

PARCELLATION OF THE HUMAN CEREBRAL CORTEX USING DIFFUSION MRI

THARINDU GANEPOLA

A thesis submitted in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy

of the

University College London

2017

Primary Supervisor

Martin I. Sereno

(Professor of Neuorimaging, Department of Psychology, UCSD)

Secondary Supervisor

Daniel C. Alexander

(Professor of Imaging Science, Department of Computer Science, UCL)



DECLARATION

I, Tharindu Ganepola, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

October, 2017

Tharindu Ganepola

ACKNOWLEDGMENTS

It would not have been possible to complete this PhD without the help and support of countless others for whom my gratitude is expressed below.

I would like to thank my supervisors, Professor Martin Sereno and Professor Daniel Alexander, for their excellent supervision over the last four years. Most of this research would not have progressed without Marty's seemingly endless expertise or his midnight additions of csurf button functionalities. Nor would it have been as enjoyable without the random and varied anecdotes regarding life in the monkey lab etc. Similarly, Danny has given constant guidance on machine learning, diffusion analysis, and support on the more general issues associated navigating the academia, providing a fun, open, and supportive research environment!

There are many colleagues, whom I would also like to thank. Zoltan Nagy has maintained an active interest in the project over the years, providing ideas and enduring many hours in the scanner. I am also grateful to everyone in the microstructure imaging group. It has been rewarding to work amongst so many talented individuals, and I have surely absorbed much of my knowledge through osmosis during our weekly meetings. However, more important than knowledge is friendship, so I also thank MIG for the countless social excursions that we have enjoyed together. I hold especially fond memories of the conferences, summer schools, away-days and Christmas parties. These sentiments extend to my fellow basement dwellers at Bedford Way, who provided excellent company during those long winter months when UCL decided that we didn't need central heating.

Fortunately, there has been plenty of time for fun during the PhD and I must thank my network of friends for this. I thank my friends from home - for allowing me to escape every now and then, the friends from my undergrad - for listening to me whinge and continually inspiring me, and all the new friends I have made along the way. Honourable mention should go to my flatmates and closest friends over the last few years, many of whom are fellow PhD students. It would have been much harder to ride the ups and downs of this journey their

fellowship, and the meals, trips, movie nights, parties, friend dates, and hugs that we have shared.

A special thanks goes to Corneliu, who deserves a medal for his unfaltering support throughout this whole process. I could not have asked for a more patient and loving partner in crime with whom to share this phase of life.

Last, and by no means least, I thank my family. My extended family in Sri Lanka have continually championed and motivated me. My brother, Ravi, has always led by example, letting me tag along on his "mad scientist" pursuits, like exploding toy cars in the back garden. I am very fortunate to have such a conscientious and kind sibling. My parents have never failed to encourage and support me in all of my endeavours. They have gone above and beyond to provide for me, and nurtured my curiosity from an early age. Completing this PhD would not have been possible without the unconditional love of these people.

PUBLICATIONS

Some ideas and figures have appeared previously in the following publications:

- Ganepola, T. et al. (2018). "Using diffusion MRI to discriminate areas of cortical grey matter."
In: *Neuroimage*, (In Press).
- Ganepola, T. et al. (2016a). "An unsupervised group average cortical parcellation using diffusion MRI to probe cytoarchitecture." In: Computational Diffusion MRI: MICCAI Workshop, Athens, Greece, October 2016, Springer, pp. 145-156.
- Ganepola, T. et al. (2016b). "Evaluation of diffusion MRI based feature sets for the classification of primary motor and somatosensory cortical areas." In: Proceedings of the 24th Annual Meeting of International Society for Magnetic Resonance Imaging, Singapore, 2016.
- Ganepola, T. et al. (2015). "An unsupervised group average cortical parcellation using HARDI data." In: Proceedings of 21st Annual Meeting of the Organization of Human Brain Mapping, Honolulu, 2015, p. 221.
- Nagy, Z. et al. (2014). "Combining HARDI datasets with more than one b-value improves diffusion MRI-based cortical parcellation." In: Proceedings of the 22nd Annual Meeting of International Society for Magnetic Resonance Imaging, Milan, 2014, p. 800.
- Lee, Y. et al. (2017). "Investigating the effects of concurrent magnetic field monitoring on high angular resolution diffusion imaging: application to cortical parcellation." In: Proceedings of the 25th Annual Meeting of International Society for Magnetic Resonance Imaging, Honolulu, 2017.

ABSTRACT

Histological methods have long been used to segment the cerebral cortex into structurally distinct cortical areas that have served as a basis for research into brain structure and function and remain in use today. There is great interest in adapting and extending these methods to be able to use non-invasive imaging, so that tighter structure-function relationships can be measured in living subjects. Whilst diffusion neuroimaging methods have been widely applied to white matter, the reduced anisotropy in the thin, complexly folded grey matter of the cortex has so far limited its study. *In vivo* parcellation pipelines have instead focussed on T1 and T2 weighted MRI. Recent advances in imaging hardware have reignited interest in grey matter diffusion MRI as a viable candidate for characterising architectonic domains. This Thesis explores the capabilities of dMRI as a measure of cortical microstructure using *in vivo* datasets from healthy adult participants. A cortical parcellation pipeline was developed in which both unsupervised and supervised algorithms were explored. Results were presented at both the group level and single subject level across the entire cortical sheet. The diffusion-based feature space characterised the known variation in cellular composition and fibre density relative to the local cortical surface normal. Thus they remain invariant to the confounding orientation changes associated with cortical folding, which usually inhibit studies of cortical microstructure. The features were compared to the alternative T1w/T2w myelin mapping methods to demonstrate that the diffusion MRI signal provides a complementary mode of contrast. A series of classification experiments were used to determine the most effective methods for utilising diffusion in grey matter applications. Several additional methods from the dMRI literature were compared to highlight the benefit of higher-order tissue representations. Similarly, classification tasks were used to corroborate the benefits of sampling multiple b-values in cortical studies. The experimental chapters provide strong evidence in favour of the future use of diffusion MRI as a measure of the varying microstructure that defines cortical areas.

CONTENTS

I	INTRODUCTION	1
1	MOTIVATION	2
1.1	Problem Statement	5
1.2	Project Aims	5
1.3	Contributions	6
II	BACKGROUND	8
2	THE CORTEX AND PARCELLATION	9
2.1	Histological studies of the cortex	10
2.1.1	Classical cortical parcellation	12
2.1.2	Probabilistic atlases	14
2.2	MRI studies of the cortex	14
2.2.1	Basic principals of MRI	15
2.2.2	Surface based analysis	16
2.2.3	Image based myeloarchitectonics	19
3	DIFFUSION MRI	23
3.1	Principals of diffusion MRI	23
3.1.1	PGSE sequence	24
3.2	Estimating the ADC	26
3.2.1	The Diffusion Tensor	26
3.2.2	Beyond the Diffusion Tensor	27
3.3	Diffusion based cortical imaging	28
3.4	HARDI model for discriminating cortical areas	31
4	MACHINE LEARNING	35
4.1	K-means	36
4.2	Random forests	38

III	EXPERIMENTS	43
5	UNSUPERVISED GROUP AVERAGE PARCELLATION	44
5.1	Methods	45
5.1.1	Human Connectome Project data	45
5.1.2	Surface reconstruction	46
5.1.3	Feature space	47
5.1.4	Classification	48
5.2	Results	50
5.2.1	Central sulcus	50
5.2.2	Broca's region	51
5.2.3	Auditory areas	52
5.2.4	Occipital areas	52
5.2.5	Gyrification	53
5.3	Discussion	54
6	FEATURE SETS AND SUPERVISED PARCELLATION	60
6.1	Methods	61
6.1.1	Data and pre-processing:	61
6.1.2	Feature sets	61
6.1.3	Classification Experiments	63
6.1.4	Searchlight Cluster Count	66
6.2	Results	67
6.2.1	Binary classification of M1 vs S1	67
6.2.2	Group average whole hemisphere classifier	68
6.2.3	Single subject whole hemisphere classifier	73
6.3	Discussion	74
7	MULTIPLE B-VALUES FOR CORTICAL IMAGING	83
7.1	Methods	84
7.1.1	Data and processing	84
7.1.2	Correlation analysis	85
7.1.3	Classification	85
7.2	Results	88

7.2.1 Correlation maps	88
7.2.2 Classification	89
7.3 Discussion	95
8 CONCLUSION	99
8.1 Future Work	101
IV APPENDIX	104
A EVALUATION OF CONCURRENT FIELD MONITOR	105
B CHAPTER 5 SUPPLEMENTARY MATERIAL	112
C CHAPTER 6 SUPPLEMENTARY MATERIAL	117
BIBLIOGRAPHY	119

LIST OF FIGURES

Figure 2.1	The cyto- and myeloarchitectonic layers of the human cerebral cortex	10
Figure 2.2	Cell-body and myelin stains for the sensorimotor areas of the central sulcus and the visual areas of the calcarine.	11
Figure 2.3	The classical cytoarchitectonic maps of Brodmann (1909) and Economo and Koskinas (1925).	13
Figure 2.4	Example T1w and T2w image slices from the same subject.	16
Figure 2.5	(A) Examples of probabilistic Brodmann area maps after translation into cortical surface representation. (B) Freesurfer automatic cortical parcellation based on cortical folding patterns.	18
Figure 2.6	(A) Example of group average myelin density map generated from T1w/T2w ratio and (B) the HCP multi-modal parcellation.	21
Figure 3.1	Pulsed gradient spin echo (PGSE) sequence.	24
Figure 3.2	Diffusion-based cortical imaging methodology and results of Nagy et al. (2013)	33
Figure 4.1	Illustration of the k-means algorithm.	37
Figure 4.2	Example of a decision tree for a simple 2-D clustering problem that contained two classes of equal size.	39
Figure 4.3	Illustration of information gain within a decision tree. (Criminisi and Shotton 2013)	40
Figure 4.4	Example of how forest size affects prediction confidence and over-fitting (Criminisi and Shotton 2013).	42
Figure 5.1	The optimisation of k, in k-means classification, using the trade-off between the best sum of total distances for each clustering solution and the total runtime of the algorithm in seconds.	49
Figure 5.2	The lateral view of the group average diffusion-based parcellation results and group average myelin map.	50

Figure 5.3	The distributions of myelin measurements corresponding to regions of interest from the dMRI parcellation result.	51
Figure 5.4	The medial view of the group average diffusion-based parcellation result and the group average myelin map.	52
Figure 5.5	Plots of diffusion MRI features against myelin density for regions of high and low myelination.	54
Figure 5.6	Unsupervised cortical parcellations using dMRI features combined with spatial or neighbourhood information.	58
Figure 5.7	Cortical parcellation results using a spatially constrained hierarchical clustering algorithm (Blumensath et al. 2013).	59
Figure 6.1	(A) The classification training labels generated from the Human Connectome Project multi-modal parcellation. (B) An example of a neighbourhood of areas, in this case, for the classification of V1. (C) An example of a classification result where instead of the neighbourhood approach, a 180 area multiclass classification is attempted . . .	64
Figure 6.2	Results for binary classification between M1 and S1 using different feature sets.	67
Figure 6.3	Maps of the group average whole hemisphere parcellation result for feature sets DT6, DT9, and SH27 (left to right).	69
Figure 6.4	Group average searchlight cluster coherence comparisons.	70
Figure 6.5	Bar charts comparing classification accuracy across all labels in the group average whole hemisphere parcellation.	71
Figure 6.6	Bar chart comparing classification performances of DT6, DT9 and SH27 in auditory areas.	72
Figure 6.7	Single subject full hemisphere supervised parcellation results.	73
Figure 6.8	The mean DWI signal intensity in areas 3b and 4 (left) and V1 and V2 (right) for a single subject.	75
Figure 6.9	Winner takes all group average parcellation results for DT6, DT9 and SH27 and the single subject winner takes all results for SH27.	77
Figure 6.10	Confirmation that misclassification by the 4T36 feature set is driven by heterogeneity in myelin density within area 3b.	79

Figure 7.1	Regions of interest used in classification experiments to test different b-value combinations.	86
Figure 7.2	The squared correlation coefficients between different feature sets from different b-values.	88
Figure 7.3	Comparison of the performance of each dataset in the binary classification experiments for the triple b-value combinations.	90
Figure 7.4	(A) Maps of the binary classification result between areas 3b and 4/4p using mixed or repeated b-value combinations. (B) The mean F1 score for each ROI using mixed or repeated b-value combinations.	91
Figure 7.5	Maps of the binary classification results using mixed or repeated b-value combinations for A1 vs LBelt and V1 vs V2, shown for a single HCP subject.	92
Figure 7.6	Comparison of the performance of each dataset in the binary classification experiments for the paired b-value combinations.	93
Figure A.1	The percentage difference between HARDI ^{1st} and HARDI ^{3rd} for $b = 0$ and $b=1000$ s/mm ² images and the voxel signal intensity across all diffusion directions for HARDI ^{1st} and HARDI ^{3rd} datasets.	108
Figure A.2	The difference in the SD maps of the HARDI ^{1st} and HARDI ^{3rd} datasets at 6 different image slices.	109
Figure A.3	Median \pm interquartile range of voxel-wise absolute diffusion tensor model residuals across the whole brain.	110
Figure B.1	HCP diffusion data gradient deviation: analysis and correction.	112
Figure B.2	Comparison of the normal FreeSurfer recon-all surface estimation to the HCP surface estimation.	113
Figure B.3	Group average cortical parcellations using different combinations of spherical harmonic features.	114
Figure B.4	Unsupervised cortical parcellations using different group sizes.	115
Figure B.5	Alternative averaging method in which the mode cluster is selected after classification, rather than surface based averaging of the feature vectors prior to classification.	116
Figure B.6	K-means clustering results using different values of k.	116

Figure C.1	Optimisation of random forest parameters for whole hemisphere parcellation experiments.	117
Figure C.2	Group average classification results for the combination of SH27 and DT9 feature sets.	118

Part I

INTRODUCTION

MOTIVATION

Scientific investigation over many centuries has resulted in an in-depth understanding of many of the systems within the human body. However, our knowledge of the brain and the exact workings of its various parenchyma remains relatively sparse and inconsistent in comparison to other organs. This is particularly true of the cerebral cortex, which is associated with many highly complex parallel processes, making it an important structure for neuroscientific inspection.

Early studies of the microstructure of the human cerebral cortex revealed a laminar pattern comprising six layers of varying thickness, and cellular and axonal fibre compositions (cytoarchitecture) (Lewis and Clarke 1878; Berlin 1858). The heterogeneous appearance of these layers as well as differences in vertical and horizontal fibre densities (myeloarchitecture) in different parts of the cortical sheet suggested that there might be a relationship between microstructural organisation and local functional specificity. Pioneers in this field (Campbell 1905; Brodmann 1909; Economo and Koskinas 1925; Vogt 1919) published hemisphere-wide maps based on sectioning and histological staining of cadaver brains, demarcating the boundaries of cyto- and myelo-architectonic domains. Those maps divided the cortical sheet into a complex mosaic based on radial (vertical) and tangential (horizontal) variations in the composition of tissue columns taken perpendicular to the pial and white matter surfaces. It is reasonable to assume that uncovering the mechanisms behind the structure-function relationships of the cortex will involve an understanding of this varying architectural organisation. Yet, these somewhat incompatible parcellations were subject to many methodological criticisms. Their labour intensive nature limited sample size, which was problematic given: (1) inter-subject variability of cytoarchitectonic boundaries (Amunts et al. 1999) but also within-area variation (Serenó et al. 2015), (2) the unavoidable artefacts of the histological process, such as idiosyncratic plastic deformation and tearing of sections, (3) observer bias, and (4) limitation to a single tissue contrast per sample. Despite these

limitations, the hypothesis that architectonic borders are related to function has prevailed and these areas are still widely adopted, for example, to localise activation foci in functional imaging studies.

In vivo image-based methods for analysis of the grey matter have the potential to alleviate or eliminate some of the limitations associated with histological parcellation. These methods are able to account for inter-subject variability through the comparable ease of *in vivo* data collection, can be combined with additional multi-modal data from the same subject to directly assess structure-function relationships, and lend themselves gracefully to observer-free algorithmic analyses. Thus far, image-based studies of the cortex have focussed mainly on the analysis of myelin density via quantitative T1/R1 mapping (Fischl et al. 2004b; Sigalovsky et al. 2006; Geyer et al. 2011; Dinse et al. 2015; Waehnert et al. 2016), R1 mapping in relation to map structure (Dick et al. 2012; Sereno et al. 2013; Lutti et al. 2014), T2* mapping (Cohen-Adad 2014; Sánchez-Panchuelo et al. 2012), MRT (Sánchez-Panchuelo et al. 2014) and the T1-weighted over T2-weighted ratio (Glasser and Van Essen 2011; Glasser et al. 2014). However, these measures as a proxy for myelin density, provide only a single-dimensional description of the multifaceted variations within the cortex. Using myelin density as a sole marker is less informative in areas with low, relatively uniform myelination that are found outside primary and secondary sensory and motor cortices (Glasser et al. 2014). Robust *in vivo* methods for analysing cytoarchitecture, as opposed to these well studied myeloarchitecture methods, have yet to be realised.

Diffusion magnetic resonance imaging (dMRI) has become ubiquitous in the study of white matter (WM) (Le Bihan et al. 2001; Le Bihan 2003) and has been successfully adopted into a plethora of scientific and clinical applications. By measuring the displacement of water molecules within tissue compartments, dMRI offers *in vivo* insight into structural properties of microenvironments. It is sensitive to many different tissue properties, for example, fibre orientation (Doeke et al. 1991), fibre fanning and dispersion (Sotiropoulos et al. 2012; Zhang et al. 2011; Zhang et al. 2012), tissue volume fractions (Jeurissen et al. 2014) and axon diameter (Alexander et al. 2010; Assaf et al. 2008). Although dMRI and T1 (or T2, T2* etc.) are affected by similar structures, e.g., myelinated axons, it is evident that dMRI can provide a richer, multi-dimensional, feature space that has increased potential to distinguish differences in local architecture. Two different cortical regions or layers might contain the same

total amount of myelin, but that myelin might be arrayed differently; for instance, one area might have more radial than tangential fibres.

Initial grey matter (GM) applications of dMRI focused on the developing brain due to its increased anisotropy (Gupta et al. 2005; McKinstry et al. 2002; Mukherjee et al. 2002). Others have used tractography to subdivide the cortex based on the WM connectivity between regions (Anwander et al. 2006; Johansen-Berg et al. 2004; Beckmann et al. 2009). Compellingly, several papers have demonstrated a good correspondence between cortical histology and dMRI of the cortex using high-resolution, *ex vivo* data (Aggarwal et al. 2015; Heidemann et al. 2010; Leuze et al. 2014; McNab et al. 2009; Bastiani et al. 2016). Recent advancements such as simultaneous multi-slice acquisition, improved gradient systems, better motion/eddy current correction algorithms, and ultra-high field MRI have facilitated the use of dMRI in several adult, *in vivo*, cortical studies. Some papers have combined diffusion tensor imaging with cortical surface-based analysis to successfully demonstrate differences between the primary motor (M1) and somatosensory (S1) cortices, and to relate them to histological findings (Anwander et al. 2010; McNab et al. 2013). Others have extended these findings by applying similar features to the medial surface of the cortex, with the aim of understanding how the microstructure of the cortex adapts when it folds (Kleinnijenhuis et al. 2015). Calamante et al. (2017) estimated the apparent fibre density across the cortical sheet, reporting distributions that correlated with known patterns of myeloarchitecture. However, there have been very few attempts to characterise the small but detailed changes in the high angular resolution diffusion imaging (HARDI) signal that would be expected to result from the different architectonic tissue configurations found in different cortical areas (Haroon et al. 2010; Nagy et al. 2013). Notably, Nagy et al. (2013) developed a surface-based cortical classification pipeline that used machine learning to distinguish between several cortical areas based on their dMRI signal, however this study was limited to low resolution data and a small set of test regions.

1.1 PROBLEM STATEMENT

For over a century we have known that the cortex can be divided into a mosaic of distinct units that are well defined by their microstructural properties. The ability to accurately map these areas on live participants has the potential to greatly expand our mechanistic understanding of the brain and even reveal the aetiology of complex neurological disorders. Yet, there are currently no methods that are able to reliably parcellate the whole cortical sheet, based on its fine-grained cytoarchitectonic properties, and thus reproduce cortical area definitions using *in vivo* data.

Diffusion MRI has emerged as a powerful and versatile tool for examining microstructure, particularly within the brain sciences. We ask to what extent can this modality be used in the field of cortical parcellation? We focus on key aspects, such as, the availability of area-specific contrast throughout the cortical sheet and how can we best utilise diffusion data sets to probe grey matter.

1.2 PROJECT AIMS

- To thoroughly explore the utility of diffusion MRI for cortical parcellation at *in vivo* resolutions by using high-quality datasets.
- To use machine learning experiments to test and optimise aspects of a diffusion-based cortical pipeline, answering various questions regarding e.g., pre-processing steps, choice of b-factor, feature representation.
- To determine if a full hemisphere parcellation is plausible using dMRI data and determine if/where this modality is beneficial to future studies.

1.3 CONTRIBUTIONS

EXPERIMENT 1 In Chapter 5 we applied the cortical parcellation pipeline of Nagy et al. (2013) to state-of-the-art, 3T, Human Connectome Project (HCP) datasets. We refined several aspects of the pipeline and applied unsupervised clustering to significantly improve on the full hemisphere parcellation result demonstrated in that paper. This was achieved by adopting an improved surface reconstruction pipeline, reducing the feature space to a more compact set, and performing population averaging. Importantly, we compared our results against a more commonly used myelin mapping technique to demonstrate the advantages of diffusion MRI and potentially complementary information content that this modality can provide.

EXPERIMENT 2: In Chapter 6 we asked which of the many diffusion-based microstructure methods is most appropriate for cortical parcellation? This, to our knowledge, was the first study to compare multiple methods in this domain. We tested the classification framework on several feature sets, including popular WM methods, e.g., diffusion tensor, as well as higher-order non-parametric approaches, e.g., the spherical harmonic features used in Experiment 1. In addition we determined which cortical areas can be reliably distinguished from their neighbours by developing a supervised learning classification framework that utilised the 180 cortical areas defined by Glasser et al. (2016) as plausible training labels. By attempting to classify each area against its neighbourhood, we objectively quantify regional differences across the whole cortical surface, whereas previous works have focussed on a smaller selection of areas. We presented results at both the group level and single subject level across the entire cortical hemisphere to demonstrate that good to moderate area-specific contrast is available in diffusion datasets for the majority of the cortex. Classification performance was evaluated based on the overlap with training labels and a newly developed searchlight cluster coherence method that provides a quantitative measure of parcellation quality.

EXPERIMENT 3: In Chapter 7 we considered the most important aspect of a diffusion MRI acquisition, the b-value. Specifically we explored whether sampling at multiple b-shells provided better contrast between cortical areas than using a single b-shell. Analysis was performed on both the 3T HCP data and locally collected 1.5T data, each comprising of 3 evenly spaced b-shells. We assessed how correlated the information content of different b-shells were and performed a series of binary classification experiments to reveal that datasets containing multiple different b-values yielded the higher classification accuracy than those containing repeated acquisitions of the same b-value.

EXPERIMENT 4: An additional experiment, which does not form part of the main body of this thesis, can be found in Appendix A. Here we applied the cortical parcellation pipeline to test a new hardware system for eddy current correction. The system uses concurrent field monitoring to eliminate diffusion imaging artefacts. The data quality was assessed at several levels with the final stage involving cortical area classification, where improvements in classification accuracy were taken as an indication of the efficacy of field camera. However, the results of this study were deemed inconclusive as the prototype system was unable to deliver data of adequate quality, resulting in the use of post processing steps that obscured interpretation of the final result.

Part II

BACKGROUND

THE CORTEX AND PARCELLATION

The cerebral cortex is a thin sheet, approximately 2-4mm thick, forming the outer surface of the brain. It is estimated to contain 10^{11} neurons and around 10^{14} synaptic connections. It is associated with the performance of complex cognitive functions, such as speech, and vision. The human cortex, unlike many in the animal kingdom, is gyrencephalic - becoming highly convoluted during development. This folding process creates deep crevasses and bulbous ridges known as sulci and gyri respectively. There is no consensus on the exact mechanism that drives cortical folding. It has been theorised that evolutionary selection pressures over time demanded larger surface area to volume ratios and cortical folding is a side effect of the mechanical instability caused by tangential expansion of grey matter (Smart and McSherry 1986; Xu et al. 2010; Tallinen et al. 2014; Tallinen et al. 2016). This is supported by the observation that, unlike primates, small mammals such as rats have lissencephalic brains. Others suspect that mechanical tension between interconnected areas draws spatially distant regions together to create faster, more compact neural circuitry (Van Essen 1997).

The thin, highly folded geometry of the cortex achieves an efficient neuronal system, facilitating complex cognitive processes. However, these exact features also inhibit the direct study of this brain structure. Much of the grey matter is buried deep within the randomly oriented and compact sulci, making it difficult to access. Cortical folding patterns are variable amongst individuals (Zilles et al. 1997; Fischl 2013), making cross subject comparisons problematic. Furthermore, with the exception of the Stria of Gennari (Gennari 1782), the length scales required to resolve architectonic features of the cortex are beyond the capabilities of the human eye.

The following subsections review the study of cortical microstructure, starting with early histological efforts and ending with the current image based techniques. Some of the contents of these sections relied on extensive consultation of Geyer and Turner (2015) which can be referred to for an excellent review of cortical parcellation.

2.1 HISTOLOGICAL STUDIES OF THE CORTEX

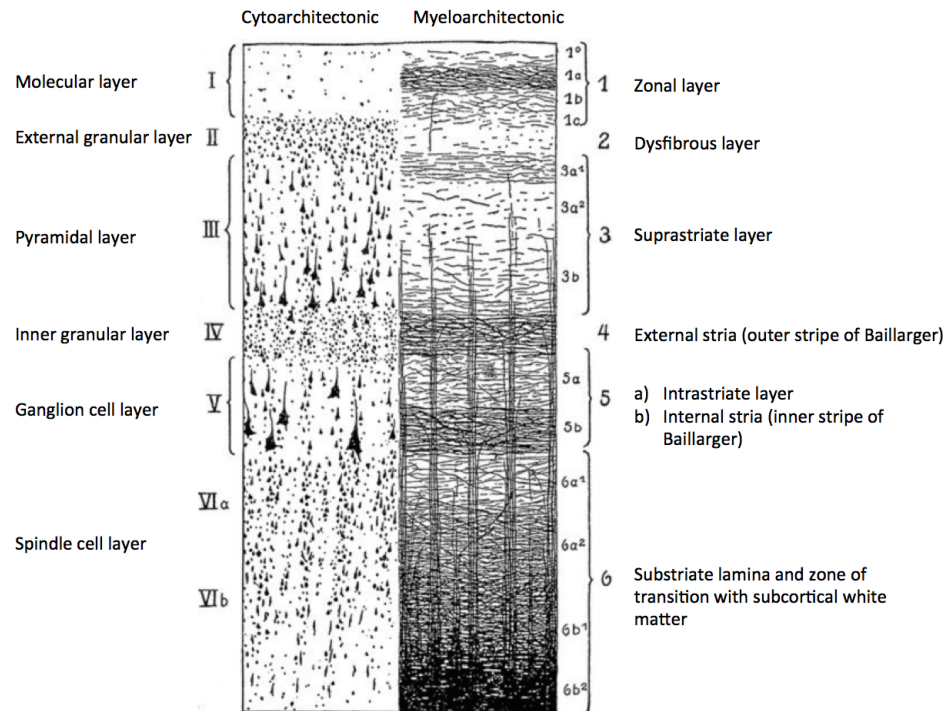


Figure 2.1: The cyto- and myeloarchitectonic layers of the human cerebral cortex. Image adapted from Vogt and Vogt (1919).

Early studies involving brain lesions (Broca 1861; Ferrier 1873; Fritsch and Hitzig 1870) provided evidence to the theory of functional localisation – the notion that the cortex is divisible into functionally distinct units. This was further supported using electrical stimulation, which revealed that neighbouring neurons shared a similar receptive field, e.g., adjacent patches of skin (Ferrier 1875; Shipp 2007). Anatomical evidence to support this hypothesis wasn't provided until the late 19th century when advancements in microscope technology finally facilitated the study of cortical microstructure. Investigators unveiled a laminar organisation comprising of six layers of varying appearance (Berlin 1858; Lewis and Clarke 1878).

The nature of the histological process allowed only a single contrast agent to be applied to a given tissue sample. This resulted in two schools of practice. The first considered the distribution of cell bodies via Nissl staining and is referred to as cytoarchitecture. The

second, known as myeloarchitecture, involved the study of tangential and radial fibre populations using Weigert staining. Illustrations of the six commonly accepted cyto- and myeloarchitectonic layers are given in Figure 2.1. Many researchers reported sharp transitions in the relative appearance of these layers when comparing adjacent radial columns of cortical tissue samples (Brodmann 1909; Lewis and Clarke 1878; Smith 1907; Vogt 1919; Economo and Koskinas 1925). This discovery bolstered the hypothesis of functional localisation with anatomical evidence and motivated efforts to create maps of the neocortex based on its architectonic variations, i.e., cortical parcellations, in the hope understanding the synergy between structure and function.

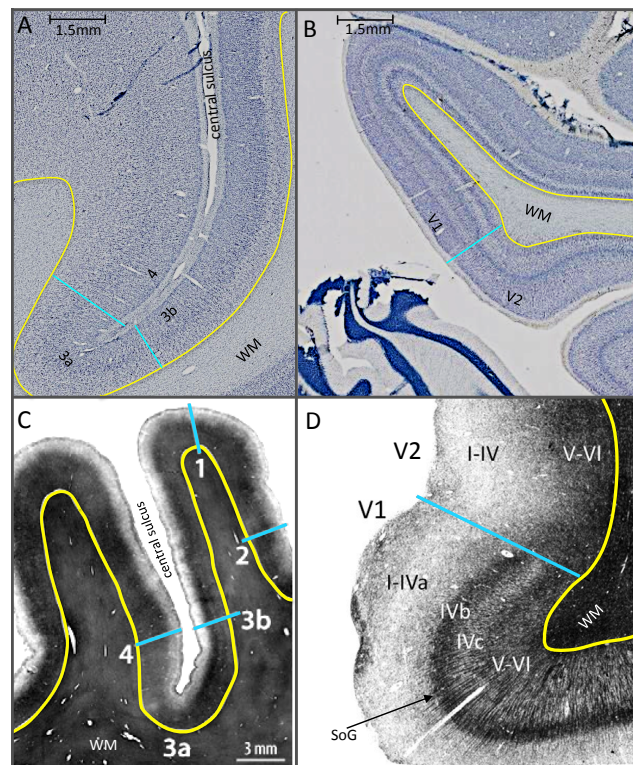


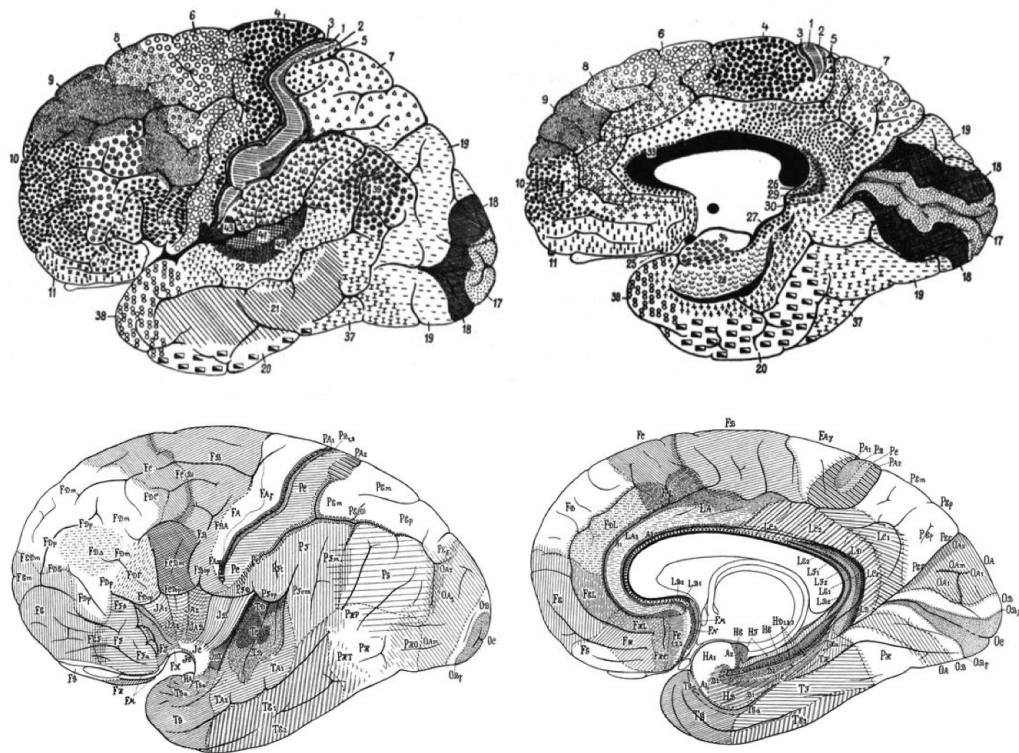
Figure 2.2: Cell body and myelin stains for the sensorimotor areas of the central sulcus and the visual areas of the calcarine. (A) Nissl (cell body) stain in which areas 4, 3a and 3b are visible. (B) Nissl stain showing the transition between V1 and V2. Both A and B were adapted from the Allen Human Brain Atlas (Ding et al. 2017). (C) Myelin stain of the sensorimotor areas 4, 3a, 3b, 1 and 2 adapted from Dinse et al. (2015). (D) Myelin stain showing the transition between V1 and V2 adapted from Amunts and Zilles (2015). The Stria of Gennari (SoG) clearly visible in V1 but absent in V2.

Figure 2.2 shows examples of laminar transitions for the sensorimotor areas in the central sulcus and the visual areas of the calcarine - both sets of areas will be discussed routinely throughout this Thesis.

2.1.1 *Classical cortical parcellation*

In 1909 Korbinian Brodmann published his monograph detailing a full hemisphere cytoarchitectonic map of the human cerebral cortex (Brodmann 1909). His work was so influential that the Brodmann Areas remain a standard and are still used to navigate and interpret studies, over a century later. Whilst many researchers were studying microstructural variation of the cortex around the same time (Campbell 1905; Betz 1874; Hammarberg 1895; Lewis and Clarke 1878; Cajal 1900; Smith 1907), Brodmann surpassed the efforts of his peers by incorporating the theory of evolution to produce a detailed coherent, cross species framework (Zilles and Amunts 2010). Brodmann identified 43 regions (see Figure 2.3) in the human cerebral hemisphere by searching for cellular and laminar variation in a small sample of cadaver brains. He described the location, extent and appearance of the areas in detail, and specified which transitions were easy or difficult to distinguish. By keeping homologous regions consistently named across different species he was able to identify several areas that were present in non-human primates, such as insular cortical areas 14-16, but absent in humans. The thoroughness of his work led to it becoming widely adopted and popularised the notion of anatomical specificity as a driving force behind functional localisation.

Despite their notoriety, the Brodmann Areas cannot be considered a definitive and final delineation of the cortical mosaic. Several of his peers pointed to inaccuracies and discrepancies in his findings compared to their own investigations of cortical architecture. For example, Economo and Koskinas (1925) reported 107 cytoarchitectonic areas using their modified method. Crucially, their work addressed the confounding effects of varying gyral/sulcal orientation by ensuring that tissue samples were taken in columns perpendicular to the local cortical surface. Their atlas is considered by some to be even more comprehensive than Brodmann's work (Triarhou 2012).



(1925).

tonic delineations of Oskar and Cécile Vogt (1903; 1919). By examining variations in the tangential and radial myelin distributions the Vogts determined as many as 200 histological units in the cortex. Over their lifetime, they and their numerous students performed further comprehensive studies of laminar variation to build an understanding of the different microcircuitry within this organ. Some of the boundaries exhibited good correspondence with Brodmann's cytoarchitectonic areas, however in general the results suggested that the cortical sheet could be much more finely partitioned.

clear consensus on the number of areas, or the placement of the cortical subdivisions. Some of the inconsistencies between these works can be explained by limitations of their methodologies. All of the above results were subject to observer biases and human error. The onerous and time-consuming processes allowed only a small sample of brains to be exam-

ined. As such, they fail to account for the large degree of inter-subject variability of cortical boundaries (Amunts et al. 1999; Fischl et al. 2008) and within-area variation (Serenio et al. 2015). Furthermore, these methods were purely qualitative and the resulting illustrations are difficult to accurately adopt in subsequent studies.

2.1.2 Probabilistic atlases

Modern histological works have alleviated the limitations of classical parcellation methods by using observer independent, probabilistic approaches (Schleicher et al. 2005; Zilles et al. 2002). These methods perform histological sectioning and subsequent observer-free intensity/density analysis to define boundaries of areas on multiple subjects. The extent and locations of areas are then translated into stereotaxic space providing probabilistic labels (Geyer et al. 1996; Geyer et al. 1997; Amunts et al. 1999; Amunts et al. 2000; Morosan et al. 2005; Eickhoff et al. 2006a; Eickhoff et al. 2006b). These works have provided valuable atlases that have been widely adopted in modern imaging studies for localisation of activation foci. Contemporary histological studies have even revealed inaccuracies in the Brodmann map, suggesting that certain areas can be subdivided into smaller units (Geyer et al. 2000; Orban et al. 2004; Wandell et al. 2007). They generally accommodate a modest sample size (usually 10 brains) over a limited set of brain areas. The Big Brain initiative has more recently achieved full brain coverage for a single subject (Amunts et al. 2013).

These methods benefit from high spatial resolutions but are subject to artefacts of the histological process, such as idiosyncratic plastic deformation and tearing of sections. They are also limited to a single tissue contrast per sample and inaccuracies may be incurred when translating to 3D image space. Histological methods are invaluable for validation purposes but lack the flexibility that modern *in vivo* imaging could bring to the study of cortical structure and function.

2.2 MRI STUDIES OF THE CORTEX

Magnetic resonance imaging (MRI) provides powerful and versatile machinery for non-invasive investigations of otherwise inaccessible tissues and structures. It is a flourishing

area of research and development that has countless, wide spread clinical and scientific applications, particularly in the neuroscientific domain, where ethical *in vivo* investigations are otherwise limited.

