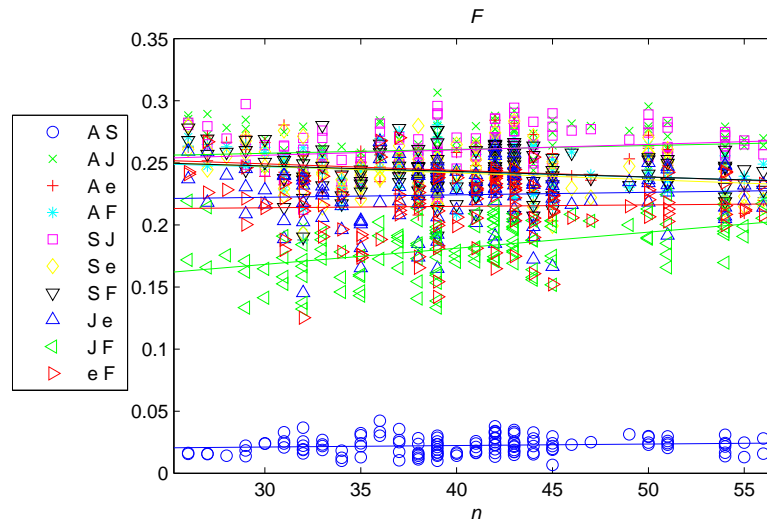


Figure 7.10: ANOVA with n as a covariate for the F metric applied to epochs of 60 s (A: AMUSE, S: SOBI, J: JADE, e: extended Infomax, F: FastICA).

and ‘SOBI–JADE’ provided the least similar signal decompositions. This information is depicted in Fig. 7.10.

Finally, the decompositions computed for epochs of 90 s were studied. Similarly to the 60 s case, only the grouping factor ‘Pair of Algorithms’ and the interaction of this factor with the number of components have significant p values in the ANOVA ($p \ll 0.0001$). The ‘JADE–FastICA’ and ‘JADE–eInfoMax’ tended to offer more different BSSs when more components were estimated ($p < 0.05$). Only the pairs ‘AMUSE–eInfoMax’, ‘AMUSE–FastICA’ and ‘SOBI–eInfoMax’ had regression slopes that were significantly ($p < 0.05$) lower than zero. For this epoch length, ‘AMUSE–JADE’ and ‘SOBI–JADE’ computed the least consistent separations. However, four pairs of algorithms had significantly different population marginal means for the F values different from the rest of pairs. These were, from more consistent to more dissimilar separations: ‘AMUSE–SOBI’, ‘JADE–FastICA’, ‘eInfoMax–FastICA’ and ‘JADE–eInfoMax’. These results are depicted in Fig. 7.11.

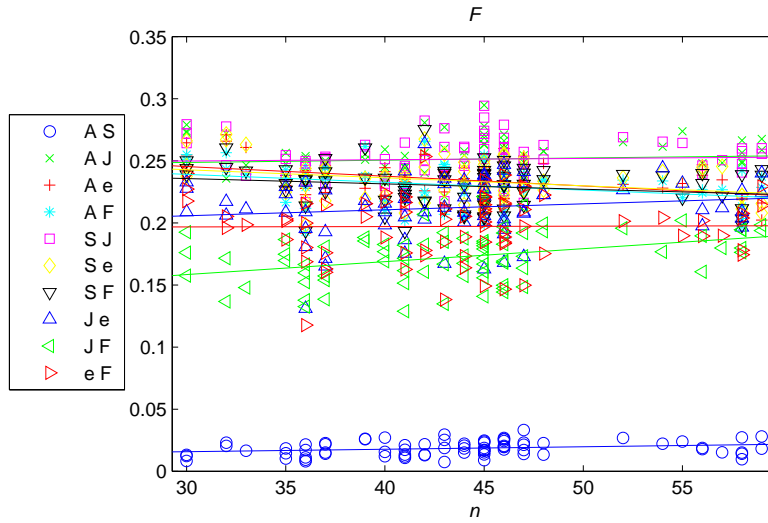


Figure 7.11: ANOVA with n as a covariate for the F metric applied to epochs of 90 s (A: AMUSE, S: SOBI, J: JADE, e: extended Infomax, F: FastICA).

7.3 Artefact Removal

This Section of the Doctoral Thesis presents the results compiled to objectively and automatically evaluate the removal of CA, OA and PLN based on BSS techniques. Similarly to the consistency analysis performed in Section 7.2, signal epochs of 10 s, 20 s, 40 s, 60 s and 90 s were considered to be decomposed with the following BSS techniques: AMUSE, SOBI, JADE, eInfoMax and FastICA. Additionally, a cBSS algorithm was used in the CA rejection with a reference of the cardiac activity computed as an average of all MEG channels.

7.3.1 Artefacts in the Original Recordings

First of all, the overall influence of CA, OA and PLN in the recordings was assessed with the measurements described in Section 6.5.2. In this way, it was possible to have *references to objectively compare* the average amount of contamination in the MEG signals before and after the BSS-based procedures (Escudero et al., 2007b). Moreover, this assessment was automatically carried out. The metrics intended to measure those contaminations without relying

Table 7.1: Average values of A_{PtP} and RMS indicating the importance of the CA in the original MEG recordings for the epoch lengths considered in this study. Data are shown as mean \pm SD.

Epoch length	A_{PtP} (pT)	RMS (pT)
10 s	0.756 \pm 0.337	0.122 \pm 0.073
20 s	0.759 \pm 0.340	0.123 \pm 0.073
40 s	0.757 \pm 0.342	0.121 \pm 0.074
60 s	0.744 \pm 0.363	0.119 \pm 0.076
90 s	0.745 \pm 0.364	0.120 \pm 0.077

on any manual marking of visual revision of the recordings. On the one hand, this might limit the reliability of the results in the sense that the quality of the processed signals was not visually reviewed on a one-by-one basis. On the other, this scheme avoided the time-consuming and laborious process of manual labelling and allowed objective comparisons to be performed.

First of all, the presence of the CA was evaluated considering the amplitude and power of the average QRS complex that could be found in the MEG signals. R-peaks were located in the average of all 148 MEG channels before any kind of signal processing (Escudero et al., 2007b). This could be done thanks to the fact the only synchronous activities in the MEG recordings are the CA and the PLN. Then, an average QRS-complex was calculated and the power of this artefact was characterised with its A_{PtP} and RMS values. These results appear in Table 7.1 for the epoch lengths considered in this Doctoral Thesis.

Additionally, in order to illustrate the waveform associated with the CA, Fig. 7.12 plots the average QRS complex computed from all control subjects. The Q, R and S peaks can be clearly observed, whereas the P and T waves are more difficult to make out.

The evaluation of the OAs is more complex since this activity does not always appear in the recordings and it is not as well localised in the spectrum as the PLN. Following a similar procedure to that proposed by Mantini et al. (2008), an amplitude threshold of ± 2 pT was set to count the number of ‘high-amplitude’ peaks in the recordings. Moreover, most of the ocular ac-

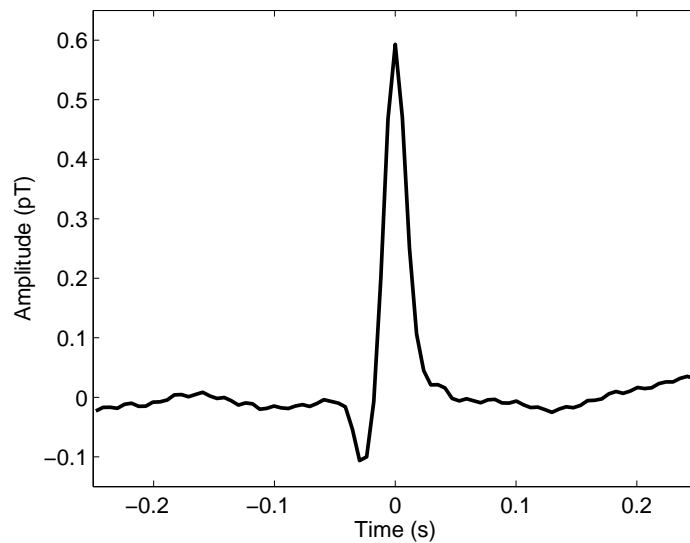
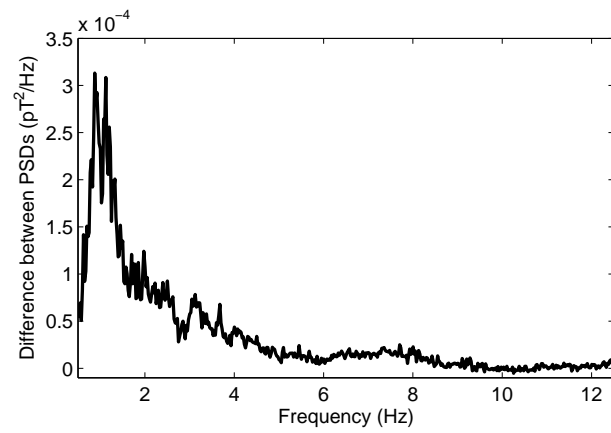


Figure 7.12: Average QRS complex that appears in the original recordings for epoch length equal to 10 s.

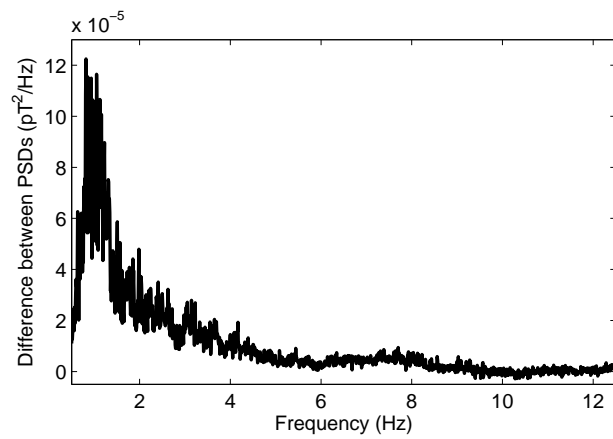
tivity power is located near the eyes. Hence, an average PSD was computed from a subset of 18 MEG channels located over the anterior part of the head and close to the eyes while another average PSD was calculated from the remaining set of 130 channels (see Fig. 6.5). By subtracting the latter PSD from the former, the difference in power at each specific frequency between the activity acquired near the eyes and the average signal recorded over the rest of head is revealed. Then, the total power of this ‘discrepancy’ measure between 0.5 Hz and 6.5 Hz was calculated. Table 7.2 contains the values of these OA-related metrics.

In order to illustrate, and later compare, how much power there is in low frequencies near the eyes in comparison with the sensors located over the rest of the scalp, Fig. 7.13 depicts the amount of power of the previously described ‘discrepancy’ measure for epoch lengths of 20 s, 60 s and 90 s.

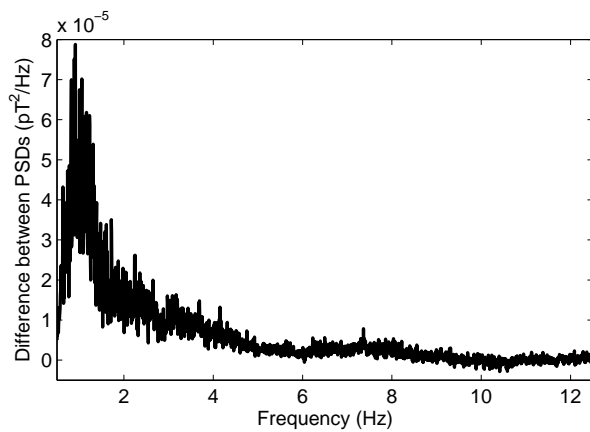
Finally, the calculation of the PLN power contaminating the MEG recordings can be easily done in the frequency domain. Therefore, the power of the PSD and PSD_n between 49 Hz and 51 Hz were computed. These results appear in Table 7.3 and an illustration of this part of the spectra is shown in Fig. 7.14.



(a) Epochs of 20 s.



(b) Epochs of 60 s.



(c) Epochs of 90 s.

Figure 7.13: Median of the difference between the PSDs for the two subsets of MEG channels used to compare the impact of OAs in the recordings.

Table 7.2: Number of peaks in the signals above ± 2 pT and the difference between the PSDs computed at 18 MEG channels close to the eyes and at the rest of MEG sensors indicating the importance of the OA in the original recordings for the epoch lengths considered in this study. Data are shown as median \pm SD.

Epoch length	Peaks above ± 2 pT	Spectral difference (pT^2)
10 s	1303 \pm 19959	0.019 \pm 1.105
20 s	1299 \pm 19925	0.019 \pm 1.016
40 s	1191 \pm 18714	0.019 \pm 1.082
60 s	1300 \pm 19889	0.019 \pm 1.087
90 s	1142 \pm 18118	0.018 \pm 1.276

Table 7.3: Average values of total and normalised spectral power in the band from 49 Hz to 51 Hz indicating the importance of the PLN in the original MEG recordings for the epoch lengths considered in this study. Data are shown as mean \pm SD.

Epoch length	Total power ($\text{pT}^2 \cdot 10^3$)	Normalised power (10^2)
10 s	2.272 \pm 3.803	3.476 \pm 4.758
20 s	2.281 \pm 3.823	3.453 \pm 4.760
40 s	2.299 \pm 3.836	3.426 \pm 4.738
60 s	2.284 \pm 3.835	3.445 \pm 4.776
90 s	2.311 \pm 3.844	3.405 \pm 4.729

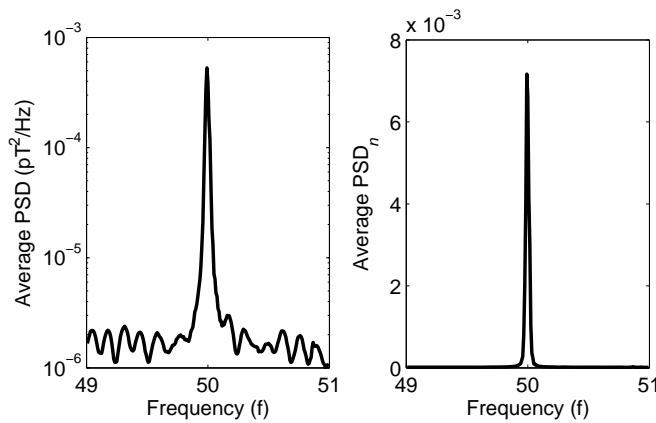


Figure 7.14: Average spectra centred at the PLN frequency in the original signals computed from epochs of 60 s.

7.3.2 Blind Source Separation

First of all, the preprocessing and dimension reduction techniques detailed in Section 6.2 were applied to the real MEG signals divided into epochs of 10 s, 20 s, 40 s, 60 s and 90 s. Once the number of optimum number of inner components (n) had been estimated with the MDL, the algorithms AMUSE, SOBI, JADE, eInfoMax and FastICA were used to decompose the recordings. Afterwards, the underlying components were characterised with the metrics described in Section 6.5.1 and the automatic artefact rejection criteria were applied. Finally, a reconstruction of the MEG signals was calculated by means of Eq. (6.15) (Mantini et al., 2008; Ting et al., 2006).

In addition to the procedure described in the previous paragraph, a reference signal following the R-peaks that appear in the MEG activity was fed into a cBSS algorithm to extract the cardiac contamination in the recordings and then cancel it out by subtracting its projection from the MEG channels (James and Gibson, 2003).

7.3.3 Artefact Detection

The labelling of a particular BSS component as an artefact or a useful signal was automatically done. Thus, this Section contains a summary of the number of components selected for rejection by each of the proposed metrics for every case (i.e., BSS algorithm and epoch length).

Cardiac Artefact

Several metrics have been used to detect the BSS with possible cardiac contamination. First of all, the Skew pinpoints components with asymmetrical probability distributions (Escudero et al., 2007b). Table 7.4 summarises the average number of BSS components for every case (i.e., epoch length and BSS technique) marked for removal within each epoch by that metric.

Additionally, the component with the minimum value of VarSc for every BSS was picked up for removal (Shao et al., 2008). Hence, only one component was removed from the signals in each case, as Table 7.5 indicates.

Table 7.4: Average number of components marked for rejection for the epoch lengths and algorithms considered in this study with the metric Skew.

#	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	0.98	0.93	1.07	0.93	1.08
20 s	0.92	0.92	1.18	1.00	1.13
40 s	0.99	0.93	1.25	1.20	1.22
60 s	1.08	0.93	1.29	1.30	1.24
90 s	1.08	0.92	1.30	1.33	1.25

Table 7.5: Average number of components marked for rejection for the epoch lengths and algorithms considered in this study with the metric VarSc.

#	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	1.00	1.00	1.00	1.00	1.00
20 s	1.00	1.00	1.00	1.00	1.00
40 s	1.00	1.00	1.00	1.00	1.00
60 s	1.00	1.00	1.00	1.00	1.00
90 s	1.00	1.00	1.00	1.00	1.00

The remaining metrics used to detect CAs in the BSS components are not specific to this kind of artefact. These include a high-order statistic moment (KrE) and three entropic measures (H_{Sh} , $H_{R\acute{e}}$ and ApEn) (Barbati et al., 2004; Escudero et al., 2007b; Greco et al., 2006; Mammone and Morabito, 2008; Mantini et al., 2008). Therefore, it is important to note that the number of components marked for removal by these metrics does not only include components with hints of CAs, but also BSS sources with possible OAs and PLN.

Table 7.6 indicates the mean number of components that were selected for removal for every signal epoch with KrE, considering that this metric was applied to detect all types of artefacts included in this Doctoral Thesis: CA, OA and PLN.

Likewise KrE, H_{Sh} , $H_{R\acute{e}}$ and ApEn were used to label BSS components suspicious of containing any of the artefacts studied. Thus, the average numbers of rejected components shown in Table 7.7 for H_{Sh} , Table 7.8 for

$H_{R\acute{e}}$ and Table 7.9 for ApEn include components with possible CAs, OAs and PLN.

Ocular Artefact

Diverse criteria were applied to point out BSS components which might isolate OAs. Firstly, the P_{EYES} metric marked components whose scalp distribution is mainly gathered near the eyes (Escudero et al., 2007b). Table 7.10 presents the mean number of components labelled as OAs by this metric for each case (i.e., for each combination of epoch length and BSS algorithm).

Another metric which is specifically proposed for the OA is P_{LF} . This is based on the assumption that most of the OA energy is concentrated in low frequencies (Escudero et al., 2007b). The average number of BSS sources identified by this criterion is summarised in Table 7.11.

Finally, the criteria KrE, H_{Sh} , $H_{R\acute{e}}$ and ApEn were also used to try to find components with ocular contamination. As these metrics are not specific to a particular kind of artefact, the mean number of components marked for rejection by each of these criteria has already been recorded in Table 7.6, Table 7.7, Table 7.8 and Table 7.9.

Power Line Noise

The last contamination studied in this Doctoral Thesis is the PLN. Thanks to the fact that this artefact is very well located in the frequency domain, the metric P_{50Hz} was proposed to detect components which are mainly composed by a spectral harmonic at 50 Hz (Escudero et al., 2007b). The mean number of components marked by this criterion for each case is indicated in Table 7.12.

Finally, the metrics KrE, H_{Sh} , $H_{R\acute{e}}$ and ApEn were also tested in the PLN removal. As it has been already explained, the average numbers of components with potential artefactual activity identified by these criteria appear in Table 7.6, Table 7.7, Table 7.8 and Table 7.9, respectively.

Table 7.6: Average number of components marked for rejection for the epoch lengths and algorithms considered in this study with the metric KrE.

#	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	0.67	0.61	0.97	0.86	0.95
20 s	0.68	0.64	1.03	0.91	1.05
40 s	0.72	0.65	1.10	1.07	1.10
60 s	0.70	0.65	1.12	1.14	1.10
90 s	0.76	0.66	1.14	1.18	1.12

Table 7.7: Average number of components marked for rejection for the epoch lengths and algorithms considered in this study with the metric H_{Sh} .

#	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	0.87	0.81	1.01	0.91	1.04
20 s	0.84	0.82	1.14	0.97	1.12
40 s	0.84	0.82	1.23	1.18	1.20
60 s	0.88	0.81	1.25	1.27	1.25
90 s	0.95	0.80	1.26	1.37	1.29

Table 7.8: Average number of components marked for rejection for the epoch lengths and algorithms considered in this study with the metric H_{Re} .

#	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	0.93	0.83	1.01	0.91	1.02
20 s	0.92	0.85	1.11	0.97	1.11
40 s	0.96	0.87	1.24	1.17	1.20
60 s	1.04	0.83	1.27	1.31	1.29
90 s	1.13	0.95	1.29	1.36	1.29

Table 7.9: Average number of components marked for rejection for the epoch lengths and algorithms considered in this study with the metric ApEn.

#	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	1.02	1.02	1.00	1.13	1.04
20 s	1.03	0.98	1.04	1.03	1.06
40 s	1.06	1.05	1.19	1.15	1.17
60 s	1.05	1.03	1.32	1.23	1.25
90 s	1.05	1.08	1.29	1.24	1.26

Table 7.10: Average number of components marked for rejection for the epoch lengths and algorithms considered in this study with the metric P_{EYES} .

#	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	0.37	0.40	0.32	0.50	0.37
20 s	0.48	0.46	0.42	0.50	0.42
40 s	0.54	0.52	0.42	0.51	0.51
60 s	0.56	0.59	0.48	0.53	0.54
90 s	0.58	0.63	0.54	0.59	0.55

Table 7.11: Average number of components marked for rejection for the epoch lengths and algorithms considered in this study with the metric P_{LF} .

#	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	0.77	0.79	0.49	0.51	0.65
20 s	0.90	0.90	0.68	0.70	0.79
40 s	1.03	1.04	0.84	0.96	0.89
60 s	1.10	1.14	0.90	1.03	0.96
90 s	1.18	1.14	1.01	1.17	1.04

7.3.4 Evaluation of the Artefact Removal

This Section contains the results obtained from the automatic evaluation of the artefact rejection performed by removing the activity isolated in the components marked for rejection with Eq. (6.15). It is important to bear in mind the results presented in Section 7.3.1, since they provided us with information about the importance of the artefacts in the original MEG recordings. There-

Table 7.12: Average number of components marked for rejection for the epoch lengths and algorithms considered in this study with the metric P_{50Hz} .

#	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	0.49	0.49	0.41	0.28	0.46
20 s	0.51	0.52	0.43	0.24	0.47
40 s	0.53	0.54	0.43	0.22	0.47
60 s	0.54	0.55	0.44	0.16	0.50
90 s	0.58	0.58	0.47	0.18	0.51

Table 7.13: Average values of A_{PtP} for the epoch lengths and BSS algorithms considered in this study after the CA removal based on the Skew. Data are shown as mean \pm SD.

(pT)	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	0.430 \pm 0.201	0.422 \pm 0.200	0.150 \pm 0.149	0.160 \pm 0.163	0.136 \pm 0.153
20 s	0.432 \pm 0.208	0.411 \pm 0.210	0.139 \pm 0.152	0.159 \pm 0.164	0.133 \pm 0.157
40 s	0.422 \pm 0.199	0.402 \pm 0.196	0.128 \pm 0.145	0.126 \pm 0.149	0.121 \pm 0.152
60 s	0.400 \pm 0.210	0.382 \pm 0.205	0.124 \pm 0.149	0.115 \pm 0.151	0.121 \pm 0.155
90 s	0.403 \pm 0.206	0.379 \pm 0.207	0.124 \pm 0.155	0.117 \pm 0.154	0.119 \pm 0.152

fore, they allowed us to *objectively* compare the reduction in the artefactual power.

Cardiac Artefact

The assessment of the CA removal is based on the computation of an average QRS complex per signal epoch. Previously, the location of the R-peaks in the raw MEG recordings had to be stored. After the BSS processing, the same locations of the MEG signals as in Section 7.3.1 were inspected to obtain the corresponding values of A_{PtP} and RMS. Therefore, this Section will introduce the values of A_{PtP} and RMS computed for each combination of CA detection metric, epoch length and BSS algorithm. Additionally, the mean QRS complex will be plotted in several representative cases.

First of all, the metrics of Skew and VarSc were used to specifically detect the CA. The corresponding values of A_{PtP} and RMS for them appear in Table 7.13 (metric Skew evaluated with A_{PtP}), Table 7.14 (criterion Skew assessed with RMS), Table 7.15 (criterion VarSc evaluated with A_{PtP}) and Table 7.16 (metric VarSc assessed with RMS).

Secondly, the criterion KrE was applied to pinpoint BSS components with abnormal probability distributions. In the case of CAs, their histograms may be characterised very high values of KrE. The values of A_{PtP} and RMS computed for this artefact rejection criterion are summarised in Table 7.17 and Table 7.18, respectively.

Additionally, entropic measures were also used for the same reason as

Table 7.14: Average values of RMS for the epoch lengths and BSS algorithms considered in this study after the CA removal based on the Skew. Data are shown as mean \pm SD.

(pT)	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	0.067 \pm 0.045	0.066 \pm 0.045	0.035 \pm 0.048	0.035 \pm 0.049	0.033 \pm 0.048
20 s	0.069 \pm 0.049	0.066 \pm 0.051	0.035 \pm 0.051	0.037 \pm 0.052	0.034 \pm 0.052
40 s	0.066 \pm 0.048	0.063 \pm 0.048	0.030 \pm 0.047	0.030 \pm 0.047	0.029 \pm 0.048
60 s	0.063 \pm 0.049	0.061 \pm 0.049	0.029 \pm 0.047	0.028 \pm 0.047	0.029 \pm 0.048
90 s	0.063 \pm 0.050	0.061 \pm 0.051	0.029 \pm 0.050	0.029 \pm 0.049	0.030 \pm 0.049

Table 7.15: Average values of A_{PtP} for the epoch lengths and BSS algorithms considered in this study after the CA removal based on the VarSc. Data are shown as mean \pm SD.

(pT)	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	0.454 \pm 0.209	0.429 \pm 0.205	0.166 \pm 0.221	0.129 \pm 0.115	0.119 \pm 0.118
20 s	0.461 \pm 0.227	0.428 \pm 0.226	0.167 \pm 0.260	0.153 \pm 0.192	0.130 \pm 0.138
40 s	0.459 \pm 0.222	0.431 \pm 0.220	0.183 \pm 0.273	0.176 \pm 0.293	0.173 \pm 0.317
60 s	0.445 \pm 0.243	0.392 \pm 0.235	0.176 \pm 0.278	0.175 \pm 0.289	0.180 \pm 0.318
90 s	0.470 \pm 0.249	0.400 \pm 0.242	0.190 \pm 0.282	0.174 \pm 0.293	0.175 \pm 0.325

Table 7.16: Average values of RMS for the epoch lengths and BSS algorithms considered in this study after the CA removal based on the VarSc. Data are shown as mean \pm SD.

(pT)	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	0.069 \pm 0.039	0.068 \pm 0.046	0.040 \pm 0.059	0.031 \pm 0.042	0.031 \pm 0.046
20 s	0.072 \pm 0.046	0.068 \pm 0.052	0.039 \pm 0.064	0.038 \pm 0.058	0.035 \pm 0.051
40 s	0.070 \pm 0.044	0.066 \pm 0.044	0.040 \pm 0.066	0.040 \pm 0.070	0.041 \pm 0.079
60 s	0.068 \pm 0.045	0.060 \pm 0.045	0.040 \pm 0.067	0.040 \pm 0.070	0.041 \pm 0.077
90 s	0.075 \pm 0.056	0.063 \pm 0.053	0.042 \pm 0.069	0.041 \pm 0.071	0.042 \pm 0.080

Table 7.17: Average values of A_{PtP} for the epoch lengths and BSS algorithms considered in this study after the CA removal based on the KrE. Data are shown as mean \pm SD.

(pT)	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	0.463 \pm 0.220	0.451 \pm 0.213	0.142 \pm 0.126	0.147 \pm 0.135	0.123 \pm 0.123
20 s	0.450 \pm 0.220	0.431 \pm 0.221	0.136 \pm 0.121	0.143 \pm 0.142	0.114 \pm 0.123
40 s	0.440 \pm 0.205	0.426 \pm 0.212	0.122 \pm 0.122	0.113 \pm 0.134	0.103 \pm 0.114
60 s	0.433 \pm 0.238	0.403 \pm 0.216	0.122 \pm 0.126	0.105 \pm 0.118	0.105 \pm 0.122
90 s	0.430 \pm 0.222	0.399 \pm 0.213	0.120 \pm 0.118	0.101 \pm 0.113	0.101 \pm 0.111

Table 7.18: Average values of RMS for the epoch lengths and BSS algorithms considered in this study after the CA removal based on the KrE. Data are shown as mean \pm SD.

(pT)	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	0.073 \pm 0.049	0.071 \pm 0.049	0.035 \pm 0.048	0.034 \pm 0.048	0.032 \pm 0.048
20 s	0.072 \pm 0.051	0.069 \pm 0.051	0.035 \pm 0.050	0.035 \pm 0.052	0.032 \pm 0.051
40 s	0.068 \pm 0.048	0.066 \pm 0.049	0.030 \pm 0.048	0.028 \pm 0.049	0.027 \pm 0.048
60 s	0.068 \pm 0.051	0.064 \pm 0.050	0.029 \pm 0.048	0.027 \pm 0.047	0.028 \pm 0.048
90 s	0.068 \pm 0.051	0.063 \pm 0.051	0.029 \pm 0.048	0.027 \pm 0.048	0.028 \pm 0.048

KrE. These metrics can detect probability distributions typical of artefactual activity. The most straightforward entropic measure is H_{Sh} . The evaluation of the CA removal based on this criterion is shown in Table 7.19 for A_{PtP} and Table 7.20 for RMS.

$H_{R\acute{e}}$ is a modification of H_{Sh} . Hence, it may be helpful to recognise the artefactual activities similar to those detected by H_{Sh} . The results of A_{PtP} and RMS computed to assess the quality of the CA removal with $H_{R\acute{e}}$ are depicted in Table 7.21 and Table 7.22, respectively.

Moreover, ApEn is an entropic measurement helpful to assess irregularity in biomedical time series. Hence, this criterion was tested in the removal of CA as well as $H_{R\acute{e}}$ and H_{Sh} . The values of A_{PtP} for this case appear in Table 7.23 whilst the evaluation of this CA rejection by means of RMS is shown in Table 7.24.

The previously introduced Tables offer quantitative data about the reduc-

Table 7.19: Average values of A_{PtP} for the epoch lengths considered in this study after the CA removal based on the metric H_{Sh} . Data are shown as mean \pm SD.

(pT)	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	0.437 \pm 0.214	0.426 \pm 0.207	0.136 \pm 0.122	0.140 \pm 0.133	0.114 \pm 0.116
20 s	0.429 \pm 0.214	0.409 \pm 0.213	0.124 \pm 0.116	0.136 \pm 0.134	0.110 \pm 0.121
40 s	0.425 \pm 0.205	0.402 \pm 0.197	0.109 \pm 0.105	0.101 \pm 0.111	0.102 \pm 0.111
60 s	0.404 \pm 0.217	0.379 \pm 0.213	0.106 \pm 0.109	0.092 \pm 0.110	0.099 \pm 0.113
90 s	0.405 \pm 0.215	0.377 \pm 0.208	0.106 \pm 0.115	0.092 \pm 0.113	0.095 \pm 0.110

Table 7.20: Average values of RMS for the epoch lengths considered in this study after the CA removal based on the metric H_{Sh} . Data are shown as mean \pm SD.

(pT)	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	0.069 \pm 0.049	0.068 \pm 0.049	0.035 \pm 0.048	0.033 \pm 0.048	0.031 \pm 0.048
20 s	0.067 \pm 0.049	0.066 \pm 0.051	0.033 \pm 0.050	0.034 \pm 0.051	0.031 \pm 0.051
40 s	0.067 \pm 0.048	0.063 \pm 0.048	0.028 \pm 0.046	0.027 \pm 0.046	0.027 \pm 0.047
60 s	0.064 \pm 0.050	0.061 \pm 0.050	0.027 \pm 0.047	0.025 \pm 0.047	0.027 \pm 0.048
90 s	0.064 \pm 0.052	0.060 \pm 0.051	0.027 \pm 0.049	0.026 \pm 0.048	0.027 \pm 0.048

Table 7.21: Average values of A_{PtP} for the epoch lengths and BSS algorithms considered in this study after the CA removal based on the $H_{Ré}$. Data are shown as mean \pm SD.

(pT)	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	0.445 \pm 0.219	0.428 \pm 0.204	0.138 \pm 0.124	0.141 \pm 0.134	0.118 \pm 0.125
20 s	0.426 \pm 0.210	0.415 \pm 0.216	0.126 \pm 0.120	0.136 \pm 0.134	0.112 \pm 0.123
40 s	0.426 \pm 0.206	0.408 \pm 0.208	0.109 \pm 0.105	0.104 \pm 0.124	0.103 \pm 0.114
60 s	0.410 \pm 0.229	0.383 \pm 0.209	0.106 \pm 0.109	0.092 \pm 0.110	0.099 \pm 0.113
90 s	0.412 \pm 0.218	0.375 \pm 0.220	0.106 \pm 0.115	0.093 \pm 0.113	0.095 \pm 0.110

Table 7.22: Average values of RMS for the epoch lengths and BSS algorithms considered in this study after the CA removal based on the $H_{R\acute{e}}$. Data are shown as mean \pm SD.

(pT)	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	0.071 \pm 0.049	0.068 \pm 0.048	0.035 \pm 0.048	0.033 \pm 0.048	0.031 \pm 0.048
20 s	0.067 \pm 0.048	0.067 \pm 0.051	0.033 \pm 0.050	0.034 \pm 0.051	0.032 \pm 0.051
40 s	0.067 \pm 0.048	0.064 \pm 0.049	0.028 \pm 0.046	0.027 \pm 0.048	0.027 \pm 0.048
60 s	0.065 \pm 0.051	0.061 \pm 0.050	0.027 \pm 0.047	0.025 \pm 0.047	0.027 \pm 0.048
90 s	0.065 \pm 0.051	0.060 \pm 0.052	0.027 \pm 0.049	0.026 \pm 0.048	0.027 \pm 0.048

Table 7.23: Average values of A_{P_tP} for the epoch lengths and BSS algorithms considered in this study after the CA removal based on the ApEn. Data are shown as mean \pm SD.

(pT)	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	0.620 \pm 0.289	0.602 \pm 0.291	0.260 \pm 0.194	0.219 \pm 0.127	0.225 \pm 0.202
20 s	0.612 \pm 0.301	0.601 \pm 0.311	0.265 \pm 0.263	0.237 \pm 0.240	0.219 \pm 0.233
40 s	0.613 \pm 0.299	0.594 \pm 0.293	0.264 \pm 0.295	0.208 \pm 0.240	0.222 \pm 0.266
60 s	0.604 \pm 0.305	0.572 \pm 0.315	0.240 \pm 0.268	0.178 \pm 0.211	0.206 \pm 0.233
90 s	0.598 \pm 0.321	0.566 \pm 0.284	0.240 \pm 0.301	0.181 \pm 0.214	0.197 \pm 0.230

Table 7.24: Average values of RMS for the epoch lengths and BSS algorithms considered in this study after the CA removal based on the ApEn. Data are shown as mean \pm SD.

(pT)	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	0.087 \pm 0.040	0.084 \pm 0.040	0.049 \pm 0.050	0.037 \pm 0.026	0.041 \pm 0.047
20 s	0.087 \pm 0.041	0.085 \pm 0.043	0.051 \pm 0.063	0.043 \pm 0.054	0.041 \pm 0.054
40 s	0.087 \pm 0.041	0.084 \pm 0.040	0.052 \pm 0.068	0.040 \pm 0.055	0.042 \pm 0.058
60 s	0.087 \pm 0.042	0.081 \pm 0.043	0.048 \pm 0.063	0.037 \pm 0.053	0.039 \pm 0.055
90 s	0.085 \pm 0.044	0.080 \pm 0.039	0.049 \pm 0.070	0.038 \pm 0.055	0.039 \pm 0.057

Table 7.25: Average values of A_{PtP} and RMS for the epoch lengths considered in this study after the CA rejection based on a cBSS procedure. Data are shown as mean \pm SD.

Epoch length (s)	A_{PtP} (pT)	RMS (pT)
10	0.101 \pm 0.090	0.024 \pm 0.027
20	0.116 \pm 0.118	0.026 \pm 0.033
40	0.124 \pm 0.126	0.027 \pm 0.036
60	0.127 \pm 0.129	0.027 \pm 0.036
90	0.140 \pm 0.150	0.029 \pm 0.039

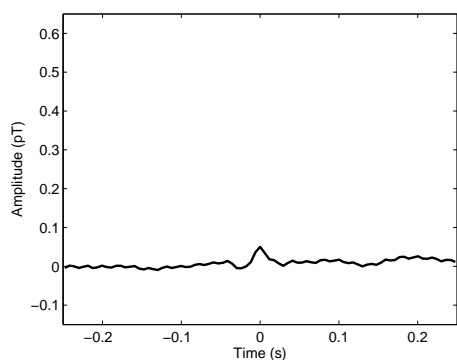
tion in the power (or amplitude) associated with the CA. In order to provide visual information about the diminution of average QRS complexes in the data, Fig. 7.15 plots the average QRS in the MEG signals after the BSS-based artefact rejection procedure for representative cases of the previous criteria.

Finally, a different approach was also tested. This consisted of using a cBSS to extract the cardiac activity from the MEGs. Then, the projection of this extracted signal was subtracted from the MEG channels. Table 7.25 contains the A_{PtP} and RMS values calculated in this case. Furthermore, Fig. 7.16 illustrates the mean QRS complex remaining in the BSS signals after the CA removal with the cBSS.

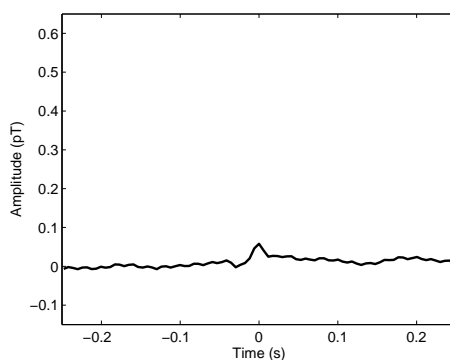
Therefore, it can be seen that certain combinations of artefact detection metric, epoch length and BSS algorithm, as well as some epochs lengths in the cBSS approach, are able to notably reduce the importance of the cardiac artefact. To be more precise, the A_{PtP} and RMS values decreased from about 0.750 and 0.120 to around 0.100 and 0.028, respectively, after the artefact removal procedure.

Ocular Artefact

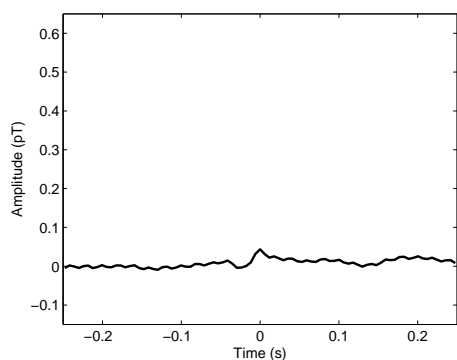
This Section is devoted to the evaluation of how the artefact removal with BSS modifies the MEG signals and the influence of the OAs on them. This evaluation was performed both in the time and spectral domains. It was based on counting the number of peaks exceeding a particular threshold (± 2 pT) and measuring the difference between two subsets of channels in a



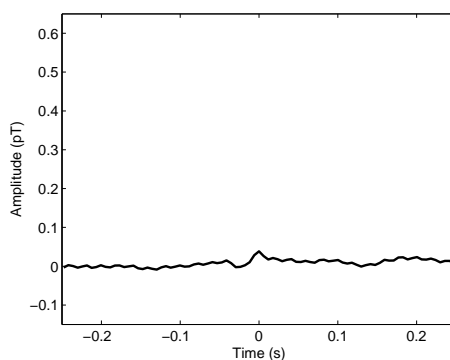
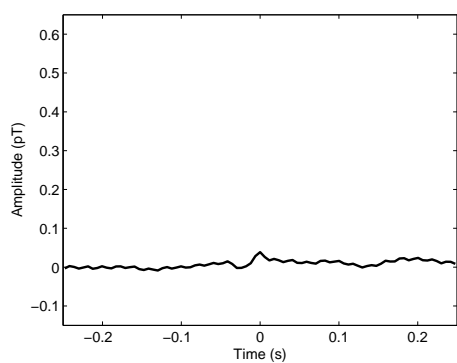
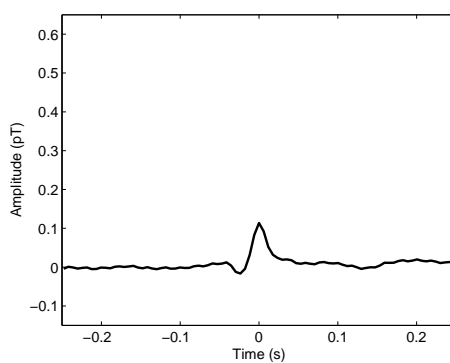
(a) eInfoMax with 60 s epochs and Skew.



(b) FastICA with 10 s epochs and VarSc.



(c) eInfoMax with 90 s epochs and KrE.

(d) eInfoMax with 60 s epochs and H_{Sh} .(e) eInfoMax with 60 s epochs and $H_{R\acute{e}}$.

(f) eInfoMax with 60 s epochs and ApEn.

Figure 7.15: Average QRS complex in the MEG recordings after diverse BSS-based CA removals.

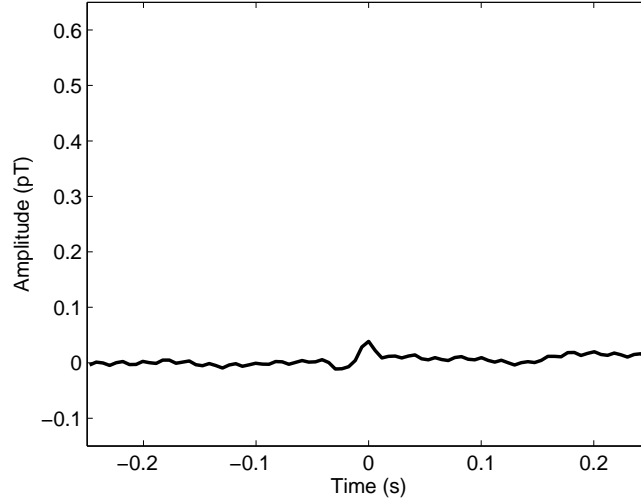


Figure 7.16: Average QRS complex in the MEG recordings after the cBSS-based CA removal for epochs of 10 s.

Table 7.26: Number of peaks in the MEG signals above ± 2 pT for the epoch lengths and BSS algorithms considered in this study after the OA removal based on the P_{EYES} . Data are shown as median \pm SD.

#	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	942 \pm 9599	931 \pm 7280	1302 \pm 29937	1303 \pm 13461	1286 \pm 29726
20 s	946 \pm 8923	930 \pm 7843	1299 \pm 29511	1299 \pm 18403	1299 \pm 29105
40 s	789 \pm 8479	806 \pm 6774	1191 \pm 29050	1139 \pm 27546	1191 \pm 28476
60 s	812 \pm 8698	801 \pm 6416	1249 \pm 30046	1288 \pm 30739	1300 \pm 30727
90 s	799 \pm 9110	798 \pm 5687	1136 \pm 31225	1142 \pm 27315	1100 \pm 28307

particular spectral band, respectively.

First of all, the P_{EYES} criterion was applied to remove ocular contamination from the MEG recordings. The values obtained for the criteria applied to evaluate the artefact rejection appear in Table 7.26, which refers to the time domain evaluation, and Table 7.27, which includes the data for the frequency domain assessment.

As well as P_{EYES} , the metric P_{LF} was proposed to detect only the OAs by considering the low frequencies of the signal. For each epoch length and BSS algorithm, the number of peaks whose amplitude exceeds ± 2 pT appears

Table 7.27: Difference between the PSDs computed at 18 MEG channels near the eyes and at the rest of MEG sensors for the epoch lengths and BSS algorithms considered in this study after the OA removal based on the P_{EYES} . Data are shown as median \pm SD.

