Abstract of " Developing and Evaluating Computational Tools for the Modeling, Visualization, and Analysis of Brain White Matter with Diffusion MRI " by Ryan P. Cabeen, Ph.D., Brown University, May 2016.

We present research developing and evaluating computational tools for brain mapping with diffusion magnetic resonance imaging (MRI). Diffusion MRI enables the *in-vivo* reconstruction of the geometric structure of brain white matter, and it has potential to improve our understanding the brain in health and disease through numerous scientific and clinical applications. The complexity of both the imaging data and the underlying anatomy make computational tools an essential part of these applications, and improvements to this toolset can allow more intricate anatomical structures to be found and population-wide variation to be mapped. Beyond these applications, such computational tools also represent a research area with many open problems of their own, as every advance made in tool development also creates a need to expand our understanding of the strengths and limitations of the tools themselves.

This thesis work represents advances made on both fronts to develop and evaluate computational tools for modeling, visualizing, and analyzing brain white matter with diffusion MRI. The contributions include voxel-based analysis, fiber bundle modeling, and model-based image processing. First, we develop an algorithm for supervoxel segmentation of brain white matter and show how this is useful for automated region-based analysis in population imaging studies. Second, we comparatively evaluate eight methods for spatial mapping in diffusion tensor studies and quantify their relationship to scan-rescan reliability and predictive models of age-related decline in an adult population of 80 subjects. Third, we develop and evaluate methods for tractography-based modeling of fiber bundles using a sparse closest point transform and its use in combination with statistical and machine learning algorithms in bundle clustering, simplification, and population selection. Fourth, we develop and evaluate a kernel regression framework for tractography and atlas construction with multi-compartment model imaging data. We show how this improves the accuracy of tractography reconstructions in crossing fiber regions and enables complex anatomy to be reliably reconstructed in both individual subjects and population averages. Finally, we apply this multi-fiber tractography method to the reconstruction of the corticospinal tract in tumor imaging applications.

Developing and Evaluating Computational Tools for the Modeling, Visualization, and Analysis of Brain White Matter with Diffusion MRI

by

Ryan P. Cabeen

Bachelors of Science, California Institute of Technology, 2005

Masters of Science, Brown University, 2012

A dissertation submitted in partial fulfillment of the
requirements for the Degree of Doctor of Philosophy
in the Department of Computer Science at Brown University

Providence, Rhode Island

May 2016

This dissertation by Ryan P. Cabeen is accepted in its present form by the Department of Computer Science as satisfying the dissertation requirement for the degree of Doctor of Philosophy.

Date _____ _____

David H. Laidlaw, Director

Recommended to the Graduate Council

Date _____ _____

Erik Sudderth, Reader

Date _____ _____

Benjamin Raphael, Reader

Date _____ _____

Carl-Fredrik Westin, Reader
Harvard Medical School

Date _____ _____

Stephen Correia, Reader

Approved by the Graduate Council

Date _____ _____

Peter M. Weber
Dean of the Graduate School

Vita

Biography

Ryan Payne Cabeen was born in Los Angeles, California in February 1983. He attended the California Institute of Technology from 2001 to 2005, where he obtained a Bachelors of Science in Engineering and Applied Science. While there, he concentrated in the area of Computation and Neural Systems and worked as a research assistant under supervision of Lihi Zelnik-Manor in Pietro Perona's research group. Following this, he spent four years as staff at the Laboratory of Neuroimaging at the University of California, Los Angeles working under the supervision of Roger P. Woods and Arthur Toga. In 2010, he started his graduate studies at the Brown University Computer Science Department in the Visualization Research Group under the direction of David H. Laidlaw. While there, he earned a Masters of Science in 2012, was the teaching assistant for CSCI 1370: Virtual Reality Design for Science in Fall 2013, won the Brown Institute for Brain Science Graduate Research Award in Spring 2014, and contributed to the development and outreach of the YURT virtual reality environment. While at Brown, he was also a visiting student at the Rhode Island Hospital 3D Laboratory under direction of Derek Merck.

Education

Brown University, Providence, Rhode Island

PhD in Computer Science, 2012 - 2016

MSc in Computer Science, 2010 - 2012

California Institute of Technology, Pasadena, California

BS in Engineering and Applied Science, 2001 - 2005

Experience

Brown Visualization Research Lab

Research Assistant **September 2010 - 2016**

Teaching Assistant **Fall 2013**

Rhode Island Hospital 3D Lab

Visiting Research Assistant **Spring 2014**

UCLA Laboratory of Neuro Imaging

Programmer Analyst II **June, 2006 - June, 2010**

Service Machines Inc.

