

Modelling the Relationship Between Structural and Functional Connectomes

Jon Clayden <j.clayden@ucl.ac.uk> and Fani Deligianni

The term **connectome** refers to the complete pattern of interconnections between cells, cell populations or regions in the brain or nervous system. Two general ways of thinking about connectomes are in terms of the axon fibre bundles that structurally connect regions together (the **structural** connectome), or in terms of coordinated activity between regions that are in communication (the **functional** connectome). Information about connectivity based on either of these perspectives can be obtained from different submodalities of magnetic resonance imaging (MRI). The aim of this project is to model and investigate the relationship between structural and functional connectomes.

Background

Diffusion MRI (dMRI) measures the mobility of water molecules in biological tissues. The brain's white matter fibre tracts, made up of bundles of axons, have a highly linear structure which restricts water diffusion more perpendicular to the tract than parallel to it. This produces a directional dependence in diffusion magnitude, called **anisotropy**, which can be harnessed to reconstruct the complete white matter pathways between coherent regions of the cortex, a process called **tractography** (see Figure 1).

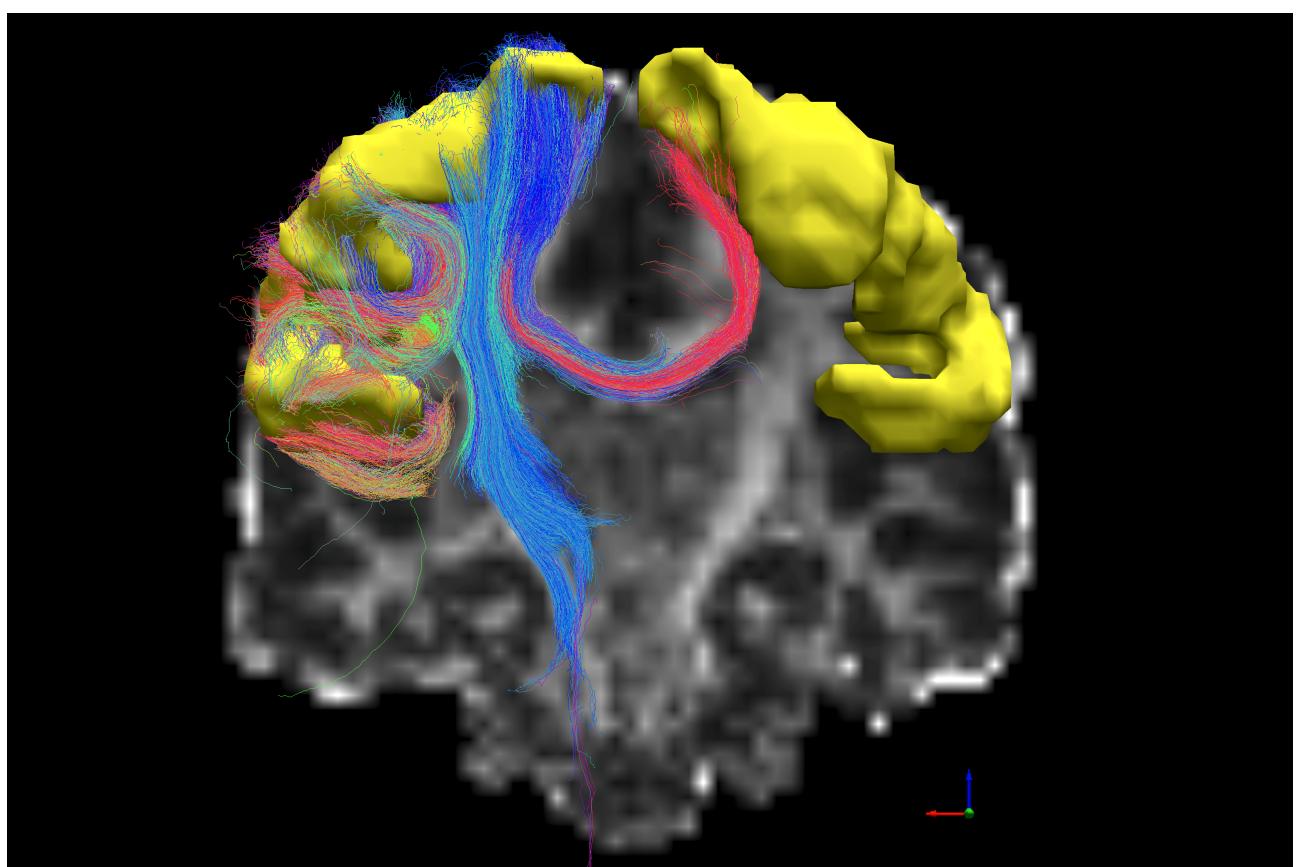
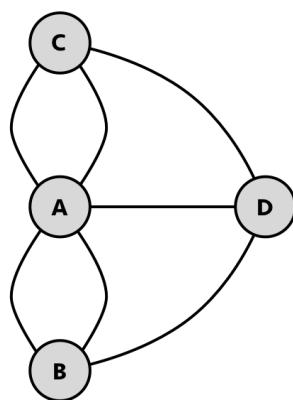
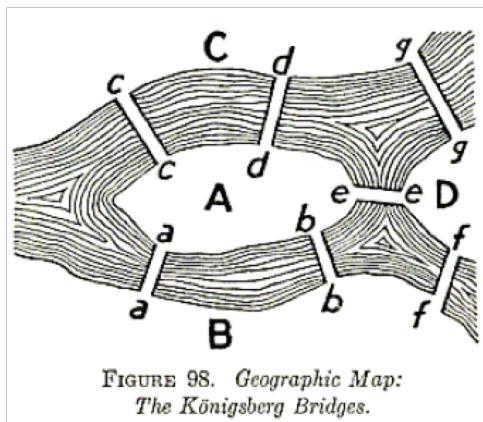