(pT ²)	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	0.014 \pm 0.191	0.013 \pm 0.088	0.017 \pm 1.333	0.017 \pm 0.533	0.017 \pm 1.216
20 s	0.011 \pm 0.153	0.014 \pm 0.109	0.019 \pm 1.233	0.019 \pm 0.888	0.019 \pm 1.268
40 s	0.008 \pm 0.192	0.008 \pm 0.095	0.017 \pm 1.381	0.019 \pm 1.462	0.019 \pm 1.363
60 s	0.008 \pm 0.183	0.008 \pm 0.059	0.019 \pm 1.299	0.019 \pm 1.443	0.019 \pm 1.431
90 s	0.008 \pm 0.315	0.010 \pm 0.043	0.015 \pm 1.360	0.018 \pm 1.471	0.018 \pm 1.350

Table 7.28: Number of peaks in the signals above ± 2 pT for the epoch lengths and BSS algorithms considered in this study after the OA removal based on the P_{LF} . Data are shown as median \pm SD.

#	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	654 \pm 6492	597 \pm 5203	975 \pm 24076	769 \pm 11873	932 \pm 24725
20 s	642 \pm 5250	537 \pm 4753	953 \pm 25468	761 \pm 12483	847 \pm 21314
40 s	546 \pm 3748	354 \pm 3659	804 \pm 23058	746 \pm 14854	745 \pm 16450
60 s	247 \pm 3815	234 \pm 3381	869 \pm 25947	797 \pm 20046	679 \pm 18929
90 s	182 \pm 2949	182 \pm 4453	751 \pm 30990	750 \pm 18921	652 \pm 15757

in Table 7.28 while the spectral difference in low frequencies between the subsets of channels is included in Table 7.29.

As it was previously described, KrE, H_{Sh} , $H_{R\acute{e}}$ and ApEn have been tested in the removal of the CA, OA and PLN. The time and spectral variables obtained for the KrE criterion are detailed in Table 7.30 and Table 7.31, respectively. Similarly to the CA, the entropies can recognise components whose probability distributions are characterised by an abnormal behaviour. Whereas the evaluation of the OA removal with H_{Sh} are detailed in Table 7.32 and Table 7.33, the results for $H_{R\acute{e}}$ appear in Table 7.34 and Table 7.35.

The last metric employed in this Doctoral Thesis in the OA removal is ApEn. The corresponding results are detailed in Table 7.36 for the time-based evaluation and in Table 7.37 for the spectral ‘discrepancy’ measure.

In addition to the quantitative information presented in the previous ta-

Table 7.29: Difference between the PSDs computed at 18 MEG channels close to the eyes and at the rest of MEG sensors for the epoch lengths and BSS algorithms considered in this study after the OA removal based on the P_{LF} . Data are shown as median \pm SD.

(pT ²)	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	0.009 \pm 0.142	0.010 \pm 0.069	0.012 \pm 0.989	0.011 \pm 0.148	0.009 \pm 0.928
20 s	0.008 \pm 0.114	0.008 \pm 0.086	0.012 \pm 1.051	0.010 \pm 0.123	0.010 \pm 0.922
40 s	0.008 \pm 0.057	0.008 \pm 0.055	0.013 \pm 0.813	0.009 \pm 0.173	0.011 \pm 0.848
60 s	0.007 \pm 0.033	0.007 \pm 0.026	0.014 \pm 0.787	0.010 \pm 0.333	0.010 \pm 0.739
90 s	0.008 \pm 0.026	0.008 \pm 0.180	0.015 \pm 0.817	0.010 \pm 0.422	0.011 \pm 0.805

Table 7.30: Number of peaks in the signals above ± 2 pT for the epoch lengths and BSS algorithms considered in this study after the OA removal based on the KrE. Data are shown as median \pm SD.

#	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	599 \pm 20608	622 \pm 20612	658 \pm 22113	635 \pm 21476	637 \pm 22072
20 s	608 \pm 20655	582 \pm 20620	652 \pm 21842	635 \pm 21737	640 \pm 21730
40 s	632 \pm 19330	637 \pm 19327	638 \pm 20363	630 \pm 20534	637 \pm 20629
60 s	735 \pm 20726	671 \pm 20784	671 \pm 21811	660 \pm 21709	657 \pm 21729
90 s	502 \pm 18959	498 \pm 19030	485 \pm 19987	482 \pm 19802	484 \pm 19904

Table 7.31: Difference between the PSDs computed at 18 MEG channels close to the eyes and at the rest of MEG sensors for the epoch lengths and BSS algorithms considered in this study after the OA removal based on the KrE. Data are shown as median \pm SD.

(pT ²)	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	0.014 \pm 1.917	0.015 \pm 1.914	0.016 \pm 1.907	0.016 \pm 1.907	0.016 \pm 1.908
20 s	0.015 \pm 1.926	0.015 \pm 1.926	0.016 \pm 1.920	0.016 \pm 1.919	0.016 \pm 1.921
40 s	0.018 \pm 1.990	0.017 \pm 1.991	0.017 \pm 1.986	0.017 \pm 1.983	0.017 \pm 1.984
60 s	0.018 \pm 1.936	0.018 \pm 1.936	0.017 \pm 1.932	0.017 \pm 1.931	0.017 \pm 1.931
90 s	0.016 \pm 2.033	0.015 \pm 2.033	0.015 \pm 2.027	0.015 \pm 2.028	0.015 \pm 2.027

Table 7.32: Number of peaks in the signals above ± 2 pT for the epoch lengths and BSS algorithms considered in this study after the OA removal based on the H_{Sh} . Data are shown as median \pm SD.

#	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	477 \pm 20708	418 \pm 20641	633 \pm 22116	539 \pm 21511	543 \pm 22082
20 s	620 \pm 20931	428 \pm 20732	652 \pm 21850	541 \pm 21743	543 \pm 21740
40 s	630 \pm 19344	637 \pm 19542	638 \pm 20353	630 \pm 20533	637 \pm 20640
60 s	522 \pm 20761	526 \pm 20812	671 \pm 21818	660 \pm 21721	657 \pm 21735
90 s	458 \pm 18979	488 \pm 19057	483 \pm 19992	479 \pm 19821	480 \pm 19922

Table 7.33: Difference between the PSDs computed at 18 MEG channels close to the eyes and at the rest of MEG sensors for the epoch lengths and BSS algorithms considered in this study after the OA removal based on the H_{Sh} . Data are shown as median \pm SD.

(pT ²)	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	0.012 \pm 1.917	0.011 \pm 1.916	0.015 \pm 1.907	0.014 \pm 1.907	0.014 \pm 1.908
20 s	0.013 \pm 1.926	0.012 \pm 1.924	0.016 \pm 1.920	0.014 \pm 1.919	0.014 \pm 1.921
40 s	0.018 \pm 1.991	0.017 \pm 1.991	0.017 \pm 1.986	0.017 \pm 1.983	0.017 \pm 1.984
60 s	0.016 \pm 1.937	0.015 \pm 1.937	0.017 \pm 1.932	0.017 \pm 1.931	0.017 \pm 1.932
90 s	0.013 \pm 2.034	0.013 \pm 2.034	0.015 \pm 2.027	0.014 \pm 2.028	0.014 \pm 2.028

Table 7.34: Number of peaks in the signals above ± 2 pT for the epoch lengths and BSS algorithms considered in this study after the OA removal based on the $H_{Ré}$. Data are shown as median \pm SD.

#	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	449 \pm 20777	414 \pm 20818	558 \pm 22123	539 \pm 21548	543 \pm 22096
20 s	566 \pm 20942	368 \pm 20763	652 \pm 21855	541 \pm 21747	543 \pm 21750
40 s	480 \pm 19347	421 \pm 19365	638 \pm 20356	630 \pm 20539	637 \pm 20645
60 s	645 \pm 20735	535 \pm 20823	671 \pm 21822	594 \pm 21739	612 \pm 21752
90 s	390 \pm 18839	373 \pm 18642	483 \pm 20005	479 \pm 19820	480 \pm 19922

Table 7.35: Difference between the PSDs computed at 18 MEG channels close to the eyes and at the rest of MEG sensors for the epoch lengths and BSS algorithms considered in this study after the OA removal based on the $H_{R\acute{e}}$. Data are shown as median \pm SD.

(pT ²)	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	0.011 \pm 1.918	0.011 \pm 1.910	0.014 \pm 1.907	0.014 \pm 1.906	0.014 \pm 1.908
20 s	0.011 \pm 1.927	0.009 \pm 1.920	0.016 \pm 1.920	0.014 \pm 1.919	0.014 \pm 1.921
40 s	0.012 \pm 1.991	0.012 \pm 1.986	0.017 \pm 1.986	0.017 \pm 1.983	0.017 \pm 1.984
60 s	0.013 \pm 1.937	0.015 \pm 1.937	0.017 \pm 1.932	0.015 \pm 1.932	0.015 \pm 1.932
90 s	0.014 \pm 2.034	0.013 \pm 2.034	0.015 \pm 2.028	0.014 \pm 2.028	0.014 \pm 2.028

Table 7.36: Number of peaks in the signals above ± 2 pT for the epoch lengths and BSS algorithms considered in this study after the OA removal based on the ApEn. Data are shown as median \pm SD.

#	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	315 \pm 3406	304 \pm 3367	560 \pm 21683	370 \pm 15016	458 \pm 21510
20 s	477 \pm 3240	370 \pm 3052	657 \pm 20774	528 \pm 17325	482 \pm 20700
40 s	460 \pm 2411	282 \pm 2571	575 \pm 19209	393 \pm 18278	408 \pm 18653
60 s	706 \pm 3230	706 \pm 3426	575 \pm 20374	420 \pm 20249	573 \pm 20609
90 s	634 \pm 3628	582 \pm 3550	453 \pm 18981	255 \pm 19133	269 \pm 19087

Table 7.37: Difference between the PSDs computed at 18 MEG channels close to the eyes and at the rest of MEG sensors for the epoch lengths and BSS algorithms considered in this study after the OA removal based on the ApEn. Data are shown as median \pm SD.

(pT ²)	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	0.009 \pm 0.011	0.010 \pm 0.012	0.013 \pm 1.883	0.011 \pm 1.055	0.014 \pm 1.737
20 s	0.008 \pm 0.012	0.007 \pm 0.013	0.016 \pm 1.924	0.014 \pm 1.252	0.012 \pm 1.905
40 s	0.008 \pm 0.016	0.008 \pm 0.013	0.017 \pm 1.989	0.011 \pm 1.841	0.011 \pm 1.992
60 s	0.010 \pm 0.014	0.009 \pm 0.015	0.017 \pm 1.934	0.015 \pm 1.935	0.015 \pm 1.909
90 s	0.011 \pm 0.011	0.011 \pm 0.013	0.013 \pm 2.029	0.013 \pm 2.000	0.013 \pm 1.145

bles, a visual evaluation of the impact of the ocular metrics in the MEG signals' spectra is depicted in Fig. 7.17 for a few representative combinations of epoch length, BSS algorithm and artefact detection criterion. This figure shows the same low frequency range as Fig. 7.13 for the sake of a straightforward comparison, since it is necessary to compare the spectra computed with the same epoch length.

By comparing Fig. 7.13 and Fig. 7.17, it is clear that the BSS-based artefact rejection procedure caused a reduction in the amount of low frequency power at the channels located near the eyes. Additionally, the number of peaks with extreme values also decreased notably.

Power Line Noise

Finally, this Doctoral Thesis also attempted to reduce the PLN in the MEG recordings. The evaluation of this artefact removal was entirely done in the frequency domain. To begin with, the power of the PSD around 50 Hz was measured. This provides information about the total power of the PLN. Then, the PSD_n was computed and the *normalised* power in the same frequency band was computed. This calculation was carried out due to the fact that some artefact removal metrics were applied to detect all kind of artefacts. Hence, it might be possible that the total power of the PLN decreased in the signals while the normalised power increased. This might happen if a certain artefact detection metric removed some portion of the PLN and a very large fraction of the CA or OA, thus removing a larger proportion of the whole recording power than the proportion of the removed PLN.

Firstly, the $P_{50\text{Hz}}$ metric was used to detect and remove the BSS components with large power at 50 Hz. The values of total and normalised power around the line frequency after the BSS-based artefact removal are given in Table 7.38 and Table 7.39, respectively.

Since the amplitude distribution of a purely harmonic signal may be relatively different from that of brain activity, KrE, H_{Sh} , $H_{R\acute{e}}$ and ApEn were also tested in the reduction of the PLN. The evaluation of the PLN power reduction by means of the KrE is given by Table 7.40 and Table 7.41.

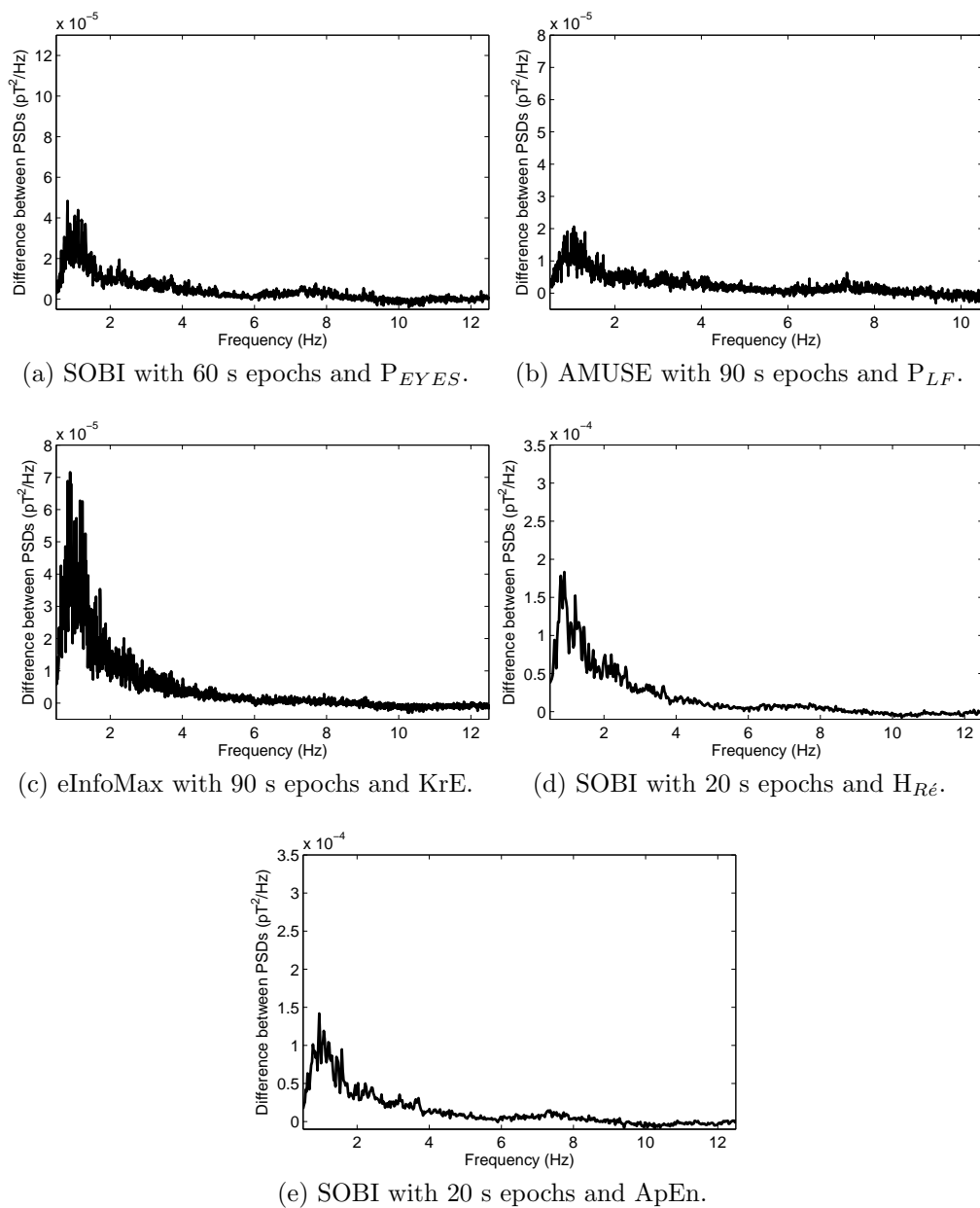


Figure 7.17: Median of the difference between the PSDs for the two subsets of MEG channels used to compare the impact of OAs in the recordings after diverse BSS-based OA removals.

Table 7.38: Average values of total spectral power in the band from 49 Hz to 51 Hz multiplied by 10^3 for the epoch lengths and BSS algorithms considered in this study after the PLN removal based on the $P_{50\text{Hz}}$. Data are shown as mean \pm SD.

$\left(\frac{\text{pT}^2}{10^3}\right)$	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	0.479 \pm 0.863	0.479 \pm 0.864	0.718 \pm 0.879	1.119 \pm 1.202	0.529 \pm 0.853
20 s	0.476 \pm 0.864	0.476 \pm 0.864	0.627 \pm 0.857	1.029 \pm 1.106	0.518 \pm 0.854
40 s	0.487 \pm 0.927	0.487 \pm 0.927	0.579 \pm 0.912	1.254 \pm 1.536	0.521 \pm 0.918
60 s	0.466 \pm 0.863	0.471 \pm 0.862	0.550 \pm 0.856	1.394 \pm 2.347	0.497 \pm 0.855
90 s	0.490 \pm 0.962	0.491 \pm 0.962	0.558 \pm 0.952	1.079 \pm 1.290	0.521 \pm 0.954

Table 7.39: Average values of normalised spectral power in the band from 49 Hz to 51 Hz multiplied by 10^2 for the epoch lengths and BSS algorithms considered in this study after the PLN removal based on the $P_{50\text{Hz}}$. Data are shown as mean \pm SD.

(10^2)	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	0.713 \pm 0.751	0.718 \pm 0.755	1.125 \pm 0.986	1.815 \pm 1.869	0.805 \pm 0.762
20 s	0.689 \pm 0.712	0.686 \pm 0.709	0.958 \pm 0.846	1.593 \pm 1.483	0.776 \pm 0.757
40 s	0.675 \pm 0.743	0.674 \pm 0.748	0.837 \pm 0.764	1.713 \pm 1.678	0.749 \pm 0.751
60 s	0.637 \pm 0.624	0.649 \pm 0.628	0.791 \pm 0.748	1.790 \pm 2.039	0.729 \pm 0.746
90 s	0.614 \pm 0.563	0.621 \pm 0.572	0.782 \pm 0.761	1.567 \pm 1.671	0.718 \pm 0.741

Table 7.40: Average values of total spectral power in the band from 49 Hz to 51 Hz multiplied by 10^3 for the epoch lengths and BSS algorithms considered in this study after the PLN removal based on the KrE. Data are shown as mean \pm SD.

$\left(\frac{\text{pT}^2}{10^3}\right)$	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	2.269 \pm 3.802	2.269 \pm 3.802	2.265 \pm 3.806	2.228 \pm 3.764	2.264 \pm 3.797
20 s	2.283 \pm 3.823	2.279 \pm 3.822	2.277 \pm 3.828	2.268 \pm 3.821	2.272 \pm 3.815
40 s	2.300 \pm 3.836	2.299 \pm 3.836	2.290 \pm 3.838	2.292 \pm 3.831	2.291 \pm 3.828
60 s	2.284 \pm 3.833	2.285 \pm 3.835	2.277 \pm 3.829	2.277 \pm 3.828	2.275 \pm 3.824
90 s	2.314 \pm 3.844	2.313 \pm 3.843	2.303 \pm 3.834	2.302 \pm 3.836	2.303 \pm 3.834

Table 7.41: Average values of normalised spectral power in the band from 49 Hz to 51 Hz multiplied by 10^2 for the epoch lengths and BSS algorithms considered in this study after the PLN removal based on the KrE. Data are shown as mean \pm SD.

(10^2)	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	4.132 \pm 5.924	4.159 \pm 6.052	4.517 \pm 6.862	4.481 \pm 6.944	4.578 \pm 6.999
20 s	4.100 \pm 5.866	4.168 \pm 6.097	4.512 \pm 6.907	4.512 \pm 6.945	4.557 \pm 7.006
40 s	4.023 \pm 5.764	4.129 \pm 6.018	4.453 \pm 6.831	4.495 \pm 6.895	4.499 \pm 6.925
60 s	4.047 \pm 5.801	4.124 \pm 6.025	4.504 \pm 6.927	4.530 \pm 6.964	4.527 \pm 6.989
90 s	4.113 \pm 6.040	4.109 \pm 6.054	4.457 \pm 6.864	4.470 \pm 6.875	4.467 \pm 6.894

Table 7.42: Average values of total spectral power in the band from 49 Hz to 51 Hz multiplied by 10^3 for the epoch lengths and BSS algorithms considered in this study after the PLN removal based on the H_{Sh} . Data are shown as mean \pm SD.

($\frac{pT^2}{10^3}$)	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	2.251 \pm 3.788	2.265 \pm 3.800	2.265 \pm 3.806	2.227 \pm 3.765	2.263 \pm 3.796
20 s	2.264 \pm 3.815	2.279 \pm 3.821	2.276 \pm 3.826	2.268 \pm 3.821	2.272 \pm 3.814
40 s	2.300 \pm 3.836	2.299 \pm 3.836	2.288 \pm 3.837	2.293 \pm 3.831	2.290 \pm 3.827
60 s	2.284 \pm 3.834	2.285 \pm 3.835	2.276 \pm 3.828	2.275 \pm 3.827	2.273 \pm 3.823
90 s	2.250 \pm 3.806	2.313 \pm 3.843	2.303 \pm 3.834	2.301 \pm 3.834	2.303 \pm 3.833

Furthermore, the entropic measurements were tested in the reduction of the PLN. The corresponding spectral metrics computed from the PSD and PSD_n are included in Table 7.42 and Table 7.43 for the H_{Sh} , while the results for $H_{R\acute{e}}$ appear in Table 7.44 and Table 7.45.

Finally, the artefact reduction results obtained with ApEn appear in Table 7.46 for the total power and Table 7.47 for the normalised power.

Considering those results, Fig. 7.18 offers visual information about the reduction in the PLN for a few representative cases of the BSS-based artefact removal. The plots show the same amplitude range as Fig. 7.14. It can be seen that the power at 50 Hz is reduced about 30 times (i.e, about 14.7 dB).

Table 7.43: Average values of normalised spectral power in the band from 49 Hz to 51 Hz multiplied by 10^2 for the epoch lengths and BSS algorithms considered in this study after the PLN removal based on the H_{Sh} . Data are shown as mean \pm SD.

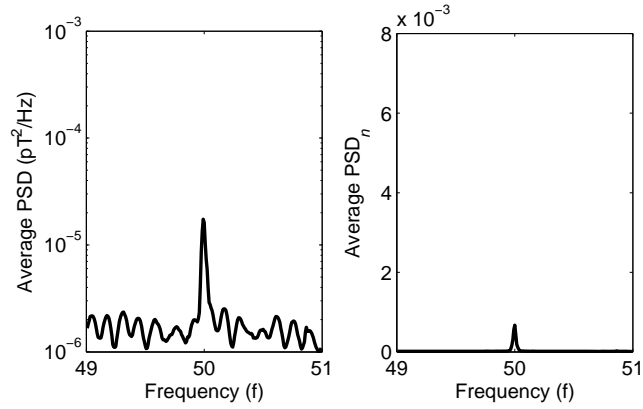
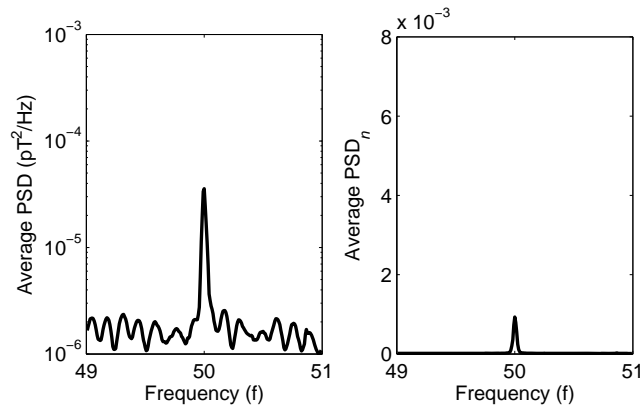
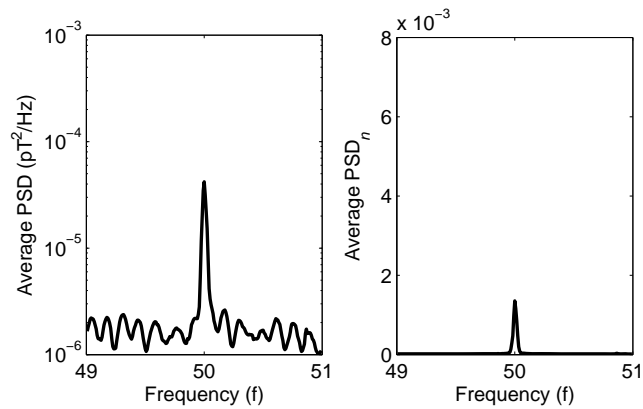
(10^2)	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	4.148 \pm 5.911	4.227 \pm 6.168	4.525 \pm 6.868	4.497 \pm 6.946	4.590 \pm 7.010
20 s	4.166 \pm 5.983	4.205 \pm 6.121	4.538 \pm 6.937	4.519 \pm 6.948	4.571 \pm 7.022
40 s	4.050 \pm 5.795	4.138 \pm 6.017	4.473 \pm 6.854	4.505 \pm 6.892	4.508 \pm 6.936
60 s	4.127 \pm 5.968	4.183 \pm 6.136	4.524 \pm 6.948	4.548 \pm 6.977	4.540 \pm 7.006
90 s	3.921 \pm 5.440	4.176 \pm 6.236	4.473 \pm 6.879	4.498 \pm 6.897	4.481 \pm 6.905

Table 7.44: Average values of total spectral power in the band from 49 Hz to 51 Hz multiplied by 10^3 for the epoch lengths and BSS algorithms considered in this study after the PLN removal based on the $H_{R\acute{e}}$. Data are shown as mean \pm SD.

$\left(\frac{pT^2}{10^3}\right)$	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	1.918 \pm 3.347	2.086 \pm 3.513	2.265 \pm 3.806	2.231 \pm 3.785	2.264 \pm 3.796
20 s	1.822 \pm 3.444	2.115 \pm 3.748	2.277 \pm 3.827	2.268 \pm 3.821	2.272 \pm 3.814
40 s	1.629 \pm 3.197	2.045 \pm 3.703	2.288 \pm 3.837	2.292 \pm 3.831	2.291 \pm 3.827
60 s	1.258 \pm 2.311	2.098 \pm 3.723	2.276 \pm 3.828	2.275 \pm 3.827	2.273 \pm 3.823
90 s	1.195 \pm 2.568	1.535 \pm 2.589	2.305 \pm 3.833	2.300 \pm 3.834	2.303 \pm 3.833

Table 7.45: Average values of normalised spectral power in the band from 49 Hz to 51 Hz multiplied by 10^2 for the epoch lengths and BSS algorithms considered in this study after the PLN removal based on the $H_{R\acute{e}}$. Data are shown as mean \pm SD.

(10^2)	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	3.357 \pm 4.467	3.859 \pm 5.338	4.524 \pm 6.867	4.496 \pm 6.939	4.587 \pm 7.006
20 s	3.226 \pm 4.271	3.740 \pm 5.211	4.534 \pm 6.937	4.519 \pm 6.947	4.565 \pm 7.015
40 s	2.470 \pm 3.356	3.422 \pm 4.635	4.474 \pm 6.853	4.501 \pm 6.893	4.509 \pm 6.935
60 s	1.927 \pm 2.528	3.576 \pm 4.852	4.524 \pm 6.948	4.551 \pm 6.975	4.543 \pm 7.005
90 s	1.713 \pm 2.540	2.518 \pm 3.202	4.476 \pm 6.878	4.497 \pm 6.898	4.481 \pm 6.905

(a) AMUSE with 60 s epochs and $P_{50\text{Hz}}$.(b) JADE with 60 s epochs and $P_{50\text{Hz}}$.

(c) JADE with 60 s epochs and ApEn.

Figure 7.18: Average spectra centred at the PLN frequency after the artefact removal with three BSS procedures.

Table 7.46: Average values of total spectral power in the band from 49 Hz to 51 Hz multiplied by 10^3 for the epoch lengths and BSS algorithms considered in this study after the PLN removal based on the ApEn. Data are shown as mean \pm SD.

$\left(\frac{\text{pT}^2}{10^3}\right)$	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	0.866 \pm 0.924	0.922 \pm 0.957	1.129 \pm 1.473	1.402 \pm 1.528	0.909 \pm 0.794
20 s	0.748 \pm 0.853	0.822 \pm 0.890	0.904 \pm 1.037	1.415 \pm 1.754	0.971 \pm 0.951
40 s	0.699 \pm 0.795	0.758 \pm 0.799	0.745 \pm 0.948	1.371 \pm 1.804	0.744 \pm 0.935
60 s	0.855 \pm 1.068	0.970 \pm 1.178	0.583 \pm 0.846	1.403 \pm 2.340	0.670 \pm 0.872
90 s	0.802 \pm 1.107	0.764 \pm 1.033	0.635 \pm 0.953	1.271 \pm 1.658	0.704 \pm 0.954

Table 7.47: Average values of normalised spectral power in the band from 49 Hz to 51 Hz multiplied by 10^2 for the epoch lengths and BSS algorithms considered in this study after the PLN removal based on the ApEn. Data are shown as mean \pm SD.

(10^2)	AMUSE	SOBI	JADE	eInfoMax	FastICA
10 s	1.997 \pm 1.912	2.137 \pm 2.048	1.998 \pm 2.245	2.700 \pm 2.541	1.954 \pm 1.845
20 s	1.606 \pm 1.706	1.790 \pm 1.892	1.737 \pm 1.760	2.342 \pm 2.290	1.913 \pm 1.660
40 s	1.510 \pm 1.614	1.670 \pm 1.770	1.341 \pm 1.173	2.298 \pm 2.271	1.410 \pm 1.323
60 s	1.522 \pm 1.653	1.841 \pm 2.245	1.095 \pm 1.090	2.289 \pm 2.761	1.306 \pm 1.285
90 s	1.394 \pm 1.593	1.370 \pm 1.475	1.092 \pm 0.936	2.074 \pm 2.092	1.249 \pm 1.244

7.4 Feature Enhancement

This Section of the Doctoral Thesis details the results obtained from the pre-processing with BSS techniques to help in the discrimination of AD patients from elderly control subjects' MEGs.

7.4.1 Preliminary Description of the Analysis

This application is based on the idea that a BSS may provide a decomposition of the MEG recordings in which some components are more sensitive to the dementia than others. Therefore, a subset of components that collects most of the differences between the subject groups can be selected to par-

tially reconstruct the MEG signals. Hopefully, this would lead to a better discrimination between signal features of AD patients and healthy elderly subjects (Escudero et al., 2008d, 2009b).

As a trade-off between the number of samples required by spectral and non-linear signal features, epochs of 10 s (1695 samples) were analysed in this Section of the Doctoral Thesis (Escudero et al., 2009b). It is important to note that, in order to avoid surplus complexity, signal epochs with minimal artefactual activity were first selected. Moreover, no dimension reduction procedure was applied due to the fact that only the most relevant components were retained to partially reconstruct the MEG signals. Regarding the application of the BSS algorithms to the actual recordings, the signal epochs of 10 s with minimal artefactual activity were fed to AMUSE and SOBI. In contrast, FastICA and eInfoMax were applied to the whole MEG recordings of five minutes in order to have enough data samples to reliably estimate the ICA decomposition. Afterwards, the same signal epochs selected for analysis in the SOS case were picked up for the HOS algorithms. Of note is that JADE could not be applied in this setting due to the high dimensionality of the data ($m = 148$).

It is important to bear in mind that the development of the preprocessing was carried out with the data from the subjects included in the training set whereas the evaluation of the classification improvement due to this procedure was done with the test set (see Table 5.3).

7.4.2 Decompositions and Ordering Criteria

Firstly, AMUSE, SOBI, FastICA and eInfoMax were applied to blindly decompose MEG background activity recorded from all AD patients and control subjects. AMUSE naturally ordered the components by decreasing linear predictability. On the other hand, the components extracted with SOBI, FastICA and eInfoMax had to be ranked to compare different subjects. These sources were ordered by increasing MF and kurtosis. These ranking processes allowed us to directly compare the values of signal features (MF, SpecEn, LZC and SampEn) computed for the ordered BSS components of both sub-

ject groups (Escudero et al., 2009b,c).

Afterwards, a one-way ANOVA with age as a covariate was used to statistically assess the differences between the MF, SpecEn, LZC and SampEn values of the subject groups for each AMUSE, SOBI, FastICA and eInfoMax component. Thanks to this information, it is possible to define subsets of components characterised by providing the most significant differences between subject groups. These subsets of components will eventually be used to partially reconstruct the MEG signals (Escudero et al., 2008d, 2009b,c).

AMUSE

AMUSE was applied to blindly decompose MEG background activity epochs of 10 s (1695 samples) recorded from all 36 AD patients and 26 controls. However, the analysis of components aiming at developing the preprocessing was only done with the training set.

Given that AMUSE orders the components by decreasing linear predictability, the MF, SpecEn, LZC and SampEn of the components for the subjects in the training set could be straightforwardly compared. Fig. 7.19 depicts these values for each AMUSE component averaged over all AD patients and control subjects included in the training set.

Overall, the values of MF, SpecEn, LZC and SampEn increased with the AMUSE component index. This relationship is particularly clear for MF, indicating that the order provided by AMUSE is related to the low- or high-frequency content of the components (Escudero et al., 2008d). For this reason, MF was selected as an ordering criterion for the other BSS algorithms (SOBI, FastICA and eInfoMax) (Escudero et al., 2009c). Moreover, the SampEn results confirm that higher AMUSE component indexes correspond to more irregular (i.e. less predictable) data. Additionally, lower MF, SpecEn, LZC and SampEn values tended to be found in the AD patients' AMUSE components, as expected (Escudero et al., 2009b).

Then, a one-way ANOVA with age as a covariate was used to statistically assess the differences between subject groups of the training set for each AMUSE component and signal feature. Fig. 7.20 depicts these results.

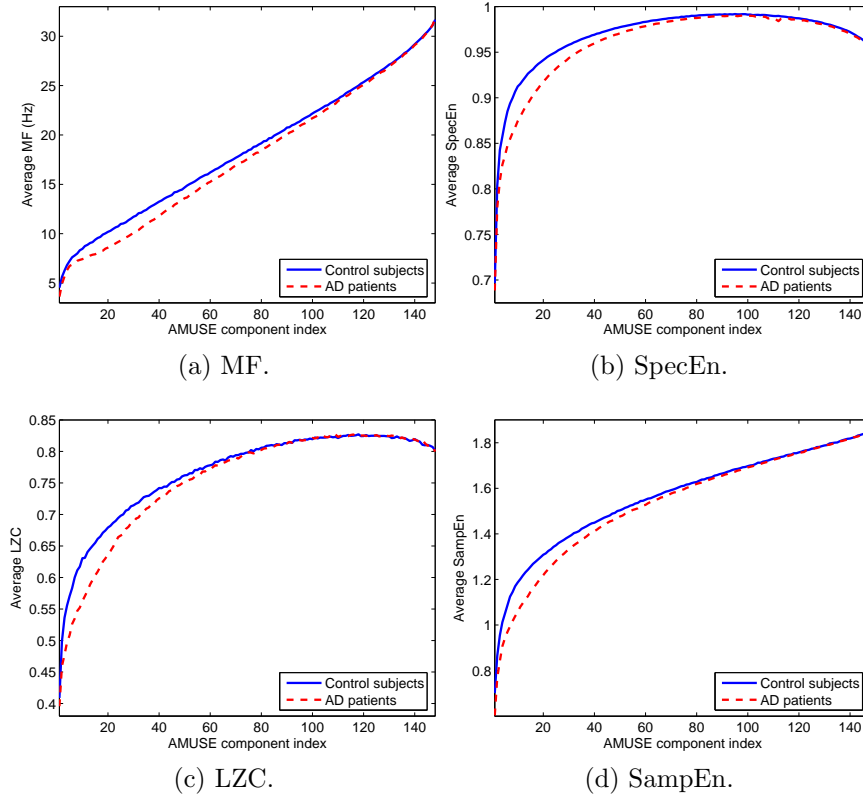


Figure 7.19: Average values of the signal features for every AMUSE component in AD patients and control subjects of the training set.

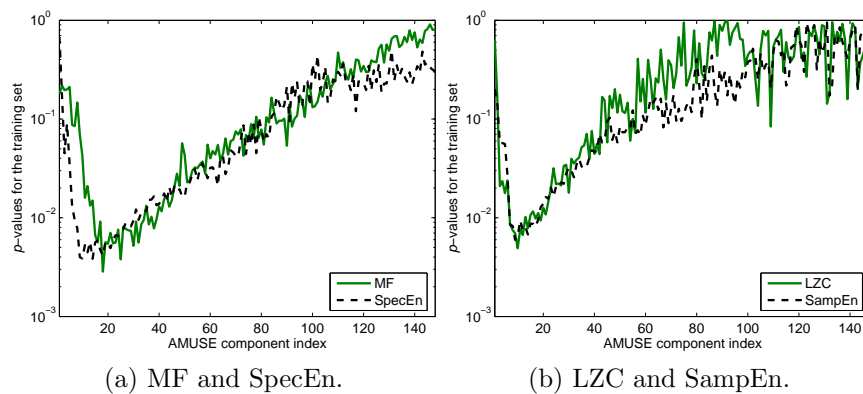


Figure 7.20: p -values of a one-way ANOVA with age as a covariate computed for each AMUSE component in the training set for (a) MF and SpecEn and (b) LZC and SampEn.

They show that the evolution of the p -values is similar for all metrics. The components with the most significant differences are gathered together and have relatively low AMUSE components indexes, although the very first components provide less differentiation between groups.

SOBI

The same epochs fed into AMUSE were decomposed with SOBI. This Section presents the study of the SOBI components derived from the training set and the corresponding signal features MF, SpecEn, LZC and SampEn.

Likewise AMUSE, SOBI is a SOS-BSS algorithm. It employs an iterative procedure to simultaneously diagonalise a set of covariance matrices. Thus, SOBI does not provide an inherent order of the extracted components. On the bases of the AMUSE results, the MF of the SOBI components was used to rank them so that comparisons can be done between the SOBI sources of different subjects (Escudero et al., 2009c). For the sake of completeness, the kurtosis of the components was also used as an ordering criteria based on HOS.

MF-Based Order Fig. 7.21 represents the values of MF, SpecEn, LZC and SampEn computed for AD patients and controls from the SOBI components once these had been ordering by increasing values of MF within each epoch.

Likewise the case of AMUSE, MF, SpecEn, LZC and SampEn increased with the ordering index. However, the evolution of these features with the index appeared ‘more noisy’ than for AMUSE. This may be due to the fact that SOBI does not impose a natural order on the decomposition. Additionally, AD patients presented lower values of the signal features.

A one-way ANOVA with age as a covariate was computed to evaluate the group differences in the training set for the features of each individual component. The results in terms of p -values appear in Fig. 7.22. It can be seen that the evolution of the statistical differences was similar to the AMUSE case.

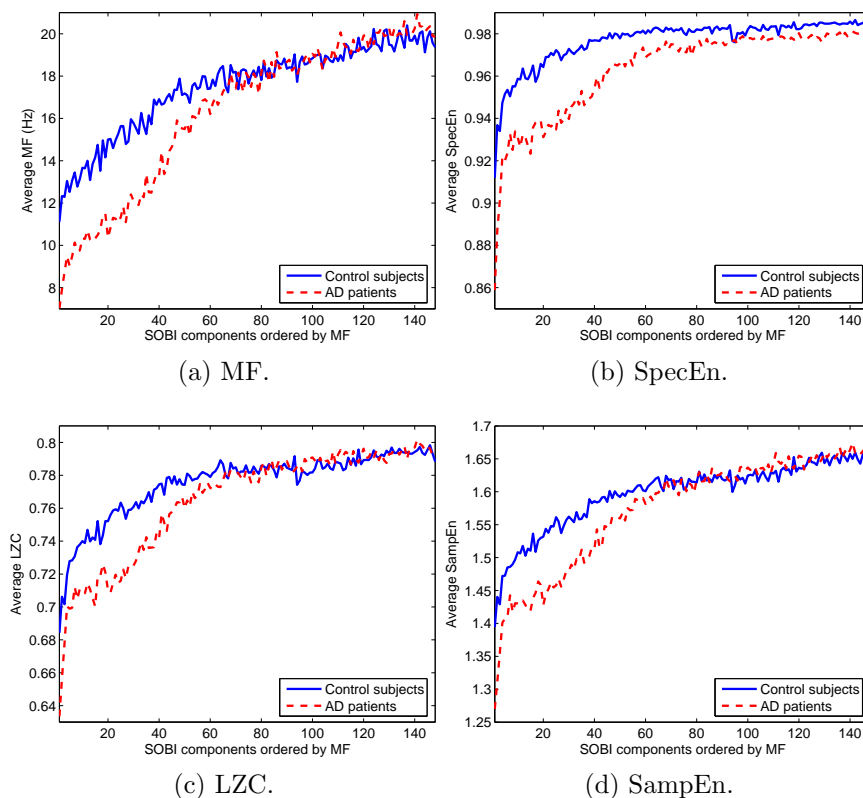


Figure 7.21: Average values of the signal features for the SOBI components orderer by their MF in AD patients and control subjects of the training set.

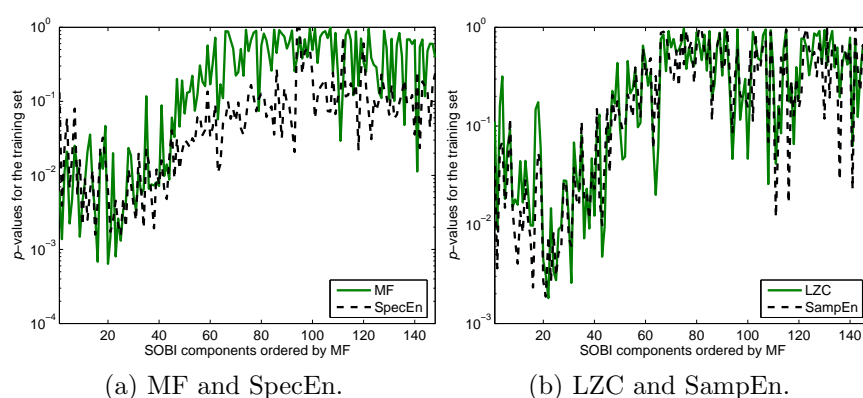


Figure 7.22: p -values of a one-way ANOVA with age as a covariate computed for SOBI components ordered by their MF in the training set for (a) MF and SpecEn and (b) LZC and SampEn.

Kurtosis-Based Order On the other hand, the SOBI components computed from every signal epoch were ordered following increasing values of kurtosis. Although the epochs selected for this analysis had minimal artefactual activity, it might be possible that some artefacts, such as the CA, appeared in the recordings. Therefore, the very first and last components according to the kurtosis ranking might contain undesired signals. Fig. 7.23 illustrates the MF, SpecEn, LZC and SampEn values for the training set after the SOBI components of each epoch were ordered by increasing kurtosis.

Fig. 7.23 indicates that AD patients had abnormally low values of the signal characterising features. However, there was not clear structure in the kurtosis ordering. Only the very first and last components, which might be related to artefactual activity, had values notably different from the rest of the components. This was corroborated by the one-way ANOVA (age as a covariate) whose p -values are illustrated in Fig. 7.24.