Software Engineer **October, 2005 - May, 2006**

Caltech Vision Lab

Research Assistant **October, 2004 - May, 2005**

Awards

Graduate Research Award from Brown Institute for Brain Science 2014

MICCAI Student Travel Award 2013

ISMRM Trainee Stipend Award 2012, 2013

Brown Graduate Fellowship 2010–11

UCLA Staff Achievement Award 2008–09

Activities

Presentations

Poster presentation at Brown University Mind-Brain Research Day 2016

Demo Presentation at *Immersion at Brown* Symposium 2015

Oral presentation at MICCAI DTI Challenge 2015

Poster presentation at MICCAI Workshop on Computational Diffusion MRI 2014

Oral presentation at MICCAI DTI Challenge 2014

Poster presentation at Brown Center for Comp. Mol. Biology Research Day 2014

Poster presentation Brown CS Industrial Partners Program Symposium 2014

Oral and poster presentation at MICCAI 2013

Oral and poster presentation at MICCAI Workshop on Medical Computer Vision 2013

Poster presentations at OHBM 2013, 2014

Poster presentations at ISMRM 2012, 2013

Teaching and Mentoring

Summer 2014: Mentored undergraduate at Rhode Island Hospital 3D Lab

Fall 2013: Teaching Assistant for Brown CSCI-1370 *Virtual Reality Design for Science*

Fall 2012: Guest lecture for Brown CSCI-2950T *Big Data: Systems, Analysis, and Vis.*

Service

Reviewer for NeuroImage 2016

Reviewer for Human Brain Mapping 2014 - 2016

Reviewer for IEEE SciVis 2015 - 2016

Reviewer for Vis. and Processing of Higher Order Descriptors for Multi-Valued Data 2014

Reviewer for ACM CHI Extended Abstracts 2014, Late-Breaking Work 2016

Reviewer for IEEE BioVis 2013

Brown University CS PhD Admission Committee Student Member 2015, 2016

Brown University CS Mac Merc 2012 - 2016

Judge for Rhode Island Science and Engineering Fair 2014

Peer–reviewed Journal Papers

RP Cabeen, ME Bastin, and DH Laidlaw. **Kernel Regression Estimation of Fiber Orientation Mixtures in Diffusion MRI.** *NeuroImage* 2016

SR Gomez, R Jianu, RP Cabeen, H Guo, and DH Laidlaw. **Fauxvea: Crowdsourcing Gaze Estimates for Visualization Analysis Tasks.** *IEEE Tranactions of Visualization and Computer Graphics* 2016

LM Baker, DH Laidlaw, RP Cabeen, E Akbudak, TE Conturo, S Correia, DF Tate, JM Heaps, MR Brier, HD Bolzenius, LE Salminen, EM Lane, AR McMichael, RH Paul. **Cognitive reserve moderates the relationship between neuropsychological performance and white matter fiber bundle length in healthy older adults** *Brain Imaging and Behavior* 2016

LE Salminen, PR Schofield, KD Pierce, Y Zhao, X Luo, YWang, DH Laidlaw, RP Cabeen, TE Conturo, DF Tate, E Akbudak, EM Lane, JM Heaps, JD Bolzenius, LM Baker, L Cagle, and RH Paul. **Neuromarkers of the common angiotensinogen polymorphism in healthy older adults: A comprehensive assessment of white matter integrity and cognition.** *Behavioural Brain Research* 2015

LE Salminen, PR Schofield, KD Pierce, TE Conturo, DH Laidlaw, RP Cabeen, EM. Lane, JM Heaps, JD Bolzenius, LM Baker, X Luo, E Akbudakd, and RH Paul. **Genetic markers of cholesterol transport and gray matter diffusion: A preliminary study of the CETP I405V polymorphism.** *Journal of Neural Transmission* 2015

LE Salminen, TE Conturo, DH Laidlaw, RP Cabeen, E Akbudak, EM Lane, JM Heaps, JD Bolzenius, LM Baker, S Cooley, S Scott, LM Cagle, S Phillips, RH Paul. **Regional age differences in gray matter diffusivity among healthy older adults**. *Brain Imaging and Behavior* 2015

JD Bolzenius, DH Laidlaw, RP Cabeen, TE Conturo, AR McMichael, EM Lane, JM Heaps,

LE Salminen, LM Baker, SE Scott, SA Cooley, J Gunstad, RH Paul. **Brain structure and cognitive correlates of body mass index in healthy older adults**. *Behavioural Brain Research* 2015

AM Behrman, C Usher, TE Conturo, S Correia, DH Laidlaw, EM Lane, J Bolzenius, JM Heaps, LE Salminen, LM Baker, RP Cabeen, E Akbudak, X Luo, P Yan, RH Paul. **Fiber bundle length and cognition: A length-based tractography MRI study**. *Brain Imaging and Behavior* 2014

SA Cooley, RP Cabeen, DH Laidlaw, EM Lane, JM Heaps, JD Bolzenius, LM Baker, LE Salminen, SE Scott, RH Paul. **Posterior brain white matter abnormalities in older adults with probable mild cognitive impairment**. *Journal of Clin. and Exp. Neuropsychology* 2014

LM Baker, DH Laidlaw, T Conturo, J Hogan, Y Zhao, X Luo, S Correia, RP Cabeen, E Lane, J Heaps, J Bolzenius, L Salminen, E Akbudak, A McMichael, C Usher, A Behrman, RH Paul. **White matter changes with age utilizing quantitative diffusion MRI**. *Neurology* 2014

