Heuristics for Connectivity-Based Brain Parcellation of SMA/pre-SMA through Force-Directed Graph Layout

Alessandro Crippa[a], Leonardo Cerliani[b], Luca Nanetti[b], Jos B. T. M. Roerdink[a,b,*]

[a] *Johann Bernoulli Institute for Mathematics and Computer Science, University of Groningen, The Netherlands*
[b] *BCN Neuroimaging Center, University Medical Center Groningen, University of Groningen, The Netherlands*

Abstract

We propose the use of force directed graph layout as an explorative tool for connectivity-based brain parcellation studies. The method can be used as an heuristic to find the number of clusters intrinsically present in the data (if any) and to investigate their organisation. It provides an intuitive representation of the structure of the data and facilitates interactive exploration of properties of single seed voxels as well as relations among (groups of) voxels.

We validate the method on synthetic datasets and we investigate the changes in connectivity in the supplementary motor cortex, a brain region whose parcellation has been previously investigated via connectivity studies. This region is supposed to present two easily distinguishable connectivity patterns, putatively denoted by SMA (supplementary motor area) and pre-SMA.

Our method provides insights with respect to the connectivity patterns of the premotor cortex. These present a substantial variation among subjects and their subdivision into two well separated clusters is not always straightforward.

Keywords: Diffusion imaging, tractography, connectivity, brain parcellation, supplementary motor cortex.

1. Introduction

Classification of structural and functional regions in human brain cortex is a key aspect of neuroscience (Passingham et al., 2002). Under the hypothesis of a relation between brain structure and brain function, several techniques for cortex parcellation are available in the literature. The golden standard for cortex structural analysis is based on cyto- or myeloarchitectonical studies of *post mortem* brains (Brodmann, 1909; Vogt and Vogt, 1919; Vogt et al., 1995), which allow parcellation of brain cortex based on variations of density and size of cell bodies at the microscopic level. Functional identification of several brain areas was achieved by using anatomical landmarks such as sulci (Geyer et al., 1996; Amunts et al., 1999), or using modern *in vivo* explorative techniques such as fMRI and PET, which improved and broadened the possibility to locate functional brain areas.

One of the imaging techniques providing insight in anatomical brain connectivity is Diffusion Tensor Imaging (DTI). This is a magnetic resonance modality for measuring white matter connectivity *in vivo*, bridging the gap between structural studies and *post mortem* anatomical analysis (Klein et al., 2007). DTI is a non-invasive method that allows quantitative assessment of directional information of water self-diffusion in biological tissues. Measuring the anisotropic self-diffusion of water, induced by the fibrous nature of white matter, allows one to infer the directional arrangement of bundles of axons (Basser et al., 1994; Pierpaoli et al., 1996). The principal direction of water diffusion in each white matter voxel is often used for tracking fiber bundles; cf. (Mori et al., 1999; Basser et al., 2000) for a reference on DTI and tracking techniques. Probabilistic tractography (Behrens et al., 2003; Parker et al., 2003; Koch et al., 2002) estimates fiber tracks using probability density functions that describe the local uncertainty of fiber bundle orientation. This technique provides confidence bounds on the locations of connections from a seed voxel to every other voxel of the brain. Such connections define the *connectivity profile* of that voxel (Johansen-Berg et al., 2004). Probabilistic tractography has been used to demonstrate different patterns of

*Corresponding author. Tel: +31-50-3633931; Fax: +31-50-3633800

Email addresses: a.crippa@rug.nl (Alessandro Crippa), l.cerliani@med.umcg.nl (Leonardo Cerliani), l.nanetti@med.umcg.nl (Luca Nanetti), j.b.t.m.roerdink@rug.nl (Jos B. T. M. Roerdink)

anatomical connectivity (Behrens and Johansen-Berg, 2005); it has also been used to compute parcellation of brain regions based on connectivity information alone (Johansen-Berg et al., 2004; Devlin et al., 2006; Anwander et al., 2007; Croxson et al., 2005; Rushworth et al., 2006; Johansen-Berg et al., 2008).

In this paper we visualise brain connectivity networks by making use of techniques from graph visualisation (Herman et al., 2000; Marshall and Herman, 2001; van Ham and van Wijk, 2004). In particular, we apply the force-directed graph layout (FDGL) technique, which models a graph as a physical system and tries to find a layout for the nodes and edges of the graph such that the total energy of the system is minimal. We show that FDGL algorithms are well suited for connectivity-based analysis of brain cortical areas. Relations among the connectivity profiles of seed voxels are used to define a graph layout where voxels (graph nodes) are drawn closer to each other when their connectivity profiles are more similar. This technique allows us to investigate the presence and the number of clusters in a dataset by analysing the positions of the graph nodes and by counting the groups of nodes drawn close to each other. Insights on the strength of similarity among voxels and among groups of voxels are also possible, and hypotheses on the reliability of the results can be verified in a follow-up analysis by mapping the density on the original anatomy, i.e., *before* clustering the data. This approach can be used as an heuristic for determining the number of clusters to be used in a (subsequent) clustering algorithm.

We applied our method to the parcellation of SMA (supplementary motor cortex) and pre-SMA, a brain region that has been previously examined via connectivity-based parcellation (CBP) (Behrens and Johansen-Berg, 2005; Anwander et al., 2007; Klein et al., 2007; Nanetti et al., 2009). On the basis of previous evidence in animal studies (Picard and Strick, 1996; Luppino et al., 1993), a sharp transition in the connectivity properties of SMA and pre-SMA is expected. Nevertheless, neither a quantitative analysis of the difference between SMA and pre-SMA in terms of connectivity patterns, nor the possibility of a smooth transition of connectivity patterns between the two functional areas, has been investigated in the literature.