Extended Infomax

In contrast to the two previous BSS algorithms, eInfoMax utilises HOS in the signal decomposition. In this case, all five minutes of recording were fed into the BSS algorithm. Afterwards, only the signal epochs previously studied with AMUSE and SOBI were analysed. eInfoMax does not estimate the components in a fixed order (Escudero et al., 2009c). Therefore, the ordering criteria based on increasing MF and kurtosis were used.

MF-Based Order The MF, SpecEn, LZC and SampEn computed from the eInfoMax components were ordered by increasing MF and represented in Fig. 7.25 for the training set. Likewise the AMUSE and SOBI cases, this ordering criterion implied that the MF, SpecEn, LZC and SampEn of the components increased with their position in the ranking. Nevertheless, all features showed a relatively steep increase in their values for the first components and relative constant values for the remaining BSS sources.

The p -values of the ANOVA with age as a covariate computed for each signal feature and component index in the training set suggested that the

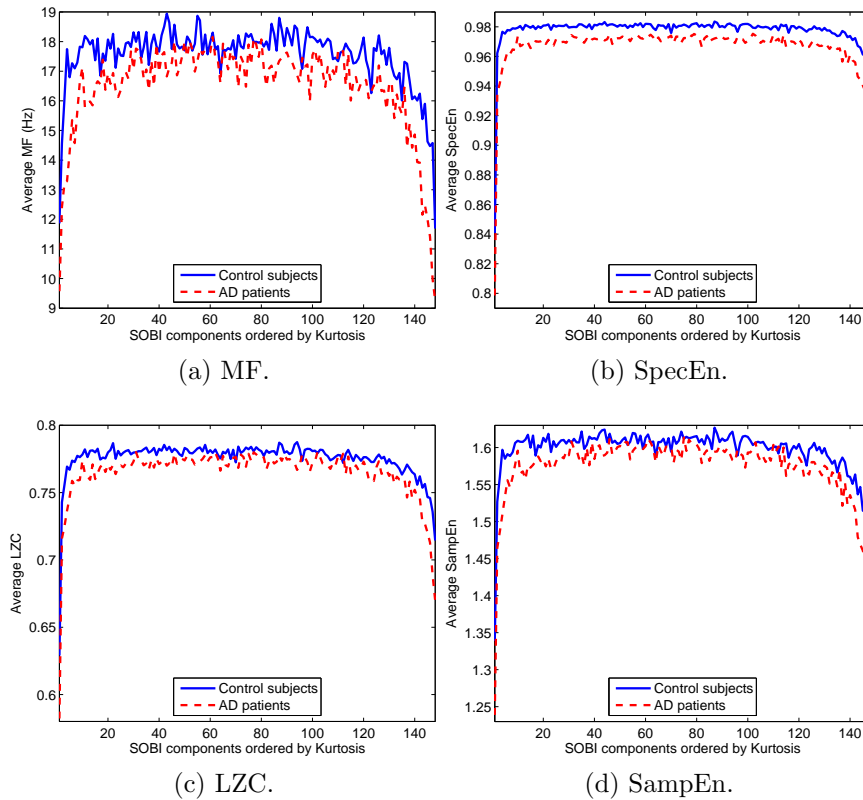


Figure 7.23: Average values of the signal features for the SOBI components orderer by their Kurtosis in AD patients and control subjects of the training set.

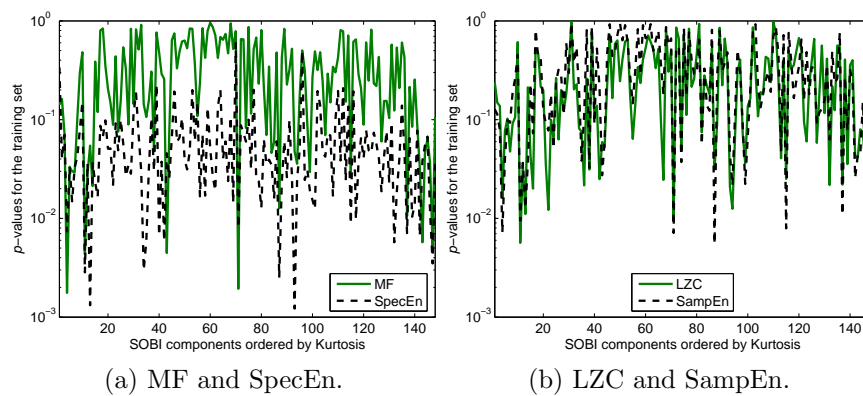


Figure 7.24: p -values of a one-way ANOVA with age as a covariate computed for SOBI components ordered by their Kurtosis in the training set for (a) MF and SpecEn and (b) LZC and SampEn.

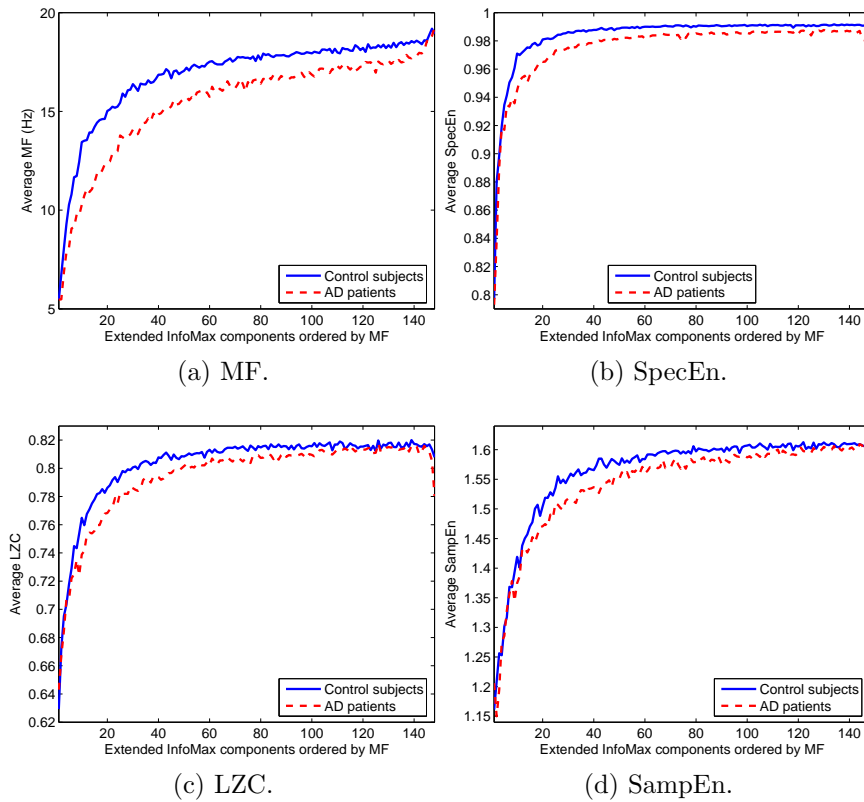


Figure 7.25: Average values of the signal features for the Extended InfoMax components ordered by their MF in AD patients and control subjects of the training set.

most significant differences usually appear in the first half of the component ranking. Those p -values are plotted in Fig. 7.26.

Kurtosis-Based Order Similarly to the ordering of the SOBI components based on kurtosis, the very first and last eInfoMax components ranked by increasing kurtosis were characterised by very low values of the signal features. The rest of the eInfoMax components had relatively stable values of the signal features. They are plotted in Fig. 7.27, which represents the MF, SpecEn, LZC and SampEn values for the training set after the eInfoMax components of each epoch were ordered by increasing kurtosis. Fig. 7.27 suggested that there was not any common structure with regard to where the main differences between both groups gathered. The corresponding p -values from the

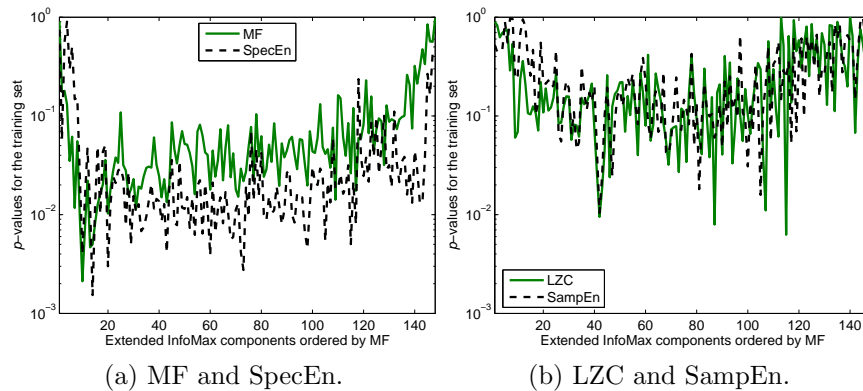


Figure 7.26: p -values of a one-way ANOVA with age as a covariate computed for the Extended InfoMax components ordered by their MF in the training set for (a) MF and SpecEn and (b) LZC and SampEn.

one-way ANOVA with age as a covariate are depicted in Fig. 7.28.

FastICA

FastICA utilises HOS in the signal decomposition. Similarly to eInfoMax, all five minutes of MEG recording were fed into the FastICA algorithm, although only the previously studied signal epochs were analysed.

MF-Based Order The values of MF, SpecEn, LZC and SampEn computed from the FastICA components ordered by increasing MF are represented in Fig. 7.29 for the training set. This ordering criterion implied that the MF, SpecEn, LZC and SampEn of the components increased with their position in the ranking. Similarly to the eInfoMax case, all features showed a relatively steep increase in their values for the first components and relative constant values for the remaining BSS sources.

The p -values of the ANOVA with age as a covariate computed for each signal feature and component index in the training set suggested that the most significant differences tended to be grouped. These p -values computed with a one-way ANOVA with age as a covariate for each signal feature and component are plotted in Fig. 7.30.

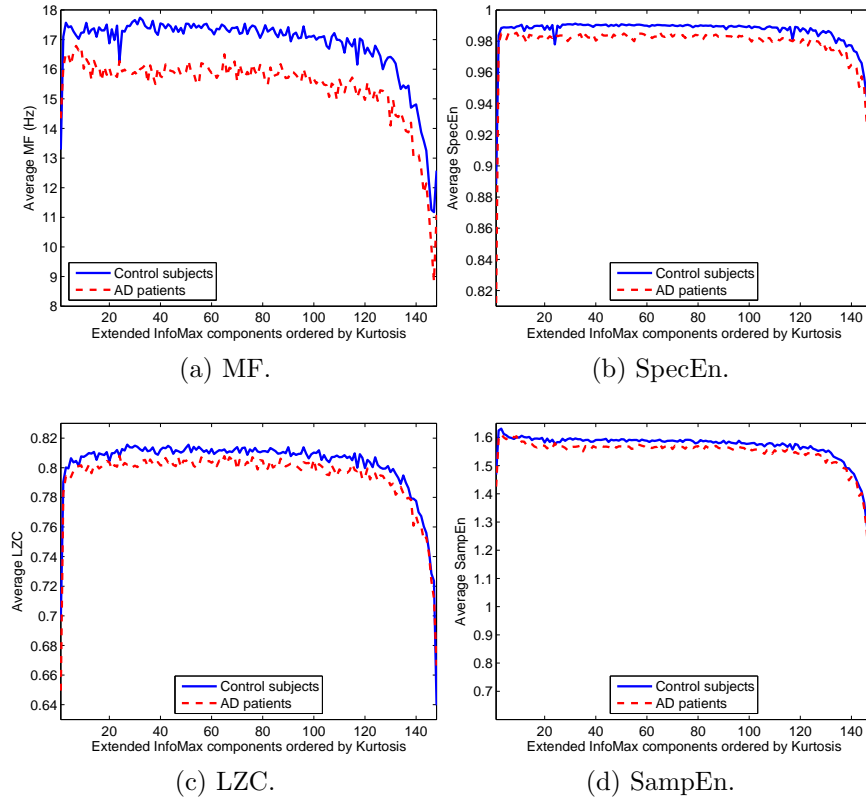


Figure 7.27: Average values of the signal features for the Extended InfoMax components orderer by their Kurtosis in AD patients and control subjects of the training set.

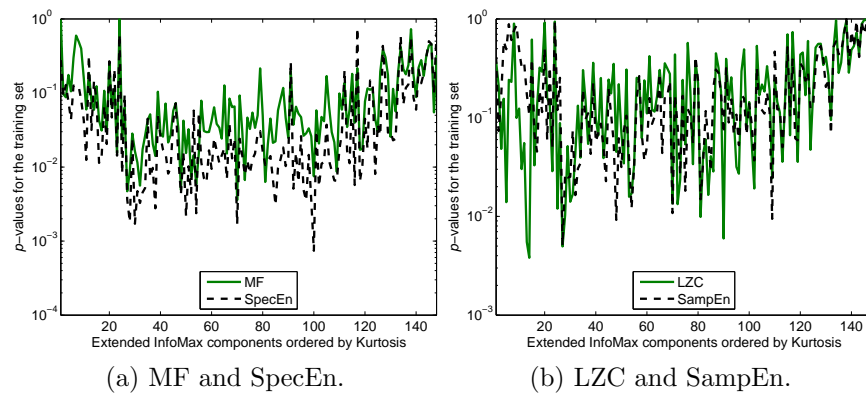


Figure 7.28: p -values of a one-way ANOVA with age as a covariate computed for the Extended InfoMax components ordered by their Kurtosis in the training set for (a) MF and SpecEn and (b) LZC and SampEn.

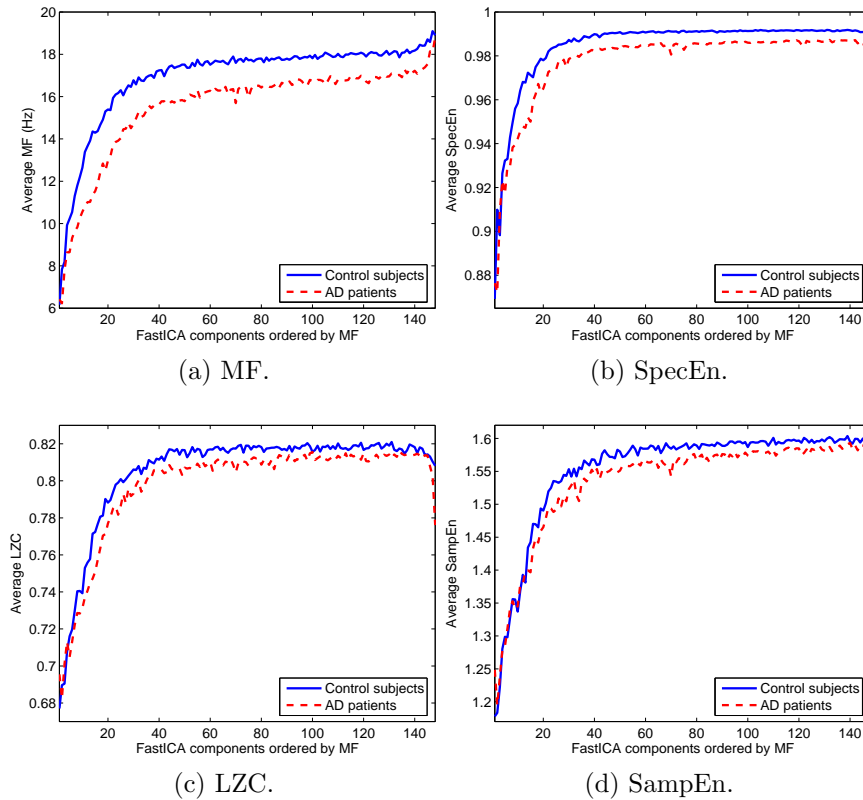


Figure 7.29: Average values of the signal features for the FastICA components orderer by their MF in AD patients and control subjects of the training set.

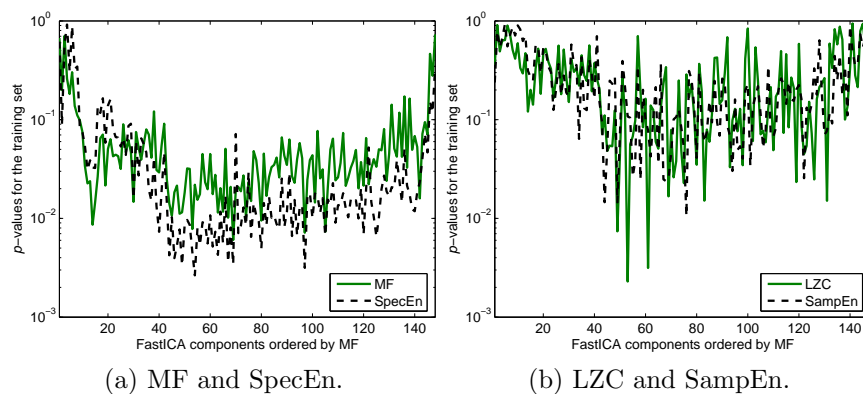


Figure 7.30: p -values of a one-way ANOVA with age as a covariate computed for the FastICA components ordered by their MF in the training set for (a) MF and SpecEn and (b) LZC and SampEn.

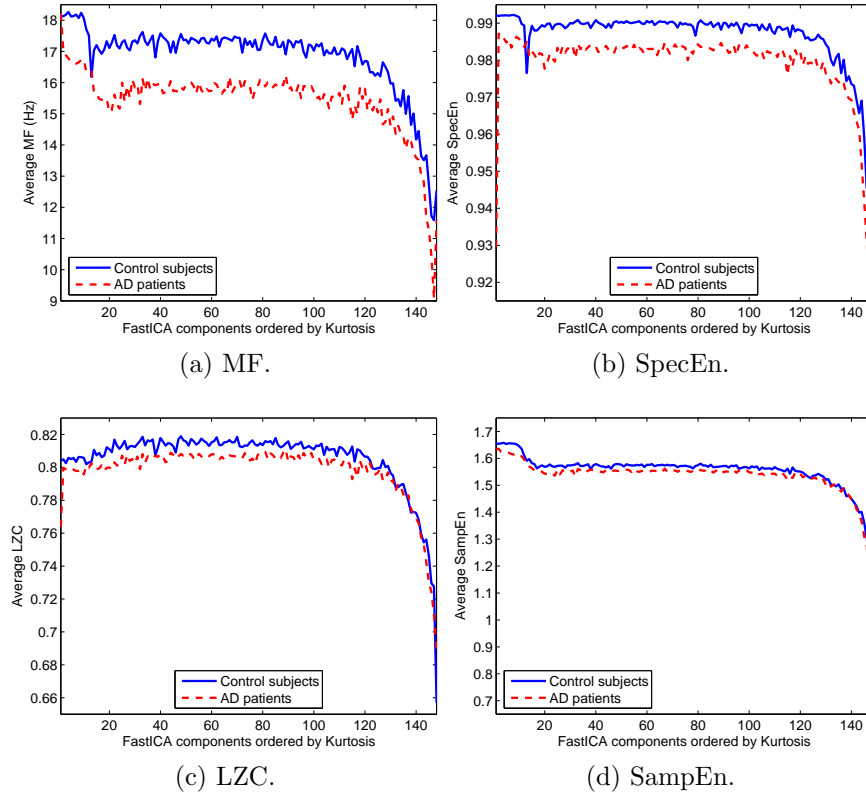


Figure 7.31: Average values of the signal features for the FastICA components order by their Kurtosis in AD patients and control subjects of the training set.

Kurtosis-Based Order Finally, the FastICA components were ordered by increasing values of kurtosis and the signal features MF, SpecEn, LZC and SampEn were calculated. These values appear in Fig. 7.31 for the subjects in the training set. Additionally, the p -values calculated with a one-way ANOVA (age as a covariate) for every component and signal feature are illustrated in Fig. 7.32.

7.4.3 Selection of the Most Significant Components

Visual inspection of the p -values shown in Fig. 7.20, Fig. 7.22, Fig. 7.24, Fig. 7.26, Fig. 7.28, Fig. 7.30 and Fig. 7.32 suggested that the most AD-sensitive components were often gathered together in relatively continuous

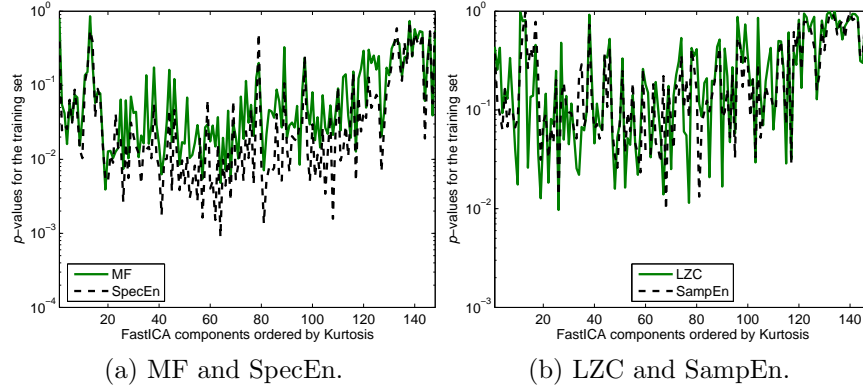


Figure 7.32: p -values of a one-way ANOVA with age as a covariate computed for the FastICA components ordered by their Kurtosis in the training set for (a) MF and SpecEn and (b) LZC and SampEn.

intervals.

Therefore, two small subsets of components — $\hat{\mathbf{s}}_{range}(t)$ — were chosen to partially reconstruct the MEG recordings — $\mathbf{x}_{partial}(t)$ — for every combination of BSS algorithm, ordering criterion and signal feature. These ranges were defined as the continuous intervals of 15 and 30 components (10% and 20% of all the 148 available BSS sources, respectively) which provided the lowest average p -value for every metric and type of preprocessing in the training set. Table 7.48 shows these ranges.

Comparing the ranges included in Table 7.48 with the MF values depicted in Fig. 7.19, Fig. 7.21, Fig. 7.25 and Fig. 7.29, it can be noticed that some of the ranges selected to partially reconstruct the MEG signals corresponded to activity centred in the α (8 Hz – 13 Hz) or β_1 (13 Hz – 19 Hz) bands. Therefore, spectral analyses of the MEG recordings focused on these bands (α and β_1) were also carried out to decide whether the application of the BSS (a spatial filtering technique) as a preprocessing offered any advantages over the most commonly used spectral filtering preprocessing. These spectral analyses consisted of an evaluation of the relative spectral power in the α and β_1 bands and the computation of the signal features MF, SpecEn, LZC and SampEn from the band-pass filtered MEG recordings in α and β_1 .

Table 7.48: Ranges of components (estimated from the training set) selected to partially reconstruct the MEG signals for each feature and type of pre-processing (BSS algorithm and ordering criterion).

Preprocessing	# comps.	MF	SpecEn	LZC	SampEn		
AMUSE	15 comps. (10%)	18 to 32	9 to 23	7 to 21	7 to 21		
	30 comps. (20%)	13 to 42	6 to 35	3 to 32	6 to 35		
Components ordered by MF	SOBI	15 comps. (10%)	20 to 34	20 to 34	20 to 34	20 to 34	
		30 comps. (20%)	5 to 34	10 to 39	9 to 38	9 to 38	
	eInfoMax	15 comps. (10%)	9 to 23	59 to 73	30 to 44	37 to 51	
		30 comps. (20%)	7 to 36	46 to 75	15 to 44	61 to 90	
	FastICA	15 comps. (10%)	62 to 76	44 to 58	42 to 56	62 to 76	
		30 comps. (20%)	45 to 74	40 to 69	43 to 72	67 to 96	
	Components ordered by Kurtosis	SOBI	15 comps. (10%)	2 to 16	79 to 93	2 to 16	2 to 16
			30 comps. (20%)	119 to 148	119 to 148	1 to 30	1 to 30
eInfoMax		15 comps. (10%)	27 to 41	96 to 110	25 to 39	41 to 55	
		30 comps. (20%)	26 to 55	57 to 86	25 to 54	27 to 56	
FastICA		15 comps. (10%)	52 to 66	48 to 62	53 to 67	52 to 66	
		30 comps. (20%)	48 to 77	39 to 68	48 to 77	39 to 68	

7.4.4 Evaluation of the Classification Improvement

This Section is devoted to comparing the classification performance achieved by means of signal features extracted from the raw MEG recordings — $\mathbf{x}(t)$ — with that obtained from the BSS preprocessed signals — $\mathbf{x}_{partial}(t)$.

The MEG signals of the training set were partially reconstructed with

Table 7.49: Classification results and AUC obtained from the MEG recordings without any preprocessing in the test set using the decision rules developed with the training set.

	MF	SpecEn	LZC	SampEn
Accuracy (%)	77.4	61.3	61.3	58.1
Sensitivity (%)	88.9	55.6	55.6	72.2
Specificity (%)	53.9	69.2	69.2	38.5
AUC	0.855	0.727	0.786	0.645

the ranges specified in Table 7.48. In each case, an average value of MF, SpecEn, LZC or SampEn per channel and subject was computed from the corresponding signals $\mathbf{x}_{\text{partial}}(t)$ or $\mathbf{x}(t)$. The 148 values of MF, SpecEn, LZC or SampEn were averaged for every subject. Owing to the high spatial density of the MEG channels, this dimensionality reduction is helpful to simplify the analysis and the interpretation of the results. Thus, every classification analysis was carried out using one mean value of MF, SpecEn, LZC or SampEn per subject (Escudero et al., 2008d, 2009b,c).

An LDA was applied to the average values of these signal features of the original MEG recordings — $\mathbf{x}(t)$ — and the partially reconstructed MEG signals — $\mathbf{x}_{\text{partial}}(t)$ — to find the optimal subject classification rules in the training set. These classification rules were evaluated, without further alteration, with the test set. Therefore, the accuracy, sensitivity, specificity and AUC results detailed in this Section were computed from the test set (Escudero et al., 2009b).

Firstly, the classification rules derived from the training data were applied to the unseen test set for the case where no BSS preprocessing and component selection procedure was used — $\mathbf{x}(t)$. These results are depicted in Table 7.49.

For the sake of comparison with more common techniques, the classifications achieved with the relative power in the α and $\beta 1$ bands are shown in Table 7.50. As it was mentioned, these bands were inspected because, in several cases, the subsets of BSS components selected to partially reconstruct the signals seemed to contain activity centred around α and $\beta 1$.

Table 7.50: Classification results and AUC from the relative power in the α (8 Hz–13 Hz) and β_1 (13 Hz–19 Hz) bands in the original MEG recordings for the test set using the decision rules developed with the training set.

	Relative power in α	Relative power in β_1
Accuracy (%)	54.8	67.7
Sensitivity (%)	50.0	72.2
Specificity (%)	61.5	61.5
AUC	0.628	0.752

Table 7.51: Classification results and AUC obtained from the α -filtered MEG signals in the test set using the decision rules developed with the training set.

α -filtered (8 Hz–13 Hz)	MF	SpecEn	LZC	SampEn
Accuracy (%)	77.4	41.9	51.6	35.5
Sensitivity (%)	72.2	44.4	61.1	33.3
Specificity (%)	84.6	38.5	38.5	38.5
AUC	0.872	0.598	0.504	0.615

Moreover, the classifications achieved with the MF, SpecEn, LZC and SampEn calculated from MEG signals filtered in the α and β_1 bands are detailed in Table 7.51 and Table 7.52, respectively. Although the accuracy or AUC values detailed in Table 7.50, Table 7.51 and Table 7.52 sometimes increased in comparison with Table 7.49, the improvement derived from studying these individual frequency bands is not completely significant.

Table 7.52: Classification results and AUC obtained from the β_1 -filtered MEG signals in the test set using the decision rules developed with the training set.

β_1 -filtered (13 Hz–19 Hz)	MF	SpecEn	LZC	SampEn
Accuracy (%)	61.3	58.1	64.5	61.3
Sensitivity (%)	66.7	55.6	61.1	55.6
Specificity (%)	53.8	61.5	69.2	69.2
AUC	0.624	0.560	0.705	0.645

Table 7.53: Classification results and AUC obtained from the partially reconstructed MEG signals with the AMUSE preprocessing in the test set using the algorithms developed with the training set (Acc: Accuracy, Sen: Sensitivity, Spe: Specificity).

A	MF		SpecEn		LZC		SampEn	
# comps. retained	30 (20%)	15 (10%)	30 (20%)	15 (10%)	30 (20%)	15 (10%)	30 (20%)	15 (10%)
Acc. (%)	77.4	71.0	71.0	67.7	67.7	74.2	80.7	80.7
Sen. (%)	77.8	66.7	77.8	72.2	72.2	72.2	83.3	83.3
Spe. (%)	76.9	76.9	61.5	61.5	61.5	76.9	76.9	76.9
AUC	0.878	0.868	0.786	0.782	0.838	0.838	0.872	0.863

On the other hand, Table 7.53 details the classification results obtained thanks to the AMUSE preprocessing. This BSS algorithm provides an inherent order for the extracted components and the ranges of components selected for the partial reconstruction were quite consistent. Therefore, this type of preprocessing improved all AUC values and the accuracy of all signal features expect for MF (Escudero et al., 2009b). Furthermore, these were the highest rises in several cases.

SOBI is similar to AMUSE in the sense that both techniques employ SOS to decompose the signals. However, SOBI does not have any intrinsic order of the estimated components. Following the results offered by AMUSE, the components were ordered by increasing MF (Escudero et al., 2009c). The ranges of selected components tended to be consistent among the four signal features. The corresponding classification results are introduced in Table 7.54. It can be seen that the accuracy and AUC improved in most cases. However, the enhancement of the classification quality was not as important as for AMUSE.

In comparison with the previous case, eInfoMax and FastICA are HOS-based BSS methods. For the sake of a complete comparison, the components computed with these techniques were also ranked by increasing MF. The classification results for the eInfoMax and FastICA preprocessings are presented in Table 7.55 and Table 7.56, in that order. Nevertheless, these BSS

Table 7.54: Classification results and AUC obtained from the partially reconstructed MEG signals with the SOBI preprocessing ordered following the MF in the test set using the algorithms developed with the training set (Acc: Accuracy, Sen: Sensitivity, Spe: Specificity).

S-MF	MF		SpecEn		LZC		SampEn	
# comps. retained	30 (20%)	15 (10%)	30 (20%)	15 (10%)	30 (20%)	15 (10%)	30 (20%)	15 (10%)
Acc. (%)	77.4	71.0	74.2	58.1	64.5	67.7	67.7	71.0
Sen. (%)	88.9	77.8	72.2	61.1	61.1	72.2	72.2	77.8
Spe. (%)	61.5	61.5	76.9	53.8	69.2	61.5	61.5	61.5
AUC	0.876	0.829	0.782	0.709	0.808	0.769	0.803	0.782

Table 7.55: Classification results and AUC obtained from the partially reconstructed MEG signals with the Extended InfoMax preprocessing ordered following the MF in the test set using the algorithms developed with the training set (Acc: Accuracy, Sen: Sensitivity, Spe: Specificity).

eI-MF	MF		SpecEn		LZC		SampEn	
# comps. retained	30 (20%)	15 (10%)	30 (20%)	15 (10%)	30 (20%)	15 (10%)	30 (20%)	15 (10%)
Acc. (%)	74.2	74.2	67.7	67.7	64.5	61.3	64.5	54.8
Sen. (%)	83.3	77.8	55.6	50.0	66.7	61.1	55.6	33.3
Spe. (%)	61.5	69.2	84.6	92.3	61.5	61.5	76.9	84.6
AUC	0.808	0.812	0.697	0.679	0.679	0.701	0.675	0.692

and component selection procedures did not produce as many improvements in the accuracy or AUC as the cases of SOBI ordered by MF or AMUSE.

Finally, the classification results obtained with a preprocessing that orders the extracted components by increasing kurtosis appear in Table 7.57, Table 7.58 and Table 7.59 for SOBI, eInfoMax and FastICA, respectively. Among these cases, the best results are offered by eInfoMax. However, the improvement is more modest than in the preprocessing done with SOBI and the MF ordering or, specially, AMUSE.

Table 7.56: Classification results and AUC obtained from the partially reconstructed MEG signals with the FastICA preprocessing ordered following the MF in the test set using the algorithms developed with the training set (Acc: Accuracy, Sen: Sensitivity, Spe: Specificity).

F-MF	MF		SpecEn		LZC		SampEn	
# comps. retained	30 (20%)	15 (10%)	30 (20%)	15 (10%)	30 (20%)	15 (10%)	30 (20%)	15 (10%)
Acc. (%)	64.5	71.0	61.3	61.3	58.1	61.3	54.8	54.8
Sen. (%)	55.6	55.6	50.0	50.0	50.0	50.0	38.9	38.9
Spe. (%)	76.9	92.3	76.9	76.9	69.2	76.9	76.9	76.9
AUC	0.825	0.816	0.722	0.701	0.688	0.701	0.650	0.679

Table 7.57: Classification results and AUC obtained from the partially reconstructed MEG signals with the SOBI preprocessing ordered following the Kurtosis in the test set using the algorithms developed with the training set (Acc: Accuracy, Sen: Sensitivity, Spe: Specificity).

S-Kurt	MF		SpecEn		LZC		SampEn	
# comps. retained	30 (20%)	15 (10%)	30 (20%)	15 (10%)	30 (20%)	15 (10%)	30 (20%)	15 (10%)
Acc. (%)	71.0	71.0	61.3	61.3	64.5	61.3	58.1	61.3
Sen. (%)	88.9	66.7	72.2	55.6	50.0	44.4	50.0	50.0
Spe. (%)	46.2	76.9	46.2	69.2	84.6	84.6	69.2	76.9
AUC	0.876	0.829	0.718	0.688	0.812	0.782	0.782	0.782

Table 7.58: Classification results and AUC obtained from the partially reconstructed MEG signals with the Extended InfoMax preprocessing ordered following the Kurtosis in the test set using the algorithms developed with the training set (Acc: Accuracy, Sen: Sensitivity, Spe: Specificity).

eI-Kurt	MF		SpecEn		LZC		SampEn	
# comps. retained	30 (20%)	15 (10%)	30 (20%)	15 (10%)	30 (20%)	15 (10%)	30 (20%)	15 (10%)
Acc. (%)	64.5	67.7	67.7	58.1	64.5	67.7	64.5	71.0
Sen. (%)	66.7	72.2	61.1	38.9	61.1	66.7	66.7	72.2
Spe. (%)	61.5	61.5	76.9	84.6	69.2	69.2	61.5	69.2
AUC	0.795	0.778	0.786	0.696	0.778	0.795	0.782	0.782

Table 7.59: Classification results and AUC obtained from the partially reconstructed MEG signals with the FastICA preprocessing ordered following the Kurtosis in the test set using the algorithms developed with the training set (Acc: Accuracy, Sen: Sensitivity, Spe: Specificity).

F-Kurt	MF		SpecEn		LZC		SampEn	
# comps. retained	30 (20%)	15 (10%)	30 (20%)	15 (10%)	30 (20%)	15 (10%)	30 (20%)	15 (10%)
Acc. (%)	64.5	61.3	67.7	64.5	61.3	61.3	64.5	64.5
Sen. (%)	72.2	66.7	66.7	66.7	61.1	61.1	72.2	72.2
Spe. (%)	53.8	53.8	69.2	61.5	61.5	61.5	53.8	53.8
AUC	0.786	0.812	0.722	0.671	0.739	0.756	0.744	0.765

7.5 Adaptive Extraction of Brain Rhythms

This Section of the Doctoral Thesis summarises the results obtained from computing a connectivity analysis of adaptively extracted brain rhythms over diverse scalp regions. This synchronisation analysis was based on the $\text{Coh}(f)$ measure, while the extraction of the signals employed a procedure composed of an EMD, a k -means clustering and a cBSS.

7.5.1 Estimation of References

First of all, MEG epochs of 10 s with minimal artefactual activity were selected for analysis. Furthermore, the cBSS-based artefact removal procedure described in Section 6.5.1 was applied to reduce the presence of the CA in the recordings. This artefact removal scheme was selected among the alternatives presented in Section 7.3 for its simplicity and good performance, which ranked among the best described in that Section.

Secondly, an EMD was applied to each single MEG channel to obtain the rhythmical components enclosed in it. To illustrate this procedure, Fig. 7.33 plots a MEG epoch recorded from a control subject at the channel A7 (LeftC) and Fig. 7.34 depicts the IMFs computed from its EMD following the order in which they are calculated.

Once all IMFs had been computed, their spectral content was charac-

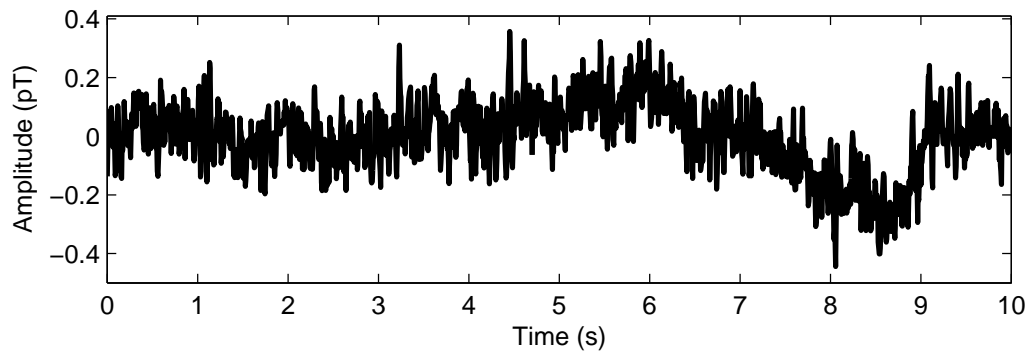


Figure 7.33: MEG channel from a control subject.

terised with the weighted average value of the instantaneous frequency, $\langle\omega\rangle$ — see Eq. (6.41). This value offered information about the average spectral localisation of every IMF. Therefore, it was possible to decide which IMFs computed from a particular scalp region (see Fig. 6.6) belong to the δ , θ , α or β band.

The cBSS might be helpful to estimate the brain rhythmical activity recorded over a particular scalp region for a specific spectral band. In order to do so, all MEG signals belonging to that region were fed into the cBSS technique. Nevertheless, a proper reference of the activity of interest was needed. This reference was selected as one of the previously computed IMFs. This selection was performed by means of an automatic k -means clustering procedure.

The $\langle\omega\rangle$ values of all IMFs of a particular region that ranged between 1 Hz and 30 Hz were fed to a k -means with the number of clusters set to four. After the convergence of this process, the IMF with the closest value of $\langle\omega\rangle$ to the centroid of each cluster was selected as the reference for the cBSS. Fig. 7.35 illustrates the distributions of the $\langle\omega\rangle$ values of the IMFs computed for the LeftC of the same control subject as in Fig. 7.33 and Fig. 7.34. Additionally, the vertical lines in Fig. 7.35 indicate the positions of the four centroids. It is clear that the $\langle\omega\rangle$ values tended to gathered together in groups relatively well localised in the δ , θ , α and β bands.

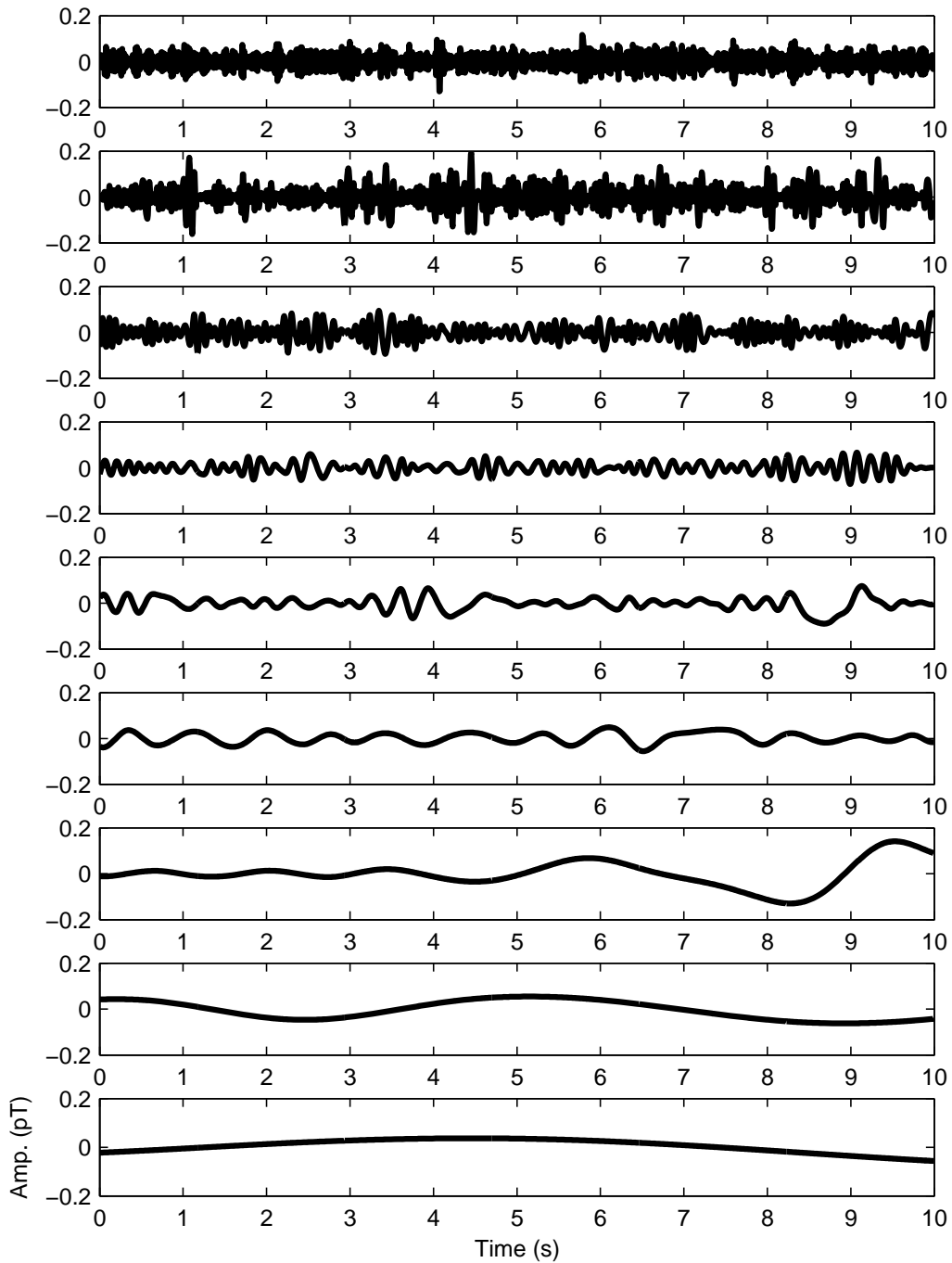


Figure 7.34: Example of the IMFs obtained from the EMD of a MEG channel acquired from a control subject. The IMFs are plotted following their order of extraction.

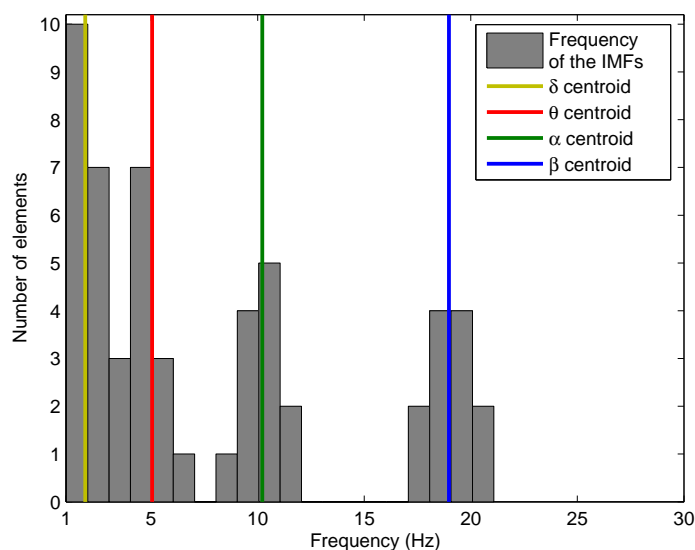


Figure 7.35: Example of the reference selection for the cBSS with the k -means clustering procedure.