JD Bolzenius, DH Laidlaw, RP Cabeen, TE Conturo, AR McMichael, EM Lane, JM Heaps, LE Salminen, LM Baker, JG Gunstad, RH Paul. **Impact of body mass index on neuronal fiber bundle lengths among healthy older adults**.g*Brain Imaging and Behavior* 2013

LE Salminen, PR Schofield, EM Lane, JM Heaps, KD Pierce, RP Cabeen, DH Laidlaw, E Akbudak, TE Conturo, S Correia, RH Paul. **Neuronal fiber bundle lengths in healthy adult carriers of the ApoE4 allele: A quantitative tractography DTI study**. *Brain Imaging and Behavior* 2013

EJ Marquez, RP Cabeen, RP Woods, D Houle. **The Measurement of Local Variation in Shape**. *Evolutionary Biology* 2012

SH Joshi, RP Cabeen, AA Joshi, B Sun, I Dinov, KL Narr, AW Toga RP Woods. **Diffeomorphic Sulcal Mapping on the Cortex**. *IEEE Trans. on Medical Imaging* 2012

OR Phillips, KH Nuechterlein, RF Asarnow, KA Clark, RP Cabeen, Y Yang, RP Woods, AW Toga, KL Narr. **Mapping Cortico–cortical Structural Integrity in Schizophrenia and Effects of Genetic Liability**. *Biological Psychiatry* 2011

Peer–reviewed Conference Papers

RP Cabeen, DH Laidlaw. **Corticospinal Tract Reconstruction with Deterministic Multi-fiber Tractography and Model-based Processing**. *MICCAI DTI Tractography Challenge* 2015

RP Cabeen, DH Laidlaw. **Bilateral Filtering of Multiple Fiber Orientations in Diffusion MRI**. *MICCAI Workshop on Computational Diffusion MRI* 2014

RP Cabeen, DH Laidlaw. **Diffusion Tractography of the Corticospinal Tract with Multi-fiber Orientation Filtering**. *MICCAI DTI Tractography Challenge* 2014

RP Cabeen, DH Laidlaw. **White Matter Supervoxel Segmentation by Axial DP–means Clustering**. *MICCAI Workshop on Medical Computer Vision* 2013

RP Cabeen, ME Bastin, DH Laidlaw. **Estimating Constrained Multi-fiber Diffusion MR Volumes by Orientation Clustering**. *MICCAI* 2013

SH Joshi, RP Cabeen, AA Joshi, RP Woods, KL Narr, AW Toga. **Diffeomorphic Sulcal Shape Analysis for Cortical Surface Registration**. *CVPR* 2010

SH Joshi, RP Cabeen, B Sun, AA Joshi, BH Gutman, A Zamanyan, S Chakrapani, ID Dinov, RP Woods, AW Toga. **Cortical Sulcal Atlas Construction Using a Diffeomorphic Mapping Approach**. *MICCAI* 2010

Peer–reviewed Conference Abstracts

RP Cabeen, DH Laidlaw. **Diffusion MR Tractography Clustering with the Sparse Closest Point Transform**. *OHBM* 2016

RP Cabeen, ME Bastin, DH Laidlaw. **Reproducibility of an Automated Regional Analysis of White Matter with Diffusion Imaging**. *OHBM* 2014

RP Cabeen, ME Bastin, DH Laidlaw. **A Diffusion MRI Resource of 80 Age-varied Subjects with Neuropsychological and Demographic Measures**. *ISMRM* 2013

RP Cabeen, ME Bastin, DH Laidlaw. **Atlas-based multi-direction tractography using tensor registration and orientation statistics**. *OHBM* 2013

RP Cabeen, D Dickstein, DH Laidlaw. **Characterization of Pediatric Bipolar Disorder with Quantitative HARDI Tractography Metrics**. *ISMRM* 2012

R Jianu, W Zhou, RP Cabeen, D Dickstein, DH Laidlaw. **Visualizing tratography metrics of cortical–connectivity integrity in diffusion imaging**. *ISMRM* 2012

SH Joshi, RP Cabeen, AA Joshi, RP Woods, KL Narr. **Diffeomorphic Sulcal Pattern Matching for Cortical Surface Registration**. *OHBM* 2010

Acknowledgements

This work was supported by the the National Institutes of Health (NIH/NINDS grant R01 NS052470, NIH/NIMH grant R01 MH085604, and NIH grant R01 EB004155), the Brown Institute for Brain Science (Graduate Research Award 2014), and Brown University fellowships. All opinions, findings, conclusions, or recommendations expressed in this document are those of the author(s) and do not necessarily reflect the views of the sponsoring agencies.

Foremost, my appreciation goes to my adviser, David Laidlaw, for the guidance, friendship, and freedom that he has provided over the years. Regardless of the topic, I look forward to our discussions and the valuable insight, not to mention a good joke. I also thank my dissertation committee, Ben Raphael, Erik Sudderth, Carl-Fredrik Westin, and Stephen Correia, for all of their support. I am also thankful for the support of my past academic mentors, Roger Woods, Arthur Toga, Pietro Perona, Lihi Zelnick-Manor, and Shantanu Joshi.