Figure 1. Connections of the left postcentral gyrus (yellow region, left), as reconstructed using fibre tractography, based on diffusion MRI. Each coloured line represents a "streamline", obtained by following local fibre orientations through the brain. Fibres connecting to the right postcentral gyrus (yellow region, right) appear red; the corticospinal tract appears blue. Also visible are shorter projections to neighbouring gyri. The underlying map shows fractional anisotropy.

By contrast, resting-state functional MRI (rs-fMRI) offers a proxy for brain activity, through the so-called blood oxygen level-dependent (BOLD) contrast, which is related to the delivery of blood to active neuronal tissue. Functional **time series** are obtained at each voxel location in the brain, at a coarse temporal resolution of two or more seconds, and connectivity is inferred from the degree of **correlation** between these time series. Regions demonstrating correlated activity are assumed to be connected, either directly or indirectly.

Whether connectivity information comes from a structural or functional source, it can be abstractly represented as a **graph** (see Figure 2), with parcellated brain regions represented as a series of nodes, or **vertices**, connected by **edges**. Since neither dMRI nor rs-fMRI can tell us about the direction of information flow, these connections are **undirected**, and the graph can be represented by a symmetric matrix. The elements of the matrix may either be binary 1s and 0s to indicate the presence or absence of the appropriate connection, or else a series of **weights** indicating the “strength” of the connection.



	A	B	C	D
A		2	2	1
B	2			1
C	2			1
D	1	1	1	

Figure 2. The seven bridges of Königsberg, a classical graph problem. The layout of the bridges connecting four regions of land (A, B, C and D) is shown in map, topographic and matrix forms.

Structural and functional connectomes of the brain represent two different but fundamentally coupled views of a single connected system. Logically, for example, a physical connection must exist between two regions if they are able to functionally communicate. However, both dMRI and rs-fMRI infer connectivity indirectly, and each technique has its own limitations. Research with MRI has therefore shown that the relationship between the two can be quite complex (e.g., Honey et al., 2009).