7.5.2 Descriptive Statistical Analysis

With the references computed for each spectral band and region, a cBSS procedure was used to extract the corresponding MEG activity for each signal epoch. Then, the $\text{Coh}(f)$ between the activity signals of all possible pairs of regions was computed and the average level of $\text{Coh}(f)$ for the band and pair of interest was computed across all epochs. Therefore, a value of $\text{Coh}(f)$ per band, pair of regions and subject was obtained.

A qualitative statistical analysis was performed to gain insight into the $\text{Coh}(f)$ values of AD patients and control subjects considering the spectral band — δ (1 Hz to 4 Hz), θ (4 Hz to 8 Hz), α (8 Hz to 13 Hz) and β (13 Hz to 30 Hz) — and the pair of regions involved.

First of all, a repeated measures ANOVA was done with the Greenhouse-Geisser correction for the degrees of freedom due to the lack of sphericity. The group was the inter-subjects factor (AD patients vs. control subjects). The two intra-subject factors were band (four levels: δ , θ , α and β) and pair of regions — 28 levels corresponding to all possible combinations of LeftC, RightC, LeftA, RightA, LeftL, RightL, LeftP and RightP.

This analysis reported significant effects involving the factors ‘Band’ ($p =$

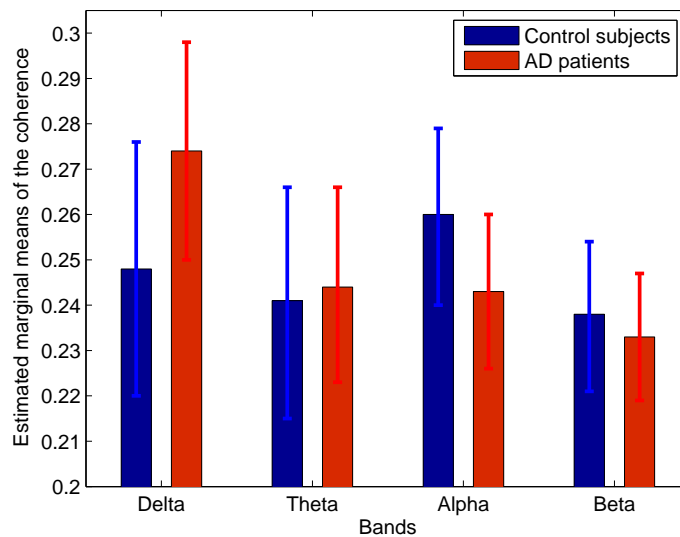


Figure 7.36: Estimated marginal means plotted with the 95% confidence interval of the $\text{Coh}(f)$ for AD patients and control subjects in each spectral band.

0.0096), ‘Band×Diagnosis’ ($p = 0.0473$), ‘Pair’ ($p < 0.0001$) and ‘Band×Pair’ ($p < 0.0001$). The inspection of the estimated marginal means for both subject groups considering each band separately revealed slight differences in the corresponding levels of $\text{Coh}(f)$. Fig. 7.36 depicts these estimated marginal means (Searle et al., 1980) with the 95% confidence interval. It suggests that AD decreased the level of $\text{Coh}(f)$ in the α and β bands while it increased the connectivity in δ .

To gain further insight, a repeated measures ANOVA with the Greenhouse-Geisser correction was applied to every individual band. Hence, the only intra-subject factor was the pair of regions (28 levels).

For the δ band, there was a significant effect due to the ‘Pair’ factor ($p < 0.0001$) but not to ‘Diagnosis’ ($p = 0.1635$). The effect of ‘Pair’ is illustrated in Fig. 7.37, which represents the estimated marginal means (without regard of group) of every pair of regions for the δ band. It can be seen that the level of $\text{Coh}(f)$ seemed to depend on the distance between the regions. For instance, the overall $\text{Coh}(f)$ level between regions located over the same hemisphere (blue) is usually higher than for regions of different hemispheres.

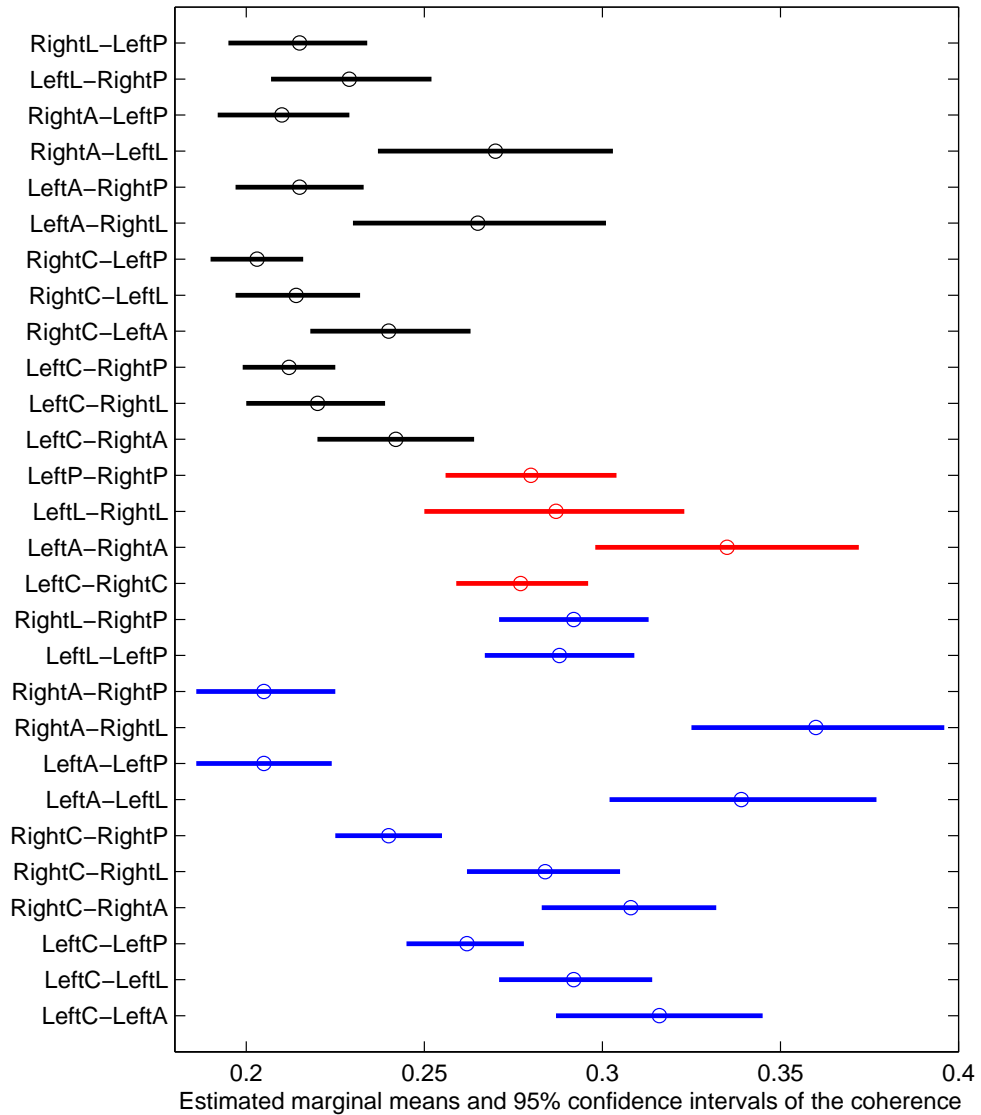


Figure 7.37: Example of the estimated marginal means and the 95% confidence interval for the $\text{Coh}(f)$ of the pairs of regions in the δ band. Pairs over the same hemisphere are plotted in blue, whereas red is used to indicate pairs composed of the same regions located over different hemispheres. The black colour represents pairs composed of totally different scalp regions and hemispheres.

Table 7.60: Classification results of the coherence measures computed from the adaptively extracted brain rhythms. Accuracy, sensitivity and specificity were calculated with a stepwise LDA. The number of variables automatically selected by this procedure and the related AUC are also shown.

	δ	θ	α	β	All bands
Accuracy (%)	67.7	46.8	69.4	43.5	96.8
Sensitivity (%)	63.9	47.2	66.7	44.4	94.4
Specificity (%)	73.1	46.2	73.1	42.3	100.0
Variables in the stepwise LDA	1	*	3	*	11
AUC	0.713	0.874	0.815	0.912	1.000

The results for the θ , α and β bands were similar to those of δ . There was a significant effect due to the ‘Pair’ ($p < 0.0001$ in all cases) but not to ‘Diagnosis’ ($p = 0.8145$ for θ , $p = 0.2143$ for α and $p = 0.6611$ for β). The dependence of the $\text{Coh}(f)$ values on the distance between regions was similar to that plotted in Fig. 7.37, also suggesting a dependence on the distance between regions.

7.5.3 Classification Analysis

A descriptive statistical analysis was presented in Section 7.5.2. Nevertheless, it is also important to quantitatively evaluate the ability of those features to distinguish AD patients from control subjects. This is due to the fact that the very subtle differences in some bands and pairs of regions might complement themselves in the subject classification. Thus, a stepwise LDA with a leave-one-out cross-validation procedure was applied to the $\text{Coh}(f)$ values computed from every band. This LDA provided the accuracy, sensitivity and specificity rates. The $\text{Coh}(f)$ values obtained for all bands were also fed into the stepwise LDA simultaneously. Table 7.60 shows these results. Additionally, the number of variables selected by the stepwise LDA and the AUC associated with the discriminant function are indicated.

In the cases where no variable fulfilled the criterion to enter into the classification rule (indicated with an ‘*’ in Table 7.60), the LDA was trained with all variables at the same time. Of note is that the ‘All bands’ case

actually included variables from all bands (δ , θ , α and β) in the classifier to achieve an accuracy over 95% with an associated AUC equal to 1.

Chapter 8

Discussion

Contents

8.1	Introduction	250
8.2	Preprocessing	252
8.3	Consistency Evaluation	254
8.4	Artefact Removal	256
8.5	Feature Enhancement	263
8.6	Adaptive Extraction of Brain Rhythms	272
8.7	Limitations of the Study	277

Bearing in mind the results reported in Chapter 7, the main findings of this Doctoral Thesis will now be discussed considering the state of the art in the application of BSS methods to MEG background activity and AD. After a brief introduction, specific sections will be devoted to revising and commenting the results related to each of the problems and applications addressed in this Doctoral Thesis. These include the preprocessing and dimension reduction procedure, the evaluation of the consistency of the decompositions computed by BSS techniques, the artefact removal based on BSS, the use of BSS as a preprocessing to enhance features associated with AD and the adaptive extraction of brain rhythms with cBSS to perform a connectivity analysis in AD. Finally, the main limitations of the studies carried out in this Doctoral Thesis will be acknowledged.

8.1 Introduction

The MEG is the non-invasive recording of the tiny magnetic fields generated by the brain activity. This neurophysiological technique offers direct measurements of the brain cortex activity without having to interpret the information on the basis of vascular or metabolic changes (Hämäläinen et al., 1993; Hari, 2004). The temporal and spatial resolution of the MEG is high. The magnetic recordings depend on neither any reference point nor the resistive properties of the extra-cerebral tissues (Hämäläinen et al., 1993; Hari et al., 2000). However, the extremely weak amplitude of the brain magnetic fields needs additional procedures to attenuate the noisy signals produced by the extra-cerebral sources affecting the MEG (Hari et al., 2000). The eye movements and blinks and the cardiac activity, together with the PLN, can contaminate the MEG recordings (Antervo et al., 1985; Hämäläinen et al., 1993; Jousmäki and Hari, 1996; Sternickel and Braginski, 2006). Therefore, it is necessary to deal with these artefacts in the MEG recordings. It is important to note that the recording of MEG signals could provide relevant information to the characterisation or diagnosis of AD (Hornero et al., 2009b; Stam, 2005, 2010).

AD is a slowly progressive neurodegenerative disorder. It produces memory loss and diverse cognitive and behavioural symptoms that progressively impair the activities of daily living (Blennow et al., 2006; Cummings, 2004). AD patients often die between seven and ten years after diagnosis (Masters et al., 2006). The symptoms of this dementia worsen gradually until, in the last stages of the dementia, the patients are utterly dependent on their caregivers (Blennow et al., 2006; Jeong, 2004; Nestor et al., 2004). AD is the most common type of dementia in the Western World. It accounts for 50% to 60% of all cases and it shows an almost exponential increase with age (Blennow et al., 2006; Cummings, 2004). Therefore, AD poses very serious economic and social problems to modern societies (Ferri et al., 2006). It is expected that diagnostic testing for AD will contribute only marginally to the cost associated with this dementia. Considering this and that a definite diagnosis of AD can only be made by necropsy (Blennow et al., 2006), the importance of developing methods to help in the diagnosis and characterisation of this disorder becomes clear. Eventually, this may lead to reduce that AD-related cost (Lahiri et al., 2002).

BSS techniques have recently become popular in the field of biomedical signal processing (James and Hesse, 2005; Vigário and Oja, 2008). These data-driven methods can decompose multi-dimensional biomedical recordings into their constituent underlying components (Cichocki and Amari, 2002; Hyvärinen et al., 2001). Hence, they are used in biomedical signal processing to separate, or extract, underlying statistically independent components of interest in biomedical signals. These techniques can provide useful perspectives to inspect the MEG recordings (James and Hesse, 2005; Vigário and Oja, 2008).

Therefore, this Doctoral Thesis aimed at applying *BSS methods to help in the analysis and interpretation of MEG background activity from AD patients and elderly control subjects*. This global objective was split into the following applications of BSS techniques to the data at hand:

- Comparison of the similarity of the decompositions provided by several BSS algorithms for real MEG signals.

- Automatic and objective assessment of the removal of artefactual cardiac and ocular activities, as well as the power electrical noise, that contaminate the MEG recordings.
- Improvement of the classification of MEG signals acquired from AD patients against those recorded from healthy elderly subjects.
- Adaptive extraction of rhythmic brain activity suitable to be analysed with connectivity measurements from the MEG recordings.

8.2 Preprocessing

This Section discusses the results obtained in the testing of the preprocessing (see Section 7.1). This preprocessing does not constitute a BSS algorithm. Actually, it relies on the classical projection technique of FA (Cao et al., 2003; Escudero et al., 2007b; Ikeda and Toyama, 2000). However, this preprocessing is important for several reasons (Escudero et al., 2007b; Hesse, 2008; Hyvärinen et al., 2001; Ikeda and Toyama, 2000; Vigário and Oja, 2008):

1. The number of inner meaningful components may be less than the number of available channels for analysis.
2. A dimensionality reduction may sometimes be necessary to avoid ‘overfitting’.
3. A dimensionality reduction may help to reduce the importance of the external noise.

The whole preprocessing included the estimation of the optimum number of inner components (n) and a quasi-whitening of the data that considered the estimated external noise. Before applying this preprocessing to the real MEG signals, it was tested on synthetic data (Escudero et al., 2007b). The results showed that the MDL and the FA-based preprocessing estimated both the value of n and the power of the external noise accurately for $\text{SNR} > 3$ dB (Escudero et al., 2007b). In the synthetic data set, n was set to the correct

value of 11 and the average E_{norm} was always lower than 3.5% (Escudero et al., 2007b).

Moreover, the performance of the MDL metric in the estimation of n was compared with three commonly used standard PCA-based criteria: “PCA: 95% limit”, “PCA: 99% limit” and “PCA: larger than 1%” (Escudero et al., 2007b; James and Hesse, 2005; Ting et al., 2006). The results clearly indicated that these PCA approaches may cause ‘overfitting’ in the BSS since they tend to over-estimate n for SNRs below a certain value. What is more, the estimation of n was neither accurate for high SNR values (Escudero et al., 2007b). On the other hand, for very high external noise power, the MDL metric tended to under-estimate the number of inner components. Obviously, the values of n should be as accurate as possible, thus avoiding both over- and under-estimations. Nevertheless, a mild under-estimation of n will cause some underlying real components to be mixed in the extracted BSS components. In some applications, this may be better than a severe ‘overfitting’ (Escudero et al., 2007b).

In summary, the *MDL provided more accurate values for the parameter n than the methodologies based on PCA*. These results confirmed that the latter not only are subjective and depend on the shape of the eigenvalue spectrum (James and Hesse, 2005), but also could be prone to over-estimation of the number of components when the external noise is not negligible (Escudero et al., 2007b). These problems are avoided with the MDL approach.

As for the real MEG recordings, the dependence of the preprocessing on the signal length was studied by analysing epochs of 10 s, 20 s, 40 s, 60 s and 90 s. As it could be expected, the value of n increased with the epoch length. This means that longer signals tend to be composed of more inner sources or, at least, need to consider more components to obtain an optimum decomposition. For each case, the corresponding power of the external noise was estimated. In most cases, it was found that the additive noise accounted between 2.8% and 5.6% of the total power in the recordings. The values of SNR associated with these noise powers range from 12.3 dB to 15.4 dB. For these SNRs, the evaluation of the preprocessing made on synthetic data suggested that the MDL was appropriate to calculate n . Finally, it must be noticed

that almost all power assigned to the inner components was considered in the following steps as shown in Fig. 7.5.

All these results are supported by the recently published studies about the model order selection by Hesse (2007, 2008). These pieces of research investigated the performance of diverse approaches based on PCA and FA to estimate the number of BSS components in real EEG and MEG recordings. Those results indicated that probabilistic PCA and FA models yielded estimations of the dimensionality that are more reliable and independent of the signal power than commonly used PCA approaches (Hesse, 2007, 2008). The estimated values of n were about one third of the measurement space dimension (Hesse, 2007). In our case, the number of components was usually lower than one third of channels, specially for the shorter epochs. Therefore, this suggested that, in the case of MEG equipment, more channels do not necessarily reflect more brain signals (Hesse, 2007). What is more, the data dimension reduction is supported by the statistical properties of the signal and the FA models may offer an appropriate description of the brain recordings. These results pinpointed the importance of considering the additive noise to be ‘anisotropic’ (Hesse, 2007, 2008).

8.3 Consistency Evaluation

This Section discusses the main results of the comparison carried out between the decompositions computed with the following BSS algorithms: AMUSE, SOBI, JADE, eInfoMax and FastICA. In addition, five different epoch lengths were considered: 10 s, 20 s, 40 s, 60 s and 90 s. The data on which this discussion relies were presented in Section 7.2.

The algorithms were compared in terms of their consistency: the similarity between their decompositions for the same input data. Of note is that only one matrix \mathbf{P}^{qr} — see Eq. (6.10) — for each pair of algorithms was analysed, instead of both \mathbf{P}^{qr} and \mathbf{P}^{rq} . This decision was supported by the fact that the average differences for the F metric — defined in Eq. (6.11) — between \mathbf{P}^{qr} and \mathbf{P}^{rq} were always lower than 1.20%. Hence, by taking this decision, we tried to reduce the surplus complexity and redundancy of the

Table 8.1: Ratios of the number of data samples for each epoch length divided by the median value of the number of elements in \mathbf{A} .

Epoch length	Data samples	Median of n	Elements in \mathbf{A} (n^2)	Ratio
10 s	1695	29	841	2.02
20 s	3390	33	1089	3.11
40 s	6780	38	1444	4.67
60 s	10170	41	1681	6.05
90 s	15255	44	1936	7.88

problem (Escudero et al., 2009d).

The visual representation of the results provided by Fig. 7.6 clearly indicated that *the pair AMUSE–SOBI provided the most similar decomposition* for the analysed MEG signals. The principle beneath these two techniques is the simultaneous diagonalisation of several time-delayed cross-covariance matrices (Hyvärinen et al., 2001; James and Hesse, 2005). It was also found that *the decompositions of JADE and FastICA were characterised by a high degree of consistency*. This might be explained by the fact that the theoretical principles of both algorithms can be related (Hyvärinen et al., 2001).

Additionally, the algorithms tended to estimate *more similar* decompositions as longer signals were decomposed (Escudero et al., 2009d). This may be due to the fact that, although the value of the parameter n increased with the epoch length, the number of data samples increased more rapidly than the number of elements to be estimated in the matrix \mathbf{A} . This fact is illustrated in Table 8.1, which represents the number of available data samples and the median of the number of elements to be estimated in \mathbf{A} for each epoch length. It is clear that, the longer the epoch, the larger the ratio of the number of samples available to estimate \mathbf{A} . Hence, the decompositions may be considered more reliable for longer epochs (Onton et al., 2006), which could explain the ‘relative similarity’ of the BSS outcomes for long epochs.

The statistical analysis carried out individually for every epoch length pointed out the statistical significance of the similarity between the decomposition computed by AMUSE–SOBI and JADE–FastICA. On the other hand, the pairs AMUSE–JADE and SOBI–JADE used to provide the most dissim-

ilar separations of the MEG signals. A relatively consistent pattern was that a larger number of components made the outcomes of the separations calculated by AMUSE–SOBI and JADE–FastICA slightly more different. Surprisingly, the pairs AMUSE–eInfoMax and SOBI–eInfoMax computed BSS decompositions that were slightly more similar for larger values of n .

Previous studies have compared the performance of several BSS techniques from different perspectives. For instance, synthetic signals have been used to assess the quality of the BSS decomposition (Klemm et al., 2009) or to evaluate whether the BSS improved the automatic detection of artefacts in the EEG (Delorme et al., 2007). Artificially mixed EEG signals have also been analysed by Fitzgibbon et al. (2007) to compare the relative performance of a few BSS algorithms to isolate artefacts. Real MEG signals were decomposed by Zavala Fernández et al. (2006) to evaluate their ability to extract artefacts by comparing the contaminated components of different algorithms with reference signals. Several of these studies used small datasets or small numbers of channels and the evaluation of the algorithms was frequently based on subjective criteria (Klemm et al., 2009).

In contrast to those studies, we analysed the decompositions of actual MEG recordings *globally* to gain insight into the similarities and relationships between some of the most commonly used BSS algorithms. Instead of comparing a manually selected subset of components (Zavala Fernández et al., 2006), the entire decomposition was assessed as the metric F (Melissant et al., 2005) was computed on the basis of the whole mixing matrices \mathbf{A}^g and \mathbf{A}^r (Escudero et al., 2009d). It is important to note that the real sources are unknown (James and Hesse, 2005). Thus, the analyses only aimed at measuring the similarity between the results of the BSS algorithms and not at evaluating the actual quality of the separation (Escudero et al., 2009d).

8.4 Artefact Removal

This Section contains the discussion about the results of the artefact removal with BSS. The relevant results were detailed in Section 7.3. It is important to note that the main objective was to *objectively measure* the artefact removal

without having to rely on visual inspections of the data, which are subjective and time-consuming (Rong and Contreras-Vidal, 2006). Hence, this evaluation was based on measures designed *ad hoc* for every type of artefact: CA, OA and PLN. In this Doctoral Thesis, the evaluation of the artefact removal was carried out for several combinations of BSS technique, epoch length and artefact detection metric.

This application is motivated by the fact that the MEG is a non-invasive neurophysiological technique to record the magnetic field generated by neural currents parallel to the skull (Hämäläinen et al., 1993; Hari, 2004). This signal carries information about the brain activity and it does not require to interpret the data as a function of metabolic changes or a reference channel. Moreover, its temporal and spatial resolution is high (Hari, 2004; Hari and Forss, 1999). Nevertheless, MEG signals are contaminated with artefacts that can bias the analyses (Antervo et al., 1985; Hämäläinen et al., 1993; Jousmäki and Hari, 1996). Hence, it is important to develop methodologies to deal with the artefacts. BSS may be a suitable tool to achieve this aim (Escudero et al., 2007b; James and Hesse, 2005).

The BSS techniques were applied after the robust preprocessing and model order procedure. This was done in order to avoid ‘over-fitting’ and to perform a robust preprocessing of the recordings (Escudero et al., 2007b). Although it is not necessary, this preprocessing was also used before the cBSS in order to ensure that all decomposition or extraction methods were applied to the same input data. After the BSS, several artefact detection metrics and criteria helped to identify the BSS components with possible artefacts.

The BSS is the key element in this artefact removal procedure. It relies on the facts that the MEG channels acquire relatively redundant information at nearby channels and that the artefacts almost always affect more than one channel (James and Hesse, 2005). On the other hand, the main hypothesis needed to apply BSS is that the underlying components are independent (Cichocki and Amari, 2002; James and Hesse, 2005; Vigário and Oja, 2008). This hypothesis is reasonable for EEG and MEG data since the useful brain signals and the artefacts are anatomically and physiologically separated processes. This should be reflected in the statistical relationships between both

kinds of activity (Jung et al., 2000; Vigário and Oja, 2000).

Several pieces of research have aimed at isolating or removing the CA in MEG data (Barbati et al., 2004; Dammers et al., 2008; Escudero et al., 2007b; James and Gibson, 2003; Mantini et al., 2008; Rong and Contreras-Vidal, 2006; Sander et al., 2002; Vigário and Oja, 2000). For instance, Sander et al. (2002) detected the subspace of the CA attending at its spectrum, temporal evolution and topographical distribution. Rong and Contreras-Vidal (2006) developed an artificial neural network methodology to detect the CA in event-related MEGs by simultaneously considering the information provided by higher-order statistics, the spectrum and the scalp topography. The kurtosis and entropy values, together with the correlation level with a reference ECG was used by Barbati et al. (2004) to detect the CA. Nevertheless, neither kurtosis nor entropy were able to uniquely identify the cardiac components (Barbati et al., 2004). Entropic measures have also been applied in this setting in other studies (Dammers et al., 2008; Mantini et al., 2008). Moreover, Dammers et al. (2008) suggested the use of phase statistics in the recognition of this artefact. The asymmetry of the CA, quantified with the Skew, can also be helpful to detect this artefact (Escudero et al., 2007b; Shao et al., 2008; Shoker et al., 2005a). Alternatively, the variability of the corresponding scalp maps have been introduced as a possibly useful metric to mark the CA (Shao et al., 2008).

Considering these studies, diverse metrics were employed in this Doctoral Thesis to label the CA. Among them, the Skew and VarSc metrics were only applied to the detection of this artefact, while the kurtosis and entropic measurements (KrE , H_{Sh} , $H_{Ré}$ and $ApEn$) were tested on the detection and removal of all types of artefacts due to the suggestions made by previous studies (Barbati et al., 2004; Dammers et al., 2008; Mammone and Morabito, 2008; Mantini et al., 2008). Despite the fact that Skew and VarSc only aimed at marking the CA, they identified more components as possible artefacts than the non-specific metrics.

On the other hand, cBSS can extract artefactual activity if a suitable reference is available (James and Gibson, 2003). In order to obtain this reference, the average of all MEG channels was computed. This approach

could be taken thanks to the fact that the CA was the only synchronous activity in the MEG background activity.

The OA removal with BSS has recently received considerable attention in both EEG (Castellanos and Makarov, 2006; Flexer et al., 2005; Frank and Frishkoff, 2007; Hoffmann and Falkenstein, 2008; Iriarte et al., 2003; Joyce et al., 2004; Jung et al., 2000; Krishnaveni et al., 2005; LeVan et al., 2006; Li et al., 2006; Nazarpour et al., 2008; Romero et al., 2008, 2009; Shao et al., 2008; Shoker et al., 2005a,b; Ting et al., 2006) and MEG (Barbati et al., 2004; Dammers et al., 2008; Escudero et al., 2007b; James and Gibson, 2003; Okada et al., 2007; Rong and Contreras-Vidal, 2006). While the goal of the first seminal studies was just to visually identify the ocular components (Jung et al., 2000; Vigário, 1997; Vigário and Oja, 2000), several pieces of research have applied metrics to detect the OA. These were based on the topographic information of the BSS (Escudero et al., 2007b; Li et al., 2006; Okada et al., 2007), the correlation with reference signals (Barbati et al., 2004; Flexer et al., 2005; Joyce et al., 2004), thresholding (Castellanos and Makarov, 2006; Ting et al., 2006) and higher-order statistics (Dammers et al., 2008; Shoker et al., 2005a). In other cases, more complex approaches involving a combination of metrics were set up (Frank and Frishkoff, 2007; LeVan et al., 2006; Rong and Contreras-Vidal, 2006; Shao et al., 2008). Nevertheless, the diverse types of signals and evaluation methodologies make the direct comparison of the results difficult.

In the case of the OA, the specific metrics P_{EYES} and P_{LF} , which used topographical and spectral information, in that order (Escudero et al., 2007b), and the general artefact detection metrics (KrE, H_{Sh} , $H_{R\acute{e}}$ and ApEn) were tested. P_{EYES} and P_{LF} identified less components as artefacts, although it must be noted that the latter were supposed to pinpoint all types of contamination.

The MEG background activity was acquired while the subjects had their eyes closed. Hence, the number of blinks and fast eye movements was reduced but the α power could have increased as a result. Moreover, the lack of a visual reference might increase the low-frequency eye movements (Croft and Barry, 2000). Despite the fact that this might induce artefacts in the

recordings, the metric P_{LF} , which detects low frequencies, might deal with this situation. Nevertheless, the application of this metric to open-eyes settings may not be straightforward. Additionally, the brain activity might be characterised by a $1/f$ spectrum and be localised over the anterior region of the head, which may affect the reliability of the P_{LF} and P_{EYES} metrics (Escudero et al., 2007b).

The reduction of the PLN power is typically done by spectral filtering. Nonetheless, if the spectral band of interest overlaps with the power line frequency, another approach is needed to avoid the removal of brain activity (Jung et al., 2000). The detection of the PLN components was done by taking into account the power around the line frequency (50 Hz) (Escudero et al., 2007b). A similar criterion was also introduced by LeVan et al. (2006) and Nicolaou and Nasuto (2007), although the artefact removal of the PLN was not individually measured.

Most of the artefact detection criteria were applied following the “trials” approach introduced by Barbati et al. (2004); Dammers et al. (2008); Greco et al. (2006) and Mammone and Morabito (2008). If a certain fraction of the trials (or segments) of a BSS component exceeds a threshold value, that component is considered an artefact (Barbati et al., 2004; Dammers et al., 2008; Greco et al., 2006; Mammone and Morabito, 2008). Usually, a component is rejected if 20% or more of its trials exceeded the threshold of ± 1.64 (Barbati et al., 2004). However, in order to minimise the amount of brain components removed by this procedure, we only marked a BSS source as an artefact when 30% or more of its segments exceeded a threshold value of ± 2.0 (Dammers et al., 2008). This criterion may decrease the amount of artefactual activity removed by the procedure but *it safeguards the brain activity* (Dammers et al., 2008). Moreover, this type of criteria seems to be appropriate to label cases for very unbalanced data. That is, the number of non-artefactual components tends to be much larger than that of artefacts. In this situation, traditional classification techniques may not be the best option to distinguish both kinds of activity (Shao et al., 2008).

It can be seen that the artefact detection metrics aiming at labelling all artefacts (KrE, H_{Sh} , $H_{R\acute{e}}$ and ApEn) seemed to have marked less components

for removal than the combination of the *metrics designed to recognise only a specific kind of contamination* (either CA, OA or PLN). This suggests that the former, by trying to recognise different artefacts, might lose some sensitivity to the undesired sources.

One problem associated with the use of BSS in the artefact removal of EEG and MEG signals is that the actual underlying components are unknown (Escudero et al., 2007b). Therefore, it is impossible to perform a completely reliable and accurate evaluation of the performance. (Barbati et al., 2004; James and Hesse, 2005). Yet, it is indeed possible to compare the brain signals *before and after* the artefact removal to gain some idea about the degree to which the artefacts have been rejected (Dammers et al., 2008; Escudero et al., 2007b; Mantini et al., 2008). In this Doctoral Thesis, this approach has been taken in order to provide objective and quantitative data about the outcome of the artefact removal.

The evaluation of the CA removal relied on the fact that the cardiac interference was the only synchronous activity in the recordings (Escudero et al., 2007b). On the other hand, the OAs do not always appear in the MEG data. Hence, their reduction is more difficult to assess. In order to do so, the temporal waveforms of the signals were inspected to count the number of peaks above a particular threshold (Mantini et al., 2008) and the low frequency range of the spectrum was considered (Escudero et al., 2007b). Finally, the reduction in the PLN was easy to quantify in the spectral domain (Escudero et al., 2007b; Iriarte et al., 2003). Thanks to these criteria, we were able to gather quantitative data about the artefact removal. On the basis of these results, Table 8.2 details some of the combinations of BSS technique, epoch length and artefact detection metric that offered the greatest artefact reduction. In each case, the data in this Table are expressed as the ratio of the values of the corresponding metric (for instance, A_{PLP}) after and before the artefact removal. It is important to note that the list of combinations detailed in Table 8.2 is not exhaustive as slight modifications, such as considering shorter or longer epoch lengths, may offer similar levels of artefact reduction. However, this Table does convey an idea of the *type of artefact reduction methodologies that are best suited* for CA, OA and PLN.

Table 8.2: Summary of the changes (expressed as ratios) in the metrics to assess the artefact removal. Lower ratios suggest better artefact removal.

Ratios of the metrics evaluating the cardiac artefact				
Metric	Length	BSS method	A_{PtP}	RMS
–	10 s	cBSS	0.1335	0.1963
Skew	60 s	eInfoMax	0.1547	0.2359
VarSc	10 s	FastICA	0.1572	0.2513
KrE	90 s	eInfoMax	0.1354	0.2258
H_{Sh}	60 s	eInfoMax	0.1238	0.2132
$H_{R\acute{e}}$	60 s	eInfoMax	0.1242	0.2131

Ratios of the metrics evaluating the ocular artefact				
Metric	Length	BSS method	Number of peaks	Spectral discrepancy
P_{EYES}	60 s	SOBI	0.6158	0.4137
P_{LF}	90 s	AMUSE	0.1589	0.4357
KrE	90 s	eInfoMax	0.4221	0.8165
$H_{R\acute{e}}$	20 s	SOBI	0.2833	0.5005
ApEn	20 s	SOBI	0.2848	0.4011

Ratios of the power from 49.5 Hz to 50.5 Hz to evaluate the power line noise				
Metric	Length	BSS method	Total power	Normalised power
P_{50Hz}	60 s	AMUSE	0.2039	0.1850
P_{50Hz}	60 s	JADE	0.2406	0.2297
ApEn	60 s	JADE	0.2551	0.3178

It can be seen that the CA is considerably reduced with a cBSS procedure or by decomposing the signals with HOS-BSS techniques. This may be due to the fact that this artefact is often characterised by having highly asymmetrical and supergaussian amplitude distributions (Escudero et al., 2007b). On the other hand, the cBSS completely avoids the need of identifying the components of interest as it only extracts the component whose activity is closer to the reference (James and Gibson, 2003). As for the OAs, the best artefact removal was usually achieved with SOS-BSS techniques combined with specific artefact detection criteria. This supports the idea suggested by Kierkels et al. (2006) and Romero et al. (2008, 2009) that this kind of artefact may be better reduced by means of SOS-BSS techniques, such as AMUSE and SOBI,

than HOS-BSS. Finally, the impact of the PLN was reduced thanks to the metric $P_{50\text{Hz}}$ which aims at marking components with high power around the line frequency (Escudero et al., 2007b). When observing the artefact removal results contained in Table 8.2, it should be noticed that the real reduction in power between 49 Hz and 51 Hz is given by the total power measure. This is due to the fact that the normalised power depends on the amount of power in the whole frequency range, which could have been affected by the removal of the CA or OA. Moreover, the actual reduction in peak at 50 Hz is higher than the ratios indicated in Table 8.2 since these metrics account for the change in power in a 2 Hz band (49 Hz to 51 Hz).

For the sake of a straightforward comparison between diverse BSS-based artefact removal studies in MEG data, Table 8.3 displays the main results regarding artefactual attenuation derived from those studies and this Doctoral Thesis.

All those data support the idea that BSS is very useful to reduce the amount of artefactual contamination in brain recordings, thus helping to properly interpret the signals in further analysis (Escudero et al., 2007b; James and Hesse, 2005; Vigário and Oja, 2008). However, care should be taken as a ‘leakage’ of brain activity into the artefactual components would mean that useful information was lost (LeVan et al., 2006). Moreover, some authors have even suggested that this kind of procedures might, in some cases, produce distortions in the signals (Castellanos and Makarov, 2006; Wallstrom et al., 2004). Nonetheless, BSS artefact removal avoids the problems of other methodologies, such as the requirement of recording reference channels and the bidirectional contamination of the regression techniques and the very limited specificity of spectral filtering (Crespo-Garcia et al., 2008; Iriarte et al., 2003; LeVan et al., 2006; Romero et al., 2008).

8.5 Feature Enhancement

In this Section, the application of BSS techniques as a preprocessing to enhance signal features of MEG recordings in AD is discussed. The corresponding results were described in Section 7.4. Due to the nature of the spectral

Table 8.3: Summary of the attenuation of artefactual activity in MEG recordings with BSS-based procedures.

Study	Database	Artefact	Evaluation
Sander et al. (2002)	1200 epochs from 1 subject.	CA.	Study of the topography of the MEG activity after the CA removal in a visual task.
James and Gibson (2003)	3 epochs from 3 subjects.	CA, OA and PLN.	Illustration of the temporal evolution and scalp maps of the MEG activity before and after the cBSS artefact removal.
Barbati et al. (2004)	3 epochs from 3 subjects.	CA, OA and PLN.	Spectral and temporal examples of MEG signals before and after the BSS artefact removal.
Rong and Contreras-Vidal (2006)	Up to 123 trials from 5 subjects.	Muscular, and OA.	CA Only data about the performance of the detection of artefactual components are given.
Okada et al. (2007)	2100 trials from 7 subjects.	OA.	Temporal examples of MEG signals before and after the BSS artefact removal
Mantini et al. (2008)	12 subjects.	Muscular, and OA.	CA Visual examples and quantitative data suggesting a clear improvement in the number of signals available for analysis and a ratio of the AP_{tP} before and after the procedure of 0.158, approximately ^a .
Dammers et al. (2008)	10 experiments from 6 subjects.	Muscular, and OA.	CA Visual examples and rejection performances in terms of RMS of about 80% to 90% around the R-peaks and blinks located in reference channels.
This Doctoral Thesis	26 subjects.	CA, OA and PLN.	The value of metric AP_{tP} decreased to about 0.13 of its original value and the peak at 50 Hz was attenuated around 14.7 dB. Suggestion that the importance of the OA decreased by, at least, an order of 2.

^aValue computed averaging the results given in the article.

and non-linear signal features computed from the MEGs, signal epochs of 10 s were analysed.

The BSS algorithms AMUSE, SOBI, eInfoMax and FastICA were applied to decompose MEG epochs with minimal artefactual activity recorded from 36 AD patients and 26 healthy elderly control subjects. It is important to note that the population was randomly divided into training and test sets to avoid the optimization of the parameters involved in the methodology (ranges of components and classification rules) on the whole dataset (Escudero et al., 2009b). Whereas AMUSE provided an inherent order for the extracted sources, the SOBI, eInfoMax and FastICA components had to be ordered (Escudero et al., 2009c). Two different ranking criteria were used: the MF and the kurtosis of the components. Then, every component was characterised with two spectral (MF and SpecEn) and two non-linear (LZC and SampEn) signal features. Of note is that the use of MF is *twofold*. On the one hand, it serves as an ordering criterion so that components extracted from different epochs and subjects can be compared. On the other, it is a signal characterising technique helpful to distinguish the subject groups (Escudero et al., 2009c).

For each case, a one-way ANOVA with age as a covariate was used to decide which components had the *most significant differences* between AD patients and controls in the training set (see Table 7.48). These ranges of components were used to compute the partial reconstructions of the MEG signals: $\mathbf{x}_{\text{partial}}(t)$. An LDA provided the classification rules for each case in the training set. Then, these rules were applied to the test set without further modification. In this way, it was compared the separation between groups achieved using the original MEG recordings with that of the partially reconstructed MEGs preprocessed with the combinations of BSS technique and ordering criterion. Overall, the results suggested that the BSS and component selection procedure improved the separation between the AD patients and control groups since this preprocessing increased both the accuracy and AUC.

BSS estimates the set of components that originated the recorded brain activity blindly (James and Hesse, 2005; Vigário and Oja, 2008). The BSS

components are not necessarily equivalent generator dipoles, but mutually independent and simultaneous electromagnetic activity measured over the scalp (Jin et al., 2002; Vorobyov and Cichocki, 2002). As it has been mentioned, one of the most common applications of BSS techniques to EEG and MEG is artefact removal (Escudero et al., 2007b; James and Hesse, 2005; Joyce et al., 2004; Melissant et al., 2005; Sander et al., 2002; Ting et al., 2006; Vorobyov and Cichocki, 2002). This is based on the assumption that the BSS methods will isolate the artefactual activity into a few BSS sources. Once these artefacts have been identified and detached, the brain components are projected back to the channels to obtain the clean recordings (Escudero et al., 2007b; Joyce et al., 2004).

On the other hand, the approach taken in this part of the Doctoral Thesis is different. The aim was not to remove artefacts from MEG recordings with minimal brain activity distortion (Escudero et al., 2008d, 2009b). Instead, we attempted to emphasise the differences between AD patients and control subjects' MEG recordings by retaining only the components which accounted for the most relevant differences between groups in each case. Thus, the partially reconstructed MEG signals *do not reflect the brain activity accurately*, but they should have more different values of MF, SpecEn, LZC and SampEn than the MEG data without any BSS preprocessing (Escudero et al., 2008d, 2009b).

It is worth noting that, previously to the BSS, the MEG epochs with the clearest artefacts were rejected. This was done in order to avoid surplus complexity and to assess the classification improvement without any other kind of previous procedure. Obviously, the same signal epochs were used to compute the MF, SpecEn, LZC and SampEn with and without the BSS and component selection procedure. Therefore, the diverse classification results could be directly compared to evaluate the relative classification improvement (Escudero et al., 2008d, 2009b).

There is a wide range of BSS algorithms available to decompose EEG or MEG data (James and Hesse, 2005). Among them, AMUSE (Tong et al., 1991), SOBI (Belouchrani et al., 1997), eInfoMax (Lee et al., 1999) and FastICA (Hyvärinen and Oja, 1997) were tested in this framework. The former

two are SOS techniques whereas the latter two are HOS-based methods. AMUSE was the first BSS technique proposed to be used in this kind of preprocessing (Cichocki et al., 2005). Afterwards, the spatial filtering technique of CSP was also tested to achieve a similar objective (Woon et al., 2007). Recently, this approach has been extended to other BSS techniques (Escudero et al., 2009c).

AMUSE might not separate the components as completely as SOBI (Tang et al., 2005). This is specially the case for short signal epochs since it was shown in Section 7.2 that their separations tend to converge for long signals. On the other hand, the computational complexity of AMUSE is lower than in most BSS algorithms (Cichocki and Amari, 2002; Ting et al., 2006). Furthermore, AMUSE has other advantages: it always offers the same separation when applied to the same input data set and orders the components by decreasing linear predictability (Cichocki et al., 2005). This *inherent order* is one of the key points of this component selection procedure and, probably, one of the reasons why the AMUSE preprocessing usually provided better classifications than the preprocessings based on other BSS techniques. AMUSE allowed to straightforwardly compare the components from different epochs and subjects. On the other hand, due to the fact that AMUSE only uses two time delays to decorrelate the signals, the AMUSE decomposition may be less robust to additive white noise than that computed using other BSS algorithms (Tang et al., 2005; Ting et al., 2006).

The other three BSS techniques tested in this framework required defining an ordering criterion to compare the components computed from different signals. On the basis of previous results (Escudero et al., 2008d, 2009b), the MF was defined as a ranking criterion that tried to mimic the ordering of AMUSE. As it was shown in Section 7.4, the combination of SOBI with the ordering of MF offered the most similar results to AMUSE. Alternatively, eInfoMax and FastICA are HOS techniques. Hence, their decompositions differ substantially from those of SOBI and AMUSE. For the sake of comparison with MF, kurtosis was also used as a ranking criterion. However, this ordering of the components failed to reveal structures in the data as clear as those pointed out by AMUSE or the combination of SOBI with the MF

ordering.

