Published in final edited form as:

Neuroimage. 2017 November 15; 162: 226–248. doi:10.1016/j.neuroimage.2017.09.005.

A Flexible Graphical Model for Multi-modal Parcellation of the Cortex

Sarah Parisot^{a,*}, Ben Glocker^a, Sofia Ira Ktena^a, Salim Arslan^a, Markus D. Schirmer^{b,c}, and Daniel Rueckert^a

^aDepartment of Computing, Biomedical Image Analysis Group, Imperial College London, 180 Queens Gate, London SW7 2AZ, UK

^bDepartment of Neurology, Massachusetts General Hospital, Harvard Medical School, Boston MA, USA

^cComputer Science and Artificial Intelligence Lab, Massachusetts Institute of Technology, Boston MA, USA

Abstract

Advances in neuroimaging have provided a tremendous amount of in-vivo information on the brain's organisation. Its anatomy and cortical organisation can be investigated from the point of view of several imaging modalities, many of which have been studied for mapping functionally specialised cortical areas. There is strong evidence that a single modality is not sufficient to fully identify the brain's cortical organisation. Combining multiple modalities in the same parcellation task has the potential to provide more accurate and robust subdivisions of the cortex. Nonetheless, existing brain parcellation methods are typically developed and tested on single modalities using a specific type of information. In this paper, we propose Graph-based Multi-modal Parcellation (GraMPa), an iterative framework designed to handle the large variety of available input modalities to tackle the multi-modal parcellation task. At each iteration, we compute a set of parcellations from different modalities and fuse them based on their local reliabilities. The fused parcellation is used to initialise the next iteration, forcing the parcellations to converge towards a set of mutually informed modality specific parcellations, where correspondences are established. We explore two different multi-modal configurations for group-wise parcellation using resting-state fMRI, diffusion MRI tractography, myelin maps and task fMRI. Quantitative and qualitative results on the Human Connectome Project database show that integrating multi-modal information yields a stronger agreement with well established atlases and more robust connectivity networks that provide a better representation of the population.

Keywords

Cortex Parcellation; functional Magnetic Resonance Imaging; diffusion Magnetic Resonance Imaging; Markov Random Fields; Connectomics; Brain connectivity

*Corresponding author. s.parisot@imperial.ac.uk.

1. Introduction

In-vivo neuroimaging and its recent advances have significantly contributed towards a thorough understanding of the brain's organisation. The brain's anatomy and cortical organisation can be investigated from the point of view of several sources of information: functional and diffusion Magnetic Resonance Imaging (fMRI and dMRI respectively) have allowed to infer the brain's structural and functional connectivity, while cortical folding or myelination patterns can be extracted from structural MRI. In particular, dMRI, fMRI and myelin maps have been largely studied for mapping functionally specialised cortical areas (Glasser et al., 2013; Craddock et al., 2012; Moreno-Dominguez et al., 2014), an objective which has been prominent for over a century (Zilles and Amunts, 2010).

There is strong evidence that a single modality is not sufficient to fully identify the brain's cortical mapping (Eickhoff et al., 2015). Indeed, cortical areas are believed to be defined by their microstructure, their connectivity and their function (Passingham et al., 2002). Because of this, a specific modality might not allow identification of the boundaries of all cortical areas. An accurate delineation of all cortical areas requires multiple modalities to exploit their complementarity and confirm the existence of certain boundaries. Yet, most existing brain parcellation methods are typically developed and tested on a specific type of information. Several popular parcellations have been derived from cortical folding (Tzourio-Mazoyer et al., 2002; Destrieux et al., 2010) or cytoarchitecture (Brodmann and Garey, 2005). More recently, connectivity-driven parcellations, in particular from resting state fMRI (rs-fMRI), have attracted a growing interest (Baldassano et al., 2015; Arslan et al., 2015; Blumensath et al., 2013; Cohen et al., 2008; Shen et al., 2013; Gordon et al., 2016; Thomas Yeo et al., 2011). The idea is to regroup vertices on the cortical surface based on how similar their connectivity profiles are. This is linked to the fact that parcellation can also be approached from a dimensionality reduction point of view for the study of brain connectivity networks. An essential step for the construction of these networks is the definition of the network nodes, which is typically done using parcellation techniques, where each parcel corresponds to a node. Connectivity-driven parcellations are expected to provide more accurate nodes than anatomical or random parcellations as they are derived directly from the connectivity data (Sporns, 2011).

In addition to the fact that a single modality cannot provide accurate cortical areas, mono-modal approaches are plagued by modality specific noise and biases which can significantly decrease the performance of parcellation algorithms. For instance, myelin maps only provide information on a subset of the cortex (highly myelinated regions), while diffusion MRI is a very indirect measurement of structural connectivity which is sensitive to the tractography algorithm used to recover the white matter fibres (Maier-Hein et al., 2016). dMRI is also subject to a gyral bias and geometrical issues associated with tractography algorithms. For these reasons, such algorithms tend to terminate fibres in gyri over sulci and align parcel boundaries with cortical folding (Van Essen et al., 2014). In contrast to this, rs-fMRI has the potential to provide reliable information across the cortex, but the modality has a poor signal to noise ratio (SNR) which can affect the accuracy and reproducibility of parcel boundaries. While the influence of noise can be strongly reduced when doing analysis on large groups, biases in the data (e.g. dropout in fMRI susceptibility regions) will constitute another

important issue. Combining multiple modalities in the same parcellation task has the potential to provide more accurate and robust parcellations, since the different modalities are intrinsically related.

Nonetheless, only a few methods have tackled the problem of combining multiple modalities in the parcellation task. A popular approach aims to combine structural and functional connectivity so as to construct a multi-modal connectivity matrix, typically with the aim of constructing less noisy fMRI connectivity matrices using dMRI information (Ng et al., 2012; Venkataraman et al., 2012). Assuming fMRI is directly dependent on structural connectivity, fMRI connections are only considered valid if they are supported by a physical connection measured by tractography. One of the main flaws of this approach is the fact that tractography itself can be unreliable and particularly prone to false positives (Maier-Hein et al., 2016), but also to false negatives. As a result, the functional connections estimated from structural support could be biased and inaccurate due to tractography errors. Furthermore, working directly on creating multi-modal connectivity matrices does not allow considering other types of information (e.g. myelin maps) for the construction of a multi-modal parcellation. Glasser et al. (2016) approached the problem differently and generated a group-wise parcellation using a semi-automated approach where an algorithm placed parcel boundaries based on expert decisions using aligned rs-fMRI, task fMRI, myelin maps, cortical thickness and topographically organised functional connectivity. Parcel boundaries were delineated if they were consistent across at least two modalities. This group level segmentation was then transferred to the single subject level using a classifier.

Markov/Conditional Random Fields (M/CRFs) offer a way of constructing very tunable models which is beneficial for the problem of parcellation in several aspects: control over the level of parcel smoothness, flexible model design, easy incorporation of prior knowledge and natural extension to multi-modal or group-wise analysis. MRFs have been used for a plethora of image processing applications, including image segmentation (Boykov and Funka-Lea, 2006), registration (Glocker et al., 2008) or image denoising (Geman and Geman, 1984). Through the MRF formulation, parcellation is cast as a labelling problem, where each label corresponds to a specific parcel and is to be assigned to a set of nodes in a graph representing brain geometry. In our setting, the vertices of the cortical surface mesh correspond to the nodes of the MRF model. This labelling problem is solved by minimising the MRF energy which comprises unary and pairwise terms. The unary terms describe the likelihood of assigning a node to a specific parcel (i.e. label) while pairwise terms model the interactions between neighbouring nodes and typically act as smoothing priors. A significant advantage of MRF models is that they do not make any assumption on the input data. In other words, several modalities can be considered and processed within the same framework. The use of MRFs for rs-fMRI driven parcellation has recently been the subject of several publications (Lashkari et al., 2010; Ryali et al., 2013; Honnorat et al., 2015; Parisot et al., 2015, 2016b). The proposed methods describe the likelihood of assigning a node to a specific parcel as the correlation to the parcel's average connectivity profile. Ryali et al. (2013) proposed an EM-like approach tailored for fMRI data that iteratively estimated the parcellation using graph cuts (Boykov and Funka-Lea, 2006) and estimated the unary cost's parameters based on the current parcellation status. Honnorat et al. (2015) introduced the notion of parcel centres which are associated with a representative connectivity profile of the

parcel. The MRF unary cost describes the correlation of a node's profile with the parcel centre's profile. This approach considers all nodes as potential parcel centres, which can be computationally expensive and subject to noise. A connectedness prior is introduced in the form of a star shape prior. Both approaches determine the number of parcels using label costs that estimate the number of necessary labels given a penalty. Despite the appeal of inferring the number of parcels from the data, this setting boils down to replacing the intuitive choice of the number of parcels with a different parameter of unknown impact.

In Parisot et al. (2015, 2016b), we introduced Graph-based Multi-modal Parcellation (GraMPa), an iterative MRF framework designed to handle the large variety of available multi-modal data: (1) We propose a set of modality specific unary cost functions that allow parcellating the brain according to different modalities and their properties. This allows to construct modality specific parcellations that can be compared without biases introduced by the use of different parcellation methods. (2) We extend the proposed framework to the context of multi-modal parcellation through the introduction of a multi-modal merging step. At each iteration, we obtain a set of parcellations from different modalities and fuse them based on prior knowledge of the modalities' reliabilities and their interactions. The fused parcellation is used to initialise the next iteration, forcing the parcellations to converge towards a set of coherent yet modality specific parcellations. This provides a framework that allows 1) to directly compare different modalities, 2) to construct a multimodal parcellation and 3) to increase the robustness of mono-modal parcellations through the introduction of additional and complementary information.

In Parisot et al. (2015, 2016b), we focused on parcellation at the single subject level. In this paper, we provide a more general and detailed formulation of the model and investigate the impact of the proposed method on group-level analysis. This allows to explore multiple multi-modal associations and to compare obtained parcellations to well established atlases. In particular, we investigate the use of task activation maps as an additional source of information, which are too noisy to be used at the single subject level. We evaluate the ability of our framework to parcellate based on different types of inputs and exploit this property to provide an experimental set-up that quantitatively evaluates multi-modal agreements. Additionally, we evaluate the impact of integrating multiple modalities on the delineation of cortical areas, as well as from the point of view of network analysis. Our experiments on data from the Human Connectome Project (HCP) database using rs-fMRI, dMRI, myelin maps and task activation maps show that GraMPa yields a stronger agreement between modalities and more robust connectivity networks that provide a better representation of the population.