2. Related Work

Automatic parcellation of brain cortex, i.e., without *a priori* knowledge of target regions, was introduced by Johansen-Berg et al. (2004). The authors clustered the connectivity profiles of the voxels of SMA/pre-SMA and recovered the supposed location of the boundary between the two areas. The cross-correlation matrix (CCM) of the connectivity profiles of SMA/pre-SMA voxels was reordered using a spectral reordering algorithm (Barnard et al., 1995), which minimizes the sum of the reordered matrix entries, multiplied by their squared distance from the diagonal. If the data contain clusters (groups of voxels with similar connectivity), these will be visually distinguishable in the reordered CCM (Johansen-Berg et al., 2004) and the investigator can visually separate matrix compartments that identify clusters with distinct connectivity patterns.

Connectivity-based brain parcellation of the premotor cortex was replicated using k-means clustering on CCM by Anwander et al. (2007), Klein et al. (2007) and by Nanetti et al. (2009). In Anwander et al. (2007), the possibility of the existence of more than two clusters was investigated. In particular, k-means clustering with k=3 was consistently found to divide the most posterior area of the premotor cortex into two clusters which were reported as possibly referring to the primary motor cortex (M1) and SMA. Klein et al. (2007) validated the use of k-means clustering for the CBP of SMA and pre-SMA with fMRI techniques. Nanetti et al. (2009) showed how the application of k-means clustering to the same dataset can lead to a variety of rather different parcellations, depending on the initial placement of the starting points of the algorithm.

The output of k-means clustering is strongly dependent on the expected number of clusters, and it may provide misleading results if this number does not correspond to the number of clusters intrinsically present in the data. Compared to k-means clustering, spectral reordering of CCMs has the advantage of not requiring *a priori* knowledge of the number of clusters. Nevertheless, spectral reordered CCMs often do not show clear separations into two (or more) compartments and visual analysis of such matrices is not always straightforward. Furthermore, any decision taken by the investigator on the partitioning of the spectral reordered CCM is solely based on the layout of the CCM itself and no anatomical reference is taken into account.

3. Materials and Methods

3.1. Data acquisition and preprocessing

Diffusion-weighted (DW) images were acquired from 13 healthy subjects (6 F, 7 M, mean age: 20,94; sd 1,95; max 25, min 18) using a single-shot pulsed gradient spin echo EPI sequence (SENSE factor = 3, TR = 6 s, and TE = 77 ms) on a 3T MR scanner (Intera, Philips Medical Systems, Best, The Netherlands).

For each subject 60 DW images were collected using 60 uniformly distributed gradient directions (Jones et al., 1999), with a maximum gradient strength of 22 mT/m and a b-value of 800 s/mm^2. In addition, 7 non-DW images (b = 0 s/mm^2) were acquired before the DW images and averaged on the scanner. Each image consisted of 51 transverse slices with an in-plane resolution of 1.875 x 1.875 mm^2 and slice thickness of 2 mm (FOV = 240 x 240 mm^2 and data matrix size = 128 x 128). The total scan time for one dataset was approximately 7 min.

Diffusion data were corrected for eddy currents and head motion by affine registration to the non-weighted image. Data from the skull was subsequently removed and the non-weighted image was registered to the MNI standard template (voxel size 2 mm^3). All these steps were performed using FSL (Smith et al., 2004; Woolrich et al., 2008).

3.2. Tractography

After preprocessing, Bayesian estimation of diffusion parameters was performed as in Behrens et al. (2003) and implemented in FSL (Smith et al., 2004). In order to perform probabilistic tractography of the medial premotor region, we defined a seed region by a binary mask at X(MNI)=-2, spanning from Y=-22 to Y=30 (MNI coordinates), and extending in the Z-direction from above the cingulate sulcus to the dorsal crest of the medial wall; cf. Johansen-Berg et al. (2004). This mask was then transformed into each subject's DW space by means of an affine transformation with nearest neighbour interpolation. Probabilistic tractography was seeded for each voxel in the mask; five thousand samples were drawn from the orientation distribution of each seed voxel to generate its connectivity profile. Connectivity profiles were then binarised, setting to unity any entry bigger than zero.

In the next step the CCM of the connectivity profiles was calculated. The CCM is a symmetric square matrix where the entry (i, j) is the Pearson's correlation between the binarised connectivity profiles of seed voxels i and j, and the i^{th} row of the CCM contains the correlation of the binarised connectivity profile of the seed voxel i with the binarised connectivity profiles of all other voxels in the seed region.

3.3. Synthetic datasets

Six synthetic datasets were created for testing our approach. Each set consisted of 250 binary vectors simulating binarised connectivity profiles. The vectors had 10^5 entries, 5‰ of which were non-zero entries (connectivity profiles of voxels in pre-SMA/SMA cover roughly 5‰ of the whole brain volume).

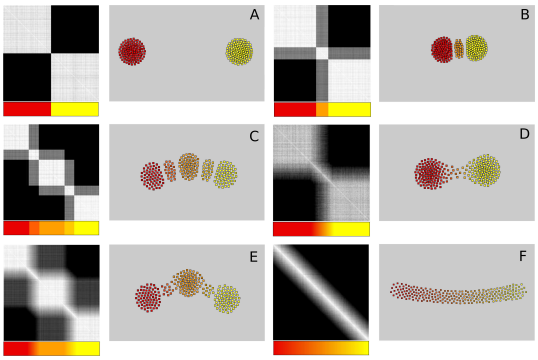

Figure 1: Six synthetic CCMs and the corresponding graph layouts. Graph layouts are coloured by the user according to visually detected clusters (A,B,C). Smooth variabilities in similarity are depicted using colour gradients (D,E,F). Bars below the CCMs show the clusters replotted on the columns of the CCMs.

The six sets of vectors were created so that the corresponding CCMs corresponded to six qualitatively different cases: (i) two distinct compartments, corresponding to the presence of two completely separated clusters; (ii) two partially overlapping compartments, corresponding to the presence of a non-empty intersection between two clusters; (iii) three pairwise overlapping compartments; (iv) two big compartments connected by a series of smaller overlapping compartments; (v) three compartments pairwise connected by a continuum of smaller overlapping compartments; (vi) a continuum of small overlapping compartments, where each matrix column (row) correlates only with its neighbouring columns (rows).