The features used for classification were MF (Poza et al., 2007b), SpecEn (Sleigh et al., 2004), LZC (Lempel and Ziv, 1976) and SampEn (Richman and Moorman, 2000). The former two try to characterise the whole spectrum with a single value, while the latter two are non-linear features. When MF and SpecEn were applied to the analysis of AD patients' EEG and MEG data, they revealed a slowing and irregularity loss in the spectrum of the signals (Abásolo et al., 2006c; Fernández et al., 2006; Jelic et al., 2000; Poza et al., 2007b; Signorino et al., 1995). Additionally, the non-linear analysis methods supported the hypothesis of decreased irregularity and complexity in AD patients' EEGs and MEGs (Abásolo et al., 2006c,e; Gómez et al., 2006a, 2009a; Hornero et al., 2009b). Nevertheless, the origin of these changes is not clear yet, due to the heterogeneity of AD (Jeong, 2004; Selkoe, 2001).

In the best cases, the BSS and component selection preprocessing increased the accuracy between 12.9% and 22.6% for all features apart from MF, for which the accuracy remained unchanged. Nevertheless, the AUCs of all features improved between 0.023 and 0.227 (for MF and SampEn, in that order). A summary of the best preprocessings in terms of improvements for the accuracy and AUC values is detailed in Table 8.4. For the sake of comparison, the classification results obtained from the raw MEG data (without any preprocessing) and from a spectral band-pass filtering are also included in Table 8.4.

It is remarkable that the accuracy and AUC results were computed in different ways. Whereas the accuracy was calculated in the test set with the classification rules developed in the training set, the AUC was estimated using data only from the test set. In contrast to the maximum accuracy value obtained for a variable, the AUC depends on the whole range of sensitivity/specificity pairs provided by that variable, thus offering an idea of how separated the groups are (Fawcett, 2006). Therefore, our results suggest that the best BSS preprocessings provided a more robust separation between groups for both kinds of features: spectral and non-linear ones (Escudero et al., 2009b). Moreover, our analyses showed that similar ranges of components contained the most significant differences for both types of features

Table 8.4: Comparison of the accuracy and AUC values reached with various preprocessing techniques with the values obtained from the MEGs without any preprocessing.

MF			
Type	Components retained	Accuracy	AUC
Without preprocessing	–	77.4%	0.855
α -filtered	–	77.4%	0.872
AMUSE	30	77.4%	0.878
SOBI - MF	30	77.4%	0.876
SOBI - Kurtosis	30	71.0%	0.876
SpecEn			
Type	Components retained	Accuracy	AUC
Without preprocessing	–	61.3%	0.727
β -filtered	–	58.1%	0.560
SOBI - MF	30	74.2%	0.782
AMUSE	30	71.0%	0.786
eInfoMax - Kurtosis	30	67.7%	0.786
LZC			
Type	Components retained	Accuracy	AUC
Without preprocessing	–	61.3%	0.786
β -filtered	–	64.5%	0.705
AMUSE	15	74.2%	0.838
AMUSE	30	67.7%	0.838
SOBI - Kurtosis	30	64.5%	0.812
SampEn			
Type	Components retained	Accuracy	AUC
Without preprocessing	–	58.1%	0.645
β -filtered	–	61.3%	0.645
AMUSE	30	80.7%	0.872
AMUSE	15	80.7%	0.863
SOBI - MF	30	67.7%	0.803

when the preprocessing employed AMUSE or SOBI with the MF ordering. Actually, these were the most consistent and accurate preprocessings. Furthermore, a spectral band-pass filtering preprocessing could not match the classification improvements provided in those cases. Therefore, it could be hypothesised that those BSS and component selection procedures could also be applied with other spectral and non-linear features and that the ranges of components with the largest differences between AD patients and healthy controls for other analysis techniques might be similar to those reported in this study when utilising AMUSE or SOBI-MF.

Previous research by [Cichocki et al. \(2005\)](#) and [Vialatte et al. \(2005\)](#) had shown the utility of the AMUSE and component selection procedure when spectral and WT features were computed from EEGs of MCI patients who later proceeded to AD. These EEG signals were characterised with the relative powers in six frequency bands ([Cichocki et al., 2005](#)). Afterwards, an LDA was applied to classify the subjects. The accuracy improvement due to this methodology was 10% ([Cichocki et al., 2005](#)). Nevertheless, the individual improvement in each variable was not measured by [Cichocki et al. \(2005\)](#). A later study used a “bump modelling” of the partially reconstructed EEG wavelet time-scale transform and a neural network classifier to further improve the subject classification ([Vialatte et al., 2005](#)).

In contrast to those studies, the classification method of this Doctoral Thesis allowed to assess the improvement in each variable (MF, SpecEn, LZC and SampEn) separately. It was also found that the BSS and component selection procedure is useful when the MEG signals are analysed with non-linear methods. Similarly, a study with the same EEGs analysed by [Cichocki et al. \(2005\)](#) and [Vialatte et al. \(2005\)](#) found that the ability of both spectral and non-linear features to distinguish the subject groups improved with the application of a different kind of spatial filters: CSP ([Woon et al., 2007](#)). CSP finds spatial filters which maximise the difference in signal power between two classes to be discriminated (e.g., patients and controls) ([Woon et al., 2007](#)). To facilitate the comparison of the results, Table 8.5 summarises the main characteristics of those studies.

Since diverse spatial filtering techniques have proven to be useful in this

Table 8.5: Summary of the results reported by studies on the improvement of subject classification by means of a spatial filtering.

Study	Subjects	Signal	Approach
Cichocki et al. (2005)	22 MCI patients who later proceeded to AD and 38 controls.	EEG	A selection of 5 AMUSE components improved the classification based on 6 spectral band powers from 70% to 80%.
Vialatte et al. (2005)	22 MCI patients who later proceeded to AD and 38 controls.	EEG	A selection of 5 AMUSE components and a classification with WT features and a neural network provided a further improvement of 13.3% in comparison with (Cichocki et al., 2005).
Woon et al. (2007)	22 MCI patients who later proceeded to AD and 38 controls.	EEG	In comparison with features computed from individual channels, the CSP improved the accuracy around 10%. The best accuracy of 88.3% was achieved with CSP and a spectral power measure.
This Doctoral Thesis	36 AD patients and 26 controls.	MEG	Evaluation of AMUSE, SOBI, eInfoMax and FastICA in this setting. The accuracy of the MF remained unchanged whilst it increased from 61.3% to 74.2% for SpecEn and LZC and from 58.1% to 80.7% for SampEn. The changes in the AUC ranged between 0.023 for the MF and 0.227 for the SampEn.

type of application, it is suggested that their common characteristics should be considered to design a preprocessing to improve the classification of AD patients' EEG or MEG signals.

8.6 Adaptive Extraction of Brain Rhythms

This Section discusses the main implications derived from the adaptive regional extraction of brain rhythms and subsequent evaluation of their connectivity carried out in Section 7.5. Similarly to the BSS preprocessing to enhance signal features of AD, the length of the signal epochs analysed in this application was 10 s.

Of note is that the CA that appeared in the signal epochs was reduced by means of a cBSS (James and Gibson, 2003). This procedure was necessary to ensure that the QRS complexes did not contaminate the IMFs that will serve as references to extract the rhythmical signals from each region. This was due to the nature of the EMD. This methodology is able to adaptively and locally decompose a signal in its rhythmic components (Huang et al., 1998). Hence, the cardiac activity may be present in some of the IMFs (Blanco-Velasco et al., 2008; Tang et al., 2008). If the CA removal was not applied, there would be chances of ending up computing the $\text{Coh}(f)$ between signals with large cardiac contamination, which might bias the results. Moreover, the same cardiac activity was estimated and removed by this procedure at all MEG channels, thus avoiding differences in the way different channels were preprocessed.

On the other hand, the OAs did not pose serious problems onto the analysis since the signals with clear artefacts were not selected for analysis. Additionally, the PLN that may contaminate the MEG recordings was contained in the first IMF computed by the EMD. This IMF had almost always a $\langle\omega\rangle$ value larger than 30 Hz. Thus, it was not fed into the subsequent clustering procedure.

The analyses carried out in this part of the Doctoral Thesis focused on specific spectral bands: δ (1 Hz – 4 Hz), θ (4 Hz – 8 Hz), α (8 Hz – 13 Hz) and β (13 Hz – 30 Hz). However, no spectral filtering had to be performed, thus

avoiding the definition of cut-off frequencies *a priori*. This was due to the fact that the IMFs have well defined spectral characteristics (Huang et al., 1998). Therefore, one can take into account the spectral content of every IMF, which was characterised with the $\langle\omega\rangle$. Then, only the IMFs whose average spectral content fell within the broadband of interest (1 Hz – 30 Hz) were considered for selection of the suitable reference for every band and scalp region.

Among all IMFs computed from the channels within a region, suitable references were selected for each spectral band. This procedure was carried out with a k -means clustering of the corresponding $\langle\omega\rangle$ values. This is an unsupervised, effective and simple algorithm that divides a set of features into k clusters (Hartigan and Wong, 1979; Sanei and Chambers, 2007). This methodology is appealing due to its adaptive nature. It grouped the IMFs with similar $\langle\omega\rangle$ together and allowed characterising each cluster by its centroid value. It must be noticed that this process marked a single IMF as a reference to extract the rhythmical activity of interest in every region. This estimation of the reference could not have been based, for example, on calculating an average of several IMFs since slightly desynchronisations in their oscillations would have altered their rhythmical nature.

Once the references for each spectral band and region had been computed, they were used with a cBSS to extract activity representative of *all* channels in a region for a particular band. This was necessary because the particular IMF selected as reference had been computed from one particular MEG channel. By means of the cBSS, the activity extracted to be analysed with the $\text{Coh}(f)$ was representative of the *whole region*. This approach was motivated by the fact that attention should be paid to the ‘domain’ where the brain activity is analysed. The most simple approach may be to analyse the signals at every particular channel — ‘channel domain’ — (Dauwels et al., 2010; Gómez et al., 2009c; Rossini et al., 2007). Alternatively, the equivalent current dipoles can be localised and analysed — ‘dipole domain’ — (Rossini et al., 2007; Stam, 2010).

As it was mentioned, the approach adopted in this study was different. The rhythmic brain activity was simultaneously extracted from all channels

in a region. This approach was relatively similar to that taken by Stam et al. (2006). It consisted of a channel-wise procedure by which the analysis was carried out channel by channel, and the the results from different sets of channels were grouped together (Stam et al., 2006). Instead, in this Doctoral Thesis, a *cBSS* was applied to extract the activity that characterises a particular scalp region for a particular spectral band ('regional domain'). This methodology may be beneficial since common sources might influence the results computed at nearby MEG channels (Stam et al., 2006). Hence, the connectivity calculated between close channels might have been produced as a result of this phenomenon rather than having been originated by true interactions (Stam, 2010; Stam et al., 2006). When analysing the signals in the 'dipole domain', there is the problem of localising the unknown sources. Furthermore, the solution to this problem is not unique and the source reconstruction algorithm could influence the interdependencies between the sources (Stam, 2010; Stam et al., 2006).

The connectivity analysis was carried out in the 'regional domain' with the $\text{Coh}(f)$ (Nunez et al., 1997, 1999). The rationale beneath this connectivity analysis is the following.

Several authors have proven that AD produces a slowdown of the electromagnetic brain activity. This dementia increases of power in the low-frequency range (δ and θ) and decreases the power in the high-frequency range (α and β) (Jeong, 2004; Rossini et al., 2007; Stam, 2010). Abnormalities in the non-linear dynamics of these recordings also appear (Hornero et al., 2009b; Stam, 2005, 2010). Nevertheless, AD could also be characterised by a neo-cortical disconnection (Jeong, 2004; Stam, 2005). However, these effects are not always easily detectable as there usually is large variability among AD patients (Dauwels et al., 2010). There is some consensus that AD is related to a lower synchronization level in α and β bands. Nevertheless, contradictory results have been found in the other frequency bands and it remains unclear whether δ and θ connectivities are decreased or increased in AD (Jelic et al., 2000; Jeong, 2004; Stam et al., 2006).

Hence, it may be interesting to study the connectivity between different areas rather than local dysfunctions (Stam et al., 2006). Large scale func-

tional integration might be essential for cognitive and behavioural functioning. This large scale functional integration may be due to the synchronization (i.e., temporal correlation) between neural activity in different brain regions, which can be assessed by analysing the connectivity of MEGs in diverse spectral bands and regions (Pereda et al., 2005; Stam et al., 2006).

$\text{Coh}(f)$ is the most widely used method to assess connectivity. It is a normalized linear measure of the correlation between two signals as a function of frequency (Nunez et al., 1997, 1999; Pereda et al., 2005). $\text{Coh}(f)$ only captures the linear interactions between pairs of signals. Thus, it might fail to detect non-linear interdependencies in the data (Gómez et al., 2009c; Pereda et al., 2005). Nevertheless, this method has already revealed changes in the connectivity of AD patients' brain recordings (Jeong, 2004). Moreover, it is not clear that alternative non-linear connectivity measures are superior to $\text{Coh}(f)$ to distinguish AD or MCI and the outcomes of $\text{Coh}(f)$ and other connectivity measurements seem to be correlated (Dauwels et al., 2010; Gómez et al., 2009c).

The statistical analysis of the $\text{Coh}(f)$ results revealed strong dependences on the spectral band and pair of regions. The estimated marginal means (Searle et al., 1980) computed for AD patients and control subjects in each spectral band supported the idea that *AD decreases the connectivity of brain signals in α and β bands* (Jeong, 2004). Moreover, the analyses seem to indicate that this dementia causes a *rise in δ coherence*. However, there were no significant differences between subject groups. Additionally, the factor 'Pair of regions' showed significant interactions in all cases. The pattern was similar for all bands. It suggested the logical idea that more distant pairs of regions tend to be characterised with lower $\text{Coh}(f)$ values.

Despite the fact that the differences in the overall level of $\text{Coh}(f)$ for AD patients and controls on every band were not significant, Fig. 7.36 indicated a certain tendency that relates alterations in connectivity to AD. The fact that not all frequencies are not equally affected by the changes in coherence suggests that these alterations are not simply due to a loss of cortical neurons (Dauwels et al., 2010). Therefore, the alterations in connectivity might be due to other mechanisms, such as anatomical disconnections

among different cortical regions or reduced cholinergic coupling between cortical neurons (Dauwels et al., 2010; de LaCoste and White, 1993; Jeong et al., 2001).

In addition to the statistical analysis, a classification study based on the $\text{Coh}(f)$ results was also performed. This employed a stepwise LDA. This technique selected the most appropriate variables to be included into the classification rule automatically. This classification used the variables from each spectral band separately and the combination of all $\text{Coh}(f)$ values. In the cases when no variable met the criterion to enter into the classification rule, the LDA was trained with all the possible variables simultaneously. However, the corresponding accuracy values were clearly lower than in the other cases.

It is important to note that a leave-one-out cross-validation procedure was used. Although this procedure typically reduces the sensitivity, specificity and accuracy, it avoids over-estimation of the true classification rates and increases the reproducibility of the results (Cichocki et al., 2005; Jobson, 1991a,b). The number of variables selected by the stepwise LDA was small in comparison with the total number of features available for classification. For instance, in the case of considering variables from all bands, only 11 out of 128 possible variables entered the classifier.

A recent review by Stam (2010) pointed out that the best classification rates obtained by different methodologies for MEGs of AD patients versus healthy elderly subjects ranged between 80% to 90%. However, cross-validation techniques were not always applied in those studies. In this study, the accuracy computed from the $\text{Coh}(f)$ values of a specific spectral band was always lower than 70%. Nonetheless, when variables of all bands were introduced into the stepwise LDA, *the accuracy rose to 96.8%*, as it is shown by Table 8.6. It must be borne in mind that this result was computed with a leave-one-out cross-validation procedure and that variables from all bands were included into the classifier.

Finally, it must be noticed that, similarly to (Dauwels et al., 2010), this study mainly tried to help in the discrimination of both subject groups. We did not aim at identifying the biophysical mechanisms that cause the

Table 8.6: Summary of the classification results obtained from the adaptively extracted brain rhythms.

	δ	α	Combination of δ , θ , α and β
Accuracy (%)	67.7	69.4	96.8
AUC	0.713	0.815	1.000

alterations of AD in the MEGs. Additionally, the subjects did not perform any specific task. That is, only MEG background activity was analysed. Hence, the performance of the classification might be improved by analysing signals acquired during specific tasks (Dauwels et al., 2010; Stam, 2010).

8.7 Limitations of the Study

After discussing the main implications of the results, this Section reviews the main limitations of the studies carried out in this Doctoral Thesis. These are detailed in the following lines:

- The significance of results detailed in this Doctoral Thesis is limited by the relatively reduced size of the database. This consisted of MEG background activity recorded from 36 AD patients and 26 control subjects. This size is relatively similar to, or even larger than, that of other studies about the application of BSS techniques to EEG (Castellanos and Makarov, 2006; Crespo-Garcia et al., 2008; Flexer et al., 2005; Halder et al., 2007; Hyvärinen et al., 2010; Iriarte et al., 2003; Krishnaveni et al., 2005; Mammone and Morabito, 2008; Melissant et al., 2005; Shao et al., 2008; Ting et al., 2006; Wallstrom et al., 2004) and MEG data (Barbati et al., 2004; Cao et al., 2003; Dammers et al., 2008; Escudero et al., 2007b, 2009d; Hesse, 2007, 2008; Ikeda and Toyama, 2000; James and Gibson, 2003; Mantini et al., 2008; Okada et al., 2007; Rong and Contreras-Vidal, 2006; Sander et al., 2002; Tang et al., 2005; Vigário and Oja, 2008; Zavala Fernández et al., 2006), including AD-related studies (Abatzoglou et al., 2007; Berendse et al., 2000; Escudero et al., 2008d, 2009b,c; Fernández et al., 2006; Gómez et al., 2006a,

2007a, 2009a,b,c; Hornero et al., 2008, 2009b; Osipova et al., 2003, 2005, 2006a,b; Pekkonen et al., 1999, 2005; Poza et al., 2007b, 2008b; Stam et al., 2006; van Cappellen van Walsum et al., 2003). However, it would be advisable to increase the dataset by adding new signals in order to further test the main results of this Doctoral Thesis.

- By the own nature of the analyses, the actual sources that originated the brain recordings are unknown. Hence, assessing the performance of the BSS analysis is not straightforward at all because the separation cannot be absolutely validated (James and Hesse, 2005; Vigário and Oja, 2008). However, in most cases, some kind of *ad hoc* criteria can be established to gain an idea of the usefulness or quality of the BSS. For instance, the BSS-processed signals can be compared with the raw recordings to estimate how much artefactual activity was removed (Dammers et al., 2008; Escudero et al., 2007b; Mantini et al., 2008) or whether the BSS improved the separation between subject groups (Escudero et al., 2008d, 2009b). On the other hand, the quality of the separation can only be evaluated by means of synthetic data. This is the reason why the analyses carried out in Section 7.2 were limited to comparing the outcomes of BSS techniques rather than assessing their quality.
- When synthetic data are used to study a particular algorithm, such as in Section 7.1.1, the results are likely to be limited by the quality of the simulated signals. Diverse types of synthetic data have been proposed and some of them are even based on actual signals recorded under specific conditions (see, for example, Barbati et al., 2004; Delorme et al., 2007; Escudero et al., 2007b; Fitzgibbon et al., 2007; Kierkels et al., 2006; Romero et al., 2009). In this sense, the efforts done by Kierkels et al. (2006) or Romero et al. (2009) to mimic real EEG signals using complex models of their generation are remarkable.
- Another limitation of this Doctoral Thesis has to do with the fact that only MEG background activity was analysed. Thus, some of the con-

clusions about the artefact removal procedures or the classification of AD patients using this kind of recordings may not be straightforwardly extended to the case of event-related brain activity. Furthermore, this ‘resting state’ might not be the most effective condition to find AD-related abnormalities in brain data (Stam et al., 2006).

- In this Doctoral Thesis, significant efforts have been devoted to dealing with MEG recordings acquired from AD patients. However, AD is not the only mental disorder known to produce alterations in the EEG and MEG activity (Jeong, 2004; Stam, 2005, 2010). In this sense, it would be particularly important to study MCI (Cichocki et al., 2005; Fernández et al., 2006; Gómez et al., 2009c; Jelic et al., 2000) or other types of dementia (Signorino et al., 1995; Stam, 2005), which have similar effects on the brain signals to those of AD.

To sum up, the findings of this Doctoral Thesis suggest the usefulness of BSS methods in the analysis of MEG signals, specially in patients with AD. Nevertheless, further investigations are needed to confirm our results.

Chapter 9

Conclusions and Future Research Lines

Contents

9.1	Introduction	282
9.2	Conclusions of the Doctoral Thesis	284
9.3	Original Contributions to the State of the Art .	287
9.4	Future Research Lines	289

This final Chapter draws the main conclusions derived from this Doctoral Thesis. In addition, the main contributions of this Doctoral Thesis to the state of the art are highlighted. Finally, possible future research lines are suggested.

9.1 Introduction

The MEG signals record the magnetic fields originated by the neural currents in the brain. They have some relevant advantages over other neurophysiologic techniques, including their high temporal and spatial resolution or the fact that they do not depend on any reference point (Hämäläinen et al., 1993; Hari, 2004). Nevertheless, their analysis is not always straightforward and suitable tools and methodologies to help in their study and interpretation could be very useful. Moreover, the MEG could offer important information about diverse brain states and disorders, including the most common dementia in the Western World: AD (Hornero et al., 2009b; Stam, 2005, 2010).

It is well-known that AD causes a progressive memory loss. Eventually, it also affects all cognitive functions of the patient (Blennow et al., 2006; Cummings, 2004; Nestor et al., 2004). Due to its high prevalence and the burden that it poses on caregivers, AD causes huge social and economic problems in Western societies (Ferri et al., 2006). The cost of the diagnostic testing for AD is expected to be lower than that associated with its treatment and management (Lahiri et al., 2002). Considering this idea and that a definite AD diagnosis can only be made by necropsy (Blennow et al., 2006), the importance of developing tools to help in the diagnosis and characterisation of this disorder becomes clear. Eventually, this may lead to reduce the AD-related costs (Lahiri et al., 2002).

The term BSS denotes a set of techniques useful to process, among others, biomedical signals (James and Hesse, 2005; Vigário and Oja, 2008). These data-driven decompositions of multi-dimensional recordings reveal the constituent inner components that generated the observed recordings through an unknown mixing process (Cichocki and Amari, 2002; Hyvärinen et al.,

2001). Their use in the analysis of biomedical data and, more particularly, brain signals has been increasing notably (James and Hesse, 2005; Vigário and Oja, 2008).

Considering these three main concepts (MEG, AD and BSS), all the analyses and studies carried out in this Doctoral Thesis have been based on the following *hypotheses*.

1. That the comparison of the decompositions estimated by diverse BSS techniques for the same real MEGs could offer relevant information about the relationships and similarities between those BSS methods.
2. That the use of BSS and some specific metrics designed to label artefacts in the BSS components could help to automatically remove some of the CA, OA and PLN that contaminate the MEG background activity.
3. That a properly designed BSS preprocessing could help to distinguish signal features (MF, SpecEn, LZC and SampEn) computed from AD patients' MEG background activity from those of control subjects.
4. That a procedure based on an EMD and a cBSS could offer an adaptive regional extraction of brain rhythms in AD patients and controls' MEG background activity. This rhythms could be characterised with connectivity techniques.

In hindsight, our results and discussions do support these hypotheses, which lay behind the methodology developed in this Doctoral Thesis. Therefore, it can be said that *the main objective of this Doctoral Thesis has been fulfilled*. This was no other than **applying BSS techniques to help in the analysis and interpretation of MEG background activity from AD patients and elderly control subjects**. The results suggested that BSS techniques are indeed extremely helpful in the analysis of MEG recordings. Moreover, this utility is not only limited to common methodological frameworks (for instance, the use of BSS to reduce the artefacts in brain signals) but it also includes more innovative ways of studying the brain activity with

BSS techniques such as their use as preprocessing techniques to enhance signal features of the MEGs or as means to perform an adaptive extraction of brain rhythms on diverse scalp regions. However, further research is needed to confirm these results.

9.2 Conclusions of the Doctoral Thesis

The methodology introduced in Chapter 6 has been used to process MEG recordings acquired from 36 patients with a diagnosis of probable AD and 26 healthy elderly subjects in order to carry out *four main studies* about the application of BSS techniques to MEG background activity:

1. The comparison of the decompositions computed by diverse BSS techniques for real MEG data helped to identify similarities between BSS methods and their evolution with the epoch length and the number of components to estimate.
2. The quantitative evaluation of the BSS-based removal of the CA, OA and PLN was useful to objectively pinpoint diverse combinations of BSS technique, epoch length and artefact-detection metric that attenuated the presence of these artefacts in MEG background activity.
3. The development and testing of a BSS preprocessing and component selection procedure showed that the application of this kind of spatial filtering techniques increased the classification accuracy and the AUC values of signal features computed from AD patients and control subjects.
4. The use of an adaptive procedure to extract rhythmical brain activity from diverse scalp regions offered information about the connectivity alterations caused by AD and revealed the possible usefulness of this analysis in the classification of AD patients versus healthy elderly subjects.

All these results are promising in the sense that they suggest the **utility of BSS in the processing of MEG background activity and in the**

identification and characterisation of AD. Therefore, BSS may be an important tool to analyse this kind of biomedical recordings. In addition to this main conclusion, the following particular conclusions can be drawn from this Doctoral Thesis:

1. **The FA-based procedure and the MDL model order selection criterion may be better suited for MEG preprocessing than classical PCA techniques.** The testing of the FA-based procedure and the MDL criterion (Cao et al., 2003; Ikeda and Toyama, 2000) on synthetic data indicated that this approach was the most appropriate to perform a first step in the data processing before BSS. This was due to the fact that the commonly used PCA criteria tended to suffer from severe overfitting (Escudero et al., 2007b). Moreover, the appropriateness of that approach was also supported by the real signal analysis.
2. **The most similar decompositions of real MEG background activity were computed by AMUSE–SOBI followed by JADE–FastICA.** The consistency evaluation of the degree to which diverse BSS techniques computed similar decompositions for real MEGs (Escudero et al., 2009d) showed that the most consistent pair of algorithms was AMUSE–SOBI, followed by JADE–FastICA. Additionally, the pairs AMUSE–JADE and SOBI–JADE used to provide the most dissimilar separations. Finally, the overall level of similarity increased as longer signals were decomposed.
3. **The quantification of the artefactual activity decrease in the MEG background activity highlighted the utility of several combinations of BSS techniques and specific artefact detection metrics.** The objective measurement of the reduction the average intensity of the CA, OA and PLN in the whole dataset pinpointed the following strategies for artefact removal:
 - For its simplicity, the cBSS scheme (James and Gibson, 2003) was considered the most suitable approach to remove the CA.

Moreover, its performance ranked among the best in the reduction of this artefact.

- The OAs reduction was the most difficult to evaluate as this artefact does not always appear in the signals and its localisation in the spectrum is not so evident as in the PLN case. However, the evaluation results suggested that the combination of artefact detection metrics such as ApEn, P_{EYES} and, specially, P_{LF} (Escudero et al., 2007b) with SOS-BSS techniques could be helpful to reduce the OA.
 - The PLN could be substantially reduced without altering the brain signal spectrum by means of the specific metric P_{50Hz} (Escudero et al., 2007b) and AMUSE.
4. **A SOS-BSS preprocessing and component selection procedure could improve the separation between AD patients and control subjects' MEG signal features.** The ordering criteria was essential to straightforwardly compare the BSS components, thus enabling us to develop the BSS preprocessing algorithms (Escudero et al., 2008d, 2009b,c). The assessment of the improvement in the separation between AD patients and controls' spectral and non-linear features of the MEGs revealed rises in the AUC values between 0.023 and 0.227 and accuracy increases of up to 22.6% in comparison with the 'raw' signals. The most appropriate algorithms for this application turned to be AMUSE and SOBI with the MF ordering.
 5. **The adaptive extraction of brain rhythms over diverse scalp regions with an adaptive procedure based on an EMD and a cBSS yielded useful information about the connectivity pattern in AD and, specially, the subject classification.** The adaptive framework introduced to extract rhythmical brain activity from diverse regions was based on an EMD (Huang et al., 1998), a k -means clustering (Hartigan and Wong, 1979) and a cBSS (Huang and Mi, 2007). The statistical analysis suggested that AD affects the connectivity between scalp regions, although the results were not statistically

significant. However, a leave-one-out classification analysis based on the $\text{Coh}(f)$ values classified the AD patients versus the controls with an accuracy of 96.8%.

Finally, it is important to note that one of the main advantages of BSS in these types of applications is that they are data-driven methods which do not need much information about the components to be estimated or the mixing process (James and Hesse, 2005; Vigário and Oja, 2008). By making a few general assumptions about the signals, these techniques are able to estimate a set of underlying components that may provide novel and fruitful perspectives to examine the data (James and Hesse, 2005; Onton et al., 2006; Vigário and Oja, 2008).

To sum up, this Doctoral Thesis has examined a few applications of BSS techniques to MEG recordings, paying special attention to AD. One of the general aims was to provide quantitative evaluation of the advantages and benefits provided by these methods. The results indicate utility of BSS in the processing of MEG signals, specially in AD. Nevertheless, further investigations should be carried out to confirm our conclusions. It must be noted that interesting reviews on the use of BSS to study brain signals have recently been published (James and Hesse, 2005; Onton et al., 2006; Vigário and Oja, 2008). Nevertheless, the possible range of application of BSS is so wide and the discipline evolves so rapidly that it is very difficult to compile most of the information available into one document.

9.3 Original Contributions to the State of the Art

The main contribution of this Doctoral Thesis is the evaluation of the performance, in terms of the previously defined objectives, of diverse BSS techniques in four different applications. In order to do so, different methodologies have been introduced. Furthermore, in some cases, this Doctoral Thesis represents the first application of the corresponding preprocessing, methodologies or evaluation techniques to MEG background activity.

The original contributions of this Doctoral Thesis to the state of the art are detailed in the following lines. Moreover, the *scientific publications* generated as a consequence of this research are listed in Appendix A.

1. The testing of the FA-based preprocessing on synthetic data indicated the suitability of this technique as a prior step to BSS over some PCA approaches. The evaluation of these preprocessing alternatives with synthetic signals and diverse SNR values represents, to the best of our knowledge, the first exhaustive assessment of the FA- and PCA-based preprocessings in this context for the analysis of brain signals (Escudero et al., 2007b).
2. The comparisons of the outcomes computed by diverse BSS algorithms for the same real MEG signals yielded information about the similarities and differences of those techniques. We could not find in the state of the art any other study about the global consistency (or similarity) of the whole BSS decompositions of real brain recordings (Escudero et al., 2009d).
3. The objective evaluation of the reduction in the amount of artefacts that contaminate the MEGs played a key role in the artefact removal application. To the best of our knowledge, this study can be considered as the first evaluation of the CA removal with BSS in terms of the average QRS (Escudero et al., 2007b). Moreover, this study also proposed the evaluation of the artefact removal in MEG based on the totality of the signals in a database by considering them as a whole dataset instead of analysing only a few examples or attending at single cases (Escudero et al., 2007b). As for the artefact detection metric, this study proposed the metrics P_{EYES} , P_{LF} and P_{50Hz} and it applied the Skew to MEG data (Escudero et al., 2006c,d, 2007b, 2008c).
4. This study evaluated the use of BSS techniques as a preprocessing to enhance signal features related to AD. Although this preprocessing had been previously used in EEG data, this Doctoral Thesis represents its

first application to MEG data (Escudero et al., 2008d), including non-linear signal features as well as spectral ones (Escudero et al., 2009b). Moreover, we could not find in the literature any evaluation of other BSS (or ICA) algorithms than AMUSE in this kind of framework (Escudero et al., 2009c).

5. The methodology developed to adaptively extract brain rhythms from diverse regions is novel. Although a division of the MEG channels into different regions had previously been used (see, for example, Stam et al., 2006), this is the first time that the EMD is applied to brain signals with the aim of characterising AD and, in particular, to MEG recordings with the aim of extracting references suitable to be fed into a cBSS whose outcome will be used in a connectivity analysis.

Additionally, it is important to note that the PhD candidate has also worked in the characterisation of brain recordings (EEG and MEG data), paying particular attention to the non-linear analysis of AD patients' signals. This has enabled him to collaborate in diverse pieces of research not directly related to the BSS (Abásolo et al., 2006a,b,d, 2007a,b, 2008a,b,c, 2009; Escudero et al., 2006b, 2007a, 2008a,b; Fernández et al., 2008, 2009; Gómez et al., 2006b, 2007a,b,c, 2008a,b, 2009a; Hornero et al., 2008, 2009a,b; Poza et al., 2007a, 2008a,b, 2009). *Of special relevance are the articles* about the multiscale entropy analysis of EEG data in AD patients (Escudero et al., 2006a) and the interpretation of the 'rate of decrease of the auto-mutual information function' in terms of straightforward signal characteristics (Escudero et al., 2009a).

9.4 Future Research Lines

In this Doctoral Thesis, MEG background activity recordings from 36 AD patients and 26 control subjects have been studied with BSS techniques. Despite the effort made to provide a comprehensive view of the topics covered by this research, there are interesting open issues that have not been stud-

ied yet. Some of these future research lines are related to the limitations acknowledged in Section 8.7.

The analyses of the BSS consistency and the artefact removal have been carried out with the recordings of the 26 control subjects. The signals acquired from 36 patients with probable AD were also included in the studies about the classification improvement due to the BSS preprocessing and the connectivity measurement of extracted brain rhythms. However, the size of the database should be increased in order to draw more reliable and significant conclusions in the future. Moreover, other disorders may cause alterations in the EEG and MEG data similar to those of AD (Stam, 2005, 2010). Hence, it would be helpful to include in the database recordings from patients with other neurodegenerative disorders. In this way, it would be possible to carry out more complex studies aiming at developing strategies for differential diagnosis. Of special importance is the comparison of MEG recordings from demented subjects with those of MCI patients. This is due to the fact that the MCI may be considered as a previous stage to dementia. Actually, a few studies have already analysed MEG activity from MCI patients and control subjects included in the database analysed in this Doctoral Thesis (Fernández et al., 2006; Gómez et al., 2009c).

As for the artefact removal, it should be checked how the BSS-based artefact removal affects the MEG signals acquired from patients with mental disorders like AD. This is due to the fact that the artefact removal was only tested with the control group. It is expected that the same methods can be applied to patients' signals but this hypothesis should be corroborated. In this sense, the pieces of research by Melissant et al. (2005) and Vialatte et al. (2009) showed that the BSS artefact removal does not cause undesired alterations in the brain recordings and that it can improve the classification rate of different subjects in comparison with that based on the noisy data with artefacts. On the other hand, Castellanos and Makarov (2006) suggested that, if the BSS-based artefact removal is done with little care, it can affect the connectivity pattern of EEG data.

Other possible research line is the application of different data-driven signal decomposition techniques. For instance, methods such as the Non-

negative Matrix Factorisation (NMF) could provide us with important information about the brain signals, even in AD (Chen et al., 2006; Cichocki et al., 2008). Alternatively, different types of constraints (e.g., spatial or spectral ones; James and Hesse, 2005) could be used instead of the temporal references needed by the cBSS applied in this Doctoral Thesis.

It is also of interest to apply the techniques included in this Doctoral Thesis to other physiological signals. For example, although the methods were designed for MEG recordings, their application to EEG recordings should be straightforward.

Although the selection of the parameters of each methodology (spectral bands and input values for the BSS algorithms and signal features) tried to be supported by the revision of the state of the art, it is possible that variations in the values of these input parameters may alter the outcomes of the signal processing. Furthermore, other signal features (Hornero et al., 2009b; Stam, 2005) or connectivity measures (Dauwels et al., 2010; Pereda et al., 2005) could also be useful in these settings and should be tested.

All in all, this Doctoral Thesis evaluated some of the most common BSS techniques in the framework of diverse applications, including the comparison of the algorithms' outcomes, the artefact removal, the development of a preprocessing to improve the classification and the adaptive and regional extraction of brain rhythms. The BSS lay at the core of the methodology. The results indicated that the BSS may be an extremely helpful tool to assist in the analysis and study of MEG background activity. Nonetheless, additional research is needed to confirm the results.

Appendix A

Publications

Contents

A.1	Peer-Reviewed International Articles in Journals Indexed in the SCI-JCR	294
A.2	International Book Chapters	297
A.3	Contributions to International Conferences . . .	297
A.4	Contributions to National Conferences	301

This Appendix details the publications in international journals indexed in the SCI–JCR, books and conferences in which the PhD candidate has collaborated.

A.1 Peer-Reviewed International Articles in Journals Indexed in the SCI–JCR

- **Javier Escudero**, Roberto Hornero, Daniel Abásolo and Alberto Fernández.
Blind source separation to enhance spectral and non-linear features of magnetoencephalogram recordings. Application to Alzheimer’s disease. *Medical Engineering and Physics*, 31(7):872–879, September 2009.
- Alberto Fernández, Javier Quintero, Roberto Hornero, Pilar Zuluaga, Marta Navas, Carlos Gómez, **Javier Escudero**, Natalia García-Campos, Joseph Biederman and Tomás Ortiz.
Complexity analysis of spontaneous brain activity in attention-deficit hyperactivity disorder: diagnostic implications. *Biological Psychiatry*, 65(7):571–577, April 2009.
- Carlos Gómez, Roberto Hornero, Daniel Abásolo, Alberto Fernández and **Javier Escudero**.
Analysis of MEG background activity in Alzheimer’s disease using non-linear methods and ANFIS. *Annals of Biomedical Engineering*, 37(3):586–594, March 2009.
- **Javier Escudero**, Roberto Hornero and Daniel Abásolo.
Interpretation of the auto mutual information rate of decrease in the context of biomedical signal analysis. Application to electroencephalogram recordings. *Physiological Measurement*, 30(2):187–199, February 2009.
- Roberto Hornero, Daniel Abásolo, **Javier Escudero** and Carlos Gómez.

Non-linear analysis of EEG and MEG in patients with Alzheimer's disease.

Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 367(1887):317–336, January 2009.

- Daniel Abásolo, **Javier Escudero**, Roberto Hornero, Carlos Gómez and Pedro Espino.

Approximate entropy and auto mutual information analysis of the electroencephalogram in Alzheimer's disease patients.

Medical & Biological Engineering & Computing, 46(10):1019–1028, October 2008.

- Daniel Abásolo, Roberto Hornero, **Javier Escudero** and Pedro Espino.

A study on the possible usefulness of detrended fluctuation analysis of the electroencephalogram background activity in Alzheimer's disease.

IEEE Transactions on Biomedical Engineering, 55(9):2171–2179, September 2008.

- Roberto Hornero, **Javier Escudero**, Alberto Fernández, Jesús Poza and Carlos Gómez.

Spectral and Non-linear Analyses of MEG Background Activity in Patients with Alzheimer's Disease.

IEEE Transactions on Biomedical Engineering, 55(6):1658–1665, June 2008.

- **Javier Escudero**, Roberto Hornero, Jesús Poza, Daniel Abásolo and Alberto Fernández.

Assessment of classification improvement in patients with Alzheimer's disease based on magnetoencephalogram blind source separation.

Artificial Intelligence in Medicine, 43(1):75–85, May 2008.

- Jesús Poza, Roberto Hornero, **Javier Escudero**, Alberto Fernández and Clara I. Sánchez.

Regional analysis of spontaneous MEG rhythms in patients with Alz-

heimer's disease using spectral entropies.

Annals of Biomedical Engineering, 36(1):141–152, January 2008.

- **Javier Escudero**, Daniel Abásolo, Roberto Hornero, Pedro Espino and Miguel López.
Reply to ‘Comment on “Analysis of electroencephalograms in Alzheimer's disease patients with multiscale entropy”’.
Physiological Measurement, 28(12):L3–L7, December 2007.
- **Javier Escudero**, Roberto Hornero, Daniel Abásolo, Alberto Fernández and Miguel López-Coronado.
Artifact removal in magnetoencephalogram background activity with independent component analysis.
IEEE Transactions on Biomedical Engineering, 54(11):1965–1973, November 2007.
- Carlos Gómez, Roberto Hornero, Daniel Abásolo, Alberto Fernández and **Javier Escudero**.
Analysis of the magnetoencephalogram background activity in Alzheimer's disease patients with auto mutual information.
Computer Methods and Programs in Biomedicine, 87(3):239–247, September 2007.
- **Javier Escudero**, Daniel Abásolo, Roberto Hornero, Pedro Espino and Miguel López.
Analysis of electroencephalograms in Alzheimer's disease patients with multiscale entropy.
Physiological Measurement, 27(11):1091–1106, November 2006.
- Tomás Ortiz, Ana María Martínez, Alberto Fernández, Fernando Maestu, Pablo Campo, Roberto Hornero, **Javier Escudero**, Joaquín Poch.
Impact of auditory stimulation at a frequency of 5 Hz in verbal memory.
Actas Españolas de Psiquiatría, 36(6):307–313, November–December 2008.

A.2 International Book Chapters

- **Javier Escudero**, Roberto Hornero, Daniel Abásolo, Jesús Poza and Alberto Fernández.
Applying Independent Component Analysis to the Artifact Detection Problem in Magnetoencephalogram Background Recordings.
In Nilmini Wickramasinghe and Eliezer Geisler, editors, *Encyclopedia of Healthcare Information Systems*, I:84–92. ISBN: 978-1-59904-889-5. IGI Global, U.S.A., 2008.
- Daniel Abásolo, **Javier Escudero**, Roberto Hornero, Pedro Espino and Carlos Gómez.
Fractal dimension of the EEG in Alzheimer’s disease.
In Nilmini Wickramasinghe and Eliezer Geisler, editors, *Encyclopedia of Healthcare Information Systems*, II:603–609. ISBN: 978-1-59904-889-5. IGI Global, U.S.A., 2008.
- Carlos Gómez, Roberto Hornero, Daniel Abásolo, Alberto Fernández and **Javier Escudero**.
Regularity Analysis of the Magnetoencephalogram Background Activity in Alzheimer’s Disease Patients Using Auto Mutual Information.
In Nilmini Wickramasinghe and Eliezer Geisler, editors, *Encyclopedia of Healthcare Information Systems*, III:1146–1153. ISBN: 978-1-59904-889-5. IGI Global, U.S.A., 2008.

A.3 Contributions to International Conferences

- **Javier Escudero**, Roberto Hornero, Daniel Abásolo, Jesús Poza and Alberto Fernández.
Comparison of the decompositions estimated using five blind source separation algorithms for magnetoencephalogram background activity.
In *Proceedings of The 6th International Workshop on Biosignal Interpretation*, 136–139, 2009.

- **Javier Escudero**, Roberto Hornero, Daniel Abásolo and Alberto Fernández.
Magnetoencephalogram preprocessing with blind source separation algorithms to improve the classification of Alzheimer's disease patients.
In *Proceedings of The 6th International Workshop on Biosignal Interpretation*, 224–227, 2009.
- Daniel Abásolo, Roberto Hornero, Carlos Gómez, **Javier Escudero** and Pedro Espino.
Electroencephalogram background activity characterization with Detrended Moving Average in Alzheimer's disease patients.
In *Proceedings of the 6th IEEE International symposium on Intelligent Signal Processing, WISP 2009*, 211–215, 2009.
- **Javier Escudero**, Roberto Hornero, Daniel Abásolo and Miguel López.
On the application of the auto mutual application rate of decrease to biomedical signals.
In *Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2137–2140, 2008.
- Jesús Poza, Roberto Hornero, **Javier Escudero**, Alberto Fernández and Carlos Gómez.
Analysis of spontaneous MEG activity in Alzheimer's disease using time-frequency parameters.
In *Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 5712–5715, 2008.
- Daniel Abásolo, Roberto Hornero and **Javier Escudero**.
Non-linear analysis of the EEG background activity in Alzheimer's disease patients.
In *Book of abstracts of the Third International Conference on Computational Intelligence in Medicine and Healthcare (CIMED 2007) - Advances In Early Diagnosis And Care For AD*, 14, 2007.