2.2.1 *Basic principals of MRI*

MRI relies on the physical phenomenon of nuclear magnetic resonance (NMR) in which certain nuclei absorb and re-emit electromagnetic radiation when exposed to an external magnetic field. For this interaction to occur, a nucleus must possess an odd number of protons or neutrons, resulting in a non-zero spin and magnetic moment. In biological tissue, such nuclei are abundant in the form of hydrogen atoms held within water and fat. When the hydrogen nuclei are placed within an external magnetic field (referred to as B_0) they precess at a resonance frequency, the Larmor frequency, which depends on the specific gyromagnetic ratio of the nucleus as well as the strength of the B_0 field. The spins align either parallel or anti-parallel to the external magnetic field with slightly larger number of spins in either orientation resulting in a net magnetisation which can be manipulated to infer details about the tissue domain. Radio frequency (RF) pulses are applied to tip the spins (and by extension, the net magnetisation) into the transverse plane where their precession induces a voltage in a receiver coil, thus creating a measurable signal. In MRI this information is spatially encoded into a set of image slices that delineate different tissue types.

Many different types of contrast are available, for example, structural information can be encoded by measuring T1 relaxation time. T1 measures how long it takes for magnetic spins to recover approximately 63% of their longitudinal magnetization after being flipped into the transverse plane i.e., the time it takes for spins to realign with the external magnetic field (Bloch 1946). This is known as spin-lattice relaxation and is dependent on the strength of the external field and the molecular properties of the local microenvironment. At 3T, the nominal value for T1 in white matter is approximately 1300ms, whereas, in grey matter T1 is typically 800ms (Wansapura et al. 1999); therefore the two tissue types can be differentiated in T1-weighted images (Figure 2.4). Similarly, T2 relaxation time relates to the transverse magnetization. It measures the time taken for spins to de-phase due to spin-spin

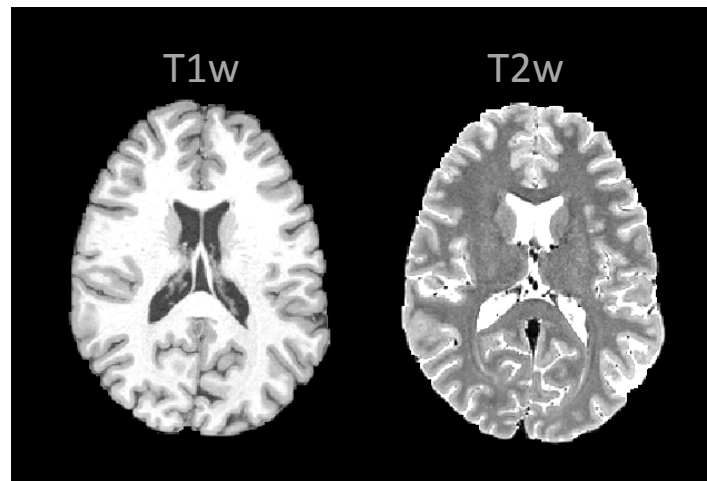


Figure 2.4: Example T1w and T2w image slices from the same subject. WM, GM and CSF have different nominal signal intensities for both contrasts.

interaction and also varies by tissue type (Bloch 1946). In a real experimental system the transverse magnetisation decays faster than predicted from spin-spin interactions due to inhomogeneities in the B_0 field, the effective relaxation time, $T2^*$, can also be measured. Other structural modalities include: R1 which is the rate of longitudinal recovery i.e. $1/T1$; and proton density imaging which measures the concentration of hydrogen atoms within a tissue (Nitz and Reimer 1999). Brain function in live participants can also be studied using MRI by measuring the blood oxygenation level-dependent (BOLD) technique (Ogawa et al. 1990), which is sensitive to the high iron content of haemoglobin in oxygenated blood.

Whilst MRI cannot match the spatial resolution of histological methods, it is extremely versatile and possesses the potential to further our understanding of the functional and structural organisation within the cortex. The following subsections review the *in vivo* estimation of Brodmann areas using MRI, the subsequent development of cortical surface based analysis, and finally, the tools specifically designed to quantify myeloarchitecture.

2.2.2 Surface based analysis

One of the primary applications of Brodmann areas has been to locate activation foci, for example, to investigate topographic organisation in functional MRI studies. However, as discussed above, the descriptions and illustrations posited by Brodmann and his contem-

poraries are not easily adapted into modern imaging frameworks. Furthermore, the highly folded geometry of the cortex and its thin sheet-like structure make it very difficult accurately study using 3D image volumes. The earliest attempt to deal with this difficulty came in the form of a common coordinate system, the Talairach atlas (Talairach and Tournoux 1988). This provided researchers with a reference frame to which they could register their own MRI data to and then estimate the location of cortical effects with relation to a reference brain. The Talairach coordinate system allows for more robust comparison across different subjects but several groups have shown that it does not accurately estimate the location of Brodmann areas (Rademacher et al. 1993; Amunts et al. 2000; Grefkes et al. 2001). Others have settled for using qualitative comparison of cortical folding patterns on image volumes to compare fMRI results across individuals and assign them to prospective cortical areas. Both methods are marred in concerns regarding their accuracy and subsequent hypothesis testing.

The break through moment for image-based study of the cerebral cortex was the development of surface based analysis pipelines (Dale et al. 1999; Fischl et al. 1999a; Fischl et al. 2001). These methods perform automated and robust reconstruction of subject specific cortical surfaces from a single structural 3D volume e.g., a T1w image. Pipelines, such as those in the FreeSurfer analysis suite, perform automatic grey/white matter segmentation and offer multiple additional functions including surface unfolding, flattening and inflation. Data can be sampled from the image voxels onto the vertices of the cortical surface. Sampling can be performed at single or multiple depths allowing laminar-like analysis, depending on the resolution of the 3D volume. Once the initial surface tessellation is generated, it is trivial to sample additional datasets of the same subject, e.g., fMRI time-series, by aligning the new data to the original T1w image. Researchers circumvent the difficulties associated with cortical folding, by moving from volumetric to surface-based analysis, which instead scrutinizes data in a way that reflects the intrinsic 2D organisation of the cortical sheet.

One important development that arose from surface based analysis was the advent of surface based averaging (Fischl et al. 1999b). This method uses the cortical folding patterns to better align data across individuals and improve the precision of any subsequent analyses including the estimation of Brodmann areas. Fischl et al. (2008) compared cortical

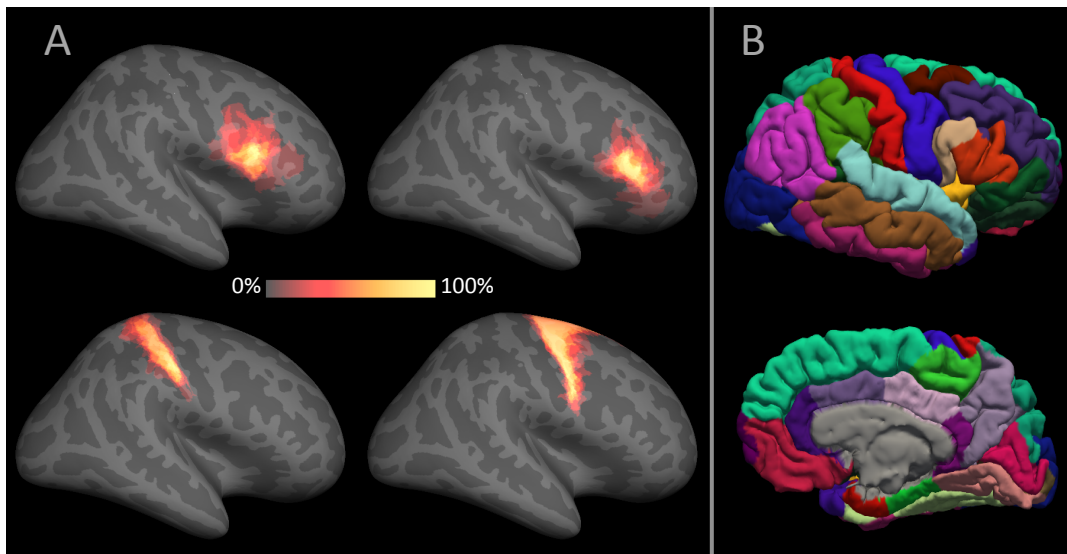


Figure 2.5: (A) Examples of probabilistic Brodmann area maps after translation into cortical surface representation. (B) Freesurfer automatic cortical parcellation based on cortical folding patterns.

folding patterns across 435 brains to confirm that variance across individuals is not uniformly distributed throughout the brain. They verified the hypothesis that the early-forming deeper folds, such as the sylvian fissure, are more stable than secondary and tertiary folds which form later in development (Chi et al. 1977; Yu et al. 2007). Using this information they developed a procedure that accurately aligns cortical surfaces by assigning higher weighting factors to the more stable folding patterns. Prior to this development, researchers relied on intensity based averaging of 3D image volumes, which was problematic because T1w image intensities are not expected to generally correlate with architectonic or functional properties of the cortex. The variance in T1w measures across cortical regions are vanishingly small compared to the variance between the different tissue types, or the myelin rich WM regions, making it unlikely for a registration algorithm to align brains using cortical reference points.

Surface-based analysis facilitates more robust hypothesis testing in cortical studies, allowing activations to be reported using surface coordinates and more reliable mapping to Brodmann areas. For example, users can automatically segment the cortex based on sulcal and gyral landmarks allowing activations to be accurately reported in relation macroscopic features (Fischl et al. 2004a). Alternatively, results can be related to Brodmann areas using the probabilistic atlases discussed in section 2.1.2, which have been translated into the cor-

tical coordinate system (Fischl et al. 2008), as shown in Figure 2.5. However architectonic boundaries are not generally closely correlated with cortical folding patterns. Methods for direct *in vivo* Brodmann mapping, using surface-based analyses, may provide the next leap in our understanding of cortical machinery and more robust analysis and interpretation of fMRI studies.

2.2.3 Image based myeloarchitectonics

Excluding deep brain structures that have high iron content, MRI contrast across the brain is dominated by myelin content (Geyer and Turner 2015). When attempting investigations of cortical microstructure using MRI, it makes sense to first target myeloarchitectonic features. The most salient of these is the Stria of Gennari (SoG), a highly myelinated tangential band along layer IV of the primary visual cortex, which is visible to the naked eye in *ex vivo* samples (Gennari 1782). Several studies have observed the presence of the SoG using *in vivo* proton density imaging (Clark et al. 1992), or T1 contrast (Barbier et al. 2002; Bridge et al. 2005; Carmichael et al. 2006; Walters et al. 2003) and *ex vivo* T2* contrast (Hinds et al., 2008). Many of the earlier contributions were performed at low field strength and low spatial resolutions using anisotropic voxels, and thus don't meet the necessary requirements to consistently observe the SoG, as outlined by Turner et al. (2008). But these findings have been validated at 7T, using high-resolution (0.5 - 0.4mm³) voxels (Trampel et al. 2011) in combination with independent retinotopic mapping of the same subjects (Sánchez-Panchuelo et al. 2012).

In addition to the SoG, investigations of *in vivo* myeloarchitectonics have revealed several other cortical features. Fischl et al. (2004b) measured myelin density via quantitative T1-mapping across the cortical surface, showing relative differences between different regions, such as significantly longer T1 in limbic and association cortices. Importantly, they demonstrated that mapping MR parameters is more useful than taking single weighted images, which cannot be reliably compared across different scanners and sequences. T2-weighted imaging (Yoshiura et al. 2000) and R1-mapping (Sigalovsky et al. 2006; Dick et al. 2012) have been used to observe intensity peaks, in Herschl's gyrus, associated with the high myelin content of the auditory core. Others have related R1-mapping to fMRI map structure,

observing correlations between myelin density and retinotopic maps of several areas (Lutti et al. 2014; Sereno et al. 2013).

The efficacy of these methods has also been tested using *ex vivo* MRI and histological samples. For example, Augustinack et al. (2005) revealed islands in layer II of the entorhinal cortex using proton density imaging and subsequent histological analysis of the same *ex vivo* samples. Bock et al. (2009) compared T1 mapping and T1-weighted imaging with myelin stains in the same marmoset brains to confirm that findings correspond to histology in several regions, e.g., MT, S1, M1, A1, and V1. Similar validation came from Geyer et al. (2011), who used *ex vivo* human samples to show correspondences between the borders of area 4 and 3b and T1 mapping data. Quantitative T1 mapping has also been combined with *ex vivo* T2*-weighted images to study the laminar structure of primary somatosensory, motor and visual areas (Waehnert et al. 2016).

Glasser and Van Essen (2011) and Glasser et al. (2014) proposed using the T1-weighted over T2-weight ratio as a proxy for myelin density. This method removes the MR intensity bias field artefacts that are associated with single T1w or T2w images and improves contrast between heavily and lightly myelinated regions. They produced group average and single subject maps of the full cortical hemisphere. The gradients of these maps were shown to correlate well with probabilistic atlas labels (Fischl et al. 2008) and other functional and anatomical evidence from the surrounding literature (Eickhoff et al. 2005). This was recently extended into a multi-modal cortical parcellation pipeline (Glasser et al. 2016) that leverages T1w/T2w ratio, resting-state and task-based fMRI, and supervised learning algorithms to segment the cortical surface into 180 areas (Figure 2.6). This is an important contribution to the field that somewhat mimics the results of the classical cortical parcellation frameworks discussed above (Vogt 1919). However, these *in vivo* measures of myeloarchitectonic gradients have limited scope across the entire cortex. Figure 2.6 gives an example group average myelin density map, as calculated from T1w/T2w ratio. It is clear here that training data may lack microstructurally driven contrast in areas of low myelin density, such as the prefrontal cortex, relying more heavily on fMRI data and the user-driven annotations that were also inputs to the supervised learning framework. A technique for *in vivo*, subject specific parcellation of the entire cortex, driven only by fine-grain cyto and myeloarchitectonic features, has yet to be accomplished. Furthermore, as Glasser et al. (2016) used the myelin density

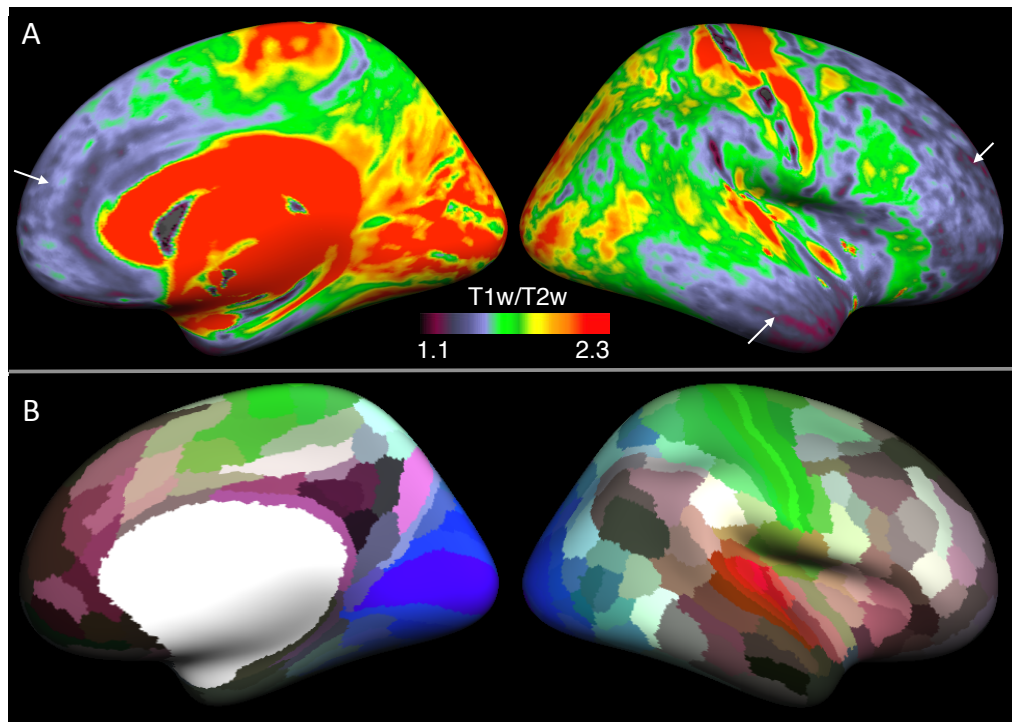


Figure 2.6: (A) Example of group average myelin density map generated from T1w/T2w ratio. T1w/T2w data was surface averaged for 17 HCP subjects, white arrows point to areas of low myelin density. (B) The 180 area definitions of the HCP multi-modal parcellation defined in Glasser et al. (2016). Areas belonging to similar functional networks are displayed in similar colours.

maps to both align and segment the data; the results might be confounded by this circular reasoning. A comparison of this new registration method to the old surface based method, which used only gyral and sulcal landmarks (Fischl et al. 1999b) to align surfaces across different subjects, could determine the extent of this effect.

This chapter reviewed the background of cortical parcellation, from its inception to state-of-the-art multi-modal results. *In vivo* investigations of cortical microstructure have focussed on myeloarchitectonics, using structural modalities e.g., T1 mapping. Whilst these techniques have successfully delineated many cortical areas, they are generally limited to myelin rich regions, and do not provide sufficient architectonic contrast to parcellate the entire cortex. This is partly because myelin density is a one dimensional measurement that cannot adequately

describe the numerous properties that define architectonic areas in traditional histological parcellations. In the next chapter, we will discuss diffusion MRI and its potential to provide a much richer, set of features that could better characterise the cyto- and myeloarchitecture of the cortex.

DIFFUSION MRI

Diffusion MRI measures the displacement of water molecules within tissue structures to infer microstructural tissue properties. It has a wide range of clinical and research applications and has become ubiquitous in the study of brain microstructure. In this chapter we overview the principals of diffusion MRI, and its use in neuroimaging studies. Some of these topics are covered in more detail in Jones (2010). We end with an overview of the cortical grey matter pipeline (Nagy et al. 2013) upon which the work of this Thesis was built.

3.1 PRINCIPALS OF DIFFUSION MRI

Brownian motion describes the random motion of molecules resulting from heat dissipation within a medium. It was first described by Robert Brown (1828) who observed that pollen molecules suspended in water randomly gyrated and migrated across the medium in a manner that could not be explained by fluid currents or any other known physical phenomenon. In an unrestricted medium the molecules are free to move in any direction with equal probability. Thus, Einstien formulated the mean displacement of molecules, in 3-dimensional space, as:

$$\langle R \rangle = \sqrt{6DT_d} \quad (3.1)$$

where T_d is the diffusion time, D is the diffusion coefficient, and $R = r - r_0$ is the average displacement of a molecule from its start position r_0 to its end position r after T_d has passed.

In biological tissues unrestricted diffusion processes are rare. Instead water populations are bound by tissue structures, such as cell walls or myelin sheaths, that restrict diffusion along certain directions. Therefore, the diffusion pattern is different to the isotropic behaviour described by 3.1 and can be used to characterise the underlying microstructure. For example, diffusion in WM fibres is restricted in the plane perpendicular to the axon, but molecules can move more freely along the length of the axon. This creates an anisotropic

diffusion profile. Another example is the water trapped in the extra-axonal space, these molecules are not restricted to a confined space, but their movement is hindered by the presence of cells. This is known as hindered diffusion and again alters the observed diffusion profile.

3.1.1 PGSE sequence

In an MRI experiment a diffusion weighted image (DWI) can be acquired by applying diffusion sensitising gradients (Hahn 1950). The most common method for achieving this is the Stejskal-Tanner pulsed gradient spin echo (PGSE) sequence (Stejskal and Tanner 1965) (Figure 3.1).

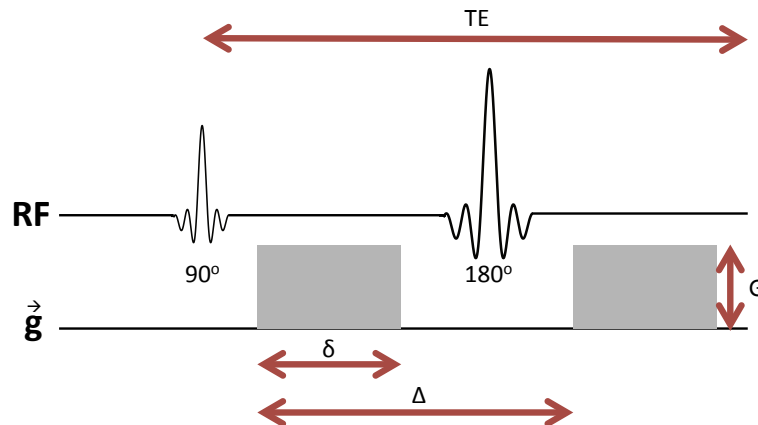


Figure 3.1: Pulsed gradient spin echo (PGSE) sequence. A 90° RF pulse is used to flip the spins into the transverse plane. This is followed by a diffusion sensitising gradient that produces a offset in the spin frequencies, and a 180° RF pulse. A second diffusion gradient, identical to the first, is then applied. The 180° should ensure that the two gradients cancel, however the random motion of the molecules, i.e., diffusion, will result in an offset that attenuates the signal.

This method places two identical gradient pulses on either side of a 180° radio frequency (RF) pulse. In the absence of diffusion, the effects of the first gradient pulse should be cancelled by the second gradient pulse. In reality, the molecules diffuse from their original position in the time between the two pulses, consequently, the phase offset induced by

the first pulse is not fully negated. As the offset encoded by the gradient is spatially dependant, the displacement of the molecules leads to reduction in the total phase accumulation and thus a reduced signal amplitude. The signal attenuation can be controlled by adjusting parameters of the PGSE sequence i.e., the gradient strength, G , the gradient direction, \mathbf{g} , the pulse duration, δ , and the pulse separation, Δ . These variables are usually combined into a single term known as the diffusion weighting factor or b-value:

$$b = \gamma^2 G^2 \delta^2 \left(\Delta - \frac{\delta}{3} \right) \quad (3.2)$$

where γ is the gyromagnetic ratio of the medium and $\Delta - \frac{\delta}{3}$ is the diffusion time, T_d , which defines the length scales probed by the experiment. Higher b-values correspond to higher diffusion-weighting. The limiting factor is generally imposed by the gradient system, whereby higher gradient strengths allow for faster diffusion encoding, shorter diffusion times and thus shorter length scales to be probed, whilst maintaining adequate SNR. The b-value is related to the signal attenuation according to:

$$S(b, D) = S_0 e^{-bD} \quad (3.3)$$

where S_0 is the signal without diffusion weighting i.e., the T2w signal, and D is the diffusion coefficient. In biological tissues the diffusion coefficient is affected by characteristics of the tissue environment, and not just the diffusivity of the medium, as in 3.1, and is usually referred to as the apparent diffusion coefficient (ADC).

SHORT GRADIENT PULSE ASSUMPTION: in the PGSE experiment it is assumed that $\delta \ll \Delta$, this is known as the short gradient pulse (SGP) approximation. Under the SGP assumption, diffusion that occurs during the gradient pulse can be ignored. In real systems it is not possible to achieve the rectangular gradient depicted in Figure 3.1. Instead, gradients are trapezoidal and the resultant slew rate needs to be considered. To achieve adequate diffusion contrast using typical clinical gradient systems, capable of maximal gradient strengths of 40-60mT/m, it is often necessary to use longer pulse durations, again violating the SGP assumption. This may lead to significant underestimation of diffusion distances (Mitra and Halperin 1995). These effects can be minimized by adopting high performing gradient coils, such as the 300mT/m system used in the Human Connectome Project.

EDDY CURRENTS: diffusion weighted images often suffer from geometric distortions caused by eddy currents. The diffusion sensitising gradients induce eddy currents, which in turn induce magnetic fields that remain beyond the duration of the gradient pulse. These field inhomogeneities alter the reconstructed image. Eddy currents can be minimized within the pulse sequence, e.g., the double refocused PGSE (Reese et al. 2003). Alternatively, they can be corrected using post processing algorithms (Andersson and Sotiropoulos 2016; Andersson and Skare 2002; Jenkinson and Smith 2001).

3.2 ESTIMATING THE ADC

3.2.1 *The Diffusion Tensor*

The most common method for estimating the ADC within a voxel is to model it as a three-dimensional, second-order tensor, the diffusion tensor (DT) (Basser et al. 1994). Under this formulation the signal attenuation is given by:

$$S(b, \mathbf{D}, \mathbf{g}) = S_0 e^{-b \mathbf{g}^T \mathbf{D} \mathbf{g}} \quad (3.4)$$

where \mathbf{g} is the gradient vector and \mathbf{D} is now the diffusion tensor which can be expressed as a positive, definite, symmetric matrix:

$$\mathbf{D} = \begin{bmatrix} d_{xx} & d_{xy} & d_{xz} \\ d_{xy} & d_{yy} & d_{yz} \\ d_{xz} & d_{yz} & d_{zz} \end{bmatrix} \quad (3.5)$$

\mathbf{D} has six degrees of freedom. The parameters, d_{xx} etc, can be estimated using a least squares approach, after acquiring a minimum of six DWIs using non-coplanar, equally spaced gradients.

After estimating \mathbf{D} from the DWIs, many characteristics of the tissue environment can be deduced via eigen-decomposition (Le Bihan et al. 2001; Le Bihan 2003). For instance, the mean diffusivity (MD), describes the average amount of diffusion within a voxel, and is defined as $\text{Tr}(\mathbf{D})/3$. Another commonly used metric is fractional anisotropy (FA) which encodes the degree of anisotropy within a voxel. CSF voxels exhibit low FA whilst a voxel

containing a single WM fibre population would have high FA. The principal eigenvector is aligned with the dominant diffusion direction, i.e., parallel to the axons in a fibre population. This makes it possible to infer tissue orientation in anisotropic voxels.

This DT has been used in many clinical applications including the identification lesions in Multiple Sclerosis (Poonawalla et al. 2008). It is generally favourable on account of its very modest acquisition requirements. However, for many applications the DT model provides an overly simplified description of diffusion processes. For instance, the DT assumes Gaussian diffusion which is not an accurate description for complex tissue structures particularly at high b-values (Clark and Le Bihan 2000). The single compartment model is not sensitive enough to characterise the tissue environment for the majority of brain voxels in which fibre crossings are present (Jeurissen et al. 2010), and the various metrics lack specificity and sensitivity to the acute pathological changes that occur with many diseases.

3.2.2 *Beyond the Diffusion Tensor*

It is now common-place to acquire a large number of DWIs- often 60 or more gradient directions are applied, sometimes using multiple b-values. Such data is known as high angular resolution diffusion imaging (HARDI) data (Tuch et al. 2002). HARDI protocols facilitate more precise descriptions of diffusion behaviours in biological and have been utilised to develop numerous, more complex, methods for estimating the ADC in brain tissue.

One school of practice aims to develop biophysical models to estimate the expected ADC using a-priori information about tissue composition. Models usually contain multiple compartments with associated volume fractions, e.g., extra-cellular and inter-cellular compartments. Many geometric representations have been proposed such as the Ball-Stick (Behrens et al. 2003) which describes restricted diffusion as a stick running parallel to the axon direction and free diffusion as a ball. Other similar models include CHARMED (Assaf and Basser 2005), DIAMOND (Scherrer et al. 2013) and multiple-stick models (Hosey et al. 2005), to name a few. With such a wide variety of options, model selection can be an issue, especially given that accurate a-priori knowledge is not available for the majority of brain voxels. Numerous models have been compared in fixed rat corpus callosum (Panagiotaki et al. 2012) and *in vivo* human corpus callosum (Ferizi et al. 2015). The rat study suggested

three compartment models best explained the data, but even two compartment models that incorporate restriction perform better than the DT. The human study indicated an upper limit, of 4-5, on the number of free parameters that can be accurately estimated using standard scanning hardware.

Multi-compartment models are often favoured because they can be designed to estimate specific neural tissue properties. For instance AxCaliber (Assaf et al. 2008) and ActiveAx (Alexander et al. 2010) map axon diameter distributions within voxels. The NODDI model (Zhang et al. 2012) and its Bingham extension (Tariq et al. 2016) provide estimates of neurite orientation and dispersion, as does Ball-and-Rackets (Sotiropoulos et al. 2012). The specificity offered by such methods may be beneficial in some applications. However, model assumptions may afflict their accuracy across the full range of widely varying neural tissue environments, rendering them difficult to generalise to more complex tissue types, such as GM. Moreover, the biophysical assumptions are inherently difficult to validate, particularly in pathological conditions.

Non-parametric approaches have been proposed for estimating more complex diffusion profiles. The high-order terms of the HARDI signal have been characterised using spherical harmonic (SH) expansion (Alexander et al. 2002), where a fourth order SH series was fit to the ADC and voxels were classified as isotropic, Gaussian or non-Gaussian. Diffusion kurtosis imaging (DKI) (Jensen et al. 2005) also characterises non-Gaussian diffusion via cumulant expansion. Spherical deconvolution (Tournier et al. 2004; Tournier et al. 2008) estimates the fibre orientation distribution, allowing fibre crossings to be identified. Higher-order methods are particularly useful for increasing the accuracy of tractography algorithms (Parker and Alexander 2005; Conturo et al. 1999; Parker et al. 2003). These methods provide more generalised instruments for imaging studies than the compartment models above, whilst also mitigating the short-comings of the diffusion tensor.

3.3 DIFFUSION BASED CORTICAL IMAGING

The methods discussed above were designed and tested with WM in mind, but diffusion MRI is less well explored in GM. As discussed in Chapter 2, the cerebral cortex is a very thin structure containing complex microcircuitry. This results in low anisotropy values, even com-

pared to complex WM fibre crossings, and issues with partial volume contamination when the spatial resolution of the data is larger than cortical thickness values. Thus microstructure imaging in the cortex has been historically limited.

Initial GM applications focussed on the developing brain (McKinstry et al. 2002; Mukherjee et al. 2002) where cortical anisotropy values increase with gestation age up to 28 weeks and decreases thereafter (Gupta et al. 2005). Others have used tractography to parcellate the cortex based on the WM connectivity between regions, building rich structural networks (Johansen-Berg et al. 2004; Sporns et al. 2005; Anwender et al. 2006; Beckmann et al. 2009). Connectivity mapping provides valuable insight into brain organisation that is complementary to Brodmann mapping techniques. Although the accuracy of tracking algorithms should be carefully considered. Descoteaux et al. (2016) developed a realistic phantom to test numerous tractography approaches and reported a high false positive rate with all methods. Furthermore, a gyral bias has been observed, wherein tracts preferentially terminate at gyral crowns - a trend that does not align with histological findings (Van Essen et al. 2013a).

Architectonic properties of the cortex have been directly measured using ultra high field MRI and *ex vivo*, fixed tissue (McNab et al. 2009; Heidemann et al. 2010; Bastiani et al. 2016). These studies benefit from high resolution datasets in which anisotropic diffusion can be observed even in adult cortical tissue. For example, Leuze et al. (2014) observed layer specific changes to diffusion orientation, including the Stria of Gennari, using fixed human tissue scanned at 242 μm isotropic resolution. Aggarwal et al. (2015) used spherical deconvolution at 90 μm resolution to show layer specific changes in diffusion orientation between different functional areas. Their comparison to histological stains of the same tissue samples revealed good agreement between the diffusion tensor orientations and cortical fibre orientations. Again, they observed a coherent tangential band in V1 (the SoG) and were able to identify the boundary between V1 and its neighbouring region, marked by the abrupt absence of this band. They report similar findings when comparing the prefrontal cortex to the primary motor cortex where tangential bands corresponding to the inner and outer stripes of Baillarger were visible in the former but not the latter across multiple subjects.

In an early *in vivo* adult study Vestergaard-Poulsen et al. (2007) developed a two compartment model to estimate a diverse set of biomarkers and describe changes that occur in brain ischemia. However, imaging of the microscopic details that define cortical areas has

only recently become plausible through advancements such as simultaneous multi-slice acquisition, improved gradient systems, better motion/eddy current correction algorithms. For example, Haroon et al. (2010) showed consistent changes between cortical regions across subjects by using HARDI acquisitions to assign a complexity value to each voxel based on the number of dominant diffusion directions that were present. Anwander et al. (2010) developed a measure of radiality to quantify the dominant diffusion direction with respect to the local tissue orientation, at much coarser *in vivo* resolutions than the above *ex vivo* studies. They revealed an association between functionally distinct regions and the radiality index in 1.5mm^3 isotropic voxels. Specifically, they found that the motor cortex and prefrontal cortex exhibit higher fractional anisotropy (FA) values and are dominated by radial diffusion processes, while the somatosensory and parietal cortices have lower FA and mostly tangential diffusion. This observation was supported by McNab et al. (2013) who again reported large changes in the dominant diffusion orientation between the primary somatosensory (S1) and motor (M1) cortices using surface-based analysis of the diffusion tensor and radiality index with 3T 1mm^3 isotropic data. Their additional partial volume analysis confirmed that these differences were not merely partial volume artefacts attributed to lower cortical thickness in S1. Kleinnijenhuis et al. (2015) used similar surface adapted DT metrics on 1mm^3 7T images of the medial wall to investigate the relationship between cortical folding and diffusion. Moving away from the DT, Calamante et al. (2017) used spherical deconvolution to estimate the apparent fibre density (AFD) across the cortical sheet. They reported region specific changes that correlate with known patterns of myeloarchitecture and importantly demonstrated that variance in the AFD cannot be fully explained by variance in myelin content or cortical thickness.

These works demonstrate that HARDI techniques can effectively capture microstructural changes between cortical regions, potentially making *in vivo* cortical parcellation possible. Crucially, they all relied on surface-based analysis to circumvent the difficulties of cortical folding, and in some cases, to provide laminar-like analysis. Yet very few attempts have been made to characterise the small but detailed changes in signal expected to result from the different architectonic tissue textures found in different cortical areas. Neither have there been any attempts to compare the multitudinous WM techniques using datasets containing high quality cortical signal.

3.4 HARDI MODEL FOR DISCRIMINATING CORTICAL AREAS

The following is an overview of the approach presented by Nagy et al. (2013), which characterises the dMRI specifically for the purpose of cortical area classification. Deoni and Jones (2006) demonstrated that the shape of the diffusion signal varies between regions in thalamic grey matter when interpreted as a time-series-like object (see Figure 3.2A). Nagy et al. (2013), expanded this to cortical grey matter by modelling the 3D shape of the diffusion profile as a per-voxel “fingerprint”. They acquired two HARDI datasets on different days from the same subject, with 61 diffusion weighted images at $b = 1000\text{s/mm}^2$ to measure effectiveness and reproducibility of this method.

To capture the increased complexity, and reduced anisotropy of cortical grey matter compared to WM, the ADC was estimated using sixth-order spherical harmonic expansion. Laplace’s spherical harmonics form a set of orthogonal basis functions which can be linearly combined to represent any complex spherical function:

$$f(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l c_l^m Y_l^m(\theta, \phi) \quad (3.6)$$

where Y_l^m are the spherical harmonics and c_l^m are the coefficients. The HARDI signal is sampled using a set of diffusion sensitising gradients described by the polar and azimuthal angles of a sphere, θ and ϕ . Therefore, the signal attenuation in a voxel is a function of these two parameters and can be represented using the spherical harmonic series. The ADC must be real valued and have antipodal symmetry; these physical constraints make the problem more tractable, requiring only even orders of l to be estimated. The fitting procedure is formulated as a least squares problem that can be efficiently computed.

Following the estimation of the ADC a total of 27 cortical features were extracted and used as descriptors for classification. These features, or subsets of these features appear throughout the experimental chapters of this Thesis and are given below.

1. The mean of the ADC (\bar{f}).

$$\bar{f} = \frac{1}{4\pi} \int f(x) dx \quad (3.7)$$

2-10. The k^{th} moments of the ADC for $k = 2, 3, \dots, 10$.

$$M_k = \int_S [f(\mathbf{x})]^k d\mathbf{x} \quad (3.8)$$

11. The value of $f(\mathbf{n})$, i.e the value of f along the local cortical surface normal \mathbf{n} .

12. The mean ADC in the plane perpendicular to \mathbf{n} .

$$\bar{f}_{\perp} = \frac{1}{2\pi} \int_{C(\mathbf{n})} f(\mathbf{x}) d\mathbf{x} \quad (3.9)$$

13-21. The k^{th} moments of f in the plane perpendicular to \mathbf{n} for $k = 2, 3, \dots, 10$.

$$M_{k,\perp} = \int_{C(\mathbf{n})} [f(\mathbf{x})]^k d\mathbf{x} \quad (3.10)$$

22-23. The two eigenvalues of the Hessian matrix of f evaluated at \mathbf{n} .

24-27. Simple rotationally invariant combinations of the SH coefficients across all indices $m = -l, \dots, l$ for each of the orders, $l = 0, 2, 4, 6$.

$$I_l = \sum_{m=-l}^l |c_l^m| \quad (3.11)$$

Features 1-10 are fully rotationally invariant descriptors of the ADC. 11-21 are similar descriptors of the ADC but are invariant in the local tangent plane and should characterise radial and tangential diffusion processes relative to the local surface normal. In 22-23 the Hessian matrix describes the curvature of the ADC profile which is sensitive to fibre dispersion. Features 24-27 represent simple additional invariant combinations of the spherical harmonic coefficients.

