



## Guiding functional connectivity estimation by structural connectivity in MEG: an application to discrimination of conditions of mild cognitive impairment



José Angel Pineda-Pardo<sup>a,b,\*</sup>, Ricardo Bruña<sup>b</sup>, Mark Woolrich<sup>c,d</sup>, Alberto Marcos<sup>d,e</sup>, Anna C. Nobre<sup>c</sup>, Fernando Maestú<sup>b</sup>, Diego Vidaurre<sup>c,d</sup>

<sup>a</sup> Laboratory of Neuroimaging, Centre for Biomedical Technology, Universidad Politécnica de Madrid, Campus de Montegancedo, 28223 Pozuelo de Alarcón, Spain

<sup>b</sup> Laboratory of Cognitive and Computational Neuroscience, Centre for Biomedical Technology, Universidad Politécnica de Madrid, Campus de Montegancedo, 28223 Pozuelo de Alarcón, Spain

<sup>c</sup> Oxford Center for Human Brain Activity (OHBA), University of Oxford, Oxford, United Kingdom

<sup>d</sup> The Oxford Centre for Functional MRI of the Brain (FMRIB), University of Oxford, Oxford, United Kingdom

<sup>e</sup> Department of Neurology, Hospital Clínico San Carlos, Madrid, Spain

### ARTICLE INFO

#### Article history:

Accepted 1 August 2014

Available online 8 August 2014

#### Keywords:

Resting state

Diffusion tensor imaging

Magnetoencephalography

Multimodal neuroimaging

Multivariate sparse regression

Graphical Lasso

Mild cognitive impairment

Machine learning

### ABSTRACT

Whole brain resting state connectivity is a promising biomarker that might help to obtain an early diagnosis in many neurological diseases, such as dementia. Inferring resting-state connectivity is often based on correlations, which are sensitive to indirect connections, leading to an inaccurate representation of the real backbone of the network. The precision matrix is a better representation for whole brain connectivity, as it considers only direct connections. The network structure can be estimated using the graphical lasso (GL), which achieves sparsity through  $l_1$ -regularization on the precision matrix. In this paper, we propose a structural connectivity adaptive version of the GL, where weaker anatomical connections are represented as stronger penalties on the corresponding functional connections. We applied beamformer source reconstruction to the resting state MEG recordings of 81 subjects, where 29 were healthy controls, 22 were single-domain amnesic Mild Cognitive Impaired (MCI), and 30 were multiple-domain amnesic MCI. An atlas-based anatomical parcellation of 66 regions was obtained for each subject, and time series were assigned to each of the regions. The fiber densities between the regions, obtained with deterministic tractography from diffusion-weighted MRI, were used to define the anatomical connectivity. Precision matrices were obtained with the region specific time series in five different frequency bands. We compared our method with the traditional GL and a functional adaptive version of the GL, in terms of log-likelihood and classification accuracies between the three groups. We conclude that introducing an anatomical prior improves the expressivity of the model and, in most cases, leads to a better classification between groups.

© 2014 Elsevier Inc. All rights reserved.

### Introduction

The pre-dementia stage of Mild Cognitive Impairment (MCI) represents an intermediate state of cognitive decline that precedes the development of Alzheimer's disease (AD) and other types of dementia (Petersen, 2011). The prevalence of MCI patients ranges from about 10% to 20% in people older than 65 years (Busse et al., 2006). From those that suffer from MCI there is a rate of progression to dementia of about 10% (Petersen, 2011). The pathophysiology of the disease

lead to a progressive loss of synapsis efficacy (Selkoe, 2002) and loss of neurons as well as damage in the white matter, due to the phosphorylation of the Tau protein affecting axon transmission, and the accumulation of the beta amyloid protein, which impairs gabaergic transmission (Garcia-Marin et al., 2009). All these lead to the view of AD as a “disconnection syndrome” (Bajo et al., 2010; Delbeuck et al., 2003; Stam et al., 2009) in which a progressive damage of global functional and structural connections are potentially the cause of the insidious cognitive impairment. The functional consequences of this “disconnection syndrome” in MCI patients has been assessed with functional magnetic resonance imaging (fMRI) (Binnewijzend et al., 2012; Toga and Thompson, 2013; Wang et al., 2013), electroencephalography (EEG) (Stam, 2003) and magnetoencephalography (MEG) (Bajo et al., 2012; Buldú et al., 2011; Zamrini et al., 2011).

The structural underpinnings associated with the functional disconnection may be studied in-vivo using diffusion-weighted magnetic

\* Corresponding author at: Center of Biomedical Technology (CTB), Parque Científico y Tecnológico de la UPM, Campus de Montegancedo, 28223 Pozuelo de Alarcón, Madrid, ES.

E-mail addresses: [joseangel.pineda@ctb.upm.es](mailto:joseangel.pineda@ctb.upm.es) (J.A. Pineda-Pardo), [ricardo.bruna@ctb.upm.es](mailto:ricardo.bruna@ctb.upm.es) (R. Bruña), [mark.woolrich@ohba.ox.ac.uk](mailto:mark.woolrich@ohba.ox.ac.uk) (M. Woolrich), [amarcosdolado@gmail.com](mailto:amarcosdolado@gmail.com) (A. Marcos), [kia.nobre@ohba.ox.ac.uk](mailto:kia.nobre@ohba.ox.ac.uk) (A.C. Nobre), [fernando.maestu@ctb.upm.es](mailto:fernando.maestu@ctb.upm.es) (F. Maestú), [diego.vidaurre@ohba.ox.ac.uk](mailto:diego.vidaurre@ohba.ox.ac.uk) (D. Vidaurre).

resonance imaging (DW-MRI). Structural disconnection in MCI has been studied in terms of white matter integrity (Medina et al., 2006; O'Dwyer et al., 2011) and connectivity strength between pre-defined brain regions (Daianu et al., 2013; Shao et al., 2012). However, little is known about the relationship between the functional and structural components of brain network organization. A recent paper (Pineda-Pardo et al., 2014) showed that the functional impairment of resting-state MEG networks in MCI patients was related to the white matter (WM) integrity of specific tracts, thus pointing to a structural-functional relationship that could provide complementary information when studying this type of pathology.

The application of machine-learning classifiers to resting-state (“off-task”) functional connectivity (FC) is rapidly spreading, and is proving to be a valuable tool in the diagnosis of neurological and psychiatric pathologies (Atluri et al., 2013; Castellanos et al., 2013). The use of resting-state fMRI in the diagnosis of amnesic MCI was applied by Wee and colleagues (Wee et al., 2012b,c). They achieved a 86% accuracy in patient classification, which was increased to 96% when including structural connectivity datasets.

The fact that structural and resting-state functional connectivity are strongly related has already been confirmed (Greicius et al., 2009; Van den Heuvel et al., 2009). However, the best way of combining these modalities to enrich our understanding of brain networks or to increase the diagnostic potential is still unclear. Computational models that simulate brain dynamics include the structural connectivity between nodes as coupling or constraining factors (Deco et al., 2013; Haimovici et al., 2013; Honey et al., 2007, 2009; Woolrich and Stephan, 2013). These models are capable of accurately reconstructing the large-scale networks derived from the slow fluctuations (<0.1Hz) of fMRI data. There is also evidence that these models explain up to 40% of the functional connections that are observed with empirical MEG and fMRI datasets (Cabral et al., 2013; Honey et al., 2009). Diffusion-tractography informed priors have also previously been shown to improve inference of effective connectivity using Dynamic Causal Modelling of fMRI data (Stephan et al., 2009). Therefore we hypothesize that the structural connectivity may also increase the accuracy in the estimation of the FC for MEG data.

Traditionally empirical FC has been computed as linear correlations between time-series. However, correlation-based approaches only measure pairwise dynamics, and are unable to provide accurate topographies of the interactions between many brain regions (Smith et al., 2011). Also, standard correlation analysis produces fully connected functional networks, which are difficult to interpret. Furthermore, it is impossible to discriminate direct from indirect functional connections between two nodes, which can be driven by third nodes (Friston, 2011; Smith et al., 2011). The best way to overcome these limitations is to use biophysical models, such as Dynamic Causal Modelling, to infer effective connectivity (Friston et al., 2003). However, this becomes unfeasible for whole-brain connectivity analyses for both computational and statistical efficiency reasons.

Simpler models, such as the multivariate Gaussian distribution, where direct connectivity is modelled by means of the precision matrix, represent a useful alternative (Marrelec et al., 2006). The precision matrix is the inverse of the covariance matrix. Zero elements in this matrix represent an absence of direct connections, i.e., the partial correlation is zero. Regularization approaches based on the  $l_1$ -norm (Vidaurre et al., 2013), such as the graphical lasso, allow one to find the zero elements in the precision matrix (Friedman et al., 2008; Smith et al., 2011), hence sparsifying the connection map, and, as a consequence, eliminating indirect functional connections. Sparse regression methods have been widely applied with fMRI time-series (Smith et al., 2011; Valdés-Sosa et al., 2005; Wee et al., 2014), however they typically do not consider DW-MRI structural connectivity information. One exception was the analysis by Hinne and colleagues (Hinne et al., 2014), where a shared adjacency matrix was estimated from an average across subjects of the structural connectivity. This matrix imposes a hard

constraint on the precision matrix, completely defining the sparsity pattern before estimating the actual precision matrix. Inclusion of this sparsity constraint was observed to produce a significant improvement in the estimation of the FC compared with non-structurally informed FC.

In this paper, we present a flexible approach for including DW-MRI structural connectivity information in FC estimation based on resting state MEG data. In our approach, the structural connectivity guides (without strictly constraining) the precision matrix structure. This means that, for example, in the case of a questionable estimation of the structural connectivity, the model still has room to depart from the structural sparsity pattern. For this purpose, we include one adaptive penalization factor per connection. Unlike the approach used by Hinne and colleagues (Hinne et al., 2014), this adaptive factor is individualised for each subject. Although a similar approach has been employed using fMRI time-series obtained from healthy subjects (Ng et al., 2012), to our knowledge structural connectivity priors have not been previously applied to neurophysiological data such as MEG. We expect this approach to yield interesting results because, compared to fMRI data, MEG data is a more direct representation of the neural activity with finer temporal resolution, thus allowing us to study functional networks on a much wider frequency spectrum. Indeed, it has recently been shown that MEG can be used to infer appropriate FC in the resting state. However, FC is not computed as correlations on the raw time-series as it usually is done in fMRI, but over band-limited power time-series, particularly in the alpha and beta bands (Brookes et al., 2012; Luckhoo et al., 2012).

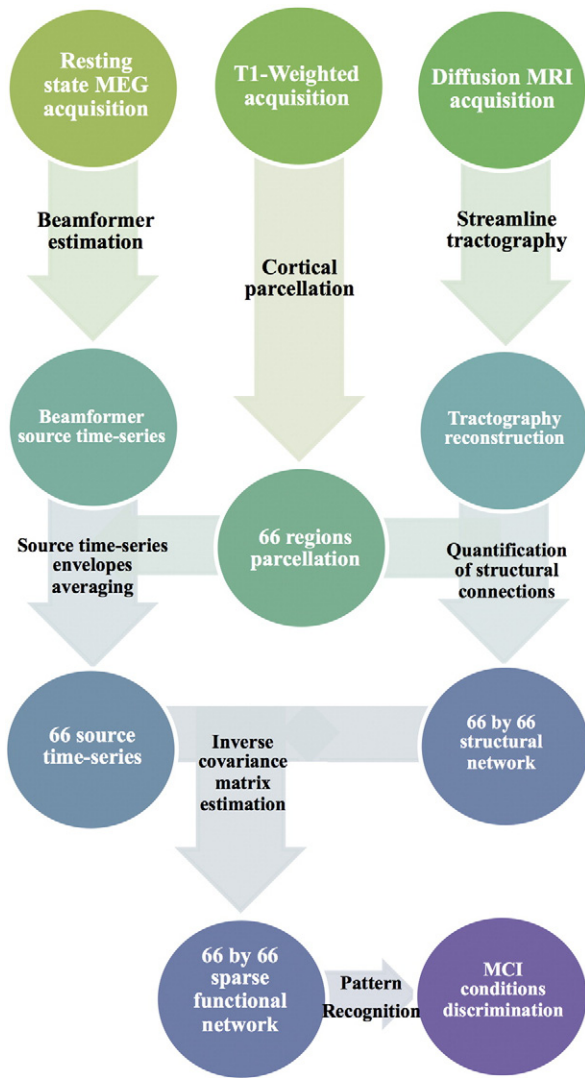
The aim of this work is hence to accurately estimate FC between brain regions, and to quantify how much the structural connectivity contributes to the estimation. We hypothesize that the element-wise adaptive penalization based on structural connectivity increases the accuracy in the estimation of the sparse networks. We compare this approach to cases in which we compute the adaptive penalization relying only on functional data, and in cases where the penalization is not adaptive. We also check whether the contribution of the structural connectivity to the functional connectivity estimation improves the accuracy of the discrimination between amnesic MCI subjects and healthy controls, and between the single-domain and multiple-domain subtypes of MCI. We carry out this analysis in different frequency bands. The accuracies are tested in a 10-fold cross-validation approach using four different classifiers: linear discriminant analysis (LDA), k-nearest neighbours (kNN), support vector machines with polynomial kernels (SVM), and support vector machines with radial basis functions kernels (SVMrbf). Consistently with the literature on other modalities, our results indicate that an appropriate inclusion of structural connectivity improves the classification.