I also gratefully acknowledge my collaborators while at Brown, including Mark Bastin, Robert Paul, Stephen Correia, Sonia Pujol, Daniel Dickstein, Derek Merck, Laurie Baker, Lauren Salminen, and all of the others in St. Louis. Our projects together were a great learning experience and an important part of extending the reach of the contributions of this work.

I am also fortunate to have shared a research group with a great group of colleagues. Thanks to Hua, Connor, Johannes, Wesley, Josh, Scott, Steve, Çağatay, Jadrian, Radu, Wenjin, Eni, and all others. Our weekly meetings have provided a stimulating environment and very helpful feedback.

The community in Brown Computer Science has also been an invaluable resource with all of the hard

work of tstaff and astaff to keep things running smoothly and support from the faculty. In particular, I would like to thank thank Lauren Clarke, Dawn Reed, Eugenia DeGouveia, Ugur Çetintemel, Andy van Dam, and James Hays.

It has also been a pleasure to be part of the community of Brown students and postdocs, who have been a wonderful source of discussion, ideas, commiseration, and friendship. Thanks Mike H, Layla, Justin, Alex, Andy, Seth, Anna, Eric, Mike B, Joe, Betsy, Gen, Geoff, Alexandra, Emanuel, Jie Ying, Krishna, Nat, Adrian, Tim, Shana, Sara, Stephen, Zach, and many others.

I also could not have made it through the last six years without support from my friends and family. Big thanks to my parents, sisters, and my wife Lucia for all of the encouragement and emotional support along the way.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Research Motivation and Significance

Our ability to interact with, understand, and appreciate the world around us is in great part due to the remarkable abilities of the human brain. Even a simple hike through the woods would be impossible without the complex coordination of our motor system to plan and execute our next steps, our visual system to avoid the rocks on the path, our auditory system to alert us to the stream up ahead, and our executive function to plan the trip. What's perhaps most remarkable though is our ability to understand the world around us. All of the neural abilities of perception, memory, and reasoning come together allowing us to appreciate complex ideas, such as the biology of the trees, geology of the mountains around us, and the fluid dynamics of the flowing stream.

While we can thank our brains for this rich experience, our understanding of brain function and structure itself is severely limited. This is in part due to the distributed nature of its function and its intricate anatomical structure, which exists across spatial scales ranging from single neurons to cortical folds. Of course, all of these aspects of the brain can vary across both individuals and time, so our understanding of the brain ultimately should include populations at various stages of development and aging. Many advances have been made to expand our knowledge of how the brain works and is organized, but clearly, these factors suggest there is a tremendous amount we do not yet know.

Fundamentally, this is a scientific process, and our scientific tools are a crucial part of making advancements. Ultimately they form the basis of how we make measurements, represent the results,

organize them, and share our findings. Many kinds of measurements of the brain can be made, for example, with single unit recordings, blunt dissection, imaging, etc. Imaging is particularly useful for characterizing the structure of the brain *in-vivo* with mathematical models that are useful for visualization and quantitative analysis. Modeling also enables the generalization of our results, for example when characterizing the variation of a particular anatomical structure among a population. Finally, such models serve as a simple way to communicate our results to others when there is a great volume of data, with much of it being irrelevant or too detailed. Mathematical models instead provide a summary of these results that hopefully captures the important and interesting parts. Because of all of this, there is much to gain by improving our tools and understanding their strengths and limitations.

Computational tools offer a particularly powerful approach for accomplishing this task, especially when combined with sophisticated imaging techniques. The interdisciplinary nature of this approach can lead to improved understanding of not only neuroscience but also of computer science, mathematics, and physics. Neuroscience has served as a catalyst for progress in the areas of graphics, visualization, computer vision, and artificial intelligence. Likewise, computational tools are essential for understanding the complex data that can be collected about the brain, and more generally, tools are a driving force in scientific progress–as Freeman Dyson observed in his book, "The Scientist as Rebel":

> The great advances in science usually result from new tools rather than from new doctrines. ... Every time we introduce a new tool, it always leads to new and unexpected discoveries, because Nature's imagination is richer than ours.

In this thesis, we explore such interdisciplinary computational tools with the goal of making contributions to not only our scientific understanding of the brain, but also to our understanding of image-based computational modeling, visualization, and analysis. In particular, we focus on the study of brain white matter with diffusion magnetic resonance imaging, a powerful technique with both scientific and clinical applications, and specifically explore a computational techniques for image processing, quantitative analysis, and visualization. We discuss the development each method by examining its theoretical basis and practical concerns for implementation. Our contributions also include the evaluation of computational methods, which consists of both controlled experiments with synthetic datasets and applied experiments with human brain imaging data. The goal of this

evaluation strategy is not only to better understand our proposed tools but also to gain insight into real driving scientific and clinical problems. To facilitate this, each project includes collaborations with domain experts, with the goal of providing driving problems to guide our development and evaluation.