2. Material and Methods

The proposed iterative multi-modal model is illustrated in Fig. 1. In this section, we first introduce the mono-modal setting where one modality is parcellated using an MRF unary cost tailored for this specific input data. We then introduce the multi-modal extension, which is designed as a step that fuses information from multiple modalities. Finally, we describe the methodological details for our evaluation set-up.

We introduce below a set of notations that will be used throughout this paper. We aim to parcellate the brain's cortical surface into a set of K distinct subregions, where K is known a-priori. We represent the cortical surface as a triangular mesh $\mathcal{S} = \{\mathcal{V}, \mathcal{E}\}$, where each vertex $v \in \mathcal{V}$ is to be assigned to a specific parcel. N_v is the number of vertices and e corresponds to the set of edges of the mesh. χ is a feature matrix that can be defined as the output of dMRI tractography algorithms or as fMRI time series and $\chi(v)$ is the row of the feature matrix at vertex v .

2.1. Dataset and Preprocessing

Our dataset comprises two sets of 100 unrelated subjects randomly selected from the S500 release (November 2014) of the Human Connectome Project (HCP) database. The first set (54 female, 46 male adults, aged 22–35) is used as our main database. The second set (50 female, 50 male adults, aged 22–35) is used for reproducibility purposes, to evaluate whether similar results can be obtained from one group to the next. Both sets of 100 subjects are pre-processed using the same strategy.

The structural, diffusion and functional data have been preprocessed following the HCPs minimal preprocessing pipelines (Glasser et al., 2013). The cortical surface mesh $\mathcal{S}$ consists of 32k nodes with a 2mm spacing. All cortices are registered to a common reference space using sulcal depth information (Robinson et al., 2014).

The diffusion MR images have been acquired using a multi-shell approach, with three shells at b-values 1000, 2000, and 3000 s/mm^2 and 90 gradient directions per shell (Sotiropoulos et al., 2013). We obtain the tractography matrix using FSL's bedpostX and probtrackX methods (Behrens et al., 2007; Jbabdi et al., (2012)). The former estimates the orientation of the fibres passing through each voxel of the brain volume while the second performs probabilistic tractography based on the estimated fibre orientations. The probabilistic tracking is performed on the native mesh (before registration) representing the gray/white matter interface. We fitted three fibre compartments per voxel and 5000 streamlines are seeded from each of the white matter surface vertices and terminated on the pial surface. We use a step length of 0.3. The obtained tractography matrix records the number of streamlines that reached the rest of the mesh. To reduce the bias towards short range connections inherent to tractography, we compute the element-wise log transform of the tractography matrix (Jbabdi et al., 2009; Moreno-Dominguez et al., 2014; Parisot et al., 2016a) which reduces the dynamic range of fibre counts.

We perform rs-fMRI driven parcellation using timeseries from a 30 minutes acquisition. The 4D rs-fMRI volumes are projected onto the cortical surfaces which associates each mesh vertex with a rs-fMRI timeseries. The dataset was preprocessed and denoised (using the HCP's ICA + FIX approach (Salimi-Khorshidi et al., 2014)) by the HCP structural and functional minimal preprocessing pipelines (Glasser et al., 2013). Subsequently, we temporally normalise each timeseries to zero-mean and unit-variance.

The remaining modalities considered in this paper (myelin maps, sulcal depth maps, task fMRI and cytoarchitectural map) are obtained as part of the HCP's S500 release. Myelin maps are computed as the ratio of T1 and T2 structural MRI intensities. Probabilistic

cytoarchitectural maps (Fischl et al., 2008; Zilles and Amunts, 2010; Amunts et al., 2007) were registered using freesurfer to each individual's cortical surface and binarised throughout the HCP pipeline.

Group average data is obtained as follows: myelin maps are obtained by averaging all subjects' maps. Similarly, the population average structural connectivity matrix is computed by averaging all the subjects' tractography matrices. The average cytoarchitectural map, illustrated in Fig. 11, is obtained by majority voting from the registered individual maps, while we compute task fMRI group-level activation maps using FSL's standard tools (FEAT) that use general linear modelling to construct activation maps (Barch et al., 2013). The analysis is carried out across sessions (single subject activation maps) and then across subjects (group-wise activation map). Task based parcellations are generated using the Z-statistics maps. Average rs-fMRI pseudo-timeseries are computed using FSL's incremental group PCA (MIGP) (Smith et al., 2014), which can efficiently approximate the PCA output applied to the concatenation of all subjects' timeseries using computationally tractable resources. We use a dimensionality of 2300 eigenmaps, which correspond to the concatenation of same day scans. Connectivity networks are then obtained by computing the Pearson's correlation coefficient between the pseudo time-series.

2.2. Modality specific Markov Random Field Model

We cast the parcellation task as a labelling problem on the mesh $\mathcal{S}$. Each surface vertex $v \in \mathcal{V}$ is to be assigned a label $l_v \in \llbracket 1, K \rrbracket$, where each label corresponds to a parcel assignment.

2.2.1. Model—A labelling (corresponding to a parcellation) can be estimated by minimising the energy of an MRF defined by:

$$E_{MRF} = \sum_{v \in \mathcal{V}} U_v(l_v) + \sum_{v \in \mathcal{V}} \sum_{w \in \mathcal{N}(v)} U_{v,w}(l_v, l_w) \quad (1)$$

Here $\mathcal{N}(v)$ represents the neighbouring vertices of v , defined from the structure of the triangular cortical surface mesh $\mathcal{S}$. The unary term $U_v(l_v)$ describes the likelihood of assigning a vertex to a specific parcel. In order to define parcel specific costs, we associate each parcel with a parcel centre c . The parcel centre associates each parcel with a characteristic that uniquely defines it and corresponds to a representative vertex which is associated with an average feature $\chi(c)$. The definition of parcel attributes and their corresponding modality specific unary costs are described in Sec. 2.2.2. The pairwise term $U_{v,w}(l_v, l_w)$ acts as a smoothing prior. Its role is to describe the pairwise interactions between neighbouring vertices. It is designed here as a Potts model, introducing a penalty β when neighbouring nodes are assigned different labels.

$$U_{v,w}(l_v, l_w) = \begin{cases} 0 & \text{if } l_v = l_w \\ \beta & \text{otherwise} \end{cases} \quad (2)$$

Intuitively, β controls the smoothness the parcel boundaries and their contiguity. A low value of β will yield more precise but also noisier boundaries as well as discontinuous parcels. A high β value can excessively smooth parcel boundaries and yield large and fewer parcels than what is specified by parameter K (number of sought parcels). We provide an example of the influence of β on rs-fMRI driven parcellations on visual and quantitative results in the supplementary material.

Adopting an iterative approach allows to identify relevant parcel centres. Given an initial parcellation, it is iteratively updated by alternating between two steps: 1) updating the parcel centre based on the current parcellation, 2) parcellating the brain through MRF optimisation of the energy introduced in Eq. 1.

2.2.2. Modality-specific cost functions—The strength of adopting an MRF model is the possibility to perform parcellation with different modalities by simply changing the unary cost. We design a set of unary cost-functions, each adapted to a specific type of modality. Each cost-function is associated to a specific parcel centre definition. We consider three different types of input data: 1) *High SNR connectivity data*, where each surface vertex is associated to brain connectivity information with limited amount of noise. An example is to associate each vertex with a high resolution dMRI tractography profile. 2) *Low SNR connectivity data* will typically be fMRI data (task or resting state), where each vertex is associated with fMRI timeseries. The low SNR requires a different, more robust cost function than the one considered in case 1). 3) *Non connectivity data* associates each vertex with data projected on the cortical surface and requires a different approach to parcellation. This data does not contain connectivity information per se but characterises the similarity/dissimilarity of neighbouring nodes.

High SNR connectivity data.: A schematic representation of this unary cost is shown in Fig. 2a. The parcel centre is defined as a vertex $v \in \mathcal{V}$. Considering the parcellation at iteration t , the parcel centre c^1 is defined as the vertex within parcel l that has the highest similarity to the rest of the parcel. The unary cost $U_v(l_v)$ is then defined as follows:

$$U_v(l_v) = 1 - \rho(\chi(v), \chi(c^1)) \quad (3)$$

Where $\rho(.,.)$ is a measure of similarity between the two rows of the connectivity matrix χ . Here, we define ρ as the Pearson's correlation coefficient, a very common measure of similarity for connectivity data. This type of unary cost has been used successfully in Parisot et al. (2015) for dMRI based parcellation. This cost has the advantage of being very simple, however, it relies on a single node to evaluate the likelihood of a parcel assignment which can make it sensitive to noise. As a result, it is well suited to robust data such as high

resolution dMRI connectivity matrices, but can be too sensitive for noisier data such as fMRI.

Low SNR connectivity data.: We design a more robust cost for noisier data such as fMRI time series. A schematic representation of this unary cost is shown in Fig. 2b. The parcel centre is again defined as the vertex with the highest similarity to the rest of the parcel. However, it is assigned an average connectivity profile $\chi(c^l)$ that is the average time series within a neighbourhood $\mathcal{N}$ around c^l :

$$\chi(c^l) = \sum_{w \in \mathcal{N}(c^l)} \frac{1}{|\mathcal{N}(c^l)|} \chi(w) \quad (4)$$

$\mathcal{N}(c^l)$ is defined as the set of N closest vertices to c^l that belong to parcel l . Closeness is determined by computing the shortest path on the surface mesh $\mathcal{S}$, where edges are weighted by the similarity between vertices. The unary cost is then defined following Eq. 3:

$$U_v(l_v) = 1 - \rho(\chi(v), \chi(c^l)) \quad (5)$$

Increasing the value of N makes the unary cost more sensitive to the current parcellation (and therefore to initialisation), while decreasing N is more sensitive to noise. In the context of rs-fMRI driven parcellation, $\chi(w)$ corresponds to the computed pseudo-timeseries as described in Sec. 2.1. A similar type of unary cost was used in Honnorat et al. (2015), where N was set as the number of vertices within the parcel. While using functional connectivity networks as connectivity profiles would provide more robust parcellations, it would significantly slow down the parcellation due to the correlation operation. A similar approach can be used for group averages with low SNR (e.g. non registered fMRI data) by computing the average correlation between the vertex to label and the N closest nodes around the parcel centre. Figure 6 shows the reduced influence of noise using this type of unary cost.