## Materials and methods

We next describe in detail our workflow, which comprises five parts: sample selection, MRI acquisition and analysis, MEG acquisition and analysis, estimation of the connectivity matrix and MCI condition discrimination. A summary in flowchart format can be found in Fig. 1.

### Sample selection

From an initial sample of 142 participants, we selected 81 due to: artifactual MEG dataset ( $N = 33$ ); presence of vascular or tumour disease after structural MRI scans ( $N = 2$ ); artifactual MRI datasets due to motion ( $N = 7$ ); or due to unmatched ages between groups ( $N = 19$ ). Twenty-nine subjects in the sample were healthy elderly controls (HC) that were recruited from the “Seniors Center of the district of Chamartín, Madrid”. The remaining fifty-two subjects were amnesic mild cognitive impairment (MCI) patients. Diagnosis of the MCI patients was reached through neuropsychological examination at the *Hospital Clínico de Madrid* and the “UPDC del Ayuntamiento de Madrid”. The diagnostic examination included: the Spanish version of the Mini Mental



**Fig. 1.** Flow-chart diagram describing the processing pipeline of the study, from the acquisition of MEG and MRI data to the statistical discrimination between healthy controls and MCI groups.

State Examination (MMSE) (Lobo et al., 1999), the Global Deterioration Scale (GDS) (Reisberg et al., 1982), the Functional assessment questionnaire (FAQ) (Pfeffer et al., 1982), the Geriatric Depression Scale (GDS) (Yesavage et al., 1982), the Hachinski Ischemic Score (Rosen et al., 1980), the questionnaire for Instrumental Activities of Daily Living (Lawton and Brody, 1969), and the Functional Assessment Staging (FAST) (Auer and Reisberg, 1997). MCI patients were classified at the stage 3 of the Global Deterioration Scale (GDS), and were diagnosed according to the criteria of Grundman et al. (2004) and Petersen (2004). All MCI patients showed memory complaints, abnormal memory functions, normal general cognitive functions (MMSE > 23), absence or minimal impairment in activities of daily living. They had no history of major psychiatric disorders or neurological diseases. None of the participants were medicated for their condition with cholinesterase inhibitors (e.g., donepezil) or other cognitive enhancing substances (e.g., memantine) before MRI and MEG scanning.

MCI patients were further divided in two groups, according to their clinical and neuropsychological profile. Single-domain MCI (sdMCI) showed isolated memory impairment, whereas multiple-domain MCI (mdMCI) showed a memory deficit accompanied by various degrees of impairment in cognitive domains such as executive functions, visuo-spatial skills, and/or language. A demographical description of the sample is included in Table 1. The groups showed no differences in age after

**Table 1**

Demographic variables including gender, age, mini-mental state examination (MMSE) scores and education scores. Education level is quantified as: 1. Illiterate; 2. Elementary school studies; 3. Secondary school studies; 4. Technical or Mid-level studies; 5. Higher-education or University studies. Data are given as mean (standard deviation). <sup>1</sup>p < 0.01 significant difference after paired t-test statistical evaluation in comparison between HC and sdMCI. <sup>2</sup>p < 0.01 significant difference after paired t-test statistical evaluation in comparison between HC and mdMCI.

Group	Gender: Male/Female	Age	MMSE <sup>1,2</sup>	Education <sup>2</sup>
HC (N = 29)	8/21	71.52 (3.36)	29.21 (0.28)	3.44 (1.24)
sdMCI (N = 22)	10/12	73.00 (5.08)	27.80 (1.87)	2.73 (1.28)
mdMCI (N = 30)	7/23	73.06 (3.42)	26.77 (1.70)	2.53 (1.14)

t-test statistical comparison (p > 0.08). Also, no statistical differences were observed between groups in gender distributions after chi-squared statistical comparison (p > 0.16). However, revealed by a paired t-test (p < 0.01), there were differences in education scores and in mini-mental state examination (MMSE) evaluation.

The research described in this report was approved by the Ethics Committee of the Hospital Clínico San Carlos, Madrid. All of the participants signed a written informed consent before completing in any research activities.

*MRI acquisition and analysis*

All images were collected using a General Electric 1.5 T magnetic resonance (MR) scanner, using a high-resolution antenna and a homogenization PURE filter. 3D T1-weighted anatomical brain MRI scans were acquired with a Fast Spoiled Gradient Echo (FSPGR) sequence with parameters: TR/TE/TI = 11.2/4.2/450 ms; flip angle 12°; 1 mm slice thickness, a 256 × 256 matrix and FOV 25 cm. Diffusion-weighted images (DWI) were acquired with a single-shot echo-planar imaging sequence with the following parameters: TE/TR 96.1/12000 ms; NEX 3 for increasing the signal to noise ratio (SNR); 2.4 mm slice thickness, 128 × 128 matrix and 30.7 cm FOV yielding an isotropic voxel of 2.4 mm; 1 image with no diffusion sensitization (i.e., T2-weighted b<sub>0</sub> images) and 25 DWI (b = 900 s/mm<sup>2</sup>).

T1-weighted images were fed to Freesurfer (version 5.1.0) in order to segment each participant’s cortex into sixty-six anatomical cortical regions (Fischl et al., 2004) (see Supplementary Table 1 for region labels). These regions constituted the nodes for the anatomical and functional networks. Because of the low SNR in sub-cortical MEG source time-series, we used only cortical nodes. All segmentations were visually inspected, concluding that there was no need to discard any of the subjects of the sample due to incorrect grey matter segmentation, even in the presence of cortical atrophy (see Supplementary Fig. 1 for an example). Although omitting the connections to subcortical structures could potentially mislead the cortico-cortical connectivity, this atlas has still been proven to be useful in previous multimodal fMRI-DWI studies (Hagmann et al., 2008; Honey et al., 2009).

Diffusion-weighted images were pre-processed with FMRIB’s Diffusion Toolbox (FDT-FMRIB Software Library v5.0). Pre-processing consisted of eddy-current correction, motion correction, and the removal of non-brain tissue using the Brain Extraction Tool (Smith, 2002). The diffusion tensor model was fit for the pre-processed diffusion images using least squares fitting with the Diffusion Toolkit Software (DTK v0.6.2). Tensor deflection tractography was applied to the diffusion tensor images to build the tractography (Lazar et al., 2003). Stopping criteria for the streamlines propagation were a maximum angle of 35° between consecutive steps and a lower fractional anisotropy, i.e. FA < 0.1 (Johansen-Berg et al., 2004). Only the tracts with a length larger than 15 mm were retained (see Supplementary Fig. 2 for an illustrative representation of the DWI temporal-signal to noise ratio (tSNR) and the performance of the tractography). The structural connection between a pair of nodes (i, j) was defined as the fiber density FD<sub>ij</sub> of the connection,

i.e. number of tracts touching the two regions divided by the total number of tracts,  $\frac{n_{ij}}{N}$ . We normalized each of the connections by the average volume of the connected nodes,  $FD_{ij} = \frac{n_{ij}}{N} \frac{2}{V_i + V_j}$ , in order to control for differences in node size (Hagmann et al., 2008).

#### MEG acquisition and analysis

MEG data were acquired with a 306-channel Vectorview system (Elekta-Neuromag) at the Center for Biomedical Technology (Madrid, Spain). The system comprises 102 magnetometers and 204 planar gradiometers on a sensor array, located inside a magnetically shielded room. Sampling frequency was 1 kHz, and an online anti-alias filter (0.1–330 Hz) was applied. A head position indicator (HPI) system and a three-dimensional digitizer (FastrakPolhemus) were used to determine the position of the head with respect to the sensor array during the recordings. Four HPI coils were attached to the subject (one on each mastoid, two on the forehead), and their position with respect to the 3 fiducials (nasion, left and right pre auricular points) was determined. We recorded vertical eye movements, using two electrodes attached above and below the left eye in a bipolar montage. Resting-state acquisitions consisted of three-minute recordings, where subjects were asked to stay calm and with their eyes closed. External noise was removed from the MEG data using the temporal extension of Signal-Space Separation (tSSS) (Taulu and Kajola, 2005) in MaxFilter (version 2.2, Elekta-Neuromag), using as parameters a window length of ten seconds and a correlation limit of 0.9. Participants' head movements were corrected using the MaxMove extension of the software.

Electronic, muscle and ocular artifacts were automatically identified, and subsequently visually confirmed. The time series were segmented into trials of four seconds avoiding the segments containing any type of artifacts. Subjects with fewer than 15 clean trials were discarded. The mean number of clean trials was  $25.9 \pm 6.9$  for the HC group,  $27.8 \pm 6.1$  for the sdMCI group and  $26.2 \pm 5.5$  for the mdMCI group.

Datasets were band-pass filtered in five frequency bands: alpha (8–13 Hz), low beta (13–20 Hz), high beta (20–30 Hz), full beta (13–30 Hz) and a broader band containing the theta, alpha and beta bands (4–30 Hz). The selection for these specific bands was performed in concordance with previous results in MEG, which show that it is the alpha band the most affected in amnesic MCI subjects (Garcés et al., 2013; Ishii et al., 2010) and in AD (Stam et al., 2009). Also, resting state fMRI large scale networks were shown to be reproducible in MEG data for alpha and beta bands (Brookes et al., 2011).

Padding segments of one second to the clean trials were included to avoid edge effects. The data covariance matrix was computed for each clean trial and then averaged for each frequency band. Instead of standard covariance matrix regularization, which is performed by adding uncorrelated noise (i.e. amplifying the diagonal of the covariance matrix) (Vrba and Robinson, 2000), we used a Bayesian principal component analysis (PCA) data dimensionality reduction, which has been demonstrated to work well in this particular domain (Woolrich et al., 2011).

The underlying currents of the time series observed in the sensor datasets were reconstructed using a linear constrained minimum variance beamformer (Van Veen et al., 1997). Lead-fields for all vertices in a 5-mm grid were obtained using a single-shell Boundary-Element-Model forward model (Mosher et al., 1999). Lead-fields for magnetometers and planar gradiometers were scaled following Mohseni et al. (2012) in order to allow the fusion of the sensor time-series. The source currents were estimated at each vertex of the grid covering the whole brain. Dipole orientations were estimated by searching for the maximum power projection of the dipole (Sekihara et al., 2001). Beamformer time-series were obtained by multiplying the beamformer coefficients by the band-pass filtered time series. These were subsequently normalized by the coefficients of variance as suggested in Hall et al. (2013) for

the purpose of connectivity analyses. We followed previous work on resting state MEG FC (Brookes et al., 2012; Luckhoo et al., 2012) to obtain a single signal for each of the sixty-six cortical regions. For this matter, in order to avoid polarity swaps, we used the Hilbert transform to obtain the power envelopes of each time-series and then we averaged them within each region.

#### Estimation of the connectivity matrix

We assume the data to have a Gaussian distribution, and we model the connectivity matrix as the estimated precision matrix, defined as the inverse of the covariance matrix. We denote the precision matrix as  $\Theta$  and the sample covariance matrix as  $S$ . A zero in the precision matrix, say  $\theta_{ij} = 0$ , indicates that the corresponding partial correlation is zero, so that channels  $i$  and  $j$  are not directly connected (i.e., they are conditionally independent). The Gaussian assumption implies that dependencies between channels are always of second order, as higher order moments are always zero under this assumption.