In this project a number of different forms of relationship between the two modalities will be explored, and compared against each other with regard to their predictive ability in unseen data.

The Datasets

This project contains two core tasks, which will make use of two different datasets. The first (*Task1Data*) is a set of multimodal imaging data from a single healthy adult, along with a series of structural connectomes derived using several different anisotropy thresholds. The second (*Task2Data*) consists of precalculated structural and functional connectome matrices from 19 healthy adults.

The first task's data is the test dataset from the [TractoR](#) software package (Clayden et al., 2011), and so it follows that package's naming conventions. Within the `tractor` subdirectory are several files, of which the following are the most important for our purposes.

<code>tractor/structural/refT1.nii.gz</code>	A high-resolution, T_1 -weighted structural image, used as the basis for cortical parcellation (Figure 3A, base layer)
<code>tractor/structural/parcellation.nii.gz</code>	A cortical parcellation, generated by FreeSurfer (Desikan et al., 2006), in which each voxel is labelled with a number corresponding to a coherent brain region (Figure 3A, overlay)
<code>tractor/structural/parcellation.lut</code>	A table of information about the labelled regions in the parcellation: their indices, names, and tissue types
<code>tractor/diffusion/data.nii.gz</code>	The dMRI dataset, a 4D volume with gradient direction varying along the fourth dimension, preprocessed using TractoR and FSL-FDT to align the individual volumes, mask the brain and calculate derived metrics such as fractional anisotropy
<code>tractor/diffusion/dti_FA.nii.gz</code>	A map of fractional anisotropy
<code>tractor/diffusion/parcellation.nii.gz</code>	The same parcellation as above, transformed into the space of the diffusion data
<code>tractor/functional/data.nii.gz</code>	The rs-fMRI dataset, a 4D volume with time varying along the fourth dimension, preprocessed using FSL-FEAT to align the individual volumes and apply temporal filtering
<code>tractor/functional/parcellation.nii.gz</code>	The same parcellation as above, transformed into the space of the functional data

Table 1. The most relevant files in the first dataset.