The synthetic CCMs are shown in Fig. 1, where grey values represent the strength of the correlation among connectivity profiles, white representing high correlation. Note that the synthetic CCMs were not spectrally reordered, but they were created this way to facilitate the visual analysis of the compartments.

4. Graph Layout

A graph $G = (V, E)$ is defined by a set V of nodes and a set E of edges, where an edge connects a pair of nodes. We use undirected graphs, so if node i is connected to node j, then j is connected to i. A graphical representation of a graph is called a graph layout. Graphs are usually depicted with their nodes as points in a plane and their edges as arcs connecting the nodes.

There are different styles of representation, suited to different types of graphs or different purposes of presentation. The arrangement of nodes and edges emphasises different characteristics of the graph, assists in the

understanding of graph properties, and reveals patterns in the data. Such graph drawing styles are known under the heading of graph aesthetics (Di Battista et al., 1999). FDGL is a consolidated visualisation method whose purpose is to position the nodes and the edges in an aesthetically pleasing way. Our main consideration here is that proximity in the layout corresponds to a certain similarity measure defined in the data space.

The purpose of FDGL is to find a position $\vec{p}_i$ for each node i of the graph such that nodes corresponding to highly correlated connectivity profiles are drawn near each other. This is achieved by using a physical analogy. Nodes are modelled as particles on which forces are acting. The optimal layout corresponds to the situation when the physical system is in equilibrium, i.e., the energy of the system is minimal. Two forces act on each node: a Hooke's spring force F_A, attracting nodes, and a Coulomb-like repulsion force F_R, repelling nodes. The total force $F_T(i)$ acting on a node i is thus:

$$F_T(i) = \sum_j F_A(i,j)\frac{(\vec{p}_i - \vec{p}_j)}{|\vec{p}_i - \vec{p}_j|} - \sum_j F_R(i,j)\frac{(\vec{p}_i - \vec{p}_j)}{|\vec{p}_i - \vec{p}_j|} \quad (1)$$

where $F_A(i,j) = CC(i,j)|\vec{p}_i - \vec{p}_j|$, $F_R(i,j) = \frac{CC(i,j)}{|\vec{p}_i - \vec{p}_j|^2}$, and $CC(i,j)$ is the cross-correlation between the connectivity profiles of seed voxels i and j.

Force directed algorithms are known to be sensitive to local minima. However, finding the global minimum is not critical in our application. It suffices to reach a good approximation of the global minimum such that the initial position of the nodes does not drastically influence the outcome, thus providing a graph representation that is roughly reproducible in all the experiments. To approximate the equilibrium we proceed via steepest descent: starting from random positions, nodes are moved along the directions of the resulting force until the total potential energy is below a certain threshold value.

For evaluating the influence of the initial position of the nodes on the final outcome of the layout, we ran our algorithm multiple times on the same dataset. In each run, the initial position of the nodes was randomly generated and the size of the smallest rectangle containing the graph layout (whose longest side is parallel to a line connecting the two most distant nodes) was computed. This allowed us to compare the size of the graph. In addition, we compared the position of the barycentre of the graph across runs (cf. Fig. 5).

The force-directed algorithm yields a layout such that visual exploration of relations among nodes is intuitive. Several techniques are at our disposal for exploring such relations: simple graphical primitives drawn by the user and data-driven approaches such as local density of nodes. Graphical primitives can be used to label nodes, or groups of nodes, and to relate them with both the corresponding entries of the CCM and their locations in the brain (cf. Fig. 1). A local density of nodes is computed using the adaptive Gaussian kernel density estimation method proposed by Pisani (1996). By interpolating the positions of the graph nodes, a continuous density function is obtained and iso-density lines can be created (cf. Fig. 3).

We use the density of graph nodes as an indicator of groups of voxels with highly correlated connectivity profiles. By counting the number of density peaks (local maxima) of the density function, the number of clusters in the data can be inferred. Similarity of the connectivity profiles of voxels within a cluster, similarity of profiles in different clusters, and relationships among clusters can also be assessed by exploring the density map and the position of the graph nodes.

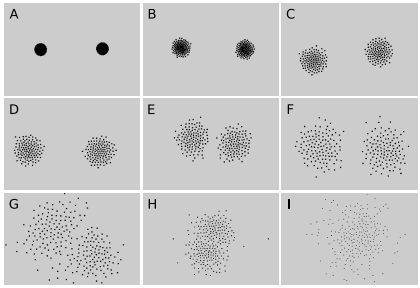

Figure 2: The effects of noise addition to the data presented in Fig. 1A. Noise ranges from 0% (A) to 80% (I). The size of the nodes differs in every picture because the spreading of the layout increases with the amount of noise.

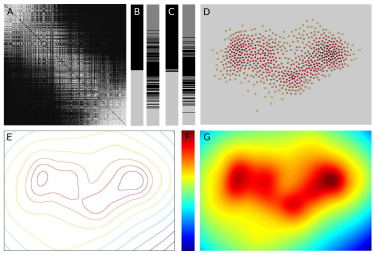

Figure 3: Comparison of spectral reordering and FDGL. A) Spectral reordered CCM of the area SMA/pre-SMA of subject 1. B) Results of k-means clustering on the CCM, for 2 and 3 clusters. C) Results of k-means clustering on the graph layout, for 2 and 3 clusters. D) The FDGL. E) Iso-density lines. F) The colour map used for mapping the density. G) The continuous density function.

5. Results

5.1. Synthetic data

Synthetic CCMs and their corresponding graph layouts are shown in Fig. 1. Nodes are labelled with gray values chosen by the user. Below each layout, the labelling of the nodes (seed voxels) is mapped along the columns of the CCMs, showing that the clusters perceived by the user reflect the actual compartments in the CCMs.

