

# A Probabilistic Framework to Infer Brain Functional Connectivity from Anatomical Connections

Fani Deligianni<sup>1</sup>, Gael Varoquaux<sup>2,3</sup>, Bertrand Thirion<sup>3</sup>, Emma Robinson<sup>1</sup>,  
David J. Sharp<sup>4</sup>, A. David Edwards<sup>5</sup>, and Daniel Rueckert<sup>1</sup>

<sup>1</sup> Department of Computing, Imperial College London, UK

<sup>2</sup> INSERM U922, Neurospin, CEA Saclay, France

<sup>3</sup> Parietal project team, INRIA, Saclay, France

<sup>4</sup> C3NL, The Division of Experimental Medicine, Imperial College London, UK

<sup>5</sup> Institute of Clinical Sciences, Imperial College London, UK

**Abstract.** We present a novel probabilistic framework to learn across several subjects a mapping from brain anatomical connectivity to functional connectivity, i.e. the covariance structure of brain activity. This prediction problem must be formulated as a structured-output learning task, as the predicted parameters are strongly correlated. We introduce a model selection framework based on cross-validation with a parametrization-independent loss function suitable to the manifold of covariance matrices. Our model is based on constraining the conditional independence structure of functional activity by the anatomical connectivity. Subsequently, we learn a linear predictor of a stationary multivariate autoregressive model. This natural parameterization of functional connectivity also enforces the positive-definiteness of the predicted covariance and thus matches the structure of the output space. Our results show that functional connectivity can be explained by anatomical connectivity on a rigorous statistical basis, and that a proper model of functional connectivity is essential to assess this link.

## 1 Introduction

The brain's power and stability depends critically on its connectional complexity. Recently, it has been suggested that consciousness depends on the brain's ability to integrate information among different thalamo-cortical areas [8]. Disruption in cortical connectivity has been also implicated in a number of disorders and pathologies, such as schizophrenia [3], autism [21], brain trauma [23] and so on. As a result there is a shift of research attention from localised brain function to network organization and dynamics.

A functional network is defined based on connections that reflect temporal dependency between spatially remote neurophysiological events. It is well established that during rest the brain shows spontaneous activity that is highly correlated between multiple brain regions. However, the neuronal basis of resting-state

functional Magnetic Resonance Imaging (r-fMRI) oscillations is not yet fully understood [20]. There is an on-going debate on whether physiological processes, such as respiration and cardiac cycle dominate the resting-state BOLD signal. Support for a neuronal basis comes from several studies that investigate the relationship between structural connectivity and functional brain connectivity [12,4,15,14]. The underlying hypothesis is that if r-fMRI has a neuronal basis then it must reflect the existence of structural connections that relate functionally linked brain regions.

To this end, significant evidence has emerged that there are strong structural connections between regions that are functionally linked, forming resting-state networks [12,13]. Furthermore, strong functional connectivity has been observed between areas with direct structural link [14]. However, these studies do not provide a systematic framework to investigate localised influences between structural connections and functional links. [30] presents a joint generative model of functional and structural connectivity based on the assumption that they follow Gaussian distributions.

In this work, we are interested in inferring the link between anatomical and functional connectivity in a data-driven way. We adapt an anatomically informed probabilistic model of functional connectivity. We describe subject-level functional connectivity as a multivariate Gaussian process [29] estimated by imposing a common structure based on the structural connectivity matrices. Subsequently, we use statistical prediction to infer functional connectivity from structural connectivity. Compared to discriminant models [14,5], which does not provide an underlying generative process, the use of a probabilistic framework assists in gaining a deeper understanding of the mechanisms underlying functional connectivity.

## 2 Problem Statement

The aim of our predictive model is to use structural connectivity across subjects to predict functional connectivity. Structural connections in the brain are formed by white matter tracts that directly interconnect large groups of spatially separated neurons via long-distance axons. Here, the structural connectivity  $\mathbf{A}$  is described as an undirected weighted graph that shows whether there is or not a link between each pair of ROIs and its strength. In this framework, each brain's connection is treated as a variable, which results in a total of  $N = n(n - 1)/2$  variables, where  $n$  is the number of ROIs. For each of the functional connection, our goal is to find a subset of structural connections to which it is strongly related.

We consider  $S$  subjects, represented by  $\mathbf{A}^s$ , the subject specific structural connectivity matrix and the correlation matrices between brain time series extracted from  $n$  ROIs:  $\{\boldsymbol{\Sigma}^s \in \mathbb{R}^{n \times n}, s = 1, \dots, M\}$ .