Non connectivity data.: The last type of input data we consider is non connectivity data projected on the cortical surface. Considering HCP data, myelin, sulcal depth or task activation maps are typical examples. In this setting, we can take inspiration from MRF-based image segmentation methods. We define the parcel centre as the geometric centre of a parcel, which we obtain by erosion. Each parcel centre c^l is assigned the average vertex value within the parcel l . The unary cost, as illustrated in Fig. 2c is the shortest path on the cortical surface mesh $\mathcal{S}$ between the centre and the vertex v . The edges are weighted by the absolute difference between the values of the parcel centre and the vertices s_i along the path:

$$U_v(l_v) = \sum_{s_i, v} |D(s_i) - \bar{D}(c^l)| \quad (6)$$

Where $D(\mathbf{v})$ is the non connectivity data value at vertex $\mathbf{v}$ and $\bar{D}(c^1)$ is the average value within parcel 1 associated with parcel centre $\mathbf{c}^1$. Computing the absolute difference between the vertices' values and the parcel centre (i.e. not using shortest paths) is also an option, but it does not yield spatially contiguous parcels.

Using the same framework allows to construct a set of parcellations for different modalities in a straightforward way and investigate their similarities. However, each parcellation converges independently and the different parcels do not have direct correspondences between modalities. In addition, each parcellation is affected by modality specific shortcomings: for instance, fMRI's low SNR, dMRI's biases with gyri and short range connections or the lack of information provided by myelin maps in large parts of the cortex.

In order to address these shortcomings, we extend the proposed framework to the multi-modal setting where the different modalities are forced to converge towards a set of mutually informed modality specific parcellations, where correspondences are established. This is done by initialising each iteration with a multi-modal parcellation obtained by fusing the individual parcellations.

2.3. Multi-modal Integration

Multi-modal integration is carried out through the introduction of an additional step in the iterative approach described in Sec. 2.2. At each iteration t , we now consider that the parcellation update is carried out in parallel for a set of N_{mod} modalities. Each modality mod is then associated with an MRF labelling solution $\mathbf{l}^{mod}$ obtained by optimising Eq. 1 with the appropriate unary cost.

We then introduce a merging step, where the set of modality-specific parcellations are fused into a joint parcellation. The main objective is to increase the robustness of the mono-modal parcellation exploiting the similarities between modalities, while at the same time relying on more stable modalities when others are too noisy and/or uninformative.

2.3.1. Merging Model—To this end, we model the merging problem as a new MRF optimisation problem so as to allow flexibility with respect to the modalities considered, their interaction and the quality of the data.

$$\underset{\mathbf{l}}{\operatorname{argmin}} E^m(\mathbf{l}) = \sum_{\mathbf{v} \in \mathcal{V}} U_{\mathbf{v}}^m(l_{\mathbf{v}}) + \sum_{\mathbf{v} \in \mathcal{V}} \sum_{\mathbf{w} \in \mathcal{N}(\mathbf{v})} U_{\mathbf{v}, \mathbf{w}}(l_{\mathbf{v}}, l_{\mathbf{w}}) \quad (7)$$

Here, the pairwise cost is again a Potts model as described in Eq. 2. MRFs allow this flexibility in the design of the unary cost function $U_{\mathbf{v}}^m(l_{\mathbf{v}})$, which describes here the merging criteria.

We associate each modality with a reliability map α^{mod} that describes how reliable and informative it is locally. A low reliability could be linked to noise in the data, or shortcomings inherent to specific modality processing (the gyral bias of dMRI tractography

for instance). We propose to define the unary cost based on the different modalities labellings I^{mod} and their reliability maps:

$$U_v(1_v) = \min_{\text{mod}} \left(1 - \alpha^{\text{mod}}(v) \delta(1_v, I_v^{\text{mod}}) \right) \quad (8)$$

Here, δ is the Kronecker delta function. This unary cost exploits the reliability maps and ensures that labels selected by a given modality are given a better or worse likelihood based on the modality's local reliability. This set-up allows one modality to have more influence than the others on the joint parcellation if it is deemed very reliable. A typical example of locally reliable data would be pre-existing atlases such as cyto-architectural maps.

2.3.2. Modality-specific merging reliability maps—In this section, we detail the design of a set of modality specific weights. We aim to merge modalities in a way that acknowledges the shortcomings and assets of each modality. Diffusion and resting-state functional MRI are the only modalities considered here that have the potential to provide information across the whole cortex, while modalities such as task fMRI and myelin maps are only informative locally.

Resting-state fMRI.: Assessing the reliability of functional MRI data is a very challenging task. The main objective is to differentiate signal from noise, and in particular to identify which regions are dominated by noise. Intuitively, comparing different rs-fMRI acquisitions could provide this kind of information. However, the temporal dynamics of rs-fMRI signals remain unclear and could introduce differences between two acquisitions that do not correspond to noise. Furthermore, regions that are heavily connected to the rest of the brain (i.e. provide more information) could contain more noise than poorly connected regions (containing very little information and signal). As a result, while comparing two acquisitions may provide an indication of local reliability, this measurement may not be directly comparable between brain regions. We therefore associate rs-fMRI data with a fixed cost C that covers the whole brain so as to avoid losing information by inaccurate noise estimations. In this set-up, rs-fMRI is defined as the default modality that will drive the multi-modal parcellation when others are deemed not reliable.

Diffusion MRI.: Many challenges associated with dMRI and tractography data could be considered and incorporated in the reliability maps. Here, we propose to focus on one specific shortcoming of the dMRI processing pipeline that has been observed to strongly influence parcellation boundaries. This major issue associated with tractography algorithms is referred to as the gyral bias (Van Essen et al., 2014): tractography streamlines terminate preferentially within gyri instead of sulci. On top of that, several geometrical effects of tractography tend to cause a preponderance of connectivity transitions to lie on gyral crowns. This has a strong influence on the boundaries of dMRI driven parcellations that tend to align with cortical folding. In order to identify regions affected by such biases, we consider probabilistic tractography and the corresponding connectivity matrix χ , where each entry $\chi(v, w)$ counts the number of streamlines sent from vertex v that reach vertex w . The obtained matrix is not symmetric, one of the main reasons being the gyral bias. A column of

the matrix will count the number of times a specific vertex v is reached, a number which will be higher for vertices located in gyri. On the other hand, a row of the matrix estimates which vertices can be reached from v and will not be sensitive to whether the location of v is in gyri or sulci. As a result, we can estimate which vertices are influenced by the gyral bias by comparing rows and columns of χ

$$\alpha^{dMRI}(v) = \frac{\sum_{w \in \mathcal{V}} \chi(v, w)}{\sum_{w \in \mathcal{V}} \chi(w, v)} \quad (9)$$

This reduces the influence of vertices where fibres tend to terminate, i.e. vertices which influence the positioning of parcel boundaries. As a result, dMRI parcellation boundaries affected by this bias will have little influence on the boundaries of the merged multi-modal parcellation. In these regions, dMRI essentially acts as a smoothing prior, providing high resolution information on which vertices should be regrouped in the same parcel based on similar connectivity.

Non connectivity data.: The main modalities of this type which we consider are myelin maps, task-fMRI activation maps and cortical thickness maps. These modalities provide information that only cover a subset of the cortical surface, where sharp variations are observed (e.g. strong activation or myelination). In order to ensure that the multi-modal parcellation aligns with these sharp transitions, the influence of non-connectivity data is designed to be proportional to the smoothed gradient maps of the non connectivity data. Gradient maps are computed using the HCP's workbench¹, which uses a regression between the positions of the vertices (unfolded onto a tangent plane) and their values. The obtained gradient maps are subsequently smoothed on the surface using a Gaussian kernel with sigma set to 8. As a result, a high reliability is associated with strong gradients (i.e. strong evidence of a transition between function and/or structure), while uninformative regions will have no influence.

The computed weights, as defined above, are defined in very different ranges and scales. In order to ensure that the reliability weights of all modalities are directly comparable, all modality specific weights are rescaled between 0 and 1. To further increase comparability, the histograms of the weights' values are equalised to match a normal distribution with mean zero and standard deviation 1.

2.3.3. Merging modalities with fusion moves—Despite the plethora of existing MRF optimisation methods, minimising multi-label MRF energies remains a difficult task with no optimality guarantees (Lempitsky et al., 2010). The concept of fusion moves, introduced in Lempitsky et al. (2010) casts the MRF optimisation problem as the task of combining an ensemble of suboptimal solutions. This is done through a set of iterative simple binary MRF sub-problems that will iteratively update the global solution with respect to one suboptimal solution. Fusion moves have the advantage of being very fast and scalable, and often yield globally optimal solutions to multi-label problems. Fusion move optimisation is particularly useful for our problem: the mono-modal parcellations can be seen as a set of suboptimal solutions to the multi-modal problem.

Each binary optimisation is called a fusion move and considers two suboptimal labellings l_i and $l_i^c \in \mathcal{L}^{\mathcal{V}}$, where l_i^c is the combined labelling obtained at the current fusion move t . We seek to label the cortical mesh $\mathcal{M}$ with a binary label $b \in \{0, 1\}^{\mathcal{V}}$ so as to update l_i^c as $l_{t+1}^c(b) = l_i^c \circ (1 - b) + l_i^c \circ b$, where $\circ$ is the Hadamard product. The fusion move is carried out by minimising the binary MRF energy $E^b(b) = E^m(l^c(b))$ which can be done with simple MRF optimisation techniques. The next fusion move considers a new modality and the current combined parcellation l^c as the suboptimal labellings to merge. Fusion moves are repeated for all modalities until the combined labelling is no longer updated. Once the joint parcellation is obtained, it is used to initialise the next iteration of the multi-modal parcellation scheme. This forces the multiple modalities to converge together towards a set of modality specific yet coherent parcellations.

2.4. Evaluation: quantitative measures

The evaluation of brain parcellation tasks is a challenge in itself since there is no ground truth to compare to. While there have been accepted boundaries for certain regions of the brain, there is still a lot of uncertainty regarding the organisation of the brain as a whole, and how accurately a parcellation driven by one modality should match a given cortical area. In addition, a parcellation method yielding very accurate cortical boundaries may not be the best suited for representing brain connectivity networks, as functional/structural organisation at coarse resolutions may differ from cortical organisation. The accuracy of MRF based parcellation methods as a means of generating homogeneous parcellations from connectivity data (fMRI and dMRI) has been investigated and demonstrated in Parisot et al. (2016b, 2015); Honnorat et al. (2015).