In order to identify the connectivity pattern, we estimate a sparse precision matrix, i.e. with a number of elements exactly equal to zero. However, even if the covariance matrix is invertible (full rank), since data are always finite and noisy, the estimated precision matrix will have all elements different from zero. A popular way to get around these problems is to use  $l_1$ -norm regularisation (Vidaurre et al., 2013), which provides both a numerically stable solution and a sparse estimation. Within the context of Gaussian inverse covariance matrix estimation, this is achieved through the graphical lasso (Friedman et al., 2008), which maximises the criterion

$$\log \det(\Theta) - \text{tr}(S\Theta) - \lambda \|\Theta\|_1,$$

where  $\lambda$  is the regularisation parameter and  $\|\cdot\|_1$  refers to the  $l_1$ -norm operator. To solve this problem, we use a coordinate-descent procedure that incorporates recent developments to accelerate computations (Witten et al., 2011).

It is well known that adaptive regularisation, which uses adaptive weights for regularising the different coefficients, improves the efficiency of the estimator and leads to more accurate sparsity patterns (Zou, 2006). An adaptive version of the graphical lasso can be readily obtained by maximising instead

$$\log \det(\Theta) - \text{tr}(S\Theta) - \lambda \|W \cdot \Theta\|_1,$$

where  $W$  is a matrix of weights with elements  $W_{ij} = 1/\sqrt{S_{ij}}$ , and the operator  $\cdot$  denotes element-wise multiplication. The same algorithm can be used to solve this problem. Thus, the adaptive setting is completely driven by the sample covariance matrix of the functional data.

In our approach, we assign values to  $W_{ij}$  based on structural connectivity information. In particular, we set  $W_{ij} = 1/\sqrt{FD_{ij}}$ . By doing this, we inform the pattern of functional connectivity by a priori structural information in order to estimate more meaningful networks. We have implemented the three aforementioned flavours of the graphical lasso: non-adaptive graphical lasso (GL), functional data based adaptive graphical lasso (GLa), and structural data based adaptive graphical lasso (GLd). The computation of the Graphical Lasso was performed in R with the *glasso* Package,<sup>1</sup> and we base on this to implement the adaptive varieties.

We used a 10-fold cross-validation to assess the methods in terms of log-likelihood and density of the networks. Within each fold we took the NIM11576 that minimizes the Bayesian Information Criteria (BIC), which amounts to choosing the model with the largest approximate posterior probability (Hastie et al., 2009). Model selection is performed

<sup>1</sup> <http://cran.r-project.org/web/packages/glasso>.

within a routine in which we define an initial sequence of  $\lambda$  values. We estimate the precision matrices for each  $\lambda$  and compute the BIC statistic. We select  $\lambda_{min}$ , which corresponds to the minimum BIC, and we define a new sequence of  $\lambda$  values within a relatively small vicinity of  $\lambda_{min}$ . Following this procedure, we sharpen the search interval up to three times to obtain a final  $\lambda_{min}$ . We limit the search to a maximum network density of 20%. This threshold is the number of non-zero links in the average of the structural networks across subjects, and is in general agreement with previous studies (Hinne et al., 2014). The final networks that were used to predict the conditions of MCI (see next Section 2.5) were obtained using the mean  $\lambda$  across the cross-validation folds.

Note that the BIC criterion needs an estimation of the effective sample size. When data are independent and identically distributed, this number equals the actual number of data points. In our case, each data point corresponds to a MEG measurement, so data is strongly autocorrelated. As a consequence, the effective sample size is lower than the number of data points. In this paper, we obtain the effective sample size by dividing the number of data points by one plus two times the sum of the autocorrelation values, for lags from 1 to a sufficiently high number – in practice, until the autocorrelation vanishes (Geyer, 1992). This is implemented in the function *ess* from the R *mcmcse* Package.<sup>2</sup>

*Pattern classification*

We chose four different machine-learning classifiers to evaluate the accuracy of the predictions for the three network estimation methods: k-nearest neighbour (k-nn), linear discriminant analysis (LDA), support vector machine with polynomial (SVM), and radial basis functions (SVMrbf) kernels. Validation of the classification algorithms was performed with 10-fold cross-validation. For each run of the classification algorithms (e.g. one per graphical lasso approach, per frequency band, and per possible between-group combination), we performed a feature selection using non-parametrical Mann-Whitney statistical comparison between groups. The number of input features and the parameters of the classification algorithms described below were chosen by a nested 10-fold cross-validation procedure. The classification results of each fold were aggregated to the confusion matrix to obtain accuracies (rate of samples correctly classified), sensitivities (rate of samples in the second group correctly classified; see tables below), and specificities (rate of samples in the first group correctly classified).

LDA assumes that different groups generate observations based on different multivariate Gaussian distributions, so that, given two given groups, it is possible to define a boundary hyperplane where the probability for an observation to belong to any of the two groups is the same (Hastie et al., 2009). This boundary is then used to assign an observation to a group. We employed a regularized variant of LDA including a variable  $\gamma$ , in the interval [0,1], that attempts to shrink the group covariance matrices towards a diagonal matrix (Guo et al., 2007).

The k-nn classifier non-parametrically assigns an observation to the group to which the majority of the *k* closest training observations (nearest neighbours) belong (Hastie et al., 2009). The *k* closest neighbours were defined in terms of Euclidean distances, and *k* was chosen within the range [2,10].

SVM also defines a separating hyperplane in the feature space. The best hyperplane in this case will be the one with the largest margin between the two groups, where the margin is the distance between the closest samples to the hyperplane (Cortes and Vapnik, 1995). For the case of non-separable datasets, the margin is transformed to a soft margin, indicating that the hyperplane separates many but not all data points. Points in the feature space are typically mapped to some convenient space by means of the function  $\phi(x)$  for which we only need to

specify a kernel so that it holds  $\kappa(x, x') = \langle \phi(x), \phi(x') \rangle$ , where  $(x, x')$  are two instances in the feature space and  $\langle \cdot \rangle$  represents the dot product. For SVM, we employed polynomial kernels  $\kappa_p(x, x') = (1 + \langle x, x' \rangle)^d$  with *d* ranging from one to six. For SVMrbf, we used radial basis functions kernels  $\kappa_{rbf}(x, x') = \exp(-\langle (x - x'), (x - x') \rangle / 2\sigma^2)$ , with  $\sigma$  taking values in  $10^{[-5, -4, \dots, 4, 5]}$ .

**Results**

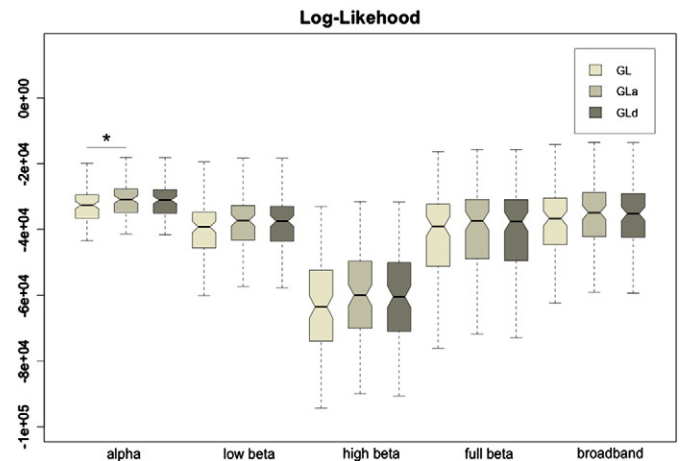
We first compare the ability of the different models to describe the data by reporting the cross-validated log-likelihood, which, assuming a Gaussian distribution, represents how well the precision matrix explains the (zero-mean) data set. Hence, the higher the log-likelihood, the more faithfully this model represents the data.

Fig. 2 shows the average log-likelihoods for all three graphical lasso approaches and frequency bands. Although the adaptive versions, which introduce penalizations based on the inverse of the sample functional connectivity matrix (GLa) or on the inverse of the structural connectivity weights (GLd), present very close likelihoods, they both are higher than the non-adaptive graphical lasso (GL). These differences are only significant in the alpha band ( $p < 0.05$  after paired t-test statistical comparisons for GLa-GL), but were not significant after multiple comparisons correction using False Discovery Rate (FDR).

Fig. 3 shows boxplots with the densities (ratio of non-zero connections in the precision matrices) for all methods and frequency bands. Interestingly, the densities were higher for the GL than for the GLa or GLd. This, along with the above likelihood results, suggests that the adaptive approaches lead to more robust network topographies with fewer spurious connections. The differences were statistically significant ( $qFDR < 0.05$ ) for GLa-GL and GLd-GL in all frequency bands, but, with the exception of alpha band, no statistical differences were observed between the GLa and GLd. In general, optimal densities for all methods and frequency bands were close to the maximum (20%).

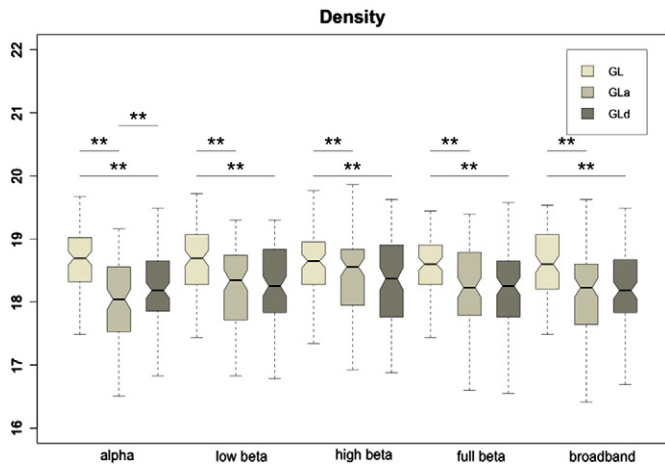
For each type of functional connectivity dataset (GL, GLa and GLd), the aforementioned machine-learning classifiers were run to estimate clinical group predictions using 10-fold cross-validation.

In the classification between HC and sdMCI, the best performance was observed for LDA and SVMrbf classifiers, achieving up to 86% of accuracy. Table 2 shows the accuracies along with the specificities and sensitivities. The best classification results were obtained in alpha, high beta and broadband ranges of frequencies. The maximum achieved



**Fig. 2.** Boxplots of the log-likelihood values for the three graphical lasso approaches: GL, GLa and GLd. The asterisk indicates that the distributions differ with significance  $p < 0.05$  after paired t-test statistical comparison.

<sup>2</sup> <http://cran.r-project.org/web/packages/mcmcse>.



**Fig. 3.** Boxplots of the density values for the three implemented graphical lasso approaches: GL, GLa and GLd. The double asterisk indicates that the distributions differ with significance  $p < 0.05$  after paired t-test statistical comparison, and survived a False Discovery Rate multiple comparisons correction ( $qFDR < 0.05$ ).

accuracy was of 86.27% (spec. 89.66% sens. 81.82%) obtained for GLd using broadband data. The classifier employed to obtain this accuracy was LDA and the chosen configuration parameters after the 10-fold cross-validation were 310 input features and  $\gamma = 0.5$ . The best accuracies for GL 82.35% (spec. 89.66% sens. 72.73%) and GLa 84.31% (spec. 89.66% sens. 77.27%) were obtained in the alpha band with SVMrbf and LDA classifiers respectively. Although these accuracies are close to the ones observed with GLd, it is clear that all four classifiers agree in that GLd with wide frequency band data offers the best predictor to discriminate between HC and sdMCI.

**Fig. 4A–C** represents, in a brain mesh, highly relevant connections that were selected by the previous LDA classifier in at least 9 of the 10-fold cross-validation testings. The width of the connections represents the median of the weights that the classifier assigned to this connection in each fold. For the sake of interpretation, only the top 1% of relevant connections are represented. **Fig. 4D** shows a matrix plot that represents in its lower part the number of folds that a specific link was selected and the median value assigned by the classifier across folds. The most important connections in this classification included

temporal regions (right and left transverse temporal gyri; left entorhinal cortex; left parahippocampal gyrus), frontal regions (left caudal middle frontal gyrus), and cingulate regions (left isthmus of cingulate gyrus).