The current practice in studying functional connectivity,  $\mathbf{F}^s$  is based on the estimation of the coefficients of the correlation matrices across subjects. This can be expressed as a univariate additive linear model on the covariance matrix:

$$\Sigma^s = \Sigma^* + \mathbf{d}\Sigma^s \quad (1)$$

where  $\Sigma^*$  is a covariance matrix representative of the group, and  $\mathbf{d}\Sigma^s$  encode subject-specific contributions. Very often some coefficients of the covariance matrix are zeroed out by thresholding, which leads to a non-positive definite matrix, hence, it does not yield any consistent signal model. Therefore, it is not possible to generalize the learned model to new subjects [29]. The simplest well-posed alternative is to model the fMRI time series  $\mathbf{U} \in \mathbb{R}^{n \times r}$  as centered multivariate Gaussian process [29], hence described by a positive definite covariance.

Let us denote  $(y_k)_{k=1 \dots N}$  the strength of functional connections, i.e. model-dependent parameters that characterize the covariance structure. These should be predicted by combinations of an unknown set of structural variables  $(x_j)_{j=1 \dots N}$ . Penalised linear regression finds a sparse solution of the under-determined linear regression problem.

It is crucial to note here that the application of a multiple regression predictive model provides localised, interpretable results. However, it infers each functional connection independently and thus it does not guarantee that the overall prediction of the covariance matrix would be an SPD matrix. The space of SPD matrices,  $Sym_p^+$ , does not form a vector space:  $\mathbf{C}, \mathbf{D} \in Sym_p^+ \not\Rightarrow \mathbf{C} - \mathbf{D} \in Sym_p^+$ . Another issue arises from the fact that the number of functional connectivity parameters is greater than the number of samples:  $n < \frac{1}{2}r(r+1)$ . This results in a large estimation error of the sample covariance matrix [16,29].

*Parameterizations and loss for Functional connectivity.* One of the challenges of predicting functional brain connectivity is that it can be described by many different parameters. Most often, the functional connectivity between a set of regions is described by the correlation matrix,  $\Sigma \in \mathbb{R}^{n \times n}$ , between the time series,  $\mathbf{X} \in \mathbb{R}^{r \times n}$ , of the mean activation in the different regions of interest. However, several studies [27,29] have shown that the precision matrix, the inverse of the covariance,  $\mathbf{K} = \Sigma^{-1}$ , has adequate properties for the estimation of brain functional networks, while Fransson and Marrelec [10] reported that partial correlations identified between-regions interaction better than correlations based on resting-state fMRI. The partial correlation matrix  $\mathbf{P}$  corresponds to the precision matrix  $\mathbf{K}$  renormalized to have unit diagonal. On our multi-subject dataset (described in section 4), we find that on average, the anatomical connectivity between a pair of regions  $i$  and  $j$ ,  $\mathbf{A}_{i,j}$ , explains 12.77% of the variance of the corresponding pair-wise functional correlation,  $\Sigma_{i,j}$ , 13.52% of the precision coefficient  $\mathbf{K}_{i,j}$ , and 14.36% of the partial correlation  $\mathbf{P}_{i,j}$ .

In this context, we choose to formalize this learning problem as a structured-output multivariate regression, suitable for predicting interdependent variables. The task is to learn a mapping  $f$  from a space  $\mathcal{A}$  describing anatomical connectivity matrices to a second space  $\mathcal{F}$  describing functional connectivity. It entails the choice of a parameterization for both input and output spaces, as well as a loss function  $l$  on the output parameterization, giving a measure of the empirical risk associated with the prediction. Typically, in an energy-based formulation [17], learning from training data is achieved by choosing  $f$  to minimize a combination of the loss and a regularization term.

In our case, we use for the input space the anatomical connectivity matrices, that is the space of symmetric matrices  $\mathcal{A} = \text{Sym}_p$ . For the output space, we work on a parametrization of the space of correlation matrices, that is symmetric definite positive matrices  $\mathcal{F} = \text{Sym}_p^+$ .  $\text{Sym}_p^+$  is not a vector space thus the standard Euclidean distance on matrices, the Frobenius norm, is ill suited to quantify errors. However,  $\text{Sym}_p^+$  can be parametrized as a Riemannian manifold using an intrinsic metric [9]:

$$d_{AI}(\mathbf{C}, \mathbf{D})^2 = \text{tr}(\log \mathbf{C}^{-\frac{1}{2}} \mathbf{D} \mathbf{C}^{-\frac{1}{2}})^2 \quad (2)$$

This metric can lead to a full statistical framework on  $\text{Sym}_p^+$  [22,19]. Most importantly, it is invariant under affine scaling and inversion of the matrices, and is thus equivalent for a wide range of parameterization for functional connectivity. In particular, if  $\mathbf{D}$  is the predicted matrix, and  $\mathbf{C}$  the ground truth, it gives the same prediction error on the correlation matrices and on the precision matrices.