Here, we focus our evaluation on the impact of the multi-modal integration, particularly how its flexibility can be exploited, and how it can improve upon the mono-modal parcellation. To this end, we focus on 1) the flexibility and reproducibility of the model (e.g. the ability to use different modalities), 2) network analysis, 3) the impact of dimensionality reduction and 4) the ability to delineate cortical areas (e.g. comparison to delineated cyto-architectural regions and other modalities). The flexibility of the model is explored by generating parcellations from broad set of modalities (dMRI, rs-fMRI, task fMRI, myelin and cortical folding), including multiple multi-modal combinations. The parcellations are compared across modalities and their quality is visually assessed. Reproducibility is assessed by computing parcellations from two different groups of 100 subjects.

2.4.1. Network analysis—We evaluate the impact of multi-modal integration on functional and structural network construction. This is assessed by 1) comparing the networks constructed using two different rs-fMRI sessions and 2) evaluating how well the group average represents the population by comparing all individual connectivity networks to their group average, similarly to what we proposed in Parisot et al. (2016a). The similarity between connectivity networks that we compute is the sum of absolute differences (SAD) between network edges, which can be seen as a graph edit distance (Raj et al., 2010) when nodes correspondences have been established and, as a result, no edit operations need to be considered for the node elements.

We expect an accurate and robust parcellation to yield similar networks between sessions, as well as an average parcellation that accurately summarises the features shared by the population.

For a given parcellation, rs-fMRI connectivity networks are computed by averaging the timeseries of all vertices in each parcel. The connectivity matrix is then obtained by computing the partial correlation between the timeseries of all pairs of nodes and applying the r-to-z Fisher transform to normalise and allow for accurate comparisons. Due to the fact that parcellation resolutions can slightly vary across methods (the method can yield empty parcels), the SAD between rs-fMRI matrix edges is normalised by the matrix size.

dMRI networks are computed from the probabilistic tractography matrix by summing the number of fibres that reach each parcel. Considering two parcels P_i and P_j , we account for the varying parcel sizes by normalising the fibre count as (Hagmann et al., 2008) :

$$w_{i,j} = \frac{2}{S_i + S_j} \sum_{v_i \in P_i} \sum_{v_j \in P_j} \chi(v_i, v_j) \quad (10)$$

where S_i and S_j are the sizes of parcels P_i and P_j respectively.

The average dMRI network is constructed by averaging all structural networks. The matrices are then normalised to sum to 1 to make the comparisons less sensitive to matrix size and the variability in fibre count. The average rs-fMRI network is constructed following the method described in (Salvador, 2004). We first compute an average connectivity matrix by averaging the partial correlation matrices of all subjects. We then identify non-zero partial correlation edges by testing the null hypothesis that the average partial correlation is zero for a given edge. This is done using one-sample t-tests for each edge of the individual partial correlation matrices using the False Discovery Rate approach to correct for multiple comparisons.

2.4.2. Comparison between modalities and parcellations—We evaluate the agreement between a parcellation and another modality (e.g. myelin or task activation maps) by generating a coarse parcellation from the considered modality using the proposed mono-modal MRF model. The parcels of both parcellations are quantitatively compared and matched using the measure of spatial overlap proposed in Bohland et al. (2009). This measure is non symmetric, and evaluates the proportion of one region i that is contained in another region j . We refer to r_i and r_j as the ensembles describing the vertices that belong to regions i and j respectively. The similarity measure P_{ij} is then defined as:

$$P_{ij} = \frac{|r_i \cap r_j|}{|r_j|} \quad (11)$$

where $|r|$ is the number of voxels in ensemble r . A symmetric measure is also defined as $O = \sqrt{P_{ij}P_{ji}}$.

We match the two parcellations by selecting the parcels that have the highest similarity scores O . It should be noted that several parcels can be matched to the same one and therefore merged into a larger parcel. We then compute the Dice Similarity Coefficient (DSC) between the matched parcels. This is more flexible than the commonly used overlap methods as it does not search for perfect overlap, but also for inclusion of a parcel in another. This approach is particularly relevant when comparing parcellations with a very different number of parcels.

Coarse parcellations (e.g. the cytoarchitectural map) are also compared to the *GraMPa* parcellation following this approach. Overlap/DSC based measures with parcel merging have been very popular for the evaluation of parcellation schemes (Eickhoff et al., 2015; Arslan et al., 2017) and are therefore used here to facilitate comparison with other methods. Nonetheless, it should be noted that the parcel merging step could artificially improve the values of the computed DSC.

2.4.3. Dimensionality reduction quality—Last but not least, we evaluate the quality of our parcels with respect to the underlying data they are computed from. Our objective is to demonstrate how multi-modal parcellations can be used reduce the overall error when the same parcellation is used to reduce the dimensionality of a set of modalities. We aim to obtain parcellations that represent the dataset as well as possible.

A parcels average value should be as close as possible to all the values of the vertices within the parcel. For each modality and a given parcellation, we evaluate how well the data can be approximated when all values within the same parcel are represented by the average parcel value.

This is done by computing, for each modality, a merged low dimensional matrix where the values within the same parcel are averaged. For dMRI, the merged matrix of size $K \times K$ is constructed as described in Sec.2.4.1 and Eq.10. The rs-fMRI timeseries and non connectivity values (size $K \times \text{dim}$, where dim is the dimensionality of the input data) are simply averaged within each parcel. An illustration of the construction of χ_{av} for dMRI data is shown in Fig. 3.

We then compute a high dimensional approximation χ_{av} of the original input data (dMRI tractography matrix, rs-fMRI timeseries and task activation and myelin maps) by assigning the parcel average values computed in the merged matrices to all vertices in the same parcel.

For dMRI data, we compute the Kullback-Leibler Divergence (KLD) between this matrix and the original tractography matrix χ that are normalised to be probability mass functions. The KLD measures how much information is lost by approximating the tractography matrix χ with χ_{av} . A low KLD corresponds to a better parcellation with respect to dMRI data.

For rs-fMRI and non-connectivity maps, the same approach cannot be used as the data can contain important negative values and cannot be assimilated to a probability mass function as naturally. Instead, we compute the Root Mean Square Error (RMSE) between the original

and approximate matrices. A low RMSE corresponds to a better parcellation with respect to the data.

3. Results

3.1. Experimental setup

In order to highlight the flexibility of our model, we compute parcellations using a set of different modalities, namely: diffusion MRI, resting-state functional MRI, myelin maps, sulcal depth and task activation maps. In particular, we show how we can use this framework to compare the different modalities, for instance to demonstrate the impact of the gyral bias on dMRI driven parcellation and to quantitatively assess the agreement of parcellations with other modalities.

We compute two different sets of group-wise multi-modal parcellations by 1) combining connectivity modalities (dMRI, rs-fMRI) and myelin (*GraMPa multi-modal parcellation*), and 2) combining myelin, resting state and task fMRI (*GraMPa multi-modal parcellation with task*). We evaluate the impact of such combinations and particularly focus on how rs-fMRI driven parcellations are impacted. In the remainder of this paper, we refer to the joint multi-modal parcellations as *GraMPa multi-modal* and *GraMPa multi-modal with task*. The corresponding modality specific parcellations are referred to as *fMRI/dMRI GraMPa* and *fMRI GraMPa with task*. The parcellations obtained from one single modality are *dMRI/fMRI mono-modal*.

The *GraMPa* parcellations are computed using rs-fMRI, dMRI and myelin maps. This set up was tested in (Parisot et al., 2016b) for single subject parcellation. The intuition behind using these three modalities is based on the fact that functional connectivity is intrinsically linked to structural connectivity, while myelin has been observed to have strong agreements with both connectivity modalities.

GraMPa with task exploits the information provided by rs-fMRI, myelin maps, and all tasks protocols provided by the HCP database. In order to avoid providing redundant information, we only select a subset of the 86 contrasts described in (Barch et al., 2013): the 2BK, 2BK-0BK, BODY-AVG, FACE-AVG, PLACE-AVG and TOOL-AVG contrasts of the WORKING MEMORY protocol; the PUNISH-REWARD contrast of the GAMBLING protocol; the LF-AVG, LH-AVG, RF-AVG, RH-AVG and T-AVG contrasts of the MOTOR protocol, the MATH and MATH-STORY contrasts of the LANGUAGE protocol; the RANDOM and RANDOM-TOM contrasts of the SOCIAL protocol; the MATCH and MATCH-REL contrasts of the RELATIONAL protocol; and the FACES and FACES-SHAPES contrasts of the EMOTION protocol.

Evaluation is performed on the multi-modal *GraMPa* and *GraMPa with task* parcellations (obtained after merging of all modalities) and the corresponding rs-fMRI and dMRI driven parcellations (obtained via the multi-modal iterative process, i.e. rs-fMRI driven parcellation informed by multi-modal data for centre definition). These parcellations are compared to the random initialisation (so as to compare results to chance), and the mono-modal parcellation using solely rs-fMRI data to evaluate the impact of multi-modal information.

For visualisation purposes, results for dMRI and rs-fMRI parcellations are reported on different graphs with corresponding axes. The merging reliability maps as well as the corresponding computed for dMRI, myelin and two example task contrasts are shown in Fig. 4. We can see that the dMRI weights align with cortical folding, with the regions with the lower weights being on the gyri.

3.2. Implementation

The method is implemented using a coarse to fine multi-resolution approach. The intuition behind the use of multiple resolutions is that the lower resolutions are less influenced by noise and allow a more efficient exploration of the search space for parcel centres updates. The cortical mesh is resampled into four different resolutions comprising 4k, 8k, 16k and 32k nodes, where the finest resolution corresponds to the original cortical surface mesh. The data projected on the cortical surface is resampled using the adaptive barycentric approach proposed in the HCP pipelines (Glasser et al., 2013). Quantitative results are computed for the left and right hemispheres separately and then averaged, with the exception of rs-fMRI connectivity network analysis which is performed on the whole cortex.

Parcellations are initialised using random parcellations of same resolution, computed using Poisson Disc Sampling. They are computed for a range of different resolutions: 50, 75, 100, 150, 175 and 200 parcels to investigate the performance of the method at different scales. The MRF smoothing parameter β is heuristically set to 0.3 for connectivity driven and GraMPa parcellations. It is adjusted to 2 for individual task driven parcellations, and 0.1 for myelin based ones. The fusion moves pairwise cost is set to 0.2. The rs-fMRI weight is set to 0.55 for the *GraMPa* parcellation (matching the average values of other modalities), and 0.65 for the *GraMPa* with task parcellation as we only want the most relevant task and myelin features to impact the multi-modal parcellation. Increasing the fMRI weight also allows to reduce the influence of locally noisy weight maps as seen in Fig. 4c. Parameters are optimised visually, so as to obtain compact parcels with smooth boundaries. The value of β has to be adapted with respect to the unary costs' average values, which is why different values are used for non-connectivity and connectivity data. The higher rs-fMRI weight for the *GraMPa with task* is linked to the fact that task and myelin provide very local information, and that they should not overpower the overall role of rs-fMRI in the parcellation scheme. Similarly, the high smoothness parameter (set to 8, as described in section 2.3.2) for non-connectivity reliability weights ensures smoothness of the merged parcellation and reduces the influence of locally noisy task activation and myelin maps.