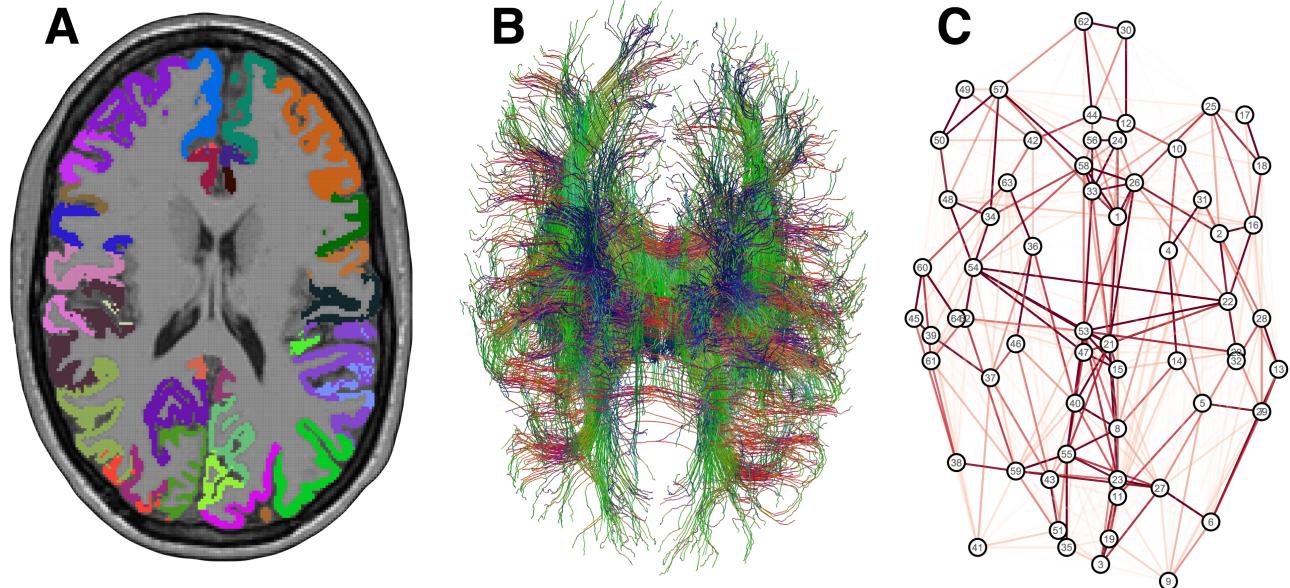


Figure 3. Stages of constructing a structural connectome, based on T_1 -weighted and diffusion MRI, and fibre tractography (from Clayden, 2013). The grey matter is parcellated into coherent regions (A), white matter pathways are reconstructed using tractography (B), and then connectivity information between each region pair is compiled into an abstract graph (C).

Figure 3 shows the process for constructing structural connectomes, which has been performed in TractoR, applying fractional anisotropy (FA) thresholds of 0.1 to 0.8, in steps of 0.1, to the brain mask to choose where to seed streamlines. 10,000 streamlines in total were generated in each case (Figure 3B), with any streamlines not reaching two parcellated cortical regions being discarded. For each threshold a binary graph was then generated (Figure 3C), with an edge value of 1 indicating that at least one streamline connected that pair of regions, and a value of 0 indicating than none did. The matrices are stored as comma-separated values (CSV) files in the *connectomes* subdirectory.

The second task's data consists of weighted structural and functional connectomes from 19 healthy adults. Subjects are numbered from 32 to 50, and structural connectomes have *WFA* in the names, while functional connectomes contain *rsfMRI*. The former were derived by taking an average of the FA values in each voxel visited by the streamlines connecting each pair of cortical regions, weighted by the number of streamlines visiting the voxel. (A value of zero indicates that no streamlines connected that pair of regions, so no FA information is available. **Note that this is not the same thing as having a mean FA of zero.**) The latter are obtained by normalising the inverse covariance matrix , or precision matrix, which is closely related to partial correlation (cf. Figure 4).