Fig. 1A shows that well-defined compartments of a CCM are represented by graph nodes that have been tightly grouped in circular regions: since the connectivity profiles within a compartment are similar, the attractive forces among the corresponding nodes have overcome the repulsive forces during the force-directed layout. A graph representation of overlapping compartments is shown in Fig. 1B, C. When two or more compartments of the CCM partially overlap, the layout shows a "bridge" between the two respective circular groups of nodes. The compactness and position of the bridges reflect the fact that the corresponding entries of the CCM are similar to each other and equally dissimilar from the two overlapping compartments. When the CCM presents a gradual change in similarities between the rows (columns) of a compartment and those of another compartment, the graph layout shows bridges that actually connect the circular regions, see Fig. 1D–F. A more gradual change in similarity (cf. Fig. 1F) is reflected by a more elongated layout.

A key feature of this method is that the topology of the graph layout reflects not only relations among voxels (i.e., presence of circular regions) but also relations between groups of voxels. For instance, as we see in Fig. 1C, E, voxels coloured in orange lay between the group of voxels coloured in red and the voxels coloured in yellow. This reflects the fact that their connectivity profiles are equally similar to those of the red voxels and those of the yellow voxels. On the other hand, the red and yellow groups are drawn far away from each other, meaning that their connectivity profiles are very different.

Figure 2 shows the influence of noise addition to the artificial dataset presented in Fig. 1A. Salt & pepper noise was added (using the Matlab© function *imnoise*) to the binary vectors before computing the CCM. Fig. 2A shows the graph layout of the original dataset without noise. In Fig. 2B–I the strength of the noise ranges from 10 to 80 percent. In this simple case it is still possible to visually detect the presence of two clusters, even when the noise affects up to 50 percent of the original data.

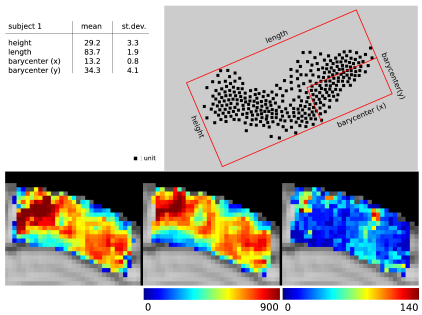

Figure 4: Reproducibility of our method. Top: statistics on the size of the smallest rectangle containing the graph layout and on the position of its barycenter. Bottom: result of a single run (left), average density over 50 runs (center) and standard deviation of the density in each voxel (right). Colour maps are shown in the bottom of the figure.

Table 1: Degree of overlap of k-means clustering results performed on the entries of the CCM and on the positions of the nodes in the graph layout, for subjects 1-5.

	subj.1	subj.2	subj.3	subj.4	subj.5
2 clusters	98.8%	95.9%	89.8%	84.6%	77.5%
3 clusters	76.2%	95.7%	84.8%	94.3%	98.7%

5.2. Subject data

Next we apply our method to data acquired from 13 healthy subjects, as described in section 3. The results of five subjects are shown in Fig. 3 and Fig. 7, while results of the remaining 8 subjects, as well as clustering and tractography results, are shown in the supplementary materials.

Figure 3 shows the results for SMA/pre-SMA data from subject 1. Figure 3A shows the spectral reordered CCM, as described in Johansen-Berg et al. (2004). Figure 3B shows results of k-means clustering of the CCM, for k=2 and for k=3 clusters. Gray values identify which rows in the CCM belong to the same cluster. Figure 3C shows results of k-means clustering of the graph layout. The clusters are mapped to the rows of the CCM. Figure 3D shows the graph layout, where nodes are coloured according to their density using the colour map shown in Fig. 3F. Figure 3E shows iso-density lines of the continuous 2D density function in Fig. 3G. The colour map in Fig. 3F was chosen to enhance the details of the 2D density function and to ease its readability.

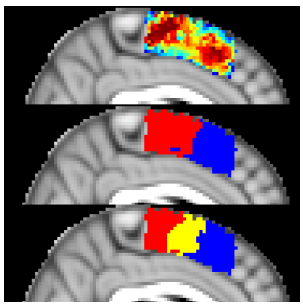

Figure 5: Top: plot of the graph densities of subject 1 on the MNI standard template. The regions with highest density overlap with the putative positions of SMA and pre-SMA. Centre: k-means clustering of SMA and pre-SMA into 2 clusters. Bottom: k-means clustering of SMA and pre-SMA into 3 clusters.

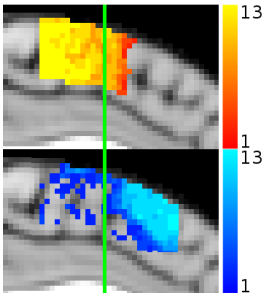

Figure 6: Frequency maps of the number of times each voxel is connected to the somatomotor area (top) and to prefrontal areas (bottom). Somatomotor area was defined by using a mask of the primary motor cortex, corticospinal tract, and lateral premotor cortex (BA6). The prefrontal area was defined by a mask of BA44, BA45, superior parietal lobe, and medial frontal gyrus (MFG). All masks except MFG were taken from Eickhoff's Anatomy Toolbox (Eickhoff et al., 2005). The mask of MFG was taken from the Harvard-Oxford Cortical Structural Atlas. Colour maps are shown on the right. The green line represents Y(MNI)=0.