The MRF optimisation is implemented using the openGM2 library (Andres et al., 2012) which provides a framework for easily testing multiple discrete MRF solvers. Figure 5 shows an example of parcellations obtained with different solvers. We can observe that the differences are minimal between the discrete solvers. While the parcellation remains very similar, we observe stronger differences using the continuous solver we introduced in (Parisot et al., 2015). This can be attributed to the fact that smoothness constraints are defined differently for both systems, and as a result do not provide the exact same parcel smoothness result. In this paper, we use fastPD (Komodakis and Tziritis, 2007) as a solver

due to its computational efficiency. Fusion moves are optimised using QPBO (Rother et al., 2007).

3.3. Parcellation results: Model flexibility

In this section, we illustrate the flexibility of our framework by presenting parcellation results from a set of different modalities. We provide visual results of parcellations and visually assess their accuracy. In addition, we highlight the usefulness of the low SNR cost function for rs-fMRI driven parcellation.

Figure 6 shows comparative results of parcellations using dMRI, and rs-fMRI with both the high SNR and low SNR unary costs. As we can see on the left hand side, the high SNR cost function appears to be adapted to dMRI data, yielding well delineated parcels that only require a limited amount of smoothing. On the other hand, this unary cost yields very noisy results and only identifies resting-state communities in a subset of the cortex. The low SNR unary cost yields a much more detailed parcellation, while we can still observe strong similarities in parcel boundaries with the parcellation obtained from the high SNR cost.

The flexibility and accuracy of our method for using non connectivity data to parcellate the brain is illustrated in Fig. 7. We provide visual examples using myelin maps, sulcal depth and task activation contrasts. We can see that our parcel boundaries tend to align with sharp transitions. We also provide comparative visual results using the non-connectivity unary cost with and without the shortest path constraint. The simple unary cost (without shortest path) yields accurate delineations of the data with very few, but discontinuous parcels (4 or 5 parcels are sufficient for myelin and task activation maps). The shortest path constraint yields compact and contiguous parcels while still following transition boundaries. However, some homogeneous regions can be split in several parcels due to the location of the parcel centres, and more parcels are necessary to accurately parcellate the cortical surface (25 parcels are computed in Fig. 2). As a result, the simpler approach is sufficient (and can be more accurate since there are no parcel splittings) for identifying boundaries and providing a reference for inter-modality comparisons. The shortest path method is more adapted to the GraMPa multi-modal integration, where compact parcels are compared across modalities.

Finally, we illustrate the impact of the gyral bias on dMRI-driven parcellation boundaries in Fig. 8. We provide in Fig. 8a a visual comparison of dMRI-driven parcellation boundaries with rs-fMRI driven ones, where we can see a much stronger agreement between dMRI and cortical folding. Quantitatively, we use a cortical folding based parcellation to compute how myelin, rs-fMRI and dMRI driven edges (100 parcels) align with gyri for all 100 subjects in our database. Each parcel of the cortical folding-driven parcellation is characterised as belonging to a gyrus or not using the average sulcal depth within each parcel. For each parcellation, we then compute the number of vertices on parcel boundaries that pass through a gyrus. As shown in Fig. 8b, this number is much higher for dMRI-driven parcellations than rs-fMRI and myelin.

3.4. Multi-modal comparisons

In this section, we provide quantitative and qualitative evaluation of the agreement of our parcellations with other modalities. We compare our results with 1) myelin maps, 2) task

activation maps, 3) cytoarchitecture, and 4) the recent semi-supervised multi-modal parcellation introduced in (Glasser et al., 2016). All overlap measures are computed using the method described in Sec. 2.4.2. Since myelin and task only provide information on a subset of the brain, we only compute overlap scores with myelin and task parcels that have high or low average values. Thresholds are set to ± 5 for task activation maps and 1.3 for myelin maps (average myelin map value). This avoids to bias the results with uninformative regions and allows to compare solely to highly myelinated and activated regions.

Comparisons between the multi-modal parcellations and myelin or task activation maps are expected to be in favour of our method as these parcellations are driven by myelin and task. Such results can be seen as a validation that the merging step is accurate. It should be noted that the boundaries of the multimodal rs-fMRI parcellations are solely driven by rs-fMRI data. Only the definition of parcel centres at each iteration is driven by multi-modal information. As a result, an improvement in agreement between rs-fMRI parcellations and task or myelin supports the assumption that there is an underlying agreement between these modalities, and that the integration of this additional information yields more accurate and robust centre definitions.

We evaluate the agreement of our parcellations with task activation maps by comparing them to a coarse task-driven parcellation (4 parcels), using the non connectivity unary cost without the shortest path constraint. Results are reported in Fig. 9. We observe a very similar performance from both the multi-modal and fMRI *GraMPa with task* parcellations. Both parcellations outperform the mono-modal approach. On the other hand, the *GraMPa* configuration with dMRI and myelin does not appear to increase the agreement with task activation maps. While the rs-fMRI *GraMPa* parcellation has a similar performance to the mono-modal approach, the *GraMPa* multi-modal parcellation has a lower agreement with task. This can be attributed to the fact that dMRI parcellations have a limited agreement with task, as can be seen in Fig. 9-dMRI group plot. We can see that dMRI parcellations perform similarly to the random initialisation.

We use the same process to compare to myelin maps. As shown in Fig. 10 we observe that both task and *GraMPa* configurations lead to an increase in performance for the multi-modal and rs-fMRI driven parcellations. While this increase is only marginal for task, the *GraMPa* configuration yields an increase in overlap of up to 6%. The good performance of multi-modal rs-fMRI driven parcellations suggests an underlying agreement between rs-fMRI and myelin maps. This increase in performance is also observed for the dMRI-driven parcellation.

We then measure the overlap of our parcels with the cytoarchitectural areas using the group-level cytoarchitectural map computed from our dataset (Brodmann and Garey, 2005). We compute overlap agreements for the primary somatosensory cortex (BA 3, 1, and 2), the primary motor cortex (BA 4), the premotor cortex (BA 6), Broca's area (BA 44, 45), the visual cortex (VI and V2), and the perirhinal cortex as shown in Fig. 11. Results are presented in Fig. 12 for all areas (Fig. 12a and 12c), and focussing on only the motor and visual cortices (Fig. 12b, 12d and 12e). While we observe a limited improvement from the *GraMPa* configuration over all areas, the advantage of using myelin and dMRI data is clear

in the motor and visual areas (i.e. BA 1–4,6 and V1-V2) where we observe a sharp increase for the *GraMPa* parcellation. This increase is also observed more marginally for the individual parcellations. The poor performance of mono-modal *rs-fMRI* and *GraMPa with task* parcellations with respect to the random initialisation might mean that the fMRI modalities we have considered in this paper do not provide sufficient information to reflect the cytoarchitectural organisation. Specialised methods such as visuotopic *rs-fMRI* and retinotopic task fMRI may be necessary to accurately delineate the sensori-motor and early visual cortices.

Last but not least, we compare our parcellations to the HCP's multi-modal parcellation, that makes use of *rs-fMRI*, task fMRI activation map, myelin maps, cortical thickness and topographically organised functional connectivity to generate a parcellation of the cortex comprising 180 parcels. We compute the Adjusted Rand Index (ARI) (Hubert and Arabie, 1985) between parcellations, a measure which has been used several times for parcellation evaluation (Eickhoff et al., 2015) and has the advantage of allowing to compare parcellations of different resolutions (Milligan and Cooper, 1986). Overall, as seen in Fig. 13, we observe a relatively poor agreement between the parcellations, with multi-modal *rs-fMRI* driven parcellations providing the best performance and improving mono-modal results. Interestingly, the *GraMPa with task* parcellation has the worst performance which may be linked to the fact that this parcellation can be noisier than the rest. These results might be attributed to the fact that the HCP's parcellation is generated from a different and larger dataset (210 HCP subjects). Our dataset consists of 100 subjects, and at this scale, our group average parcellations might still be influenced by the variability of the subjects comprising the group. In particular, anatomical differences are likely to be observed and exacerbated by the integration of additional information. In addition, our dataset uses a different registration scheme between cortical surfaces than the data used in Glasser et al. (2016): we use a simpler folding based registration, while theirs is driven by multi-modal information, including *rs-fMRI*, which will provide sharper boundaries. A more accurate comparison of our method with this multi-modal parcellation would have been to generate a group-level parcellation from the single subject ones applied to our data-set, and to use the same registration method. Unfortunately, the data was not available at the time our data was collected and experiments were run.

Comparative visual examples between our parcellations and other modalities are also provided. Comparisons to myelin, cytoarchitecture and task are shown in Fig. 14, 15 and 16 respectively, where parcel boundaries are overlayed to the different maps. The improvement with respect to myelin is particularly striking for the *GraMPa* multi-modal parcellation, similarly to what was observed quantitatively. Despite some relatively poor quantitative results, we observe that *GraMPa with task* can improve the agreement with cytoarchitecture, notably in the motor area. Similarly, the *GraMPa* parcellation also shows some improvement with respect to task activation maps.

3.5. Reproducibility analysis

We compute the agreement between parcellations and all their corresponding modalities for a second set of 100 unrelated subjects. Overlap results for myelin maps, task activation

maps, the HCP multi-modal parcellation and the cytoarchitectural map are reported in Fig. 18. The obtained results are compared to the ones obtained with the first set of 100 subjects in Fig. 9, 10, 13 and 12e.

The most direct comparison can be done with respect to the multi-modal parcellation, as both sets of parcellations are compared to the same atlas. We observe a very similar behaviour, both in terms of ranges as well as relative performance of methods. Comparing the overlaps with the cytoarchitectural maps is also natural, as we expect very few changes between the maps averaged over all subjects. Similarly to what is obtained with respect to the HCP parcellation, performances are very similar between groups. Overlaps with respect to myelin maps are also very similar, with a poorer performance of the fMRI mono-modal parcellation overall with respect to other modalities. Similarly to what is obtained for the first group, the *GraMPa* parcellations substantially outperform the other methods. Task overlaps are almost identical between groups with the exception of the *GraMPa* multi-modal parcellation which yields a performance equivalent to the random initialisation. Overall, we observe very similar performances between groups, which suggests that our method is robust with respect to the group studied.