1.2 Contributions and Dissertation Organization

The dissertation in organized into eight chapters in total. The next chapter provides the basic background material necessary for the rest of the dissertation, including a review of brain anatomy and computational diffusion magnetic resonance imaging. The following chapters each constitute the major research contributions of the thesis work and include the following:

- **Supervoxel Segmentation of Diffusion Tensor Images**:

  This chapter develops and evaluates a method for automated region-based analysis of brain white matter anatomy in diffusion tensor imaging studies. While many standard region atlases exist, these typically require manual delineation and are not typically based on specific population being studied. This work aims to expand the toolset available for deriving study-specific whole brain region segmentations with a supervoxel segmentation algorithm. This can be used in conjunction with standard atlas-driven voxel-based and skeleton-based analyses to efficiently and automatically segment regions-of-interest based on locally coherent patterns of white matter fiber orientations. Our contributions include both the development of the theoretical aspects of the method and two experiments to evaluate the method with synthetic and *in vivo* brain imaging data. The first experiment tests the algorithm's ability to segment fiber orientations in a synthetic dataset, and the results demonstrate the ability of the method to automatically identify the number of clusters present in the data under some conditions. The second experiment applies the method to a population-averaged dataset to test the effect of parameter settings, and the results show the effect of the parameters on region count and volume and demonstrate the alignment of region boundaries to anatomical boundaries between major bundles. The third experiment applies the method to segment scans of individual subjects using deformable registration, and the results demonstrate the consistency of region segmentation across individuals and scanning session. These results suggest that this is a potentially useful technique for quantifying whole-brain patterns of diffusion-based anatomical variation

and has possible applications to a variety of clinical neuroimaging studies.

  - RP Cabeen, DH Laidlaw. "White Matter Supervoxel Segmentation by Axial DP-means Clustering." In MICCAI Workshop on Medical Computer Vision 2013
  - RP Cabeen, ME Bastin, DH Laidlaw. "Reproducibility of an Automated Regional Analysis of White Matter with Diffusion Imaging." OHBM 2013

- **A Comparative Evaluation of Spatial Mapping in Diffusion Tensor Imaging**:
  This chapter presents a comparative evaluation of methods for automated voxel-based spatial mapping in diffusion tensor imaging studies. Such methods are an essential step in computational pipelines and provide anatomically comparable measurements across a population in atlas-based studies. To better understand their strengths and weaknesses, we tested a total of eight methods for voxel-based spatial mapping with four diffusion tensor indices and investigated performance in a scan-rescan reliability analysis and an application to predictive modeling of normal aging. The methods included voxel-based and skeleton-based analysis with and without smoothing, and two types of region-based analysis with and without skeletonization. The tensor metrics included fractional anisotropy, and mean, radial, and axial diffusivities. To control for other factors in the pipeline, the experiments used a common dataset, acquired at 1.5T with a single shell high angular resolution diffusion protocol, and a state of the art method for tensor-based spatial normalization and study-specific template construction. Scan-rescan reliability was assessed using the coefficient of variation (CV) and intraclass correlation (ICC) with eight subjects with three scans each. The application experiment quantitatively compared predictive models of age based on principal component regression with a population of 80 normal subjects aged 25 to 65 years. The results show that reliability depends greatly on the method for spatial mapping with less reliability in single-voxel methods (CV = 5.91 $\pm$ 0.81 %, ICC = 0.58 $\pm$ 0.04) than region-based methods (CV = 2.45 $\pm$ 0.25 % and ICC = 0.74 $\pm$ 0.02). The results also highlight how reliability depends on anatomical location and somewhat on the diffusion metric, but not significantly with skeletonization. Despite these differences, all methods performed well in age modeling (R^2 = 0.53 $\pm$ 0.02); however, skeleton-based analysis (R^2 = 0.56 $\pm$ 0.0) showed generally better performance than voxel-based analysis (R^2 = 0.50 $\pm$ 0.0). These results not only can help to interpret results from past white matter imaging studies, but also help in planning future studies to maximize reliability and modeling power

in relation to the scientific goals.

  - RP Cabeen, ME Bastin, DH Laidlaw. "A Comparative Evaluation of Spatial Mapping in Diffusion Tensor Imaging." NeuroImage (under review)