**Table 3** shows the accuracies in the classification between HC and mdMCI. The maximum achieved accuracy was of 81.36% (spec. 82.76% sens. 80.00%). This accuracy was achieved with GLd for the frequency band containing the full beta [13–30 Hz], and employing a LDA classifier with 385 input features and  $\gamma = 0.9$ . **Figs. 5A–C** depicts the most relevant selected features, following the same inclusion criteria described for **Fig. 4**.

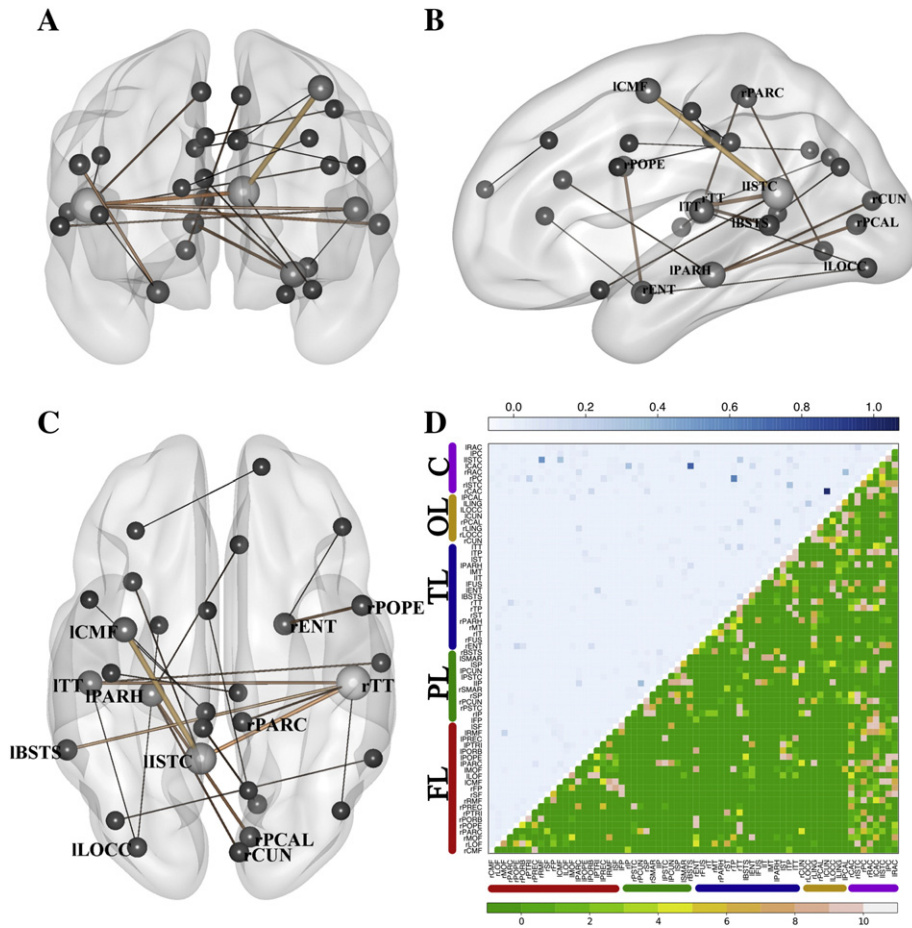
The most relevant connections in the classification between HC and mdMCI were linked with temporal regions (right and left transverse temporal gyri, left entorhinal cortex, right and left parahippocampal gyri) and cingulate regions (right and left rostral anterior cingulate cortices, right isthmus of cingulate gyrus). These regions had also an important role in the classification between HC and sdMCI, where, besides, connections with frontal regions acquired more relevance. These include the right frontal pole, the left rostral middle frontal gyrus and the right pars triangularis.

In the classification between sdMCI and mdMCI, the LDA classifier showed the best outcomes (see **Table 4**). The maximum achieved accuracy was observed for GL in the broadband 84.62% (spec. 81.82% sens. 86.67%). This outcome was obtained with a LDA classifier with ten input features and  $\gamma = 0.5$  (see **Figs. 6 A–C** for a graphical description of the most frequently selected features). In **Fig. 6**, many fewer links were selected for the classification between sdMCI and mdMCI. In this case, the maximum accuracy was achieved with only ten input features, which makes the number of depicted links (selected at least in nine of the ten folds) low. Note that the best accuracies for GLa and GLd were also obtained with LDA classifiers, and for the alpha and broadband datasets respectively.

The most important connections in the classification between sdMCI and mdMCI included occipital regions (right and left lingual gyri; right and left lateral occipital gyri), cingulate regions (right isthmus of cingulate gyrus) and frontal regions (right caudal middle frontal gyrus, right postcentral gyrus). In this scenario, connections between temporal regions were not as diagnostic as they were before for HC–sdMCI and HC–mdMCI. This result is consistent with the cognitive state of these patients, as both groups sdMCI and mdMCI have a memory impairment. While in the classification between HC–sdMCI or between HC–mdMCI

**Table 2**  
10-fold cross-validated accuracies for the classification between HC and sdMCI. The table shows the mean accuracies (specificities; sensitivities) in percentages for all five frequency bands, the three regularization methods (GL, GLa, GLd) and the four classifiers evaluated. Accuracies higher than 80% are highlighted.

HC vs sdMCI		GL	GLa	GLd
alpha	Knn	72.55 (79.31; 63.64)	72.55 (68.97; 77.27)	72.55 (72.41; 72.73)
	LDA	<b>80.39 (93.10; 63.64)</b>	<b>84.31 (89.66; 77.27)</b>	<b>82.35 (93.10; 68.18)</b>
	SVM	64.71 (75.86; 50.00)	76.47 (82.76; 68.18)	76.47 (86.21; 63.64)
	SVMrbf	<b>82.35 (89.66; 72.73)</b>	<b>80.39 (86.21; 72.73)</b>	<b>80.39 (93.10; 63.64)</b>
Low beta	Knn	72.55 (62.07; 86.36)	72.55 (65.52; 81.82)	68.63 (51.72; 90.91)
	LDA	70.59 (68.97; 72.73)	74.51 (75.86; 72.73)	72.55 (89.66; 50.00)
	SVM	66.67 (65.52; 68.18)	68.63 (75.86; 59.09)	66.67 (68.97; 63.64)
	SVMrbf	68.63 (72.41; 63.64)	72.55 (79.31; 63.64)	68.63 (72.41; 63.64)
High beta	Knn	70.59 (65.52; 77.27)	72.55 (79.31; 63.64)	74.51 (79.31; 68.18)
	LDA	<b>82.35 (86.21; 77.27)</b>	76.47 (79.31; 72.73)	76.47 (86.21; 63.64)
	SVM	66.67 (75.86; 54.55)	70.59 (75.86; 63.64)	<b>80.39 (86.21; 72.73)</b>
	SVMrbf	68.63 (93.10; 36.36)	76.47 (96.55; 50.00)	74.51 (86.21; 59.09)
Full beta	Knn	68.63 (62.07; 77.27)	70.59 (68.97; 72.73)	68.63 (62.07; 77.27)
	LDA	74.51 (82.76; 63.64)	68.63 (75.86; 59.09)	76.47 (82.76; 68.18)
	SVM	72.55 (79.31; 63.64)	64.71 (65.52; 63.64)	68.63 (75.86; 59.09)
	SVMrbf	76.47 (86.21; 63.64)	66.67 (96.55; 27.27)	72.55 (89.66; 50.00)
Broadband	Knn	74.51 (75.86; 72.73)	72.55 (79.31; 63.64)	<b>80.39 (72.41; 90.91)</b>
	LDA	78.43 (82.76; 72.73)	74.51 (82.76; 63.64)	<b>86.27 (89.66; 81.82)</b>
	SVM	74.51 (75.86; 72.73)	74.51 (75.86; 72.73)	74.51 (82.76; 63.64)
	SVMrbf	<b>80.39 (82.76; 77.27)</b>	72.55 (79.31; 63.64)	<b>82.35 (89.66; 72.73)</b>



**Fig. 4.** Selected features in the LDA classifier for HC-sdMCI for the broadband and GLd (accuracy 86.27%). A–C panels represent three views (sagittal, coronal and axial) of the selected links. The width and color (black to soft brown) of these links grows proportional to the median of the weights assigned to the links by the classifier across the ten folds of the cross-validation testing procedure. The size of the ball that represents the node is proportional to the number of links converging at this node. In panel D, the upper triangular matrix shows the median of the assigned weights and the lower triangular matrix represents the number of folds in which a specific link has been selected. The nodes were grouped according to brain lobes: FL – Frontal Lobe; PL – Parietal Lobe; TL – Temporal Lobe; OL – Occipital Lobe; C – Cingulate Cortex. See Supplementary Table 2 for a ranking of the links with the highest weights.

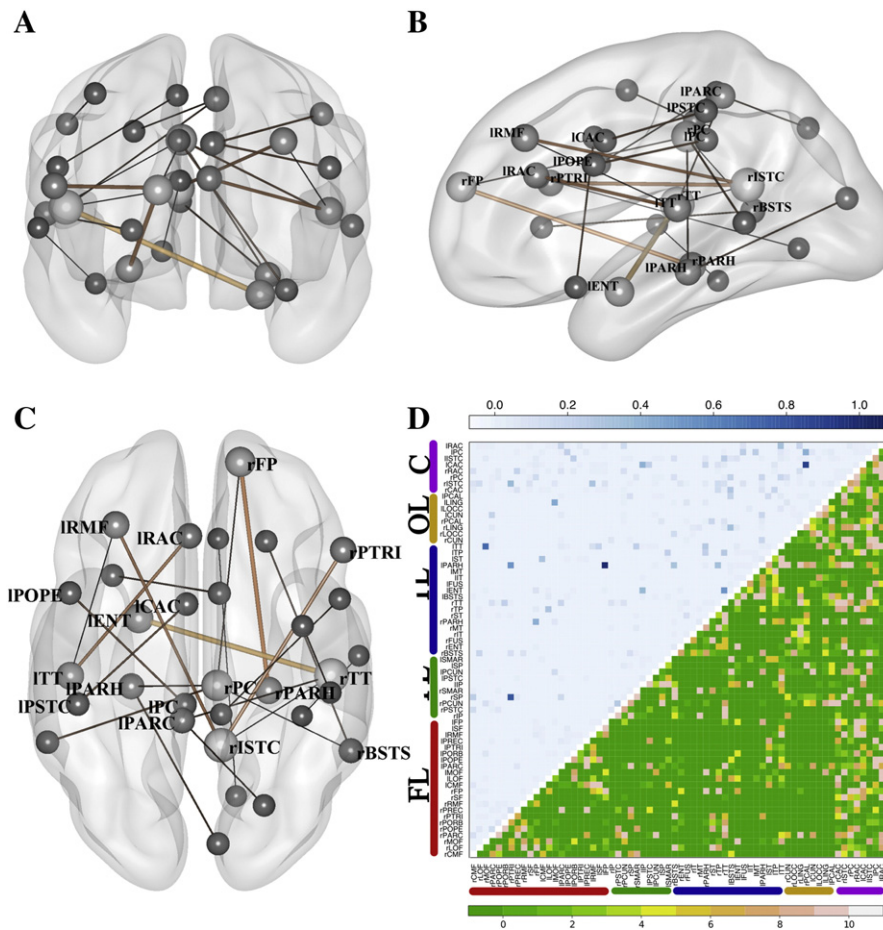
the GLd offers the best results, in the classification between sdMCI-mdMCI the best accuracies were obtained by GL (the non-adaptive case). In general, datasets in alpha and broadbands provided in general

the best classification results. Consistent with previous reports, alpha is traditionally considered to be the main affected band in these amnesic MCI patients (Garcés et al., 2013; Ishii et al., 2010). However,

**Table 3**

10-fold cross-validated accuracies for the classification between HC and mdMCI. The table shows the mean accuracies (specificities; sensitivities) in percentages for all five frequency bands, three regularization methods (GL, GLa, GLd) and the four classifiers evaluated. Accuracies higher than 80% are highlighted.