Finally, we compute the Adjusted Rand Index between the parcellations computed for all groups for all methods. The obtained values are expectedly higher than for comparisons with the HCP's multi-modal parcellation as the pairs of parcellation benefit here from using the same method and assumptions, as well as the same registration strategy (contrarily to the HCP parcellation). We observe a slightly reduced reproducibility across groups when introducing more modalities. This can be linked to the fact that the parcellations are more precisely specific to the group considered, the added complexity of the multi-modal model, or the fact that the boundaries of the mono-modal parcellations are smoother. Additionally, the reproducibility of dMRI parcellations is expectedly reduced as its influence on the final *GraMPa* parcellation is relatively limited compared to fMRI which is less reproducible than dMRI. Interestingly, the *GraMPa*, fMRI *GraMPa* and fMRI *GraMPa* with task parcellations all yield very similar performances.

3.6. Network analysis

Network consistency results comparing the distance between connectivity networks and their corresponding average are shown in Fig. 19 for functional networks and Fig. 20 for structural networks. The similarity between the functional networks computed from different rs-fMRI sessions is shown in Fig. 21. To better visualise the different performances between rs-fMRI methods, we also plot the ratio of the distances computed for a given method and the fMRI mono-modal parcellation. We observe that multi-modal integration generally improves the robustness of connectivity networks. For functional networks, the best results are obtained for the *GraMPa with task* configuration where we consistently obtain lower distances. We obtain statistically significant results ($p < 0.05$) for 50, 75 and 200 parcels for the distance to average, and for 50 and 200 parcels for the distance between rs-fMRI sessions. The improvement is less prominent for the *GraMPa* configuration: we only obtain significant results for 50 parcels and the distance to average. In addition, the rs-fMRI *GraMPa* parcellation yields slightly worse results at higher resolutions. Interestingly, for

both configurations, we obtain the best performance for the multi-modal parcellation with respect to their rs-fMRI driven counterpart. Finally, we obtain consistently better structural networks from the *GraMPa* dMRI parcellation, while the multi-modal one yields worse results.

3.7. Dimensionality reduction quality

We compute the KLD and RMSE for all computed parcellations and modalities. KLD results with respect to the dMRI tractography matrix are reported in Fig. 22. To increase clarity, we only report results related to parcellations that comprised dMRI data. Unsurprisingly, the best results are obtained for the mono-modal dMRI parcellation since the remaining methods are attempting at reducing the influence of certain regions. We still obtain a very similar performance for the dMRI *GraMPa* parcellation which decreases at high resolutions. Interestingly, the multi-modal parcellation yields a poor performance (on par with the initialisation), which may be linked to the reduced influence of dMRI on the joint parcellation.

We report RMSE values for myelin maps, task activation maps and fMRI pseudo-timeseries in Fig. 23a, 23c and 23b respectively. For myelin maps, results for the initial random parcellation are not shown for increased clarity, as they were in a very different range (i.e. approx 0.03 w.r.t to 0.005 for other modalities.).

We observe that the multi-modal parcellation RMSE/KLD are not outperforming the mono-modal RMSE/KLD of the investigated modality, but outperform the mono-modal parcellations RMSE/KLD corresponding to different modalities. This suggests that integrating multiple modalities does not alter the precision of the method, and yields, for each modality involved, parcellations that are as accurate as the corresponding mono-modal ones.

4. Discussion

In this paper, we proposed a flexible graphical model for automatic parcellation of the cortex using multiple sources of information. The proposed method can parcellate the brain using connectivity (rs-fMRI, dMRI) as well as non connectivity data projected on the cortical surface (e.g. myelin maps). We provide a multi-modal extension to this model, where our iterative approach computes modality-specific coherent parcellations as well as a multi-modal parcellation that merges modalities based on their reliabilities.

We compute two sets of multi-modal parcellations, by merging connectivity information (rs-fMRI and dMRI) and merging rs-fMRI with task fMRI. We evaluate the impact of integrating multiple modalities in the parcellation framework and more specifically compare the results obtained from the two different configurations. Our results show that multi-modal parcellations yield more accurate brain delineations with respect to well established atlases. We also observe an increase in performance for the rs-fMRI and dMRI parcellations when they are informed by multiple modalities. Last but not least, our method provides a framework for quantitatively comparing parcellations to other modalities, using coarse parcellations as a reference.

An important observation is that we obtain different performances depending on the modalities considered in the multi-modal framework. On one hand, the task based configuration appears to yield a better representation of the functional organisation, since both the agreement with task activation maps and the robustness of constructed connectivity networks increase. On the other hand, merging rs-fMRI, dMRI and myelin maps yields parcellations that provide a better representation of microstructure (i.e. a better agreement with cytoarchitecture and myelination). From a dimensionality reduction point of view, we observe that the best performances with respect to all modalities are obtained from the fMRI parcellations derived from the multi-modal scheme. This suggests that there is no unique parcellation or parcellation method that is optimally adapted to all tasks (e.g. cortical delineations and network construction). A similar observation was made in Arslan et al. (2017), where certain parcellation methods provided more accurate cortical delineations while other performed better with respect to dimensionality reduction tasks. One may need to adapt the parcellation scheme to the task at hand. For example cortical delineation requires precise boundaries and is likely to require manual inspection by experts. Several cortical areas (e.g. V1/V2) have already been well defined, and the work of Glasser et al. (2016) is an important step towards identifying a ground truth parcellation. Fully automated methods are more adapted to the dimensionality reduction problem. While our method shows promising results towards identifying cortical areas, its flexibility and multi-modal aspect makes it well adapted to data-driven parcellations for dimensionality reduction, for example to build a common parcellation suitable for multiple modalities.

While it might be interesting to consider merging all modalities together, the overall performance is likely to decrease as we would attempt to integrate too much information that can be contradictory and therefore cannot be appropriately resolved in a single parcellation. This effect can be observed for myelin maps. Despite the fact that they are integrated in both multi-modal configurations, we obtain a much better agreement using rs-fMRI and dMRI than using multiple task contrasts. The influence of myelin is considerably reduced in the task based multi-modal parcellation. Another possible interpretation is that we are currently using Z-statistics task activation maps. A better performance might be obtained using effect size maps, which are expected to reflect the biological boundaries within a population more accurately. Similarly, the data used in this paper relies on a registration based on cortical folding. This may impact the quality of the parcellations (e.g. misalignments between modalities, blurring in areas in which areal boundaries and folds are not well correlated). More robust results might be obtained using an areal feature-based alignment.

Regarding our network analysis experiments, it is interesting to see that we generally observe the best performance from the multi-modal parcellation rather than the purely rs-fMRI driven one. One can infer from this that some functional transitions that can be identified from other modalities cannot be identified robustly from rs-fMRI mono-modal data. Our results also show that the influence of multi-modal information is reduced compared to what is observed for the task of brain mapping. This is to be expected as the boundaries of the parcels have a smaller influence for this kind of task. The impact of parcellation on network analysis has the potential to be more important in the context of diseased brains, especially in the presence of lesions which can considerably affect the brain's organisation.

One particular advantage of the method is its flexibility with respect to the input data and design of the merging step. This flexibility can be particularly relevant when working on pathological data, where the quality fMRI or dMRI processing could be affected. In such settings, exploiting multiple sources of information has the potential to provide a clearer interpretation of how a disease affects the brain. Furthermore, flexibility is essential for clinical data where some modalities may be missing from the dataset.

It is important to keep in mind that comparisons between different modalities have to be considered carefully. All modalities considered utilise an ensemble of different processing steps in which multiple errors and wrong assumptions could have been made. The cytoarchitectural map is a particularly interesting case. It has been constructed from a limited set of subjects (Fischl et al., 2008) and then mapped to our dataset through a series of projections and registrations. More specifically, we focus our evaluation on the motor and visual cortices due to potential errors in registration. The cytoarchitectural atlas is mapped onto each subject using cortical folding based registration, which is optimised for the motor and visual cortices where folding is more consistent across subjects. As a result, areas located outside these regions have the potential to be misaligned. The high performance of the random parcellation with respect to data-driven parcellations suggests that this is the case, or that connectivity locally disagrees with cytoarchitecture. Another point is the use of overlap based measures, which can be affected by the size and shape of the parcels. While this bias is reduced due to the fact that we are using a common framework to carry out comparisons, it could still affect the performance of noisier parcellations such as the multi-modal ones.

Several extensions and improvements to the model could be considered. One option could be to adapt the model for group-wise parcellation, where each subject would correspond to a modality and the merged parcellation would be the group-level representation. We could obtain a set of coherent single-subject parcellations while preserving inter-subject variability. Here, the merging step could be decided based on group consistency measurements rather than reliability weights. A similar extension could be carried out for longitudinal data, which could be particularly interesting for brain development studies. We propose a fully automated model, which we believe is an important option to have considering neuroanatomists are not necessarily available for such tasks. Nonetheless, the parcellation quality could be largely increased by allowing user feedback, both in terms of reliability weights and parcel boundaries. Our model flexibility also allows an easy extension to a semi-automated setting, allowing neuroanatomists to provide feedback during our iterative and weight computation processes.

The reliability weights that we have proposed in this paper are relatively simple and aim at demonstrating the potential of the method. More accurate merging criteria could be learned from the data, or designed from prior knowledge on the agreement between modalities. In particular, the weights could be optimised for a specific task, for example with the objective of maximising the parcels' overlap with known areas or other modalities. This can be optimised using the quantitative measures of overlap provided within our framework. Another option would be to compute data-driven reliability weights by computing parcellation uncertainties using the MRF's min-marginal energies obtained from each

optimisation (Kohli and Torr, 2006). This could be particularly relevant for the construction of a more robust rs-fMRI parcellation, which could be done using our framework with multiple rs-fMRI acquisitions. In such setting, a consistency measurement also has the potential to provide more accurate parcellations.

Finally, the MRF model itself could benefit from integrating the pairwise similarities between nodes within the smoothing pairwise term. This could provide more information and help define the parcel boundaries more precisely. Another interesting extension would be to investigate the impact of using fully connected CRFs on the obtained parcellations (Krähenbühl and Koltun, 2011).