A support vector machine (SVM) was used to differentiate between several spatially separated regions of interest using the above feature space. A mean classification accuracy

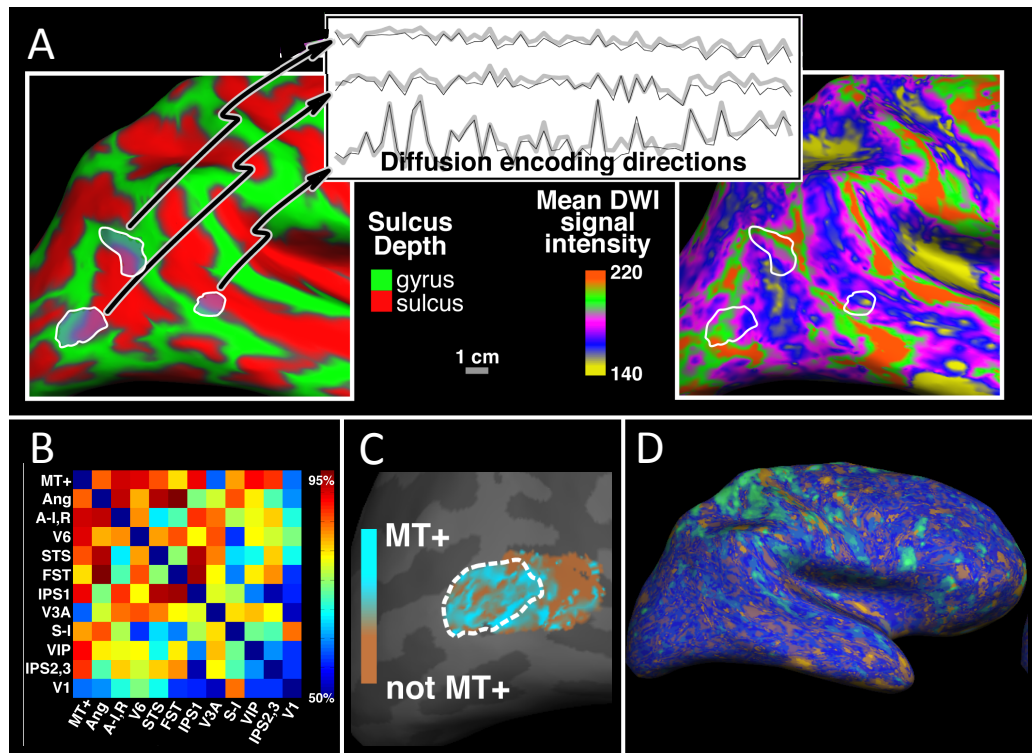


Figure 3.2: Diffusion-based cortical imaging methodology and results of Nagy et al. (2013). (A) The mean DWI data of three ROIs, plotted as a "time-series" across the diffusion encoding directions. The left panel shows the ROIs in relation to cortical folding, whilst the right displays the mean diffusion signal at a single diffusion encoding direction. This signal is clearly correlated with cortical curvature values; therefore, a rotationally invariant representation of the signal is necessary for cortical area parcellation that is not confounded by curvature effects. When comparing the "time-series" objects in the centre panel, the differences between the three ROIs are much larger than the differences between test (black lines) and re-test (grey lines) of the same ROI. This supported the hypothesis that the differences between ROIs are not simply noise related. (B) Classification accuracies for all binary SVM tests. (C) Boundary detection for MT+. The dashed white line represents the outline of the ROI. (D) Full hemisphere parcellation result using k-means on the SH feature set of a single participant.

of 75% (Figure 3.2B) and high reproducibility between the separately acquired datasets was reported. In addition, they were able to locate the boundary of the MT+ region when attempting to classify two neighbouring cortical areas (Figure 3.2C). The results suggest that microstructural changes associated with the different regions are characterised in the

estimated ADC. Figure 3.2D shows the hemisphere-wide k-means clustering for a single subject. Whilst the results do not resemble the Brodmann map (see Chapter 2), some structure is visible in the central sulcus.

The methods presented by Nagy et al. (2013) serve as a proof of concept that diffusion MRI can be harnessed for cortical parcellation applications. However, the low spatial resolution of the data, 2.3mm^3 , may have resulted in substantial partial volume effects, as such, these methods need to undergo more rigorous testing and validation using more reliable datasets. The full hemisphere parcellation result demonstrates that the method needs considerable refinement if it is to be used for *in vivo* Brodmann mapping.

In this Chapter we introduced the principals of diffusion MRI and how it can be exploited to investigate microstructure. We discussed how myelinated axons restrict the diffusion of water in the direction perpendicular to their long axis, thus there is an implicit correlation between measures of anisotropy and myelin content in neural tissue. However, as discussed above, dMRI can be sensitised to many different aspects of neural tissue environments, including neurite density, axon diameter, orientation dispersion, to name a few. Diffusion anisotropy has even been measured in nonmyelinated axons (Beaulieu 2011), and investigators believe that cortical anisotropy is more heavily influenced by the density of intact cell membranes than it is by myelin density (Beaulieu 2002). The evidence presented in this Chapter alludes to the notion that diffusion MRI can provide a much richer, multi-dimensional, feature space for cortical parcellation than the one-dimensional proxies for myelin density which were presented in Chapter 2. Compellingly, some of the studies discussed above have already successfully measured changes in cortical microstructure in between different regions of the cortex by using dMRI. In this Thesis we will focus on expanding on these findings using objective machine learning approaches to further explore the capabilities of diffusion MRI in cortical parcellation applications.

MACHINE LEARNING

Machine learning (ML) is a branch of artificial intelligence (AI) in which generalised algorithms are constructed to recognise, learn from and predict important patterns within datasets. ML methods are popular in a variety of applications and represent the leading edge of data analytics. The appeal of these techniques lies in their scalability to large, complex and multi-dimensional datasets that are readily available in the information age. Amongst the recent slew of extraordinary accomplishments is the AlphaGo project in which neural networks were trained in Go, the complex turn-based strategy game, earning landmark victories against professional human players (Silver et al. 2016). Powerful ML algorithms hold enormous potential to improve our understanding and diagnostic capabilities within the brain sciences, where rich image-based datasets are abundant. For instance, graphical models and hierarchical clustering have been utilized to identify functional connectivity networks (Blumensath et al. 2013; Rajapakse and Zhou 2007). Pattern recognition algorithms have been proposed to improve objectivity and accuracy in MR "fingerprinting" by providing more quantitative analysis (Ma et al. 2013). ML methods are particularly appealing because they allow objective and data driven analysis of MRI data.

Broadly speaking, ML algorithms fall into one of two categories, 'unsupervised' or 'supervised' learning. Unsupervised learning techniques are designed to cluster unlabelled datasets based on their hidden structures. Contrastingly, supervised learning approaches implement a training stage designed to inform the algorithm on the type of partitioning that the user is interested in. For example, if a user wishes to generate a face recognition model, the training data could consist of a selection of patches of images for which the user has manually associated a label: (1) containing a face or (2) background. The algorithm can then determine the characteristics that are common in face patches and uncommon in background patches and subsequently be used to predict the likelihood that a new unseen patch contains a face. In this thesis we adopt both unsupervised and supervised learning

approaches and the specific algorithms that were implemented are detailed below. For an excellent introduction to machine learning techniques, including those outlined below, the reader is advised to consult Bishop (2006). A more focussed discussion of random forests can be found in Criminisi and Shotton (2013).

4.1 K-MEANS

K-means is an unsupervised clustering algorithm. It is computationally efficient and therefore a popular choice for initialising more expensive algorithms but also produces sufficient results for many applications.

Given a dataset $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$ of N observations, each described by a D dimensional feature vector \mathbf{x} , k-means aims to partition the observations into K clusters, where $K < N$. A cluster is considered a set of data points which exhibit low variance between their respective feature vectors compared to other data points outside the cluster. The initial stage initialises K cluster centres, μ_k as representative data points for each cluster. The cluster centres can be randomly initialized with arbitrary but sensible values or chosen randomly from the dataset. A two steps procedure is repeated until the sum of the squared distances between each data point and its closest cluster centre is minimised or a maximum number of iterations has passed. The objective function is:

$$J = \sum_{n=1}^N \sum_{k=1}^K r_{nk} \|\mathbf{x}_n - \mu_k\|^2 \quad (4.1)$$

where $r_{nk} \in \{0, 1\}$ defines which of the K clusters the data point is assigned to.

$$r_{nk} = \begin{cases} 1 & \text{if } k = \operatorname{argmin}_j \|\mathbf{x}_n - \mu_j\|^2 \\ 0 & \text{otherwise} \end{cases} \quad (4.2)$$

First, each data point is assigned to the cluster centre, μ_k that it is closest to, thus minimizing J with respect to r_{nk} with fixed μ_k . In the second step, the cluster centres are recalculated as the mean of all data points within each cluster. Figure 4.1 depicts the two-step procedure.

Due to the random initialisation step, there is no guarantee the J will converge at the global minimum (MacQueen 1967), therefore, it is general practice to repeated the algorithm several times and select the solution that achieved the lowest sum of squared differences.

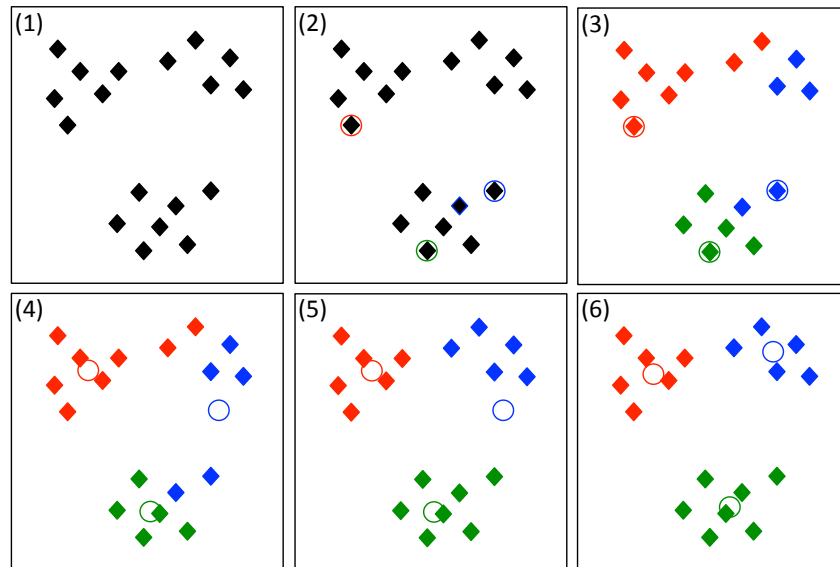


Figure 4.1: Illustration of the k-means algorithm. (1) the input data, (2) the cluster centres, depicted as coloured rings, are randomly initialised, (3) the data points are assigned to the cluster centre they are closest to, (4) the cluster centres are recalculated from the data points within each cluster, (5) data points are re-assigned to the cluster centre they are closest to, (6) the cluster centres are recalculated. At the next iteration, moving any of the data points will not reduce the sum of squared distances, therefore, this is the final solution.

The k-means algorithm is a data driven approach that is simple to compute and easy to interpret. As a consequence of this simplicity, it has some limitations. The algorithm requires a-priori estimates of K which are not always available. It does not scale very efficiently for very large datasets as the distance needs to be computed for each data point at each iteration. The similarity between data points and cluster centres is measured using Euclidean distance, which is generally appropriate for continuous variables, but cannot be used with categorical labels. Furthermore, measures of Euclidean distance are also affected by outliers. More general similarity measures are sometimes adopted, such implementations are usually referred to as *k-medoids*, and may be more computationally demanding depending on the chosen similarity metric. Finally, as it is a non-probabilistic approach, it does not provide uncertainty measures which may be preferable in some applications.

In this Thesis we use the k-means algorithm in Chapter 5 to perform unsupervised clustering of cortical diffusion data. The goal of the experiment was to determine if regionally specific information is organically available in the data without the introduction of any user defined priors. Thus, k-means clustering emerged as the fairest way to partition the data without the influence of a-prior information.

4.2 RANDOM FORESTS

Random forests are an ensemble learning method in which multiple randomly trained decision trees are combined to perform classification or regression tasks (Breiman 2001). Whilst singular decision trees are prone to overfitting to the training data, a randomised ensemble of trees overcomes this limitation to provide greater accuracy on unseen data. In this thesis we used only binary trees to perform supervised classification, the discussion below has been focused accordingly.

Decision trees

Decision trees are a type of directed graph that contain no loops and have a hierarchical structure. Each node has exactly one input edge and, in binary trees, internal nodes split the data between their two output edges. Terminal (leaf) nodes determine the final classification outcome of the data. The process can be interpreted as a series of tests at each node which are used to split a complex problem into a series of simpler ones.

The tree structure is dependant on a set of training data, $S = \{\mathbf{v}_1 \dots \mathbf{v}_N\}$, where each data point, \mathbf{v}_n , is a multidimensional feature vector, that is accompanied by a known training label. An energy function is optimised at each node to minimize the heterogeneity within each of the two subsets, S_i, S_j , that are sent to its child nodes. At the child node a new energy function is optimised and the data split again. The process is repeated until a stopping criteria is met (Figure 4.2). The parameters of the tree structure are learned automatically from the training data.

Following the training phase, a new unseen data point traverses the tree via the opti-

mized splitting functions until it lands at a leaf node, where its class label is predicted.

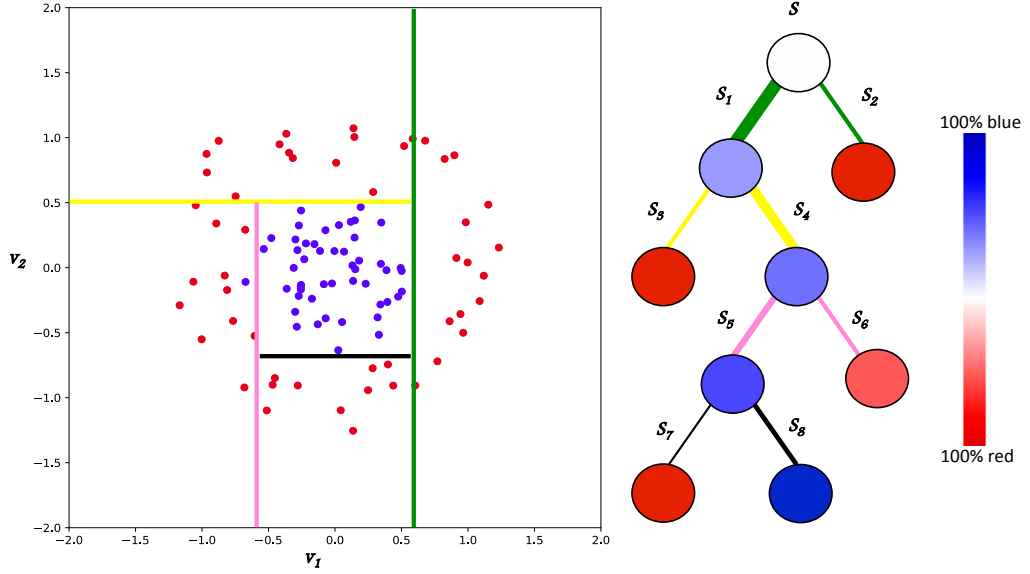


Figure 4.2: Example of a decision tree for a simple 2-D clustering problem that contained two classes of equal size. At each node (excluding leaf nodes) the data is split into subsets S_i and S_j by finding splitting planes that lower the heterogeneity of data arriving at the child nodes. Node colour indicates the distribution of class labels that have arrived at that node. Edge colours correspond to the splitting lines shown on the left. The number of datapoints existing along each edge is indicated by edge thickness.

SPLIT NODE OPTIMISATION When considering the optimisation of internal nodes it is useful to begin with a discussion of entropy. In the context of probability theory, entropy can be interpreted as a measure of uncertainty. Shannon entropy can be defined as:

$$H(S) = - \sum_{c \in C} p(c) \log(p(c)) \quad (4.3)$$

where c refers to the class labels of the training data. Information gain, in classification problems, can be formulated as:

$$I = H(S) - \sum_{i \in \{1,2\}} \frac{|S_i|}{|S|} H(S_i) \quad (4.4)$$

where S_i are subsets of S after splitting. The example given in Figure 4.3 demonstrates how splitting data can lead to a reduction in the combined entropy of the child nodes and

a corresponding information gain. In the first instance (a) we find clusters of equal size, such that a posterior distribution would estimate that a new data point would have equal probability ($p=0.25$) of belonging to any of the four classes. In other words, there is a very large level of uncertainty associated with the problem. In (b) we split the data horizontally, resulting in subsets that no longer have equally distributed points. The entropy in each subset, and the combined entropy of the child nodes is less than that of the parent node. In (c) we perform an alternative, vertical split, on the data. In this instance, data within each child node only belongs to two classes, compared to the three in (b). The entropy in each child node is lower than in (b) and the information gain is higher. Therefore, this vertical split would be a more optimal split when considering the data in this feature space. In this manner, information gain can be used to optimise the splitting planes and corresponding thresholds at each node of a decision tree.

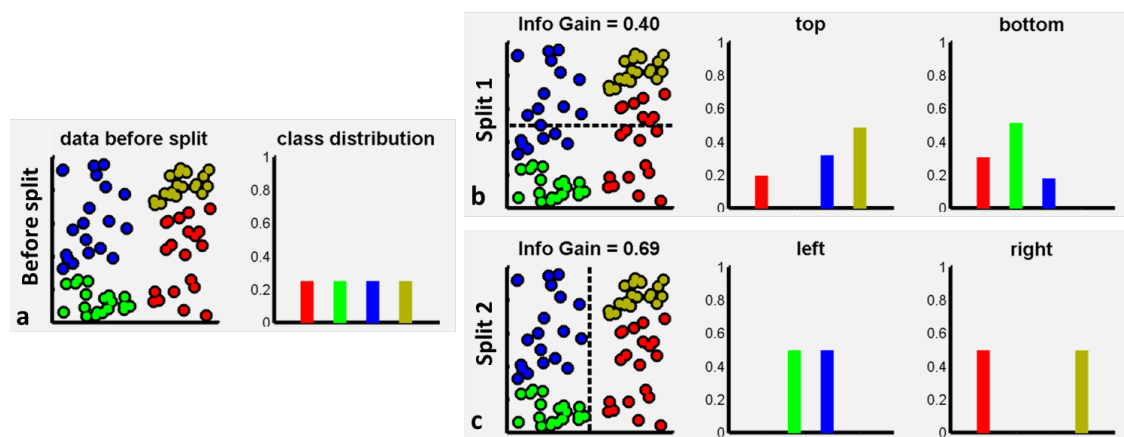


Figure 4.3: Illustration of information gain within a decision tree. (a) Input data containing 4 equally distributed classes, (b) class distribution after horizontal split, (c) class distribution after vertical split. Figure from Criminisi and Shotton (2013)

LEAF NODE PREDICTION In the case of classification, the leaf nodes are defined by the probability distribution of the training classes that arrived at that node. This distribution is used to predict the class assignment of any unseen data that terminates at the node during

the test phase.

Classification forests

In a forest, multiple trees are independently optimised. Forests are able to provide more accurate classification outcomes by introducing randomness into the training process. This helps to overcome the overfitting that is associated with single decision trees, thus creating a more generalised classifier. Below we discuss the ways in which randomness can be introduced, and other properties that affect the performance of a classification forest.

DATA BAGGING One method is to randomly sub-sample the training data that is sent to each tree. This is referred to as "bagging". This approach can be very efficient however it is not always guaranteed to find the surface that best separates the data.

FEATURE SELECTION Another option is to randomly sub-sample the features that are optimised at each node. If the entire feature vector were used at each node then all trees in the forest would be identical and the predictive outcome of test data would depend heavily on how well it is described by the training data. This limits the accuracy of the classifier for most applications. In many implementations the number of features that are selected at the nodes is the square root of the full dimensionality of the training data. Another advantage of feature sub-sampling is that the efficiency of each node optimisation is improved. This is a consequence of reducing a very high dimensional problem into a much lower dimensional one for which the maximal-margin can be found more accurately using less calculations.

FOREST SIZE Increasing the number of trees improves the accuracy of classification results. Figure 4.4 demonstrates how a single tree provides a rigid distribution of class labels in the parameter space. This may not extend well to the test data. Increasing the number of trees smooths the posterior distribution by introducing uncertainty in areas of the parameter space where the training data was sparse. This greatly reduces overfitting when compared to a single tree and increases the confidence we can place in the classification outcome.

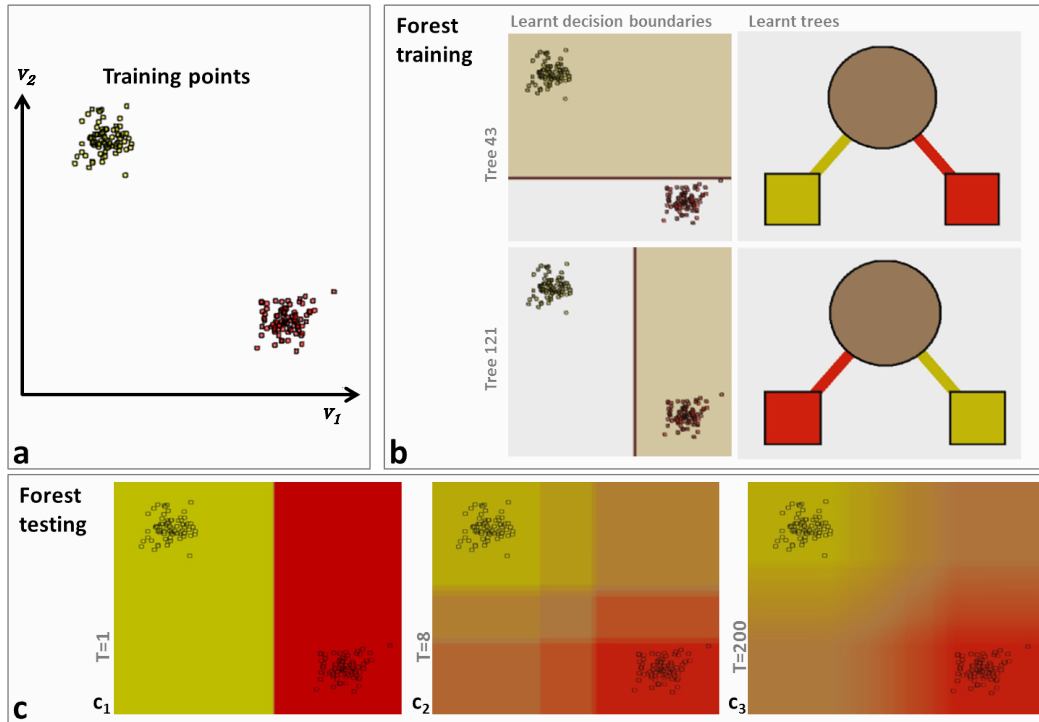


Figure 4.4: Example of how forest size effects prediction confidence and overfitting (Criminisi and Shotton 2013). (a) Training data. (b) example training trees with fixed depth. Each tree finds a slightly different partition line. (c) Posterior distributions for different forest sizes.

TREE DEPTH Varying the maximal depth of the trees can also increase prediction confidence. This is particularly true for noisy training data that is not linearly separable. In such cases, finding the optimal separation surface between classes will require several partitions. However, using trees that are too deep can result in overfitting. The complexity of the classification task should be considered when optimising tree depth.

Several classification experiments were performed in chapters 6, 7, A of this Thesis. We chose random forest classification over other approaches, such as SVM, because it scales well to large, high-dimensional datasets, can perform multilabel classification without adaptation, and has greater generalisability. This final point was particularly important to reduce overfitting in our cortical area classifications where the training labels cannot be considered to be as accurate as the ground truth labels that are commonly adopted in other machine vision tasks.

Part III

EXPERIMENTS

UNSUPERVISED GROUP AVERAGE PARCELLATION

In this chapter we aim to assess whether diffusion MRI provides area-specific information content that can be used to distinguish cortical areas. We applied and refined the framework initially developed by Nagy et al. (2013) to a large group of subjects using high quality diffusion datasets. By performing group-level analysis using surface-based averaging of the dMRI data, we improve between area contrast and obtain a hemisphere-wide map using high resolution, 3T data. As the dMRI signal is sensitive to several microstructural features, including axon diameter, neurite density, and dominant fibre direction, it potentially offers additional structural information beyond bulk myelination density. We substantiate this hypothesis by comparing the parcellation result to myelin density maps from the same set of subjects.

Some of this work has previously been presented in:

- Ganepola, T. et al. (2016a). "An unsupervised group average cortical parcellation using diffusion MRI to probe cytoarchitecture." In: Computational Diffusion MRI: MICCAI Workshop, Athens, Greece, October 2016, Springer, pp. 145-156.
- Ganepola, T. et al. (2015). "An unsupervised group average cortical parcellation using HARDI data." In: Proceedings of 21st Annual Meeting of the Organization of Human Brain Mapping, Honolulu, 2015, p. 221.

5.1 METHODS

5.1.1 *Human Connectome Project data*

The data in this Chapter and subsequent experimental chapters of this Thesis were acquired, in part, from the Human Connectome Project (HCP) 500-subject public data release. These datasets represent the state-of-the-art for *in vivo*, 3T diffusion imaging. Full descriptions of the pipeline and pre-processing steps are available in (Van Essen et al. 2013b; Sotiropoulos et al. 2013; Glasser et al. 2013; Uğurbil et al. 2013). In summary, data were collected on a 3T Siemens Skyra system with a custom, $G_{\max} = 100\text{mT/m}$, gradient coil. The use of this high performing gradient system and multi-band (MB) acceleration (MB factor = 3) allowed the spatial resolution of the data to be pushed to 1.25mm^3 without compromising greatly on SNR. This high spatial resolution lends itself to investigations of grey matter diffusion where partial volume effects and low anisotropy values are limiting factors. A total of 270 gradient directions were collected across three interleaved b-shells, $b=1000, 2000$, and 3000 s/mm^2 . An additional 18 $b=0\text{ s/mm}^2$ images were evenly interspersed throughout the acquisition. The b-values were modulated by varying G_d with matched diffusion times of $\Delta = 43.1\text{ms}$ and $\delta = 10.6\text{ms}$ and echo and repetition times of $TE=89\text{ms}$, $TR=5.5\text{s}$. Gradient directions were sampled uniquely in each shell to maximize q-space coverage (Caruyer et al. 2013).

PRE-PROCESSING Images arrived subsequent to pre-processing using *Eddy* (Andersson and Sotiropoulos 2016) which corrects DWI artefacts, including, susceptibility and eddy current distortions, whilst simultaneously co-registering the images. The HCP data suffer from subject specific spatial gradient non-linearities, resulting from irregular subject positioning in the customised hardware system (Sotiropoulos et al. 2013). B-values are estimated to deviate up to 15% from their nominal values across the brain. Voxel-specific corrections were provided to compensate for this, and were used accordingly throughout this Thesis to ensure accurate model fitting in all experimental chapters (see Appendix B.1).

COMPLIMENTARY DATA Diffusion datasets were accompanied by several other imaging modalities and behavioural information. Importantly, all subject datasets included high, 0.7mm^3 isotropic, resolution T1-weighted (T1w) and T2-weighted (T2w) structural volumes which can be used for accurate cortical surface reconstruction and myelin density estimates.

SUBJECTS In the parcellation experiment described below, diffusion, T1w and T2w data for 17 HCP subjects (10m, 7f aged 22-35) were randomly selected from a pool of 100-unrelated subjects.

5.1.2 *Surface reconstruction*

To utilise the high resolution of the available structural data we chose the HCP FreeSurfer pipeline over the standard recon-all pipeline to generate cortical surface reconstructions for each subject. This improved pipeline is deemed more accurate because it does not down-sample the T1w images to 1mm^3 , instead utilising the added precision of the higher resolution data. It also takes advantage of the additional information available in the T2w scans to reduce surface placement errors (Glasser et al. 2013). A comparison of the two methods is provided in Appendix B.2.

DATA SAMPLING the HARDI datasets of each subject were sampled onto their cortical surfaces. To achieve this, the average b0 image was registered to the T1w volume using an affine transformation matrix. The same transformation matrix was then applied to the DWIs. The signal intensity for each DWI image was nearest-neighbour sampled at the midpoint between the GM/WM boundary surface and the pial surface (i.e., cortical depth = 0.5). This depth was chosen to minimise the likelihood of either WM or CSF contamination (McNab et al. 2013; Nagy et al. 2013). Similarly, we opted for equidistant sampling to, again, minimise partial volume contamination from either the subarachnoid space or the white matter. Recent papers have argued in favour of equivolume sampling (Waehnert et al. 2014) to take into account the differential expansion/compression of cortical layers at gyral crowns/sulcal

floors (Bok 1929). We tested an approximation of the more anatomically realistic equivolume sampling, but as a consequence of the relatively low spatial resolution (compared to histology and 7T data), nearest-neighbour sampling and single depth analysis, we observed little difference between the two approaches and deemed equidistant sampling to be best practice for reducing systematic biases at crowns and fundi.

5.1.3 Feature space

A sixth order spherical harmonic series (SHS) was fit to the dMRI signal in order to characterise the apparent diffusion coefficient (ADC) of the cortical tissue. A SHS was fit separately in each b-shell at each surface vertex of the right hemisphere of each subject. The Nagy et al. (2013) paper proposed 27 cortical features (3.4), however their principal component analysis suggested only 9-10 degrees of freedom in the feature set. Therefore, we performed several ad-hoc parcellation tests (see Appendix B.3) to reduce the feature set to a subset containing only 5 features per b-shell. The chosen features, listed below, characterise the ADC in relation to the local surface normal, and therefore describe the GM tissue irrespective of the orientation differences that result from cortical folding.

1. The value of the ADC along the surface normal.
2. The mean of the ADC in the plane perpendicular to the surface normal, i.e. parallel to the cortical sheet.
- 3-5. The $k=2,3$ and 4 moments, respectively, of the ADC in the plane perpendicular to the surface normal. These features correspond to the variance, skew and kurtosis of the ADC.

The group average of each of the features was computed in turn using sulcus-based surface averaging (Fischl et al. 1999b). With this method, any given fsaverage vertex will

combine data from individual subject vertices that have surface normals in different directions. This makes it possible to detect local-surface-geometry-dependent diffusion signatures of cortical areas even though their local normal directions might differ from subject to subject. This information would be compromised if the diffusion data were to be directly averaged in 3D (folded) space. The transformation between each subject's cortical surface and the target brain space was applied to each of the diffusion features in turn. The mean across all subjects of each feature was then calculated at each vertex of the *fsaverage* surface. Finally, the averaged features were recombined into a $[n \times 15]$ group average feature space, where n is the number of vertices in the *fsaverage* surface tessellation. An alternative averaging approach in which the mode class was selected after clustering each subject was also implemented without success (Appendix B.5).

MYELIN MAPS We applied the same middle-depth, equivolume surface sampling process to the 0.7mm^3 T1w/T2w data that was also provided in the HCP datasets. As in the diffusion data we used surface-based averaging to create a group average map for the same set of subjects.

5.1.4 Classification

The unsupervised clustering algorithm, k-means, was implemented to parcellate the group average feature set. Several values of k were tested on a trial and error basis, starting with $k=40$, approximately the number of Brodmann areas. At lower values, the parcellation produced large smooth clusters, and did not provide additional structural information to the myelin density map (Appendix B.6). Results are shown for a value of $k=150$. At this value the parcellation displayed many more area-like clusters than for lower values, whereas, increasing beyond this value did not provide additional information in initial qualitative assessments. Furthermore, $k=150$ was the value after which decreases in the sum of total distances for the clustering solution began to plateau, and the total runtime began to increase rapidly (Figure. 5.1).

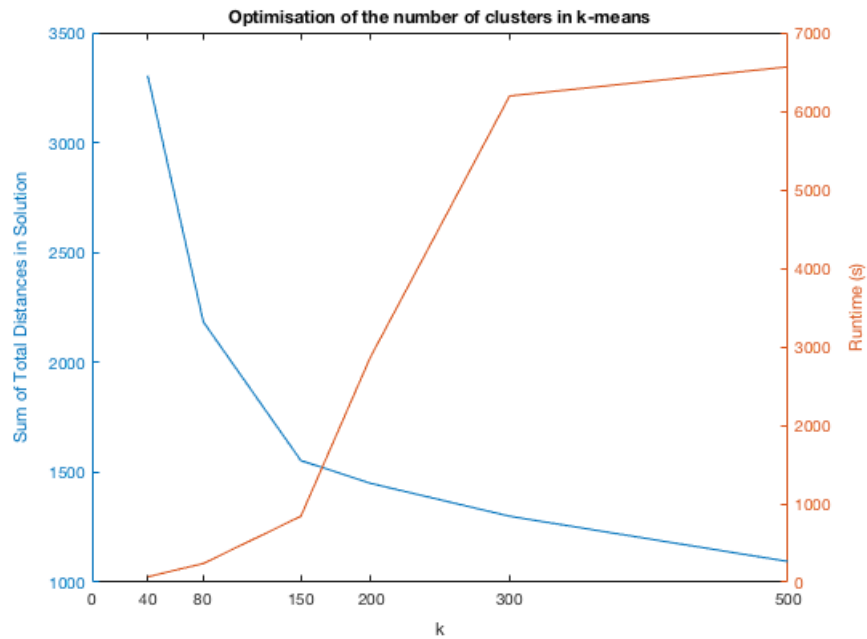


Figure 5.1: The optimisation of k , in k -means classification, using the trade-off between the best sum of total distances for each clustering solution (blue) and the total runtime of the algorithm in seconds (orange).

Since the resulting clusters are numbered arbitrarily by the algorithm, we included an additional ordering step whereby clusters were reordered according to the similarity of their cluster centres, starting with the pairing which had the highest affinity. Here, similarity was defined as the Euclidean distance between the mean feature vector of each cluster. This additional ordering stage acts to smooth the results when viewed on the surface, compared to a completely randomised cluster order by painting similar clusters in similar colours.

The resulting group average cortical parcellation was qualitatively compared to the group average myelin map, estimated from the T1w/T2w ratio of the same set of HCP subjects.

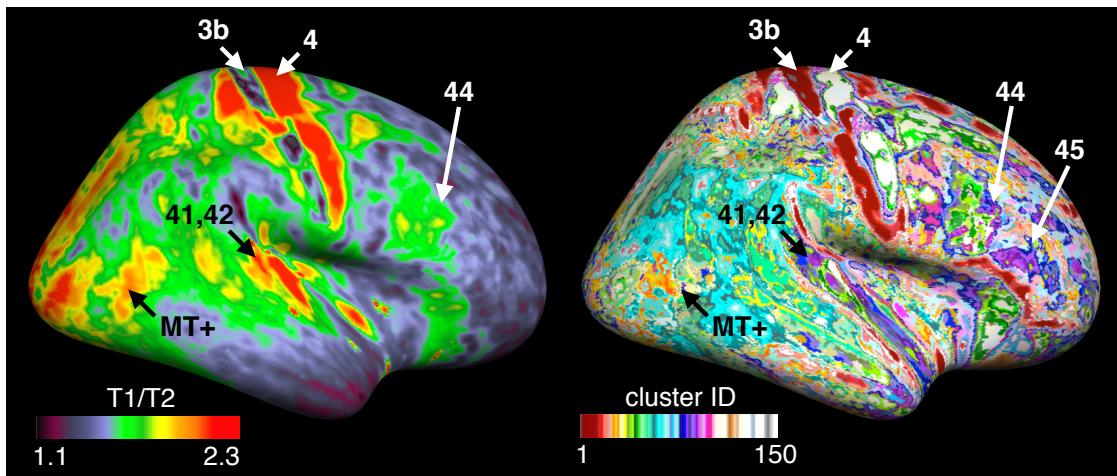


Figure 5.2: The lateral view of the group average diffusion-based parcellation (right) and the group average myelin map (left). Several areas of interest that are discussed in the results have been labelled. The clusters in the parcellation result are ordered by similarity of their cluster centres, such that similar clusters are painted in alike colours.

5.2 RESULTS

5.2.1 *Central sulcus*

Figure 5.2 shows the lateral view of the group average diffusion MRI based parcellation (right) alongside the group average myelin map (T1w/T2w) for the same set of subjects (left) on the inflated surface. The most salient feature of the parcellation is the emergence of distinct and spatially coherent clusters along the anterior (white cluster) and posterior (red cluster) banks of the central sulcus. These have been provisionally labelled 3b and 4 due to their consistency with the location and extent of Brodmann areas 3b and 4. These cortical areas correspond to part of primary somatosensory cortex (S-I) and to primary motor cortex (M-I), respectively. Comparison with the myelin map (left) indicates that the white cluster ID correlates with an area of high myelination, whereas, the red cluster correlates with a drop in myelination. Figure 5.3 confirms that both the myelin and diffusion measurements discriminate these regions. However, the independently derived diffusion based parcellation produces a more spatially coherent area along the posterior bank of the central sulcus, when compared to the variability of the myelin map in Figure 5.2.

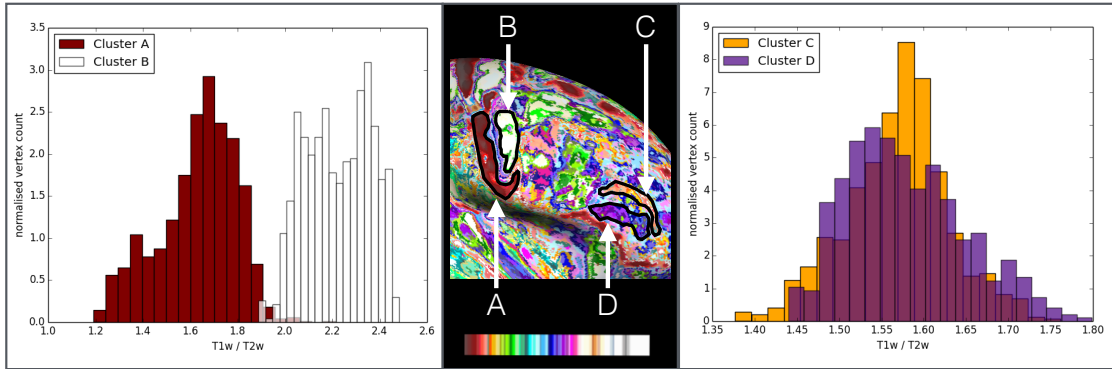


Figure 5.3: The distributions of myelin measurements corresponding to regions of interest from the dMRI parcellation result. The histograms on the left show T1w/T2w distribution from regions A and B within the central sulcus and are distinct in both modalities. The histograms on the right correspond to regions C and D which appear distinct in the parcellation but have closely overlapping myelin distributions. The outlines of the regions of interest from which the histogram data were extracted are shown in the centre image.

5.2.2 Broca's region

Areas 44 and 45 are collectively referred to as Broca's region, which has long been implicated in the production and recognition of speech. We note that both the parcellation and myelin map exhibit a distinct region that is consistent with the location of area 44. This area emerged as a coherent patch despite it not having a consistent relation to macroscopic landmarks in the majority of subjects (Amunts et al. 1999).