HC vs mdMCI		GL	GLa	GLd
alpha	Knn	62.71 (62.07; 63.33)	64.41 (75.86; 53.33)	64.41 (75.86; 53.33)
	LDA	69.49 (75.86; 63.33)	76.27 (72.41; 80.00)	71.19 (62.07; 80.00)
	SVM	66.10 (48.28; 83.33)	66.10 (55.17; 76.67)	67.80 (58.62; 76.67)
	SVMrbf	59.32 (55.17; 63.33)	67.80 (72.41; 63.33)	64.41 (65.52; 63.33)
Low beta	Knn	69.49 (58.62; 80.00)	67.80 (75.86; 60.00)	62.71 (55.17; 70.00)
	LDA	71.19 (58.62; 83.33)	71.19 (72.41; 70.00)	76.27 (86.21; 66.67)
	SVM	66.10 (68.97; 63.33)	69.49 (68.97; 70.00)	66.10 (65.52; 66.67)
	SVMrbf	66.10 (79.31; 53.33)	69.49 (68.97; 70.00)	66.10 (72.41; 60.00)
High beta	Knn	59.32 (65.52; 53.33)	66.10 (65.52; 66.67)	64.41 (62.07; 66.67)
	LDA	62.71 (62.07; 63.33)	69.49 (68.97; 70.00)	71.19 (65.52; 76.67)
	SVM	55.93 (58.62; 53.33)	62.71 (58.62; 66.67)	66.10 (62.07; 70.00)
	SVMrbf	57.63 (27.59; 86.67)	59.32 (48.28; 70.00)	69.49 (65.52; 73.33)
Full beta	Knn	62.71 (79.31; 46.67)	62.71 (72.41; 53.33)	62.71 (72.41; 53.33)
	LDA	71.19 (72.41; 70.00)	76.27 (86.21; 66.67)	<b>81.36 (82.76; 80.00)</b>
	SVM	69.49 (75.86; 63.33)	64.41 (62.07; 66.67)	69.49 (65.52; 73.33)
	SVMrbf	62.71 (65.52; 60.00)	66.10 (62.07; 70.00)	67.80 (68.97; 66.67)
Broadband	Knn	66.10 (55.17; 76.67)	61.02 (86.21; 36.67)	62.71 (72.41; 53.33)
	LDA	67.80 (75.86; 60.00)	74.58 (82.76; 66.67)	71.19 (75.86; 66.67)
	SVM	61.02 (62.07; 60.00)	66.10 (75.86; 56.67)	61.02 (58.62; 63.33)
	SVMrbf	62.71 (65.52; 60.00)	66.10 (75.86; 56.67)	59.32 (62.07; 56.67)



**Fig. 5.** Selected features in the LDA classifier for HC-mdMCI in the full beta band for GLd (accuracy 81.36%). A–C panels are represent three views (sagittal, coronal and axial) of the selected links. The width and color (black to soft brown) of these links grows proportional to the median of the weights assigned to the links by the classifier across the 10-fold of the cross-validation testing procedure. The size of the ball that represents the node is proportional to the number of links converging at it. In panel D, the upper triangular matrix shows the median of the assigned weights and in the lower triangular matrix represents the number of folds in which a specific link has been selected. The nodes were grouped according to brain lobes: FL – Frontal Lobe; PL – Parietal Lobe; TL – Temporal Lobe; OL – Occipital Lobe; C – Cingulate Cortex. See Supplementary Table 2 for a ranking of the links with the highest weights.

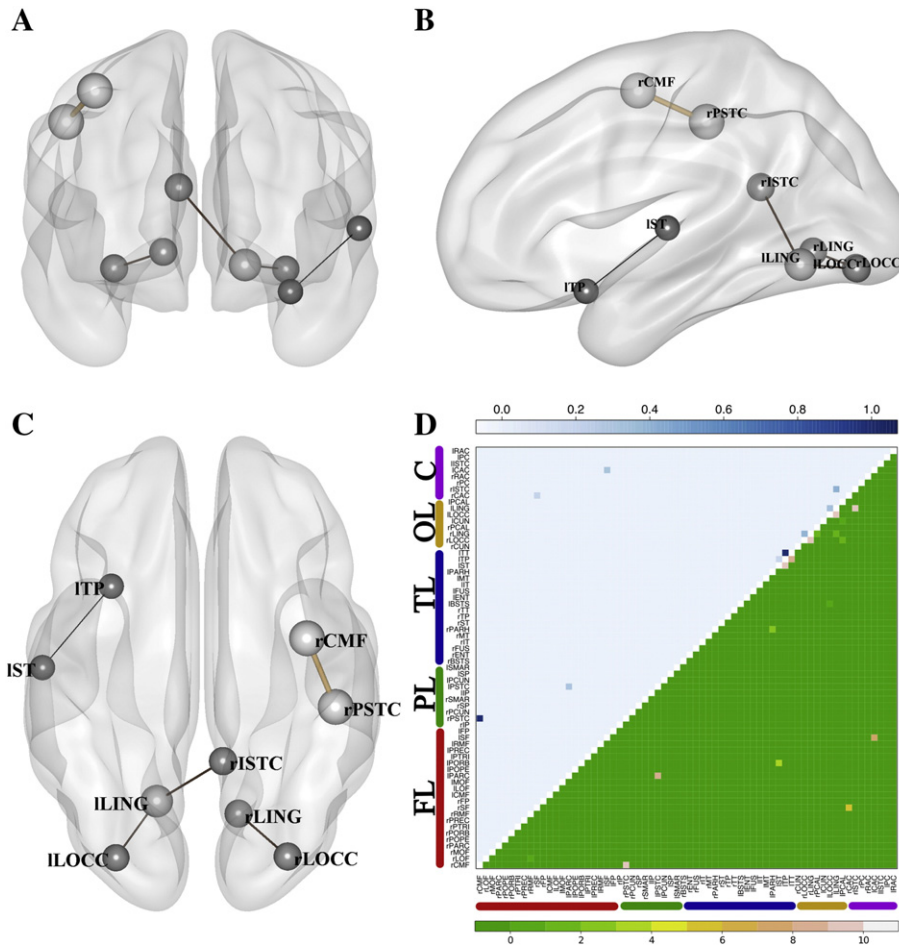
we also provide evidence that the beta band also contributes to the classification, at least in those MCI patients with multiple impaired cognitive domains.

In summary, although the accuracies obtained using SVMrbf were high in general (in particular for the classification between HC and sdMCI), the best accuracies were obtained using LDA, suggesting that

**Table 4**  
10-fold cross-validated accuracies for the classification between sdMCI and mdMCI. The table shows the mean accuracies (specificities; sensitivities) in percentages for all five frequency bands, the three regularization methods (GL, GLa, GLd) and the four classifiers evaluated. Accuracies higher than 80% are highlighted.

sdMCI vs mdMCI		GL	GLa	GLd
alpha	Knn	59.62 (68.18; 53.33)	67.31 (54.55; 76.67)	65.38 (36.36; 86.67)
	LDA	71.15 (63.64; 76.67)	<b>82.69 (63.64; 96.67)</b>	78.85 (54.55; 96.67)
	SVM	57.69 (4.55; 96.67)	69.23 (59.09; 76.67)	63.46 (59.09; 66.67)
	SVMrbf	63.46 (18.18; 96.67)	76.92 (72.73; 80.00)	63.46 (27.27; 90.00)
Low beta	Knn	69.23 (50.00; 83.33)	67.31 (81.82; 56.67)	67.31 (77.27; 60.00)
	LDA	67.31 (45.45; 83.33)	69.23 (36.36; 93.33)	69.23 (54.55; 80.00)
	SVM	65.38 (50.00; 76.67)	59.62 (13.64; 93.33)	61.54 (22.73; 90.00)
	SVMrbf	65.38 (36.36; 86.67)	69.23 (59.09; 76.67)	61.54 (36.36; 80.00)
High beta	Knn	65.38 (45.45; 80.00)	63.46 (100.00; 36.67)	57.69 (90.91; 33.33)
	LDA	71.15 (68.18; 73.33)	69.23 (77.27; 63.33)	71.15 (72.73; 70.00)
	SVM	71.15 (36.36; 96.67)	67.31 (68.18; 66.67)	63.46 (50.00; 73.33)
	SVMrbf	67.31 (54.55; 76.67)	65.38 (27.27; 93.33)	61.54 (13.64; 96.67)
Full beta	Knn	63.46 (45.45; 76.67)	59.62 (40.91; 73.33)	69.23 (50.00; 83.33)
	LDA	67.31 (50.00; 80.00)	67.31 (36.36; 90.00)	69.23 (45.45; 86.67)
	SVM	63.46 (36.36; 83.33)	61.54 (18.18; 93.33)	65.38 (50.00; 76.67)
	SVMrbf	63.46 (36.36; 83.33)	59.62 (4.55; 100.00)	61.54 (9.09; 100.00)
Broadband	Knn	78.85 (72.73; 83.33)	71.15 (68.18; 73.33)	71.15 (59.09; 80.00)
	LDA	<b>84.62 (81.82; 86.67)</b>	78.85 (68.18; 86.67)	<b>80.77 (68.18; 90.00)</b>
	SVM	67.31 (54.55; 76.67)	63.46 (45.45; 76.67)	65.38 (45.45; 80.00)
	SVMrbf	78.85 (72.73; 83.33)	73.08 (50.00; 90.00)	76.92 (77.27; 76.67)





**Fig. 6.** Selected features in the LDA classifier for sdMCI-mdMCI in broadband and with GL (accuracy 84.62%). Panels A-C represent three views (sagittal, coronal and axial) of the selected links. The width and color (black to soft brown) of these links grows proportional to the median of the weights assigned to the links by the classifier across the 10-fold of the cross-validation testing procedure. The size of the ball that represents the node is proportional to the number of links converging at it. In panel D, the upper triangular matrix shows the median of the assigned weights and the lower triangular matrix represents the number of folds in which a specific link has been selected. The nodes were grouped according to brain lobes: FL – Frontal Lobe; PL – Parietal Lobe; TL – Temporal Lobe; OL – Occipital Lobe; C – Cingulate Cortex. See Supplementary Table 2 for a ranking of the links with the highest weights.

linearity is the best, most robust choice in this context. When it comes to the frequency contents, the alpha and broadband datasets got the best accuracies in the discrimination of sdMCI subjects, and the beta band gave the best performance in the classification between HC and mdMCI.

Besides the FC classification accuracies, we included the accuracies obtained from the power of the Hilbert-envelope time-series (see Supplementary Table 3) and from the FD matrices (see Supplementary Table 4). None of these results were better than 80% accuracy. The best results were obtained using the power of the envelopes in the broadband data as input features, reaching 78.43% (spec. 72.41%; sens. 86.36%) for the discrimination between HC and sdMCI. When using the FD matrices, the best results were 70.59% (spec. 72.41%; sens. 68.18%) for HC-sdMCI and 71.19% (68.97%; 73.33%) for HC-mdMCI.

**Discussion**

In this paper, we have used three different varieties of the graphical lasso (Friedman et al., 2008), the non-adaptive graphical lasso, the functional adaptive graphical lasso, and the structural adaptive graphical lasso, to obtain sparse precision matrices explaining the direct functional connectivity between pairs of brain regions. To our knowledge, this is the first time that an adaptive sparse estimation has been applied to MEG data. The three approaches were evaluated in terms of log-likelihood, the network density, and the performance in the classification between three groups: healthy controls, amnesic MCI with a single (sdMCI) or multiple domains (mdMCI) affected.

In summary, we observed that, by including structural soft constraints, the classification accuracies were improved between MCI and HC groups in most cases. The best classification accuracies were obtained for the alpha and broadband datasets when discriminating sdMCI subjects. For HC and mdMCI, the full beta band achieved the best classification. Among the functional connections that were selected to discriminate between sdMCI and HC, we found several connections to temporal regions and to the posterior part of the cingulate cortex (Davatzikos et al., 2011; Sorg et al., 2007). The engagement of connections to frontal regions in the classification of mdMCI subjects agrees with the symptomatology of this condition (Schroeter et al., 2012).