Non-structured output predictors, such as independent prediction of each coefficient of the functional connectivity matrix, cannot guarantee that prediction  $\mathbf{F}^{\text{pred}}$  will be positive definite. To quantify their prediction error in a parametrization-independent way, we develop the metric given in Eq. (2) around the target matrix  $\mathbf{F}^{\text{target}}$ , that can be a correlation matrix, a precision matrix, or a partial correlation matrix, but is by construction definite positive (note that  $d$  is no longer a distance however):

$$d_{AI}(\mathbf{F}^{\text{pred}}, \mathbf{F}^{\text{target}}) \sim d(\mathbf{F}^{\text{pred}}, \mathbf{F}^{\text{target}}) = \|(\mathbf{F}^{\text{target}})^{-1}(\mathbf{F}^{\text{target}} - \mathbf{F}^{\text{pred}})\| \quad (3)$$

$\|\cdot\|$  is the Frobenius norm.

### 3 From Structure to Functional Correlations: A Probabilistic Framework

*Generative model* We use multivariate autoregressive models (MAR) to describe the generative process of fMRI time series. If  $\mathbf{x} \in \mathbb{R}^n$  is the multivariate vector of observations at a given time  $r$ ,

$$\mathbf{x}(r+1) = \mathbf{A}\mathbf{x}(r) + \mathbf{e}(r+1) \quad (4)$$

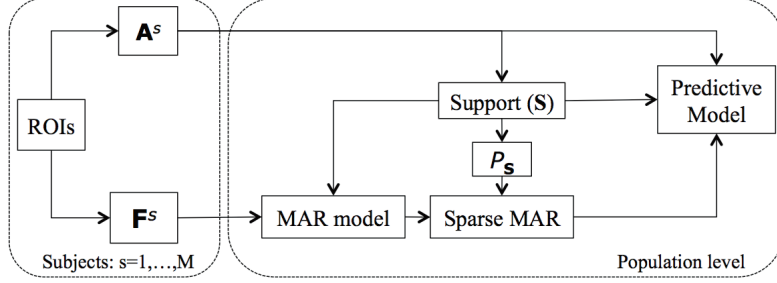
with  $\mathbf{A} \in \mathbb{R}^{n \times n}$  a matrix of the connection between variables and  $\mathbf{e}$  additive Gaussian noise between variables with zero mean and identity covariance.

We consider the ongoing brain activity in resting-state as a stationary process. If  $\mathbf{U} \in \mathbb{R}^{n \times r}$  is the matrix of the observed time series and  $\mathbf{E} \in \mathbb{R}^{n \times r}$  then

$$\mathbf{U} = \mathbf{A}\mathbf{U} + \mathbf{E} \Leftrightarrow \mathbf{U} = (\mathbf{I} - \mathbf{A})^{-1}\mathbf{E} \quad (5)$$

where  $\mathbf{I}$  is the  $n \times n$  identity matrix. Thus, the covariance of the observed time series,  $\mathbf{\Sigma}$ , is given by:

$$\mathbf{\Sigma} = \frac{1}{r} \mathbf{U}\mathbf{U}^T = (\mathbf{I} - \mathbf{A})^{-1} \text{cov} \mathbf{E} (\mathbf{I} - \mathbf{A})^T = ((\mathbf{I} - \mathbf{A})^T (\mathbf{I} - \mathbf{A}))^{-1} \quad (6)$$



**Fig. 1.** Probabilistic framework for predicting functional connectivity from structural brain connectivity.  $\mathbf{A}^s$  and  $\mathbf{F}^s$  represent structural and functional connectivity for subject  $s$ , respectively.  $\mathbf{S}$  represents the support, the sparsity pattern over all structural connectivity matrices.  $P_{\mathbf{S}}$  represents the permutation that provides with a sparser Cholesky factor of the support matrix,  $\mathbf{S}$ .

Note that  $\text{cov}\mathbf{E} = \mathbf{I}$  and  $\mathbf{B} = \mathbf{I} - \mathbf{A}$  is a matrix square root of the inverse covariance, which we call the interaction matrix.

$$\Sigma^{-1} = \mathbf{B}^T \mathbf{B}. \quad (7)$$

Conditional independence between variables is given by the zeros in the precision matrix  $\mathbf{K} = \Sigma^{-1}$ . In order to reduce the large estimation error of the sample covariance matrix, covariance selection is achieved by imposing a sparse support for the estimated precision matrix [29]. We adopt the sparsity pattern of the structural data  $\mathbf{S}$  to represent the support, see Fig. 1.