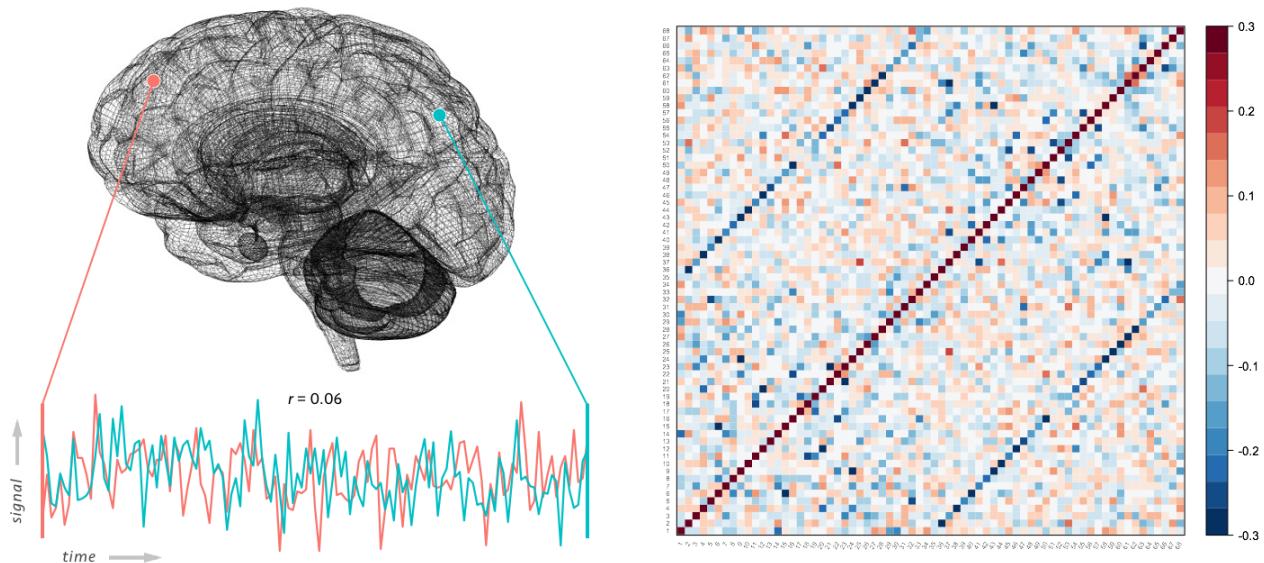


Figure 4. Calculation of functional connectivity. Time courses from cortical regions are obtained and their correlations or partial correlations with each other are calculated (left). This is repeated for each region pair to produce a full correlation matrix, which is directly interpreted as the functional connectome (right).

Three Matlab toolboxes have also been provided, to help with aspects of the tasks: [Jimmy Shen's NIfTI toolbox](#) for reading NIfTI-format images, the [Brain Connectivity Toolbox](#) for calculating graph metrics, and [Kevin Murphy's Matlab port](#) of the *corpcor* R package for estimating correlation using a shrinkage approach. (The tasks could also be completed using TractoR and R, but you may find using Matlab more convenient and/or more familiar.)

Core Task 1: Graph properties of structural and functional connectomes

The purpose of this task is to explore the graph properties of connectomes, and how they are affected by certain parameters chosen in the course of generating them. You will need to complete and report on the following subtasks.

- Explore the dataset within the *tractor* directory, using Table 1 as a guide¹. Convince yourselves that you understand the information needed to reconstruct a structural connectome, and the process for doing so.
- Manually threshold the FA map at different values between 0.1 and 0.8, setting subthreshold voxels to zero. Discuss the spatial distribution of voxels remaining above threshold as the latter increases.
- Load the structural connectomes in the *connectomes* directory² and use the Brain Connectivity Toolbox (BCT) to calculate metrics such as edge density, mean shortest path, efficiency and mean clustering coefficient, for each FA threshold³ (Rubinov & Sporns, 2010). Carefully consider how to handle missing edges and disconnected vertices, where they occur. Discuss the effects of varying the FA threshold on the graphs, and the reasons for them.
- Calculate a weighted functional connectome, by first extracting an average time series in each cortical region, and then deriving a correlation matrix using the shrinkage approach of Schäfer & Strimmer (2005). The provided *corshrink* function provides a starting point for this, but it uses an algorithm to select the shrinkage parameter, *lambda*, whereas you should vary it manually between 0.1 and 0.8, in steps of 0.1. In each case, threshold the resulting matrix at a correlation of 0.1, binarise it, and then calculate graph metrics using the BCT. Discuss the effects of varying the shrinkage parameter on the graphs, and the reasons for them. Consider the effects of retaining or discarding negative correlations with absolute value greater than 0.1, and what significance these edges may have in the context of brain activity.

Core Task 2: Modelling the relationships between connectomes

In this task, we will explore the relationship between structural and functional connectomes, by predicting functional connectivity weights, f_{ij} , between regions i and j , from the associated structural connectivity weights, s_{ij} . Weighted graphs will therefore be used here.

Since the structural connectome is much sparser than the functional one, and we can only model edge weights where edges exist, we will also consider indirect structural connectivity weights, t_{ij} , representing a connection between two regions via exactly one other region.

A series of simple linear models will be explored and compared in this core task, both including and excluding indirect connectivity information. The following subtasks need to be completed.

- Calculate an indirect structural connectivity matrix for each subject, which is defined for our purposes as the greatest minimum weight in all available two-step chains:

$$t_{ij} = \max_k \{ \min \{ s_{ik}, s_{kj} \} \} \text{ s.t. } s_{ik}, s_{kj} \neq 0.$$

¹ The *load_nii* function, from the NIfTI toolbox, can be used to read *.nii.gz* files, and *view_nii* is a simple viewer. The *.lut* files are simple text files and can be examined in any text editor.