- Daniel Abásolo, Roberto Hornero and **Javier Escudero**.
Non-linear analysis of the EEG background activity in Alzheimer's disease patients (Special Invited Session).
In *Book of abstracts of the Third International Conference on Computational Intelligence in Medicine and Healthcare (CIMED 2007) - Advances In Early Diagnosis And Care For AD - special invited session*, 19, 2007.
- **Javier Escudero**, Roberto Hornero, Daniel Abásolo, Alberto Fernández and Jesús Poza.
Magnetoencephalogram blind source separation and component selection procedure to improve the diagnosis of Alzheimer's disease patients¹.
In *Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 5437–5440, 2007.
- Jesús Poza, Roberto Hornero, Daniel Abásolo, Alberto Fernández and **Javier Escudero**.
Analysis of spontaneous MEG activity in patients with Alzheimer's disease using spectral entropies.
In *Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 6179–6182, 2007.
- Carlos Gómez, Roberto Hornero, Daniel Abásolo, Alberto Fernández and **Javier Escudero**.
Analysis of MEG recordings from Alzheimer's disease patients with sample and multiscale entropies.
In *Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 6183–6186, 2007.
- Daniel Abásolo, Roberto Hornero, Pedro Espino, **Javier Escudero** and Carlos Gómez.

¹This contribution was awarded the Third Prize, and selected as the European Finalist, in the International Student Paper Competition of the IEEE-EMBC07.

Electroencephalogram background activity characterization with approximate entropy and auto mutual information in Alzheimer's disease patients.

In *Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 6191–6194, 2007.

- **Javier Escudero**, Roberto Hornero, Jesús Poza, Daniel Abásolo, Alberto Fernández and Miguel López.

Applying blind source separation to magnetoencephalogram recording in the diagnosis of Alzheimer's disease.

In *Telemedicine in Future Health*, 2006.

- Daniel Abásolo, Roberto Hornero, Carlos Gómez, **Javier Escudero** and Miguel López.

Fluctuation analysis of the electroencephalogram background activity in patients with Alzheimer's disease.

In *Telemedicine in Future Health*, 2006.

- **Javier Escudero**, Roberto Hornero, Daniel Abásolo, Jesús Poza, Alberto Fernández and Miguel López.

Rejection of artifact sources in magnetoencephalogram background activity using independent component analysis.

In *Proceedings of the 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 5282–5285, 2006.

- Carlos Gómez, Roberto Hornero, Alberto Fernández, Daniel Abásolo, **Javier Escudero** and Miguel López.

Magnetoencephalogram background activity analysis in Alzheimer's disease patients using auto mutual information.

In *Proceedings of the 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 6181–6184, 2006.

- **Javier Escudero**, Daniel Abásolo, Roberto Hornero, Pedro Espino and Miguel López.

Electroencephalograms multiscale entropy analysis of Alzheimer's disease patients.

In *Proceedings of the 3rd Institution of Engineering and Technology International Conference on Advances in Medical, Signal and Information Processing, MEDSIP 2006*, 2006.

- Daniel Abásolo, Roberto Hornero, **Javier Escudero**, Carlos Gómez, María García and Miguel López.

Approximate entropy and mutual information analysis of the electroencephalogram in Alzheimer's disease patients.

In *Proceedings of the 3rd Institution of Engineering and Technology International Conference on Advances in Medical, Signal and Information Processing, MEDSIP 2006*, 2006.

A.4 Contributions to National Conferences

- Jesús Poza, Roberto Hornero, María García, **Javier Escudero** and Alberto Fernández.

Análisis de la irregularidad en la actividad MEG de enfermos de Alzheimer mediante representaciones tiempo-frecuencia.

In *Libro de Actas del XXVII Congreso Anual de la Sociedad Española de Ingeniería Biomédica*, 61–64, 2009.

- Roberto Hornero, Daniel Abásolo, **Javier Escudero** and Carlos Gómez.

Análisis no lineal de EEG y MEG para la ayuda en el diagnóstico de la enfermedad de Alzheimer.

In *Libro de Actas del Simposio CEA de BIOINGENIERÍA 2009*, 40–49, 2009.

- Carlos Gómez, Roberto Hornero, Ángela Mediavilla, Alberto Fernández and **Javier Escudero**.

Análisis de la predictibilidad de registros MEG pertenecientes a enfermos de Alzheimer.

In *Libro de Actas del XXVI Congreso Anual de la Sociedad Española de Ingeniería Biomédica*, 37–40, 2008.

- **Javier Escudero**, Roberto Hornero, Daniel Abásolo, Alberto Fernández and Carlos Gómez.
Análisis complementario del magnetoencefalograma en pacientes con la enfermedad de Alzheimer mediante técnicas espectrales y no lineales.
In *Libro de Actas del XXVI Congreso Anual de la Sociedad Española de Ingeniería Biomédica*, 162–165, 2008.
- Alberto Fernández, Javier Quintero, Roberto Hornero, Carlos Gómez, **Javier Escudero**, Pilar Zuluaga and Tomás Ortiz.
Lempel-Ziv complexity analysis of spontaneous brain activity in attention-deficit hyperactivity disorder.
In *Libro de Actas del XXVI Congreso Anual de la Sociedad Española de Ingeniería Biomédica*, 525–528, 2008.
- Carlos Gómez, Ángela Mediavilla, Roberto Hornero, Alberto Fernández and **Javier Escudero**.
Estimación de la dimensión fractal de Higuchi en magnetoencefalogramas de enfermos de Alzheimer.
In *Libro de Actas del XXV Congreso Anual de la Sociedad Española de Ingeniería Biomédica*, 168–171, 2007.
- **Javier Escudero**, Roberto Hornero, Daniel Abásolo, Carlos Gómez and Alberto Fernández.
Separación ciega de fuentes para realzar características espectrales y no lineales del magnetoencefalograma en pacientes de la enfermedad de Alzheimer.
In *Libro de Actas del XXV Congreso Anual de la Sociedad Española de Ingeniería Biomédica*, 222–225, 2007.
- Daniel Abásolo, **Javier Escudero**, Roberto Hornero, Pedro Espino and Carlos Gómez.
Análisis de la actividad base del EEG con la dimensión fractal. Aplicación a la enfermedad de Alzheimer.
In *Libro de Actas del XXIV Congreso Anual de la Sociedad Española de Ingeniería Biomédica*, 347–350, 2006.

- **Javier Escudero**, Roberto Hornero, Jesús Poza, Alberto Fernández and Miguel López.

Estudio de la actividad base del magnetoencefalograma en pacientes con la enfermedad de Alzheimer mediante separación ciega de fuentes y frecuencia media.

In *Libro de Actas del XXIV Congreso Anual de la Sociedad Española de Ingeniería Biomédica*, 527–530, 2006.

- **Javier Escudero**, Roberto Hornero, Daniel Abásolo, Jesús Poza, Alberto Fernández and Miguel López.

Rechazo de artefactos cardiacos y oculares en magnetoencefalografía con análisis de componentes independientes.

In *Libro de actas del XXI Simposium Nacional de la Unión Científica Internacional de Radio, URSI 2006*, 1271–1275, 2006.

Apéndice B

Resúmenes en Español

Contents

B.1	Capítulo 1: Introducción	306
B.2	Capítulo 2: Magnetoencefalografía	307
B.3	Capítulo 3: Enfermedad de Alzheimer	309
B.4	Capítulo 4: Separación ciega de fuentes	310
B.5	Capítulo 5: Registro de MEG	314
B.6	Capítulo 6: Métodos	316
B.7	Capítulo 7: Resultados	318
B.8	Capítulo 8: Discusión	323

Este Apéndice recoge los correspondientes resúmenes en Español de todos los Capítulos que componen esta Tesis Doctoral salvo el Capítulo 9, cuya traducción íntegra al Español se encuentra en el Apéndice C.

B.1 Capítulo 1: Introducción

Esta Tesis Doctoral se enmarca en la disciplina científica del Procesado de Señales Biomédicas. Este área se preocupa de la adquisición y procesado matemático de señales informativas derivadas de sistemas biológicos (Onaral, 2006). En la mayoría de las aplicaciones, el mero registro de estas señales no es suficiente para obtener la información contenida en ellas. Por ello, deben emplearse procedimientos específicos para obtener tal información. Estos pueden incluir diversos tratamientos de señal para realzar sus características más importantes o para reducir la cantidad de ruido externo o artefactos que afectan a los registros (Laguna and Sörnmo, 2009).

Las señales biomédicas que se analizarán en esta Tesis Doctoral son conocidas como registros de MEG. Éstos reflejan los campos magnéticos generados por la actividad neuronal en el cerebro. Diversas técnicas de BSS serán empleadas para procesar señales de MEG de pacientes con un diagnóstico de AD probable y sujetos sanos de control de edad avanzada con los siguientes objetivos:

- Comparar la similitud de las descomposiciones calculadas con diversas técnicas de BSS para señales reales de MEG.
- Evaluar de un modo objetivo y automático la eliminación basada en BSS de los artefactos cardíaco y ocular, así como de la interferencia de red eléctrica, que contaminan los registros de MEG.
- Mejorar la clasificación de las señales de MEG adquiridas de pacientes con AD frente a aquellas de sujetos de control mediante el uso de un preprocesado de BSS.
- Extraer adaptativamente de los registros de MEG con técnicas de BSS actividad rítmica cerebral adecuada para ser analizada con medidas de

conectividad.

Estas aplicaciones se basan en las siguientes hipótesis:

- Que la comparación de las descomposiciones estimadas con varias técnicas de BSS para los mismos registros de MEG puede ofrecer información relevante acerca de las relaciones y similitudes existentes entre aquellas.
- Que el uso conjunto de la BSS y métricas especialmente diseñadas para detectar artefactos en las componentes BSS puede ayudar a eliminar automáticamente la contaminación por CA, OA y PLN presente en la actividad basal de MEG.
- Que un preprocesado adecuadamente diseñado con BSS puede ser útil para distinguir características calculadas a partir de los registros de MEG de pacientes con AD de aquellas computadas para señales de sujetos de control.
- Que una metodología basada en EMD y cBSS puede proporcionar una extracción adaptativa de ritmos cerebrales sobre diversas regiones del cuero cabelludo con información relevante acerca de la conectividad de pacientes con la AD en comparación sujetos de control.

En resumen, podemos decir que el principal objetivo de esta Tesis Doctoral es aplicar técnicas de BSS para ayudar en el análisis e interpretación de actividad basal de MEG registrada en pacientes con la AD y sujetos de control.

B.2 Capítulo 2: Magnetoencefalografía

El MEG es el registro no invasivo de los pequeños campos magnéticos generados por la actividad neuronal en el cerebro. Junto con el EEG, es la única técnica neurofisiológica capaz de medir la actividad de la corteza cerebral directamente, sin tener que interpretar la información en función de cambios metabólicos o vasculares (Hari, 2004).

Las señales MEG se registran sobre la superficie del cuero cabelludo. Aunque reflejan la actividad de las corrientes generadas en la parte activa del cerebro en cada momento, la mayor parte de la señal procede de la corteza cerebral, la cual contiene células piramidales alineadas. Las neuronas se componen de dendritas, cuerpo celular (o soma) y axón. Un conjunto de axones forma las fibras nerviosas, destinadas a conectar diversas regiones del cerebro. Dentro de cada neurona, la información se transmite en forma de impulso eléctrico, mientras que la transferencia de información entre neuronas se realiza por medio de neurotransmisores químicos en las sinapsis ([Vrba and Robinson, 2001](#)).

La adquisición del MEG es una tarea extremadamente compleja. Necesita de la utilización de materiales superconductores en los sensores para registrar las diminutas variaciones del campo magnético cerebral. Puesto que los materiales superconductores deben operar a temperaturas próximas al cero absoluto, éstos deben estar inmersos en un baño de He líquido para mantenerlos refrigerados. Asimismo, es necesario realizar las medidas en una habitación magnéticamente aislada para así disminuir el impacto de fuentes externas de ruido magnético (como vehículos) en la adquisición de las señales ([Vrba and Robinson, 2001](#)).

No obstante, los artefactos constituyen otra importante fuente de actividad indeseada en el MEG. Éstos son actividad no deseada que aparece en los registros y puede afectar los análisis. Pueden estar producidos por órganos del propio sujeto como los ojos o el corazón ([Hämäläinen et al., 1993](#)). Por ejemplo, el corazón genera un campo magnético ostensiblemente más potente que el creado por las neuronas ([Jousmäki and Hari, 1996](#)) y los parpados o movimientos oculares pueden también causar una contaminación significativa de las señales MEG ([Antervo et al., 1985](#)). Es obvio que el aislamiento magnético no ayuda a reducir estos artefactos puesto que el sujeto se encuentra dentro de la habitación magnéticamente aislada. Además, la interferencia de red eléctrica a 50 Hz puede aparecer también en este tipo de señales ([Hämäläinen et al., 1993](#)).

B.3 Capítulo 3: Enfermedad de Alzheimer

La AD es una enfermedad neurodegenerativa progresiva que conduce a la muerte del paciente (Blennow et al., 2006). Además de la famosa pérdida de memoria asociada con esta patología, la AD también causa otros déficits cognitivos (por ejemplo: desorientación, confusión y disminución de la capacidad para razonar) y conductuales (como agitación, ansiedad, delirio, depresión, alucinaciones o insomnio) que merman notablemente la capacidad del paciente para realizar tareas cotidianas (Blennow et al., 2006; Cummings, 2004; Lahiri et al., 2002).

El factor de riesgo más importante en la AD es la edad. Asimismo, diversos estudios epidemiológicos han sugerido otros factores de riesgo relacionados con enfermedades vasculares o hábitos alimenticios, aunque los resultados no son concluyentes (Blennow et al., 2006; Lahiri et al., 2002; Purnell et al., 2009). Además, la AD tiene un significativo componente genético en sus dos formas: familiar y esporádica (Blennow et al., 2006). La primera ha sido asociada con mutaciones en los genes APP, PSEN1 y PSEN2 (Blennow et al., 2006; van Es and van den Berg, 2009), mientras que la forma esporádica, mucho más común, se ha relacionado con el gen APOE (Blennow et al., 2006) y, más recientemente, con SORL1, CLU, PICALM y CR1 (Harold et al., 2009; Lambert et al., 2009; National Institute on Aging, 2008).

A nivel microscópico, las lesiones más representativas de la AD son las placas seniles y los ovillos neurofibrilares. Además, la AD produce degeneración celular y sináptica (Mattson, 2004). El mecanismo por el cual aquellas lesiones provocan la neurodegeneración no se conoce con total exactitud (van Es and van den Berg, 2009). Sin embargo, sí existe un cierto consenso en cuanto a que la producción y acumulación de $A\beta$ juega un papel esencial en la patogénesis de esta demencia puesto que se ha relacionado con la formación de placas seniles y ovillos neurofibrilares y con procesos de oxidación, inflamación y fomento de muerte celular, entre otros (Cummings, 2004; Masters et al., 2006).

En lo referente a su impacto en la sociedad, cabe destacar que la AD es el tipo más común de demencia en Occidente. Representa entre el 50 %

y el 60% de todos los casos (Blennow et al., 2006). Aunque su prevalencia se sitúa por debajo del 1% en sujetos de entre 60 y 64 años, su incidencia se dobla, aproximadamente, con cada 5 años de edad (Cummings, 2004). En consecuencia, la AD afecta a entre el 24% y el 33% de la población mayor de 85 años (Blennow et al., 2006). Teniendo en cuenta el aumento de la población y de la esperanza de vida, se estima que unos 81 millones de personas padecerán demencia en 2040 (Blennow et al., 2006). Por todo ello, resultan evidentes los elevados costes de los sistemas de salud asociados a esta demencia (Ferri et al., 2006).

Puesto que cabe pensar que los gastos relacionados con las pruebas diagnósticas de la AD contribuirán tan sólo de forma marginal a esos costes, resulta deseable desarrollar nuevas alternativas que puedan ayudar en la detección de esta demencia (Lahiri et al., 2002).

B.4 Capítulo 4: Separación ciega de fuentes

El uso de la BSS en el procesado de señales biomédicas ha ganado popularidad recientemente (James and Hesse, 2005; Onton et al., 2006; Vigário and Oja, 2008). Estas técnicas son útiles para separar registros multidimensionales en las componentes internas que los constituyeron (James and Hesse, 2005; Vigário and Oja, 2008). Su uso implica la separación o extracción de fuentes estadísticamente independientes que dieron lugar a la actividad registrada. Para ello, estos métodos definen un modelo de generación de las señales observadas, las cuales se asumen relacionadas entre sí (James and Hesse, 2005; Vigário and Oja, 2008).

Se necesitan ciertas hipótesis para calcular la BSS (James and Hesse, 2005). La más importante de todas ellas es que las fuentes, o componentes, son mutuamente independientes entre sí o, de forma alternativa, que están incorreladas para cualquier retardo temporal. La BSS extrae estas fuentes explotando esta hipótesis de independencia o decorrelación espacio-temporal (Hyvärinen et al., 2001; James and Hesse, 2005; Onton et al., 2006). De hecho, BSS es una técnica más potente que otros métodos clásicos de descomposición de señales como PCA y FA (Hyvärinen et al., 2001; James and Hesse, 2005).

El problema básico que la BSS (Cichocki and Amari, 2002; Hyvärinen et al., 2001) intenta resolver puede enunciarse como sigue. Supongamos que $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_m(t)]^T$ denota un conjunto de m muestras temporales en un instante de tiempo t . $\mathbf{x}(t)$ se asume compuesto de una combinación lineal de n componentes internas $\mathbf{s}(t) = [s_1(t), s_2(t), \dots, s_n(t)]^T$. En el caso más sencillo, el proceso de mezcla se considera lineal y se obvia el ruido aditivo externo, resultando en:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t), \quad (\text{B.1})$$

donde \mathbf{A} es una matriz de mezcla de rango completo $m \times n$. Se asume que $m \geq n$ (Hyvärinen et al., 2001; James and Hesse, 2005; Vigário and Oja, 2008).

Los algoritmos de BSS calculan una estimación, $\hat{\mathbf{s}}(t)$, de las componentes reales subyacentes en el proceso, $\mathbf{s}(t)$, a partir de las medidas, $\mathbf{x}(t)$, por medio de:

$$\hat{\mathbf{s}}(t) = \mathbf{B}\mathbf{x}(t), \quad (\text{B.2})$$

donde \mathbf{B} es la matriz de separación estimada, la cual se corresponde con la pseudo-inversa de \mathbf{A} (Cichocki and Amari, 2002; Hyvärinen et al., 2001).

Para encontrar la solución se adoptan las hipótesis de mezcla lineal e instantánea, estacionaria e independencia estadística de las componentes (Cichocki and Amari, 2002; Hyvärinen et al., 2001).

Existen diversas metodologías para estimar la BSS (Cichocki and Amari, 2002; Hyvärinen et al., 2001). Mientras algunos métodos se basan en HOS de las distribuciones de amplitud de las fuentes (comúnmente conocidos como ICA; Vigário and Oja, 2008), otras técnicas consideran la estructura temporal (o espectral) de las señales. Éstos se denominan técnicas basadas en SOS (James and Hesse, 2005).

La metodología ICA obvia la información temporal de los datos (Cichocki and Amari, 2002; Hyvärinen et al., 2001; James and Hesse, 2005). Estos algoritmos buscan combinaciones de señales que generen estimaciones de las fuentes lo menos gaussianas posibles ya que buscar componentes independientes equivale a buscar componentes no gaussianas (Hyvärinen et al., 2001;

James and Hesse, 2005). Ejemplos de estas técnicas son FastICA (Hyvärinen, 1999), InfoMax (Bell and Sejnowski, 1995) y JADE (Cardoso and Souloumiac, 1993).

Las matrices de separación y mezcla pueden estimarse a partir de un conjunto de matrices de covarianza calculadas en diversos retardos temporales (Hyvärinen et al., 2001). Si este procedimiento se realiza considerando únicamente dos retardos temporales, se obtiene el algoritmo AMUSE (Tong et al., 1991). Por el contrario, si la técnica se extiende a más de dos retardos, el procedimiento requiere de técnicas iterativas resultando en el método SOBI (Belouchrani et al., 1997; Ziehe and Muller, 1998).

Todos los anteriores algoritmos son capaces de extraer componentes caracterizadas por ser lo más independientes posibles desde diversos puntos de vista. Sin embargo, en ciertas aplicaciones, puede ser aconsejable extraer una única componente caracterizada por guardar similitud con cierta señal de referencia. En tales situaciones, las técnicas de cBSS resultan extremadamente útiles (Huang and Mi, 2007; James and Gibson, 2003; Lu and Rajapakse, 2005).

Como puede observarse, existe una amplia variedad de técnicas de BSS y no todos los procedimientos emplean los mismos principios (Cichocki and Amari, 2002; Hyvärinen et al., 2001; James and Hesse, 2005). Por ello, puede resultar difícil en ocasiones seleccionar a priori el algoritmo más adecuado para una aplicación concreta (Fitzgibbon et al., 2007; Zavala Fernández et al., 2006). Por ello, resulta importante estudiar la consistencia, o similitud, de las separaciones calculadas por diversos algoritmos para registros reales de actividad electromagnética cerebral de tal modo que las relaciones entre técnicas BSS puedan entenderse en mayor profundidad.

Asimismo, existen numerosas aplicaciones de la BSS al análisis de EEG y MEG relacionadas con los estudios llevados a cabo en esta Tesis Doctoral.

La primera de ellas es el rechazo de artefactos en EEG y MEG con BSS. A pesar de la relativa popularidad de estos procedimientos (Barbati et al., 2004; Castellanos and Makarov, 2006; Dammers et al., 2008; Delorme et al., 2007; Escudero et al., 2006c,d, 2007b, 2008c; Joyce et al., 2004; LeVan et al., 2006; Li et al., 2006; Mantini et al., 2008; Okada et al., 2007; Romero et al., 2008;

Shao et al., 2008; Ting et al., 2006), existen pocos estudios comparativos de técnicas en la Literatura. Además, la evaluación del rechazo de artefactos se realiza habitualmente mediante una inspección visual de los resultados, existiendo pocos estudios que realicen una evaluación cuantitativa de éstos (Dammers et al., 2008; Escudero et al., 2007b; Mantini et al., 2008). Además, la mayor parte de los esfuerzos se han centrado en EEG (Castellanos and Makarov, 2006; Fitzgibbon et al., 2007; Joyce et al., 2004; Romero et al., 2008, 2009; Shao et al., 2008; Ting et al., 2006), con el MEG recibiendo menos atención (Dammers et al., 2008; Escudero et al., 2007b; Mantini et al., 2008). Por ello, esta Tesis Doctoral pretende evaluar cuantitativamente diversas técnicas de rechazo de artefactos con BSS en señales de MEG.

Los registros cerebrales también pueden procesarse con métodos de BSS para ayudar en la detección de diversas enfermedades neurológicas (Hung et al., 2007; Jarchi et al., 2009; Kobayashi et al., 1999). Teniendo en cuenta la propia complejidad de los registros cerebrales, puede asumirse que la aplicación de una BSS, junto con técnicas de caracterización de señal, puede mejorar la detección de sujetos con patologías como la AD (Cichocki et al., 2005). Esto se debe a que algunas componentes BSS podrían ser más sensibles a esos estados cerebrales que otras (Cichocki et al., 2005; Escudero et al., 2008d, 2009b; Vialatte et al., 2005). Así pues, podrían seleccionarse tales componentes para mejorar la separación basada en el MEG entre, por ejemplo, pacientes con AD y sujetos sanos de edad avanzada. A pesar de las ventajas que este procedimiento puede reportar, pocos estudios lo han empleado (Escudero et al., 2008d, 2009b). En consecuencia, parte de esta Tesis Doctoral se dedica a explorar las posibilidades que este planteamiento ofrece.

La AD causa, además de alteraciones en el espectro y las características no lineales de los registros de EEG y MEG (Hornero et al., 2009b; Jeong, 2004; Stam, 2005, 2010), cambios en el patrón de sincronización de los ritmos cerebrales (Jeong, 2004). Sin embargo, estos efectos no son siempre fáciles de detectar (Dauwels et al., 2010). Además, en la Literatura se han reportado resultados contradictorios en las bandas espectrales δ y θ (Jeong, 2004). Por ello, esta Tesis Doctoral plantea un procedimiento alternativo para estudiar la sincronización entre regiones. Éste se basa en las propiedades adaptativas

de la EMD (Huang et al., 1998) y la extracción de señales de la cBSS. Esta metodología buscará analizar diversas bandas espectrales de interés (δ , θ , α y β) y regiones del cuero cabelludo con el fin de analizar los patrones de conectividad de señales en la AD.

B.5 Capítulo 5: Registro de MEG

En esta Tesis Doctoral se ha analizado la actividad basal de MEG registrada en 62 sujetos. De ellos, 36 eran pacientes diagnosticados con AD probable y 26 eran sujetos de control de edad avanzada. Los principales datos sociodemográficos de ambos grupos de sujetos, así como de sus divisiones en conjuntos de entrenamiento y validación se muestran en la Tabla B.1.

Todos los pacientes cumplían los criterios de AD probable de acuerdo con las guías de la NINCDS-ADRDA (McKhann et al., 1984) y fueron reclutados de la “Asociación de Familiares de Enfermos de Alzheimer” en Madrid. Cabe destacar que no hay diferencias significativas entre la edad media de ambos grupos de sujetos ($p = 0,1911 > 0,05$). Se obtuvo consentimiento informado de todos los sujetos de control y de los cuidadores de los pacientes de AD para su participación en este estudio, el cual fue aprobado por el correspondiente comité ético.

Los registros de MEG se realizaron con un magnetómetro de 148 canales (MAGNES 2500WH, 4D Neuroimaging) ubicado en una habitación magnéticamente aislada del “Centro de Magnetoencefalografía Dr. Pérez-Modrego” de la Universidad Complutense de Madrid. Para minimizar la cantidad de artefactos en los registros, éstos se adquirieron mientras los sujetos yacían en una camilla, despiertos, en un estado relajado y con los ojos cerrados.

Para cada sujeto, se realizó un registro MEG de cinco minutos de duración con una frecuencia de muestreo de 678.17 Hz. Posteriormente, las señales se submuestrearon por un factor cuatro para reducir su longitud, siendo la frecuencia de muestreo resultante de 169.54 Hz. En los casos necesarios, se empleó la herramienta “*Filter Design & Analysis*” de Matlab para diseñar filtros FIR paso-banda con ventana de Hamming para limitar el ancho de banda de las señales.

Tabla B.1: Principales datos sociodemográficos y clínicos de todo los participantes y la divisiones en grupos de entrenamiento y validación. Los datos se expresan como $\text{media} \pm \text{SD}$.

Todos los sujetos		
	Pacientes con AD	Sujetos de Control
Número de sujetos	36	26
Número de mujeres	24	17
Edad	74.06 ± 6.95	71.77 ± 6.38
MMSE	18.06 ± 3.36	28.88 ± 1.18
GDA/FAST	4.17 ± 0.45	1.73 ± 0.45
Grupo de entrenamiento		
	Pacientes con AD	Sujetos de Control
Número de sujetos	18	13
Número de mujeres	12	9
Edad	74.11 ± 7.38	71.38 ± 4.84
MMSE	17.72 ± 3.63	28.92 ± 1.04
GDA/FAST	4.17 ± 0.38	1.85 ± 0.38
Grupo de validación		
	Pacientes con AD	Sujetos de Control
Número de sujetos	18	13
Número de mujeres	12	8
Edad	74.00 ± 6.70	72.15 ± 7.82
MMSE	18.39 ± 3.15	28.85 ± 1.34
GDA/FAST	4.17 ± 0.51	1.62 ± 0.51

En las aplicaciones que emplean la BSS como un preprocesado para aumentar la precisión diagnóstica basada en el MEG (Sección 6.6) o en la extracción adaptativa de actividad rítmica (Sección 6.7) las épocas de MEG con los artefactos más evidentes fueron rechazadas. Se empleó este método de rechazo de artefactos para evitar complejidad adicional en los análisis y para evaluar el rendimiento de estas aplicaciones de BSS sin ningún otro tipo de preprocesado. En estos casos, se empleó un método semi-automático que combinó una inspección visual de los registros con un método de umbralización en amplitud para detectar actividad espuria de manera similar a otros

estudios de actividad de MEG (Osipova et al., 2003, 2005, 2006a,b; Pekkonen et al., 1999, 2005).

B.6 Capítulo 6: Métodos

Diversos estudios han sido realizados en esta Tesis Doctoral. No obstante, en el caso del análisis de la consistencia entre técnicas de BSS y del rechazo de artefactos, fue necesario emplear una etapa de preprocesado. Se evaluó la eficacia de aquella propuesta por Cao et al. (2003), la cual se basa en FA. Este método tiene en cuenta los elementos en la diagonal principal de la matriz de covarianza de ruido. Mediante un procedimiento iterativo (Cao et al., 2003; Escudero et al., 2007b), se calculan las matrices que darán lugar a la matriz de preprocesado de los datos. Asimismo, para determinar el valor óptimo de n , se empleó un método de FA basado en un modelado estadístico de los datos: el MDL (Escudero et al., 2007b; Ikeda and Toyama, 2000). Para evaluar todo este preprocesado se emplearon señales sintéticas a las que se superpuso ruido aditivo externo de diversa intensidad (Escudero et al., 2007b).

La metodología básica de los diversos estudios llevados a cabo en esta Tesis Doctoral se resume a continuación:

1. La aplicación destinada a evaluar la consistencia (o similitud) de diversos algoritmos de BSS se fundamentó en comparar las descomposiciones calculadas para las mismas señales reales de MEG. Para ello, dos algoritmos diferentes de BSS (algoritmo q y algoritmo r) se aplican al mismo conjunto de datos para estimar las correspondientes matrices de mezcla : \mathbf{A}^q y \mathbf{A}^r (Cichocki et al., 2009). Posteriormente, las columnas de estas matrices se normalizan y se calcula la matriz \mathbf{P}^{qr} según:

$$\mathbf{P}^{qr} = (\mathbf{A}^q)^{-1} \mathbf{A}^r. \quad (\text{B.3})$$

Si los dos algoritmos, q y r , proporcionan exactamente la misma separación, \mathbf{P}^{qr} tendrá la forma de una matriz de permutación generalizada. Del mismo modo, cuanto más similar sea \mathbf{P}^{qr} a este tipo de matriz, más

consistentes (similares) serán las separaciones de los algoritmos q y r (Cichocki et al., 2009). En consecuencia, empleando una métrica adecuada puede medirse la similitud de las descomposiciones computadas con distintos algoritmos (Escudero et al., 2009d).

2. La segunda aplicación busca evaluar la utilidad de la BSS en el rechazo de artefactos. Esta aproximación al problema basada en BSS puede ser útil puesto que las señales se separan en componentes con mínima interdependencia. De este modo, se pueden inspeccionar las componentes para encontrar cuáles son responsables de la presencia de artefactos en los registros (Barbati et al., 2004; Escudero et al., 2007b; Ting et al., 2006; Vigário, 1997). Estas componentes con artefactos se pueden eliminar para reducir la influencia de la contaminación en las señales. Existen muy diversos criterios para decidir qué componentes se corresponden con actividad de artefactos. Por ello, se pretende evaluar y comparar de forma objetiva el rendimiento de diversas combinaciones de técnica de BSS, longitud de época y criterio de detección de artefactos. Esta aplicación se encuentra motivada por el hecho de que, en la mayoría de los casos, la utilidad de la BSS en el rechazo de artefactos se ha medido sólo con datos sintéticos (Barbati et al., 2004; Delorme et al., 2007; Kierkels et al., 2006; Romero et al., 2008, 2009; Wallstrom et al., 2004) o por medio de una inspección visual de los registros (Flexer et al., 2005; Greco et al., 2006; Joyce et al., 2004; Jung et al., 2000; LeVan et al., 2006; Li et al., 2006; Mammone and Morabito, 2008; Shoker et al., 2005a; Ting et al., 2006; Vigário, 1997; Vorobyov and Cichocki, 2002). Sin embargo, sería deseable desarrollar evaluaciones objetivas de la eliminación de los artefactos, de forma similar a (Dammers et al., 2008; Escudero et al., 2007b; Mantini et al., 2008).
3. Es tercer estudio de esta Tesis Doctoral intenta mejorar la capacidad de características espectrales y no lineales calculadas del MEG para distinguir pacientes con AD de sujetos de control mediante un preprocesado basado en BSS. Para ello, se emplearon los algoritmos AMUSE (Tong

et al., 1991), SOBI (Belouchrani et al., 1997), FastICA (Hyvärinen and Oja, 1997) y eInfoMax (Lee et al., 1999). Por sencillez, se ha asumido $m = n$ gracias a que únicamente las componentes con diferencias más significativas entre ambos grupos de sujetos se utilizarán para calcular una reconstrucción parcial de las señales de MEG (Escudero et al., 2008d, 2009b). En este estudio resulta clave ordenar las componentes extraídas para poder realizar comparaciones entre sujetos encaminadas a discernir dónde se concentran las diferencias más significativas entre grupos de sujetos. Esta ordenación se basó en dos criterios distintos: MF y curtosis. Las características de señal MF, SpecEn, LZC y SpecEn se calcularon para las señales parcialmente reconstruidas con las componentes más significativas y para los registros originales, de tal modo que pudo compararse la clasificación obtenida en cada caso.

4. La última parte de la Tesis Doctoral busca extraer de forma adaptativa, haciendo uso de una cBSS, actividad rítmica en las bandas espectrales típicas (δ , θ , α y β) sobre diversas regiones del cuero cabelludo. Este procedimiento, en lugar de fijar bandas espectrales a priori, emplea una EMD (Huang et al., 1998) para descomponer la actividad cerebral registrada en cada canal en IMFs. Después, las IMFs calculadas a partir de todos los canales en la misma región se introducen en un procedimiento de clúster para seleccionar automáticamente una referencia en cada una de las bandas (δ , θ , α y β). Esta referencia se emplea junto con la cBSS para extraer actividad de interés simultáneamente en todos los canales de la región. Por último, se calcula la $\text{Coh}(f)$ para cada par de señales extraídas con el fin de evaluar la conectividad en pacientes con la AD frente a sujetos de control.

B.7 Capítulo 7: Resultados

Para evaluar el grado en el cual el preprocesado basado en FA (Cao et al., 2003; Escudero et al., 2007b; Ikeda and Toyama, 2000) era capaz de estimar adecuadamente el número de componentes internas y la potencia del ruido

externo añadido a las mezclas, se crearon 1000 mezclas sintéticas distintas con diversos niveles de ruido externo. A continuación, el preprocesado se aplicó diez veces a cada una de estas mezclas (Escudero et al., 2007b). Cabe reseñar que los valores de SNR de las mezclas sintéticas estuvieron comprendidos entre -14 dB y 22 dB. Además, el preprocesado basado en FA y MDL se comparó con otras metodologías que empleaban PCA.

Como cabía esperar, los criterios que empleaban PCA sobre-estimaban el valor de n cuando la potencia de ruido externo no era despreciable (Escudero et al., 2007b). Por el contrario, el criterio MDL proporcionó estimaciones más precisas del número de componentes internas para los datos sintéticos. Además, la aplicación del preprocesado a las señales reales apoyó la elección del FA en lugar de PCA en el sentido de que los valores de potencia de ruido externo estimada se encontraban en el rango de SNR para el cual la MDL proporcionó estimaciones precisas en las señales sintéticas.

En lo referente a la comparación de las descomposiciones estimadas con diversas técnicas de BSS, las matrices \mathbf{P}^{qr} se calcularon para cada par de algoritmos y diversas longitudes de época (10 s, 20 s, 40 s, 60 s y 90 s). Estas matrices se caracterizaron con una métrica F destinada a proporcionar una idea acerca de la similitud de las separaciones calculadas con q y r . Para reducir la cantidad de datos a analizar, sólo se tuvo en cuenta la matriz \mathbf{P}^{qr} en lugar de ambas \mathbf{P}^{qr} y \mathbf{P}^{rq} (Escudero et al., 2009d). Estos resultados se muestran en la Fig. B.1, donde menores valores se corresponden con pares de algoritmos más consistentes (similares).

Puede observarse que los pares de algoritmos más consistentes fueron AMUSE–SOBI (métodos SOS), seguidos del par JADE–FastICA (HOS). Estos resultados se corroboraron con los pertinentes análisis estadísticos. Además, el nivel global de consistencia mejoró conforme se analizaron épocas de mayor longitud.

En la parte de esta Tesis Doctoral destinada a evaluar de forma objetiva el rechazo del CA, OA y PLN con técnicas de BSS se descompusieron épocas de 10 s, 20 s, 40 s, 60 s y 90 s con las siguientes técnicas BSS: AMUSE, SOBI, JADE, eInfoMax y FastICA. Además, un método de cBSS se empleó para reducir el CA con una referencia calculada como el promedio de todos los

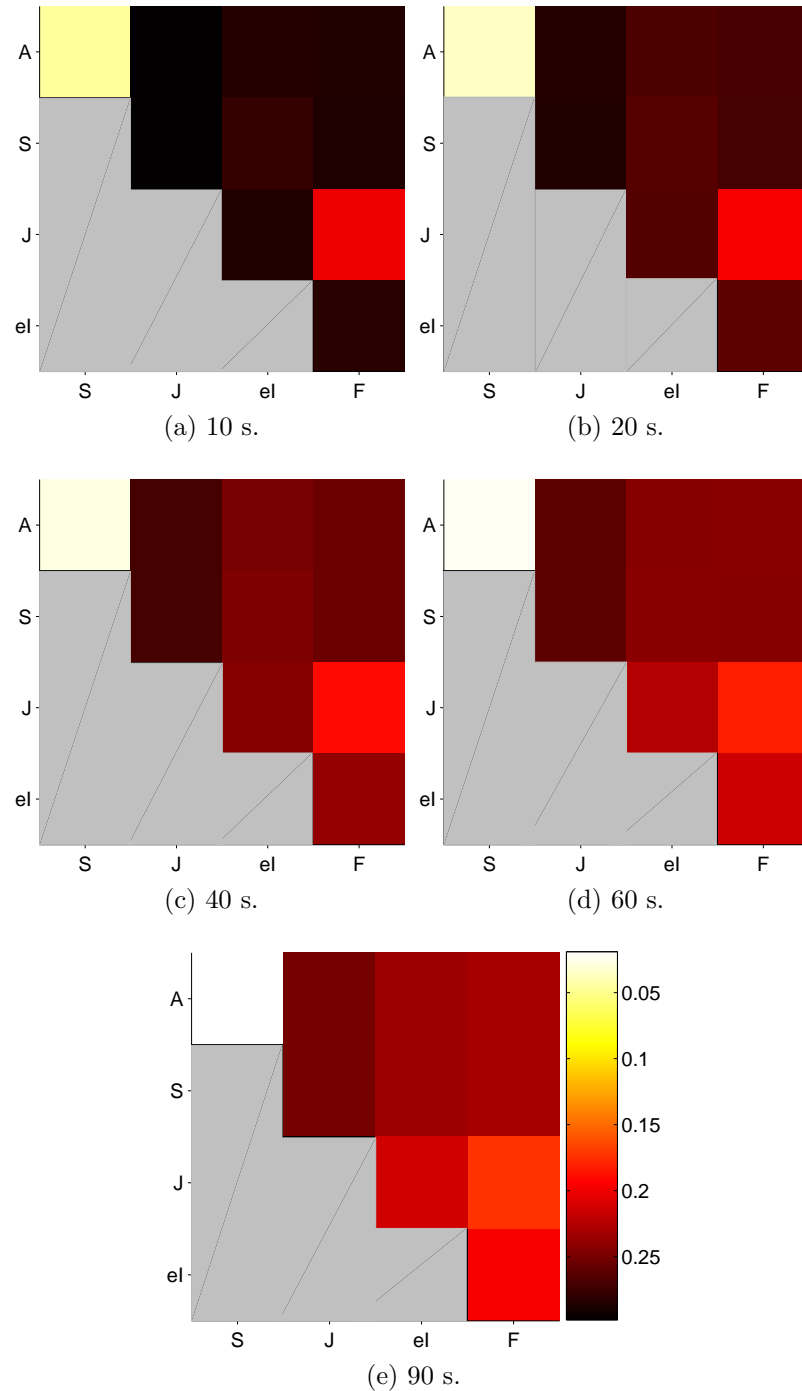


Figura B.1: Promedio de la métrica F para cada par de algoritmos BSS (A: AMUSE, S: SOBI, J: JADE, el: extended Infomax, F: FastICA) y longitud de época.

canales de MEG.

En primer lugar, se evaluó el impacto global de estos artefactos en toda la base de datos con el fin de obtener referencias válidas para medir la disminución de los artefactos después de la BSS (Escudero et al., 2007b). Tras la aplicación de las técnicas de BSS y de diversas métricas de detección de artefactos en componentes, se comprobó que algunas combinaciones de métodos BSS y criterios de detección fueron capaces de reducir de forma notable el impacto del CA: los valores de A_{PtP} y RMS disminuyeron, aproximadamente, de 0.750 y 0.120 a 0.100 y 0.028, respectivamente, después del rechazo de artefactos. Por otro lado, la evaluación de la disminución del OA se realizó tanto en el dominio temporal (contando el número de picos de la señal de MEG por encima de un determinado valor) como espectral (midiendo la diferencia en bajas frecuencias entre espectros calculados en diferentes subconjuntos de canales). Los resultados indicaron que el rechazo de OAs con BSS disminuía la potencia de señal localizada en bajas frecuencias cerca de los ojos, lo cual puede interpretarse como un decrecimiento de la actividad ocular, y reducía el número de picos en las señales con valores extremos. Por último, la evaluación de la reducción en el PLN arrojó tasas de reducción tras el rechazo de artefactos con BSS de unos 14.7 dB.

La tercera aplicación de esta Tesis Doctoral se basa en la idea de que la BSS del MEG puede contener componentes especialmente adecuadas para distinguir los pacientes con AD de los sujetos de control. De este modo, un subconjunto de tales componentes podría servir para realizar reconstrucciones parciales del MEG que condujeran a una mejor distinción de ambos grupos de sujetos (Escudero et al., 2008d, 2009b). En primer lugar, los algoritmos AMUSE, SOBI, FastICA y eInfoMax se emplearon para descomponer las señales. Mientras AMUSE proporciona un orden natural para las componentes, éstas tuvieron que ordenarse según su MF o curtosis en los casos de SOBI, FastICA y eInfoMax para poder realizar comparaciones entre sujetos (Escudero et al., 2009b,c). Posteriormente, un análisis estadístico sirvió para indicar qué subconjuntos de componentes proporcionaban las diferencias más significativas entre sujetos (Escudero et al., 2008d, 2009b,c). Estos subconjuntos, definidos como intervalos continuos de 15 ó 30 compo-

Tabla B.2: Resultados de la clasificación con las medidas de coherencia calculadas a partir del procedimiento adaptativo de extracción de ritmos cerebrales.

	δ	θ	α	β	Todas las bandas
Precisión (%)	67.7	46.8	69.4	43.5	96.8
Sensibilidad (%)	63.9	47.2	66.7	44.4	94.4
Especificidad (%)	73.1	46.2	73.1	42.3	100.0
AUC	0.713	0.874	0.815	0.912	1.000

nentes con el menor p -valor promedio, se utilizaron para calcular reconstrucciones parciales del MEG.

Los resultados indicaron que estos rangos de componentes seleccionadas tendían a ser consistentes entre las cuatro características de señal consideradas (MF, SpecEn, LZC y SampEn). Es más, este procedimiento fue capaz de mejorar el valor de AUC y precisión en múltiples casos.