Acknowledgements

The research leading to these results has received funding from NIH grant P41EB015902 and the European Research Council under the European Union's Seventh Framework Programme (FP/2007–2013) / ERC Grant Agreement no. 319456. Data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

References

- Amunts K, Schleicher A, Zilles K, 2007 Cytoarchitecture of the cerebral cortex more than localization. *Neuroimage* 37, 1061–1065. [PubMed: 17870622]
- Andres B, Beier T, Kappes J, 2012 OpenGM: A C++ Library for Discrete Graphical Models. *CoRR* abs/1206.0111.
- Arslan S, Ktena SI, Makropoulos A, Robinson EC, Rueckert D, Parisot S, 2017 Human brain mapping: A systematic comparison of parcellation methods for the human cerebral cortex. *Neuroimage*, –URL: http://www.sciencedirect.com/science/article/pii/S1053811917303026, doi:10.1016/j.neuroimage.2017.04.014.
- Arslan S, Parisot S, Rueckert D, 2015 Joint Spectral Decomposition for the Parcellation of the Human Cerebral Cortex Using Resting-State fMRI, in: *IPMI*, pp. 85–97.
- Baldassano C, Beck DM, Fei-Fei L, 2015 Parcellating Connectivity in Spatial Maps. *PeerJ* 3, e784. doi:10.7717/peerj.784. [PubMed: 25737822]
- Barch DM, Burgess GC, Harms MP, Petersen SE, Schlaggar BL, Corbetta M, Glasser MF, Curtiss S, Dixit S, Feldt C, Nolan D, Bryant E, Hartley T, Footer O, Bjork JM, Poldrack R, Smith S, Johansen-Berg H, Snyder AZ, Van Essen DC, 2013 Function in the human connectome: Task-fMRI and individual differences in behavior. *Neuroimage* 80, 169–189. doi:10.1016/j.neuroimage.2013.05.033. [PubMed: 23684877]
- Behrens T, Berg HJ, Jbabdi S, Rushworth M, Woolrich M, 2007 Probabilistic Diffusion Tractography with Multiple Fibre Orientations: What Can We Gain? *Neuroimage* 34, 144–155. doi:10.1016/j.neuroimage.2006.09.018. [PubMed: 17070705]
- Blumensath T, Jbabdi S, Glasser MF, Van Essen DC, Ugurbil K, Behrens TE, Smith SM, 2013 Spatially Constrained Hierarchical Parcellation of the Brain with Resting-State fMRI. *Neuroimage* 76, 313–324. doi:10.1016/j.neuroimage.2013.03.024. [PubMed: 23523803]
- Bohland JW, Bokil H, Allen CB, Mitra PP, 2009 The Brain Atlas Concordance Problem: Quantitative Comparison of Anatomical Parcellations. *PLoS ONE* 4, e7200. doi:10.1371/journal.pone.0007200. [PubMed: 19787067]
- Boykov Y, Funka-Lea G, 2006 Graph Cuts and Efficient N-D Image Segmentation. *International Journal of Computer Vision* 70, 109–131. doi:10.1007/s11263-006-7934-5.
- Brodmann K, Garey LJ, 2005 *Brodmann's Localisation in the Cerebral Cortex*. Springer US, Boston, MA.

- Cohen AL, Fair DA, Dosenbach NU, Miezin FM, Dierker D, Van Essen DC, Schlaggar BL, Petersen SE, 2008 Defining Functional Areas in Individual Human Brains Using Resting Functional Connectivity MRI. *Neuroimage* 41, 45–57. doi:10.1016/j.neuroimage.2008.01.066. [PubMed: 18367410]
- Craddock RC, James G, Holtzheimer PE, Hu XP, Mayberg HS, 2012 A Whole Brain fMRI Atlas Generated via Spatially Constrained Spectral Clustering. *Human Brain Mapping* 33, 1914–1928. doi:10.1002/hbm.21333. [PubMed: 21769991]
- Destrieux C, Fischl B, Dale A, Halgren E, 2010 Automatic Parcellation of Human Cortical Gyri and Sulci Using Standard Anatomical Nomenclature. *Neuroimage* 53, 1–15. doi:10.1016/j.neuroimage.2010.06.010. [PubMed: 20547229]
- Eickhoff SB, Thirion B, Varoquaux G, Bzdok D, 2015 Connectivity-based parcellation: Critique and implications: Connectivity-Based Parcellation. *Human Brain Mapping* 36, 4771–4792. doi:10.1002/hbm.22933. [PubMed: 26409749]
- Fischl B, Rajendran N, Busa E, Augustinack J, Hinds O, Yeo BTT, Mohlberg H, Amunts K, Zilles K, 2008 Cortical Folding Patterns and Predicting Cytoarchitecture. *Cerebral Cortex* 18, 1973–1980. doi:10.1093/cercor/bhm225. [PubMed: 18079129]
- Geman S, Geman D, 1984 Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-6*, 721–741. doi:10.1109/TPAMI.1984.4767596.
- Glasser MF, Coalson TS, Robinson EC, Hacker CD, Harwell J, Yacoub E, Ugurbil K, Andersson J, Beckmann CF, Jenkinson M, Smith SM, Van Essen DC, 2016 A Multi-Modal Parcellation of Human Cerebral Cortex. *Nature* 536, 171–178. doi:10.1038/nature18933. [PubMed: 27437579]
- Glasser MF, Sotiropoulos SN, Wilson JA, Coalson TS, Fischl B, Andersson JL, Xu J, Jbabdi S, Webster M, Polimeni JR, Van Essen DC, Jenkinson M, 2013 The Minimal Preprocessing Pipelines for the Human Connectome Project. *Neuroimage* 80, 105–124. doi:10.1016/j.neuroimage.2013.04.127. [PubMed: 23668970]
- Glocker B, Komodakis N, Tziritas G, Navab N, Paragios N, 2008 Dense Image Registration through MRFs and Efficient Linear Programming^{*}. *Medical Image Analysis* 12, 731–741. doi:10.1016/j.media.2008.03.006. [PubMed: 18482858]
- Gordon EM, Laumann TO, Adeyemo B, Huckins JF, Kelley WM, Petersen SE, 2016 Generation and Evaluation of a Cortical Area Parcellation from Resting-State Correlations. *Cerebral Cortex* 26, 288–303. doi:10.1093/cercor/bhu239. [PubMed: 25316338]
- Hagmann P, Cammoun L, Gigandet X, Meuli R, Honey CJ, Wedeen VJ, Sporns O, 2008 Mapping the Structural Core of Human Cerebral Cortex. *PLoS Biology* 6, e159. doi:10.1371/journal.pbio.0060159.
- Honnorat N, Eavani H, Satterthwaite T, Gur R, Gur R, Davatzikos C, 2015 GraSP: Geodesic Graph-Based Segmentation with Shape Priors for the Functional Parcellation of the Cortex. *Neuroimage* 106, 207–221. doi:10.1016/j.neuroimage.2014.11.008. [PubMed: 25462796]
- Hubert L, Arabie P, 1985 Comparing Partitions. *Journal of Classification* 2, 193–218. doi:10.1007/BF01908075.
- Jbabdi S, Sotiropoulos SN, Savio AM, Grana M, Behrens TEJ, 2012 Model-Based Analysis of Multishell Diffusion MR Data for Tractography: How to Get over Fitting Problems. *Magnetic Resonance in Medicine* 68, 1846–1855. doi:10.1002/mrm.24204. [PubMed: 22334356]
- Jbabdi S, Woolrich M, Behrens T, 2009 Multiple-Subjects Connectivity-Based Parcellation Using Hierarchical Dirichlet Process Mixture Models. *Neuroimage* 44, 373–384. doi:10.1016/j.neuroimage.2008.08.044. [PubMed: 18845262]
- Kohli P, Torr PHS, 2006 Measuring Uncertainty in Graph Cut Solutions – Efficiently Computing Min-Marginal Energies Using Dynamic Graph Cuts, in: Hutchison D, Kanade T, Kittler J, Kleinberg JM, Mattem F, Mitchell JC, Naor M, Nierstrasz O, Pandu Rangan C, Steffen B, Sudan M, Terzopoulos D, Tygar D, Vardi MY, Weikum G, Leonardis A, Bischof H, Pinz A (Eds.), *Computer Vision – ECCV 2006*. Springer Berlin Heidelberg, Berlin, Heidelberg, volume 3952, pp. 30–43.
- Komodakis N, Tziritas G, 2007 Approximate Labeling via Graph Cuts Based on Linear Programming. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29, 1436–1453. doi:10.1109/TPAMI.2007.1061. [PubMed: 17568146]