- **Fiber Bundle Modeling with the Sparse Closest Point Transform**:
  This chapter develops and evaluates methods to model brain white matter fiber bundles in diffusion MR tractography datasets using a sparse closest point transform. Such techniques are important in characterizing white matter anatomy in scientific imaging studies; however, most existing techniques include a tight coupling between the representation of tractography data and the statistical methods for bundle modeling. The proposed sparse closest point transform aims to provide a representation that is more generally useful in combination with existing statistical and machine learning algorithms, and this paper evaluates this by quantitatively measuring how well several such combinations solve common tasks in fiber bundle modeling. The proposed representation is obtained by first extracting sparse and representative landmarks from a tractography dataset, and then transforming curves relative to these landmarks using a closest point transform. The evaluation includes three tasks: fiber bundle clustering, simplification, and selection across a population. The bundle clustering algorithm groups fibers from single whole-brain datasets using the DP-means algorithm. It was evaluated with the adjusted Rand index in reference to manually delineated bundles across four tractography datasets, and the results show performance as good as or better than hierarchical and spectral clustering. The simplification algorithm removes redundant curves to improve interactive visualization and similarly using the DP-means to find a representative subsets of curves. It was evaluated with the Dice coefficient relative to the full data, and the results show performance generally superior to random subsampling. The bundle selection algorithm extracts similar bundles across a population given a prototypical example in a group template using a one-class Gaussian classifier. It was evaluated in terms of scan-rescan reliability and an application to normal aging in 80 adult subjects, and the results show comparable reliability to manual region-based bundle selection, as well as comparable or better sensitivity to age-related anatomical changes in the forceps minor. In general, the results show that the sparse closest point transform is efficient and generally useful in fiber bundle modeling tasks common to diffusion MR tractography studies. While the methods tested here are based on

simple machine learning techniques, it could also open the door for a variety of statistical and machine learning techniques to be tested to determine what best suits the needs of a particular scientific study.

  - RP Cabeen, DH Laidlaw. "Diffusion MR Tractography Clustering with the Sparse Closest Point Transform." OHBM 2016
  - RP Cabeen, DH Laidlaw. "Diffusion MR Fiber Bundle Modeling with the Sparse Closest Point Transform." IEEE Transactions on Medical Imaging 2016 (in preparation)

- **Kernel Regression Estimation of Fiber Orientation Mixtures in Diffusion MRI**: This chapter presents and evaluates a method for kernel regression estimation of fiber orientations and associated volume fractions for diffusion MR tractography and population-based atlas construction in clinical imaging studies of brain white matter. This is a model-based image processing technique in which representative fiber models are estimated from collections of component fiber models in model-valued image data. This extends prior work in nonparametric image processing and multi-compartment processing to provide computational tools for image interpolation, smoothing, and fusion with fiber orientation mixtures. In contrast to related work on multi-compartment processing, this approach is based on directional measures of divergence and includes data-adaptive extensions for model selection and bilateral filtering. This is useful for reconstructing complex anatomical features in clinical datasets analyzed with the ball-and-sticks model, and our framework's data-adaptive extensions are potentially useful for general multi-compartment image processing. We experimentally evaluate our approach with both synthetic data from computational phantoms and *in vivo* clinical data from human subjects. With synthetic data experiments, we evaluate performance based on errors in fiber orientation, volume fraction, compartment count, and tractography-based connectivity. With *in vivo* data experiments, we first show improved scan-rescan reproducibility and reliability of quantitative fiber bundle metrics, including mean length, volume, streamline count, and mean volume fraction. We then demonstrate the creation of a multi-fiber tractography atlas from a population of 80 human subjects. In comparison to single tensor atlasing, our multi-fiber atlas shows more complete features of known fiber bundles and includes reconstructions of the lateral projections of the corpus callosum and complex fronto-parietal connections of the superior longitudinal fasciculus I, II, and III.

  - RP Cabeen, ME Bastin, DH Laidlaw. "Kernel Regression Estimation of Fiber Orientation Mixtures in Diffusion MRI." NeuroImage 2016
  - RP Cabeen, DH Laidlaw. Bilateral Filtering of Multiple Fiber Orientations in Diffusion MRI. MICCAI Workshop on Computational Diffusion MRI 2014
  - RP Cabeen, ME Bastin, DH Laidlaw. "Estimating Constrained Multi-fiber Diffusion MR Volumes by Orientation Clustering." MICCAI 2013

- **An Evaluation of Diffusion MR Tractography-based Reconstruction of the Corticospinal Tract**:

  This chapter evaluates several methods for tractography-based reconstruction of the corticospinal tract in clinical neuroimaging. In particular, we examine the reconstruction of fiber bundles in the presence of brain tumors. This is an important task for surgical planning, as it can help a surgeon understand not only how the tumor is interacting with eloquent anatomical structures but also how best to resect tumor tissue without a loss of function. This work is based on submissions to the 2014 and 2015 workshops of the MICCAI DTI Challenge, which is an annual competition with the goal of evaluating tractography algorithms in collaboration with neurosurgeons and other domain experts. For this, we applied the multi-fiber tractography methods presented in this thesis and created reconstructions of the corticospinal tract in several cases and the arcuate fasciculus in a single case. The paper describes this process, presents the results, and discusses the strengths and weaknesses of the proposed method in contrast to standard approaches. Overall, the results show the ability of the multi-fiber modeling framework to recover complex features of white matter in clinical datasets, suggesting this is a potentially useful technique for future applications in surgical planning.

  - RP Cabeen, DH Laidlaw. "Corticospinal Tract Reconstruction with Deterministic Multi-fiber Tractography and Model-based Processing." MICCAI DTI Challenge 2015
  - RP Cabeen, DH Laidlaw. "Diffusion Tractography of the Corticospinal Tract with Multi-fiber Orientation Filtering." MICCAI DTI Challenge 2014

Finally, the dissertation is concluded with a discussion of the results, thoughts about open problems, and reflections on the research process.