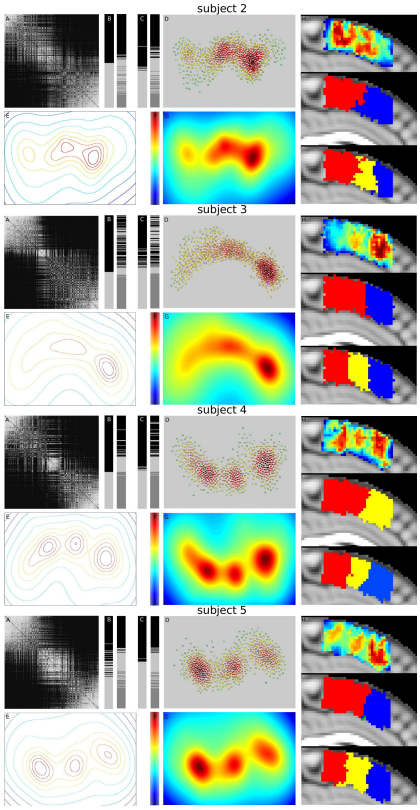

Figure 7: Spectral reordering, graph layout and MNI mapping of subjects 2-5.

Figure 4 shows statistics on the reproducibility of our method. The smallest rectangle containing the graph layout of Fig. 3 is shown in the top image of Fig. 4. This rectangle has its longest side parallel to the imaginary line connecting the two most distant nodes in the layout. Since orientation of the rectangle depends on the initial random position of the nodes and usually differs from one run to the other, the coordinates of its barycentre are computed according to the node of the rectangle most close to the barycentre itself. Average and standard deviations were computed running our algorithm fifty times on the data from subject 1. Figure 4 (bottom) shows statistics on the mapping of the densities to the anatomy. The result of a single run is shown bottom left. The average of the density maps over fifty runs is shown in the center. The corresponding standard deviations are shown on the right. We see that the position of the density peaks is similar across runs, and the standard deviation is small, although there is a little higher variability in the anterior region, dorsal with respect to the cingulate sulcus. These results indicate that our method is consistent across runs and highly reproducible.

Quantitative comparisons between k-means clustering of graph layout and k-means clustering of the CCM for subjects 1-5 are shown in Table 1. Both quantitative and qualitative comparisons of clustering results for the remaining 8 subjects are shown in the supplementary materials. We observe high consistency of the clustering results. This confirms the soundness of our approach and validates the hypothesis that high similarity of connectivity profiles corresponds to proximity in the graph layout.

Figure 5 shows the results of our method (top figure) and the result of k-means clustering on the graph (center and bottom figures) for the SMA/pre-SMA of subject 1. The density peaks shown by our method correspond to two groups of voxels, one in the posterior and one in the frontal premotor cortex, situated in the putative location of SMA and pre-SMA (the colour map is the same as Fig. 3). In agreement with column "subj. 1" of Table 1, the position of the density peaks best matches the results of k-means clustering with k equal to 2 rather than k equal to 3.

Figure 6 shows a group-wise comparison of tractography results on each voxel of the region of interest (ROI). Two ROIs were used as targets for probabilistic tractography. These ROIs are the somatomotor area (M1 and cortico-spinal tract) and the prefrontal areas BA44, BA45, superior parietal lobe, and medial frontal gyrus. The figure shows the frequency of connections from each voxel to the somatomotor area (top) and to the prefrontal brain regions (bottom) in the group of subjects. The voxels in the posterior regions of the ROI (putatively SMA) are connected to the somatomotor area in all subjects. The frequency of this connectivity pattern decreases as we move toward the front of the ROI (putatively pre-SMA). Accordingly, the region showing most frequent connectivity with prefrontal brain regions lays in the frontal part of our ROI. An important result is that these two connectivity patterns do not separate SMA/pre-SMA in two clear regions, but overlap in the central region of the ROI.

6. Discussion

The medial frontal cortex of the hemispheres, above the cingulate sulcus, is composed of different cortical regions featuring different architecture, functionality and anatomical connections with the rest of the brain. The most posterior sector is part of the primary motor cortex M1 (BA4, (Brodmann, 1909)). SMA, adjacent to M1, is involved in practical issues such as response selection or production. Pre-SMA is involved in the establishment or retrieval of sensory-motor associations (Tanji and Hoshi, 2008; Picard and Strick, 2001).

Tracer injection studies in macaque showed that SMA and pre-SMA are also distinct in terms of their anatomical connections with other parts of the brain: SMA has direct connections to the spinal cord and to M1; pre-SMA is not directly connected with final motor pathways but mainly with other regions of the prefrontal cortex. The two regions also differ in their projections to the cingulate, parietal and temporal cortex (Luppino et al., 1993).

Using results from animal studies on the macaque, Picard and Strick (1996) showed the presence of a relatively sharp subdivision between SMA and pre-SMA, localised around the plane through the anterior commissure (Y(MNI)=0). However, this boundary is idealised and based on evidence from animal studies; in a single subject examination no anatomical landmark can be used to effectively discriminate SMA from pre-SMA.

Connectivity based parcellation of SMA/pre-SMA was performed by Johansen-Berg et al. (2004), Anwander et al. (2007), Klein et al. (2007), and Nanetti et al. (2009). In Johansen-Berg et al. (2004), Klein et al. (2007) and Nanetti et al. (2009) the location of the boundary between SMA and pre-SMA was investigated. In Anwander et al. (2007) results of k-means clustering with k>2 were investigated and the possibility that k-means with k=3 could distinguish M1, SMA and pre-SMA was discussed.

Given the recovered similarity in the tractography results between the two posterior clusters, the lack

of a macroscopic landmark to distinguish SMA from M1, and the inter-subject variability on the location of the border M1/SMA (cf. probabilistic maps for M1 in the cytoarchitectonic atlas of Juelich (Eickhoff et al., 2005)), it is possible that at least part of M1 is included in the ROI. Nevertheless, the posterior boundary of the mask used for these parcellation studies (MNI Y=-22), introduced by Johansen-Berg et al. (2004) and used in several SMA/pre-SMA parcellation studies, localises the boundary between medial M1 and medial SMA as described in the probabilistic cytoarchitectonic atlas of Juelich (Eickhoff et al., 2005).

With our method we were not able to find well defined clusters covering the areas of SMA and pre-SMA, but rather a continuum of similarity among connectivity profiles when moving from the posterior to the anterior premotor area. Like in Nanetti et al. (2009) a substantial variability among the connectivity patterns of the subjects was found (cf. Fig. 6).