In addition, anterior to area 44, we note the presence of a spatially coherent purple/blue cluster, provisionally labelled area 45. It seems that area 45 has no counterpart in the myelin map, supporting the notion that diffusion based cortical imaging may be able to provide additional information to myelin mapping, particularly in areas of lower myelination such as the premotor and prefrontal cortex. This notion is further supported in Figure 5.3 where it is clear that the distribution of myelin values between area 45 and the adjacent ROI are very similar. In contrast, these two regions could be differentiated in the diffusion based feature set. The frontal lobe of the diffusion-based parcellation appears more like a patchwork of distinct clusters, whereas, the myelin map in this region appears more homo-

geneous.

5.2.3 Auditory areas

On the temporal lobe we note that both the myelin map and diffusion-based parcellation exhibited distinct patches that roughly coincide with primary auditory core and belt areas (BA 41 and 42). This suggests that some structural information in the dataset is maintained despite Heschl's gyrus exhibiting a markedly variable folding pattern across subjects (Leonard et al. 1998).

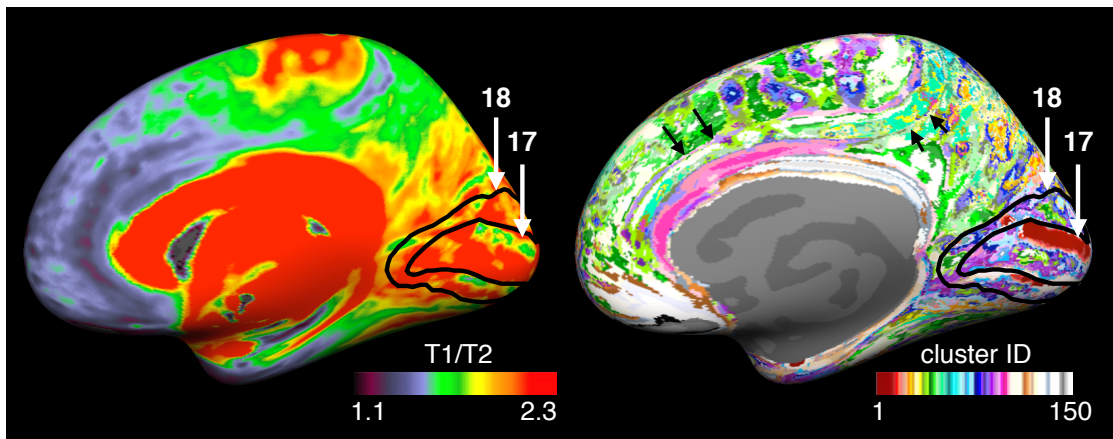


Figure 5.4: The medial view of the diffusion-based parcellation (right) and the group average myelin map (left). Again, areas of interest that are discussed in the results have been highlighted with labels, contours and arrows and similar clusters are painted in alike colours in the parcellation result.

5.2.4 Occipital areas

On the right hand side of Figure 5.4 the posterior occipital lobe contained prominent purple, blue, and red clusters, giving it a distinct appearance compared to much of the rest of the medial cortical sheet, which was assigned predominantly to green and white clusters. This generally correlates with the higher myelination of this region, seen in the T1w/T2w data at the left of Fig. 3; however it is worth noting that the region of apparent heavy myelination in

V2 just below the tip of the "18" arrowhead projected further in the superior direction than the purple region in the diffusion-based result.

The black contours outline the inner and outer extents of area 18 from the FreeSurfer probabilistic atlas, i.e the secondary visual area V2. The inner boundary of this contour corresponds to the neighbouring primary visual cortex, V1 (area 17), within the calcarine sulcus. Despite V1 possessing a prominent tangential band in layer 4B that is lacking in extrastriate area V2 (Aggarwal et al. 2015; Amunts et al. 2000), we did not observe distinct coherent clusters corresponding to the full extent of these two areas in either the T1/T2 data or the diffusion data. Instead the most salient feature of this region was the red cluster, which is located near the upper vertical meridian of V1.

Finally, returning to the lateral surface (see Figure 5.2) in the middle of the myelin-dense region of MT+ it is possible to distinguish a border between a posterior orange cluster and an anterior white/tan cluster. A study examining the relation between quantitative T1 and retinotopy (Serenó et al. 2013) surprisingly showed that the heavily myelinated oval in the lateral occipital cortex does not directly correspond to MT; instead, MT proper only accounts for the posterior part of that oval. The anterior part may correspond to FST, which represents parts of the visual field already mapped in MT, and which responds to the ipsilateral visual field unlike MT. Once again, this suggests that diffusion data may help distinguish regions not easily separated by using myelin density alone.

5.2.5 Gyrification

The unlabelled black arrows on the right of Figure 5.4 indicate curvature-like features in the parcellation. These lines follow the fundus and crown of the cingulate sulcus and gyrus respectively.

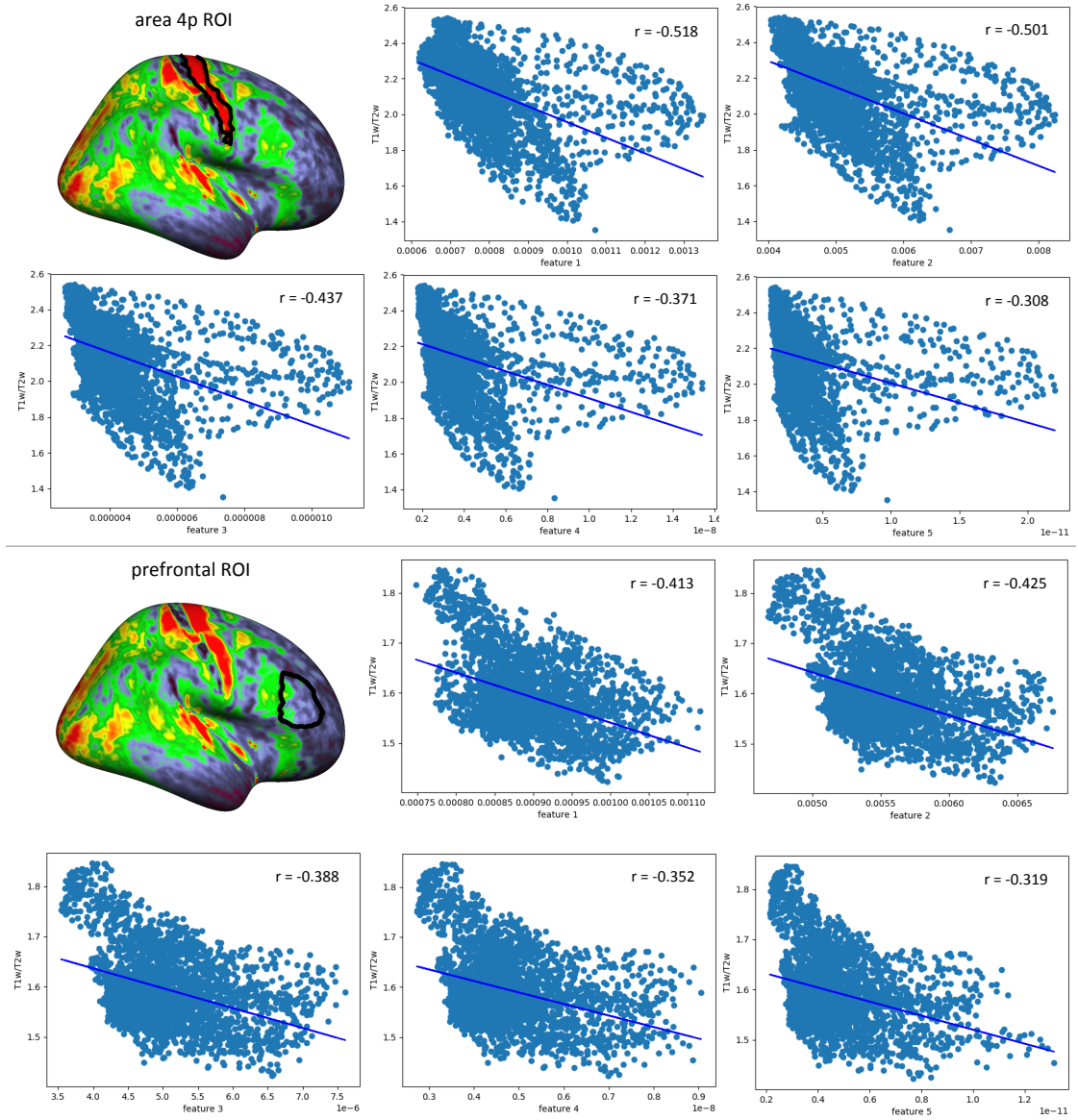


Figure 5.5: Plots of diffusion MRI features against myelin density for regions of high and low myelination. Diffusion features are the 5 features detailed in methods taken from the $b=1000$ s/mm^2 shell. The top two rows take data from heavily myelinated area 4p within the central sulcus; the bottom two rows take data from an arbitrary region of low myelination in the prefrontal cortex. Both axes display arbitrary units and the pearson correlation coefficient between each feature and the myelin density is given as r .

5.3 DISCUSSION

We employed group averaging and several refinements to a cortical parcellation pipeline including improved surface reconstruction, and a refined feature set. The resulting parcellation

exhibits several clusters which correspond to well known Brodmann areas. Importantly, we found that regionally specific information in the diffusion-based feature set was not limited to areas of high myelin density. Diffusion MRI signal attenuation is expected to correlated with myelin content because myelinated axons restrict and trap water, creating anisotropic diffusion patterns; however other factors such as membrane density have been suggested to have a larger contribution to the dMRI signal than myelin density (Beaulieu 2002). Our results, as in (Calamante et al. 2017), support the hypothesis that diffusion data is not solely modulated by myelin density but is also characterised by other aspects of cortical cyto and myeloarchitecture. Infact our dMRI features demonstrate only moderate correlation with myelination, even in areas of high myelin density such as area 4p (see Figure 5.5). The results strongly indicate that diffusion MRI provides microstructural contrast in the cortex which is complementary to the traditionally used myelin mapping methods.

CENTRAL SULCUS: We found particularly close agreement between the group average parcellation and the Brodmann areas within the central sulcus. This may arise from several factors, for example, the boundary between S-I and M-I represents one of the most prominent transitions within the cortex. The input layers of S-I possess many small cell bodies giving it a granular appearance. In contrast, long cortico-spinal projections in M-I result in large pyramidal cell bodies known (in the foot representation) as Betz Cells, giving an agranular appearance. S-I also exhibits highly myelinated tangential bands of Baillarger, which are less prominent in M-I. Furthermore, *in vivo* studies support the hypothesis that differences in the laminar composition between these two regions are manifested in dMRI signal (McNab et al. 2013). Another factor which may have improved the detection of these areas is that they demonstrate relatively low intersubject variability. S-I occupies the posterior bank of the central sulcus and extends back into the postcentral gyrus and M-I occupies anterior bank of the central sulcus, extending forwards into the precentral gyrus, with their transition consistently located at the fundus of the central sulcus (White et al. 1997; Amunts et al. 2000) near the location of area 3a, which receives predominant input from muscle receptors. This is consistent with the location and extent of the red and white clusters in the parcellation. Therefore, it is likely that architectural differences between these regions, as

characterised by their feature vectors, are reinforced by averaging across multiple subjects.

PRIMARY VISUAL CORTEX: The parcellation did not exhibit a cluster that corresponds to the full extent of the primary visual cortex, despite the prominent architectural properties which distinguish this area from its surroundings. Instead we found that the boundary between V1 and V2 was characterised uniquely by the dMRI feature set. The myelin map also showed a significant decrease in myelination in the same location which is not consistent with the underlying anatomy. This suggests the presence of a systematic surface placement error that may have resulted in CSF partial voluming in both data sets, causing a region near the upper and lower vertical meridian border between V1 and V2 to stand out.

GYRIFICATION: Clusters that closely matched cortical folding patterns were also observed. The emergence of these macroscopic landmarks could be associated with sampling errors at areas of high curvature, where partial volume is more prevalent. Alternatively, it may reflect a relationship between gyrification and diffusion anisotropy, as suggested by other groups – e.g., (Kleinnijenhuis et al. 2015). Deeper cortical layers appear to thin in sulci, which has been suggested to be a way of maintaining equal local volume by folding-induced tangential stretching (Bok 1929; Waehnert et al. 2014). By contrast, upper cortical layers appear to puff out and become more myelinated on gyri. These systematic folding-correlated effects may give rise to detectable differences in grey matter diffusion patterns. The initial detection of a correlation between gyrification and T1 was similarly initially dismissed as a depth-sampling artifact, but then subsequently suggested to be partly due to real myelination differences between sulcal and gyral cortex. An additional complication is that partial volume errors may be detecting systematic differences in fiber direction near the grey/white matter border; for example, the dominant diffusion direction is expected to be highly radial in areas of high curvature, such as gyral crowns, and more tangential in along the banks of gyri due to the angle at which u-fibres project into the cortex (McNab et al. 2013; Kleinnijenhuis et al. 2015).

Limitations and future work

In the above, we implemented a global, unconstrained segmentation algorithm, k-means. The aim was to determine what degree area specific information was present in the dMRI signal of the high resolution HCP data sets; as such, unsupervised classification was deemed the most appropriate for this exploratory study. In reality, segmenting the cortex is not a global problem, but rather a set of spatially constrained classification tasks, i.e., each area needs only be recognised against its neighbouring areas. Moreover, it is plausible that many spatially distant areas of the cortex share similar architectonic properties, such as lightly myelinated, higher-order non-primate areas; an ideal parcellation pipeline would identify these as separate clusters despite their microstructural similarities. The next obvious step for improving the classification result would be to introduce spatial constraints to the system. This could be achieved implicitly by including either spatial information or neighbourhood information in the feature vector at each vertex (see Figure 5.6 in the supplementary material). Alternatively, constraints could be explicitly enforced in the segmentation algorithm. An example of this is shown in the supplementary material, Figure 5.7, where the same group average diffusion feature set was parcellated using a spatially constrained hierarchical clustering algorithm (Blumensath et al. 2013). Such methods use a region growing approach to provide noise free cluster definitions. Another approach would be to use supervised learning methods with the aid of a set of training labels. This is explored in Chapter 6.

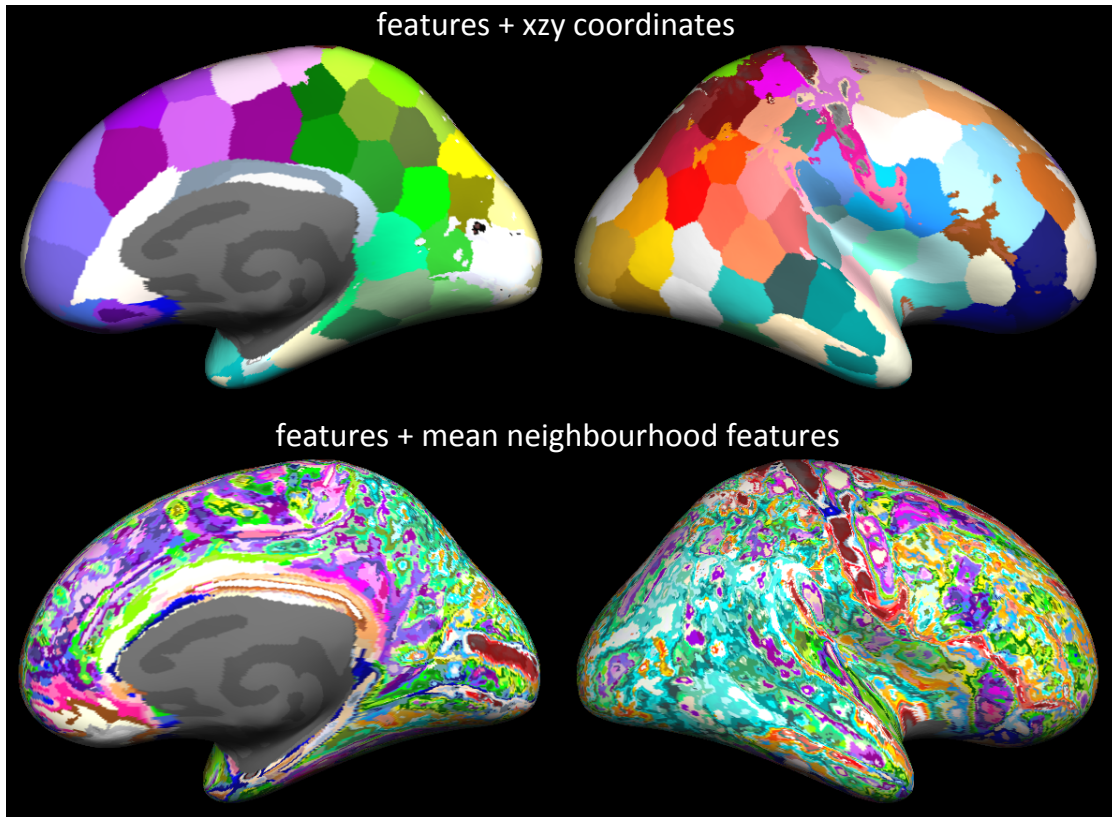


Figure 5.6: Unsupervised cortical parcellations using dMRI features combined with spatial or neighbourhood information. Top row shows k-means results using the 5 diffusion-based features plus the x, y, z surface coordinates at each vertex. The variance in spatial coordinates dominate the feature vector, creating a voronoi effect for much of the cortex. However, a cluster which follows the posterior bank of the central sulcus is clearly visible, suggesting that the dMRI information in this region is particularly distinct. The bottom row shows the k-means result after combining the 5 dMRI features at each vertex with the mean feature vector of the immediately neighbouring vertices. The results look similar to those presented in Figures 5.2 and 5.4. It is possible that a larger set of neighbours could produce a more spatially coherent clustering as many neighbouring vertices are likely to sample from the same image voxel, therefore the neighbourhood information in the nearest neighbours is likely to be somewhat redundant.

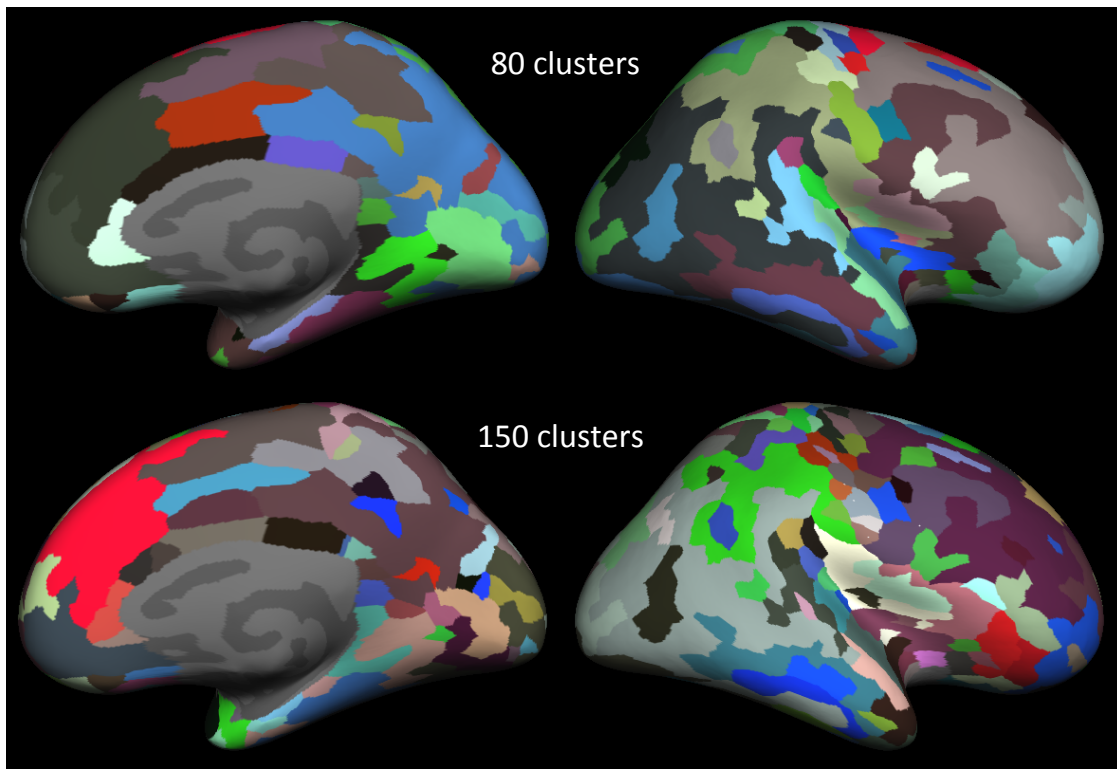


Figure 5.7: Cortical parcellation results using a spatially constrained hierarchical clustering algorithm (Blumensath et al. 2013). The top row shows the result for 80 clusters and the bottom for 150 clusters. This region growing approach, adapted from a resting-state fMRI pipeline, produces spatially coherent clusters across the cortex. The resultant clusters do not seem to adhere closely to the known architectonic areas of the cortex. They suggest that the diffusion data is sensitive to fine grained within area fluctuations of densely myelinated, primary areas. For example we note several sub-divisions along the expected locations of M1 and S1 which may reflect the varying microstructure within these areas, for example the hand, foot and face subdivisions. Similar sets of small clusters are observed in the insular cortex and visual cortex.

FEATURE SETS AND SUPERVISED PARCELLATION

Although a large variety of diffusion imaging methods have been developed in white matter studies, the efficacy of these methods have not been compared in grey matter. In the *in vivo* diffusion-based grey matter literature, investigators have either adapted the simple diffusion tensor for surface-based analysis (Anwander et al. 2010; McNab et al. 2013; Kleinnijenhuis et al. 2015) or opted for higher-order decomposition methods (Nagy et al. 2013; Calamante et al. 2017). The higher-order approaches may better characterize the subtle changes in diffusion that are associated with microscopic architectonic properties of cortical areas. In this chapter we aimed to test this hypothesis.

Several different dMRI methods were compared in a cortical parcellation pipeline that employed random forest classification (RFC) (Breiman 2001). Initially we focussed on the sensorimotor areas (cf. (McNab et al. 2013)) but then extended the analysis to the entire cortical surface by utilising the HCP multi-modal parcellation labels (Glasser et al. 2016). In the previous chapter we saw that our surface-based diffusion pipeline produced area-like clusters in both areas of high and low myelin density. Here, we expand on this finding by developing a neighbourhood-based supervised learning approach, which allowed us to test dMRI more robustly across the cortical sheet to quantify the degree of area specific contrast that is available within the 3T HCP data. We performed both individual and group-level analyses.

Some of the work in this chapter has been previously published in:

Ganepola, T. et al. (2018). "Using diffusion MRI to discriminate areas of cortical grey matter." In: *Neuroimage*, (In Press).

Ganepola, T. et al. (2016b). "Evaluation of diffusion MRI based feature sets for the classification of primary motor and somatosensory cortical areas." In: Proceedings of the 24th Annual Meeting of International Society for Magnetic Resonance Imaging, Singapore, 2016.

6.1 METHODS

6.1.1 *Data and pre-processing:*

All data sets were obtained from Human Connectome Project (HCP) Q3 500-Subjects release (Van Essen et al. 2013b; Sotiropoulos et al. 2013; Glasser et al. 2013; Uğurbil et al. 2013). Refer to Section 5.1.1 for in depth descriptions of these data. In summary, the 3T diffusion datasets comprised of 270 gradient direction spread across three b-shells, $b=1000$, 2000 and 3000 s/mm^2 , and acquired at an isotropic spatial resolution of 1.25mm^3

As in Chapter 5, we utilized the improved HCP Freesurfer Pipeline for surface reconstruction. Data were processed for total of 40 subjects. The same surface-based pipeline (5.1.2) as in the previous Chapter was applied to sample the diffusion data at the midpoint between the GM/WM boundary surface and the pial surface.

6.1.2 *Feature sets*

Four different diffusion MRI analysis techniques were used to generate six feature sets that were tested in the classification experiments below. The feature sets were chosen to address: (a) whether projecting the HARDI signal characteristics into the local frame of reference (indexed by the local surface normal) is beneficial when differentiating cortical areas, and (b) whether an increase in model complexity is beneficial when differentiating cortical areas. The four techniques are described below, Table 6.1 provides the names and details of all the feature sets that were generated from the different techniques.

1. The diffusion tensor (DT) (Basser et al. 1994; Basser and Pierpaoli 2011) provides two scalar metrics, mean diffusivity (MD) and fractional anisotropy (FA), as well as directional information stored in its 3 eigenvectors. One simple way to incorporate surface information is to measure the dot product of the primary eigenvector and the local surface normal (the radially index, RI, (McNab et al. 2013)), to generate another

scalar metric that indicates the extent to which diffusion is radial (e.g., along apical dendrites of pyramidal cells).

2. The neurite orientation, dispersion and density imaging (NODDI) (Zhang et al. 2012) is a popular WM model which aims to increase specificity to microstructural properties. It disentangles FA into the possibly more anatomically relevant neurite density index (NDI), and orientation dispersion index (ODI). In addition, NODDI provides the isotropic volume fraction (Viso), representing free-water content.
3. In the third approach, we use the spherical harmonic model of the previous chapter to more finely characterise the shape of the local diffusion surface. A total of 9 features were generated for each b-shell. Features 1-4 were the $k=1,2,3,4$ th moments of the ADC and fully rotationally invariant. The 5th feature was the mean of the ADC in the direction of the surface normal, and the remaining 4 features were the $k=1,2,3,4$ th moments of the ADC in the plane that is parallel to the cortical surface and are invariant in the local tangent plane.
4. A final approach used a 4th order tensor representation of the ADC and all its informationally (or functionally) complete and irreducible invariants. These invariants fully describe the geometric characteristics of the ADC up to any orientation or pose in 3D. Higher order tensors were introduced as an alternative (and bijective) mathematical basis to spherical harmonics in Özarslan and Mareci (2003). The tensors were estimated using the ternary quartic (TQ) framework to ensure positive ADC (Barmpoutis et al. 2009; Barmpoutis and Vemuri 2010; Ghosh et al. 2012), while the invariants were computed following the method proposed in Papadopoulos et al. (2014). The invariants were found by progressively projecting the TQ coefficients via an orthogonal transform and a rotation transform to a canonical representation with 12 degrees of freedom. Unlike the spherical harmonic approach, this 4-tensor method does not attempt to explicitly describe the radial and tangential diffusion processes within the

cortex; however, such information may be implicitly encoded in the 4-tensor.

The latter two approaches better characterise the precise shape of the HARDI signal due to their sensitivity to higher-order details. Because GM regions have much lower anisotropy than WM, which often consists of coherently organised WM fascicles, we hypothesised that the higher-order feature sets will be better suited to capturing the varying dendritic arborization patterns and axonal mesh patterns that are known to distinguish different cortical areas. Furthermore, much of the information content captured in the lower order features should also be encoded in the higher order feature sets alongside the additional degrees of freedom that they offer.

Feature Sets	Description
DT3	[MD, FA, RI], calculated after fitting the diffusion tensor to the $b=1000\text{s/mm}^2$ data.
DT9	[MD, FA, RI] $\times 3$, after fitting the DT to each b-shell separately and concatenating the 3 metrics from each shell
DT6	[MD, FA] $\times 3$, same as DT9 with the surface specific component (RI) omitted.
ND3	[NDI, ODI, V_{iso}], after fitting the NODDI model to the full multi-shell dataset.
SH27	9 features per b-shell calculated from the SH series, which characterise properties of the ADC that are either fully rotationally invariant or invariant in the plane perpendicular to the local surface normal.
4T36	12 features per b-shell, calculated from the 4-tensor, which together provide a functionally complete set of invariants.

Table 6.1: The names and descriptions of each of the diffusion-based cortical feature sets tested in the classification experiments. The prefix signifies which method was used to estimate the ADC and the number at the end of each name signifies the total dimensionality (length) of each feature set.

6.1.3 Classification Experiments

BINARY CLASSIFICATION OF M1 AND S1 The S1 and M1 areas are very distinct from each other and consistently located across subjects. Therefore, classification between these two regions was chosen as a robust initial test-bed. Training labels were defined for each subject using the HCP multi-modal parcellation (HCP-MMP) atlas (Glasser et al. 2016),

see Figure 6.1A. Specifically, the labels 3b, and 4 were registered from the fsaverage surface to each subject's surface tessellation (Fischl et al. 1999b).

Random forest classification (RFC) (Breiman 2001) was implemented in sk-learn (forest size=20, tree depth=7, other parameters at default values) to distinguish the two areas using a pool of 20 subjects. Data under the two regions were extracted for each feature set in each subject. The classifier was trained on a set of subjects for several training group sizes (TS), to help determine the lower limits on group size for future experiments. TS ranged from 1-19 subjects, and the classifier was then tested on an unseen subject. The same training group was maintained until all feature sets were tested. A leave-one-out approach was implemented to ensure that all of the available subjects were tested in turn for each TS.

We hypothesised that the low-order features would dominate the classification because they have previously been shown to be very distinct between S1 and M1 (Anwander et al. 2010; McNab et al. 2009).

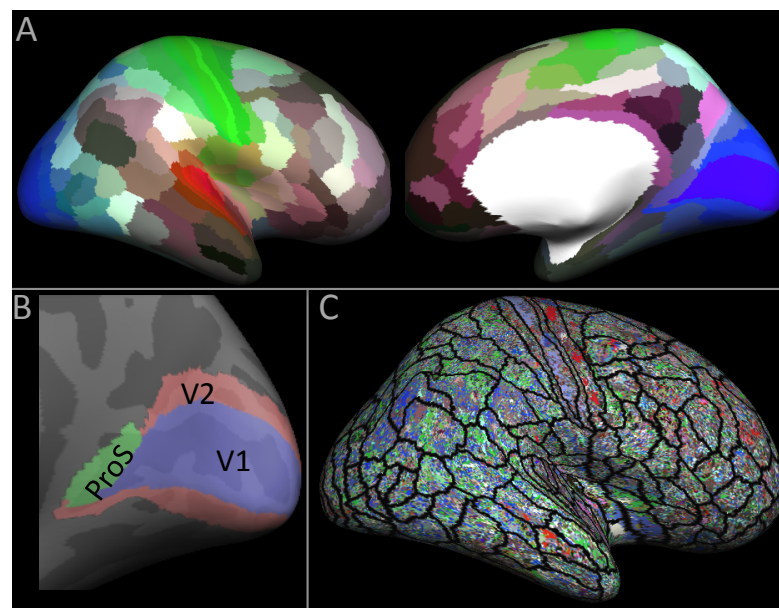


Figure 6.1: (A) The classification training labels generated from the Human Connectome Project multi-modal parcellation. (B) An example of a neighbourhood of areas, in this case, for the classification of V1. (C) An example of a classification result where instead of the neighbourhood approach, a 180 area multiclass classification is attempted. The result was generated using the DT9 feature set, and shows very little structure.

GROUP AVERAGE WHOLE HEMISPHERE PARCELLATION A hemisphere-wide parcellation pipeline was developed to test the DT6, DT9 and SH27 feature sets across a broader range of cortical areas. These three feature sets were prioritised because they encompass the differences we intended to test, i.e., DT9 includes surface specific features where DT6 does not, and SH27 is a higher-order method compared to DT9 and DT6. We also employed population averaging for these classification tests to increase the contrast-to-noise ratio between cortical areas.

Of the 40 HCP subjects, 30 subjects were selected as the training pool and the remaining 10 were assigned as the test pool. 6 of the 30 training subjects were randomly selected at a time, to generate a total of 20 group average training sets. As in Chapter 5, surface-based averaging (Fischl et al. 1999b) was performed on each dimension (column) of the feature sets. We chose a reduced group size for these experiments as initial testing suggested that 6 subjects provides sufficient between area contrast (B.4). The same process was repeated to create test datasets from the test subject pool, also consisting of data from 6 subjects each.

The 180 areas of the HCP-MMP were utilised as a set of prospective hemisphere-wide, training labels. Given a test average, the cortical area marked by each of the training labels was tested in turn, using a multi-label RFC, against its neighbouring cortical areas. For example, when trying to predict the class of the data marked by the V1 label, the classifier was trained on data taken from 3 labels in the training averages: ProS (prostriata), V1 and V2 (Figure 6.1B). This neighbourhood approach mimicked traditional parcellation techniques that define areas based on local transitions in laminar appearance, and also reduced the number of classes within a single test to a relevant set. Figure 6.1C demonstrates the ineffectiveness of implementing a global 180-label classification experiment.

SINGLE SUBJECT WHOLE HEMISPHERE PARCELLATION A major goal of cortical parcellation applications is to generate subject-specific maps. Therefore, single subject, whole hemisphere parcellation was performed to assess whether between-area contrasts can be detected without population averaging.

The two methods presented above were combined to produce a whole hemisphere

parcellation on an individual subject. The neighbourhood multi-label RFC approach was used to classify the data marked by each of the 180 HCP-MMP areas for a single unseen subject. The training data was generated by concatenating the data for each of the neighbourhood labels from a group of training subjects, as in the binary classification test. In this instance the training group was reduced to 10 subjects in order to reduce runtime and memory requirements of the classifier. Note, when the fsaverage HCP-MMP labels are surface-registered to each subject's surface tessellation there is no guarantee that some neighbouring regions will share some overlapping vertices. During registration of the training labels, a small number of vertices (1-5%) were assigned to multiple classes. In such cases, the vertex was assigned to its mode class label across all of the classification tasks in which it appeared; in the absence of a mode class, the vertex assignment was selected randomly from its predicted classes.

6.1.4 *Searchlight Cluster Count*

A quantitative vertex-wise comparison method was developed to evaluate the quality of different full hemisphere parcellation results. This tool was used to bolster the qualitative analysis of the parcellation maps. Here, quality was defined as the local spatial coherence of the parcellation. Given two corresponding parcellation results, A and B, the number of unique cluster IDs within a 90-nearest neighbour surface searchlight surrounding each vertex were counted, and the resulting cluster counts were subtracted from each other ($A - B$). A positive value ($A < B$, orange in Figures 6.4, 6.7, C.2) denotes that A did better (i.e., local regions had fewer different cluster IDs) while a negative value ($A > B$, blue) indicates that B did better. The searchlight diameter was chosen to be somewhat smaller than the width of a typical cortical area.

6.2 RESULTS

6.2.1 Binary classification of M1 vs S1

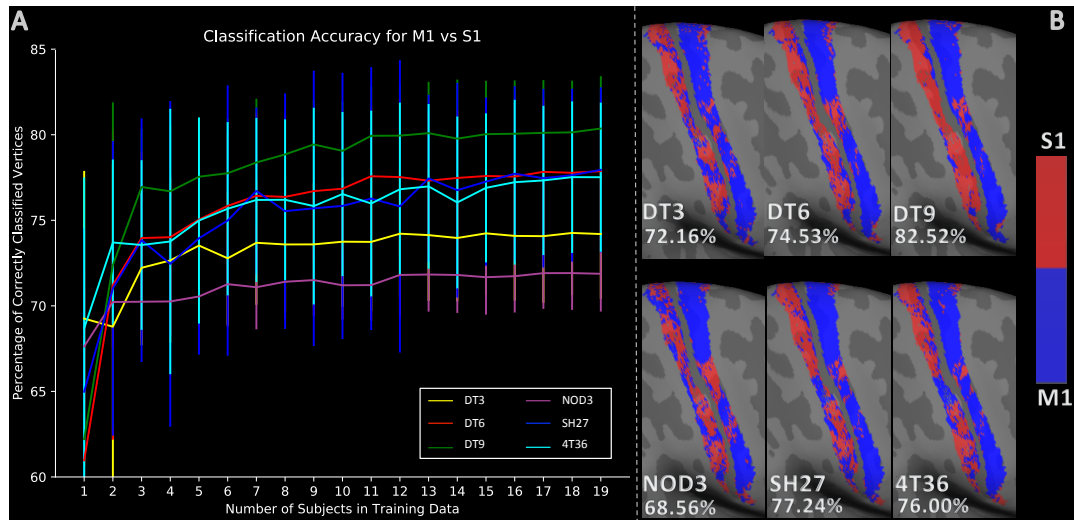


Figure 6.2: Results for binary classification between M1 and S1 using different feature sets. (A) The mean classification accuracy for each feature set, at each training group size (TS). Error bars are the standard deviation in classification accuracy across the repeated tests for each TS. (B) M1 vs S1 classification results for a typical subject are shown on the cortical surface for TS=19. Accuracy scores were defined as the percentage of correctly classified vertices. Red corresponds to the S1 class label and blue to the M1 class label.

The results for binary classification between S1 and M1 using different training group sizes are shown in Figure 6.2A. The feature sets DT6, DT9, SH27 and 4T36 demonstrate a similar trend with steep improvement in accuracy from TS=1 to TS=3 followed by more gradual, improvement up to TS=19. The DT3 and ND3 feature sets exhibit a more modest rate of improvement in accuracy between TS=1 and TS=3, and significantly lower plateaus. It is evident that using fewer than 3 training subjects does not provide a broad enough set of training examples to account for inter-subject variability; incorporating at least 10 subjects is beneficial.

The DT9 feature set performs best in the classification between M1 and S1 for all values of TS larger than 2, achieving a mean classification accuracy of **80.36%** at TS=19. SH27, 4T36, DT6, DT3, and ND3 gave mean accuracies of **77.94%**, **77.51%**, **77.87%**,

73.98%, and **71.87%** respectively, for the same TS. A Wilcoxon signed rank test between each set of results found the differences between DT9 and all other feature sets to be significant ($p < 0.02$), whereas the performances of SH27, 4T36 and DT6 were not significantly different from each other. ND3 and DT3 were also found to be significantly different from each other and all of the other feature sets ($p < 0.001$). Aside from the reasons given in the methods, we chose not to include DT3 and ND3 in subsequent experiments because they performed comparatively poorly here. 4T36 was also omitted because its performance was highly correlated with that of SH27.

In Figure 6.2B it is evident that DT9 provides the most spatially coherent result, particularly within area S1.

6.2.2 Group average whole hemisphere classifier

QUALITATIVE ASSESSMENT The lateral and medial views of the parcellation result for the DT6, DT9, and SH27 feature sets are displayed in Figure 6.3. In general, early sensory and motor areas showed a strong resemblance to the training labels, exhibiting spatially locally coherent clusters. Moving away from those easy-to-distinguish areas, the spatial coherence of the classification results was reduced, and a number of training areas contained a speckling of multiple cluster IDs. One can also observe several coherent clusters that contain several training regions (black brackets).