More precisely, we impose the same support *supp* across individuals, which is the set of supra-threshold t-tests statistics on  $\{\mathbf{A}^s, s = 1, \dots, M\}$ :  $t = \frac{\bar{\mathbf{A}}}{\text{std}(\mathbf{A})\sqrt{M}}$ ; the choice of the threshold is discussed later. We estimate the non-zero coefficients using the iterative proportional scaling (IPS) algorithm [16].

Instead of predicting the precision matrix  $\Sigma^{-1}$  directly, we aim at recovering the interaction matrix  $\mathbf{B}$ . This guarantees that the predicted precision matrix  $\hat{\Sigma}^{-1} = \hat{\mathbf{B}}^T \hat{\mathbf{B}}$  is SPD. Here, we use the Cholesky decomposition of the precision matrix  $\Sigma^{-1}$  to estimate the interaction matrix  $\mathbf{B}$ . Since the Cholesky decomposition is not invariant to a permutation of the rows and columns of the input matrix  $\Sigma^{-1}$ , we adapt a permutation  $P_{\mathbf{S}}(\mathbf{B})$  that tends to yield sparser Cholesky factors  $LU$  of the support matrix  $\mathbf{S}$ . To find the ordering we construct a matrix  $\mathbf{M}$ , such that  $\text{supp}(\mathbf{L}^T \mathbf{L}) = \text{supp}(\mathbf{B})$ , where *supp* represents the support, i.e. the sparsity pattern of the underlying matrix. To reduce the number of non-zero elements of the Cholesky factor the column approximate minimum degree permutation is utilised. Note that we are interested in predicting correlation and not covariance, thus, we rescale the diagonal of  $\mathbf{B}$  to ones by multiplying with the inverse of its diagonal matrix:

$$\tilde{\mathbf{B}} = \mathbf{B} \text{diag}(\mathbf{B})^{-1} \quad (8)$$

---

**Algorithm 1.** Inferring functional from structural connectivity based on the sparse MAR model described in Section 3, Fig. 1. Note that it is embedded in a leave-one-subject-out cross-validation loop, in order to measure the performance of the model.

---

**Input:**  $\{\mathbf{A}^s, s = 1, \dots, M\}$  is the structural connectivity matrix for each subject  $s$ ,  
 $\{\mathbf{F}^s, s = 1, \dots, M\}$  is the observed functional correlation matrix  
**Output:** Predictive model of the  $\{\mathbf{F}^s, s = 1, \dots, M\}$  based on  $\{\mathbf{A}^s, s = 1, \dots, M\}$

- 1: Estimate the support  $\mathbf{S}$  as the set of supra-threshold t-tests statistics on  $\{\mathbf{A}^s, s = 1, \dots, M\}$
- 2: **for**  $s=1$  to  $M$  **do**
- 3:   For all functional correlation matrices estimate the precision matrix  $\mathbf{K}$  based on IPS:  $\mathbf{K}^s = \text{IPS}(\mathbf{F}^s, \mathbf{S})$
- 4:   Find the sparser permutation of the support  $\mathbf{S}$  and re-order  $\mathbf{K}$
- 5:   Estimate the interaction matrix  $\tilde{\mathbf{B}}$  based on the sparse Cholesky decomposition
- 6: **end for**
- 7: **for**  $k=1$  to  $N$  **do**
- 8:   Apply coefficient-wise LASSO to predict  $\tilde{\mathbf{B}}_k^s$  from  $\{\mathbf{A}^s, s = 1, \dots, M\}$
- 9: **end for**
- 10: The output is  $\{\hat{\mathbf{F}}^s = (\hat{\mathbf{B}}^s(\hat{\mathbf{B}}^s)^T)^{-1}, s = 1, \dots, M\}$

---

This corresponds to fixing the variance of innovation terms of the MAR model.

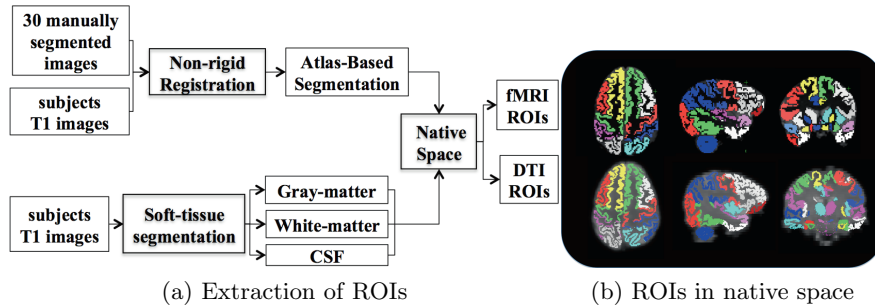
*Statistical Inference.* The inference problem takes the form of multiple regression:

$$y_k = b_{k,0} + \sum_{j=1}^N b_{k,j}x_j, \quad (9)$$

where  $b_0$  is the intercept,  $b_j$  are the coefficients that encode the relationship between each functional connection  $y_k$  and structural connectivity,  $(x_j)_{j=1 \dots N}$ .