² CSV files may be loaded using the *csvread* Matlab function. You will need to use second and third arguments to skip the first three lines in each case, since they contain metadata.

³ Remember that these graphs are binary and undirected. There are variants of most BCT functions which are specialised to particular types of graph.

- Fit each of the following linear models to the data in turn. Notice that the coefficients (α_{ij} , β_{ij} and γ_{ij}) are estimated separately for each edge for which there is data. In each case the regression should be performed across all 19 subjects.

1. $f_{ij} = \alpha_{ij} + \beta_{ij} s_{ij}$
2. $f_{ij} = \alpha_{ij} + \beta_{ij} s_{ij} + \gamma_{ij} s_{ij}^2$
3. $f_{ij} = \alpha_{ij} + \beta_{ij} t_{ij}$
4. $f_{ij} = \alpha_{ij} + \beta_{ij} t_{ij} + \gamma_{ij} t_{ij}^2$
5. $f_{ij} = \alpha_{ij} + \beta_{ij} s_{ij} + \gamma_{ij} t_{ij}$

- For each model, evaluate the Akaike and Bayesian information criteria (AIC and BIC), and perform leave-one-out (or k -fold) cross-validation to evaluate its predictive value in terms of the sum-of-squared-errors. Consider the variation in parameter estimates and fit quality across the different edges, and plot the relationship described by the best-fitting models, for a small number of individual connections where it holds particularly strongly or weakly.
- Repeat the previous subtask, but using a single set of coefficients for all edges, for each model. Explore and discuss how the results compare.
- Devise an additional model for the relationship between the structural and the functional connectivity data, and fit it. Be prepared to explain your choice of relationship in terms of what is plausible in the brain.

Advanced Tasks

- Estimate the structural connectivity density for each vertex (the sum across rows or columns of each structural connectivity matrix). Do the same for the functional connectivity matrices, and then model functional connectivity density based on structural connectivity density, independently for each vertex. Do you find any strong associations between the modalities using this approach?
- Repeat task 2, fitting a single set of coefficients over all edges, but for each subject independently. Do the results generalise well, when you use the fitted parameters to estimate the functional connectivity of other subjects?
- Use sparse linear models based on LASSO to relate each functional connection to a subset of structural connections (cf. Deligianni et al., 2013). Once again, you should use cross-validation to examine how well this model performs. (There are several packages you could use to implement this approach.)

References

Clayden J.D. (2013). Imaging connectivity: MRI and the structural networks of the brain. *Functional Neurology* **28**(3):197-203.

Clayden J.D., Muñoz Maniega S., Storkey A.J., King M.D., Bastin M.E. & Clark C.A. (2011). TractoR: Magnetic resonance imaging and tractography with R. *Journal of Statistical Software* **44**(8):1-18.

Deligianni F., Varoquaux G., Thirion B., Sharp D.J., Ledig C., Leech R. & Rueckert D. (2013). A framework for inter-subject prediction of functional connectivity from structural networks. *IEEE Transactions on Medical Imaging* **32**(12):2200-2214.

Desikan R.S., Segonne F., Fischl B., Quinn B.T., Dickerson B.C., Blacker D., Buckner R.L., Dale A.M., Maguire R.P., Hyman B.T., Albert M.S. & Killiany R.J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* **31**:968-980.

Honey C.J., Sporns O., Cammoun L., Gigandet X., Thiran J.P., Meuli R. & Hagmann P. (2009). Predicting human resting-state functional connectivity from structural connectivity. *Proceedings of the National Academy of Sciences of the United States of America* **106**(6):2035-2040.

Rubinov M. & Sporns O. (2010). Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage* **52**:1059-1069.

Schäfer J. & Strimmer K. (2005). A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. *Statistical Applications in Genetics and Molecular Biology* **4**(1):32.