The overall trend comparing DT6, DT9, and SH27 (left to right) is an observable decrease in the granularity of the parcellation in some areas (e.g., area 3b). This suggests that inclusion of a surface normal component (DT9 & SH27 compared to DT6) and the use of higher order features (SH27 compared to DT6 & DT9) both provide useful additional information to the training and classification process.

Qualitative assessment of the lateral surface indicates that area 3b (part of S1) has the most distinct HARDI signal profile compared to its neighbouring areas, as all of the feature sets achieve a reasonable correspondence to the training label for this area. Some regions show markedly different classification outcomes between the different feature sets (dotted arrows). For example, the MT region is subdivided by SH27, but recovered more

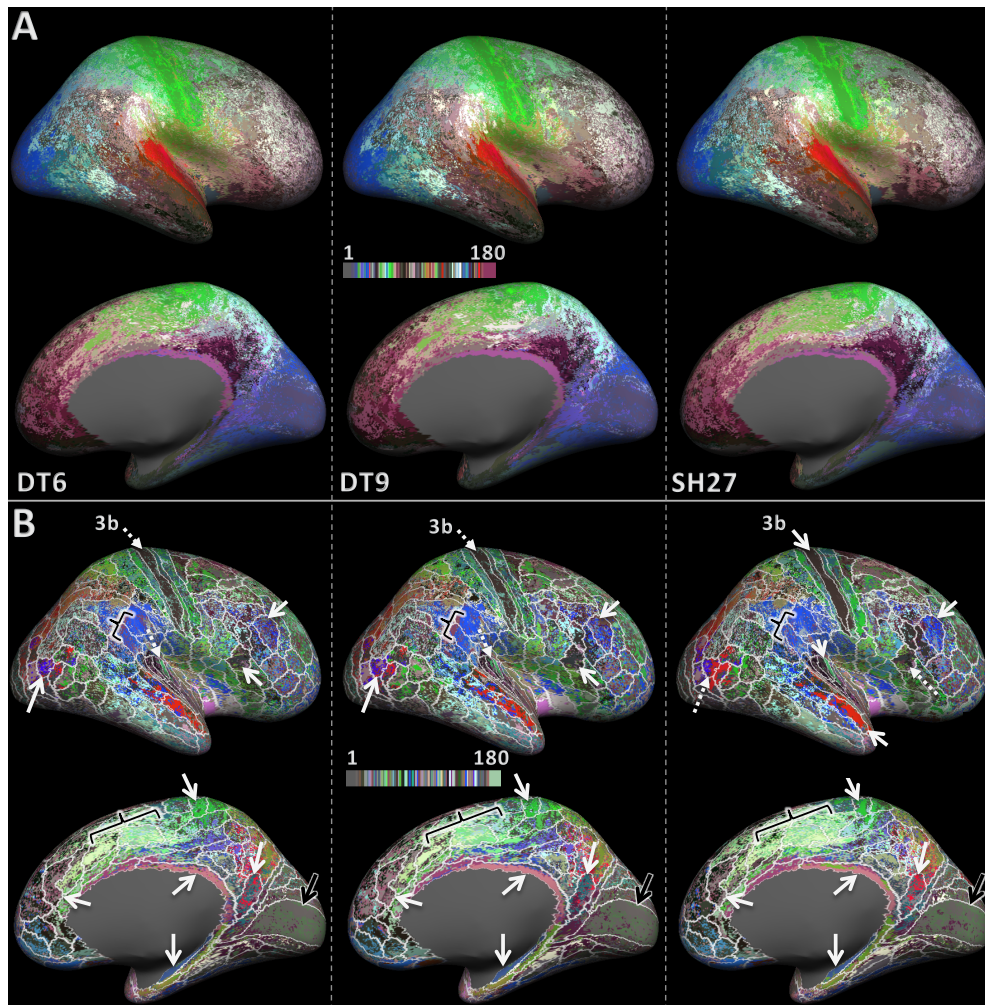


Figure 6.3: Maps of the group average whole hemisphere parcellation result for feature sets DT6, DT9, and SH27 (left to right). (A) Shows the original colour scheme from the HCP-MMP. (B) Shows the same results as A but with the colour scheme shuffled to achieve better contrast between neighbouring areas. In addition, the boundaries of the training areas are overlaid in white. The solid white arrows signify areas that have a large overlap with the training labels. The dotted white arrows indicate that an area is subdivided or not as well classified as it was for another feature set. The black arrows point to the V1 area that did not classify as well, despite its distinct architecture. The black brackets point out regions in which one cluster expands over several training labels.

completely by DT6 and DT9. The SH27 subdivision may reflect inter-subject variability in the location of MT proper (Bridge et al. 2014).

Inspection of the medial surface surprisingly reveals that none of the feature vec-

tors strongly differentiated the primary (V1) and secondary (V2) visual areas (black arrows). Areas close to the medial surface interface with the corpus callosum, e.g., retrosplenial complex and hippocampus, can be accurately classified by all feature sets and are known to be architectonically distinct from most other medial and lateral cortical areas.

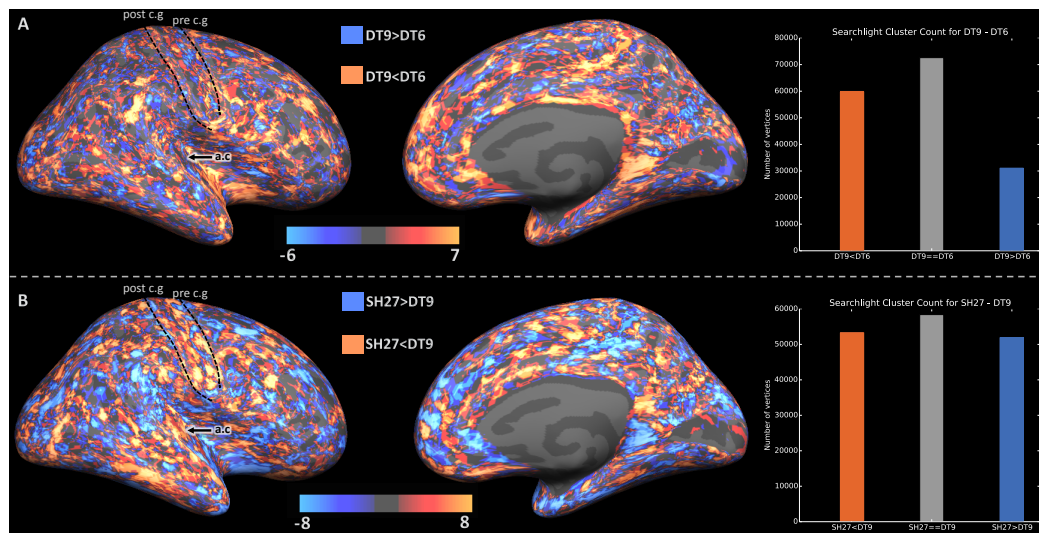


Figure 6.4: Group average searchlight cluster coherence comparisons. (A) DT9 vs DT6: orange indicates that the parcellation was more spatially coherent in DT9 and blue indicates the reverse effect. (B) SH27 vs DT9: orange indicates that the parcellation was more spatially coherent in SH27 and blue indicates the reverse. Bar charts to the right show the number of vertices satisfying each condition across the whole hemisphere. The dotted contours highlight the position of the pre-central gyrus (pre c.g) and post-central gyrus (post c.g), and the black arrow points to the auditory core (a.c).

QUANTITATIVE ASSESSMENT The searchlight comparison of the results is shown in Figure 6.4. Overall, DT6 and DT9 are similar to each other in terms of cluster coherence, as a large number of vertices (over 70000) had equal cluster counts for both feature sets. Where they differ, DT9 tends to out-perform DT6, with 60000 wins for DT9 (orange) compared to 30000. The local cluster counts of DT9 and SH27 are equal for a smaller portion of the cortex (just under 60,000 vertices). The number of vertices in which SH27 outperforms (orange) or underperforms (blue) DT9 are relatively even. However, the distribution of these results indicate that SH27 provides more spatially coherent clusters in the central sulcus,

auditory core, MT, cingulate sulcus and the temporal lobe. In contrast, DT9 performs better in the inferior parietal lobe and posterior default mode network areas.

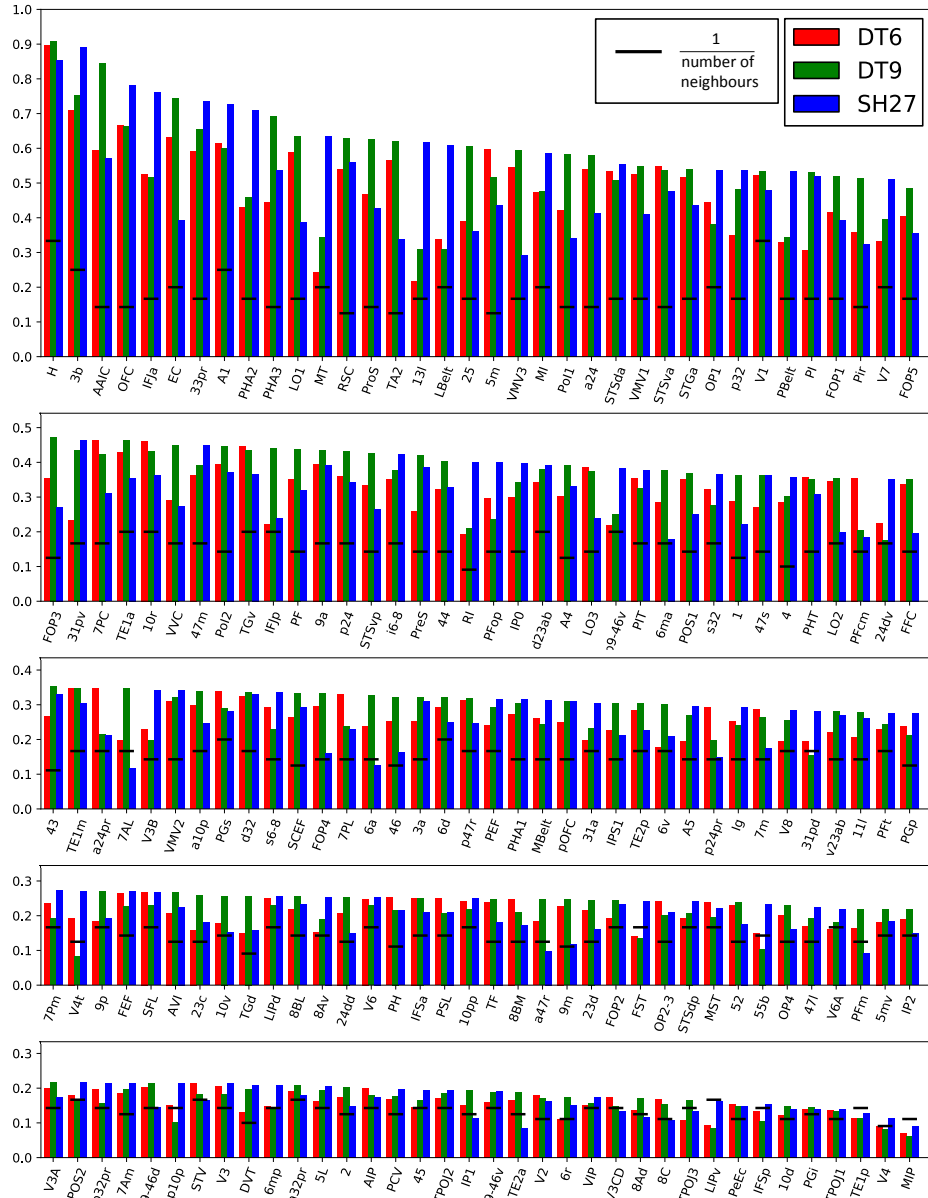


Figure 6.5: Bar charts comparing classification accuracy across all labels in the group average whole hemisphere parcellation. As in Figure 6.6, the bars correspond to the proportion of correctly classified vertices in each ROI for feature sets DT6 (red), DT9 (green) and SH27 (blue). The ROIs have been ordered according to the overall highest classification accuracy across any of the feature sets. A comprehensive legend containing area names and descriptions can be found in the supplementary material of Glasser et al. (2016).

The bar plots in Figure 6.5 of the supplementary material display the classification accuracy of each feature set in each of the 180 areas. 125 of the 180 areas were reproduced with a greater than chance accuracy for all three feature sets. Many areas were highly reproducible, for example, the Hippocampus, area 33pr and MT to name a few. Areas such as LIPv and MIP could not be distinguished from their neighbours.

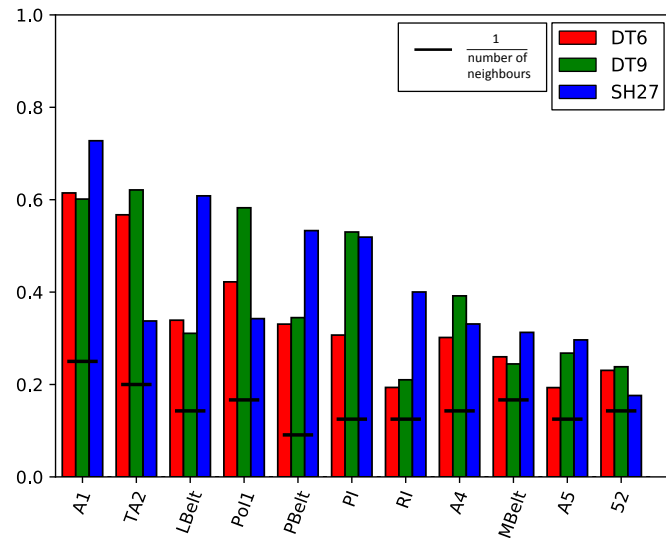


Figure 6.6: Bar chart comparing classification performances in auditory areas. The bar heights indicate the proportion of correctly classified vertices in each ROI for the DT6 (red), DT9 (green) and SH27 (blue) feature sets. The black lines indicate the chance outcome for each ROI, i.e., 1/the number of neighbours for each ROI.

Figure 6.6 takes a closer look at performance for a subgroup of areas belonging to the auditory network in the insular cortex. Many of the areas in the auditory subgroup are classified with a much higher than chance accuracy. SH27 is the winning feature set for just over half of the areas, whereas DT9 wins in the remaining portion. The performance of SH27 is better within the auditory core (A1, RI) and surrounding belt areas (LBelt, PBelt, MBelt), which have previously been shown to have extremely distinct myelin characteristics (Serenó et al. 2013). In contrast, DT9 yields higher accuracy in areas that are generally more architectonically uniform (areas outside of primary, secondary, and tertiary visual, auditory, somatosensory, and motor areas).

6.2.3 Single subject whole hemisphere classifier

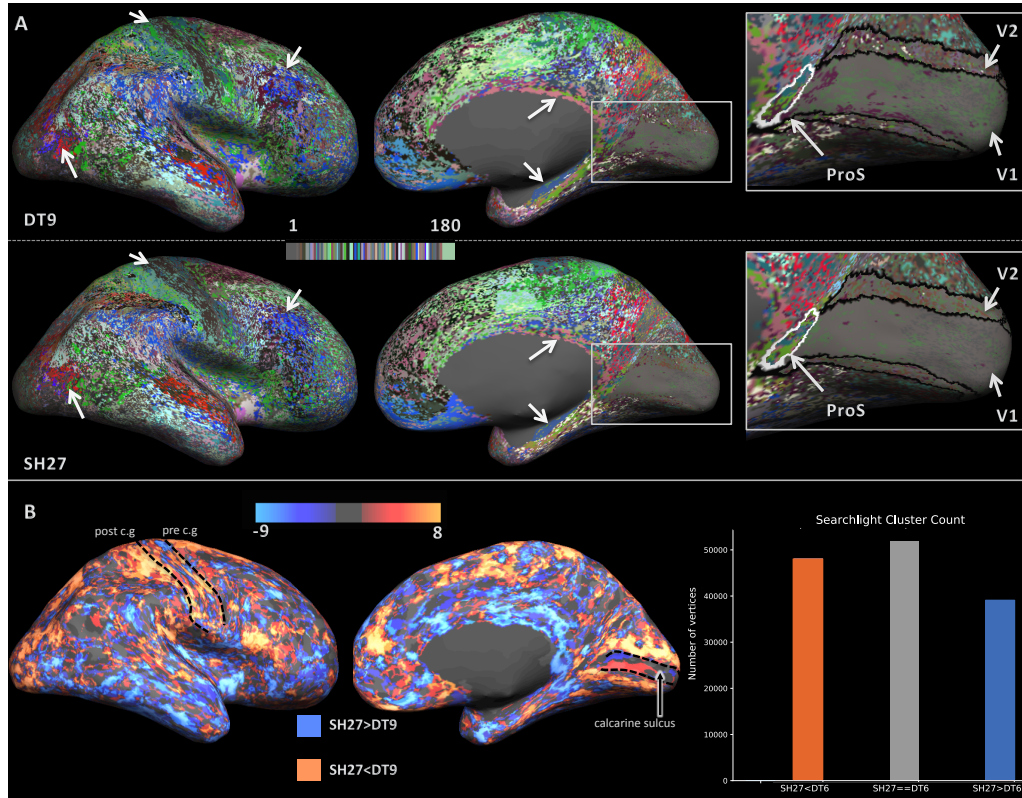


Figure 6.7: Single subject full hemisphere parcellation results. A) Medial and lateral views of the parcellation for DT9 (top) and SH27 (bottom). The white arrows highlight areas that exhibit a good correspondence to the training labels. The right panel provides a close up view of the primary visual cortex. (B) Maps comparing searchlight cluster coherency between the single subject DT9 and SH27 parcellations. Orange indicates that the parcellation was more spatially coherent in SH27 and blue indicates the reverse. Bars on the right show the number of vertices belonging to each condition across the whole hemisphere. The dotted contours mark out the pre-central gyrus (pre c.g.), post-central gyrus (post c.g.) and calcarine sulcus.

Figure 6.7A shows single subject whole hemisphere classification results for DT9 and SH27. The results for DT6 are not shown but the trends were similar to the group average result with DT6 giving the most granular, least accurate classifications. The overall spatial coherence is lower for both feature sets than the group average results (see Figure 6.3). However, area-like clusters can still be observed in both (white arrows). The map for SH27 is qual-

itatively smoother than that of DT9. In particular, SH27 provides a much more coherent definition of V1 than does DT9, or indeed any of the group average results above. But again, neither of the feature sets manages to properly differentiate V1 from V2.

The quality of each parcellation was scrutinised more rigorously using the searchlight cluster counting method (Figure 6.7B). Again, as in the group average findings, SH27 provides a quantitatively smoother parcellation overall. The distribution of these results is also similar to the group average result, i.e., SH27 is more consistent in regions such as the sensorimotor areas of the central sulcus and the primary visual areas.

6.3 DISCUSSION

Areas of high reproducibility

This work provides evidence that diffusion MRI is a sensitive and anatomically meaningful contrast for identifying differences between cortical areas at 3T resolutions. We demonstrated that the M1 and S1 cortices can be reliably distinguished from each other using a simple white matter model, the diffusion tensor, as previously suggested (McNab et al. 2013). The transition between S1 and M1, within the central sulcus, is one of the most distinct in the cortical sheet (Brodmann 1909; Geyer et al. 1997; White et al. 1997). Furthermore, the associated Brodmann Areas, 3b (part of S1) and 4 (M1), are consistently located along the posterior and anterior banks of the central sulcus respectively. Thus, it is understandable that these areas were easy to reliably distinguish. However, accurate classification outcomes were not limited to these two regions. We found that 3b and several other areas can be reliably differentiated from their surrounding cortical tissue using group average data. For example, A1 and auditory belt areas demonstrated a large overlap with their corresponding training labels. High reproducibility in these areas suggests the ability of the classification method and feature sets to overcome the confounding effects of inter-subject variations in idiosyncratic cortical folding patterns within Heschl's gyrus (Leonard et al. 1998).

Areas of high reproducibility are not limited to myelin rich areas; for example, 78%

of the vertices in the inferior frontal sulcus area, IFJa, were correctly assigned using the SH27 feature set despite this area having six neighbours (see Figure 6.5). Such examples indicate that dMRI provides useful contrast in regions where myelin density is not as informative. However, as the fidelity of the training labels is questionable in these areas, further investigation will be required to illuminate what is driving classification outcomes in these regions. Nevertheless, the results suggests that dMRI could be a useful modality to incorporate in future studies that aim to non-invasively fingerprint the differing microstructure in cortical units. This hypothesis was further supported by the analysis of the single subject, whole hemisphere parcellation in which area-like clusters were demonstrated in similar regions to the group average results.

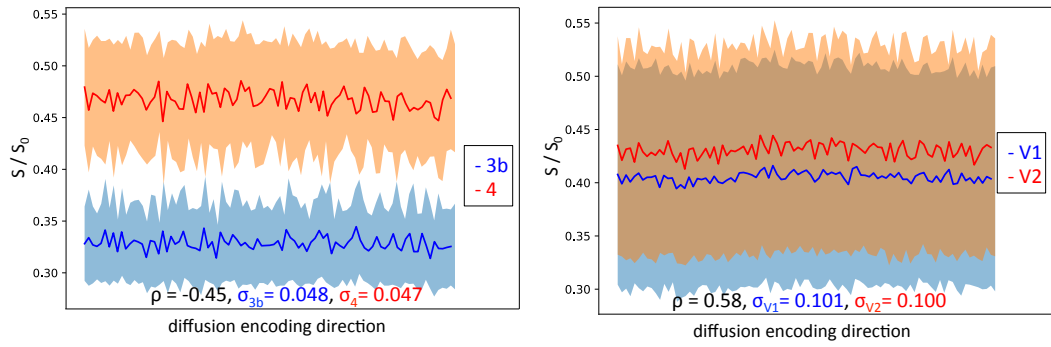


Figure 6.8: The mean DWI signal intensity in areas 3b and 4 (left) and V1 and V2 (right) for a single subject. The subject is the same as the one for which results were shown in Figure 6.7. The signal intensities have been normalised by the mean $b=0\text{s/mm}^3$ and the shaded regions indicate the standard deviation within each ROI. σ_{area} is the mean std across the DWIs. ρ is the Pearson correlation coefficient between the mean signals for each pair of areas.

Areas of low reproducibility

All feature sets failed to clearly distinguish the V1 and V2 areas in the group average classifier, despite the marked differences between these areas (Amunts et al. 2000; Hinds et al. 2008; Mountcastle 1997). It is possible that inter-subject variability regarding the exact

boundary between these two regions causes mixing of data when the averaging is performed which in turn obscures the contrast between these regions in both the training and test data. Although the horizontal meridian of V1 always resides within the calcarine sulcus, V1's extension onto the surrounding gyrus, and therefore its boundary with V2, shows considerable variation across subjects (Amunts et al. 2000). It is also likely that the relatively low resolution of the HCP data is insufficient to delineate defining characteristics in the extremely thin V1 region. Turner et al. (2008) have suggested that an isotropic resolution below 0.6mm³ is required at 3T to consistently image Stria of Gennari. It may also be that the interacting effects of orientation dispersion and microstructural composition (Kaden et al. 2016; Reisert et al. 2017) diminishes differences in the dMRI signal between these two regions. Figure 6.8 indicates that the signal intensity across different gradient directions is more correlated between V1 and V2 compared to S1 and M1.

In some regions, multiple training areas were classified as the same cluster. It is possible that the dMRI signal is not sensitive to subtle differences between these regions or that the multi-modal training labels do not correspond to their architectonic subdivisions.

Cluster coherence

It should be noted that no smoothing steps were implemented to enforce spatial coherence in any of the cortical maps illustrated so far. Of course, a much cleaner result can be obtained by adding additional post-processing steps. For example, a winner-take-all approach (Figure 6.9) results in a significantly less noisy parcellation and closer correspondence to the training atlas. Though this may be beneficial in some applications, we felt that it was critical to illustrate the unaltered, vertex-wise results that will eventually form the basis for more complex, knowledge-based pipelines. Furthermore, the unsmoothed results allowed us to use cluster coherence as a measure of performance (Figures 6.4 and 6.7B). Importantly, this analysis suggested that the higher-order SH27 method provides a more spatially coherent result in areas where the training labels were defined using architectonic information, such as heavily myelinated primary cortices.

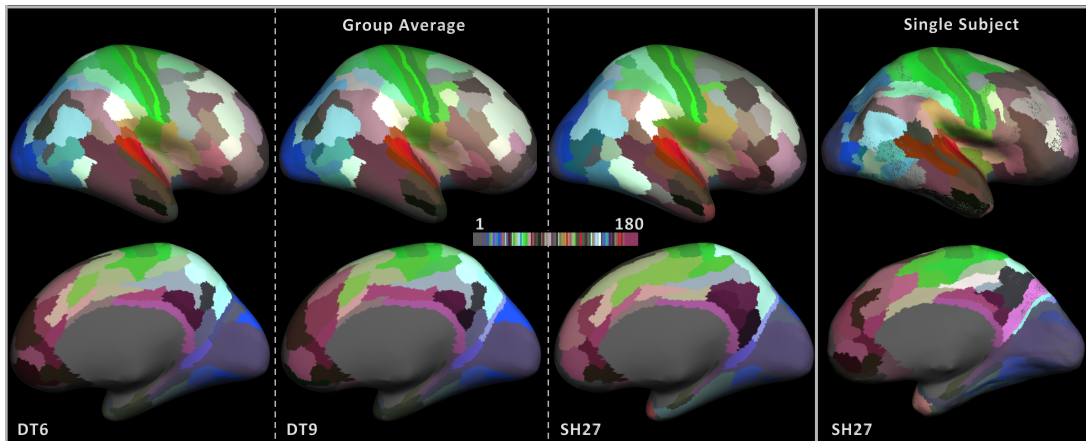


Figure 6.9: Winner takes all group average parcellation results for DT6, DT9 and SH27 and the single subject winner takes all results for SH27. For each feature set, the results were generated by calculating the most frequent class under each training label and assigning the entire region to that class ID. The resulting labels match the training labels for 106, 113, 105, and 99 (from left to right) out of the 180 areas.

Training labels

The HCP-MMP labels were generated from the simultaneous analysis of myelin density, resting state fMRI, and task fMRI. As such, the boundaries found by combining these modalities are not necessarily correlated with the cortical features that are captured by the dMRI signal attenuation. These provisional labels cannot be expected to exactly correspond to the underlying cortical areas, as defined by their architectonic properties alone. In addition, the uniform parcels of our training labels are themselves an abstraction from the real neocortex. For example, many of the best-defined cortical areas (e.g., MT, S1, M1) contain internal architectonic boundaries that are just as striking as any between-area boundary (Kuehn et al. (2017) on M1/S1; Sereno et al. (2015) on MT). This is particularly important to consider when interpreting the results, especially given the supervised nature of the classification methods. Care should be taken not to over interpret the above results in regions where the training labels were heavily influenced by functional MRI modalities, such as prefrontal areas. Whilst the boundaries in such regions may not be congruent with architectonic domains, they were still useful in demonstrating that regional variance can be observed across the cortex using dMRI. Further analysis, involving high-resolution architectonic mapping is

needed to shed light on what is driving the contrast in these areas. Such studies may also provide better training labels that can minimise the circular reasoning associated with supervised classification.

It should also be considered that we used only one of the many competing atlases. We remain some distance from being able to generate a definitive in vivo, cyto- and myeloarchitectonic reference map of the entire cortical surface. However, the method adequately demonstrated that diffusion MRI represents a complementary modality for future studies of cortical microstructure.

Feature sets

The classification efficacy of several diffusion-based feature sets were assessed. In particular, we wanted to determine whether either (a) the explicit inclusion of radial and tangential diffusion properties (via the local surface normal) or (b) the use of higher-order feature sets, improves between-area contrast compared to commonly used scalar metrics, such as FA.

BINARY M1/S1 CLASSIFICATION Initial tests on S1/M1 classification found relatively poor performance in lower order feature sets that only used a single b-shell or combined b-shells before classification (DT3, ND3). The consistently poor performance of DT3 compared to DT6 and DT9 supports the notion that different b-values can probe different aspects of cortical microstructure (Nagy et al. 2013). The relatively weak performance of ND3 suggests that the three-shell data contains more useful information than is captured by the NODDI model, which imposes biophysical assumptions regarding the underlying tissue composition. Crucially, the improvement in classification when adopting the second-order diffusion tensor (DT9), compared to the higher-order 4-tensor (4T36) or spherical harmonics (SH27), relies on the inclusion of surface specific metric, i.e., the radially index, which was omitted from DT6.

The DT-based feature sets perform well despite diffusion at higher b-values ($b > 1000$) not respecting the Gaussian assumptions of this model (Alexander et al. 2002; Le Bihan et

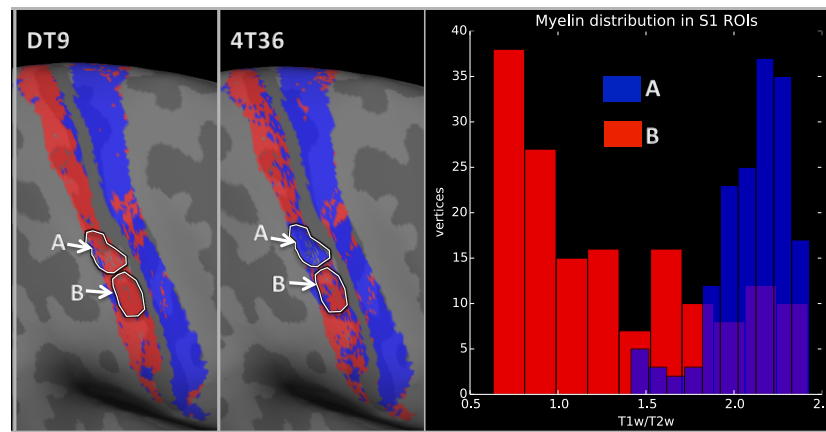


Figure 6.10: Confirmation that misclassification by the 4T36 feature set is driven by heterogeneity in myelin density within area 3b. The binary classification result for 4T36 is shown (right) with the outlines of the two ROIs that were selected. The myelin distribution within each ROI is also shown (left), where myelin density is measured from the T1w/T2w ratio provided with the HCP structural images.

al. 2006). One possible interpretation is that the higher-order features may be driven more by noise and inter-subject variation than intrinsic features of distinct grey-matter regions, at least at our current resolution. Alternatively, we hypothesised that the low-order features are so distinct between these regions that they dominate the classification. The DT-based features also create a smoothing effect that results from the coarser description of the microstructure that they provide. As such, it is possible that DT model is insensitive to real, within area, microstructural variation. For example, there are myelin density changes corresponding to the boundaries of individual digits in 3b (Serenó and Tootell 2005), or the hand, foot, and face subdivisions of areas 3b and 4 (Kuehn et al. 2017). Figure 6.10 indicates that areas of misclassification by 4T36 correspond to large variations in the underlying myelin density (and thus architecture) within the tissue. These variations are not reflected in the DT results.

FULL HEMISPHERE PARCELLATION The full hemisphere, group average, and single subject results provide evidence in favour of using the more generalised, higher-order spherical harmonic feature sets. In contrast to the trend observed in the binary experiment, we

found that in this multi-class problem SH27, not DT9, obtains the most accurate definition of 3b. It is possible that the DT-based features are not sensitive enough to describe differences between the larger set of classes. S1 has 3 neighbours in this 4-way classification meaning the three times as many features that SH27 provides are beneficial.

The more generalized features of SH27 provide a more coherent clustering result across the cortex, particularly for the single subject result. This again suggests that these features describe cortical tissue domains more effectively. Perhaps the higher order feature sets are less influenced by confounding effects, such as changes in global diffusivity between subjects (Veenith et al. 2013), than the DT feature set as a result of their increased dimensionality. Support for the higher order feature sets is again evidenced in the distribution of areas in which SH27 out performs DT6 and DT9. It achieves better results for areas of high myelination, primary areas, or areas for which the training labels can be considered more reliable. On the other hand, it is possible that the SH27 feature set is more susceptible to overfitting in regions where the diffusion signatures of neighbouring areas are less distinct from each other. The combination of DT9 and SH27 can also be used as a feature set to maximise performance across the cortex (see Figure A2). A few regions are still better classified with either of the reduced sets (DT9 or SH27) than with the combined feature set (SH27+DT9). For example, area 3b is better classified using SH27 alone, suggesting that combining it with DT9, which performed comparatively poorly in this area (see Figure 6.3B), adds noise to the RFC.

INCLUSION OF SURFACE SPECIFIC FEATURES The inclusion of features that take the orientation of the local cortical surface into account (i.e., the radially index in DT9, several features in the SH27 set), consistently offered an advantage over the scalar DT6. This is clear when comparing DT6 to DT9 in both the binary classification and group average experiments. However, we cannot conclusively say that such features are always necessary. Comparing the results of SH27 to 4T36 indicates that explicit reference to the local tissue orientation might not be required if the feature set provides a functionally complete description of the ADC. Further testing of 4T36 across a broader set of areas is required to confirm

this.

Ultimately, deciding which feature set to use requires a nuanced approach that considers the specific aims of future studies. If attempting to delineate architectonic domains, the above results indicate that high order decomposition approaches might be more appropriate when describing the texture of the 'fabric' of the neuropil at an intermediate scale. On the other hand, an advantage of biophysical models is that they provide features that are more readily interpreted. For example, they are more useful if one wishes to understand the specific microstructural changes at the level of single fibres that can affect abnormal brain function.

The above works omit a set of recently emerging techniques, which aim to separate the contribution of microstructural tissue composition from the mesoscopic orientation distribution within the dMRI signal (Kaden et al. 2016; Reisert et al. 2017). These methods remain to be tested in a cortical parcellation framework and may be particularly beneficial in extrastriate or other non-primary areas that do not exhibit distinct tangential or radial laminar properties.

Limitations and future work

One of the limiting factors of the above work was the relatively coarse resolution of the diffusion data compared to the thickness of cortical laminae. This only allowed data to be sampled at a single cortical depth, which may have led to a failure to sufficiently capture variations in laminar structure, particularly for thicker areas of the cortex i.e., gyral crowns or area 4 (M1). It was also insufficient for reliably characterising the properties of V1. The relatively large voxels may have introduced noise by differential mixing of signals from different laminae in different locations. A finer sampling of different depths in each cortical column has the potential to provide a closer approach to the classical histological analysis of the cortex. The low resolution is also likely to incur partial volume effects, which we aimed to minimise by sampling at the middle cortical depth. However, further analysis is needed to determine the extent to which partial voluming from WM and CSF impact the results. The higher spatial resolutions that can be obtained at 7T have great potential for resolving these

issues and validating the above findings.

The HCP presents a unique opportunity to validate the above findings at higher resolution using the recently released 7T datasets. 184 of the 3T participants were also scanned at 7T with an increased isotropic spatial resolution of 1.05mm^3 . Analysis of the 7T data has revealed that several additional voxels are present between the GM/WM boundary and pial surface compared to the 3T counterparts (Vu et al. 2015). In addition, the 7T data shows improved contrast in cortical areas with extremely low FA, and the authors were able to resolve details such as WM fibres changing in orientation as they entered the cortical sheet along gyral walls (*ibid.*). The first step would involve repeating the 3T pipeline that we presented to see if the trends we observed are maintained in the better quality data. An interesting next step would be to extend our method to a more laminar-like analysis pipeline. For instance, the feature vectors of several voxels, sampled at different depths along the local surface normal, could be concatenated to represent the diffusion within a cortical column. Although, given that cortical thickness can be as little as 2mm, it is possible that even the HCP 7T data have insufficient spatial resolution for this type of analysis in very thin cortical areas.

There are several additional avenues that could be explored in more depth in future studies. For example, the labels were back-projected from a reference brain, but it would be interesting to investigate whether generating labels on individual HCP subjects can improve accuracy. Furthermore, we found that the feature ranking information provided by the RFC was highly variable between different regions – additional analysis could shed light on which features are more discriminative for which areas.

MULTIPLE B-VALUES FOR CORTICAL IMAGING

So far we have utilized the HCP multi-shell HARDI datasets in our classification experiments. These datasets benefit from very rich q-space sampling, high spatial resolution and use of multiple diffusion weightings. In the previous chapter, we saw that features based on a single b-shell were less effective than their triple shell counter-parts. In this chapter we build on this to explicitly demonstrate that multi-shell HARDI protocols are advantageous in cortical grey matter applications.

Previous investigators have optimized b-values for a single or a range of diffusion constants (Le Bihan and Warach 1995; Papadakis et al. 1999; Kingsley and Monahan 2004; Armitage and Bastin 2001; Alexander and Barker 2005). However, a single b-value will not be optimal for the specific tissue in all voxels and all orientations of that tissue relative to the diffusion-encoding directions (Skare et al. 2000). Furthermore, a given b-value will probe different aspects of the tissue microstructure depending on the timing and amplitude of the gradient pulses. These factors may be particularly relevant in cortical grey matter applications, where local tissue orientation is highly variable due to cortical folding.

We investigated whether repeated measurements at a single b-value or multiple different b-values perform better in cortical area classification tasks. We hypothesized that, analogously to multi-echo fMRI methods (Posse et al. 1999; Poser et al. 2006), collecting data at additional b-values would improve classification accuracy due to the fact that there is a higher probability of reaching the optimal b-value within each voxel. We tested this hypothesis on the HCP data and an additional local dataset acquired at 1.5T, first comparing the correlation coefficients across data from different shells and then testing different b-value combinations in a series of binary classification experiments.

Some of the ideas in this chapter have been previously presented in:

Nagy, Z. et al. (2014). "Combining HARDI datasets with more than one b-value improves diffusion MRI-based cortical parcellation." In: Proceedings of the 22nd Annual Meeting of International Society for Magnetic Resonance Imaging, Milan, 2014, p. 800.

7.1 METHODS

7.1.1 *Data and processing*

HCP 3T DATA Data from three HCP participants were used in the below experiments, detailed descriptions of these datasets are provided in Section 5.1.1 and the relevant HCP publications (Van Essen et al. 2013b; Sotiropoulos et al. 2013; Glasser et al. 2013; Uğurbil et al. 2013). In brief, the datasets have a spatial resolution of 1.25mm^3 and contain three b-shells, $b=1000, 2000$ and 3000 s/mm^2 , collected with matched diffusion and echo times.