We turn now to the solution of the multiple regression problem Eq. (9), where the quantities to be predicted are the coefficients of  $\mathbf{B}$ . We use the Least Absolute Shrinkage and Selective Operator (Lasso), which performs both variable selection and prediction [28]. Over classical least square regression Lasso offers two major advantages that are very useful in modeling brain connectivity: Firstly, it improves prediction by setting some coefficients to zero. This results in removing noisy and irrelevant variables and thus reducing the total variance. Secondly, it allows the selection of the most relevant variables and thus it links each functional connection with a subset of structural connection in a data driven way. The correct predictors are identified with high probability even when the number of variables is higher than the number of observation under the assumption that the true model is sparse [6]. The Lasso estimate is defined as:

$$b_L = \arg \min_b \left\{ \sum_{i=i}^M (y_i^k - b_0 - \sum_j^N b_j x_{i,j})^2 + \lambda \sum_{j=1}^N |b_j| \right\} \quad (10)$$



**Fig. 2.** a) The combination of atlas-based segmentation and soft-tissue segmentation allows the delineation of anatomically sensible ROIs in gray matter. Atlas-based segmentation is based on established work [1], where 30 manual segmentations are propagated to the new subjects by non-rigid registration [25]. SPM has been used for soft-tissue segmentation [11]. b) The resulted 83 ROIs are shown both in diffusion space (top row) as well as in fMRI space (bottom row). Note that here we used only cortical regions.

Because of the  $\ell_1$  lasso penalty, most of the coefficients are set to zero. We used the LARS implementation of the Lasso algorithm in R statistics [7], which computes the complete lasso solution simultaneously for all values of the shrinkage parameter,  $\lambda$ .

#### 4 Experimental Results: Inter-subject Prediction of Resting-State Functional Connectivity

Brain connectivity analysis was performed in 22 normal adults (11 females, 11 males, mean age  $33.4 \pm 11.37$  years). rs-fMRI: T2\*-weighted gradient EPI sequence, TR/TE=2000/30, 31 ascending slices with thickness  $3.25\text{mm}$ , gap  $0.75\text{mm}$ , voxel size  $2.5 \times 2.5 \times 4\text{mm}$ , flip angle  $90^\circ$ , FOV  $280 \times 220 \times 123\text{mm}$ , matrix  $112 \times 87$ . DWI: 64 non-collinear directions, in 72 slices, slice thickness  $2\text{mm}$ , FOV  $224\text{mm}$ , matrix  $128 \times 128$ , voxel size  $1.75 \times 1.75 \times 2\text{mm}^3$ , b-value  $1000\text{ s/mm}^2$ . High resolution T1-weighted whole-brain structural images were also obtained in all subjects.

FSL was used for image pre-processing of both diffusion weighted (DWI) and fMRI images [26]. This involved eddy current correction of DWI and motion correction as well as spatial smoothing and whitening of fMRI images. Brain extraction was performed originally with FSL and it was manually refined later. Bias correction was applied to T1 and B0 images to improve the robustness of the non-rigid registration tools [25].

BOLD fluctuations are profound in gray matter, while DTI is more reliable in delineating white matter fibers. Hence, we are interested in defining cortical regions of interest (ROIs) that are located in gray matter and they are defined according to widely used anatomical atlas [1]. Cortical parcellation is obtained with the fusion of atlas-based [1] and tissue based segmentation [11], Fig. 2.

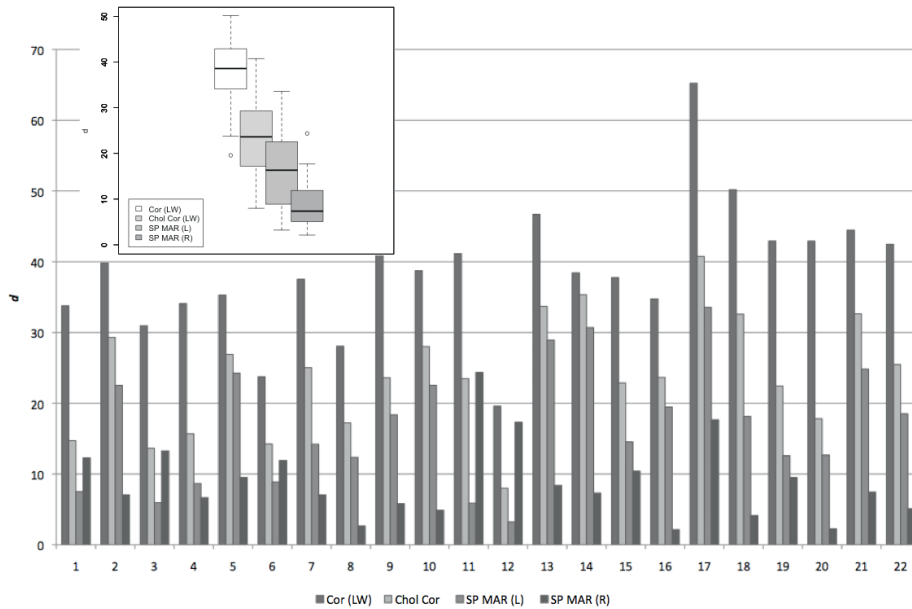
**Table 1.** Prediction performance under different scenarios