*The contribution of structural connectivity*

Single-modality biomarkers have been widely employed for the diagnosis of MCI or AD. As an advantage, they rely on simple imaging protocols, requiring less acquisition effort and costs. Nevertheless, we believe that the integration of information from different imaging biomarkers can considerably improve diagnosis and prognosis efficacy. There is evidence that structurally informed functional connectivity provides a more accurate representation of the real transfer of information that takes place in brain networks (Hinze et al., 2014; Ng et al., 2012). These approaches, however, have only been tested in healthy subjects, and have never been applied to the diagnosis of brain diseases. We demonstrate here that the inclusion of structural connectivity in the estimation model of functional connectivity improves the accuracy in

the classification between MCI and HC: up to 3% of improvement for the sdMCI group (see Table 3) and up to 10% of improvement for the mdMCI group (see Table 4). In the classification between sdMCI and mdMCI, however, GL provides the best accuracy, perhaps because the structural connectivity differences are too heterogeneous, reflecting the variety of symptoms of the mdMCI group.

Our initial hypotheses were: first, that using the structural connectivity to guide the estimation of the network would improve the estimation of the functional connectivity; and second, that the classification accuracies between groups would benefit from this multimodal fusion. As mentioned above, the log-likelihoods quantify how well the model (in this case, the estimated precision matrix) describes the data under a certain distribution of probability (in this case, the multivariate Gaussian distribution). In our experiments, the log-likelihood increased when guided by the structural connectivity in comparison to the case of a homogeneous regularization. When comparing the log-likelihoods between the adaptive cases GLa and GLd, the results are mostly indistinguishable. GLa uses the empirical covariance matrix computed from the functional data in order to guide the estimation of the network and, hence, in terms of the ability of describing the functional data, it has a head start. The fact that the log-likelihood is roughly equal for GLd and GLa (and higher than GL) is indeed encouraging, considering that the adaptive penalty of GLa is defined using the same data modality (functional data) that we use for computing the likelihood. This result suggests a genuine link between functional and anatomical connectivity, consistent with findings elsewhere (Cabral et al., 2013; Greicius et al., 2009; Hinne et al., 2014; Honey et al., 2009).

#### Disruption of functional networks in MCI

The classification of MCI using whole-brain connectivity is of growing clinical promise. Previous attempts distinguished between MCI that developed Alzheimer's disease and HC with a 90% of accuracy (Shao et al., 2012) by using just structural connectivity. Here we obtained an accuracy of around 70% in the discrimination between groups when using only SC between cortical regions. This contrast of accuracies might be in part due to differences in the methodological pipeline. For example, we used deterministic tractography, and we did not include sub-cortical connections. Most likely, however, this is attributable to the advanced disease stage of the subjects, who would probably have had considerable cortical and sub-cortical atrophy that decreased the tract density, thus enhancing the discriminability power of the tractography.

In another study using graph-theory metrics from fMRI functional connectivity networks, Wee et al. classified amnesic MCI subjects versus HC with an accuracy of 86% (Wee et al., 2012b). This accuracy was increased to 96% when they concatenated functional and structural connectivity features (Wee et al., 2012c). These two studies had clear limitations: first, the reduced sample size (ten subjects in the sdMCI group) hinders the interpretation of the results; second, it is not clear whether the feature selection step was carried out on the entire data (including the testing data which is used to evaluate the model). In this case, this would result in an overestimation of the obtained accuracies. Also, leave-one-out cross-validation is known to be a high variance estimator of the classification accuracy (Hastie et al., 2009), and 10-fold cross-validation is a more reliable alternative (Kohavi, 1995).

The same group evaluated the classification performance in a similar sample using a group-constrained sparse estimation of fMRI connectivity through  $l_2$ -regularization, achieving accuracies of 84% in the best case (Wee et al., 2014). In a third paper, they use sparse multivariate autoregressive modelling to compute an effective connectivity measure similar to Granger causality (Li et al., 2014). This time they achieved a mean accuracy of 91%, which turned out to be higher than those obtained using full-connected matrices (Pearson correlations). This result suggested that sparse functional networks provide a more effective representation of the whole brain connectivity. Note that fMRI is sensitive

to other factors than brain activity, as the cerebral perfusion quantified as the cerebral blood volume/flow. Hypo-perfusion in regions such as the posterior cingulate has been related to cortical atrophy and cognitive decline in MCI patients (Chen et al., 2011; Lacalle-Auriales et al., 2014). Thus, it is certainly possible that cerebral perfusion could also have contributed to the differences observed with the mentioned fMRI studies.

Note that not all of the patients included in our sample will necessarily develop Alzheimer's dementia (non-converters), and, from those that will develop Alzheimer's dementia (converters), the time to the conversion will likely vary. These two factors, as well as the possible development of other kinds of dementia, should be monitored on this sample, and the classifiers could be used to predict between an imminent development of dementia, a delayed development of dementia, and a return to the healthy condition.

The connections that most frequently distinguished both groups from the healthy control group include regions of the cingulate cortex that agree with those of previous findings in similar populations (Davatzikos et al., 2011). These regions were previously observed to have decreased functional connectivity in resting state networks derived from independent component analysis (Sorg et al., 2007). Connections with regions of frontal cortex were often selected, including caudal middle frontal gyri, rostral middle frontal gyri, frontal poles, and the pars triangularis (Grady et al., 2003). Interestingly, both HC-sdMCI and HC-mdMCI classifications assigned a similar importance to connections with regions of the temporal cortex, including transverse temporal gyri, entorhinal cortices and parahippocampal gyri. This seems reasonable, as the memory cognitive decline that both groups present might be related to a malfunctioning of regions in the temporal cortex. However, the importance in the classification of connections with frontal regions was much higher in the classification between HC-mdMCI. The fact that the mdMCI group presents a major executive impairment in comparison to the sdMCI group can motivate the role of this frontal connections as a clinical biomarker (Schroeter et al., 2012). It has also been reported that connections between frontal and temporal cortices are diminished in Alzheimer's patients (Stam et al., 2009), i.e., at a more advanced disease stage. The medial frontal and temporal regions have been implicated in a "default-mode" network. This network comprises a set of regions with high connectivity during rest (Buckner et al., 2008), which become deactivated when performing any attention-demanding task. The level of activity in this network at rest has been observed to be lowered in individuals at risk of Alzheimer's (Petrella, 2013; Vogelaere et al., 2012; Wang et al., 2013).

In the classification between sdMCI-mdMCI, the maximum accuracy was obtained with ten input features. These connections included regions of the occipital cortex (i.e. lateral occipital and lingual gyri), cingulate cortex (isthmus of cingulate gyrus) and frontal cortex (caudal middle frontal and postcentral gyri). In other studies, the posterior part of the cingulate, including the isthmus cingulate, showed reduced metabolic activity in a set of patients that later developed Alzheimer's dementia (Minoshima et al., 1997). Additionally to the posterior cingulate, the medial temporal lobe and the inferior parietal lobe showed decreased metabolic activity in AD patients (Gusnard and Raichle, 2001). This decreased metabolism has been found to correlate with pathological atrophy of the entorhinal and hippocampal cortices (Huang et al., 2002). According to Zhou et al. (2008), disruption of the anatomical connections from the posterior cingulate and the hippocampus to the whole brain in early AD patients were related to immediate recall memory scores. All these findings indicate that the posterior cingulate cortex is strongly involved in the course of AD, and the fact that connections with the isthmus cingulate are present in the classification between sdMCI-mdMCI might indicate that the connectivity with this region, as revealed by the proposed approaches, could be a useful connectivity-based biomarker to define the risk of converting to dementia.

For the sake of comparison, we have included classifiers using the power of the Hilbert envelopes time-series as input features. We

observed that, in general, the accuracies were lower than those obtained using the graphical lasso approaches.

#### Methodological issues and limitations

We should note a few limitations of this study. First, it is still unknown whether the MCI patients of this sample will convert to AD. This is important, as the classifier could lose predictive accuracy due to high heterogeneity within the groups. As an example, we evaluated this heterogeneity in terms of MMSE, as it seems that it could be a differentiating factor between groups (see Table 1). The misclassified subjects, however, were not significantly different from the rest of the group (i.e. the correctly classified subjects) according to a non-parametrical Mann-Whitney statistical comparison ( $p > 0.05$ ) for any of the classifications.

Second, we have restricted ourselves to the alpha and beta bands. The selection of the frequency bands was made according to previous findings in MEG resting-state networks (Brookes et al., 2011) and to previous MEG literature in MCI (Garcés et al., 2013; Ishii et al., 2010). However, the impact of introducing structural priors for estimating functional connectivity in other classical frequency bands, such as delta (0.5–4Hz), theta (4–8Hz) and gamma (30–50Hz), was not evaluated and is still interesting. Future work should assess thoroughly the effect of structural priors in these frequency bands as well as the corresponding classification accuracies. The paper is also restricted to stationary measures of functional connectivity. Recent work using Hidden Markov Models (HMMs) has shown MEG resting state networks switching on very fast time-scales, on the order of 200 ms (Baker et al., 2014). Future work could assess the use of structural priors to constrain the cross-region interactions in the HMM observation models.

Third, the quantification of structural connectivity is prone to errors. Fiber crossing, bending or kissing are unresolved issues when using single DTI as a model of water diffusion. More complex models, such as q-ball (Tuch, 2004) or spherical deconvolution (Tournier et al., 2004) are potentially useful alternatives. However, the limited number of encoding directions in our data hinders this estimation. Future MRI protocols will include higher angular resolution diffusion schemes in order to achieve a better characterization of the structural connectivity.

Fourth, no standardized brain parcellation is fully agreed by the scientific community. There is evidence that different parcellation schemes give rise to different network topographies (Zalesky et al., 2010). The effects that this parcellation would have on the integration of structural and functional information is still an open question. Here, we employed a cortical anatomical parcellation of sixty-six regions which has been previously employed in multi-modal whole-brain network studies (Hagmann et al., 2008; Honey et al., 2009). In future analyses, we will test the impact of structural connectivity on the functional connectivity for different functional or anatomical parcellation schemes.

Fifth, we have not applied any leakage correction. Although there may be some benefits to using leakage correction due to removal of false positives (Maldjian et al., 2014), state-of-the-art methods are over-conservative in the sense that they remove all the zero-lag correlations (including the genuine ones), resulting in an (unknown) extent of true positives decrement. For this reason we considered that standard leakage correction could have unpredictable results in the context of precision matrix estimation and therefore chose not to apply it in these experiments. The inclusion of a structural prior, which is not contaminated by leakage, could have mitigated this effect as well. Note also that, despite this caveat (which is equally applicable to all the methods we compared), the results of the classification were still quite accurate. In the future, we will investigate the trade-off between true and false positives decrement by leakage correction, and, potentially, we will consider how to integrate leakage correction within the precision matrix estimation, perhaps by an inclusion of another adaptive penalty where closer links are more heavily penalised. This could be considered as a

softer (less conservative) alternative to the usual orthogonalizations (Brookes et al., 2012; Maldjian et al., 2014).

Finally, the sample size of this study, although higher than previous investigations on HC-MCI discrimination (Shao et al., 2012; Wee et al., 2014, 2012c), is still moderate. Nevertheless, the presented results remain promising, and illustrate the potential of combining different modalities, for which we provide a simple but efficient method to carry through this integration.

#### Conclusion

In this paper, we provide for the first time a multimodal integration of MEG-DWI for the estimation of sparse whole-brain networks. This integration has been compared in terms of classification accuracy between HC and MCI to two single-modality approaches: using the power of the Hilbert envelope time-series and single-modality SC. The conclusion is that GL-based techniques for estimating FC can yield higher accuracies than single modality approaches. The inclusion of SC in our FC estimation further improves the accuracies for HC-MCI discrimination but not for MCI subgroups classification. The reported accuracies are analogous to other results obtained from samples of amnesic MCI subjects using whole brain fMRI connectivity approaches (Wee et al., 2014, 2012a,b). We conclude that whole brain MEG connectivity is a powerful biomarker of MCI and that the inclusion of SC increases the classification potential. Future studies with larger datasets and clinical follow-up will contribute to validate the effectiveness of the proposed integration.

#### Acknowledgments

This work was supported in part by funds from the Spanish Ministry of Economy and Competitiveness (PSI2012-38375-C03-01) and the Spanish Ministry of Science (PSI2009-14415-C03-01). The authors are grateful to the research participants and their families for their participation in this study. The authors are also grateful to María Eugenia López and Sara Aurretetxe for the acquisition of the resting-state MEG data. José Angel Pineda-Pardo was supported by the Spanish Ministry of Education through the National Program FPU (AP2010-1317). Diego Vidaurre and Mark Woolrich were supported by a Wellcome Trust Strategic Award (WT103552). Mark Woolrich was also supported by an Equipment Grant from the Wellcome Trust (092753/Z/10/Z), the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre based at Oxford University Hospitals Trust Oxford University, and by an MRC UK MEG Partnership Grant (MR/K005464/1).