To facilitate fair comparison between the repeated and multiple b-value combinations we created size-matched datasets by sub-sampling the gradient directions from each b-shell. Shells were split into three evenly spaced subsets, containing 30 directions each, using an electrostatic repulsion algorithm (Cook et al. 2007). This provided a mixed b-value dataset containing one 30-direction sub-volume each from the $b=1000, 2000$, and 3000 s/mm^2 shells, and a repeated b-value dataset containing the full set of 90 $b=2000\text{ s/mm}^2$ DWIs. The middle b-value was selected for the repeated b-value combination to minimize signal-to-noise ratio (SNR) differences between the two datasets.

LOCAL 1.5T DATA Additional data were collected for three healthy adult volunteers in accordance with local ethics approval. A 1.5T Siemens Avanto scanner with a 32-channel head coil was used. Three b-values were sampled, with the middle b-value repeated 3 times. The order of acquisition was $b=1400, 800, 1400, 2000, 1400\text{ s/mm}^2$ and each image set was collected in both the blip-up and blip-down phase encode directions, resulting in a total of ten DWI volumes per subject. A multiband factor=2 was used with $\text{TE}=86, 94, 101\text{ms}$ and $\text{TR}=5647, 5980, 6224\text{ms}$ in the $b=800, 1400$ and 2000s/mm^2 shells, respectively. The local data had spatial a resolution of 1.7mm^3 .

All 10 volumes were simultaneously pre-processed using the Eddy (Andersson and Sotiropoulos 2016) pipeline to remove susceptibility and eddy current distortions whilst ensuring that all volumes were co-registered. Subsequent to pre-processing two diffusion datasets were generated for each subject. The first containing a mixture of b-values, comprised of $b=800$, 1400 , and 2000 s/mm^2 image volumes. The second contained all three $b=1400 \text{ s/mm}^2$ volumes, again ensuring that the datasets were size matched and roughly SNR matched.

SAMPLING AND FEATURE REPRESENTATION The surface-based pipeline of the previous chapters was again adopted to sample the diffusion-weighted data at the middle cortical depth for each dataset. To perform classification experiments nine spherical harmonic features were calculated per b-shell (same as Chapter 6), resulting in a 1×27 feature vector at each vertex for each dataset.

7.1.2 *Correlation analysis*

We examined the correlation coefficients between the feature sets generated from each b-value to determine whether the different b-values provide additional information content. At each surface vertex the squared correlation coefficient was calculated between the different b-shells for both the HCP and local data. Features sets generated using repeated b-values were expected to be highly correlated with each other, whereas feature sets from different b-values should exhibit lower correlation coefficients.

7.1.3 *Classification*

REGIONS OF INTEREST For the 1.5T local datasets seven regions of interest were defined on the lateral surface. Three ROIs were defined using the FreeSurfer Brodmann Atlas labels (Fischl et al. 2008; Zilles et al. 2002) for areas 3b, 4p, and 45, corresponding to the primary somatosensory and motor regions and Broca's language area, respectively.

The remaining four ROIs: MT, VIP, AnG and Insular, were defined in a previous study using multimodal data (Nagy et al. 2013), and correspond to the middle temporal area, ventral intraparietal area, a section of the angular gyrus belonging to the default mode network and a small region in the auditory core within the insular. The regions were selected to cover a wide range of cortical tissues and are displayed for a single subject in Figure 7.1.

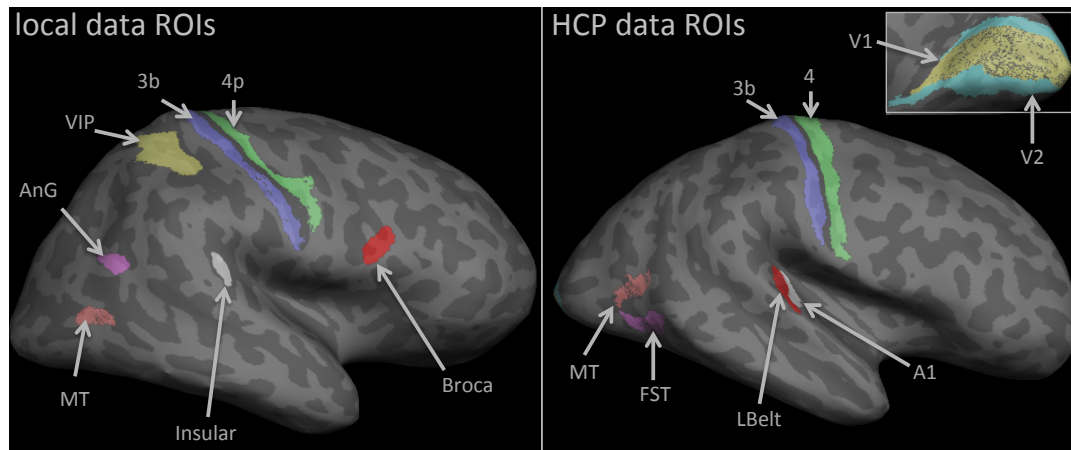


Figure 7.1: Regions of interest used in classification experiments to test different b-value combinations. Left shows the ROI for the local 1.5T data, generated using the probabilistic Brodmann Atlas labels, and fMRI data. Right shows the ROI for the HCP data, generated from the HCP multi-modal parcellation labels.

Regions of interest for the HCP data were defined using the HCP Multi-Modal Parcellation (HCP-MMP) (Glasser et al. 2016). A total of eight area labels were selected and registered from the fsaverage surface to the subject specific surface tessellations using surface based registration (Fischl et al. 1999b). Some of the ROIs were equivalent to those above, i.e., 3b, 4, MT and A1. The remaining regions were V1, V2, FST, and Lbelt corresponding to the primary and secondary visual areas, the fundus of superior temporal sulcus and the lateral auditory belt region, respectively. The regions were selected to encompass a range of cortical functions and also include some neighbouring regions (V1/V2 and A1/Lbelt) to assess the performance of the feature sets in boundary detection (see Figure 7.1). However, we limited analysis to a small subset, of the full 180 areas, choosing specifically ROIs that were well supported by the surrounding anatomical literature, therefore, increasing the

reliability of the training labels.

CLASSIFICATION TESTS The appropriate set of ROIs for each dataset (local or HCP) were tested against each other using random forest classification (Breiman 2001). Within each binary classification experiment, test/train datasets were generated using leave-one-out cross validation, in which the classifier is trained on all but one of the data-points and then tested on the unseen data-point, this process is repeated until all data-points have been tested. It is possible for geographically distant areas to have similar architecture, which may confound the classifier, making it difficult to find an optimal splitting plane in a multi-class problem. Using, instead, a series of binary classifications, limits this confounding effect to a small subset of the classification tests, allowing one to easily identify which regions are most similar.

Apart from comparing the performance of the data with three different b-values against the data with three identical b-values, further comparisons were made on partial data sets that contained only two acquisitions: either different or repeated b-values. Namely, binary classification experiments were performed on features extracted from a) the lowest and middle b-values, b) the two middle b-values and c) the middle and high b-values. For example, in the local data this would result in dataset1: b=800, 1400b; dataset2: b=1400a, 1400b; dataset3: b=1400b, 2000 s/mm². The middle b-value volume was kept constant for all three datasets. These double b-value datasets were tested using the same classification experiments as the triple b-value datasets, described in the previous paragraph. These additional pairings helped to determine if the inclusion of a higher or lower b-value was more beneficial.

PERFORMANCE METRICS Two performance metrics were used to assess the quality and accuracy of the classification experiments. First, the proportion of correctly classified vertices from the two ROIs in a binary experiment was taken to indicate overall performance accuracy (as in Chapter 6). Secondly, an aggregated F1-score was used as a measure of accuracy within each region of interest. The F1-score provides the harmonic mean between precision and recall, and is commonly used to measure the accuracy in binary classification.

It quantifies the proportion of true positives in relation to both false positives and false negatives, giving a value ranging from 0-1, where 1 means the classification result was 100% accurate (Powers 2011). A high, F1 score across all tests would indicate that the region is particularly distinct, compared to the others, and well characterised by the feature set.

7.2 RESULTS

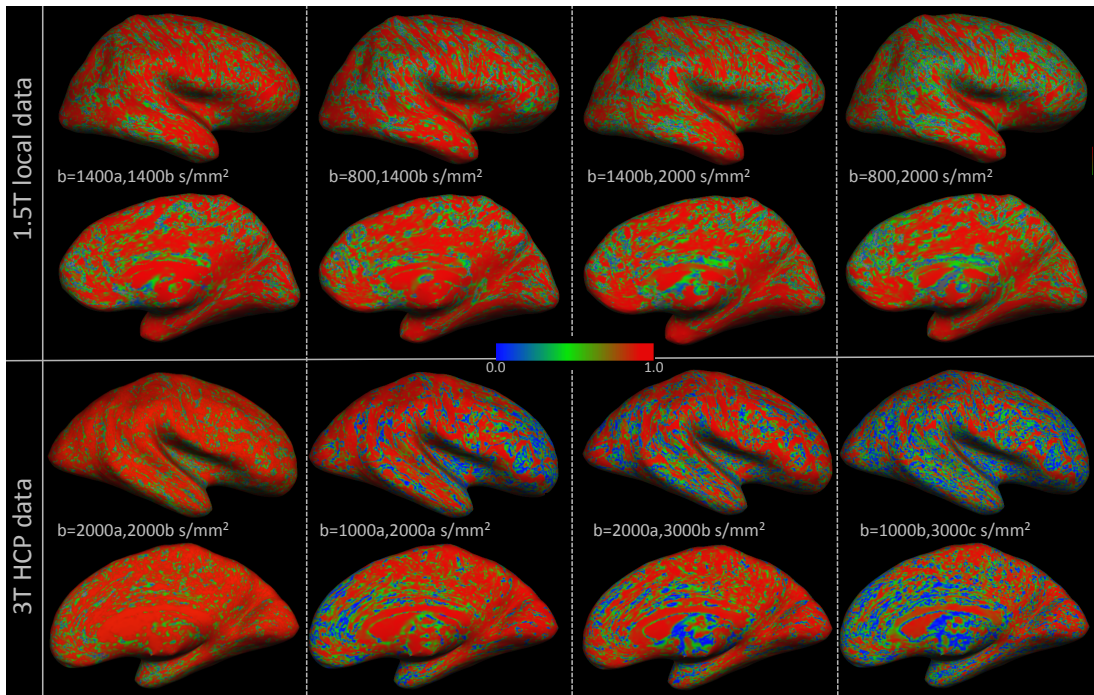


Figure 7.2: The squared correlation coefficients between different feature sets from different b-values.

The top two rows show the correlations across the medial and lateral surfaces of a local 1.5T subject. The bottom rows show the equivalent results for a HCP subject. In each case the first column displays the correlation between two repeated b-values, and the rest of the columns display the correlation between two differing b-values.

7.2.1 Correlation maps

Both datasets show a similar trend in Figure 7.2. The correlation drops as the difference in b-values increases. The effect is less pronounced in the local data, however it is still clear

that the first column, showing the correlations between two $b=1400\text{s/mm}^2$ shells, has a higher proportion of red, highly correlated vertices, than the other columns, particularly with comparison to the correlation map on the far right which is calculated between the lowest ($b=800\text{s/mm}^2$) and highest ($b=2000\text{s/mm}^2$) b-value datasets. Analysis of these correlation maps suggests that the varying information captured by the differing b-values does indeed manifest in the spherical harmonic features used for classification and that using multiple different b-values may provide a richer feature vector for classification between cortical tissue domains.

7.2.2 Classification

TRIPLE B-VALUE EXPERIMENTS The mean performance of the classification experiments in data that contained either three different b-values or a single b-value repeated three times is provided in Figure 7.3. For both the local and the HCP data, the feature set obtained from the mixed b-values (pink) outperforms that of the repeated b-values (blue) in all of the tests. This supports the hypothesis that measuring diffusion MRI data with multiple b-values will contain more discriminative power to differentiate the cortical areas. There was an average difference of $5.6\% \pm 2.6\%$ between the mixed and repeated feature sets across tests for the local data and $4.6\% \pm 4.2\%$ for the HCP data. In both cases, the differences were found to be significant, with $p < 0.001$, according to the Wilcoxon signed rank test.

An aggregated F1-score was obtained for each ROI by taking the average F1 value across all binary classification tests in which that region was used (Figure 7.4B). For all regions, in both the local and HCP datasets, the mixed b-values result in better classification. For the local 1.5T data the differences were consistent with a 4-10 point difference between the two b-value conditions, across all the regions. However, for the HCP data the effect is stronger for the smaller areas i.e., MT, FST, A1, and LBelt, which in general are classified with less accuracy than the larger, primary regions of V1, V2, 3b and 4. The dataset with mixed b-values provides a much larger increase in accuracy when classifying the smaller regions i.e., a 10-20 point difference over the repeated b-values datasets. This suggests that multiple b-values are particularly beneficial at characterizing small cortical areas.

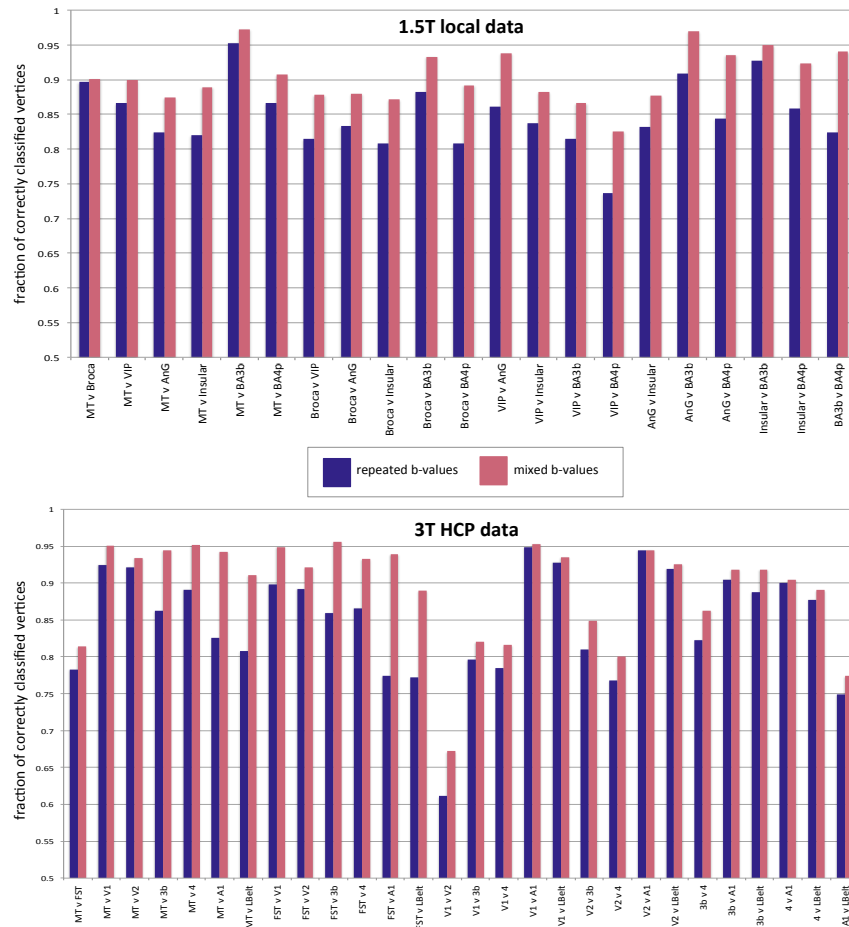


Figure 7.3: Comparison of the performance of each dataset in the binary classification experiments for the triple b-value combinations. Results for the local data are displayed at the top and results for the HCP data are displayed at the bottom. In both bar charts, the dataset that combined three different b-values is shown in pink bars, whereas the dataset that combined repeated acquisitions of a single b-value correspond to blue bars. Bar height corresponds to the mean classification accuracy, i.e., the fraction of correctly classified vertices after averaging across the volunteers.

The result of classification between areas 3b (S1 region) and 4/4p (M1 region) are shown in 7.4A for a single subject from the local (top) and HCP (bottom) data sources. These two regions are known to have widely differing cytoarchitecture (Brodmann 1909; White et al. 1997) and we demonstrated that these differences can be reliably detected using our feature set in the previous chapters. Therefore, this pair of ROIs is likely to be the most reliable test-bed amongst the available regions. For the local data we found that the

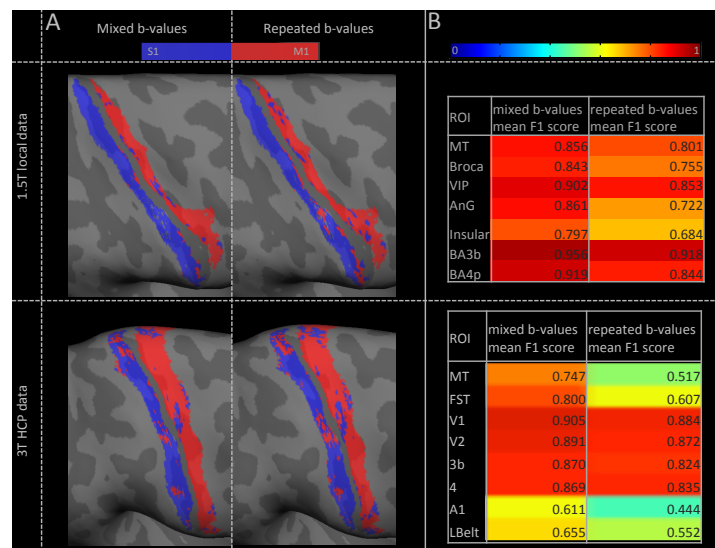


Figure 7.4: (A) Maps of the binary classification result between areas 3b and 4/4p using mixed or repeated b-value combinations. Vertices that have been coloured red have been assigned to area 4/4p by the classifier. Blue vertices have been identified as belonging to area 3b by the classifier. (B) The mean F1 score for each ROI using mixed or repeated b-value combinations. The mean was taken across all binary experiments in which the ROI was tested.

mixed b-values condition out performed the repeated b-value condition with an average improvement of $4.0\% \pm 1.1\%$ across subjects. Similarly, for the HCP data, the mixed b-values condition performed $11.6\% \pm 3.6\%$ better than the repeated b-values (see Figure 7.3). Qualitative assessment of these results on the cortical surface of the example subject reveals that the mixed b-value data produces more spatially continuous areas that bare the closest resemblance to the training labels.

Additional classification maps are displayed for the HCP binary experiments that tested immediately neighbouring ROIs against each other (Figure 7.5). We find that the mixed b-values data adheres much more closely to the training boundary compared to the repeated b-values for the classification between A1 and the LBelt. The effect is far less obvious in the classification between V1 and V2. Here, we find that for the mixed b-value dataset the majority of misclassification within the V1 area is more central in the region and there is a more consistent transition from blue to red across the border. In contrast, the repeated b-value result is generally far more noisy, displaying seemingly more random patches of

misclassification across the full extent of the ROIs.

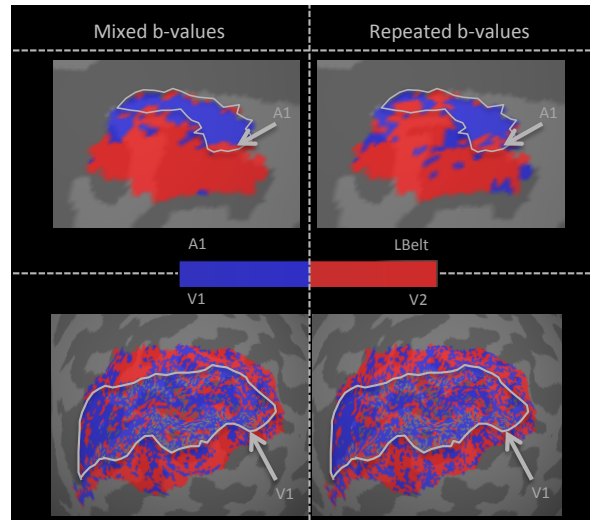


Figure 7.5: Maps of the binary classification results using mixed or repeated b-value combinations for A1 vs LBelt and V1 vs V2, shown for a single HCP subject. The grey contours show the boundaries of the A1 and V1 training labels to help compare the quality of boundary detection between the two b-value combinations.

PAIRED B-VALUE EXPERIMENTS The classification performances of the feature sets derived from the paired b-value combinations are compared in Figure 7.6. The mean differences between the various feature sets are discussed in the text below, but are also summarised in Table 7.1

For the HCP data, a mix of two different b-values outperforms two repeated b-values in the majority of tests. The repeated ($2 \times b=2000\text{s/mm}^2$) data (dark blue) produced greater classification accuracy than the combination of a low and middle b-value ($b=1000, 2000\text{s/mm}^2$), only when differentiating area 3b from the LBelt region. On average, the combination of $b=1000\text{s/mm}^2$ and $b=2000\text{s/mm}^2$ shells (light blue) performed $3.2\% \pm 3.1\%$ ($p < 0.001$) better than the repeated $b=2000\text{s/mm}^2$ data. The combination of $b=2000\text{s/mm}^2$ and $b=3000\text{s/mm}^2$ shells (green) also outperformed the repeatedly measured $b=2000\text{s/mm}^2$ for the majority of tests (excluding MT vs. FST, V1 vs. A1, V1 vs. LBelt, V2 vs. LBelt, and 3b vs. LBelt). Here, a mean improvement of $3.4\% \pm 4.2\%$ ($p < 0.001$) was observed. Combining the lowest and high-

est b-value (yellow) produced a mean classification improvement of $4.6\% \pm 3.9\%$, ($p < 0.001$) over the repeated set.

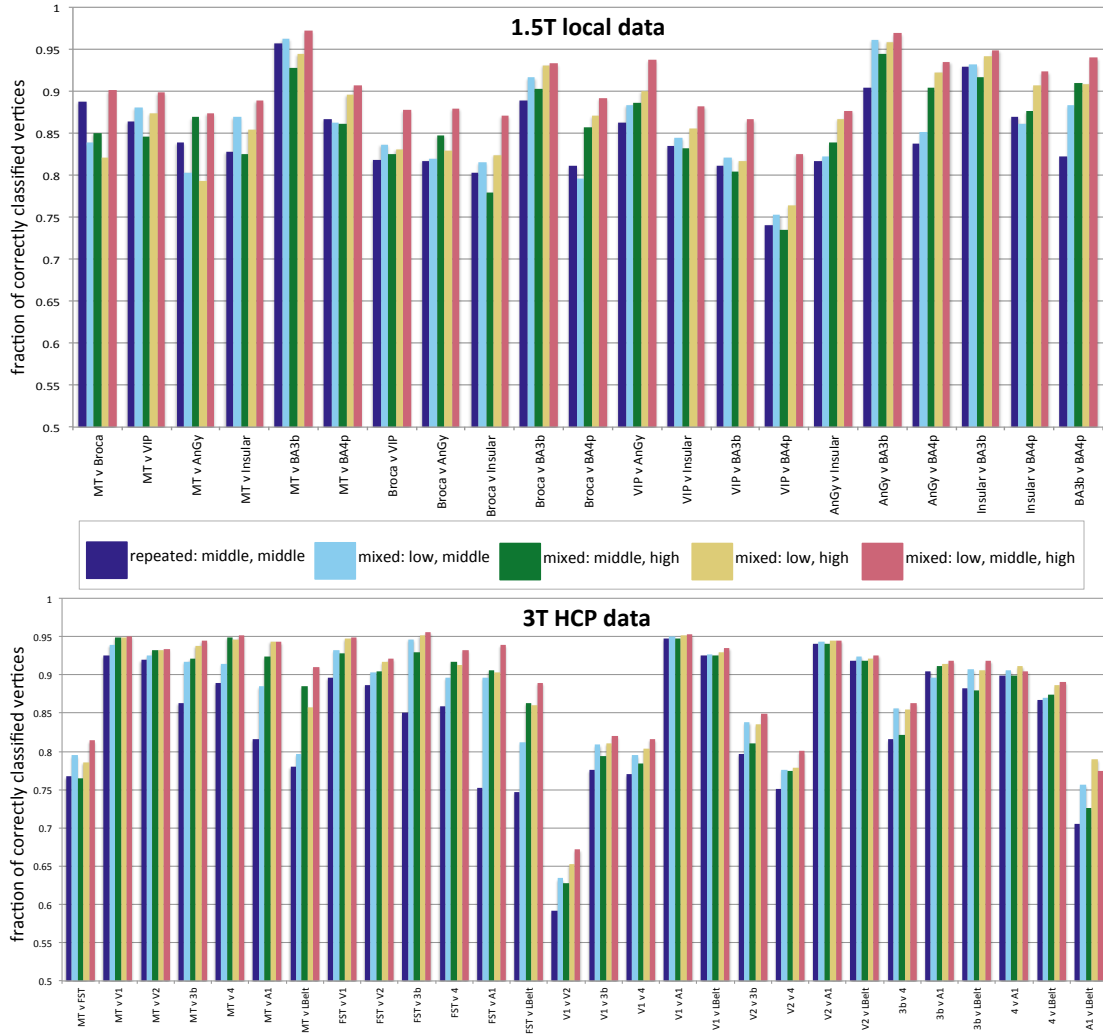


Figure 7.6: Comparison of the performance of each dataset in the binary classification experiments for the paired b-value combinations. The colours correspond to the following b-value combinations: dark blue - a pair of middle b-values ; light blue - a pair containing 1 middle and 1 low b-value; green - a pair containing 1 middle and 1 high b-value; yellow - a pair containing 1 low and 1 high b-value and pink - the combination of all three different b-values (as shown in Figure 7.3). Bar height corresponds to the mean classification accuracy, i.e., the fraction of correctly classified vertices after averaging across the volunteers.

The performance of the three mixed b-value pairs was also compared against each other. The differences between these feature sets were much smaller than when comparing

to the repeated feature set. There was a small mean improvement of $1.4\% \pm 1.8\%$ when comparing the accuracy for $b=1000+3000 \text{ s/mm}^2$ (yellow) against that for $b=1000+2000 \text{ s/mm}^2$ (green). Similarly, the combination $b=1000+3000 \text{ s/mm}^2$ was $1.2\% \pm 1.6\%$ better than $b=1000+2000 \text{ s/mm}^2$ (light blue). We found no significant differences between the two combinations that used the middle b-values, i.e., $b=1000+2000 \text{ s/mm}^2$ and $b=2000+3000 \text{ s/mm}^2$.

HCP data					
	paired	middle+low	middle+high	low+high	triple
paired		3.2 ± 3.1	3.4 ± 4.2	4.6 ± 3.9	5.6 ± 4.6
middle+low			0.2 ± 2.6	1.4 ± 1.8	2.4 ± 2.4
middle+high				1.2 ± 1.6	2.2 ± 1.5
low+high					1.0 ± 1.4
triple					
Local data					
paired		0.9 ± 2.5	1.1 ± 3.1	2.4 ± 3.6	5.7 ± 2.6
middle+low			0.1 ± 3.1	1.4 ± 2.6	4.7 ± 2.4
middle+high				1.3 ± 2.6	4.6 ± 2.0
low+high					3.3 ± 2.2
triple					

Table 7.1: The average percentage differences between the feature sets that are compared in the paired b-value experiments. The values indicate the mean differences across all of the binary tests in Figure 7.6 \pm the standard deviation in differences. The winning feature vector corresponds to the column heading. Differences that were statistically significant according to the Wilcoxon rank test ($\alpha = 0.0125$ to correct for multiple comparisons) are displayed in red.

The results for the triple mixed b-value combination from Figure 7.3 are also displayed in Figure 7.6 (pink) to reinforce that they provide highest classification accuracy overall. This three b-value combination outperforms all of the paired combinations for twenty-one out of twenty-eight tests, with a mean improvement of $5.6\% \pm 4.5\%$ over the repeated pair and 1-2% over the mixed pairs combinations.

The results for the local, lower resolution, data exhibit similar trends to the HCP

data when considering pairs of b-values against the repeated b-value. However the improvement gained from including different b-values was only significant when using the two furthest apart b-values, at $2.4\% \pm 3.6\%$, and the combination of all three b-values, at $5.7\% \pm 2.6\%$. There were no significant differences between the mixed paired combinations i.e., $b=800+1400$, $b=1400+2000$ and $b=800+2000\text{s/mm}^2$. The inclusion of a third b-value was more beneficial at this lower field strength and resolution with accuracy improvements ranging from 5.7-3.3% and proportionately smaller standard deviations ranging from 2.6%-2.0%. Again, the third shell provided the least improvement when compared to the pair that combined the shells that were furthest apart.

7.3 DISCUSSION

CORRELATION MAPS AND SNR We confirmed that the feature sets of different b-values are far less correlated across the cortical surface than those generated from independent acquisitions of the same b-value. This supports our hypothesis that different b-values provide a richer description of the cortical tissue.

Despite the directions being uniquely sampled in the HCP data, the features sets from two $b=2000\text{s/mm}^2$ subsets are highly correlated through most of the cortical surface. Spots of lower correlation are confined to areas that are prone to sampling errors i.e., gyral ridges and sulcal fundi. We can therefore be more confident that the electrostatic repulsion method, used to sub-sample the gradient directions, worked adequately without introducing significant biases. This in turn, increases our confidence that the effects observed in the HCP datasets are driven predominantly by differences in b-value rather than differences in the uniquely sampled gradient directions.

The local data exhibits relatively large patches of lower correlation (green) between the two repeated b-values (first column). This may be a result of the reduced SNR available at 1.5T compared to 3T. The SNR of the local data should be considered when interpreting the classification results (below). We attempted to minimize the SNR differences between mixed and repeated b-value feature sets by equally spacing the different b-values and repeating the middle b-value to ensure that the total amount of data was identical. But, as TE

was different for each shell, the increase in transverse relaxation will inevitably result in lower SNR in the higher shells. Comparison of the first and second column of the correlation maps suggests that differences in local data are not solely caused by SNR issues as the combination of $b=800, 1400$ s/mm² has higher total SNR than that of $b=1400a, 1400b$ s/mm² but the former still exhibits lower correlation coefficients across the cortical sheet. Furthermore, in their simulation study, Alexander and Barker (2005) found that the optimal b -value was largely independent of the noise level at WM crossings.

OVERALL BENEFITS OF MULTI-SHELL We demonstrated that sampling at different b -values provides a consistent improvement in cortical area classification experiments. The effect was found to be small, 5.6-4.6% depending on the field strength and resolution, but statistically significant. This finding further supports our hypothesis that multi-shell acquisitions better characterise cortical areas and should be considered when acquiring data for future cortical applications.

The F1 scores of each ROI exhibited a small but consistent improvement for all of the tested areas. Although the training labels cannot be considered the perfect ground truth, we limited ourselves to regions that are myelin rich, and well reported in the architectonic literature. Therefore, this universal improvement in accuracy when employing multiple b -values supports our hypothesis that multi-shell data has increased potential to probe the optimal b -value over the range of complex tissue structures and varying orientations in the cortical sheet. For the HCP data, the effect was far more pronounced for the smaller ROIs. We attribute this to the fact that the larger regions are inherently more heterogeneous and hence harder to classify.

Boundary detection is the ultimate goal of a parcellation pipeline which aims to replicate the delineations discovered by histological methods such as Brodmann (Brodmann 1909). In the experiment in which we attempted to differentiate the neighbouring A1 and LBelt areas (Figure 7.5) we found that the mixed b -value feature set produced spatially continuous clusters that closely aligned with the expected positions of these two regions. Vertices that were identified as belonging to A1 (blue) are almost exclusively found in a patch that is superior to the large red patch corresponding to the LBelt class. In contrast

the repeated b-value features produced a much grainier classification result. As in previous experiments, we found relatively low classification accuracy between V1 and V2, suggesting once again that the resolution of the HCP dataset is insufficient in these areas. Nevertheless, there appears to be a small improvement in the boundary detection between the two areas.

PAIRED B-VALUE EXPERIMENTS By repeating our classification experiments using pairs of b-values we were able to demonstrate that including a third b-value at 3T HCP resolution is only marginally beneficial, and might not warrant the increase in acquisition time. The improvement earned by the third b-value was greater for the lower resolution 1.5T data, where the additional data might be compensating for lower SNR. In both datasets we found that the best way to maximise classification accuracy was to combine the two most distant b-values. This is reflected in the correlation maps which confirmed, as expected, that these are the least correlated feature sets. It may be best practice in future cortical studies to acquire 2 b-shells that have larger separation, whilst ensuring that the SNR of the higher shell is sufficient.

Limitations and future work

We did not attempt to find a specific range of optimal b-values for cortical grey matter, as has been achieved in the WM literature (Le Bihan and Warach 1995; Papadakis et al. 1999; Kingsley and Monahan 2004; Armitage and Bastin 2001; Alexander and Barker 2005). Instead, we focussed on demonstrating that using more than one b-shell is beneficial for discriminating cortical areas. Having an optimal range of b-values may be particularly important in future clinical applications of cortical parcellation where scan times need to be minimized to ensure patient comfort. As such, this would be an interesting future direction that could determine the best practice for grey matter diffusion imaging. Such studies are best initiated using simulations to better control the many variables and more robustly test the specific hypothesis. This will require careful consideration of the structure of the simu-

lated data, and how best to reflect the full range of cortical microenvironments.

CONCLUSION

Architecturally driven cortical parcellation remains in avid pursuit in the field of neuroscience. Such tools will eventually facilitate our burgeoning understanding of the structure-function relations that drive complex brain processes. *In vivo* Brodmann mapping could even provide insight into the aetiology of developmental disorders, such as Autism, which are currently poorly understood but suspected to relate, in part, to cortical organisation (Blatt 2012). This Thesis aimed to establish the extent to which diffusion MRI can be utilised in cortical parcellation applications and what the best practices are in this domain. We hoped to achieve this by using machine learning techniques to evaluate several aspects of a cortical parcellation pipeline with high quality, 3T datasets, at *in vivo* resolutions.

In Chapter 5 we demonstrated that diffusion-based measurements provide area-specific contrast that could potentially improve the performance of future parcellation studies. We adopted unsupervised learning, via the k-means algorithm, and several refinements to a parcellation pipeline, including population averaging, and generated a cortical map. The results exhibited several spatially coherent clusters in regions corresponding to well known functional anatomical areas. The feature set was most discriminative in primary areas that were consistently located across subjects, such as S1 and M1. This supported previous findings that have also demonstrated structural difference between these two regions using diffusion MRI (Anwander et al. 2010; McNab et al. 2013; Calamante et al. 2017). However, we additionally observed clusters that may correspond to non-primary areas, such as area 44 and 45, and the several clusters in the myelin poor prefrontal cortex. We reported moderately low correlation between the dMRI features and myelin density measurements derived from the T1w/T2w ratio. These findings lead us to conclude that the dMRI features provided rich information content and could be useful in future cortical studies.

In Chapter 6 we introduced supervised learning and further demonstrated the utility of diffusion MRI in cortical parcellation applications. By developing a neighbourhood

specific classification pipeline, that employed random forest classification, we attempted to distinguish a full range of cortical areas against their neighbours. We were able to classify the majority of areas (125/180) with a higher than chance outcome. As in Chapter 5, the diffusion-based features were highly effective in primary sensorimotor areas. We also observed high classification accuracy in many additional areas including the auditory core, middle temporal area (MT), and several other secondary and tertiary functional regions. However, we found only moderately good reproducibility of area V1, again, this was also reflected in the results of Chapter 5; suggesting the spatial resolution of the data was insufficient in this particularly thin region. Overall the results indicated that while diffusion MRI can characterise a broad range of cortical areas, it cannot currently be used to reliably parcellate the entire cortex at *in vivo* 3T resolutions.

Aside from generally establishing the utility of dMRI as a measure of cortical microstructure, we also set out to investigate more specific methodological questions. For example, given that there are copious different ways in which diffusion can be used to probe microstructure, what is the most appropriate technique for grey matter? To answer this question, our classification experiments in Chapter 6 were used to compare several different methods. Feature sets were derived from the diffusion tensor (DT), a more specific compartment model, NODDI, our higher-order spherical harmonic (SH) technique and a set of 4-tensor (4T) invariants. The results demonstrated that higher-order decomposition methods provide a more consistent characterisation of grey matter microenvironments in regions for which the training labels could be considered most reliable. However, even simple lower order models such as the diffusion tensor provide contrast between cortical areas; in particular, combining the traditional diffusion tensor metrics of FA and MD with the surface specific radiality index was very powerful in binary classification between M1 and S1. This suggests that using surface-specific features that describe tissue properties within the local frame of reference, as indexed by the surface normal, are likely to provide the most discriminative feature sets.

In Chapter 7 we investigated b-values, a very important parameter of diffusion MRI experiments. Again by using objective classification tasks, we were able to demonstrate that acquiring multiple b-values is beneficial for cortical imaging. The improvement in classification accuracy was consistent in both the 3T HCP data, and 1.5T local data, which more

closely resembles what is currently tenable in clinical applications. These results should be considered when acquiring data for future cortical applications.

In this work we robustly demonstrated the utility of diffusion in cortical MRI studies by employing objective machine learning approaches. Our findings motivate further investigation of this modality and encourage future investigators to include diffusion measurements alongside myelin mapping or other imaging modalities. It is unlikely that a single set of areas will ever emerge as the definitive and final parcellation of the cerebral cortex. Instead, as the study of this complex system continues, the community is likely to land on various different maps, each reflecting different contrasts, e.g., cytoarchitecture, myeloarchitecture, functional or structural connectivity. There is even potential for disagreement on what constitutes a cortical transition. In fact, we have already witnessed several Brodmann areas being further subdivided as new methods of observation were developed (Geyer et al. 1996; Geyer et al. 1997; Geyer et al. 2000; Orban et al. 2004; Sereno and Tootell 2005; Wandell et al. 2007), as well as the definition of new regions, such as the fusiform face area - defined by its BOLD activity rather than its architecture (Kanwisher et al. 1997). We expect that with advances in data acquisition methods, surface-based analysis of grey matter diffusion will become a new powerful tool in the large and varied arsenal that will help eventually uncover the synergy between structure and function in the human neocortex.