Methods	Cor(LW)	Chol Cor(LW)	SP MAR (L)	SP MAR(R)
log-likelihood	NA	194.57	185.13	<b>232.4128</b>
$d$	38.61	22.32	16.74	<b>8.96</b>

To construct corresponding functional networks the fMRI signal was averaged across voxels within each area. The signal in CSF and white matter has been also averaged and the six motion correction parameters were estimated with FEAT, FSL [26]. All these eight parameters were accounted in the estimation of the covariance matrix.

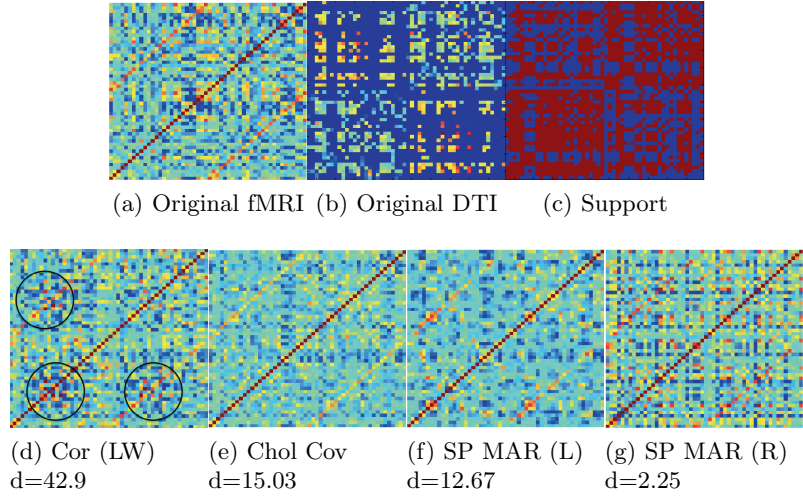
Tracts between regions are identified using a standard probabilistic algorithm available as part of FSL [26,2]. However, measurements of connection probability are difficult to interpret as the probability measure is very sensitive to noise in the data, as well as the size and separation of the ROIs. Instead, we estimate the local diffusion anisotropy by determining the diffusive transfer between voxels using the orientation distribution function (ODF) [24].

*Model selection framework.* Here we utilise the probabilistic framework described in section 3 to learn the link between anatomical connectivity and functional brain connectivity. We use leave-one-out cross validation and the algorithm described in Alg. 1 to measure the performance of the model. The results are



**Fig. 3.** Performance of the different probabilistic models under leave-one-out validation. Inside the main figure is the Box-and-Whisker diagram over all subjects.





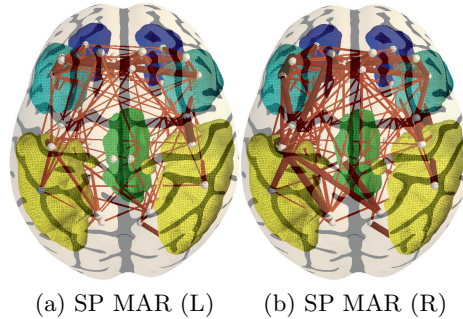
**Fig. 4.** Qualitative results for one leave-one-out subject when its structural connectivity matrix is used as a predictor. a) Sampled covariance matrix of the original fMRI signal, b) Structural connectivity matrix used as the input to the trained model. c) Sparsity pattern over all other subjects, d) Prediction based on Cor (LW). Black circles point out to relatively large errors in prediction for several connections, e) Prediction based on the Chol cov, f) Prediction based on the SP MAR (L), g) Prediction based on the SP MAR (R).

compared to the univariate linear model described in Eq. (1) with and without the Cholesky decomposition. Quantitative analysis is performed based on both the log-likelihood measure Eq. (11) as well as the  $d$  metric described in Eq. (2). The log-likelihood gives the likelihood to observe  $\Sigma^{\text{test}}$  given the model  $\Sigma^{\text{pred}}$ , and it relies on the underlying model of functional connectivity.

$$L(\Sigma^{\text{pred}}|\Sigma^{\text{test}}) = -\log \det(\Sigma^{\text{pred}}) + \text{tr}((\Sigma^{\text{pred}})^{-1}\Sigma^{\text{test}}) \quad (11)$$

When the prediction model do not use a matrix square root then the output it is not a SPD matrix,  $\mathcal{S}ym_p^+$ .