Subjects 1, 3, 6, 8, 11, 12 and 13 showed the presence of two density peaks. These divide the premotor cortex into two sagittal regions that could correspond to SMA and pre-SMA. The other datasets (subjects 2, 4, 5, 7, 9, and 10) showed the presence of three density peaks. Our speculative explanation is that these peaks, mapped sagitally along the premotor cortex, could refer to the presence of two groups of voxels with similar connectivity profiles (possibly SMA/pre-SMA) and a third region (a transitional area) where connectivity profiles are equally similar to those in SMA and to those in pre-SMA (cf. Fig. 1B, Fig. 6). This solution would be consistent with the results proposed in Nanetti et al. (2009), which indicate both a substantial inter-subject variability in the location of the border between SMA and pre-SMA, and the existence of a gradual change in connectivity between the posterior and the anterior regions of the ROI.

Another possible explanation is that either SMA or pre-SMA can be divided into two smaller regions with different connectivity patterns. Also, as proposed in Anwander et al. (2007), this could be due to the inclusion of M1 in the ROI.

A comparative study of structural and functional data (i.e., functional localization of M1, SMA, and pre-SMA) in each subject would be an asset for the understanding of the relation between functional regions and similarity in connectivity patterns.

7. Acknowledgements

We thank Valeria Gazzola and Nikola Valchev (Neuroimaging Center, University Medical Center Groningen) for allowing us to make use of datasets 6-13 which were acquired by them. V.G. was supported by a Veni grant from NWO; N.V. was supported by the Foundation for Science and Technology, Portugal SFRH/BD/47576/2008. All subjects gave written informed consent for their participation in the study.

References

Amunts, K., Schleicher, A., Brgel, U., Mohlberg, H., Uylings, H. B., Zilles, K., Sep 1999. Broca's region revisited: cytoarchitecture and intersubject variability. J. Comp. Neurol. 412, 319–341.

Anwander, A., Tittgemeyer, M., von Cramon, D. Y., Friederici, A. D., Knösche, T. R., 2007. Connectivity-based parcellation of Broca's area. Cereb. Cortex 17 (4), 816–825.

Barnard, S. T., Pothen, A., Simon, H. D., 1995. A spectral algorithm for envelope reduction of sparse matrices. Numerical Linear Algebra With Applications, 317–334.

Basser, P. J., Mattiello, J., Bihan, D. L., 1994. Estimation of the effective self-diffusion tensor from the NMR spin-echo. J. Magn. Reson. 103 (3), 247–254.

Basser, P. J., Pajevic, S., Pierpaoli, C., Duda, J., Aldroubi, A., 2000. In vivo fiber tractography using DT-MRI data. Magn. Reson. Med. 44 (4), 625–632.

Behrens, T. E., Johansen-Berg, H., May 2005. Relating connectional architecture to grey matter function using diffusion imaging. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 360, 903–911.

Behrens, T. E. J., Woolrich, M. W., Jenkinson, M., Johansen-Berg, H., Nunes, R. G., Clare, S., Matthews, P. M., Brady, J. M., Smith, S. M., 2003. Characterization and propagation of uncertainty in diffusion-weighted MR imaging. Magn. Reson. in Medicine 50 (5), 1077–1088.

Brodmann, K., 1909. Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues. Johann Ambrosius Barth Verlag, Leipzig.

Croxson, P. L., Johansen-Berg, H., Behrens, T. E., Robson, M. D., Pinsk, M. A., Gross, C. G., Richter, W., Richter, M. C., Kastner, S., Rushworth, M. F., Sep 2005. Quantitative investigation of connections of the prefrontal cortex in the human and macaque using probabilistic diffusion tractography. J. Neurosci. 25, 8854–8866.

Devlin, J. T., Sillery, E. L., Hall, D. A., Hobden, P., Behrens, T. E., Nunes, R. G., Clare, S., Matthews, P. M., Moore, D. R., Johansen-Berg, H., May 2006. Reliable identification of the auditory thalamus using multi-modal structural analyses. Neuroimage 30, 1112–1120.

Di Battista, G., Tollis, I. G., Eades, P., Tamassia, R., 1999. Graph drawing : algorithms for the visualization of graphs. Prentice Hall, Englewood Cliffs, NJ.

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., Zilles, K., May 2005. A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. Neuroimage 25, 1325–1335.

Geyer, S., Ledberg, A., Schleicher, A., Kinomura, S., Schormann, T., Brgel, U., Klingberg, T., Larsson, J., Zilles, K., Roland, P. E., Aug 1996. Two different areas within the primary motor cortex of man. Nature 382, 805–807.

Herman, I., Society, I. C., Melanon, G., Marshall, M. S., 2000. Graph visualization and navigation in information visualization: a survey. IEEE Transactions on Visualization and Computer Graphics 6, 24–43.

Johansen-Berg, H., Behrens, T. E., Robson, M. D., Drobnjak, I., Rushworth, M. F., Brady, J. M., Smith, S. M., Higham, D. J., Matthews, P. M., Sep 2004. Changes in connectivity profiles define function-

ally distinct regions in human medial frontal cortex. Proc. Natl. Acad. Sci. U.S.A. 101, 13335–13340.

Johansen-Berg, H., Gutman, D. A., Behrens, T. E., Matthews, P. M., Rushworth, M. F., Katz, E., Lozano, A. M., Mayberg, H. S., Jun 2008. Anatomical connectivity of the subgenual cingulate region targeted with deep brain stimulation for treatment-resistant depression. Cereb. Cortex 18, 1374–1383.

Jones, D. K., Simmons, A., Williams, S. C. R., Horsfield, M. A., 1999. Non-invasive assessment of axonal fiber connectivity in the human brain via diffusion tensor MRI. Magn. Res. Med. 42 (1), 37–41.