La última aplicación de analizada en esta Tesis Doctoral pretendía realizar un análisis de la conectividad de actividad rítmica de MEG extraída de forma adaptativa sobre diversas regiones del cuero cabelludo. Este análisis de sincronización se basó en una medida de $\text{Coh}(f)$. Por su parte, la extracción de señales empleó un procedimiento compuesto de una EMD, un clúster k -medias y una cBSS. Las pruebas estadísticas mostraron que existían efectos significativos en los valores de $\text{Coh}(f)$ en función de la banda espectral analizada y del par de regiones considerado. Sin embargo, no aparecieron diferencias significativas entre grupos de sujetos. A pesar de esto, las medias marginales estimadas sugirieron que AD parecía disminuir la $\text{Coh}(f)$ en bandas α y β y aumentarla en δ .

Además de este análisis, se cuantificó la precisión de ambos grupos de sujetos con un LDA con introducción de variables por pasos junto con un procedimiento de validación cruzada dejando uno fuera. Los resultados para cada banda, y cuando éstas se consideran de forma conjunta, aparecen recogidos en la Tabla B.2, así como los valores de AUC correspondientes.

B.8 Capítulo 8: Discusión

El objetivo de esta Tesis Doctoral era aplicar métodos de BSS para ayudar en el análisis e interpretación de actividad basal de MEG adquirida en pacientes con AD y sujetos de control de edad avanzada. Este objetivo global se evaluó atendiendo a cuatro aplicaciones concretas.

Aunque el preprocesado basado en FA y MDL no es una técnica de BSS, su uso sí puede ser útil como un paso previo a este tipo de descomposiciones en diversas situaciones. El criterio MDL proporcionó estimaciones más precisas del valor n que otras metodologías que empleaban PCA. Estos resultados confirmaron que estas últimas no sólo poseen un cierto grado de subjetividad y dependen de la forma del espectro de autovalores (James and Hesse, 2005), sino que también pueden tender a sobre-estimar n cuando el ruido externo no es despreciable (Escudero et al., 2007b). Estos problemas se evitan con el MDL de FA.

Estos resultados corroboran los hallazgos recientemente publicados por Hesse (2007, 2008) acerca de la selección del orden adecuado para el modelo. Estos estudios investigaron el rendimiento de diversos preprocesados con PCA y FA para estimar el número de componentes BSS en registros reales de EEG y MEG. Tales resultados indicaron que un PCA probabilístico o un modelo de FA proporcionaba estimaciones de la dimensionalidad para fiables e independientes de la potencia de la señal que técnicas de PCA comúnmente empleadas (Hesse, 2007, 2008).

Los resultados acerca de la comparación de técnicas BSS mostrados en la Fig. B.1 indicaron claramente que el par de algoritmos AMUSE-SOBI computaba las descomposiciones más similares para las señales de MEG. El principio de descomposición que comparten ambas técnicas es la diagonalización simultánea de varias matrices de covarianza calculadas para varios retardos temporales (Hyvärinen et al., 2001; James and Hesse, 2005). Los resultados también indicaron que las descomposiciones de JADE y FastICA se caracterizaban por un elevado grado de consistencia. Esto podría deberse a que los principios teóricos sobre los que se asientan ambos algoritmos pueden relacionarse (Hyvärinen et al., 2001).

Asimismo, los algoritmos de BSS tendían a estimar descomposiciones más similares para épocas de señal de mayor longitud (Escudero et al., 2009d). Esto puede deberse al hecho de que, a pesar de que el valor de n aumentó con la longitud de la señal, el número de muestras disponibles para la estimación de \mathbf{A} creció más rápidamente que el tamaño de esta matriz. Cabe reseñar que este estudio comparó las descomposiciones calculadas para registros de MEG reales desde un punto de vista global en lugar de comparar componentes manualmente (Zavala Fernández et al., 2006). No obstante, las fuentes reales son desconocidas (James and Hesse, 2005), así que resulta imposible conocer a ciencia cierta la verdadera calidad de la separación BSS (Escudero et al., 2009d).

La evaluación objetiva del rechazo de artefactos empleó medidas especialmente diseñadas para cada tipo de artefacto, evitando el proceso relativamente subjetivo, laborioso y tedioso de identificar manualmente las componentes (Rong and Contreras-Vidal, 2006). Gracias a estas métricas, se pudieron evaluar automáticamente diversas combinaciones de algoritmo de BSS, longitud de época y métrica de detección de artefactos.

La evaluación del rechazo del CA se efectuó gracias al hecho de que la actividad cardíaca es la única actividad sincrónica, junto con el PLN, en los registros (Escudero et al., 2007b). Por otro lado, los OAs no aparecen siempre en las señales de MEG. Es por esta razón que su reducción es más difícil de medir. No obstante, se consideraron características temporales, como el número de picos por encima de un cierto umbral de amplitud (Mantini et al., 2008), y espectrales de las señales, estudiando las frecuencias más bajas de la señal (Escudero et al., 2007b). Por último, la reducción en el PLN pudo medirse fácilmente en el dominio espectral (Escudero et al., 2007b; Iriarte et al., 2003). Con estos criterios y medidas, pudo recopilarse información cuantitativa acerca de la calidad del rechazo de artefactos. La Tabla B.3 detalla algunas de las combinaciones más exitosas en la eliminación de cada tipo de contaminación. En cada caso, los datos en la Tabla B.3 están expresados como el ratio de las métricas correspondientes después y antes del rechazo de artefactos. De este modo, es posible proporcionar una idea del tipo de metodologías más adecuadas para cada caso.

Tabla B.3: Resumen de los cambios (expresados como ratios) en las métricas de evaluación del rechazo de artefactos. Ratios menores indican mejores rechazos de artefactos.

Ratios para las métricas del artefacto cardíaco				
Métrica	Época	Tipo de BSS	A_{PtP}	RMS
–	10 s	cBSS	0.1335	0.1963
Skew	60 s	eInfoMax	0.1547	0.2359
VarSc	10 s	FastICA	0.1572	0.2513
KrE	90 s	eInfoMax	0.1354	0.2258
H_{Sh}	60 s	eInfoMax	0.1238	0.2132
$H_{Ré}$	60 s	eInfoMax	0.1242	0.2131
Ratios para las métricas del artefacto ocular				
Métrica	Época	Tipo de BSS	Número de picos	Diferencia espectral
P_{EYES}	60 s	SOBI	0.6158	0.4137
P_{LF}	90 s	AMUSE	0.1589	0.4357
KrE	90 s	eInfoMax	0.4221	0.8165
$H_{Ré}$	20 s	SOBI	0.2833	0.5005
ApEn	20 s	SOBI	0.2848	0.4011
Ratios de potencia entre 49.5 Hz to 50.5 Hz para evaluar el ruido eléctrico				
Métrica	Época	Tipo de BSS	Total	Normalizada
P_{50Hz}	60 s	AMUSE	0.2039	0.1850
P_{50Hz}	60 s	JADE	0.2406	0.2297
ApEn	60 s	JADE	0.2551	0.3178

Los algoritmos de BSS AMUSE, SOBI, eInfoMax y FastICA fueron aplicados para descomponer épocas MEG de 36 pacientes con AD y 26 sujetos de control de edad avanzada. Es destacable que la población se dividió de manera aleatoria en conjuntos de entrenamiento y validación para evitar la optimización de los parámetros incluidos en la metodologías sobre toda la base de datos (Escudero et al., 2009b). Mientras AMUSE proporcionó un orden para las componentes de manera intrínseca, las componentes SOBI, eInfoMax y FastICA tuvieron que ser ordenadas (Escudero et al., 2009c). Para ello se emplearon dos criterios: MF y curtosis. Posteriormente, cada componente se caracterizó con características de señal espectrales y no lineales (Escudero

et al., 2009b).

Tras comparar las componentes, pudieron evaluarse las mejoras en la separación de sujetos reportadas por este tipo de metodología. En los mejores casos, la BSS y el procedimiento de selección de componentes aumentó la precisión entre un 12.9 % y un 22.6 % para todas las características salvo MF, en cuyo caso la precisión no se modificó. Sin embargo, las AUCs de todas las características mejoraron entre 0.023 y 0.227 (para la MF y la SampEn, respectivamente). Un resumen de los mejores preprocesados en términos de precisión y AUC aparece en Tabla B.4.

Los análisis acerca de la extracción adaptativa de ritmos cerebrales se centró en las bandas espectrales típicamente consideradas en la Literatura: δ (1 Hz – 4 Hz), θ (4 Hz – 8 Hz), α (8 Hz – 13 Hz) y β (13 Hz – 30 Hz). Sin embargo, no hubo la necesidad de realizar un filtrado espectral a priori, evitando tener que definir frecuencias de corte. Esto se debe a que las IMFs tienen características espectrales bien definidas (Huang et al., 1998). Por este motivo, puede considerarse en los análisis el contenido espectral de cada IMF, el cual se caracterizó con los valores de frecuencia instantánea promedio. Posteriormente, sólo las IMFs cuyo contenido espectral se encontraba en la banda espectral de interés (1 Hz – 30 Hz) fueron introducidas en el procedimiento de clúster que proporcionó las referencias para el cBSS.

El análisis estadístico sugirió que la AD afecta el patrón de conectividad de la actividad electromagnética cerebral (Dauwels et al., 2010; Jeong, 2004; Stam, 2010), aunque los resultados no fueron significativos. Sin embargo, el análisis de la capacidad de estos resultados para clasificar los pacientes con AD frente a sujetos de control proporcionó unos resultados más que notables, alcanzado un 96.8 % de precisión, el cual puede compararse favorablemente con otros estudios citados en la reciente revisión del estado del arte publicada por Stam (2010).

Deben tenerse en cuenta diversas limitaciones de los estudios realizados en esta Tesis Doctoral. Las principales se deben al relativamente reducido tamaño de la base de datos analizada, lo que limita los análisis realizados y las conclusiones extraídas. Además, por la propia naturaleza de los análisis, se desconocen las componentes reales que dieron lugar a los registros de MEG.

Tabla B.4: Comparación de los valores de precisión y AUC alcanzados mediante diversos preprocesados con los valores obtenidos a partir de las señales de MEG sin preprocesar.

MF				
Tipo	Componentes seleccionadas	Precisión	AUC	
Sin preprocesado	–	77.4 %	0.855	
Filtrado en α	–	77.4 %	0.872	
AMUSE	30	77.4 %	0.878	
SOBI - MF	30	77.4 %	0.876	
SpecEn				
Tipo	Componentes seleccionadas	Precisión	AUC	
Sin preprocesado	–	61.3 %	0.727	
Filtrado en β	–	58.1 %	0.560	
SOBI - MF	30	74.2 %	0.782	
AMUSE	30	71.0 %	0.786	
LZC				
Tipo	Componentes seleccionadas	Precisión	AUC	
Sin preprocesado	–	61.3 %	0.786	
Filtrado en β	–	64.5 %	0.705	
AMUSE	15	74.2 %	0.838	
AMUSE	30	67.7 %	0.838	
SampEn				
Tipo	Componentes seleccionadas	Precisión	AUC	
Sin preprocesado	–	58.1 %	0.645	
Filtrado en β	–	61.3 %	0.645	
AMUSE	30	80.7 %	0.872	
AMUSE	15	80.7 %	0.863	

Por ello, la evaluación del rendimiento de las técnicas de BSS no es simple y las descomposiciones no pueden validarse por completo (James and Hesse, 2005; Vigário and Oja, 2008).

Apéndice C

Conclusiones en Español

Contents

C.1	Introducción	330
C.2	Conclusiones de la Tesis Doctoral	332
C.3	Contribuciones originales al estado del arte . . .	336
C.4	Futuras líneas de investigación	338

Este Apéndice constituye la traducción al español del Capítulo 9 donde se recogen las principales conclusiones extraídas de esta Tesis Doctoral. Además de éstas, este Apéndice detalla las principales contribuciones al estado del arte realizadas en esta Tesis Doctoral y sugiere posibles líneas futuras de investigación.

C.1 Introducción

El MEG recoge los campos magnéticos cerebrales originados por las corrientes eléctricas de las neuronas. Esta técnica posee ciertas ventajas sobre otros métodos de neuro-imagen, como son su elevada resolución espacial y temporal o el hecho de que el MEG no depende de ningún punto de referencia (Hämäläinen et al., 1993; Hari, 2004). Sin embargo, el análisis del MEG no resulta siempre sencillo, por lo que es deseable desarrollar herramientas y métodos adecuados para ayudar en esta tarea. Además, el MEG podría ofrecer información relevante acerca de diversos estados y enfermedades cerebrales, incluyendo la causa más frecuente de demencia en Occidente: la AD (Hornero et al., 2009b; Stam, 2005, 2010).

Es por todos conocido que la AD causa una pérdida progresiva de memoria. Eventualmente, esta demencia termina afectando a todas las funciones cognitivas del paciente (Blennow et al., 2006; Cummings, 2004; Nestor et al., 2004). Debido a su elevada prevalencia y a la pesada carga que esta demencia impone sobre los cuidadores y familiares de los pacientes, la AD causa considerables gastos y problemas sociales (Ferri et al., 2006). Cabe considerar que el coste del diagnóstico clínico de la AD es bajo en comparación con los costes asociados con el tratamiento de esta demencia y el cuidado de los enfermos (Lahiri et al., 2002). Por ello, y teniendo en cuenta que un diagnóstico definitivo de la AD sólo puede conseguirse mediante necropsia, resulta evidente la importancia de desarrollar técnicas que ayuden en el diagnóstico y caracterización de esta enfermedad. Esto podría repercutir en una reducción de los costes producidos por esta patología (Lahiri et al., 2002).

El término BSS denota un conjunto de técnicas estadísticas útiles para el procesado, entre otras, de señales biomédicas (James and Hesse, 2005; Vigário

and Oja, 2008). Estas herramientas generan transformaciones de registros multidimensionales sin asumir una base para proyectar los datos a priori y permiten revelar las componentes internas que dieron lugar a señales observadas a través de un proceso de mezcla que se asume desconocido (Cichocki and Amari, 2002; Hyvärinen et al., 2001). Hoy en día, se está generalizando su uso en el estudio de datos biomédicos y, de forma más particular, en el procesado de señales cerebrales (James and Hesse, 2005; Vigário and Oja, 2008).

Teniendo en cuenta estos tres conceptos principales sobre los que versa esta Tesis Doctoral (MEG, AD y BSS), los estudios desarrollados en ella se han basado en las siguientes *hipótesis*:

1. Que la comparación de las descomposiciones estimadas con varias técnicas de BSS para los mismos registros de MEG podría ofrecer información relevante acerca de las relaciones y similitudes existentes entre aquellas técnicas.
2. Que el uso conjunto de la BSS y métricas especialmente diseñadas para detectar artefactos en las componentes BSS podría ayudar a eliminar automáticamente la contaminación por CA, OA y PLN presente en la actividad basal de MEG.
3. Que un preprocesado adecuadamente diseñado basado en BSS podría ser útil para distinguir características (MF, SpecEn, LZC y SampEn) calculadas a partir de los registros de MEG de pacientes con AD de aquellas computadas para los registros de los sujetos de control.
4. Que una metodología basada en EMD y cBSS podría proporcionar una extracción adaptativa de ritmos cerebrales sobre diversas regiones del cuero cabelludo con información relevante acerca de la conectividad de pacientes con la AD en comparación sujetos de control.

Viendo en perspectiva los resultados y la discusión de esta Tesis Doctoral, estas hipótesis parecen ser correctas, lo que corrobora la idoneidad de la

metodología empleada. En consecuencia, puede afirmarse que *se ha conseguido el principal objetivo de esta Tesis Doctoral*. Éste no era otro que **aplicar técnicas de BSS para ayudar en el análisis e interpretación de actividad basal de MEG registrada en pacientes con la AD y sujetos de control de edad avanzada**. Los resultados han resaltado la idoneidad de las técnicas BSS para procesar señales de MEG. Además, la utilidad de estas técnicas no se limita únicamente a los marcos metodológicos en donde podrían tener una aplicación más obvia (por ejemplo, el uso de la BSS en el rechazo de artefactos en señales cerebrales) sino que también son útiles para estudiar la actividad cerebral desde puntos de vista más novedosos como el uso de la BSS como un preprocesado para resaltar características de las señales relacionadas con la AD o como un medio de realizar una extracción adaptativa de ritmos cerebrales sobre diversas agrupaciones de canales.

C.2 Conclusiones de la Tesis Doctoral

Con la metodología presentada en el Capítulo 6 se han procesado registros de MEG adquiridos de 36 pacientes con un diagnóstico de AD probable y 26 sujetos sanos de edad avanzada. La aplicación de tal metodología perseguía la consecución de *cuatro estudios principales* acerca de la aplicación de la BSS a actividad basal de MEG:

1. La comparación de las descomposiciones estimadas mediante diversas técnicas de BSS para registros reales de MEG ayudó a identificar similitudes entre los métodos de BSS y su evolución con el número de componentes a estimar y la longitud de las épocas de señal analizadas.
2. La evaluación cuantitativa del rechazo del CA, OA y PLN basado en técnicas de BSS resultó útil para señalar las combinaciones de método de BSS, longitud de época y métrica de detección de artefactos que proporcionaban una mayor atenuación de estas actividades indeseadas en la actividad basal de MEG.
3. El desarrollo, y posterior evaluación, de un procedimiento de BSS y

selección de componentes significativas mostró que este tipo de filtrado espacial de los datos era capaz de aumentar la precisión de la clasificación y los valores de AUC de características de señal calculadas para los MEGs de pacientes de AD y sujetos de control.

4. El uso de un procesado adaptativo para extraer actividad rítmica cerebral sobre diversas regiones del cuero cabelludo proporcionó información acerca de las alteraciones que la AD causa en la sincronización de las señales cerebrales y reveló la posible utilidad de este tipo de procesado en la clasificación de pacientes con la AD versus controles de edad avanzada.

Todos estos estudios resultan alentadores en el sentido de que sugieren que **la BSS es una herramienta útil en el procesado de actividad basal de MEG y en la identificación y caracterización de la AD. En consecuencia, la BSS puede representar un conjunto de técnicas importantes en el estudio de este tipo de registros biomédicos.** Además de esta conclusión general, las siguientes conclusiones particulares pueden derivarse también de esta Tesis Doctoral:

1. **El procedimiento basado en un FA, junto con el criterio MDL de selección del orden del modelo, resultó ser más adecuado para preprocesar los registros de MEG que PCA.** La evaluación de procedimiento basado en FA y el criterio MDL (Cao et al., 2003; Ikeda and Toyama, 2000) en los datos sintéticos sugirió que este esquema era el más apropiado para realizar un acondicionamiento de las señales previo a BSS. Esto se debe a que otros criterios comúnmente empleados, los cuales se basan en PCA, tendían a sufrir de sobreajuste (Escudero et al., 2007b). Además, la utilidad de esta metodología fue corroborada por el análisis de las señales reales.
2. **Los algoritmos AMUSE–SOBI, seguidos por el par JADE–FastICA, ofrecieron las descomposiciones de registros reales de actividad basal de MEG más similares.** La evaluación de la consistencia de las separaciones calculadas con diversas técnicas de BSS

para señales reales de MEG (Escudero et al., 2009d) mostró que los pares de algoritmos más consistentes fueron AMUSE–SOBI y JADE–FastICA. Asimismo, los pares AMUSE–JADE y SOBI–JADE ofrecieron las separaciones más diferentes. Por último, cabe reseñar que el nivel global de similitud aumentó conforme se descompusieron señales de mayor longitud.

3. **La cuantificación de la reducción en los artefactos presentes en la actividad basal de MEG destacó la utilidad de varias combinaciones de técnicas de BSS y métricas específicas de detección de artefactos.** La medida de la reducción de la intensidad promedio del CA, OA y PLN a lo largo de todo el conjunto de datos disponible destacó las siguientes estrategias para reducir los artefactos:

- Por su sencillez, se consideró el procedimiento de cBSS (James and Gibson, 2003) como el más adecuado para eliminar el CA. Es más, su rendimiento resultó ser equiparable al de las mejores combinaciones de técnicas basadas en BSS.
- La reducción en los OAs fue la más difícil de evaluar puesto que este artefacto aparece de forma intermitente en los registros y su energía no se encuentra tan localizada en el espectro como en el caso del PLN. No obstante, la evaluación indicó que la combinación de métricas de detección de artefactos como ApEn, P_{EYES} y, especialmente, P_{LF} (Escudero et al., 2007b) con métodos SOS-BSS podría ser útil para reducir este artefacto.
- El PLN pudo reducirse sustancialmente sin alterar el espectro de la señal cerebral por medio de la métrica específica P_{50Hz} (Escudero et al., 2007b) y AMUSE.

4. **Un preprocesado con SOS-BSS y un procedimiento de selección de componentes pudo mejorar la separación entre características de la señal de MEG de pacientes con AD y sujetos de control.** Los criterios de ordenación resultaron esenciales para comparar de forma directa las componentes de BSS, permitiendo de este

modo desarrollar las metodologías de preprocesado con BSS (Escudero et al., 2008d, 2009b,c). La evaluación de la mejora en la separación entre características espectrales y no lineales de los MEGs de pacientes con la AD y controles reveló aumentos en los valores de AUC de entre 0.023 y 0.227 y mejoras en la precisión de hasta un 22.6 % en comparación con las señales “crudas”. Los algoritmos más adecuados para este fin resultaron ser AMUSE y SOBI con una ordenación basada en la MF.

5. **La extracción adaptativa de ritmos cerebrales sobre diversas regiones del cuero cabelludo con un procedimiento adaptativo basado en una EMD y una cBSS ofreció información útil acerca del patrón de conectividad en la AD y, especialmente, y la clasificación de sujetos.** La metodología adaptativa introducida para extraer actividad rítmica cerebral sobre diversas regiones del cuero cabelludo se basó en una EMD (Huang et al., 1998), un procedimiento clúster de k -medias (Hartigan and Wong, 1979) y una cBSS (Huang and Mi, 2007). El análisis estadístico sugirió que la AD afecta a la conectividad entre regiones, a pesar de que los resultados no fueron estadísticamente significativos. Por el contrario, un análisis de la clasificación con una validación dejando uno fuera empleando los valores de $\text{Coh}(f)$ permitió clasificar los pacientes con la AD frente a los controles con una precisión del 96.8 %.

Por último, cabe reseñar que una de las principales ventajas de la BSS en estos tipos de aplicaciones reside en el hecho de que la BSS representa un método de descomposición de señales que no necesita apenas información acerca de las componentes a estimar o del proceso de mezcla al que fueron sujetas (James and Hesse, 2005; Vigário and Oja, 2008). Asumiendo unas pocas hipótesis generales acerca de las señales, estas técnicas son capaces de estimar un conjunto de componentes internas que pueden proporcionar novedosa información para examinar los datos bajo estudio (James and Hesse, 2005; Onton et al., 2006; Vigário and Oja, 2008).

En resumen, esta Tesis Doctoral ha examinado algunas aplicaciones de técnicas de BSS a registros de MEG, prestando especial atención a la AD. Uno de los objetivos generales fue evaluar cuantitativamente las ventajas y beneficios obtenidos gracias al uso de estos métodos. Debe tenerse en cuenta que interesantes revisiones acerca del estado del arte sobre la utilización de la BSS en el estudio de señales cerebrales han sido publicadas recientemente (James and Hesse, 2005; Onton et al., 2006; Vigário and Oja, 2008). Sin embargo, el posible rango de aplicaciones de las técnicas de BSS es tan extenso y evoluciona tan rápidamente que es difícil recopilar la mayor parte de información disponible al respecto en un único documento.

C.3 Contribuciones originales al estado del arte

La principal contribución de esta Tesis Doctoral es la evaluación del rendimiento, en términos de los objetivos definidos previamente, de diversas técnicas de BSS en cuatro aplicaciones diferentes. Para realizar estas investigaciones, se han propuesto diversas metodologías. Es más, en algunos casos, esta Tesis Doctoral representa la primera aplicación de los correspondientes preprocesados, métricas o técnicas de evaluación a la actividad basal de MEG.

Las contribuciones originales al estado del arte aportadas por esta Tesis Doctoral se detallan a continuación. Además, las *publicaciones científicas* generadas como consecuencia de estas investigaciones se encuentran listadas en el Apéndice A.

1. La evaluación del preprocesado basado en FA mediante datos sintéticos permitió comprobar lo adecuado de este procedimiento como un paso previo a BSS en contraposición a otras alternativas basadas en PCA. La evaluación de estos tipos de preprocesado con señales sintéticas diversos valores de SNR representa, hasta lo que conocemos, la primera evaluación exhaustiva de preprocesados basados en FA y PCA en este contexto para señales cerebrales (Escudero et al., 2007b).

2. Las comparaciones de las descomposiciones estimadas mediante diversos algoritmos de BSS para el mismo conjunto de señales de MEG reales reportaron información acerca de las similitudes y diferencias de estas técnicas. Hasta lo que sabemos, éste es el primer estudio acerca de la consistencia (desde un punto de vista global) de descomposiciones BSS de señales cerebrales reales (Escudero et al., 2009d).
3. La evaluación objetiva de la reducción en los artefactos que contaminan los registros de MEG representa una parte esencial en la aplicación de rechazo de artefactos. Hasta donde tenemos conocimiento, este estudio representa la primera evaluación de la eliminación del CA con BSS medido en términos del complejo QRS promedio (Escudero et al., 2007b). Además, se ha propuesto la evaluación de la eliminación de artefactos en MEG considerando de forma conjunta la totalidad de las señales en la base de datos para obtener medidas de la atenuación en lugar de considerar ejemplos concretos o atender a señales individuales (Escudero et al., 2007b). En lo que respecta a las métricas de detección de artefactos, este estudio propuso las métricas P_{EYES} , P_{LF} y P_{50Hz} y aplicó Skew a señales de MEG (Escudero et al., 2006c,d, 2007b, 2008c).
4. Este estudio ha evaluado el uso de técnicas de BSS como un preprocesado para realzar características de señal relacionadas con la AD. Hasta donde sabemos, este estudio constituye la primera aplicación de esta metodología a señales de MEG (Escudero et al., 2008d), incluyendo el estudio de características no lineales así como espectrales (Escudero et al., 2009b). Asimismo, este trabajo constituye la primera introducción de otros algoritmos de BSS (o ICA) distintos a AMUSE en este tipo de procesado (Escudero et al., 2009c).
5. Hasta lo que conocemos, la metodología desarrollada para extraer ritmos cerebrales sobre diversas regiones del cuero cabelludo de manera adaptativa es nueva. A pesar de que estudios previos (ver, por ejemplo, Stam et al., 2006) ya habían propuesto previamente agrupar canales en función de su localización espacial, ésta es la primera vez que la EMD

se ha aplicado a señales cerebrales con el fin de caracterizar la AD y, en particular, a señales de MEG para extraer referencias adecuadas para emplearlas en un cBSS cuyos resultados se emplearán en un análisis de conectividad.

Además, cabe reseñar que el doctorando también ha trabajado en la caracterización de registros cerebrales (EEG y MEG) prestando especial atención al análisis no lineal de señales adquiridas en pacientes con la AD. Esta labor le ha permitido colaborar en diversas investigaciones las cuales no están directamente relacionadas con la BSS (Abásolo et al., 2006a,b,d, 2007a,b, 2008a,b,c, 2009; Escudero et al., 2006b, 2007a, 2008a,b; Fernández et al., 2008, 2009; Gómez et al., 2006b, 2007a,b,c, 2008a,b, 2009a; Hornero et al., 2008, 2009a,b; Poza et al., 2007a, 2008a,b, 2009). *Especialmente relevantes son los artículos* acerca del análisis de señales de EEG en pacientes con la AD mediante la entropía multiescala EEG (Escudero et al., 2006a) y la interpretación de la ‘tasa de descenso de la función de auto-información mutua’ en términos de conceptos sencillos de procesado de señal (Escudero et al., 2009a).

C.4 Futuras líneas de investigación

En esta Tesis Doctoral, se han estudiado registros de actividad basal de MEG adquiridos de 36 pacientes con un diagnóstico de AD probable y 26 sujetos de control haciendo uso de técnicas de BSS. A pesar del esfuerzo realizado para realizar un estudio lo más detallado posible acerca de los temas abordados en ella, existen cuestiones interesantes directamente relacionadas con esta Tesis Doctoral que no han sido tratadas aún. Algunas de estas líneas futuras de investigación están relacionadas con las limitaciones de la Tesis Doctoral reconocidas en la Sección 8.7.

El análisis de la consistencia y del rechazo de artefactos se llevó a cabo empleando los registros de los 26 sujetos de control. Las señales adquiridas de los 36 pacientes con un diagnóstico de AD probable también fueron incluidas en los estudios acerca de la mejora de la clasificación debida al preprocesado

con BSS y en la medida de la conectividad de los ritmos cerebrales. Sin embargo, debería aumentarse el tamaño de la base de datos con el fin de obtener conclusiones más fiables y significativas en el futuro. Asimismo, otras patologías pueden causar alteraciones en la actividad de EEG y MEG similares a las inducidas por la AD (Stam, 2005, 2010). Por ello, sería deseable incluir en la base de datos registros de pacientes con otras enfermedades neurodegenerativas. De este modo, sería posible llevar a cabo estudios más complejos cuyo objetivo fuera desarrollar estrategias útiles para elaborar un diagnóstico diferencial. En este sentido, merece destacarse la comparativa de los registros de MEG adquiridos de pacientes con demencia con aquellos registrados en pacientes con MCI. Esto se debe al hecho de que el MCI puede considerarse como un estado previo a la demencia. De hecho, algunos estudios ya han analizado actividad MEG en pacientes con MCI y sujetos de control incluidos en la base de datos analizada en esta Tesis Doctoral (Fernández et al., 2006; Gómez et al., 2009c).

Con respecto al rechazo de artefactos, debería comprobarse cómo éste afecta a los registros de MEG adquiridos de pacientes con enfermedades como la AD. Esto se debe a que el rechazo de artefactos se evaluó empleando únicamente el grupo de control. Cabe esperar que los mismos métodos de rechazo de artefactos puedan aplicarse a señales registradas en pacientes, pero esta hipótesis debería corroborarse. En este sentido, son remarcables las investigaciones realizadas por Melissant et al. (2005) y Vialatte et al. (2009), las cuales mostraron que un rechazo de artefactos con BSS no causa alteraciones indeseadas en los registros cerebrales y que puede mejorar la clasificación entre grupos de sujetos en comparación con aquella basada en el análisis de señales ruidosas con artefactos. Por el contrario, Castellanos and Makarov (2006) sugirieron que, si el rechazo de artefactos no se realiza con el cuidado necesario, podría alterar el patrón de conectividad del EEG.

Otra posible línea de investigación es la aplicación de otros métodos de descomposición de señales multidimensionales que tampoco definen una base para proyectar las señales a priori. Por ejemplo, métodos como la NMF podrían ofrecer información importante acerca de las señales cerebrales, incluso en la AD (Chen et al., 2006; Cichocki et al., 2008). Por otro lado,

diversos tipos de restricciones (por ejemplo, espaciales o espectrales; [James and Hesse, 2005](#)) podrían emplearse en lugar de las referencias temporales necesarias en el cBSS aplicado en esta Tesis Doctoral.

Otra posible línea de investigación futura es la aplicación de las técnicas incluidas en esta Tesis Doctoral a otras señales fisiológicas. Por ejemplo, a pesar de que los métodos se diseñaron específicamente para registros de MEG, su aplicación a registros de EEG debería ser prácticamente directa.

Aunque la selección de los parámetros de cada método (bandas espectrales y valores de los parámetros de entrada a los algoritmos BSS y características de señal) se basó en la revisión realizada del estado del arte, es posible que variaciones en estos valores puedan alterar los resultados del procesado. Es más, otras medidas de características de señal ([Hornero et al., 2009b](#); [Stam, 2005](#)) o conectividad ([Dauwels et al., 2010](#); [Pereda et al., 2005](#)) podrían ser también útiles en este tipo de análisis y deberían ser evaluados.

En resumen, esta Tesis Doctoral ha evaluado algunos de los métodos más comunes de BSS en el marco de diversas aplicaciones, incluyendo la comparación de los resultados de diversos algoritmos, el rechazo de artefactos, el desarrollo de un preprocesado para mejorar la clasificación de sujetos y la extracción adaptativa de ritmos cerebrales. La BSS es una parte esencial de esta metodología. Los resultados indican que la BSS puede ser extremadamente útil para ayudar en el análisis y estudio de la actividad basal de MEG.

Acronyms

$A\beta$	Amyloid β peptide.
A_{PtP}	Peak-to-Peak Amplitude.
AD	Alzheimer's Disease.
AMUSE	Algorithm for Multiple Unknown Signals Ex- traction.
ANOVA	ANalysis Of VAriance.
ApEn	Approximate Entropy.
APOE	Apolipoprotein E.
APP	Amyloid Precursor Protein.
ASCII	American Standard Code for Information In- terchange.
AUC	Area Under the ROC Curve.
BSS	Blind Source Separation.
CA	Cardiac Artefact.
CAT	Computerised Axial Tomography.
cBSS	Constrained Blind Source Separation.
CJD	Creutzfeldt-Jakob Disease.
CLU	Clusterin.
CNS	Central Nervous System.
$\text{Coh}(f)$	Spectral Coherence.
CR1	Complement Component (3b/4b) Receptor 1.
CSF	CerebroSpinal Fluid.

CSP	Common Spatial Patterns.
ECG	Electrocardiogram.
EEG	Electroencephalogram.
eInfoMax	Lee & Sejnowski's Extended Information Maximisation Algorithm.
EMD	Empirical Mode Decomposition.
EOG	Electrooculogram.
EPSP	Excitatory Post-Synaptic Potential.
FA	Factor Analysis.
FastICA	Hyvärinen–Oja's FastICA Algorithm.
FIR	Finite Impulse Response.
fMRI	Functional Magnetic Resonance Imaging.
GDA/FAST	Global Deterioration Scale / Functional Assessment Staging.
$H_{Ré}$	Rényi Entropy.
H_{Sh}	Shannon Entropy.
HOS	Higher-Order Statistics.
HPI	Head-Position Indicator.
ICA	Independent Component Analysis.
IMF	Intrinsic Mode Function.
InfoMax	Bell & Sejnowski's Information Maximisation Algorithm.
IPSP	Inhibitory Post-Synaptic Potential.
JADE	Joint Approximate Diagonalization of Eigenmatrices.
KrE	Kurtosis Excess.

LDA	Linear Discriminant Analysis.
LeftA	Left Anterior Region.
LeftC	Left Central Region.
LeftL	Left Lateral Region.
LeftP	Left Posterior Region.
LZC	Lempel-Ziv Complexity.
MCI	Mild Cognitive Impairment.
MDL	Minimum Description Length.
MEG	Magnetoencephalogram.
MF	Median Frequency.
MMSE	Mini-Mental State Examination.
MRI	Magnetic Resonance Imaging.
MRS	Magnetic Resonance Spectroscopy.
NINCDS-ADRDA	National Institute of Neurological and Communicative Disorders and Stroke – Alzheimer’s Disease and Related Disorders Association.
NMDA	<i>N</i> -methyl-D-aspartate.
NMF	Non-negative Matrix Factorisation.
OA	Ocular Artefact.
$P_{50\text{Hz}}$	Power at the Line Frequency.
P_{EYES}	Power near the Eyes.
P_{LF}	Power in Low Frequencies.
PCA	Principal Component Analysis.
PET	Positron Emission Tomography.
PICALM	Phosphatidylinositol-binding Clathrin Assembly Protein.
PLN	Power Line Noise.

PNS	Peripheral Nervous System.
PSD	Power Spectral Density.
PSD _n	Normalised Power Spectral Density.
PSEN1	Presenilin 1.
PSEN2	Presenilin 2.
RightA	Right Anterior Region.
RightC	Right Central Region.
RightL	Right Lateral Region.
RightP	Right Posterior Region.
RMS	Root Mean Square.
ROC	Receiver Operating Characteristic.
SampEn	Sample Entropy.
SCI–JCR	Science Citation Index – Journal Citation Reports®.
SD	Standard Deviation.
Skew	Skewness.
SNR	Signal-to-Noise Ratio.
SOBI	Second-Order Blind Identification.
SORL1	SOrtilin-Related receptor.
SOS	Second-Order Statistics.
SpecEn	Spectral Entropy.
SPECT	Single Photon Emission Computed Tomography.
SQUID	Superconductive QUantum Interference Device.
VaD	Vascular Dementia.
VarSc	Variance of the Scalp Distribution.
WT	Wavelet Transform.

Bibliography

- D. Abásolo, J. Escudero, R. Hornero, P. Espino, and C. Gómez. Análisis de la actividad base del EEG con la dimensión fractal. Aplicación a la enfermedad de Alzheimer. In *Libro de Actas del XXIV Congreso Anual de la Sociedad Española de Ingeniería Biomédica*, pages 347–350, 2006a.
- D. Abásolo, R. Hornero, J. Escudero, C. Gómez, M. García, and M. López. Approximate entropy and mutual information analysis of the electroencephalogram in Alzheimer’s disease patients. In *Proceedings of the 3rd Institution of Engineering and Technology International Conference on Advances in Medical, Signal and Information Processing*, 2006b.
- D. Abásolo, R. Hornero, P. Espino, D. Álvarez, and J. Poza. Entropy analysis of EEG background activity in Alzheimer’s disease patients. *Physiological Measurement*, 27(3):241–253, 2006c.
- D. Abásolo, R. Hornero, C. Gómez, J. Escudero, and M. López. Fluctuation analysis of the electroencephalogram background activity in patients with Alzheimer’s disease, 2006d.
- D. Abásolo, R. Hornero, C. Gómez, M. García, and M. López. Analysis of EEG background activity in Alzheimer’s disease patients with Lempel-Ziv complexity and central tendency measure. *Medical Engineering and Physics*, 28(4):315–322, 2006e.
- D. Abásolo, R. Hornero, and J. Escudero. Non-linear analysis of the EEG background activity in Alzheimer’s disease patients. In *Book of abstracts of the Third International Conference on Computational Intelligence in Medicine and Healthcare (CIMED 2007) - Workshop On Advances In Early Diagnosis And Care For Alzheimer’s Disease*, page 14, 2007a.

- D. Abásolo, R. Hornero, and J. Escudero. Non-linear analysis of the EEG background activity in Alzheimer's disease patients (Special Invited Session). In *Book of abstracts of the Third International Conference on Computational Intelligence in Medicine and Healthcare (CIMED 2007) - Workshop On Advances In Early Diagnosis And Care For Alzheimer's Disease - Special invited Session*, page 19, 2007b.
- D. Abásolo, J. Escudero, R. Hornero, P. Espino, and C. Gómez. Fractal dimension of the EEG in Alzheimer's disease. In N. Wickramasinghe and E. Geisler, editors, *Encyclopedia of Healthcare Information Systems*, volume II, pages 603–609. IGI Global, U.S.A., 2008a.
- D. Abásolo, J. Escudero, R. Hornero, C. Gómez, and P. Espino. Approximate entropy and auto mutual information analysis of the electroencephalogram in Alzheimer's disease patients. *Medical and Biological Engineering and Computing*, 46(10):1019–1028, 2008b.
- D. Abásolo, R. Hornero, J. Escudero, and P. Espino. A study on the possible usefulness of detrended fluctuation analysis of the electroencephalogram background activity in Alzheimer's disease. *IEEE Transactions on Biomedical Engineering*, 55(9):2171–2179, 2008c.
- D. Abásolo, R. Hornero, C. Gómez, J. Escudero, and P. Espino. Electroencephalogram background activity characterization with Detrended Moving Average in Alzheimer's disease patients. In *Proceedings of the 6th IEEE International symposium on Intelligent Signal Processing (WISP 2009)*, pages 211–215, 2009.
- I. Abatzoglou, P. Anninos, A. Adamopoulos, and M. Koukourakis. Nonlinear analysis of brain magnetoencephalographic activity in Alzheimer disease patients. *Acta Neurologica Belgica*, 107(2):34–39, 2007.
- M. Aboy, R. Hornero, D. Abásolo, and D. Álvarez. Interpretation of the Lempel-Ziv complexity measure in the context of biomedical signal analysis. *IEEE Transactions on Biomedical Engineering*, 53(11):2282–2288, 2006.
- A. Andrade, S. Nasuto, P. Kyberd, C. Sweeney-Reed, and F. Van Kanijn. EMG signal filtering based on Empirical Mode Decomposition. *Biomedical Signal Processing and Control*, 1(1):44–55, 2006.
- A. Antervo, R. Hari, T. Katila, T. Ryhänen, and M. Seppänen. Magnetic fields produced by eye blinking. *Electroencephalography and clinical Neurophysiology*, 61(4):247–253, 1985.

- G. Barbati, C. Porcaro, F. Zappasodi, P. Rossini, and F. Tecchio. Optimization of an independent component analysis approach for artifact identification and removal in magnetoencephalographic signals. *Clinical Neurophysiology*, 115(5):1220–1232, 2004.
- A. Bell and T. Sejnowski. An information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, 7(6):1129–1159, 1995.
- A. Belouchrani, K. Abed-Meraim, J. Cardoso, and E. Moulines. A blind source separation technique using second-order statistics. *IEEE Transactions on Signal Processing*, 45(2):434–444, 1997.
- H. Berendse, J. Verbunt, P. Scheltens, B. Van Dijk, and E. Jonkman. Magnetoencephalographic analysis of cortical activity in Alzheimer’s disease: a pilot study. *Clinical Neurophysiology*, 111(4):604–612, 2000.
- M. Blanco-Velasco, B. Weng, and K. Barner. ECG signal denoising and baseline wander correction based on the empirical mode decomposition. *Computers in Biology and Medicine*, 38(1):1–13, 2008.
- K. Blennow, de Leon M.J., and H. Zetterberg. Alzheimer’s disease. *The Lancet*, 368(9533):387–403, 2006.
- J. Bronzino. *The biomedical engineering handbook*. Springer, Boca Raton, FL, U.S.A., 3rd, edition, 2006.
- F. Campos Viola, J. Thorne, B. Edmonds, T. Schneider, T. Eichele, and S. Debener. Semi-automatic identification of independent components representing EEG artifact. *Clinical Neurophysiology*, 120(5):868–877, 2009.
- J. Cao, N. Murata, S. Amari, A. Cichocki, and T. Takeda. A robust approach to independent component analysis of signals with high-level noise measurements. *IEEE Transactions on Neural Networks*, 14(3):631–645, 2003.
- J. Cardoso and A. Souloumiac. Blind Beamforming for Non Gaussian Signals. *IEE proceedings. Part F. Radar and Signal Processing*, 140(6):362–370, 1993.
- J. Cardoso and A. Souloumiac. Jacobi angles for simultaneous diagonalization. *SIAM Journal on Matrix Analysis and Applications*, 17(1):161–164, 1996.

- J. Carr and J. Brown. *Introduction to biomedical equipment technology*. Prentice-Hall, New Jersey, USA, 4th edition, 2001.
- M. Carrasquillo, F. Zou, V. Pankratz, S. Wilcox, L. Ma, L. Walker, S. Younkin, C. Younkin, L. Younkin, G. Bisceglia, et al. Genetic variation in PCDH11X is associated with susceptibility to late-onset Alzheimer's disease. *Nature Genetics*, 41(2):192–198, 2009.
- N. Castellanos and V. Makarov. Recovering EEG brain signals: Artifact suppression with wavelet enhanced independent component analysis. *Journal of Neuroscience Methods*, 158(2):300–312, 2006.
- Z. Chen, A. Cichocki, and T. Rutkowski. Constrained non-negative matrix factorization method for EEG analysis in early detection of Alzheimer's disease. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, (ICASSP2006)*, pages 893–896, 2006.
- A. Cichocki and S. Amari. *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*. Wiley, Cichester, U.K., 2002.
- A. Cichocki, S. Shishkin, T. Musha, Z. Leonowicz, T. Asada, and T. Kurachi. EEG filtering based on blind source separation (BSS) for early detection of Alzheimer's disease. *Clinical Neurophysiology*, 116(3):729–737, 2005.
- A. Cichocki, R. Zdunek, and S. Amari. Nonnegative matrix and tensor factorization. *IEEE Signal Processing Magazine*, 25(1):142–145, 2008.
- A. Cichocki, S. Amari, K. Siwek, T. Tanaka, and A. Huy Phan. ICALAB for Signal Processing. website, August 2009. URL <http://www.bsp.brain.riken.jp/ICALAB/ICALABSignalProc/>. [online] <http://www.bsp.brain.riken.jp/ICALAB/ICALABSignalProc/>.
- A. Cohen. Biomedical signals: origin and dynamic characteristics; frequency-domain analysis. In J. Bronzino, editor, *The biomedical engineering handbook*, pages 1–1–1–22. Springer, Boca Raton, FL, U.S.A., 3rd, edition, 2006.
- D. Cohen. Magnetoencephalography: evidence of magnetic fields produced by alpha-rhythm currents. *Science*, 161(3843):784–786, 1968.