In Table 1, the mean values over all leave-one-out cross validations are given. Cor (LW) corresponds to predicting directly the functional correlation matrix based on Ledoit-Wolf regularisation [18], which provides a more accurate and well-conditioned covariance matrix. Chol Cor (LW) corresponds to predicting the square root matrix of the correlation matrix, Cor (LW). The original correlation matrix has been reordered to be consistent with our methodology (see section 3). Elements on the diagonal are set to one by multiplying with the inverse of the diagonal from the right, Eq. 8. Subsequently, results of our approach, SP MAR (R), as well as a variation, SP MAR (L), where Eq. 8 takes the form ( $\tilde{\mathbf{B}} = \text{diag}(\mathbf{B})^{-1}\mathbf{B}$ ), are presented. Fig. 3 shows a more detailed summary of the results over all subjects. The probabilistic framework described in section 3 outperforms all other approaches. The difference between SP MAR (R) and SP



**Fig. 5.** Identifying structural connections associated with the default mode network. With yellow is represented the lateral parietal cortex, green areas represent the posterior cingulate gyrus (PCC), blue and light blue represent the medial prefrontal and orbito-frontal areas, respectively.

MAR (L) is that one scales the innovation terms of the MAR model, whereas the other scales the relative weights of the observed time series. The innovation terms are not related to connectivity and hence it is not possible to predict them from anatomical connectivity. Fig. 4 shows results distribution obtained in leave-one-out cross-validation scheme. In the simple correlation model, Cor (LW), we use black circles to point out that there is a relatively large error in prediction for several connections.

Finally, in Fig. 5, an application of the predictive framework in identifying structural connections associated with the the default mode network is shown. Connections are plotted when they appear in more than 15 out of 22 folds of the leave-one-subject-out cross validation loop. The diameter of the tubular shapes depends on the average of the absolute value of the coefficient associated with each connection. We note that stronger and more dense connections are identified in the SP MAR (R) relative to the SP MAR (L) model. This indicates that the earlier identifies connections and the associated coefficients more consistently. Also in SP MAR (L) model the PCC is not as well interconnected as in SP MAR (R) model.

## 5 Conclusions

We have developed a novel probabilistic framework to learn the mapping  $f$  from the space of anatomical connectivity  $\mathcal{A}$  to the space that describes functional connectivity  $\mathcal{F}$ . Functional connectivity is modeled as the covariance structure of functional brain signals. We assume that all the individual precision matrices share the same structure of conditional independence. Cholesky decomposition is used to enforce a SPD,  $Sym_p^+$ , prediction output of the multiple regression. Leave-one-subject-out cross validation demonstrates the efficiency of the technique. We also depict the default mode network and the most influential structural connections associated to it.