Klein, J. C., Behrens, T. E., Robson, M. D., Mackay, C. E., Higham, D. J., Johansen-Berg, H., Jan 2007. Connectivity-based parcellation of human cortex using diffusion MRI: Establishing reproducibility, validity and observer independence in BA 44/45 and SMA/pre-SMA. Neuroimage 34, 204–211.

Koch, M. A., Norris, D. G., Hund-Georgiadis, M., 2002. An investigation of functional and anatomical connectivity using magnetic resonance imaging. NeuroImage 16 (1), 241–250.

Luppino, G., Matelli, M., Camarda, R., Rizzolatti, G., Dec 1993. Corticocortical connections of area F3 (SMA-proper) and area F6 (pre-SMA) in the macaque monkey. J. Comp. Neurol. 338 (1), 114–40.

Marshall, M. S., Herman, I., 2001. An object-oriented design for graph visualization. Software - Practice and Experience 31, 739–756.

Mori, S., Crain, B. J., Chacko, V. P., vanZijl, P. C., 1999. Three dimensional tracking of axonal projections in the brain by magnetic resonance imaging. Ann. Neurol. 45 (2), 265–269.

Nanetti, L., Cerliani, L., Gazzola, V., Renken, R., Keysers, C., 2009. Group analyses of connectivity-based cortical parcellation using repeated k-means clustering. NeuroImage 47 (4), 1666–1677.

Parker, G. J., Haroon, H. A., Wheeler-Kingshott, C. A., Aug 2003. A framework for a streamline-based probabilistic index of connectivity (PICo) using a structural interpretation of MRI diffusion measurements. J Magn Reson Imaging 18, 242–254.

Passingham, R. E., Stephan, K. E., Kotter, R., Aug 2002. The anatomical basis of functional localization in the cortex. Nat. Rev. Neurosci. 3, 606–616.

Picard, N., Strick, P. L., Jan 1996. Motor areas of the medial wall: a review of their location and functional activation. Cereb Cortex 6 (3), 342–53.

Picard, N., Strick, P. L., Dec 2001. Imaging the premotor areas. Current Opinion in Neurobiology 11 (6), 663–72.

Pierpaoli, C., Jezzard, P., Basser, P. J., Barnett, A., DiChiro, G., 1996. Diffusion tensor MR imaging of the human brain. Radiology 201 (3), 637–648.

Pisani, A., 1996. A non-parametric and scale-independent method for cluster analysis – II. The multivariate case. Monthly Notices of the Royal Astronomical Society 278, 697–726.

Rushworth, M. F., Behrens, T. E., Johansen-Berg, H., Oct 2006. Connection patterns distinguish 3 regions of human parietal cortex. Cereb. Cortex 16, 1418–1430.

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., Bannister, P. R., Luca, M. D., Drobnjak, I., Flitney, D. E., Niazy, R. K., Saunders, J., Vickers, J., Zhang, Y., Stefano, N. D., Brady, J. M., Matthews, P. M., 2004. Advances in functional and structural MR image analysis and implementation as FSL. NeuroImage 23 (Supplement 1), S208 – S219, Mathematics in Brain Imaging.

Tanji, J., Hoshi, E., Jan 2008. Role of the lateral prefrontal cortex in executive behavioral control. Physiol Rev 88 (1), 37–57.

van Ham, F., van Wijk, J. J., 2004. Interactive visualization of small world graphs. In: Proc. IEEE Symposium on Information Visualization. pp. 199–206.

Vogt, B. A., Nimchinsky, E. A., Vogt, L. J., Hof, P. R., Aug 1995. Human cingulate cortex: surface features, flat maps, and cytoarchitecture. J. Comp. Neurol. 359, 490–506.

Vogt, O., Vogt, C., 1919. Ergebnisse unserer Hirnforschung. J. Psychol. Neurology 25, 277–462.

Woolrich, M., Jbabdi, S., Patenaude, B., Chappell, M., Makni, S., Behrens, T., Beckmann, C., Jenkinson, M., Smith, S., Nov 2008. Bayesian analysis of neuroimaging data in FSL. NeuroImage, 173–186.

A. Supplementary Material

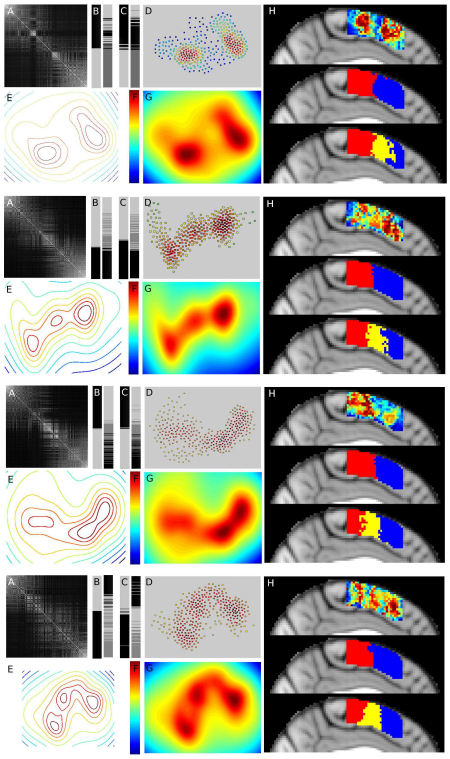

Figure A.1: Data of subjects 6, 7, 8, 9. Spectral reordering of the CCM(A). Comparison between k-means clustering with 2 and 3 clusters on CCM (B) and on graph layout (C); clusters are indicated by colours and mapped along the rows of the CCM. Graph layout (D). Iso-density lines (E) of the continuous density function (G). Images D, E, G are colour-coded according to the colour map F. MNI mapping of densities and k-means results (k=2 and k=3) on graph layout (H).