- D. Cohen. Magnetoencephalography: detection of the brain's electrical activity with a superconducting magnetometer. *Science*, 175(4022):664–666, 1972.
- P. Comon. Independent component analysis, A new concept? *Signal Processing*, 36(3):287–314, 1994.
- M. Crespo-Garcia, M. Atienza, and J. Cantero. Muscle artifact removal from human sleep EEG by using independent component analysis. *Annals of Biomedical Engineering*, 36(3):467–475, 2008.
- R. Croft and R. Barry. Removal of ocular artifact from the EEG: A review. *Neurophysiologie Clinique/Clinical Neurophysiology*, 30(1):5–19, 2000.
- J. Cummings. Drug therapy: Alzheimer's disease. *New England Journal of Medicine*, 351(1):56–67, 2004.
- J. Dammers, M. Schiek, F. Boers, C. Silex, M. Zvyagintsev, U. Pietrzyk, and K. Mathiak. Integration of amplitude and phase statistics for complete artifact removal in independent components of neuromagnetic recordings. *IEEE Transactions on Biomedical Engineering*, 55(10):2353–2362, 2008.
- G. Dassios, A. Fokas, and D. Hadjiloizi. On the complementarity of electroencephalography and magnetoencephalography. *Inverse Problems*, 23(6):2541–2549, 2007.
- J. Dauwels, F. Vialatte, T. Musha, and A. Cichocki. A comparative study of synchrony measures for the early diagnosis of Alzheimer's disease based on EEG. *Neuroimage*, 49(1):668–693, 2010.
- M. de LaCoste and C. White. The role of cortical connectivity in Alzheimer's disease pathogenesis: a review and model system. *Neurobiology of Aging*, 14(1):1–16, 1993.
- X. Delbeuck, M. Van der Linden, and F. Collette. Alzheimer's Disease as a Disconnection Syndrome? *Neuropsychology Review*, 13(2):79–92, 2003.
- A. Delorme and S. Makeig. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1):9–21, 2004.

- A. Delorme, T. Sejnowski, and S. Makeig. Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. *NeuroImage*, 34(4):1443–1449, 2007.
- R. Doody, J. Stevens, C. Beck, R. Dubinsky, J. Kaye, L. Gwyther, R. Mohs, L. Thal, P. Whitehouse, S. DeKosky, et al. Practice parameter: management of dementia (an evidence-based review). Report of the Quality Standards Subcommittee of the American Academy of Neurology. *Neurology*, 56(9):1154–1166, 2001.
- J. Escudero, D. Abásolo, R. Hornero, P. Espino, and M. López. Analysis of electroencephalograms in Alzheimer’s disease patients with multiscale entropy. *Physiological Measurement*, 27(11):1091–1106, 2006a.
- J. Escudero, D. Abásolo, R. Hornero, P. Espino, and M. López. Electroencephalograms multiscale entropy analysis of Alzheimer’s disease patients. In *Proceedings of the 3rd Institution of Engineering and Technology International Conference on Advances in Medical, Signal and Information Processing*, 2006b.
- J. Escudero, R. Hornero, D. Abásolo, J. Poza, A. Fernández, and M. López. Rejection of artifact sources in magnetoencephalogram background activity using independent component analysis. In *Proceedings of the 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 5282–5285, 2006c.
- J. Escudero, R. Hornero, D. Abásolo, J. Poza, A. Fernández, and M. López. Rechazo de artefactos cardiacos y oculares en magnetoencefalografía con análisis de componentes independientes. In *Libro de actas del XXI Simposium Nacional de la Unión Científica Internacional de Radio, URSI 2006*, pages 1271–1275, 2006d.
- J. Escudero, R. Hornero, J. Poza, A. Fernández, and M. López. Estudio de la actividad base del magnetoencefalograma en pacientes con la enfermedad de Alzheimer mediante separación ciega de fuentes y frecuencia media. In *Libro de Actas del XXIV Congreso Anual de la Sociedad Española de Ingeniería Biomédica*, pages 527–530, 2006e.
- J. Escudero, D. Abásolo, R. Hornero, P. Espino, and M. López. Reply to ” Comment on ’Analysis of electroencephalograms in Alzheimer’s disease patients with multiscale entropy’ ”. *Physiological Measurement*, 28(12):L3–L7, 2007a.

- J. Escudero, R. Hornero, D. Abásolo, A. Fernández, and M. López-Coronado. Artifact removal in magnetoencephalogram background activity with independent component analysis. *IEEE Transactions on Biomedical Engineering*, 54(11):1965–1973, 2007b.
- J. Escudero, R. Hornero, D. Abásolo, A. Fernández, and J. Poza. Magnetoencephalogram blind source separation and component selection procedure to improve the diagnosis of Alzheimer’s disease patients. In *Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 5437–5440, 2007c.
- J. Escudero, R. Hornero, D. Abásolo, C. Gómez, and A. Fernández. Separación ciega de fuentes para realzar características espectrales y no lineales del magnetoencefalograma en pacientes de la enfermedad de Alzheimer. In *Libro de Actas del XXV Congreso Anual de la Sociedad Española de Ingeniería Biomédica*, pages 222–225, 2007d.
- J. Escudero, R. Hornero, D. Abásolo, A. Fernández, and C. Gómez. Análisis complementario del magnetoencefalograma en pacientes con la enfermedad de Alzheimer mediante técnicas espectrales y no lineales. In *Libro de Actas del XXVI Congreso Anual de la Sociedad Española de Ingeniería Biomédica*, pages 162–165, 2008a.
- J. Escudero, R. Hornero, D. Abásolo, and M. López. On the application of the auto mutual application rate of decrease to biomedical signals. In *Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 2137–2140, 2008b.
- J. Escudero, R. Hornero, D. Abásolo, J. Poza, and A. Fernández. Applying Independent Component Analysis to the Artifact Detection Problem in Magnetoencephalogram Background Recordings. In N. Wickramasinghe and E. Geisler, editors, *Encyclopedia of Healthcare Information Systems*, volume I, pages 84–92. IGI Global, U.S.A., 2008c.
- J. Escudero, R. Hornero, J. Poza, D. Abásolo, and A. Fernández. Assessment of classification improvement in patients with Alzheimer’s disease based on magnetoencephalogram blind source separation. *Artificial Intelligence in Medicine*, 43(1):75–85, 2008d.
- J. Escudero, R. Hornero, and D. Abásolo. Interpretation of the auto-mutual information rate of decrease in the context of biomedical signal analysis. Application to electroencephalogram recordings. *Physiological Measurement*, 30(2):187–199, 2009a.

- J. Escudero, R. Hornero, D. Abásolo, and A. Fernández. Blind source separation to enhance spectral and non-linear features of magnetoencephalogram recordings. Application to Alzheimer's disease. *Medical Engineering and Physics*, 31(7):872–879, 2009b.
- J. Escudero, R. Hornero, D. Abásolo, and A. Fernández. Magnetoencephalogram preprocessing with blind source separation algorithms to improve the classification of Alzheimer's disease patients. In *Proceedings of The 6th International Workshop on Biosignal Interpretation*, pages 224–227, 2009c.
- J. Escudero, R. Hornero, D. Abásolo, J. Poza, and A. Fernández. Comparison of the decompositions estimated using five blind source separation algorithms for magnetoencephalogram background activity. In *Proceedings of The 6th International Workshop on Biosignal Interpretation*, pages 136–139, 2009d.
- M. Fatourehchi, A. Bashashati, R. Ward, and G. Birch. EMG and EOG artifacts in brain computer interface systems: a survey. *Clinical Neurophysiology*, 118(3):480–494, 2007.
- T. Fawcett. An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8):861–874, 2006.
- R. Ferenets, T. Lipping, A. Anier, V. Jäntti, S. Melto, and S. Hovilehto. Comparison of entropy and complexity measures for the assessment of depth of sedation. *IEEE Transactions on Biomedical Engineering*, 53(6):1067–1077, 2006.
- A. Fernández, R. Hornero, A. Mayo, J. Poza, P. Gil-Gregorio, and T. Ortiz. MEG spectral profile in Alzheimer's disease and mild cognitive impairment. *Clinical Neurophysiology*, 117(2):306–314, 2006.
- A. Fernández, J. Quintero, R. Hornero, C. Gómez, J. Escudero, P. Zuluaga, and T. Ortiz. Lempel-Ziv complexity analysis of spontaneous brain activity in attention-deficit hyperactivity disorder. In *Libro de Actas del XXVI Congreso Anual de la Sociedad Española de Ingeniería Biomédica*, pages 525–528, 2008.
- A. Fernández, J. Quintero, R. Hornero, P. Zuluaga, M. Navas, C. Gómez, J. Escudero, N. García-Campos, J. Biederman, and T. Ortiz. Complexity Analysis of Spontaneous Brain Activity in Attention-Deficit/Hyperactivity Disorder: Diagnostic Implications. *Biological Psychiatry*, 65(7):571–577, 2009.

- C. Ferri, M. Prince, C. Brayne, H. Brodaty, L. Fratiglioni, M. Ganguli, K. Hall, K. Hasegawa, H. Hendrie, Y. Huang, et al. Global prevalence of dementia: a Delphi consensus study. *The Lancet*, 366(9503): 2112–2117, 2006.
- S. Fitzgibbon, D. Powers, K. Pope, and C. Clark. Removal of EEG Noise and Artifact Using Blind Source Separation. *Journal of Clinical Neurophysiology*, 24(3):232–243, 2007.
- P. Flandrin, G. Rilling, and P. Gonçalvès. Empirical mode decomposition as a filter bank. *IEEE Signal Processing Letters*, 11(2):112–114, 2004.
- A. Flexer, H. Bauer, J. Pripfl, and G. Dorffner. Using ICA for removal of ocular artifacts in EEG recorded from blind subjects. *Neural Networks*, 18(7):998–1005, 2005.
- M. Folstein, S. Folstein, and P. McHugh. “Mini-mental state”. A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3):189–198, 1975.
- R. Frank and G. Frishkoff. Automated protocol for evaluation of electromagnetic component separation (APECS): Application of a framework for evaluating statistical methods of blink extraction from multichannel EEG. *Clinical Neurophysiology*, 118(1):80–97, 2007.
- H. Gävert, J. Hurri, J. Särelä, and A. Hyvärinen. FastICA Toolbox. website, August 2009. URL <http://www.cis.hut.fi/projects/ica/fastica/>. [online] <http://www.cis.hut.fi/projects/ica/fastica/>.
- C. Gómez, R. Hornero, D. Abásolo, A. Fernández, and M. López. Complexity analysis of the magnetoencephalogram background activity in Alzheimer’s disease patients. *Medical Engineering and Physics*, 28(9): 851–859, 2006a.
- C. Gómez, R. Hornero, A. Fernández, D. Abásolo, J. Escudero, and M. López. Magnetoencephalogram background activity analysis in Alzheimer’s disease patients using auto mutual information. In *Proceedings of the 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 6181–6184, 2006b.
- C. Gómez, R. Hornero, D. Abásolo, A. Fernández, and J. Escudero. Analysis of the magnetoencephalogram background activity in Alzheimer’s disease patients with auto-mutual information. *Computer Methods and Programs in Biomedicine*, 87(3):239–247, 2007a.

- C. Gómez, R. Hornero, D. Abásolo, A. Fernández, and J. Escudero. Analysis of MEG recordings from Alzheimer's disease patients with sample and multiscale entropies. In *Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 6183–6186, 2007b.
- C. Gómez, Á. Mediavilla, R. Hornero, A. Fernández, and J. Escudero. Estimación de la dimensión fractal de Higuchi en magnetoencefalogramas de enfermos de Alzheimer. In *Libro de Actas del XXV Congreso Anual de la Sociedad Española de Ingeniería Biomédica*, pages 168–171, 2007c.
- C. Gómez, R. Hornero, M. Á., A. Fernández, and J. Escudero. Análisis de la predictibilidad de registros MEG pertenecientes a enfermos de Alzheimer. In *Libro de Actas del XXVI Congreso Anual de la Sociedad Española de Ingeniería Biomédica*, pages 37–40, 2008a.
- C. Gómez, R. Hornero, D. Abásolo, A. Fernández, and J. Escudero. Regularity Analysis of the Magnetoencephalogram Background Activity in Alzheimer's Disease Patients Using Auto Mutual Information. In N. Wickramasinghe and E. Geisler, editors, *Encyclopedia of Healthcare Information Systems*, volume III, pages 1146–1152. IGI Global, U.S.A., 2008b.
- C. Gómez, R. Hornero, D. Abásolo, A. Fernández, and J. Escudero. Analysis of MEG background activity in Alzheimer's disease using nonlinear methods and ANFIS. *Annals of Biomedical Engineering*, 37(3):586–594, 2009a.
- C. Gómez, Á. Mediavilla, R. Hornero, D. Abásolo, and A. Fernández. Use of the Higuchi's fractal dimension for the analysis of MEG recordings from Alzheimer's disease patients. *Medical Engineering and Physics*, 31(3):306–313, 2009b.
- C. Gómez, C. Stam, R. Hornero, A. Fernández, and F. Maestú. Disturbed Beta Band Functional Connectivity in Patients With Mild Cognitive Impairment: An MEG Study. *IEEE Transactions on Biomedical Engineering*, 56(6), 2009c.
- A. Greco, N. Mammone, F. Morabito, and M. Versaci. Kurtosis, renyis entropy and independent component scalp maps for the automatic artifact rejection from eeg data. *International Journal of Signal Processing*, 2(4):240–244, 2006.

- L. Gwyther. *Caring for People With Alzheimer's Disease: A manual for Facility Staff*. American Health Care Association, 2nd edition, 2001.
- S. Halder, M. Bensch, J. Mellinger, M. Bogdan, A. Kubler, N. Birbaumer, and W. Rosenstiel. Online artifact removal for brain-computer interfaces using support vector machines and blind source separation. *Computational Intelligence and Neuroscience*, 2007:82069, 2007.
- M. Hämäläinen and R. Hari. Magnetoencephalographic (MEG) Characterization of Dynamic Brain Activation: Basic Principles and Methods of Data Collection and Source Analysis. In A. Toga and J. Mazziotta, editors, *Brain mapping: The methods*, chapter 10. Academic Press, 2nd, edition, 2004.
- M. Hämäläinen, R. Hari, R. Ilmoniemi, J. Knuutila, and O. Lounasmaa. Magnetoencephalography – theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics*, 65(2):413–497, 1993.
- R. Hari. Magnetoencephalography in clinical neurophysiological assessment of human cortical functions. In E. Niedermeyer and F. Lopes da Silva, editors, *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*, pages 1165–1197. Lippincott Williams and Wilkins, Philadelphia, U.S.A., 2004.
- R. Hari and N. Forss. Magnetoencephalography in the study of human somatosensory cortical processing. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 354(1387): 1145–1154, 1999.
- R. Hari, S. Levänen, and T. Raij. Timing of human cortical functions during cognition: role of MEG. *Trends in Cognitive Sciences*, 4(12):455–462, 2000.
- D. Harold, R. Abraham, P. Hollingworth, R. Sims, A. Gerrish, M. Hamshere, J. Pahwa, V. Moskvina, K. Dowzell, A. Williams, et al. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. *Nature Genetics*, 41(10):1088–1093, 2009.
- J. Hartigan and M. Wong. A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28(1):100–108, 1979.

- G. Henderson, E. Ifeachor, N. Hudson, C. Goh, N. Outram, S. Wimalaratna, C. Del Percio, and F. Vecchio. Development and assessment of methods for detecting dementia using the human electroencephalogram. *IEEE Transactions on Biomedical Engineering*, 53(8):1557–1568, 2006.
- C. Hesse. On Estimating the Signal Subspace Dimension of High-Density Multichannel Magnetoencephalogram Measurements. In *Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 6227–6230, 2007.
- C. Hesse. Model order estimation for blind source separation of multichannel magnetoencephalogram and electroencephalogram signals. In *Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 3348–3351, 2008.
- S. Hoffmann and M. Falkenstein. The correction of eye blink artefacts in the EEG: a comparison of two prominent methods. *PLoS ONE*, 3(8):e3004, 2008.
- R. Hornero, J. Escudero, A. Fernández, J. Poza, and C. Gómez. Spectral and nonlinear analyses of MEG background activity in patients with Alzheimer’s disease. *IEEE Transactions on Biomedical Engineering*, 55(6):1658–1665, 2008.
- R. Hornero, D. Abásolo, J. Escudero, and C. Gómez. Análisis no lineal de EEG y MEG para la ayuda en el diagnóstico de la enfermedad de Alzheimer. In *Libro de Actas del Simposio CEA de Bioingeniería 2009*, pages 40–49, Elche, Spain, 2009a.
- R. Hornero, D. Abásolo, J. Escudero, and C. Gómez. Nonlinear analysis of electroencephalogram and magnetoencephalogram recordings in patients with Alzheimer’s disease. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 367(1887):317–336, 2009b.
- D. Huang and J. Mi. A new constrained independent component analysis method. *IEEE Transactions on Neural Networks*, 18(5):1532–1535, 2007.
- N. Huang, Z. Shen, S. Long, M. Wu, H. Shih, Q. Zheng, N. Yen, C. Tung, and H. Liu. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of*

- the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454(1971):903–995, 1998.
- C. Hung, P. Wang, B. Soong, S. Teng, J. Hsieh, and Y. Wu. Blind Source Separation of Concurrent Disease-Related Patterns from EEG in Creutzfeldt–Jakob Disease for Assisting Early Diagnosis. *Annals of Biomedical Engineering*, 35(12):2168–2179, 2007.
- A. Hyvärinen. Fast and Robust Fixed-Point Algorithms for Independent Component Analysis. *IEEE Transactions on Neural Networks*, 10(3):626–634, 1999.
- A. Hyvärinen and E. Oja. A fast fixed-point algorithm for independent component analysis. *Neural Computation*, 9(7):1483–1492, 1997.
- A. Hyvärinen, J. Karhunen, and E. Oja. *Independent Component Analysis*. John Wiley & Sons, NJ, U.S.A., 2001.
- A. Hyvärinen, P. Ramkumar, L. Parkkonen, and R. Hari. Independent component analysis of short-time Fourier transforms for spontaneous EEG/MEG analysis. *NeuroImage*, 49(1):257–271, 2010.
- S. Ikeda and K. Toyama. Independent component analysis for noisy data—MEG data analysis. *Neural Networks*, 13(10):1063–1074, 2000.
- J. Iriarte, E. Urrestarazu, M. Valencia, M. Alegre, A. Malanda, C. Viteri, and J. Artieda. Independent component analysis as a tool to eliminate artifacts in EEG: a quantitative study. *Journal of clinical neurophysiology*, 20(4):249–257, 2003.
- C. James and O. Gibson. Temporally constrained ICA: an application to artifact rejection in electromagnetic brain signal analysis. *IEEE Transactions on Biomedical Engineering*, 50(9):1108–1116, 2003.
- C. James and C. Hesse. Independent component analysis for biomedical signals. *Physiological Measurement*, 26(1):R15–R39, 2005.
- D. Jarchi, R. Boostani, M. Taheri, and S. Sanei. Seizure source localization using a hybrid second order blind identification and extended rival penalized competitive learning algorithm. *Biomedical Signal Processing and Control*, 4(2):108–117, 2009.
- V. Jelic, S. Johansson, O. Almkvist, M. Shigeta, P. Julin, A. Nordberg, B. Winblad, and L. Wahlund. Quantitative electroencephalography

- in mild cognitive impairment: longitudinal changes and possible prediction of Alzheimer's disease. *Neurobiology of aging*, 21(4):533–540, 2000.
- J. Jeong. EEG dynamics in patients with Alzheimer's disease. *Clinical Neurophysiology*, 115(7):1490–1505, 2004.
- J. Jeong, J. Gore, and B. Peterson. Mutual information analysis of the EEG in patients with Alzheimer's disease. *Clinical neurophysiology*, 112(5):827–835, 2001.
- S. Jin, J. Jeong, D. Jeong, D. Kim, and S. Kim. Nonlinear dynamics of the EEG separated by independent component analysis after sound and light stimulation. *Biological Cybernetics*, 86(5):395–401, 2002.
- J. Jobson. *Applied Multivariate Data Analysis: Regression and Experimental Design*, volume 1. Springer Verlag, 1991a.
- J. Jobson. *Applied Multivariate Data Analysis: Categorical and Multivariate Methods*, volume 2. Springer Verlag, 1991b.
- V. Jousmäki and R. Hari. Cardiac Artifacts in Magnetoencephalogram. *Journal of Clinical Neurophysiology*, 13(2):172–176, 1996.
- C. Joyce, I. Gorodnitsky, and M. Kutas. Automatic removal of eye movement and blink artifacts from EEG data using blind component separation. *Psychophysiology*, 41(2):313–325, 2004.
- T. Jung, S. Makeig, C. Humphries, T. Lee, M. Mckeown, V. Iragui, and T. Sejnowski. Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, 37(2):163–178, 2000.
- J. Kierkels, G. van Boxtel, and L. Vogten. A model-based objective evaluation of eye movement correction in EEG recordings. *IEEE Transactions on Biomedical Engineering*, 53(2):246–253, 2006.
- M. Klemm, J. Haueisen, and G. Ivanova. Independent component analysis: comparison of algorithms for the investigation of surface electrical brain activity. *Medical & Biological Engineering & Computing*, 47(4):413–423, 2009.
- G. Knoll. Single-photon emission computed tomography. *Proceedings of the IEEE*, 71(3):320–329, 1983.

- D. Knopman, S. DeKosky, J. Cummings, H. Chui, J. Corey-Bloom, N. Relkin, G. Small, B. Miller, and J. Stevens. Practice parameter: Diagnosis of dementia (an evidence-based review). Report of the Quality Standards Subcommittee of the American Academy of Neurology. *Neurology*, 56(9):1143–1153, 2001.
- K. Kobayashi, C. James, T. Nakahori, T. Akiyama, and J. Gotman. Isolation of epileptiform discharges from unaveraged EEG by independent component analysis. *Clinical Neurophysiology*, 110(10):1755–1763, 1999.
- V. Krishnaveni, S. Jayaraman, P. Kumar, K. Shivakumar, and K. Ramadoss. Comparison of Independent Component Analysis Algorithms for removal of ocular artifacts from Electroencephalogram. *Measurement Science Review Journal*, 5(2):67–79, 2005.
- V. Krishnaveni, S. Jayaraman, and K. Ramadoss. Application of mutual information based least dependent component analysis (MILCA) for removal of ocular artifacts from electroencephalogram. *International Journal of Biological and Medical Sciences*, 1(1):63–74, 2006.
- P. Laguna and L. Sörnmo. Introduction. Editorial on Signal processing in vital rhythms and signs. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 367(1887):207–211, 2009.
- D. Lahiri, M. Farlow, N. Greig, and K. Sambamurti. Current drug targets for Alzheimer’s disease treatment. *Drug Development Research*, 56(3):267–281, 2002.
- J. Lambert, S. Heath, G. Even, D. Champion, K. Sleegers, M. Hiltunen, O. Combarros, D. Zelenika, M. Bullido, B. Tavernier, et al. Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer’s disease. *Nature Genetics*, 41(10):1094–1099, 2009.
- T. Lee, M. Girolami, and T. Sejnowski. Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural Computation*, 11(2):417–441, 1999.
- A. Lempel and J. Ziv. On the complexity of finite sequences. *IEEE Transactions on Information Theory*, 22(1):75–81, 1976.
- P. LeVan, E. Urrestarazu, and J. Gotman. A system for automatic artifact removal in ictal scalp EEG based on independent component analysis

- and Bayesian classification. *Clinical Neurophysiology*, 117(4):912–927, 2006.
- Y. Li, D. Powers, and J. Peach. Comparison of blind source separation algorithms. *Advances in Neural Networks and Applications*, pages 18–23, 2001.
- Y. Li, Z. Ma, W. Lu, and Y. Li. Automatic removal of the eye blink artifact from EEG using an ICA-based template matching approach. *Physiological Measurement*, 27(4):425–436, 2006.
- A. Lobo, J. Ezquerro, B. Gómez, J. Sala, and D. Seva. Cognitive mini-test (a simple practical test to detect intellectual changes in medical patients). *Actas Luso-Españolas de Neurología, Psiquiatría y Ciencias Afines*, 7(3):189–202, 1979.
- W. Lu and J. Rajapakse. Approach and applications of constrained ICA. *IEEE Transactions on Neural Networks*, 16(1):203–212, 2005.
- W. Lu and J. Rajapakse. ICA with reference. *Neurocomputing*, 69(16–18):2244–2257, 2006.
- C. Maestú, E. Gómez-Utrero, R. Piñeiro, and R. Sola. Magnetoencephalography: A new functional diagnostic technique for the neurosciences. *Revista de Neurología*, 28(11):1077–1090, 1999.
- N. Mammone and F. Morabito. Enhanced automatic artifact detection based on independent component analysis and Renyi’s entropy. *Neural Networks*, 21(7):1029–1040, 2008.
- D. Mantini, R. Franciotti, G. Romani, and V. Pizzella. Improving MEG source localizations: An automated method for complete artifact removal based on independent component analysis. *NeuroImage*, 40(1):160–173, 2008.
- C. Masters, R. Cappai, K. Barnham, and V. Villemagne. Molecular mechanisms for Alzheimer’s disease: implications for neuroimaging and therapeutics. *Journal of Neurochemistry*, 97(6):1700–1725, 2006.
- M. Mattson. Pathways towards and away from Alzheimer’s disease. *Nature*, 430(7000):631–639, 2004.
- G. McKhann, D. Drachman, M. Folstein, R. Katzman, D. Price, and E. Stadlan. Clinical diagnosis of Alzheimer’s disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and

- Human Services Task Force on Alzheimer's Disease. *Neurology*, 34(7): 939–944, 1984.
- C. Melissant, A. Ypma, E. Frietman, and C. Stam. A method for detection of Alzheimer's disease using ICA-enhanced EEG measurements. *Artificial Intelligence in Medicine*, 33(3):209–222, 2005.
- K. Najarian and R. Splinter. *Biomedical signal and image processing*. CRC Press, Boca Raton, FL, U.S.A., 2006.
- National Institute on Aging. *Alzheimer's Disease. Unraveling the Mystery*, volume 08-3782 of the National Institutes of Health (NIH) Publication. U.S. Department of Health and Human Services, 2008. www.nia.nih.gov/Alzheimers/Publications/UnravelingTheMystery/, last accessed: 7 October 2009.
- K. Nazarpour, H. Mohseni, C. Hesse, J. Chambers, and S. Sanei. A novel semiblind signal extraction approach for the removal of eye-blink artifact from EEGs. *EURASIP Journal on Advances in Signal Processing*, 2008:857459, 2008.
- P. Nestor, P. Scheltens, and J. Hodges. Advances in the early detection of Alzheimer's disease. *Nature Reviews Neuroscience*, 5(Supplement): S34–S41, 2004.
- N. Nicolaou and S. Nasuto. Automatic artefact removal from event-related potentials via clustering. *The Journal of VLSI Signal Processing*, 48(1):173–183, 2007.
- P. Nunez, R. Srinivasan, A. Westdorp, R. Wijesinghe, D. Tucker, R. Silberstein, and P. Cadusch. EEG coherency I: statistics, reference electrode, volume conduction, Laplacians, cortical imaging, and interpretation at multiple scales. *Electroencephalography and clinical Neurophysiology*, 103(5):499–515, 1997.
- P. Nunez, R. Silberstein, Z. Shi, M. Carpenter, R. Srinivasan, D. Tucker, S. Doran, P. Cadusch, and R. Wijesinghe. EEG coherency II: experimental comparisons of multiple measures. *Clinical Neurophysiology*, 110(3):469–486, 1999.
- Y. Okada, J. Jung, and T. Kobayashi. An automatic identification and removal method for eye-blink artifacts in event-related magnetoencephalographic measurements. *Physiological Measurement*, 28(12): 1523–1532, 2007.

- B. Onaral. Biomedical signal analysis. In J. Bronzino, editor, *The biomedical engineering handbook*, chapter I-1. Springer, Boca Raton, FL, U.S.A., 3rd, edition, 2006.
- J. Onton, M. Westerfield, J. Townsend, and S. Makeig. Imaging human EEG dynamics using independent component analysis. *Neuroscience and Biobehavioral Reviews*, 30(6):808–822, 2006.
- D. Osipova, J. Ahveninen, S. Kaakkola, I. Jääskeläinen, J. Huttunen, and E. Pekkonen. Effects of scopolamine on MEG spectral power and coherence in elderly subjects. *Clinical Neurophysiology*, 114(10):1902–1907, 2003.
- D. Osipova, J. Ahveninen, O. Jensen, A. Ylikoski, and E. Pekkonen. Altered generation of spontaneous oscillations in Alzheimer’s disease. *Neuroimage*, 27(4):835–841, 2005.
- D. Osipova, E. Pekkonen, and J. Ahveninen. Enhanced magnetic auditory steady-state response in early Alzheimer’s disease. *Clinical Neurophysiology*, 117(9):1990–1995, 2006a.
- D. Osipova, K. Rantanen, J. Ahveninen, R. Ylikoski, O. Häppölä, T. Strandberg, and E. Pekkonen. Source estimation of spontaneous MEG oscillations in mild cognitive impairment. *Neuroscience letters*, 405(1–2):57–61, 2006b.
- E. Pekkonen, I. Jääskeläinen, M. Hietanen, M. Huotilainen, R. Näätänen, R. Ilmoniemi, and T. Erkinjuntti. Impaired preconscious auditory processing and cognitive functions in Alzheimer’s disease. *Clinical Neurophysiology*, 110(11):1942–1947, 1999.
- E. Pekkonen, I. Jääskeläinen, S. Kaakkola, and J. Ahveninen. Cholinergic modulation of preattentive auditory processing in aging. *Neuroimage*, 27(2):387–392, 2005.
- E. Pereda, R. Quiroga, and J. Bhattacharya. Nonlinear multivariate analysis of neurophysiological signals. *Progress in Neurobiology*, 77(1–2):1–37, 2005.
- S. Pincus. Assessing Serial Irregularity and Its Implications for Health. *Annals of the New York Academy of Sciences*, 954(1):245–267, 2001.
- J. Poza, R. Hornero, D. Abásolo, A. Fernández, and J. Escudero. Analysis of spontaneous MEG activity in patients with Alzheimer’s disease using

- spectral entropies. In *Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 6179–6182, 2007a.
- J. Poza, R. Hornero, D. Abásolo, A. Fernández, and M. García. Extraction of spectral based measures from MEG background oscillations in Alzheimer's disease. *Medical Engineering and Physics*, 29(10):1073–1083, 2007b.
- J. Poza, R. Hornero, J. Escudero, A. Fernández, and C. Gómez. Analysis of spontaneous MEG activity in Alzheimer's disease using time-frequency parameters. In *Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 5712–5715, 2008a.
- J. Poza, R. Hornero, J. Escudero, A. Fernández, and C. Sánchez. Regional Analysis of Spontaneous MEG Rhythms in Patients with Alzheimer's Disease Using Spectral Entropies. *Annals of Biomedical Engineering*, 36(1):141–152, 2008b.
- J. Poza, R. Hornero, M. García, J. Escudero, and A. Fernández. Análisis de la irregularidad en la actividad MEG de enfermos de Alzheimer mediante representaciones tiempo-frecuencia. In *Libro de Actas del XXVII Congreso Anual de la Sociedad Española de Ingeniería Biomédica*, pages 61–64, 2009.
- C. Purnell, S. Gao, C. Callahan, and H. Hendrie. Cardiovascular Risk Factors and Incident Alzheimer Disease: A Systematic Review of the Literature. *Alzheimer Disease & Associated Disorders*, 23(1):1–10, 2009.
- S. Rampp and H. Stefan. On the opposition of EEG and MEG. *Clinical Neurophysiology*, 118(8):1658–1659, 2007.
- B. Reisberg. Functional assessment staging (FAST). *Psychopharmacology Bulletin*, 24(4):653–659, 1988.
- C. Reitz and R. Mayeux. Endophenotypes in normal brain morphology and Alzheimer's disease: A review. *Neuroscience*, 164(1):174–190, 2009.
- J. Richman and J. Moorman. Physiological time-series analysis using approximate entropy and sample entropy. *American journal of physiology. Heart and circulatory physiology*, 278(6):H2039–H2049, 2000.

- G. Rilling, P. Flandrin, and P. Gonçalves. On empirical mode decomposition and its algorithms. In *IEEEURASIP Workshop on Nonlinear Signal and Image Processing NSIP-03, Grado (I)*, 2003.
- S. Romero, M. Mañanas, and M. Barbanøj. A comparative study of automatic techniques for ocular artifact reduction in spontaneous EEG signals based on clinical target variables: A simulation case. *Computers in Biology and Medicine*, 38(3):348–360, 2008.
- S. Romero, M. Mañanas, and M. Barbanøj. Ocular Reduction in EEG Signals Based on Adaptive Filtering, Regression and Blind Source Separation. *Annals of Biomedical Engineering*, 37(1):176–191, 2009.
- F. Rong and J. Contreras-Vidal. Magnetoencephalographic artifact identification and automatic removal based on independent component analysis and categorization approaches. *Journal of neuroscience methods*, 157(2):337–354, 2006.
- P. Rossini, S. Rossi, C. Babiloni, and J. Polich. Clinical neurophysiology of aging brain: from normal aging to neurodegeneration. *Progress in Neurobiology*, 83(6):375–400, 2007.
- P. Sadasivan and D. Dutt. SVD based technique for noise reduction in electroencephalographic signals. *Signal Processing*, 55(2):179–189, 1996.
- M. Samanta, B. Wilson, K. Santhi, K. Sampath Kumar, and B. Suresh. Alzheimer disease and its management: A review. *American Journal of Therapeutics*, 13(6):516–526, 2006.
- M. Samonas, M. Petrou, and A. Ioannides. Identification and elimination of cardiac contribution insingle-trial magnetoencephalographic signals. *IEEE Transactions on Biomedical Engineering*, 44(5):386–393, 1997.
- T. Sander, G. Wubbeler, A. Lueschow, G. Curio, and L. Trahms. Cardiac artifact subspace identification and elimination in cognitive MEG data using time-delayed decorrelation. *IEEE Transactions on Biomedical Engineering*, 49(4):345–354, 2002.
- S. Sanei and J. Chambers. *EEG Signal Processing*. Wiley-Interscience, Chichester, UK, 2007.
- A. Schloegl, A. Ziehe, and K. Müller. Automated ocular artifact removal: comparing regression and component-based methods. Available from Nature Precedings at <http://hdl.handle.net/10101/npre.2009.3446.1>, 2009.

- A. Schlögl, C. Keinrath, D. Zimmermann, R. Scherer, R. Leeb, and G. Pfurtscheller. A fully automated correction method of EOG artifacts in EEG recordings. *Clinical Neurophysiology*, 118(1):98–104, 2007.
- S. Searle, F. Speed, and G. Milliken. Population marginal means in the linear model: an alternative to least squares means. *The American Statistician*, 34(4):216–221, 1980.
- D. Selkoe. Alzheimer’s disease: genes, proteins, and therapy. *Physiological Reviews*, 81(2):741–766, 2001.
- S. Shao, K. Shen, C. Ong, E. Wilder-Smith, and X. Li. Automatic EEG Artifact Removal: A Weighted Support-Vector-Machine Approach With Error Correlation. *IEEE Transactions on Biomedical Engineering*, 56(2):336–344, 2008.
- L. Shoker, S. Sanei, and J. Chambers. Artifact removal from electroencephalograms using a hybrid BSS-SVM algorithm. *IEEE Signal Processing Letters*, 12(10):721–724, 2005a.
- L. Shoker, S. Sanei, W. Wang, and J. Chambers. Removal of eye blinking artifact from the electro-encephalogram, incorporating a new constrained blind source separation algorithm. *Medical and Biological Engineering and Computing*, 43(2):290–295, 2005b.
- M. Signorino, E. Pucci, N. Belardinelli, G. Nolfe, and F. Angeleri. EEG spectral analysis in vascular and Alzheimer dementia. *Electroencephalography and clinical neurophysiology*, 94(5):313–325, 1995.
- J. Sleigh, D. Steyn-Ross, M. Steyn-Ross, C. Grant, and G. Ludbrook. Cortical entropy changes with general anaesthesia: theory and experiment. *Physiological Measurement*, 25(4):921–934, 2004.
- C. Stam. Nonlinear dynamical analysis of EEG and MEG: review of an emerging field. *Clinical Neurophysiology*, 116(10):2266–2301, 2005.
- C. Stam. Use of magnetoencephalography (MEG) to study functional brain networks in neurodegenerative disorders. *Journal of the Neurological Sciences*, 2010.
- C. Stam, B. Jones, I. Manshanden, A. van Cappellen van Walsum, T. Montez, J. Verbunt, J. de Munck, B. van Dijk, H. Berendse, and P. Scheltens. Magnetoencephalographic evaluation of resting-state functional connectivity in Alzheimer’s disease. *NeuroImage*, 32(3):1335–1344, 2006.

- K. Sternickel and A. Braginski. Biomagnetism using SQUIDs: status and perspectives. *Superconductor Science and Technology*, 19(3):S160–S171, 2006.
- S. Stufflebeam, N. Tanaka, and S. Ahlfors. Clinical applications of magnetoencephalography. *Human Brain Mapping*, 30(6):1813–1823, 2009.
- M. Sutherland, J. Liu, and A. Tang. Temporal delays in blind identification of primary somatosensory cortex. In *Proceedings of the 3rd IEEE International Conference on Machine Learning and Cybernetics (ICMLC 2004)*, Shanghai, China, pages 4222–4227, 2004.
- C. Sweeney-Reed and S. Nasuto. Detection of neural correlates of self-paced motor activity using empirical mode decomposition phase locking analysis. *Journal of Neuroscience Methods*, 184(1):54–70, 2009.
- A. Tang, J. Liu, and M. Sutherland. Recovery of correlated neuronal sources from EEG: The good and bad ways of using SOBI. *NeuroImage*, 28(2):507–519, 2005.
- J. Tang, X. Yang, J. Xu, Y. Tang, Q. Zou, and X. Zhang. The Algorithm of R Peak Detection in ECG Based on Empirical Mode Decomposition. In *Proceedings of the 2008 Fourth International Conference on Natural Computation*, volume 5, pages 624–627, 2008.
- K. Ting, P. Fung, C. Chang, and F. Chan. Automatic correction of artifact from single-trial event-related potentials by blind source separation using second order statistics only. *Medical Engineering and Physics*, 28(8):780–794, 2006.
- L. Tong, R. Liu, V. Soon, and Y. Huang. Indeterminacy and identifiability of blind identification. *IEEE Transactions on Circuits and Systems*, 38(5):499–509, 1991.
- J. Tukey. *Exploratory data analysis*. 1977.
- A. van Cappellen van Walsum, Y. Pijnenburg, H. Berendse, B. van Dijk, D. Knol, P. Scheltens, and C. Stam. A neural complexity measure applied to MEG data in Alzheimer’s disease. *Clinical Neurophysiology*, 114(6):1034–1040, 2003.
- M. van Es and L. van den Berg. Alzheimer’s disease beyond APOE. *Nature Genetics*, 41(10):1047–1048, 2009.

- F. Vialatte, A. Cichocki, G. Dreyfus, T. Musha, S. Shishkin, and R. Gervais. Early detection of Alzheimer's disease by blind source separation, time frequency representation, and bump modeling of EEG signals. *Lecture Notes in Computer Science*, 3696:683–692, 2005.
- F. Vialatte, J. Solé-Casals, M. Maurice, C. Latchoumane, N. Hudson, S. Wimalaratna, J. Jeong, and A. Cichocki. Improving the Quality of EEG Data in Patients with Alzheimer's Disease Using ICA. *Lecture Notes in Computer Science*, 5507:979–986, 2009.
- R. Vigário. Extraction of ocular artefacts from EEG using independent component analysis. *Electroencephalography and Clinical Neurophysiology*, 103(3):395–404, 1997.
- R. Vigário and E. Oja. Independence: A new criterion for the analysis of the electromagnetic fields in the global brain? *Neural Networks*, 13(8–9): 891–907, 2000.
- R. Vigário and E. Oja. BSS and ICA in Neuroinformatics: From Current Practices to Open Challenges. *IEEE Reviews in Biomedical Engineering*, 1:50–61, 2008.
- E. Vincent, R. Gribonval, and C. Févotte. Performance measurement in blind audio source separation. *IEEE Trans. on Audio, Speech and Language Processing*, 14(4):1462–1469, 2006.
- S. Vorobyov and A. Cichocki. Blind noise reduction for multisensory signals using ICA and subspace filtering, with application to EEG analysis. *Biological Cybernetics*, 86(4):293–303, 2002.
- J. Vrba and S. Robinson. Signal processing in magnetoencephalography. *Methods*, 25(2):249–271, 2001.
- G. Wallstrom, R. Kass, A. Miller, J. Cohn, and N. Fox. Automatic correction of ocular artifacts in the EEG: a comparison of regression-based and component-based methods. *International journal of psychophysiology*, 53(2):105–119, 2004.
- K. Wesson, R. Ochshorn, and B. Land. Low-cost, high-fidelity, adaptive cancellation of periodic 60Hz noise. *Journal of Neuroscience Methods*, 185(1):50–55, 2009.
- W. Woon, A. Cichocki, F. Vialatte, and T. Musha. Techniques for early detection of Alzheimer's disease using spontaneous EEG recordings. *Physiological Measurement*, 28(4):335–348, 2007.

- M. Wu and C. Hu. Empirical mode decomposition and synchrogram approach to cardiorespiratory synchronization. *Physical Review E*, 73(5):51917, 2006.
- Y. Ye, J. Garcia-Casado, J. Martinez-de Juan, and J. Ponce. Empirical mode decomposition: a method to reduce low frequency interferences from surface electroenterogram. *Medical & Biological Engineering & Computing*, 45(6):541–551, 2007.
- H. Zavala Fernández, T. Sander, M. Burghoff, R. Orglmeister, and L. Trahms. Comparison of ICA Algorithms for the Isolation of Biological Artifacts in Magnetoencephalography. *Lecture Notes in Computer Science*, 3889:511–518, 2006.
- D. Zhang, X. Wu, and X. Guo. The EEG Signal Preprocessing Based on Empirical Mode Decomposition. In *The 2nd International Conference on Bioinformatics and Biomedical Engineering, 2008 (ICBBE 2008)*, pages 2131–2134, 2008.
- X. Zhang, R. Roy, and E. Jensen. EEG complexity as a measure of depth of anesthesia for patients. *IEEE Transactions on Biomedical Engineering*, 48(12):1424–1433, 2001.
- A. Ziehe and K. Muller. TDSEP—an efficient algorithm for blind separation using time structure. In *Proc. Int. Conf. on Artificial Neural Networks (ICANN'98)*, pages 675–680, 1998.
- J. Zimmerman, P. Thiene, and J. Harding. Design and Operation of Stable rf-Biased Superconducting Point-Contact Quantum Devices, and a Note on the Properties of Perfectly Clean Metal Contacts. *Journal of Applied Physics*, 41:1572–1580, 1970.