Chapter 2

Background

In this chapter, we provide a high level introduction to human brain anatomy and computational diffusion magnetic resonance imaging (MRI). The goal is not to provide a comprehensive description of these topics, but instead to provide the reader with the background information necessary for the rest of the dissertation. The first section will cover human brain anatomy, including anatomical nomenclature, gross brain anatomy, and the major white matter structures discussed later manuscript. The second section covers computational diffusion MRI, including the basics of MR physics, diffusion MR image formation, computational and mathematical methods in diffusion MRI.

2.1 Human Brain Anatomy

In this section, we provide a high level introduction to human brain anatomy. We'll first cover the basic terms for describing the spatial layout of anatomy, and then we'll describe the large-scale structure of the brain. Following this, we'll go into more detail and discuss a number of well-known fiber bundles and show visualizations derived from a brain atlas.

2.1.1 Anatomical Nomenclature

When talking about anatomical structures, it is important to describe their spatial layout in an accurate and specific way. Our colloquial language for describing positions and orientations of objects is unfortunately not up to this task, so biologists instead use anatomical coordinate systems that aim to be standardized and unambiguous [247]. The history of these systems can be traced as

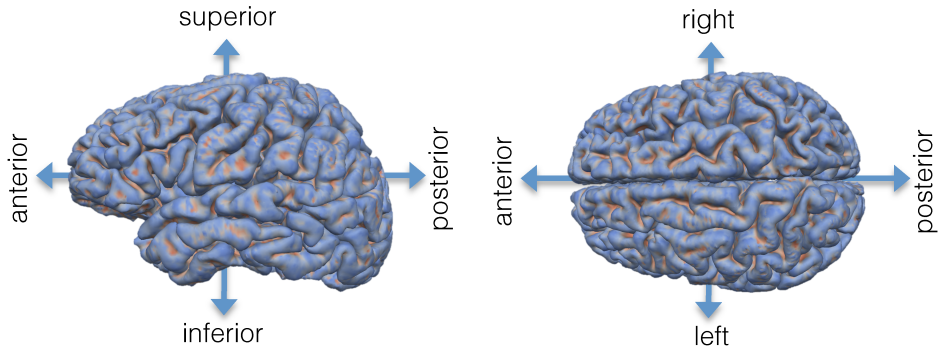


Figure 2.1: The anatomical axes of the brain shown in two views of the cortical surfaces.

far back as the ancient Greeks, so much of the language is derived from Latin and Greek [193] [65]. The system used today for human anatomy is an international standard known as the Terminologia Anatomica [240], although similar terminology is used for most other vertebrates. In the rest of this section, we'll describe the elements of these systems that are useful for discussing human brain anatomy and how it relates to imaging data.

When describing the position of anatomical structures, they are usually placed in a coordinate system consisting of *axes*, which are defined relative to the animal in some standard anatomical position [133]. This ensures that the description does not change when an observer changes their viewing position or when the animal moves its limbs. This system is essentially a three-dimensional coordinates system where each of the axes is assigned a pair of names for the its extreme ends, as shown in Fig. 2.1. While a similar set of axes can be used for most vertebrates, one potentially confusing issue is that humans have a *neuraxis* that curves at the head. The *anterior-posterior axis* (or *rostral-caudal axis*) can be imagined as running from the front to the back of the brain. The *inferior-superior axis* (or *ventral-dorsal axis*) can be imagined as running from the bottom to the top of the brain. The *left-right axis* is similar to the common usage; however, when describing positions along that axis, the term *medial* is used to refer to objects closer to the midpoint and *lateral* for objects away from the midpoint.

In addition to describing anatomy in terms of axes, it is also useful to think of anatomical *planes* [133]. These planes are defined relative to the same standard anatomical position and can be imagined as separating the extreme ends of each axis, as shown in Fig. 2.2. The *transverse plane* (or *axial plane*) separates the inferior from the superior, the *coronal plane* (or *frontal plane*) separates

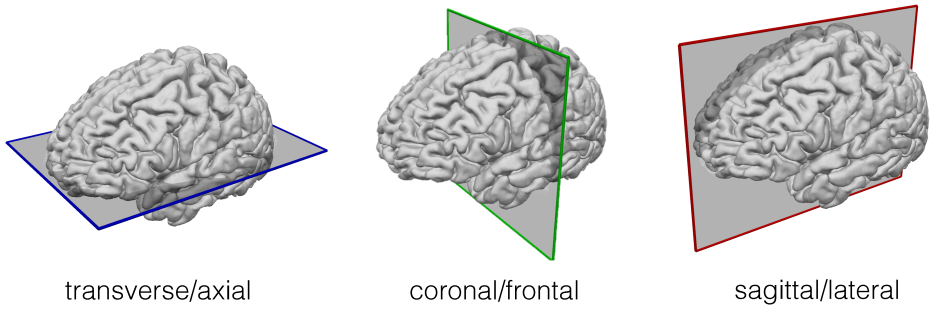


Figure 2.2: The anatomical reference planes of the brain shown in three views of the cortical surfaces.