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.neuroimage.2014.08.002>.

#### References

- Atluri, G., Padmanabhan, K., Fang, G., Steinbach, M., Petrella, J.R., Lim, K., Macdonald, A., Samatova, N.F., Doraiswamy, P.M., Kumar, V., 2013. Complex biomarker discovery in neuroimaging data: Finding a needle in a haystack. *NeuroImage Clin.* 3, 123–131.
- Auer, S., Reisberg, B., 1997. The GDS/FAST staging system. *Int. Psychogeriatr.* 9 (Suppl. 1), 167–171.
- Bajo, R., Maestú, F., Nevado, A., Sancho, M., Gutiérrez, R., Campo, P., Castellanos, N.P., Gil, P., Moratti, S., Pereda, E., Del-Pozo, F., 2010. Functional connectivity in mild cognitive impairment during a memory task: implications for the disconnection hypothesis. *J. Alzheimers Dis.* 22, 183–193.
- Bajo, R., Castellanos, N.P., Cuesta, P., Aurretetxe, S., Garcia-Prieto, J., Gil-Gregorio, P., del-Pozo, F., Maestu, F., 2012. Differential patterns of connectivity in progressive mild cognitive impairment. *Brain Connect.* 2, 21–24.
- Baker, A.P., Brookes, M.J., Redek, I.A., Smith, S.S., Behrens, T., Probert Smith, P.J., Woolrich, M.W., 2014. Fast transient networks in spontaneous human brain activity. *Elife* 3, e01867.
- Binnewijzend, M.A.A., Schoonheim, M.M., Sanz-Arigitia, E., Wink, A.M., van der Flier, W.M., Tolboom, N., Adriaanse, S.M., Damoiseaux, J.S., Scheltens, P., van Berckel, B.N.M.,

- Barkhof, F., 2012. Resting-state fMRI changes in Alzheimer's disease and mild cognitive impairment. *Neurobiol. Aging* 33, 2018–2028.
- Brookes, M.J., Woolrich, M., Luckhoo, H., Price, D., Hale, J.R., Stephenson, M.C., Barnes, G.R., Smith, S.M., Morris, P.G., 2011. Investigating the electrophysiological basis of resting state networks using magnetoencephalography. *Proc. Natl. Acad. Sci. U. S. A.* 108, 16783–16788.
- Brookes, M.J., Woolrich, M.W., Barnes, G.R., 2012. Measuring functional connectivity in MEG: A multivariate approach insensitive to linear source leakage. *NeuroImage* 63, 910–920.
- Buckner, R.L., Andrews-Hanna, J.R., Schacter, D.L., 2008. The brain's default network: anatomy, function, and relevance to disease. *Ann. N. Y. Acad. Sci.* 1124, 1–38.
- Buldú, J.M., Bajo, R., Maestú, F., Castellanos, N., Leyva, I., Gil, P., Sendiña-Nadal, I., Almendral, J.A., Nevado, A., del-Pozo, F., Boccaletti, S., 2011. Reorganization of functional networks in mild cognitive impairment. *PLoS ONE* 6, e19584.
- Busse, A., Hensel, A., Gühne, U., Angermeyer, M.C., Riedel-Heller, S.G., 2006. Mild cognitive impairment: long-term course of four clinical subtypes. *Neurology* 67, 2176–2185.
- Cabral, J., Luckhoo, H., Woolrich, M., Joensson, M., Mohseni, H., Baker, A., Kringelbach, M.L., Deco, G., 2013. Exploring mechanisms of spontaneous functional connectivity in MEG: How delayed network interactions lead to structured amplitude envelopes of band-pass filtered oscillations. *NeuroImage* 90C, 423–435.
- Castellanos, F.X., Di Martino, A., Craddock, R.C., Mehta, A.D., Milham, M.P., 2013. Clinical applications of the functional connectome. *NeuroImage* 80, 527–540.
- Chen, Y., Wolk, D.A., Reddin, J.S., Korczykowski, M., Martinez, P.M., Musiek, E.S., Newberg, A.B., Julin, P., Arnold, S.E., Greenberg, J.H., Detre, J.A., 2011. Voxel-level comparison of arterial spin-labeled perfusion MRI and FDG-PET in Alzheimer disease. *Neurology* 77, 1977–1985.
- Cortes, C., Vapnik, V., 1995. Support-vector networks. *Mach. Learn.* 20 (3), 273–297.
- Daianu, M., Jahanshad, N., Nir, T.M., Toga, A.W., Jack, C.R., Weiner, M.W., Thompson for the Alzheimer's Disease Neuroimaging Initiative, P.M., 2013. Breakdown of brain connectivity between normal aging and Alzheimer's disease: a structural k-core network analysis. *Brain Connect.* 3, 407–422.
- Davatzikos, C., Bhatt, P., Shaw, L.M., Batmanghelich, K.N., Trojanowski, J.Q., 2011. Prediction of MCI to AD conversion, via MRI, CSF biomarkers, and pattern classification. *Neurobiol. Aging* 32.
- Deco, G., Ponce-Alvarez, A., Mantini, D., Romani, G.L., Hagmann, P., Corbetta, M., 2013. Resting-state functional connectivity emerges from structurally and dynamically shaped slow linear fluctuations. *J. Neurosci.* 33, 11239–11252.
- Delbeck, X., Linden, M., Van Der, Collette, F., 2003. Alzheimer's disease as a disconnection syndrome. 13.
- Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Ségonne, F., Salat, D.H., Busa, E., Seidman, L.J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B., Dale, A.M., 2004. Automatically parcellating the human cerebral cortex. *Cereb. Cortex* 14, 11–22.
- Friedman, J., Hastie, T., Tibshirani, R., 2008. Sparse inverse covariance estimation with the graphical lasso. *Biostatistics* 9, 432–441.
- Friston, K.J., 2011. Functional and effective connectivity: a review. *Brain Connect.* 1, 13–36.
- Friston, K.J., Harrison, L., Penny, W., 2003. Dynamic causal modelling. *NeuroImage* 19, 1273–1302.
- Garcés, P., Vicente, R., Wibrál, M., Pineda-Pardo, J.A., López, M.E., Aurtentex, S., Marcos, A., de Andrés, M.E., Yus, M., Sancho, M., Maestú, F., Fernández, A., 2013. Brain-wide slowing of spontaneous alpha rhythms in mild cognitive impairment. *Front. Aging Neurosci.* 5, 1–7.
- García-Marin, V., Blazquez-Llorca, L., Rodríguez, J.-R., Boluda, S., Muntane, G., Ferrer, I., Defelipe, J., 2009. Diminished perisomatic GABAergic terminals on cortical neurons adjacent to amyloid plaques. *Front. Neuroanat.* 3, 28.
- Geyer, C.J., 1992. Practical Markov chain Monte Carlo. *Stat. Sci.* 7, 473–483.
- Grady, C.L., McIntosh, A.R., Beig, S., Keightley, M.L., Burian, H., Black, S.E., 2003. Evidence from functional neuroimaging of a compensatory prefrontal network in Alzheimer's disease. *J. Neurosci.* 23 (3), 986–993.
- Greicius, M.D., Supekar, K., Menon, V., Dougherty, R.F., 2009. Resting-state functional connectivity reflects structural connectivity in the default mode network. *Cereb. Cortex* 19, 72–78.
- Grundman, M., Petersen, R.C., Ferris, S.H., Thomas, R.G., Aisen, P.S., Bennett, D.A., Foster, N.L., Jack, C.R., Galasko, D.R., Doody, R., Kaye, J., Sano, M., Mohs, R., Gauthier, S., Kim, H.T., Jin, S., Schultz, A.N., Schafer, K., Mulnard, R., van Dyck, C.H., Mintzer, J., Zamrini, E.Y., Cahn-Weiner, D., Thal, L.J., 2004. Mild cognitive impairment can be distinguished from Alzheimer disease and normal aging for clinical trials. *Arch. Neurol.* 61 (1), 59–66.
- Guo, Y., Hastie, T., Tibshirani, R., 2007. Regularized linear discriminant analysis and its application in microarrays. *Biostatistics* 8, 86–100.
- Gusnard, D.A., Raichle, M.E., 2001. Searching for a baseline: functional imaging and the resting human brain. *Nat. Rev. Neurosci.* 2, 685–694.
- Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C.J., Wedeen, V.J., Sporns, O., 2008. Mapping the structural core of human cerebral cortex. *PLoS Biol.* 6, e159.
- Haimovici, A., Tagliazucchi, E., Balenzuela, P., Chialvo, D.R., 2013. Brain organization into resting state networks emerges at criticality on a model of the human connectome. *Phys. Rev. Lett.* 110, 178101.
- Hall, E.L., Woolrich, M.W., Thomaz, C.E., Morris, P.G., Brookes, M.J., 2013. Using variance information in magnetoencephalography measures of functional connectivity. *NeuroImage* 67, 203–212.
- Hastie, T., Tibshirani, R., Friedman, J., 2009. The elements of statistical learning: data mining, inference, and prediction, 2nd ed. Springer, Berlin Heidelberg.
- Hinne, M., Ambrogioni, L., Janssen, R.J., Heskes, T., van Gerven, M.A.J., 2014. Structurally-informed Bayesian functional connectivity analysis. *NeuroImage* 86, 294–305.
- Honey, C.J., Kötter, R., Breakspear, M., Sporns, O., 2007. Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proc. Natl. Acad. Sci. U. S. A.* 104, 10240–10245.
- Honey, C.J., Sporns, O., Cammoun, L., Gigandet, X., Thiran, J.P., Meuli, R., Hagmann, P., 2009. Predicting human resting-state functional connectivity. *Proc. Natl. Acad. Sci. U. S. A.* 106, 2035–2040.
- Huang, C., Wahlund, L.-O., Svensson, L., Winblad, B., Julin, P., 2002. Cingulate cortex hypoperfusion predicts Alzheimer's disease in mild cognitive impairment. *BMC Neurol.* 12, 2–9.
- Ishii, R., Canuet, L., Kurimoto, R., Ikezawa, K., Aoki, Y., Azechi, M., Takahashi, H., Nakahachi, T., Iwase, M., Kazui, H., Takeda, M., 2010. Frontal shift of posterior alpha activity is correlated with cognitive impairment in early Alzheimer's disease: a magnetoencephalography-beamformer study. *Psychogeriatrics* 10, 138–143.
- Johansen-Berg, H., Behrens, T.E.J., Robson, M.D., Drobnyak, I., Rushworth, M.F.S., Brady, J.M., Smith, S.M., Higham, D.J., Matthews, P.M., 2004. Changes in connectivity profiles define functionally distinct regions in human medial frontal cortex. *Proc. Natl. Acad. Sci. U. S. A.* 101, 13335–13340.
- Kohavi, R., 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. *International Joint Conference on Artificial Intelligence*, pp. 1137–1143.
- Lacalle-Auriales, M., Mateos-Pérez, J.M., Guzmán-De-Villoria, J.A., Olazarán, J., Cruz-Orduña, I., Alemán-Gómez, Y., Martino, M.-E., Descio, M., 2014. Cerebral blood flow is an earlier indicator of perfusion abnormalities than cerebral blood volume in Alzheimer's disease. *J. Cereb. Blood Flow Metab.* 654–659.
- Lawton, M.P., Brody, E.M., 1969. Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist* 9, 179–186.
- Lazar, M., Weinstein, D.M., Tsuruda, J.S., Hasan, K.M., Arfanakis, K., Meyerand, M.E., Badie, B., Rowley, H.A., Haughton, V., Field, A., Alexander, A.L., 2003. White matter tractography using diffusion tensor deflection. *Hum. Brain Mapp.* 18, 306–321.
- Li, Y., Wee, C.-Y., Jie, B., Peng, Z., Shen, D., 2014. Sparse multivariate autoregressive modeling for mild cognitive impairment classification. *Neuroinformatics* 12 (3), 455–469.
- Lobo, A., Saz, P., Marcos, G., Día, J.L., de la Cámara, C., Ventura, T., Morales Asín, F., Fernando Pascual, L., Montañés, J.A., Aznar, S., 1999. Revalidation and standardization of the cognition mini-exam (first Spanish version of the Mini-Mental Status Examination) in the general geriatric population. *Med. Clin. (Barc.)* 112, 767–774.
- Luckhoo, H., Hale, J.R., Stokes, M.G., Nobre, A.C., Morris, P.G., Brookes, M.J., Woolrich, M.W., 2012. Inferring task-related networks using independent component analysis in magnetoencephalography. *NeuroImage* 62, 530–541.
- Maldjian, J.A., Davenport, E.M., Whitlow, C.T., 2014. Graph theoretical analysis of resting-state MEG data: Identifying interhemispheric connectivity and the default mode. *NeuroImage* 96, 88–94.
- Marrelec, G., Krainik, A., Duffau, H., Péligrini-Issac, M., Léhericy, S., Doyon, J., Benali, H., 2006. Partial correlation for functional brain interactivity investigation in functional MRI. *NeuroImage* 32, 228–237.
- Medina, D., DeToledo-Morrell, L., Urresta, F., Gabrieli, J.D.E., Moseley, M., Fleischman, D., Bennett, D.A., Leurgans, S., Turner, D.A., Stebbins, G.T., 2006. White matter changes in mild cognitive impairment and AD: A diffusion tensor imaging study. *Neurobiol. Aging* 27, 663–672.
- Minoshima, S., Giordani, B., Berent, S., Frey, K.A., Foster, N.L., Kuhl, D.E., 1997. Metabolic reduction in the posterior cingulate cortex in very early Alzheimer's disease. *Ann. Neurol.* 42, 85–94.
- Mohseni, H.R., Woolrich, M.W., Kringelbach, M.L., Luckhoo, H., Smith, P.P., Aziz, T.Z., 2012. Fusion of magnetometer and gradiometer sensors of MEG in the presence of multiplicative error. *IEEE Trans. Biomed. Eng.* 59 (7), 1951–1961.
- Mosher, J.C., Leahy, R.M., Lewis, P.S., 1999. EEG and MEG: forward solutions for inverse methods. *IEEE Trans. Biomed. Eng.* 46, 245–259.
- Ng, B., Varoquaux, G., Poline, J.-B., Thirion, B., 2012. A novel sparse graphical approach for multimodal brain connectivity inference. *Med. Image Comput. Comput. Assist. Interv.* 15, 707–714.
- O'Dwyer, L., Lamberton, F., Bokde, A.L.W., Ewers, M., Faluy, Y.O., Tanner, C., Mazoyer, B., O'Neill, D., Bartley, M., Collins, D.R., Coughlan, T., Prvulovic, D., Hampel, H., 2011. Multiple indices of diffusion identifies white matter damage in mild cognitive impairment and Alzheimer's disease. *PLoS ONE* 6, e21745.
- Petersen, R.C., 2004. Mild cognitive impairment as a diagnostic entity. *J. Intern. Med.* 256, 183–194.
- Petersen, R.C., 2011. Mild cognitive impairment. *N. Engl. J. Med.* 364, 2227–2234.
- Petrella, J.R., 2013. Neuroimaging and the search for a cure for Alzheimer disease. *Radiology* 269, 671–691.
- Pfeffer, R.I., Kurosaki, T.T., Harrah, C.H., Chance, J.M., Filos, S., 1982. Measurement of functional activities in older adults in the community. *J. Gerontol.* 37, 323–329.
- Pineda-Pardo, J.A., Garcés, P., López, M.E., Aurtentex, S., Cuesta, P., Marcos, A., Montejó, P., Yus, M., Hernández-Tamames, J.A., Del-Pozo, F., Becker, J.T., Maestu, F., 2014. White matter damage disorganizes brain functional networks in amnesic Mild Cognitive Impairment. *Brain Connect.* 4 (5), 312–322.
- Reisberg, B., Ferris, S.H., de Leon, M.J., Crook, T., 1982. The Global Deterioration Scale for assessment of primary degenerative dementia. *Am. J. Psychiatry* 139, 1136–1139.
- Rosen, W.G., Terry, R.D., Fuld, P.A., Katzman, R., Peck, A., 1980. Pathological verification of ischemic score in differentiation of dementias. *Ann. Neurol.* 7, 486–488.
- Schroeter, M.L., Vogt, B., Frisch, S., Becker, G., Barthel, H., Mueller, K., Villringer, A., Sabri, O., 2012. Executive deficits are related to the inferior frontal junction in early dementia. *Brain* 135, 201–215.
- Sekihara, K., Nagarajan, S., Poeppel, D., Miyashita, Y., 2001. Reconstructing spatio-temporal activities of neural sources from magnetoencephalographic data using a vector beamformer. *IEEE Int. Conf. Acoust. Speech, Signal Process. Proc. (Cat. No.01CH37221)*, 3).
- Selkoe, D.J., 2002. Alzheimer's disease is a synaptic failure. *Science* 298, 789–791.
- Shao, J., Myers, N., Yang, Q., Feng, J., Plant, C., Böhm, C., Förstl, H., Kurz, A., Zimmer, C., Meng, C., Riedl, V., Wohlschläger, A., Sorg, C., 2012. Prediction of Alzheimer's disease using individual structural connectivity networks. *Neurobiol. Aging* 33, 2756–2765.