## References

1. Aljabar, P., Heckemann, R., Hammers, A., Hajnal, J., Rueckert, D.: Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy. *Neuroimage* 46(3), 726–738 (2009)
2. Behrens, T., Woolrich, M., Jenkinson, M., Johansen-Berg, H., Nunes, R., Clare, S., Matthews, P., Brady, J., Smith, S.: Characterization and propagation of uncertainty in diffusion-weighted mr imaging. *Magnet Reson Med.* 50(5), 1077–1088 (2003)
3. Burns, J.: An evolutionary theory of schizophrenia: Cortical connectivity, metarepresentation, and the social brain. *Behavioral and Brain Sciences* 27(6), 831 (2004)
4. Damoiseaux, J., Greicius, M.: Greater than the sum of its parts: a review of studies combining structural connectivity and resting-state functional connectivity. *Brain Struct. Funct.* 213(6), 525–533 (2009)
5. Deligianni, F., Robinson, E., Beckmann, C., Sharp, D., Edwards, A., Rueckert, D.: Inference of functional connectivity from structural brain connectivity. In: ISBI, pp. 1113–1116 (2010)
6. Donoho, D.: For most large underdetermined systems of linear equations the minimal  $\ell_1$ -norm solution is also the sparsest solution. *Comm. Pure Appl. Math.* 59(6), 797–829 (2006)
7. Efron, B., Hastie, T., Johnstone, I., Tibshirani, R.: Least angle regression. *Ann. Stat.* 32(2), 407–499 (2004)
8. Ferrarelli, F., Massimini, M., Sarasso, S., Casali, A., Riedner, B., Angelini, G., Tononi, G., Pearce, R.: Breakdown in cortical effective connectivity during midazolam-induced loss of consciousness. *P. Natl. Acad. Sci. Usa* 107(6), 2681–2686 (2010)
9. Förstner, W., Moonen, B.: A metric for covariance matrices. *Qua. Vadis Geodesia*, 113–128 (1999)
10. Fransson, P., Marrelec, G.: The precuneus/posterior cingulate cortex plays a pivotal role in the default mode network: Evidence from a partial correlation network analysis. *Neuroimage* 42, 1178–1184 (2008)
11. Friston, K.: *Statistical parametric mapping: the analysis of functional brain images.* Academic Press, London (2007)
12. Greicius, M., Supekar, K., Menon, V., Dougherty, R.: Resting-state functional connectivity reflects structural connectivity in the default mode network. *Cereb. Cortex* (2008)
13. van den Heuvel, M.P., Mandl, R.C.W., Kahn, R.S., Pol, H.E.H.: Functionally linked resting-state networks reflect the underlying structural connectivity architecture of the human brain. *Human Brain Mapping* 30(10), 3127–3141 (2009)
14. Honey, C., Sporns, O., Cammoun, L., Gigandet, X., Thiran, J., Meuli, R., Hagmann, P.: Predicting human resting-state functional connectivity from structural connectivity. *P. Natl. Acad. Sci. Usa* 106(6), 2035–2040 (2009)
15. Honey, C., Kötter, R., Breakspear, M., Sporns, O.: Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *P. Natl. Acad. Sci. Usa* 104(24), 10240 (2007)
16. Lauritzen, S.: *Graphical models.* Oxford University Press, USA (1996)
17. LeCun, Y., Chopra, S., Hadsell, R., Ranzato, M., Huang, F.L.: Energy-based models. In: BakIr, G., Hofmann, T., Schölkopf, B. (eds.) *Predicting Structured Data*, pp. 191–245. MIT Press, Cambridge (2007)

18. Ledoit, O., Wolf, M.: A well-conditioned estimator for large-dimensional covariance matrices. *J. Multivar. Anal.* 88, 365–411 (2004)
19. Lenglet, C., Rousson, M., Deriche, R., Faugeras, O.: Statistics on the manifold of multivariate normal distributions: Theory and application to diffusion tensor MRI processing. *J. Math. Imaging Vis.* 25, 423–444 (2006)
20. Morcom, A., Fletcher, P.: Does the brain have a baseline? why we should be resisting a rest. *Neuroimage* 37(4), 1073–1082 (2007)
21. Müller, R.: The study of autism as a distributed disorder. *Ment. Retard. Dev. Disabil. Res.* 13(1), 85–95 (2007)
22. Pennec, X., Fillard, P., Ayache, N.: A Riemannian framework for tensor computing. *Int. J. Comput. Vision* 66, 41–66 (2006)
23. Pollonini, L., Pophale, S., Situ, N., Wu, M.H., Frye, R., Leon-Carrion, J., Zouridakis, G.: Information communication networks in severe traumatic brain injury. *Brain Topogr.* 23(2), 221–226 (2010)
24. Robinson, E., Hammers, A., Ericsson, A., Edwards, A., Rueckert, D.: Identifying population differences in whole-brain structural networks: a machine learning approach. *Neuroimage* 50(3), 910–919 (2010)
25. Rueckert, D., Sonoda, L., Hayes, C., Hill, D.: Non-rigid registration using free-form deformations: application to breast mr images. *IEEE Trans. Med. Imag.* 18, 712–721 (1999)
26. Smith, S., Jenkinson, M., Woolrich, M., Beckmann, C., Behrens, T., Johansen-Berg, H., Bannister, P., Luca, M.D., Drobnjak, I., Flitney, D., Niazy, R., Saunders, J., Vickers, J., Zhang, Y., Stefano, N.D., Brady, J., Matthews, P.: Advances in functional and structural mr image analysis and implementation as fsl. *Neuroimage* 23, 208–219 (2004)
27. Smith, S., Miller, K., Salimi-Khorshidi, G., Webster, M., Beckmann, C., Nichols, T., Ramsey, J., Woolrich, M.: Network modelling methods for fMRI. *Neuroimage* (2010) (in press)
28. Tibshirani, R.: Regression shrinkage and selection via the lasso. *J. Roy. Stat. Soc. B* 58(1), 267–288 (1996)
29. Varoquaux, G., Gramfort, A., Poline, J.B., Thirion, B.: Brain covariance selection: better individual functional connectivity models using population prior. In: *NIPS* (2010)
30. Venkataraman, A., Rathi, Y., Kubicki, M., Westin, C.F., Golland, P.: Joint generative model for fmri/dwi and its application to population studies. *Med. Image Comput. Comput. Assist Interv.* 13(pt. 1), 191–199 (2010)