8.1 FUTURE WORK

We presented findings using exemplar 3T HCP datasets; however, even this data offers low spatial resolution in relation to the thickness of the cortical ribbon. This is major limiting factor of the above work. Further work at higher resolutions and improved SNR will likely enhance the performance of these methods. For example, one could utilize the 7T HCP datasets, as discussed in Chapter 6, or even higher resolution, fixed tissue datasets. Such data could provide a valuable platform for validating the methods proposed in this Thesis and reveal, with more certainty, which of the cortical areas contain architectonic features that manifest in the dMRI signal. Furthermore, We only sampled data at the middle cortical depth using equidistant sampling. This decision was taken to minimize partial volume effects. With high resolution 7T data it may be possible to extend our methods to a laminar-like analysis

pipeline, and possibly even adhere to the more anatomically realistic equivolume sampling (Waehnert et al. 2016). Such an approach would more closely mimic the traditional histological approaches of Brodmann etc., (Brodmann 1909; Vogt 1919). It would be interesting to see if the trends we report, such as higher-order models performing more reliably, are maintained when using a higher resolution laminar approach.

We did not attempt to provide a transferable set of area definitions which could be adopted into future studies, such as the 180 HCP multi-modal parcellation areas. Given that diffusion imaging of grey matter is still in its infancy, this was beyond the scope of this project, which instead focussed on exploring the capabilities of this modality. Although we chose to present our unmodified, noisy, parcellation results we did discuss a method for adapting our neighbourhood based parcellation scheme to generate coherent area definitions by using a winner-takes-all approach. However, the accuracy of the training labels must be carefully considered when performing such post-processing steps. Alternatively, future studies could combine increasingly powerful machine learning approaches such as deep learning, with the surface-based pipeline that we have presented. Again the limitations of the training labels should be carefully considered, as unlike in most machine vision applications, accurately annotated training data is not readily available. Nevertheless, deep learning has produced remarkable results in a wide range of applications and could potentially improve *in vivo* parcellation prospects.

The results we presented concentrated on one hemisphere, but there are many interesting questions that will be able to be addressed in future studies where homologous regions are compared across hemispheres, given the well-documented differences in function between them. Similarly, future cortical parcellation pipelines could assess group level differences between healthy and diseased populations. It remains to be seen if differences can be observed, for example, when comparing two group average parcellations (similar to chapter 5, or 6) from different cohorts. We only reported reliable contrast in a few cortical areas using this approach, so it is very unlikely that clinically available data, acquired at much lower gradient strengths and scan times, would have sufficient contrast and resolution to perform such analyses. One future approach could be to employ sophisticated super resolution, methods, such as image quality transfer (Alexander et al. 2017), to enhance low quality diffusion datasets and test clinical hypotheses. Once established, *in vivo* cortical par-

cellation pipelines will hold enormous potential diagnostic potential in neurological disorders. This should be the ultimate goal of future pipelines.

Part IV

APPENDIX

EVALUATION OF CONCURRENT FIELD MONITOR

In the main body of this thesis we explored several aspects of diffusion-based cortical parcellation. However diffusion data could not be utilized in this domain without the recent advancements in hardware and pre-processing that have led to increased image quality. Diffusion weighted data acquisitions are notoriously susceptible to artifacts due to the long lasting eddy currents that are introduced by the diffusion encoding gradients. These eddy currents effect the imaging gradients of the EPI readout train causing spatial distortions in the DWIs. These artefacts are commonly dealt with using complex post processing algorithms (Jenkinson and Smith 2001; Andersson and Skare 2002), which may also produced unwanted smoothing effects on the underlying data (Graham et al. 2016) or have time consuming aquisition requirements such as the reverse phase encoding requirements of eddy (Andersson and Sotiropoulos 2016).

An alternative way to eliminate eddy current distortions is to use magnetic field probes (De Zanche et al. 2008) to measure the dynamic field variations and incorporate this information into the image reconstruction pipeline (Barmet et al. 2008; Wilm et al. 2015). In this work we investigated the effects of a concurrent field monitoring system on HARDI data and applied the improved HARDI data to cortical parcellation. We showed that the field monitoring improved the HARDI data quality especially in anterior/posterior poles of the brain and air tissue interfaces. This regional improvement was also clear in the cortical classification results, where the field monitoring improved the accuracy of the V1/V2 classification by 3%, compared to 0.5% in the motorstrip.

My contributions to this work were the analysis of the diffusion tensor model residuals using Explore DTI and performing the cortical parcellation pipeline with binary classification tests between V1/V2 and M1/S1. Data acquisition, reconstruction, motion correction, DWI analysis and DT model fitting were conducted by the co-authors.

This work has been previously presented in:

Lee, Y. et al. (2017). "Investigating the effects of concurrent magnetic field monitoring on high angular resolution diffusion imaging: application to cortical parcellation." In: Proceedings of the 25th Annual Meeting of International Society for Magnetic Resonance Imaging, Honolulu, 2017.

METHODS

Data Type	TE (ms)	TR (ms)	Flip Angle	Acquisition Resolution (mm ³)	Reconstructed Resolution (mm ³)	SENSE/ GRAPPA
HARDI	65	7500	90	2.0x2.0x2.0	1.3x1.3x2.0	3
MPRAGE	3.57	8.4	7	1.0	1.0	none

Table A.1: Acquisition parameters for HARDI and high-resolution MPRAGE image volumes.

DATA AND PROCESSING

DATA ACQUISITION Using a 3T Philips Achieva scanner and 8 channel receive only head coil we collected HARDI data with 1 $b=0$ s/mm² image and 61 DWIs ($b=1000$ s/mm²) on a healthy adult male volunteer under local ethics approval. Full Acquisition parameters are given in Table A.1. Concurrently, the imaging gradient fields were measured by 16 magnetic field probes distributed around the head.

IMAGE RECONSTRUCTION The HARDI data were reconstructed offline in two different ways:

1. The 0th and 1st order trajectory from the first slice of b_0 image was estimated from the field probe data and used to reconstruct the b_0 and HARDI data (HARDI^{1st}).

2. Trajectories up to 3rd order were individually estimated for each slice of the b0 image and DWIs and used to correct for the higher order field fluctuations and eddy currents from diffusion gradients (HARDI^{3rd}).

The prototype field monitor system resulted in a long acquisition time, leading to unavoidable subject head motion between DWIs. As such, the two HARDI datasets were re-aligned for subject motion separately within Explore DTI.

DATA ANALYSIS Three different processing pipelines were used to investigate the improvement in the HARDI^{3rd} data quality as compared to HARDI^{1st}:

1. We calculated the standard deviation (SD) across the diffusion encoding direction and compared the signal intensities across the DWIs for individual voxels.
2. We calculated the voxel-wise residuals after fitting the diffusion tensor model to the data.
3. We carried out cortical area classification. Here, the DT parameters from each HARDI dataset were sampled onto an existing FreeSurfer surface tessellation of the same volunteer, resulting in a feature vector at each vertex comprised of FA, MD, RI. We then extracted regions of interest, corresponding to the primary motor (M1), somatosensory (S1) and visual (V1) areas and secondary (V2) visual area using the FreeSurfer Brodmann Atlas labels. In addition, two ROI corresponding to the M1 and S1 hand subdivisions, were generated from the group average somatotopic data of 20 subjects. To assess whether improvements in data quality provide higher classification accuracies, binary classification experiments of the above regions were performed, using random forest classification with leave one out cross validation.

RESULTS

Figure A.1 shows the percentage difference map calculated between the HARDI^{1st} and HARDI^{3rd} datasets for (a) a $b = 0 \text{ s/mm}^2$ image slice and (b) a $b = 1000 \text{ s/mm}^2$ image slice. The signal intensity across the changing diffusion encoding directions of a voxel near the subcortical gray/white matter boundary is also shown in (c). We observe differences of

$\pm 10\%$ across the $b=1000\text{s/mm}^2$ image slice and differences between the signal intensities of the two datasets at several diffusion directions for the elected voxel.

Figure A.2 shows the difference between HARDI^{1st} and HARDI^{3rd} standard deviation maps for 6 different slices and the intensity variations from example voxels with large difference, e.g., edge of the brain and air-tissue interface. The differences in signal intensity between the two datasets for these voxels is markedly more pronounced than that of the subcortical voxel in Figure A.1.

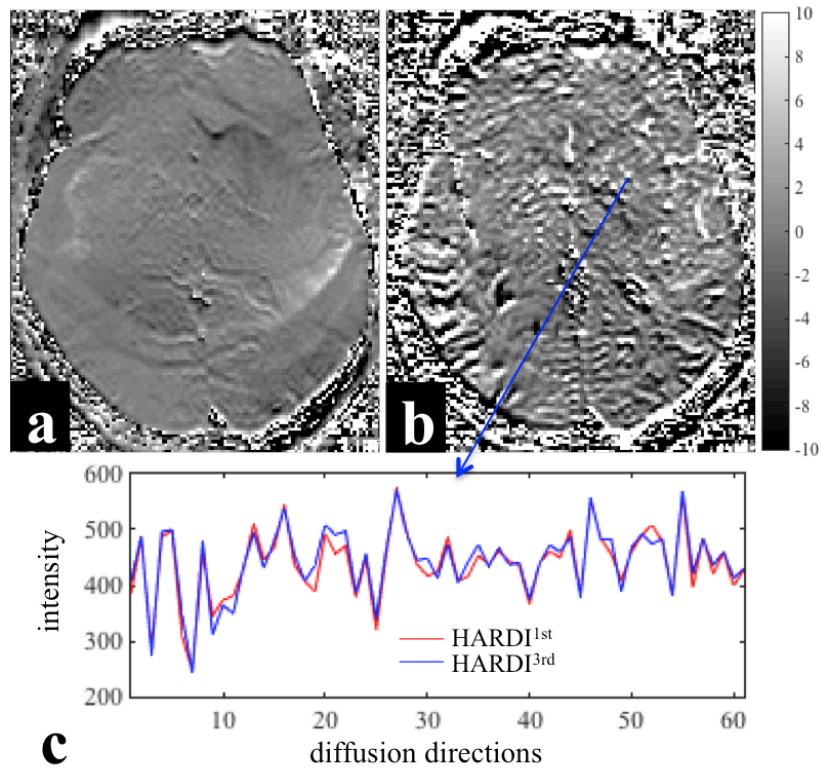


Figure A.1: The percentage difference between HARDI^{1st} and HARDI^{3rd} for (a) $b = 0$ and (b) $b=1000\text{ s/mm}^2$ images at diffusion direction vector (0.87,0.038,0.033). (c) The voxel signal intensity across all diffusion directions for HARDI^{1st} (red) and HARDI^{3rd} (blue) datasets.

The median \pm interquartile ranges of voxel-wise absolute residuals for voxels containing brain tissue are shown in Figure A.3 for each diffusion direction. HARDI^{1st} has larger mean residuals than HARDI^{3rd} for 45 of the 60 diffusion directions. The accuracies of the binary classification experiments are given in Table A.2 as the percentage of correctly clas-

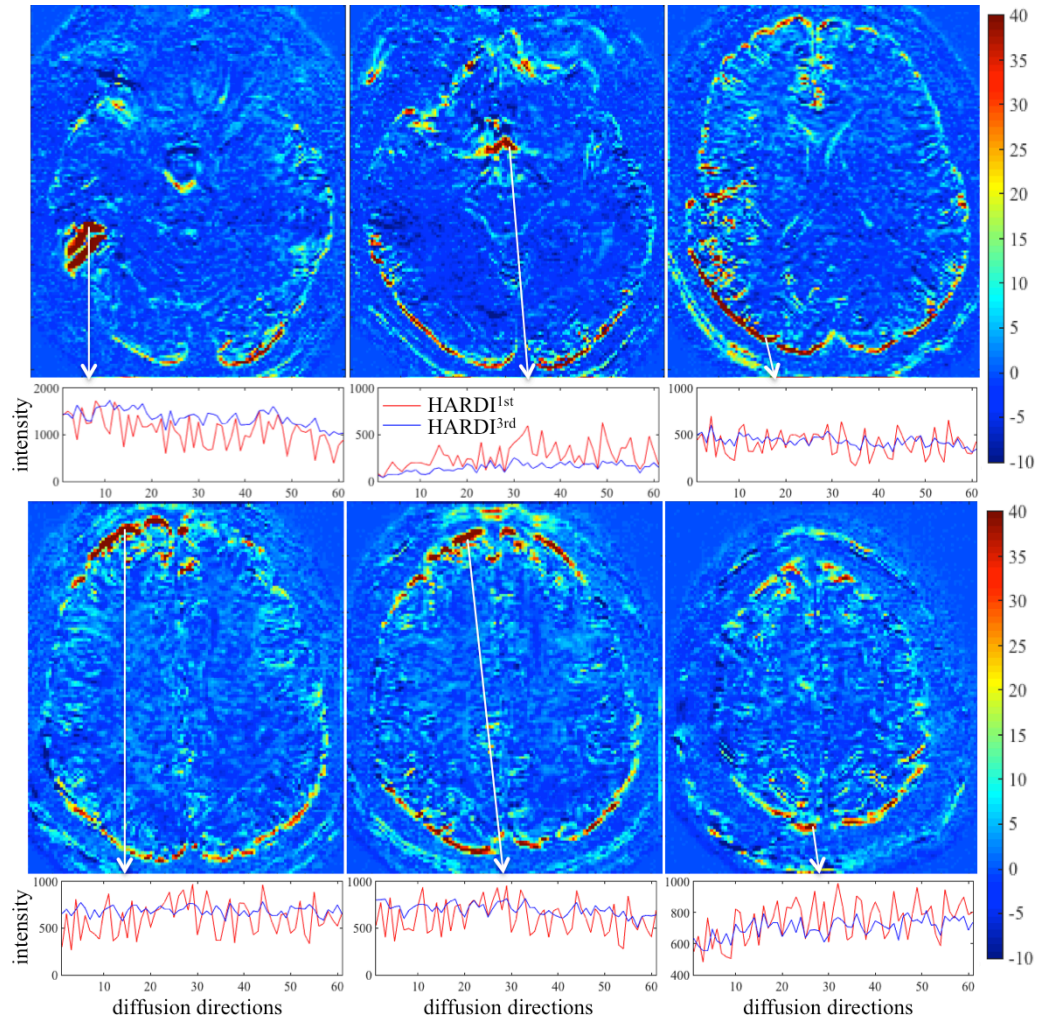


Figure A.2: The difference in the SD maps of the HARDI^{1st} and HARDI^{3rd} datasets at 6 different image slices. Below each difference map the signal intensity across all diffusion directions is displayed for a representative voxel.

sified vertices. HARDI^{3rd} provides an approximate improvement of only 0.5% in the classification of M1 and S1, for both hemispheres, even within the smaller, less variable hand subdivisions. The improvement in classification between V1 and V2 is larger, between 2-3%. This reflects the results of the SD maps (Figure A.2) which show that the largest differences between the two datasets is within anterior and posterior voxels.

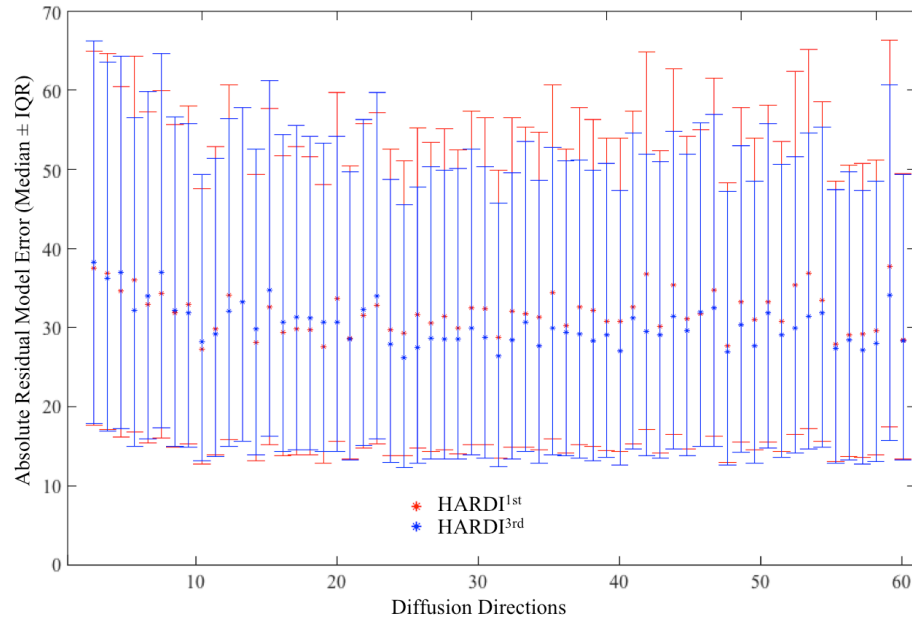


Figure A.3: Median \pm interquartile range of voxel-wise absolute diffusion tensor model residuals across the whole brain.

ROI	Left Hemisphere		Right Hemisphere	
	HARDI1st	HARDI3rd	HARDI1st	HARDI3rd
M1/S1	88.96	89.47	81.00	81.58
V1/V2	66.83	68.01	64.21	67.21
M1 hand/ S1 hand	n/a	n/a	67.96	67.25

Table A.2: The percentage of correctly classified vertices in each binary classification test for the HARDI^{1st} and HARDI^{3rd} datasets.

DISCUSSION

Using higher order image reconstruction, as afforded by the data collected with magnetic field probes, has benefits for HARDI data quality. The benefits are regionally specific with the anterior/posterior poles of the brain and air-tissue interface regions benefiting the most. This regional effect is again clear in the cortical classification results, where the use of the field probe improves the accuracy of V1/V2 classification by 3%, compared to 0.5% in the

motor strip.

In voxels containing white matter, the changing diffusion directions modulate the signal extensively, therefore, calculating the SD may not be a good measure of data quality. The individually subtracted DWIs between HARDI^{1st} and HARDI^{3rd} highlight this aspect (Figure A.1). For some diffusion directions the error can be up to 8%.

Note that image artifacts not related to the diffusion-encoding gradients are already corrected in all images through the use of 0th and 1st order trajectory. This step was taken to correct for a significant amount of head motion during the long scan time. One corollary of this processing step may have been a reduction in eddy artefacts; therefore further testing of the final hardware system, which does not require excessive scan times, will be necessary to confirm the true extent of data quality improvement that the field monitor provides without additional post processing.

Future work will also need to compare concurrent field monitoring against the current state-of-the-art processing pipelines, such as Eddy. In addition the effects of the data quality improvement in other applications, such as tractography and group-level analyses needs to be investigated. Both of these additional avenues will help determine if the upfront cost of field monitoring systems is a worthy investment for imaging centres.

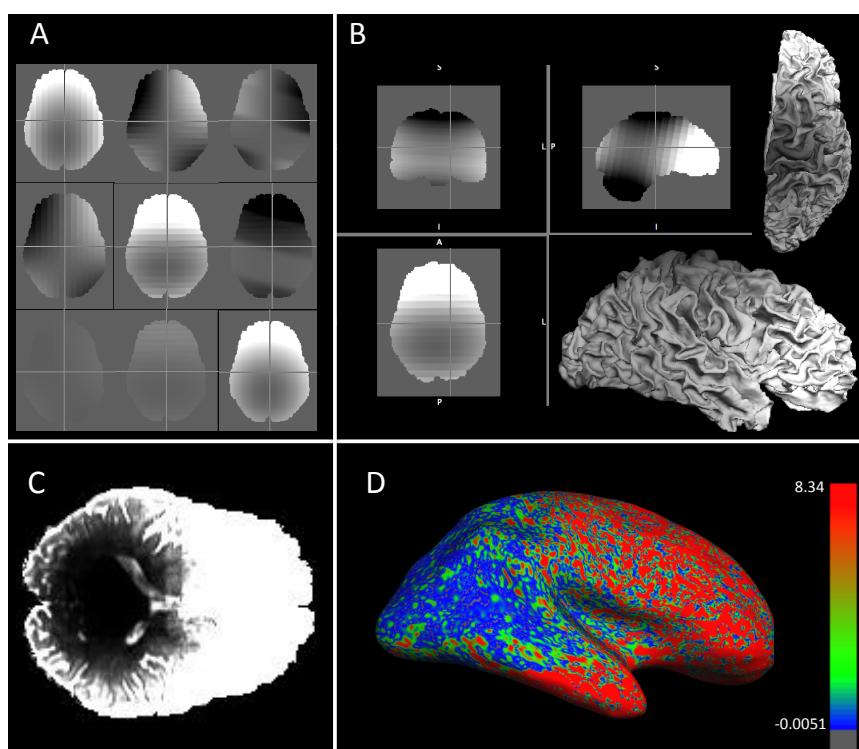


Figure B.1: HCP diffusion data gradient deviation: analysis and correction. (A) 9-component gradient deviation data, shown for a single image slice. The offsets range from +15% (white) to -15% (black), and vary spatially. A corrective function was implemented in *camino* to modify the gradient information at each voxel by applying the 3-dimensional adjustment encoded in the 9-components. (B) Demonstration that the gradient deviation data is correctly sampled onto the cortical surface. As the cortical parcellation pipeline performs model fitting after surface sampling, the gradient deviation data needed to be sampled accurately to the surface vertices. (C) Sum of squared differences (SSD) between a diffusion tensor component fit with and without the gradient correction applied. Cortical voxels are the generally the most severely effected. (D) SSD between the a spherical harmonic feature calculated with and without the gradient correction.

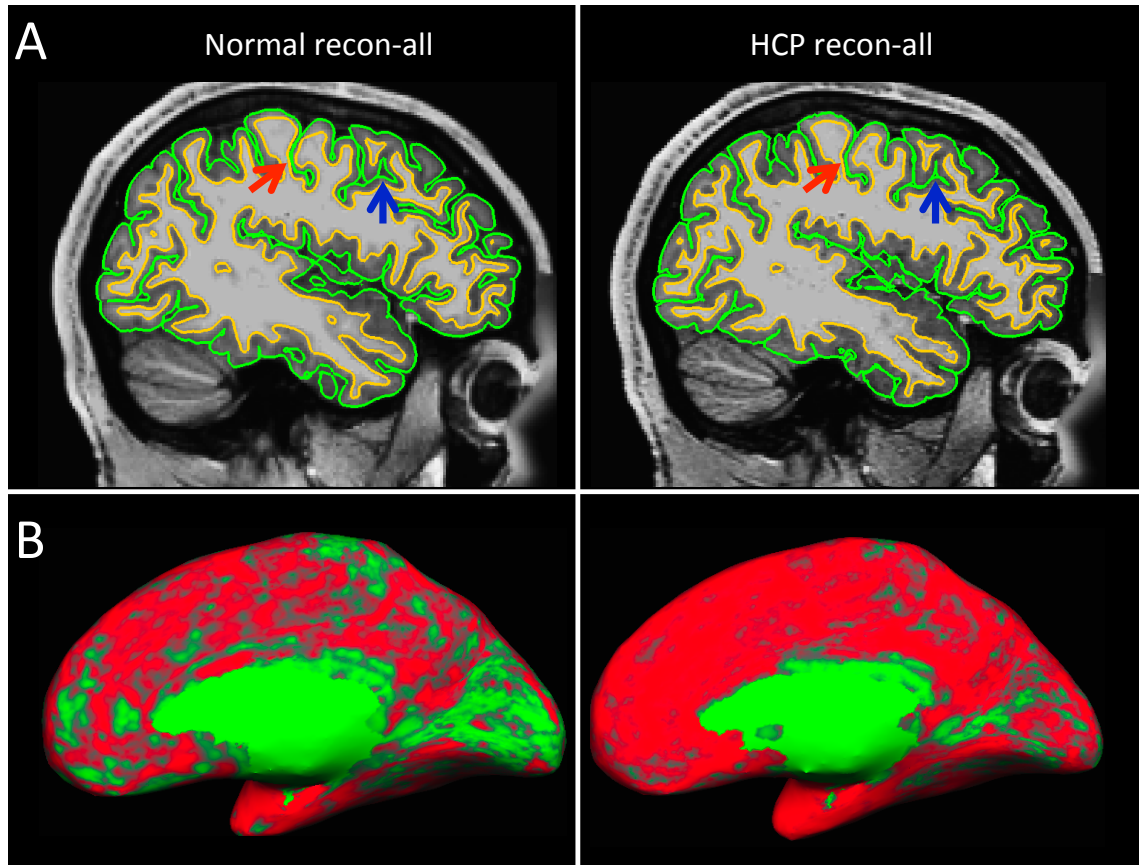


Figure B.2: Comparison of the normal FreeSurfer recon-all surface estimation (left) to the HCP surface estimation (right). The HCP version is reported to provide better surface estimation due to two improvements. (1) It makes use of the high resolution datasets, whereas the traditional recon-all algorithm down samples data to 1mm^3 isotropic. (2) The HCP version leveraged information in both the T1w and T2w images. (A) Shows the bounding lines of the white matter (yellow) and pial (green) surface estimations. In general, the HCP version makes a more conservative estimate of the WM surface, (e.g., red arrow) and a less conservative estimate of the pial surface (e.g., blue arrow). (B) shows the cortical thickness estimates for both algorithms, painted using the same scale. Overall, the HCP algorithm estimates a thicker cortical ribbon.

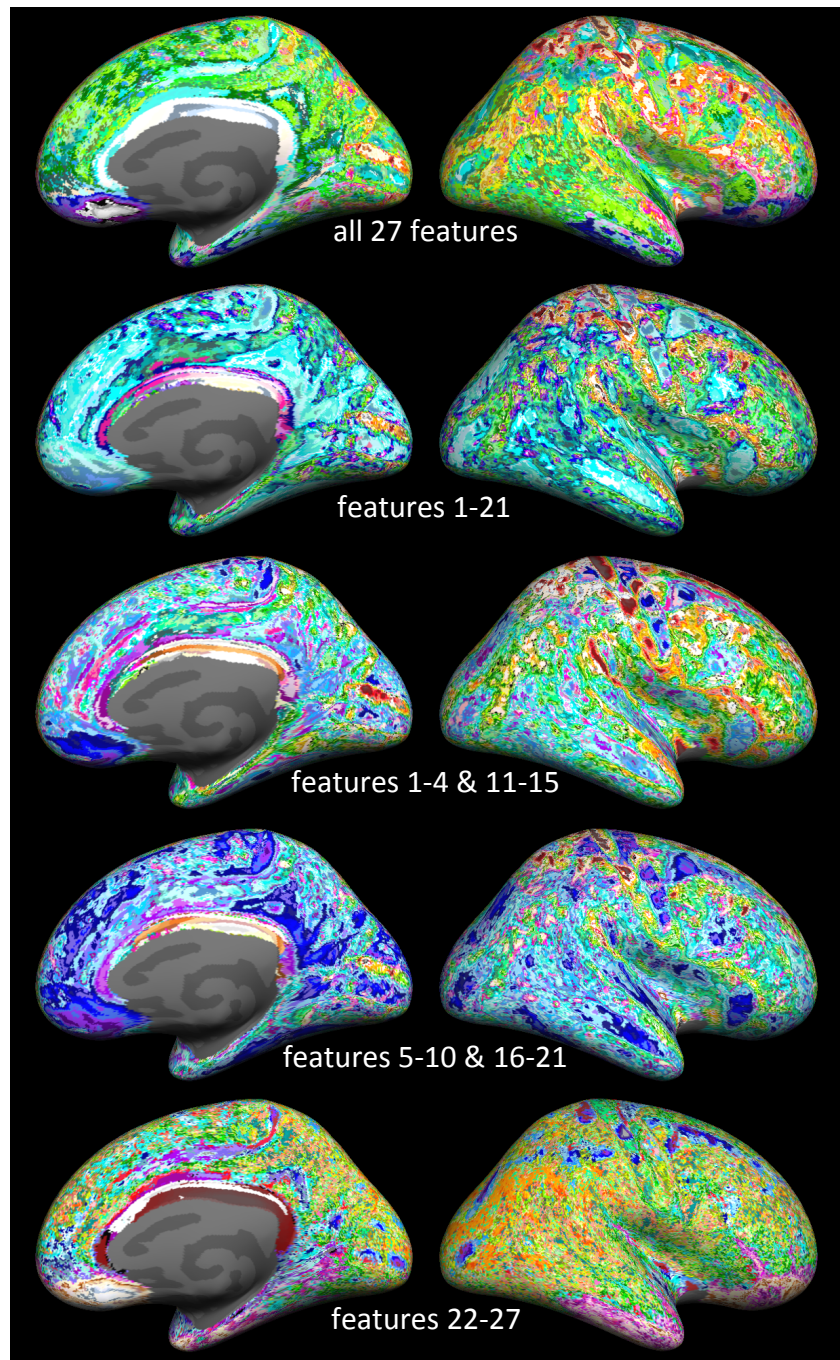


Figure B.3: Group average cortical parcellations using different combinations of spherical harmonic features. Clusters are ordered within each parcellation such that similar clusters are displayed in alike colours. The difference in cluster colours between the different parcellations is arbitrary. Features numbers match those given in chapter 3. Qualitative assessment of these parcellation results demonstrate that none of the combinations provide as coherent delineations of the areas 3b and 4 as the feature set presented in Chapter 5. The results for features 22-27 are particularly noisy and contain very little area-like clusters, suggesting that these features are unlikely to add useful information for cortical area classification.

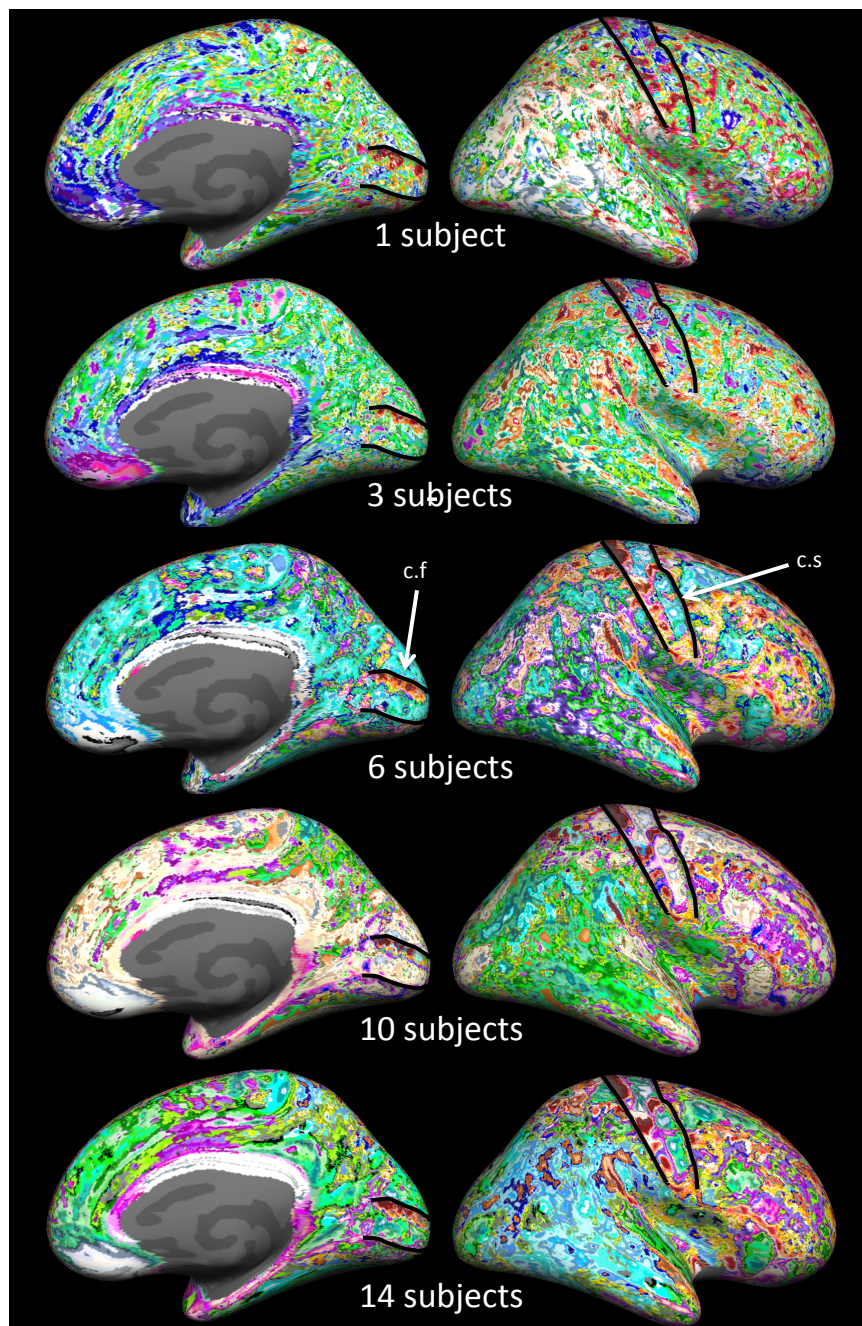


Figure B.4: Unsupervised cortical parcellations using different group sizes. All results were generated using the five features presented in Chapter 5. Contours show the borders of the central sulcus (c.s) and the calcarine fissure (c.f). Clusters corresponding to M-I can be observed by using as few as six subjects. All the results show a distinct cluster at the upper vertical meridian of V1.

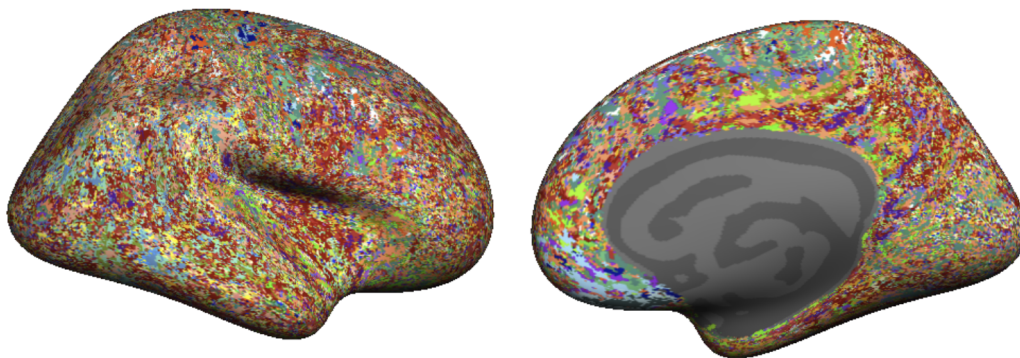


Figure B.5: Alternative averaging method in which the mode cluster is selected after classification, rather than surface based averaging of the feature vectors prior to classification. The result was generated using the 27 spherical harmonic features, and k-means clustering with $k=150$.

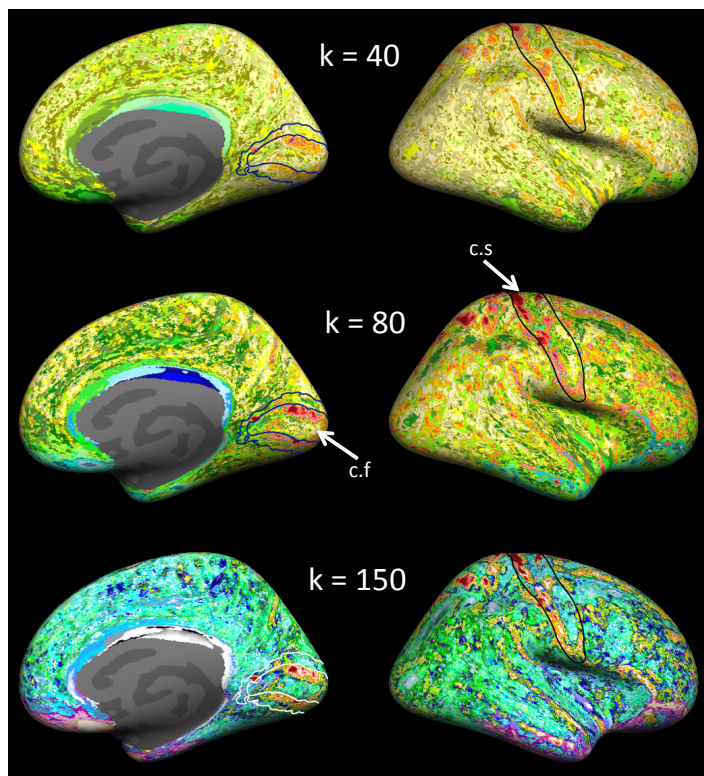


Figure B.6: K-means clustering results using different values of k . Results all generated with a 6-subject average and the full spherical harmonic feature set, containing 27 features. c.s points to the central sulcus and c.f. points to the calcarine fissure.