- Smith, S.M., 2002. Fast robust automated brain extraction. *Hum. Brain Mapp.* 17, 143–155.
- Smith, S.M., Miller, K.L., Salimi-Khorshidi, G., Webster, M., Beckmann, C.F., Nichols, T.E., Ramsey, J.D., Woolrich, M.W., 2011. Network modelling methods for fMRI. *NeuroImage* 54, 875–891.
- Sorg, C., Riedl, V., Mühlau, M., Calhoun, V.D., Eichele, T., Läer, L., Drzezga, A., Förstl, H., Kurz, A., Zimmer, C., Wohlschläger, A.M., 2007. Selective changes of resting-state networks in individuals at risk for Alzheimer's disease. *Proc. Natl. Acad. Sci. U. S. A.* 104, 18760–18765.
- Stam, C., 2003. EEG synchronization in mild cognitive impairment and Alzheimer's disease. *Acta Neurol.* 108 (92), 90–96.
- Stam, C.J., de Haan, W., Daffertshofer, A., Jones, B.F., Manshanden, I., van Cappellen, van Walsum, A.M., Montez, T., Verbunt, J.P.A., de Munck, J.C., van Dijk, B.W., Berendse, H.W., Scheltens, P., 2009. Graph theoretical analysis of magnetoencephalographic functional connectivity in Alzheimer's disease. *Brain* 132, 213–224.
- Stephan, K.E., Tittgemeyer, M., Knösche, T.R., Moran, R.J., Friston, K.J., 2009. Tractography-based priors for dynamic causal models. *NeuroImage* 47, 1628–1638.
- Taulu, S., Kajola, M., 2005. Presentation of electromagnetic multichannel data: The signal space separation method. *J. Appl. Phys.* 97, 124905.
- Toga, A.W., Thompson, P.M., 2013. Connectomics sheds new light on Alzheimer's disease. *Biol. Psychiatry* 73, 390–392.
- Tournier, J.-D., Calamante, F., Gadian, D.G., Connelly, A., 2004. Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. *NeuroImage* 23, 1176–1185.
- Tuch, D.S., 2004. Q-ball imaging. *Magn. Reson. Med.* 52, 1358–1372.
- Valdés-Sosa, P.A., Sánchez-Bornot, J.M., Lage-Castellanos, A., Vega-Hernández, M., Bosch-Bayard, J., Melie-García, L., Canales-Rodríguez, E., 2005. Estimating brain functional connectivity with sparse multivariate autoregression. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 969–981.
- Van den Heuvel, M.P., Mandl, R.C.W., Kahn, R.S., Hulshoff Pol, H.E., 2009. Functionally linked resting-state networks reflect the underlying structural connectivity architecture of the human brain. *Hum. Brain Mapp.* 30, 3127–3141.
- Van Veen, B.D., van Drongelen, W., Yuchtman, M., Suzuki, A., 1997. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Trans. Biomed. Eng.* 44, 867–880.
- Vidaurre, D., Bielza, C., Larrañaga, P., 2013. A survey of L1 regression. *Int. Stat. Rev.* 81 (3), 361–387.
- Vogelaere, F., Santens, P., Achten, E., Boon, P., Vingerhoets, G., Vogelaere, F., Santens, P., Achten, E., Boon, P., Vingerhoets, G., 2012. Altered default-mode network activation in mild cognitive impairment compared with healthy aging. *Neuroradiology* 54 (11), 1195–1206.
- Vrba, J., Robinson, S.E., 2000. Linearly constrained minimum variance beamformers, synthetic aperture magnetometry, and MUSIC in MEG applications. *Conf. Rec. Thirty-fourth asilomar conf. Signals, Syst. Comput.*, p. 1 (Cat. No.00CH37154).
- Wang, J., Zuo, X., Dai, Z., Xia, M., Zhao, Z., Zhao, X., Jia, J., Han, Y., He, Y., 2013a. Disrupted functional brain connectome in individuals at risk for Alzheimer's disease. *Biol. Psychiatry* 73, 472–481.
- Wang, L., Li, H., Liang, Y., Zhang, J., Li, X., Shu, N., Wang, Y.Y., Zhang, Z., 2013b. Amnesic mild cognitive impairment: topological reorganization of the default-mode network. *Radiology* 268, 501–514.
- Wee, C.-Y., Shen, D., 2012a. Resting-state multi-spectrum functional connectivity networks for identification of MCI patients. *PLoS ONE* 7, e37828.
- Wee, C.-Y., Yap, P.-T., Zhang, D., Denny, K., Browndyke, J.N., Potter, G.G., Welsh-Bohmer, K.A., Wang, L., Shen, D., 2012b. Identification of MCI individuals using structural and functional connectivity networks. *NeuroImage* 59, 2045–2056.
- Wee, C.-Y., Yap, P., Zhang, D., Wang, L., Shen, D., 2012c. Constrained sparse functional connectivity networks for MCI classification. *Med. Image Comput. Comput. Assist. Interv.* 15, 212–219.
- Wee, C.-Y., Yap, P.-T., Zhang, D., Wang, L., Shen, D., 2014. Group-constrained sparse fMRI connectivity modeling for mild cognitive impairment identification. *Brain Struct. Funct.* 219 (2), 641–656.
- Witten, D.M., Friedman, J.H., Simon, N., 2011. New insights and faster computations for the graphical lasso. *J. Comput. Graph. Stat.* 20 (4), 892–900.
- Woolrich, M.W., Stephan, K.E., 2013. Biophysical network models and the human connectome. *NeuroImage* 80, 330–338.
- Woolrich, M., Hunt, L., Groves, A., Barnes, G., 2011. MEG beamforming using Bayesian PCA for adaptive data covariance matrix regularization. *NeuroImage* 57, 1466–1479.
- Yesavage, J.A., Brink, T.L., Rose, T.L., Lum, O., Huang, V., Adey, M., Leirer, V.O., 1982. Development and validation of a geriatric depression screening scale: A preliminary report. *J. Psychiatr. Res.* 17 (1), 37–49.
- Zalesky, A., Fornito, A., Harding, I.H., Cocchi, L., Yücel, M., Pantelis, C., Bullmore, E.T., 2010. Whole-brain anatomical networks: does the choice of nodes matter? *NeuroImage* 50, 970–983.
- Zamrini, E., Maestu, F., Pekkonen, E., Funke, M., Makela, J., Riley, M., Bajo, R., Sudre, G., Fernandez, A., Castellanos, N., Del Pozo, F., Stam, C.J., van Dijk, B.W., Bagic, A., Becker, J.T., 2011. Magnetoencephalography as a putative biomarker for Alzheimer's disease. *Int. J. Alzheimers Dis.* 10, 280289.
- Zhou, Y., Dougherty, J.H., Hubner, K.F., Bai, B., Cannon, R.L., Hutson, R.K., 2008. Abnormal connectivity in the posterior cingulate and hippocampus in early Alzheimer's disease and mild cognitive impairment. *Alzheimers Dement.* 4 (4), 265–270.
- Zou, H., 2006. The adaptive lasso and its oracle properties. *J. Am. Stat. Assoc.* 101 (476), 1418–1429.