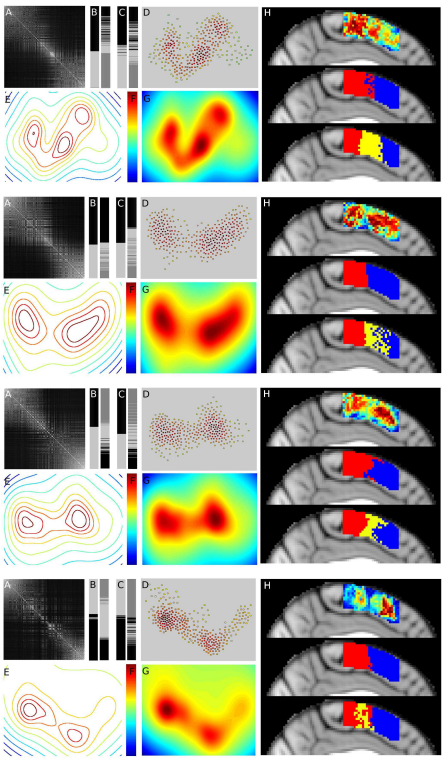

Figure A.2: Data of subjects 10, 11, 12, 13. Spectral reordering of the CCM(A). Comparison between k-means clustering with 2 and 3 clusters on CCM (B) and on graph layout (C); clusters are indicated by colours and mapped along the rows of the CCM. Graph layout (D). Iso-density lines (E) of the continuous density function (G). Images D, E, G are colour-coded according to the colour map F. MNI mapping of densities and k-means results (k=2 and k=3) on graph layout (H).

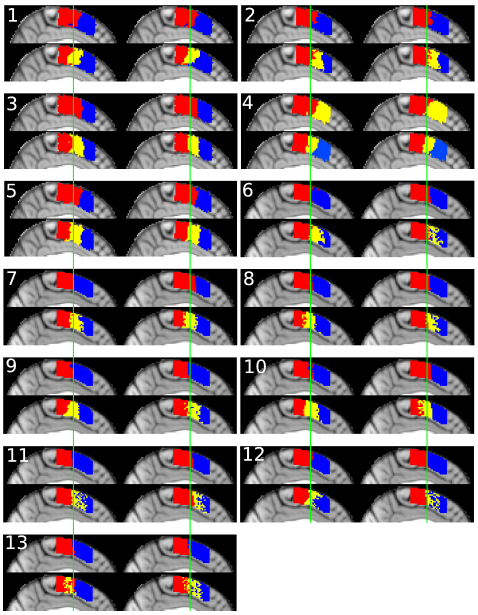

Figure A.3: Visual comparison between clustering of graph nodes and clustering of CCM for both 2 and 3 clusters for subjects 1-13. For each subject the four subfigures show: clustering of graph nodes for 2 clusters (top left); clustering of CCM for 2 clusters (top right); clustering of graph nodes for 3 clusters (bottom left); clustering of CCM for 3 clusters (bottom right). The green line represents Y(MNI)=0. See also Table 1 of the paper and Table A.1 below for quantitative results.

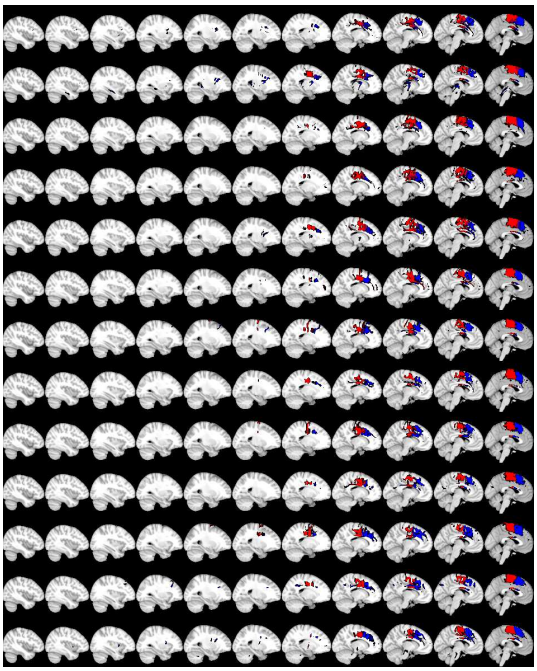

Figure A.4: Probabilistic tracking results seeded in k-means clusters found with k=2.

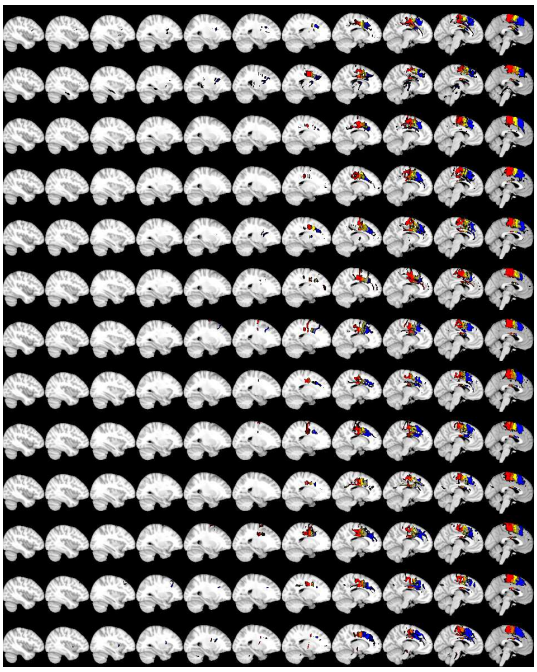

Figure A.5: Probabilistic tracking results seeded in k-means clusters found with k=3.

Table A.1: Overlap of k-means clustering results performed on the entries of the CCM and on the positions of the nodes in the graph layout, for subjects 6-13.

	subj.6	subj.7	subj.8	subj.9	subj.10	subj.11	subj.12	subj.13
2 clusters	96.1%	91.4%	98.3%	93.9%	87.4%	95.7%	96.1%	96.3%
3 clusters	77.5%	87.1%	74.7%	71.9%	85.3%	71.9%	83.7%	69.3%