CHAPTER 6 SUPPLEMENTARY MATERIAL

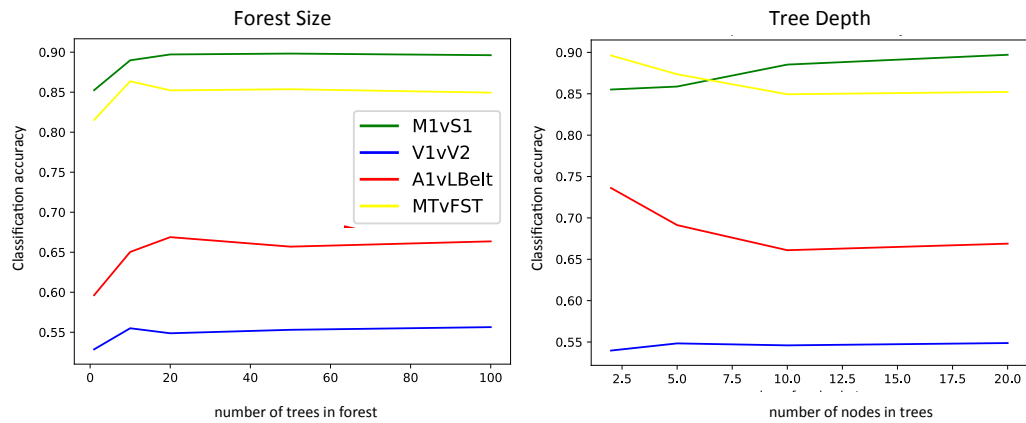


Figure C.1: Optimisation of random forest parameters for whole hemisphere parcellation experiments. The left shows the change in accuracy as the number of trees are increased with a fixed tree depth. Accuracy improvements are generally small and plateau after 10-20 trees for all of the tested areas, as such 20 trees was deemed more than sufficient for subsequent classifications. The right shows the change in accuracy as the tree depth is increased at fixed forest size. A1 vs LBelt was the only test that showed larger differences, with accuracy dropping off as depth increase, suggesting that overfitting may be an issue for some areas. a tree depth of 5 was selected for subsequent classifications

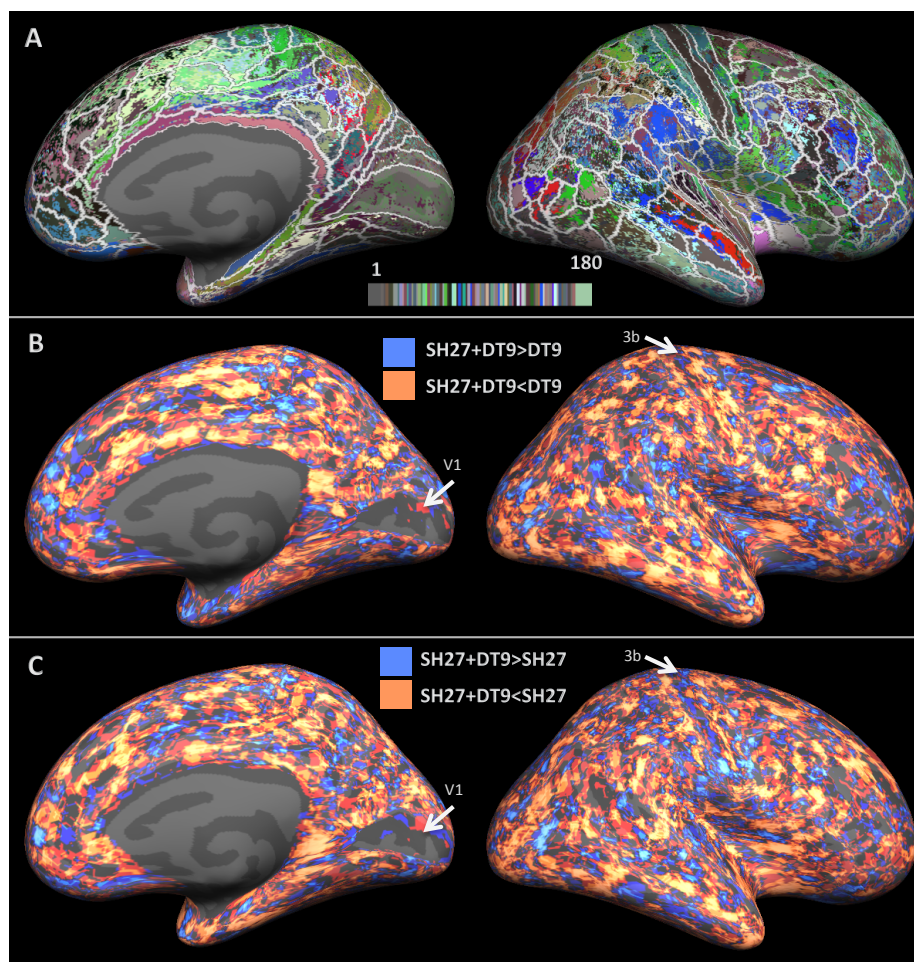


Figure C.2: Group average classification results for the combination of SH27 and DT9 feature sets.

(A) The parcellation result with the same colour assignment as Figure 6.3B. Again, the HCP training labels have been overlaid in white. (B) The searchlight cluster coherency comparison between the combined feature set and the DT9 feature set. (C) The same as B but with comparison to the SH27 feature set. In both B and C, a lower cluster count indicates a more spatially coherent parcellation, therefore orange (lower value in SH27+DT9) indicates that the combined feature set performed better.

BIBLIOGRAPHY

- Aggarwal, Manisha et al. (2015). "Probing region-specific microstructure of human cortical areas using high angular and spatial resolution diffusion MRI." In: *Neuroimage* 105, pp. 198–207.
- Alexander, DC, GJ Barker, and SR Arridge (2002). "Detection and modeling of non-Gaussian apparent diffusion coefficient profiles in human brain data." In: *Magnetic Resonance in Medicine* 48.2, pp. 331–340.
- Alexander, Daniel C and Gareth J Barker (2005). "Optimal imaging parameters for fiber-orientation estimation in diffusion MRI." In: *Neuroimage* 27.2, pp. 357–367.
- Alexander, Daniel C et al. (2010). "Orientationally invariant indices of axon diameter and density from diffusion MRI." In: *Neuroimage* 52.4, pp. 1374–1389.
- Alexander, Daniel C et al. (2017). "Image quality transfer and applications in diffusion MRI." In: *NeuroImage* 152, pp. 283–298.
- Amunts, Katrin and Karl Zilles (2015). "Architectonic mapping of the human brain beyond Brodmann." In: *Neuron* 88.6, pp. 1086–1107.
- Amunts, Katrin et al. (1999). "Broca's region revisited: cytoarchitecture and intersubject variability." In: *Journal of Comparative Neurology* 412.2, pp. 319–341.
- Amunts, Katrin et al. (2000). "Brodmann's areas 17 and 18 brought into stereotaxic space — where and how variable?" In: *Neuroimage* 11.1, pp. 66–84.
- Amunts, Katrin et al. (2013). "BigBrain: an ultrahigh-resolution 3D human brain model." In: *Science* 340.6139, pp. 1472–1475.
- Andersson, Jesper LR and Stefan Skare (2002). "A model-based method for retrospective correction of geometric distortions in diffusion-weighted EPI." In: *Neuroimage* 16.1, pp. 177–199.
- Andersson, Jesper LR and Stamatios N Sotiropoulos (2016). "An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging." In: *Neuroimage* 125, pp. 1063–1078.

- Anwander, A, A Pampel, and TR Knosche (2010). "In vivo measurement of cortical anisotropy by diffusion-weighted imaging correlates with cortex type." In: *Proc. Int. Soc. Magn. Reson. Med.* Vol. 18, p. 109.
- Anwander, Alfred et al. (2006). "Connectivity-based parcellation of Broca's area." In: *Cerebral cortex* 17.4, pp. 816–825.
- Armitage, PA and ME Bastin (2001). "Utilizing the diffusion-to-noise ratio to optimize magnetic resonance diffusion tensor acquisition strategies for improving measurements of diffusion anisotropy." In: *Magnetic resonance in medicine* 45.6, pp. 1056–1065.
- Assaf, Yaniv and Peter J Basser (2005). "Composite hindered and restricted model of diffusion (CHARMED) MR imaging of the human brain." In: *Neuroimage* 27.1, pp. 48–58.
- Assaf, Yaniv et al. (2008). "AxCaliber: a method for measuring axon diameter distribution from diffusion MRI." In: *Magnetic resonance in medicine* 59.6, pp. 1347–1354.
- Augustinack, Jean C et al. (2005). "Detection of entorhinal layer II using Tesla magnetic resonance imaging." In: *Annals of neurology* 57.4, pp. 489–494.
- Barbier, Emmanuel L et al. (2002). "Imaging cortical anatomy by high-resolution MR at 3.0 T: Detection of the stripe of Gennari in visual area 17." In: *Magnetic resonance in medicine* 48.4, pp. 735–738.
- Barmet, Christoph, Nicola De Zanche, and Klaas P Pruessmann (2008). "Spatiotemporal magnetic field monitoring for MR." In: *Magnetic Resonance in Medicine* 60.1, pp. 187–197.
- Barpoutis, Angelos and Baba C Vemuri (2010). "A unified framework for estimating diffusion tensors of any order with symmetric positive-definite constraints." In: *Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on.* IEEE, pp. 1385–1388.
- Barpoutis, Angelos et al. (2009). "Regularized positive-definite fourth order tensor field estimation from DW-MRI." In: *NeuroImage* 45.1, S153–S162.
- Basser, Peter J and Carlo Pierpaoli (2011). "Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI." In: *Journal of magnetic resonance* 213.2, pp. 560–570.
- Basser, Peter J, James Mattiello, and Denis LeBihan (1994). "MR diffusion tensor spectroscopy and imaging." In: *Biophysical journal* 66.1, pp. 259–267.

- Bastiani, Matteo et al. (2016). "Automatic segmentation of human cortical layer-complexes and architectural areas using ex vivo diffusion MRI and its validation." In: *Frontiers in neuroscience* 10.
- Beaulieu, C (2011). "What makes diffusion anisotropic in the nervous system." In: *Diffusion MRI: Theory, methods, and applications*. Oxford University Press Oxford, England, pp. 92–109.
- Beaulieu, Christian (2002). "The basis of anisotropic water diffusion in the nervous system—a technical review." In: *NMR in Biomedicine* 15.7-8, pp. 435–455.
- Beckmann, Matthias, Heidi Johansen-Berg, and Matthew FS Rushworth (2009). "Connectivity-based parcellation of human cingulate cortex and its relation to functional specialization." In: *Journal of Neuroscience* 29.4, pp. 1175–1190.
- Behrens, Timothy EJ et al. (2003). "Characterization and propagation of uncertainty in diffusion-weighted MR imaging." In: *Magnetic resonance in medicine* 50.5, pp. 1077–1088.
- Berlin, Rudolf (1858). *Beitrag zur structurlehre der grosshirnwindungen*. AE Junge'sche universitätsbuchdr.
- Betz, W (1874). "Anatomischer nachweis zweier gehirncentra." In: *Zentralbl Med Wiss* 12.578,595.
- Bishop, Christopher M (2006). *Pattern recognition and machine learning*. springer.
- Blatt, Gene J (2012). "The neuropathology of autism." In: *Scientifica* 2012.
- Bloch, Felix (1946). "Nuclear induction." In: *Physical review* 70.7-8, p. 460.
- Blumensath, Thomas et al. (2013). "Spatially constrained hierarchical parcellation of the brain with resting-state fMRI." In: *Neuroimage* 76, pp. 313–324.
- Bock, Nicholas A et al. (2009). "Visualizing the entire cortical myelination pattern in marmosets with magnetic resonance imaging." In: *Journal of neuroscience methods* 185.1, pp. 15–22.
- Bok, Siegfried Thomas (1929). "Der Einfluß der in den Furchen und Windungen auftretenden Krümmungen der Großhirnrinde auf die Rindenarchitektur." In: *Zeitschrift für die gesamte Neurologie und Psychiatrie* 121.1, pp. 682–750.
- Breiman, Leo (2001). "Random forests." In: *Machine learning* 45.1, pp. 5–32.
- Bridge, Holly et al. (2005). "Independent anatomical and functional measures of the V1/V2 boundary in human visual cortex." In: *Journal of Vision* 5.2, pp. 1–1.

- Bridge, Holly, Stuart Clare, and Kristine Krug (2014). "Delineating extrastriate visual area MT (V5) using cortical myeloarchitecture." In: *Neuroimage* 93, pp. 231–236.
- Broca, Paul (1861). "Loss of speech, chronic softening and partial destruction of the anterior left lobe of the brain." In: *Bull Soc Anthropol* 2.1, pp. 235 –238.
- Brodmann, Korbinian (1909). *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. Barth.
- Brown, Robert (1828). "XXVII. A brief account of microscopical observations made in the months of June, July and August 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies." In: *Philosophical Magazine Series 2* 4.21, pp. 161–173.
- Cajal, Santiago Ramón y (1900). *Studien über die Hirnrinde des Menschen*. Vol. 1. Johann Ambrosius Barth.
- Calamante, Fernando et al. (2017). "The role of whole-brain diffusion MRI as a tool for studying human in vivo cortical segregation based on a measure of neurite density." In: *Magnetic Resonance in Medicine*.
- Campbell, Alfred Walter (1905). *Histological studies on the localisation of cerebral function*. University Press.
- Carmichael, David W et al. (2006). "Improving whole brain structural MRI at 4.7 Tesla using 4 irregularly shaped receiver coils." In: *Neuroimage* 32.3, pp. 1176–1184.
- Caruyer, Emmanuel et al. (2013). "Design of multishell sampling schemes with uniform coverage in diffusion MRI." In: *Magnetic resonance in medicine* 69.6, pp. 1534–1540.
- Chi, Je G, Elizabeth C Dooling, and Floyd H Gilles (1977). "Gyral development of the human brain." In: *Annals of neurology* 1.1, pp. 86–93.
- Clark, Chris A and Denis Le Bihan (2000). "Water diffusion compartmentation and anisotropy at high b values in the human brain." In: *Magnetic Resonance in Medicine* 44.6, pp. 852–859.
- Clark, Vincent P, Eric Courchesne, and Marjorie Grafe (1992). "In vivo myeloarchitectonic analysis of human striate and extrastriate cortex using magnetic resonance imaging." In: *Cerebral Cortex* 2.5, pp. 417–424.
- Cohen-Adad, Julien (2014). "What can we learn from T2* maps of the cortex?" In: *Neuroimage* 93, pp. 189–200.

- Conturo, Thomas E et al. (1999). "Tracking neuronal fiber pathways in the living human brain." In: *Proceedings of the National Academy of Sciences* 96.18, pp. 10422–10427.
- Cook, Philip A et al. (2007). "Optimal acquisition orders of diffusion-weighted MRI measurements." In: *Journal of magnetic resonance imaging* 25.5, pp. 1051–1058.
- Criminisi, Antonio and Jamie Shotton (2013). *Decision forests for computer vision and medical image analysis*. Springer Science & Business Media.
- Dale, Anders M, Bruce Fischl, and Martin I Sereno (1999). "Cortical surface-based analysis: I. Segmentation and surface reconstruction." In: *Neuroimage* 9.2, pp. 179–194.
- De Zanche, Nicola et al. (2008). "NMR probes for measuring magnetic fields and field dynamics in MR systems." In: *Magnetic Resonance in Medicine* 60.1, pp. 176–186.
- Deoni, J and J Jones (2006). "Time–Series Analysis of the Diffusion Weighted Signal as a Model–Free Approach to Segmenting Tissue." In: *Proceedings of ISMRM 14th Annual Meeting*. Vol. 2734.
- Descoteaux, Maxime et al. (2016). "False positive bundles in tractography." In: ISMRM.
- Dick, Frederic et al. (2012). "In vivo functional and myeloarchitectonic mapping of human primary auditory areas." In: *Journal of Neuroscience* 32.46, pp. 16095–16105.
- Ding, Song-Lin et al. (2017). "Comprehensive cellular-resolution atlas of the adult human brain." In: *Journal of Comparative Neurology* 525.2, pp. 407–407.
- Dinse, Juliane et al. (2015). "A cytoarchitecture-driven myelin model reveals area-specific signatures in human primary and secondary areas using ultra-high resolution in-vivo brain MRI." In: *Neuroimage* 114, pp. 71–87.
- Douek, Philippe et al. (1991). "MR color mapping of myelin fiber orientation." In: *J Comput Assist Tomogr* 15.6, pp. 923–929.
- Economo, Constantin Freiherr von and Georg N Koskinas (1925). *Die cytoarchitektonik der hirnrinde des erwachsenen menschen*. J. Springer.
- Eickhoff, Simon B et al. (2005). "A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data." In: *Neuroimage* 25.4, pp. 1325–1335.
- Eickhoff, Simon B et al. (2006a). "Testing anatomically specified hypotheses in functional imaging using cytoarchitectonic maps." In: *Neuroimage* 32.2, pp. 570–582.
- Eickhoff, Simon B et al. (2006b). "The somatotopic organization of cytoarchitectonic areas on the human parietal operculum." In: *Cerebral Cortex* 17.8, pp. 1800–1811.

- Ferizi, Uran et al. (2015). "White matter compartment models for in vivo diffusion MRI at 300mT/m." In: *NeuroImage* 118, pp. 468–483.
- Ferrier, David (1873). "Experimental researches in cerebral physiology and pathology." In: *Journal of anatomy and physiology* 8.Pt 1, p. 152.
- (1875). "The Croonian Lecture: experiments on the brain of monkeys (second series)." In: *Philosophical Transactions of the Royal Society of London* 165, pp. 433–488.
- Fischl, Bruce (2013). "Estimating the location of Brodmann areas from cortical folding patterns using histology and ex vivo MRI." In: *Microstructural Parcellation of the Human Cerebral Cortex*. Springer, pp. 129–156.
- Fischl, Bruce, Martin I Sereno, and Anders M Dale (1999a). "Cortical surface-based analysis: II: inflation, flattening, and a surface-based coordinate system." In: *Neuroimage* 9.2, pp. 195–207.
- Fischl, Bruce et al. (1999b). "High-resolution intersubject averaging and a coordinate system for the cortical surface." In: *Human brain mapping* 8.4, pp. 272–284.
- Fischl, Bruce, Arthur Liu, and Anders M Dale (2001). "Automated manifold surgery: constructing geometrically accurate and topologically correct models of the human cerebral cortex." In: *IEEE transactions on medical imaging* 20.1, pp. 70–80.
- Fischl, Bruce et al. (2004a). "Automatically Parcellating the Human Cerebral Cortex." In: *Cerebral Cortex* 14.1, pp. 11–22.
- Fischl, Bruce et al. (2004b). "Sequence-independent segmentation of magnetic resonance images." In: *Neuroimage* 23, S69–S84.
- Fischl, Bruce et al. (2008). "Cortical folding patterns and predicting cytoarchitecture." In: *Cerebral cortex* 18.8, pp. 1973–1980.
- Fritsch, GT and Hitzig (1870). "JE Über die elektrische Erregbarkeit des Grosshirns. Arch. f. Anat. u." In: *Physiol. Wiss. Med* 370, pp. 300–332.
- Gennari, Francesco (1782). *De peculiari structura cerebri nonnullisque ejus morbis. ex regio typographeo*.
- Geyer, Stefan and Robert Turner (2015). *Microstructural parcellation of the human cerebral cortex*. Springer.
- Geyer, Stefan et al. (1996). "Two different areas within the primary motor cortex of man." In: *Nature* 382.6594, pp. 805–807.

- Geyer, Stefan, Axel Schleicher, and Karl Zilles (1997). "The somatosensory cortex of human: cytoarchitecture and regional distributions of receptor-binding sites." In: *Neuroimage* 6.1, pp. 27–45.
- Geyer, Stefan et al. (2000). "Functional neuroanatomy of the primate isocortical motor system." In: *Anatomy and embryology* 202.6, pp. 443–474.
- Geyer, Stefan et al. (2011). "Microstructural parcellation of the human cerebral cortex—from Brodmann's post-mortem map to in vivo mapping with high-field magnetic resonance imaging." In: *Frontiers in human neuroscience* 5.
- Ghosh, Aurobrata, Théodore Papadopoulos, and Rachid Deriche (2012). "Generalized Invariants of a 4th order tensor: Building blocks for new biomarkers in dMRI." In: *Proceedings of the Computation Diffusion MRI Workshop at the MICCAI Conference*.
- Glasser, MF et al. (2016). "A Multi-modal parcellation of human cerebral cortex." In: *Nature* 536, 171—178.
- Glasser, Matthew F and David C Van Essen (2011). "Mapping human cortical areas in vivo based on myelin content as revealed by T1-and T2-weighted MRI." In: *The Journal of Neuroscience* 31.32, pp. 11597–11616.
- Glasser, Matthew F et al. (2013). "The minimal preprocessing pipelines for the Human Connectome Project." In: *Neuroimage* 80, pp. 105–124.
- Glasser, Matthew F et al. (2014). "Trends and properties of human cerebral cortex: correlations with cortical myelin content." In: *Neuroimage* 93, pp. 165–175.
- Graham, Mark S, Ivana Drobnjak, and Hui Zhang (2016). "Realistic simulation of artefacts in diffusion MRI for validating post-processing correction techniques." In: *NeuroImage* 125, pp. 1079–1094.
- Grefkes, Christian et al. (2001). "Human somatosensory area 2: observer-independent cytoarchitectonic mapping, interindividual variability, and population map." In: *Neuroimage* 14.3, pp. 617–631.
- Gupta, Rakesh K et al. (2005). "Diffusion tensor imaging of the developing human cerebrum." In: *Journal of neuroscience research* 81.2, pp. 172–178.
- Hahn, Erwin L (1950). "Spin echoes." In: *Physical review* 80.4, p. 580.
- Hammarberg, Carl (1895). *Studien über Klinik und Pathologie der Idiotie, nebst Untersuchungen über die normale Anatomie der Hirnrinde*. Berling.

- Haroon, HA, RJ Binney, and GJ Parker (2010). "Probabilistic quantification of regional cortical microstructural complexity." In: *Proceedings Intl Soc Magn Res Med*, p. 578.
- Heidemann, Robin M et al. (2010). "Diffusion imaging in humans at 7T using readout-segmented EPI and GRAPPA." In: *Magnetic Resonance in Medicine* 64.1, pp. 9–14.
- Hinds, Oliver P et al. (2008). "Accurate prediction of V1 location from cortical folds in a surface coordinate system." In: *Neuroimage* 39.4, pp. 1585–1599.
- Hosey, Tim, Guy Williams, and Richard Ansorge (2005). "Inference of multiple fiber orientations in high angular resolution diffusion imaging." In: *Magnetic Resonance in Medicine* 54.6, pp. 1480–1489.
- Jenkinson, Mark and Stephen Smith (2001). "A global optimisation method for robust affine registration of brain images." In: *Medical image analysis* 5.2, pp. 143–156.
- Jensen, Jens H et al. (2005). "Diffusional kurtosis imaging: The quantification of non-gaussian water diffusion by means of magnetic resonance imaging." In: *Magnetic resonance in medicine* 53.6, pp. 1432–1440.
- Jeurissen, B et al. (2010). "Estimating the number of fiber orientations in diffusion MRI voxels: a constrained spherical deconvolution study." In: *International Society for Magnetic Resonance in Medicine (ISMRM)*, p. 573.
- Jeurissen, Ben et al. (2014). "Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion MRI data." In: *NeuroImage* 103, pp. 411–426.
- Johansen-Berg, Heidi et al. (2004). "Changes in connectivity profiles define functionally distinct regions in human medial frontal cortex." In: *Proceedings of the National Academy of Sciences of the United States of America* 101.36, pp. 13335–13340.
- Jones, Derek K (2010). *Diffusion mri*. Oxford University Press.
- Kaden, Enrico et al. (2016). "Multi-compartment microscopic diffusion imaging." In: *NeuroImage* 139, pp. 346–359.
- Kanwisher, Nancy, Josh McDermott, and Marvin M Chun (1997). "The fusiform face area: a module in human extrastriate cortex specialized for face perception." In: *Journal of neuroscience* 17.11, pp. 4302–4311.
- Kingsley, Peter B and W Gordon Monahan (2004). "Selection of the optimum b factor for diffusion-weighted magnetic resonance imaging assessment of ischemic stroke." In: *Magnetic resonance in medicine* 51.5, pp. 996–1001.

- Kleinnijenhuis, Michiel et al. (2015). "Diffusion tensor characteristics of gyrencephaly using high resolution diffusion MRI in vivo at 7T." In: *NeuroImage* 109, pp. 378–387.
- Kuehn, Esther et al. (2017). "Body Topography Parcellates Human Sensory and Motor Cortex." In: *Cerebral Cortex*, pp. 1–16.
- Le Bihan, Denis (2003). "Looking into the functional architecture of the brain with diffusion MRI." In: *Nature Reviews Neuroscience* 4.6, pp. 469–480.
- Le Bihan, Denis and Steven J Warach (1995). *Diffusion and Perfusion Magnetic Resonance Imaging: Applications to Functional MRI*.
- Le Bihan, Denis et al. (2001). "Diffusion tensor imaging: concepts and applications." In: *Journal of magnetic resonance imaging* 13.4, pp. 534–546.
- Le Bihan, Denis et al. (2006). "Artifacts and pitfalls in diffusion MRI." In: *Journal of magnetic resonance imaging* 24.3, pp. 478–488.
- Leonard, Christiana M et al. (1998). "Normal variation in the frequency and location of human auditory cortex landmarks. Heschl's gyrus: where is it?" In: *Cerebral Cortex* 8.5, pp. 397–406.
- Leuze, Christoph WU et al. (2014). "Layer-specific intracortical connectivity revealed with diffusion MRI." In: *Cerebral Cortex* 24.2, pp. 328–339.
- Lewis, Bevan and Henry Clarke (1878). "The cortical lamination of the motor area of the brain." In: *Proceedings of the Royal Society of London* 27.185-189, pp. 38–49.
- Lutti, Antoine et al. (2014). "Using high-resolution quantitative mapping of R1 as an index of cortical myelination." In: *Neuroimage* 93, pp. 176–188.
- Ma, Dan et al. (2013). "Magnetic resonance fingerprinting." In: *Nature* 495.7440, p. 187.
- MacQueen, James et al. (1967). "Some methods for classification and analysis of multivariate observations." In: *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*. Vol. 1. 14. Oakland, CA, USA., pp. 281–297.
- McKinstry, Robert C et al. (2002). "Radial organization of developing preterm human cerebral cortex revealed by non-invasive water diffusion anisotropy MRI." In: *Cerebral Cortex* 12.12, pp. 1237–1243.
- McNab, Jennifer A et al. (2009). "High resolution diffusion-weighted imaging in fixed human brain using diffusion-weighted steady state free precession." In: *Neuroimage* 46.3, pp. 775–785.

- McNab, Jennifer A et al. (2013). "Surface based analysis of diffusion orientation for identifying architectonic domains in the in vivo human cortex." In: *Neuroimage* 69, pp. 87–100.
- Mitra, Partha P and Bertrand I Halperin (1995). "Effects of finite gradient-pulse widths in pulsed-field-gradient diffusion measurements." In: *Journal of Magnetic Resonance, Series A* 113.1, pp. 94–101.
- Morosan, P et al. (2005). "Multimodal architectonic mapping of human superior temporal gyrus." In: *Anatomy and embryology* 210.5-6, pp. 401–406.
- Mountcastle, Vernon B (1997). "The columnar organization of the neocortex." In: *Brain: a journal of neurology* 120.4, pp. 701–722.
- Mukherjee, Pratik et al. (2002). "Diffusion-tensor MR imaging of gray and white matter development during normal human brain maturation." In: *American Journal of Neuroradiology* 23.9, pp. 1445–1456.
- Nagy, Zoltan et al. (2013). "Using high angular resolution diffusion imaging data to discriminate cortical regions." In: *PloS one* 8.5, e63842.
- Nitz, WR and P Reimer (1999). "Contrast mechanisms in MR imaging." In: *European radiology* 9.6, pp. 1032–1046.
- Ogawa, Seiji et al. (1990). "Brain magnetic resonance imaging with contrast dependent on blood oxygenation." In: *Proceedings of the National Academy of Sciences* 87.24, pp. 9868–9872.
- Orban, Guy A, David Van Essen, and Wim Vanduffel (2004). "Comparative mapping of higher visual areas in monkeys and humans." In: *Trends in cognitive sciences* 8.7, pp. 315–324.
- Özarslan, Evren and Thomas H Mareci (2003). "Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution diffusion imaging." In: *Magnetic resonance in Medicine* 50.5, pp. 955–965.
- Panagiotaki, Eleftheria et al. (2012). "Compartment models of the diffusion MR signal in brain white matter: a taxonomy and comparison." In: *Neuroimage* 59.3, pp. 2241–2254.
- Papadakis, Nikolaos G et al. (1999). "A comparative study of acquisition schemes for diffusion tensor imaging using MRI." In: *Journal of Magnetic Resonance* 137.1, pp. 67–82.

- Papadopoulos, Théo, Aurobrata Ghosh, and Rachid Deriche (2014). "Complete set of invariants of a 4th order tensor: the 12 tasks of HARDI from ternary quartics." In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, pp. 233–240.
- Parker, Geoffrey JM and Daniel C Alexander (2005). "Probabilistic anatomical connectivity derived from the microscopic persistent angular structure of cerebral tissue." In: *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 360.1457, pp. 893–902.
- Parker, Geoffrey JM, Hamied A Haroon, and Claudia AM Wheeler-Kingshott (2003). "A framework for a streamline-based probabilistic index of connectivity (PICO) using a structural interpretation of MRI diffusion measurements." In: *Journal of Magnetic Resonance Imaging* 18.2, pp. 242–254.
- Poonawalla, Aziz H et al. (2008). "Diffusion-tensor MR imaging of cortical lesions in multiple sclerosis: initial findings." In: *Radiology* 246.3, pp. 880–886.
- Poser, Benedikt A et al. (2006). "BOLD contrast sensitivity enhancement and artifact reduction with multiecho EPI: parallel-acquired inhomogeneity-desensitized fMRI." In: *Magnetic Resonance in Medicine* 55.6, pp. 1227–1235.
- Posse, Stefan et al. (1999). "Enhancement of BOLD-contrast sensitivity by single-shot multi-echo functional MR imaging." In: *Magnetic resonance in medicine* 42.1, pp. 87–97.
- Powers, David Martin (2011). "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation." In:
- Rademacher, J et al. (1993). "Topographical variation of the human primary cortices: implications for neuroimaging, brain mapping, and neurobiology." In: *Cerebral Cortex* 3.4, pp. 313–329.
- Rajapakse, Jagath C and Juan Zhou (2007). "Learning effective brain connectivity with dynamic Bayesian networks." In: *Neuroimage* 37.3, pp. 749–760.
- Reese, TG et al. (2003). "Reduction of eddy-current-induced distortion in diffusion MRI using a twice-refocused spin echo." In: *Magnetic resonance in medicine* 49.1, pp. 177–182.
- Reisert, Marco et al. (2017). "Disentangling micro from mesostructure by diffusion MRI: A Bayesian approach." In: *NeuroImage* 147, pp. 964–975.

- Sánchez-Panchuelo, Rosa M et al. (2012). "Correspondence of human visual areas identified using functional and anatomical MRI in vivo at 7 T." In: *Journal of Magnetic Resonance Imaging* 35.2, pp. 287–299.
- Sánchez-Panchuelo, Rosa-María et al. (2014). "Regional structural differences across functionally parcellated Brodmann areas of human primary somatosensory cortex." In: *Neuroimage* 93, pp. 221–230.
- Scherrer, Benoit et al. (2013). "Characterizing the distribution of anisotropic micro-structural environments with diffusion-weighted imaging (DIAMOND)." In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, pp. 518–526.
- Schleicher, A et al. (2005). "Quantitative architectural analysis: a new approach to cortical mapping." In: *Anatomy and embryology* 210.5-6, pp. 373–386.
- Sereno, Martin I and Roger BH Tootell (2005). "From monkeys to humans: what do we now know about brain homologies?" In: *Current opinion in neurobiology* 15.2, pp. 135–144.
- Sereno, Martin I et al. (2013). "Mapping the human cortical surface by combining quantitative T1 with retinotopy." In: *Cerebral Cortex* 23.9, pp. 2261–2268.
- Sereno, Martin I, Colin T McDonald, and John M Allman (2015). "Retinotopic organization of extrastriate cortex in the owl monkey—dorsal and lateral areas." In: *Visual neuroscience* 32.
- Shipp, Stewart (2007). "Structure and function of the cerebral cortex." In: *Current Biology* 17.12, R443–R449.
- Sigalovsky, Irina S, Bruce Fischl, and Jennifer R Melcher (2006). "Mapping an intrinsic MR property of gray matter in auditory cortex of living humans: a possible marker for primary cortex and hemispheric differences." In: *Neuroimage* 32.4, pp. 1524–1537.
- Silver, David et al. (2016). "Mastering the game of Go with deep neural networks and tree search." In: *Nature* 529.7587, pp. 484–489.
- Skare, Stefan et al. (2000). "Condition number as a measure of noise performance of diffusion tensor data acquisition schemes with MRI." In: *Journal of Magnetic Resonance* 147.2, pp. 340–352.
- Smart, IH and GM McSherry (1986). "Gyrus formation in the cerebral cortex of the ferret. II. Description of the internal histological changes." In: *Journal of anatomy* 147, p. 27.

- Smith, G Elliot (1907). "A new topographical survey of the human cerebral cortex, being an account of the distribution of the anatomically distinct cortical areas and their relationship to the cerebral sulci." In: *Journal of anatomy and physiology* 41.Pt 4, p. 237.
- Sotiropoulos, Stamatios N, Timothy EJ Behrens, and Saad Jbabdi (2012). "Ball and rackets: inferring fiber fanning from diffusion-weighted MRI." In: *NeuroImage* 60.2, pp. 1412–1425.
- Sotiropoulos, Stamatios N et al. (2013). "Advances in diffusion MRI acquisition and processing in the Human Connectome Project." In: *Neuroimage* 80, pp. 125–143.
- Sporns, Olaf, Giulio Tononi, and Rolf Kötter (2005). "The human connectome: a structural description of the human brain." In: *PLoS computational biology* 1.4, e42.
- Stejskal, Edward O and John E Tanner (1965). "Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient." In: *The journal of chemical physics* 42.1, pp. 288–292.
- Talairach, Jean and Pierre Tournoux (1988). "Co-planar stereotaxic atlas of the human brain. 3-Dimensional proportional system: an approach to cerebral imaging." In:
- Tallinen, Tuomas et al. (2014). "Gyrification from constrained cortical expansion." In: *Proceedings of the National Academy of Sciences* 111.35, pp. 12667–12672.
- Tallinen, Tuomas et al. (2016). "On the growth and form of cortical convolutions." In: *Nature Physics* 12.6, pp. 588–593.
- Tariq, Maira et al. (2016). "Bingham–noddli: Mapping anisotropic orientation dispersion of neurites using diffusion mri." In: *NeuroImage* 133, pp. 207–223.
- Tournier, J-Donald et al. (2004). "Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution." In: *NeuroImage* 23.3, pp. 1176–1185.
- Tournier, J-Donald et al. (2008). "Resolving crossing fibres using constrained spherical deconvolution: validation using diffusion-weighted imaging phantom data." In: *Neuroimage* 42.2, pp. 617–625.
- Trampel, Robert, Derek VM Ott, and Robert Turner (2011). "Do the congenitally blind have a stria of Gennari? First intracortical insights in vivo." In: *Cerebral Cortex* 21.9, pp. 2075–2081.

- Triarhou, Lazaros C (2012). "Cytoarchitectonics of the Human Cerebral Cortex: The 1926 Presentation by Georg N. Koskinas (1885-1975) to the Athens Medical Society." In: *Neuroimaging-Cognitive and Clinical Neuroscience*. InTech.
- Tuch, David S et al. (2002). "High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity." In: *Magnetic Resonance in Medicine* 48.4, pp. 577–582.
- Turner, Robert et al. (2008). "Optimised in vivo visualisation of cortical structures in the human brain at 3 T using IR-TSE." In: *Magnetic resonance imaging* 26.7, pp. 935–942.
- Uğurbil, Kamil et al. (2013). "Pushing spatial and temporal resolution for functional and diffusion MRI in the Human Connectome Project." In: *Neuroimage* 80, pp. 80–104.
- Van Essen, David C (1997). "A tension-based theory of morphogenesis and compact wiring in the central nervous system." In: *Nature* 385.6614, p. 313.
- Van Essen, David C et al. (2013a). "Mapping connections in humans and non-human primates: aspirations and challenges for diffusion imaging." In: *Diffusion MRI: From Quantitative Measurement to In vivo Neuroanatomy*, pp. 337–358.
- Van Essen, David C et al. (2013b). "The WU-Minn human connectome project: an overview." In: *Neuroimage* 80, pp. 62–79.
- Veenith, Tonny V et al. (2013). "Inter subject variability and reproducibility of diffusion tensor imaging within and between different imaging sessions." In: *PloS one* 8.6, e65941.
- Vestergaard-Poulsen, Peter et al. (2007). "Microstructural changes in ischemic cortical gray matter predicted by a model of diffusion-weighted MRI." In: *Journal of Magnetic Resonance Imaging* 26.3, pp. 529–540.
- Vogt, C (1919). "Allgemeine Ergebnisse unserer Hirnforschung I–IV." In: *J. Psychol. Neurol.(Lpz.)* 25, Erg. heft 1, pp. 279–462.
- Vogt, O and C Vogt (1903). "Zur anatomischen Gliederung des Cortex cerebri." In: *J Psychol Neurol* 2, pp. 160–180.
- Vu, An T et al. (2015). "High resolution whole brain diffusion imaging at 7T for the Human Connectome Project." In: *Neuroimage* 122, pp. 318–331.
- Waehnert, MD et al. (2014). "Anatomically motivated modeling of cortical laminae." In: *Neuroimage* 93, pp. 210–220.

- Waehnert, Miriam D et al. (2016). "A subject-specific framework for in vivo myeloarchitectonic analysis using high resolution quantitative MRI." In: *Neuroimage* 125, pp. 94–107.
- Walters, Nathan B et al. (2003). "In vivo identification of human cortical areas using high-resolution MRI: an approach to cerebral structure–function correlation." In: *Proceedings of the National Academy of Sciences* 100.5, pp. 2981–2986.
- Wandell, Brian A, Serge O Dumoulin, and Alyssa A Brewer (2007). "Visual field maps in human cortex." In: *Neuron* 56.2, pp. 366–383.
- Wansapura, Janaka P et al. (1999). "NMR relaxation times in the human brain at 3.0 tesla." In: *Journal of magnetic resonance imaging* 9.4, pp. 531–538.
- White, LE et al. (1997). "Structure of the human sensorimotor system. I: Morphology and cytoarchitecture of the central sulcus." In: *Cerebral Cortex* 7.1, pp. 18–30.
- Wilm, Bertram J et al. (2015). "Diffusion MRI with concurrent magnetic field monitoring." In: *Magnetic resonance in medicine* 74.4, pp. 925–933.
- Xu, Gang et al. (2010). "Axons pull on the brain, but tension does not drive cortical folding." In: *Journal of biomechanical engineering* 132.7, p. 071013.
- Yoshiura, Takashi et al. (2000). "Heschl and superior temporal gyri: low signal intensity of the cortex on T2-weighted MR images of the normal brain." In: *Radiology* 214.1, pp. 217–221.
- Yu, Peng et al. (2007). "Cortical surface shape analysis based on spherical wavelets." In: *IEEE transactions on medical imaging* 26.4, pp. 582–597.
- Zhang, Hui et al. (2011). "Axon diameter mapping in the presence of orientation dispersion with diffusion MRI." In: *Neuroimage* 56.3, pp. 1301–1315.
- Zhang, Hui et al. (2012). "NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain." In: *Neuroimage* 61.4, pp. 1000–1016.
- Zilles, K et al. (2002). "Quantitative analysis of cyto- and receptor architecture of the human brain." In: *Brain mapping: the methods* 2, pp. 573–602.
- Zilles, Karl and Katrin Amunts (2010). "Centenary of Brodmann's map - conception and fate." In: *Nature Reviews Neuroscience* 11.2, pp. 139–145.
- Zilles, Karl et al. (1997). "Quantitative analysis of sulci in the human cerebral cortex: development, regional heterogeneity, gender difference, asymmetry, intersubject variability and cortical architecture." In: *Human brain mapping* 5.4, pp. 218–221.