- Krähenbühl P, Koltun V, 2011 Efficient Inference in Fully Connected CRFs with Gaussian Edge Potentials, in: Shawe-Taylor J, Zemel RS, Bartlett PL, Pereira F, Weinberger KQ (Eds.), *Advances in Neural Information Processing Systems 24*. Curran Associates, Inc., pp. 109–117.
- Lashkari D, Vul E, Kanwisher N, Golland P, 2010 Discovering Structure in the Space of fMRI Selectivity Profiles. *Neuroimage* 50, 1085–1098. doi:10.1016/j.neuroimage.2009.12.106. [PubMed: 20053382]
- Lempitsky V, Rother C, Roth S, Blake A, 2010 Fusion Moves for Markov Random Field Optimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32, 1392–1405. doi:10.1109/TPAMI.2009.143. [PubMed: 20558873]
- Maier-Hein K, Neher P, Houde JC, Cote MA, Garyfallidis E, Zhong J, Chamberland M, Yeh FC, Lin YC, Ji Q, Reddick WE, Glass JO, Chen DQ, Feng Y, Gao C, Wu Y, Ma J, Renjie H, Li Q, Westin CF, Deslauriers-Gauthier S, Gonzalez JOO, Paquette M, St-Jean S, Girard G, Rheault F, Sidhu J, Tax CMW, Guo F, Mesri HY, David S, Froeling M, Heemskerk AM, Leemans A, Bore A, Pinsard B, Bedetti C, Desrosiers M, Brambati S, Doyon J, Sarica A, Vasta R, Cerasa A, Quattrone A, Yeatman J, Khan AR, Hodges W, Alexander S, Romascano D, Barakovic M, Auria A, Esteban O, Lemkaddem A, Thiran JP, Cetingul HE, Odry BL, Mailhe B, Nadar M, Pizzagalli F, Prasad G, Villalon-Reina J, Galvis J, Thompson P, Requejo F, Laguna P, Lacerda L, Barrett R, Dell'Acqua F, Catani M, Petit L, Caruyer E, Daducci A, Dyrby T, Holland-Letz T, Hilgetag C, Stieltjes B, Descoteaux M, 2016 Tractography-Based Connectomes Are Dominated by False-Positive Connections. Technical Report.
- Milligan GW, Cooper MC, 1986 A Study of the Comparability of External Criteria for Hierarchical Cluster Analysis. *Multivariate Behavioral Research* 21, 441–458. doi:10.1207/sl5327906mbr21045. [PubMed: 26828221]
- Moreno-Dominguez D, Anwender A, Knösche TR, 2014 A hierarchical method for whole-brain connectivity-based parcellation: Whole-Brain Connectivity-Based Parcellation. *Human Brain Mapping* 35, 5000–5025. doi:10.1002/hbm.22528. [PubMed: 24740833]
- Ng B, Varoquaux G, Poline JB, Thirion B, 2012 A Novel Sparse Graphical Approach for Multimodal Brain Connectivity Inference. *Medical image computing and computer-assisted intervention: MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention* 15, 707–714.
- Parisot S, Arslan S, Passerat-Palmbach J, Wells WM, Rueckert D, 2016a Group-wise parcellation of the cortex through multi-scale spectral clustering. *Neuroimage* 136, 68–83. doi:10.1016/j.neuroimage.2016.05.035. [PubMed: 27192437]
- Parisot S, Glocker B, Schirner MD, Rueckert D, 2016b GraMPa: Graph-Based Multi-modal Parcellation of the Cortex Using Fusion Moves, in: Ourselin S, Joskowicz L, Sabuncu MR, Unal G, Wells W (Eds.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*. Springer International Publishing, Cham, volume 9900, pp. 148–156.
- Parisot S, Rajchl M, Passerat-Palmbach J, Rueckert D, 2015 A Continuous Flow-Maximisation Approach to Connectivity-Driven Cortical Parcellation, in: Navab N, Hornegger J, Wells WM, Frangi AF (Eds.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Springer International Publishing, Cham, volume 9351, pp. 165–172.
- Passingham RE, Stephan KE, Kötter R, 2002 The Anatomical Basis of Functional Localization in the Cortex. *Nature Reviews Neuroscience* 3, 606–616. doi:10.1038/nrn893. [PubMed: 12154362]
- Raj A, Mueller S, Young K, Laxer K, Weiner M, 2010 Network-Level Analysis of Cortical Thickness of the Epileptic Brain. *Neuroimage* 52, 1302–1313. doi:10.1016/j.neuroimage.2010.05.045. [PubMed: 20553893]
- Robinson EC, Jbabdi S, Glasser MF, Andersson J, Burgess GC, Harms MP, Smith SM, Van Essen DC, Jenkinson M, 2014 MSM: A new flexible framework for Multimodal Surface Matching. *Neuroimage* 100, 414–426. doi:10.1016/j.neuroimage.2014.05.069. [PubMed: 24939340]
- Rother C, Kolmogorov V, Lempitsky V, Szummer M, 2007 Optimizing Binary MRFs via Extended Roof Duality, *IEEE*, pp. 1–8. doi:10.1109/CVPR.2007.383203.
- Ryali S, Chen T, Supekar K, Menon V, 2013 A parcellation scheme based on von Mises-Fisher distributions and Markov random fields for segmenting brain regions using resting-state fMRI. *Neuroimage* 65, 83–96. doi:10.1016/j.neuroimage.2012.09.067. [PubMed: 23041530]

- Salimi-Khorshidi G, Douaud G, Beckmann CF, Glasser MF, Griffanti L, Smith SM, 2014 Automatic denoising of functional mri data: combining independent component analysis and hierarchical fusion of classifiers. *Neuroimage* 90, 449–468. [PubMed: 24389422]
- Salvador R, 2004 Neurophysiological Architecture of Functional Magnetic Resonance Images of Human Brain. *Cerebral Cortex* 15, 1332–1342. doi:10.1093/cercor/bhi016.
- Shen X, Tokoglu F, Papademetris X, Constable R, 2013 Groupwise Whole-Brain Parcellation from Resting-State fMRI Data for Network Node Identification. *Neuroimage* 82, 403–415. doi:10.1016/j.neuroimage.2013.05.081. [PubMed: 23747961]
- Smith SM, Hyvärinen A, Varoquaux G, Miller KL, Beckmann CF, 2014 Group-PCA for very large fMRI datasets. *Neuroimage* 101, 738–749. doi:10.1016/j.neuroimage.2014.07.051. [PubMed: 25094018]
- Sotiropoulos SN, Jbabdi S, Xu J, Andersson JL, Moeller S, Auerbach EJ, Glasser MF, Hernandez M, Sapiro G, Jenkinson M, Feinberg DA, Yacoub E, Lenglet C, Van Essen DC, Ugurbil K, Behrens TE, 2013 Advances in diffusion MRI acquisition and processing in the Human Connectome Project. *Neuroimage* 80, 125–143. doi:10.1016/j.neuroimage.2013.05.057. [PubMed: 23702418]
- Sporns O, 2011 The Human Connectome: A Complex Network. *Annals of the New York Academy of Sciences* 1224, 109–125. doi:10.1111/j.1749-6632.2010.05888.x. [PubMed: 21251014]
- Thomas Yeo BT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, Roffman JL, Smoller JW, Zolkei L, Polimeni JR, Fischl B, Liu H, Buckner RL, 2011 The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology* 106, 1125–1165. doi:10.1152/jn.00338.2011. [PubMed: 21653723]
- Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M, 2002 Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain. *Neuroimage* 15, 273–289. doi:10.1006/nimg.2001.0978. [PubMed: 11771995]
- Van Essen DC, Jbabdi S, Sotiropoulos SN, Chen C, Dikranian K, Coalson T, Harwell J, Behrens TE, Glasser MF, 2014 Mapping Connections in Humans and Non-Human Primates, in: *Diffusion MRI*. Elsevier, pp. 337–358.
- Venkataraman A, Rathi Y, Kubicki M, Westin C, Golland P, 2012 Joint Modeling of Anatomical and Functional Connectivity for Population Studies. *IEEE Transactions on Medical Imaging* 31, 164–182. doi:10.1109/TMI.2011.2166083. [PubMed: 21878411]
- Zilles K, Amunts K, 2010 Centenary of Brodmann's Map — Conception and Fate. *Nature Reviews Neuroscience* 11, 139–145. doi:10.1038/nrn2776.

Figure 1:

Overview of the GraMPa method. Each iteration updates an initial parcellation into a set of modality specific parcellations using an MRF model. The parcellations are then merged based on the modalities' relative influences into a multi-modal parcellation, which initialises the next iteration.

Figure 2:

Illustration of the modality specific unary costs for high (a) and low (a) SNR connectivity data and (c) non connectivity data.

Figure 3:

Illustration of the process to construct a $K \times K$ low dimensionality representation of dMRI data, followed by a high resolution approximate matrix.

Figure 4:

Merging reliability maps and corresponding maps obtained from dMRI, myelin and task activation maps. Higher values correspond to giving a higher influence to the modality.

Figure 5:

Visual comparison of the impact of optimisation methods on the parcellation. The red arrows highlight the main differences between the continuous optimisation and discrete ones.

Figure 6:

Parcellation results using the low and high SNR unary costs for connectivity data (dMRI and rs-fMRI) for 50 (top row) and 150 parcels (bottom row). (a,c) shows parcel assignments obtained by simply minimising the unary cost over all potential labels without MRF smoothing constraint. (b,d) are the corresponding parcellations obtained after optimisation. Each colour corresponds to a parcel. The middle column shows how the high SNR unary cost fails for rs-fMRI data.

Figure 7: Visual examples of parcellations on non-connectivity data. For the different considered maps, we compare the parcellations obtained with (right) and without the shortest path constraint (left).

Figure 8:

Illustration of the agreement between dMRI parcel boundaries and cortical folding. **(a)** Visual assessment of the alignment of dMRI parcel boundaries with cortical folding. The boundaries of the parcellations are superimposed to a sulcal depth map, where lighter colours correspond to gyri. This is shown alongside a rs-fMRI driven parcellation for comparative purposes. **(b)** Quantitative assessment of the alignment of dMRI parcel boundaries with cortical folding.

Figure 9: Overlap between task activation maps and our parcellations. dMRI based parcellations are reported with dashed lines.

Figure 10: Overlap between myelin maps and our parcellations. dMRI based parcellations are reported with dashed lines.

Figure 11:
Illustration of the cytoarchitectural areas that are being matched with our parcellations.

Figure 12:
Overlap between cytoarchitectural areas and our parcellations.

Figure 13: Overlap between the HCP's multi-modal parcellation and our parcellations. dMRI based parcellations are reported with dashed lines.

Figure 14:
Comparison between the average myelin maps and the mono-modal, GraMPa and GraMPa with task *rs-fMRI*, *dMRI*, myelin driven and multi-modal parcellations (100 parcels). Parcel boundaries are superimposed to the average myelin map (red: highly myelinated regions.)

Figure 15:
Comparison between the average cytoarchitectural map and the mono-modal, GraMPa and GraMPa with task *rs-fMRI*, *dMRI*, myelin driven and multi-modal parcellations (75 parcel).

Figure 16:
Comparison between the average task activation maps and the mono-modal, GraMPa and GraMPa with task *rs-fMRI* and multi-modal parcellations.

(a) Agreement with myelin maps.

(b) Agreement with task activation maps.

(c) Agreement with the HCP multi-modal parcellation.

(d) Agreement with cytoarchitectural motor and visual areas.

Figure 17:

Reproducibility analysis: overlap between our parcellations and the different modalities for a second group. *dMRI* based parcellations are reported with dashed lines.

Figure 18:

Reproducibility analysis: Adjusted Rand Index between the parcellations obtained from two different groups of 100 subjects, (a) Edge distance between networks and their average for all subjects (b) Ratio between the mono-modal distance and the other methods' distances

Figure 19: similarity between the functional networks and their corresponding average networks.

Figure 20:
Similarity between the structural networks and their corresponding average networks.

Figure 21:

Similarity between two functional networks obtained from two different *rs-fMRI* sessions.

(a) Edge diBtance between the two sessions' net-work for all Subjects, (b) Ratio between the mono-modal distance and the other methods' distances

Figure 22:
Kullback Liebler Divergence between the original tractography matrix and the merged one for dMRI driven parcellations

(a) *RMSE between the original myelin maps and approximated one using the parcellations.*

(b) *RMSE between the original fMRI timeseries and approximation using the parcellations.*

(c) *RMSE between the original task activation maps and approximation using the parcellations.*

Figure 23:

Impact of parcellation based dimensionality reduction: error and loss of reconstruction with respect to the different modalities.