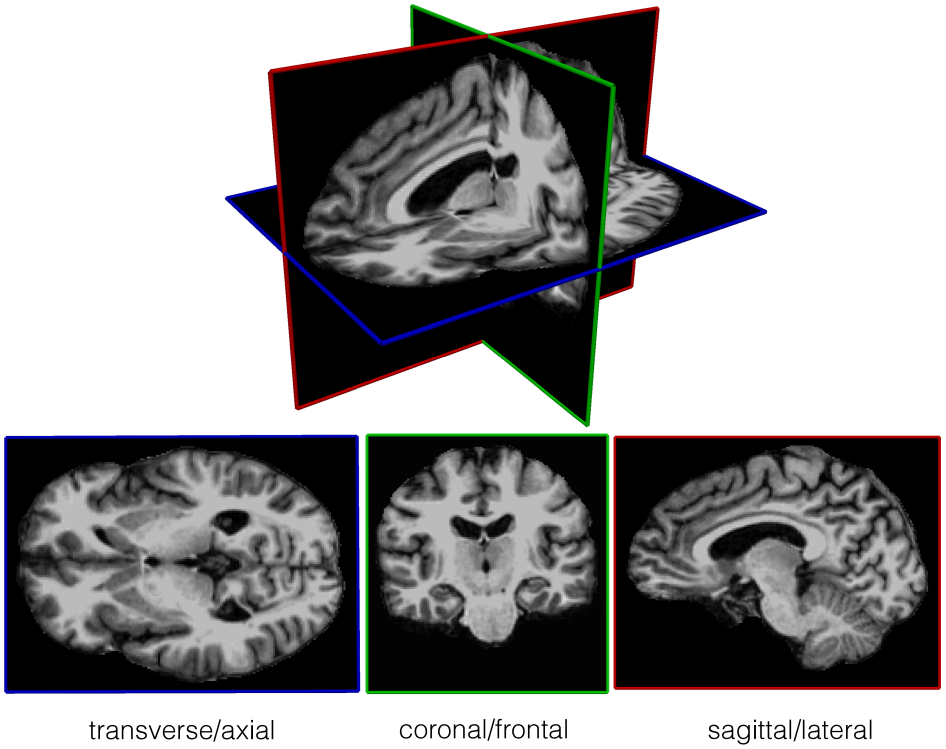


Figure 2.3: Common image volume data and naming conventions for each of the three slices. The top panel shows a 3D view of slices taken through the volumetric imaging data. The bottom panels show 2D views of each slice and their common names.

the anterior from the posterior, and the *sagittal plane* (or *lateral plane*) separates the left from the right. These planes are especially important when talking about imaging data. The axes of the image volume are often aligned with the anatomical axes, so the displayed images are commonly named by the closest anatomical reference plane, as shown in Fig. 2.3.

2.1.2 Gross Brain Anatomy

Next, we'll look at the gross anatomy of the human brain with the goal of understanding the context of white matter anatomy. The brain is included in a larger hierarchical system that starts with the the nervous system at the highest level. The nervous system is then subdivided into the peripheral nervous system (PNS) and the central nervous system (CNS). The PNS is distributed collection of spinal and cranial nerves allow communication between the CNS and the rest of the body. The CNS includes the spinal cord and the brain. The brain is then subdivided into the cerebrum, cerebellum, and the brainstem [153], as shown in Fig. 2.4. While there are many interesting aspects of other parts of the nervous system, we'll focus on the cerebrum here.

The cerebrum is a complex structure that is responsible for much of our sensory processing, decision making, and initiation of voluntary action [133]. It can be subdivided into the diencephalon and the left and right hemispheres, which have a very similar structure. Each hemisphere can be subdivided according to several criteria, such as function, histology, and morphology. Cerebral lobes can be defined by large-scale functional properties [240], resulting in frontal, parietal, occipital, and temporal areas, which are shown in Fig. 2.4. The surface of the brain can also be subdivided according to cytoarchitectonic properties, such as Brodmann areas [29]. Finally, the hemispheres can be divided according to the morphology of the outer cortical surface, which folds to produce gyri (the peaks) and sulci (the valleys), as shown in Fig. 2.6.

The other important way to describe the brain is in terms of tissue types, which include the cerebro-spinal fluid, gray matter, white matter, as shown in Fig. 2.6. The cerebro-spinal fluid is a clear fluid produced in the ventricles and bathes the rest of the brain and spinal cord. It not only protects the CNS with a physical buffer, and also serves immunological and regulatory roles. The gray matter consists of neuronal cell bodies that lie on the outer cortical surface of the brain. The cortex is roughly 2 to 4 millimeters in thickness and contains a complex vertical organization into a number of distinct layers based on neuronal cell type [153]. The white matter consists of the axonal projections of neurons and supporting structures, such as oligodendrocytes and astrocytes.

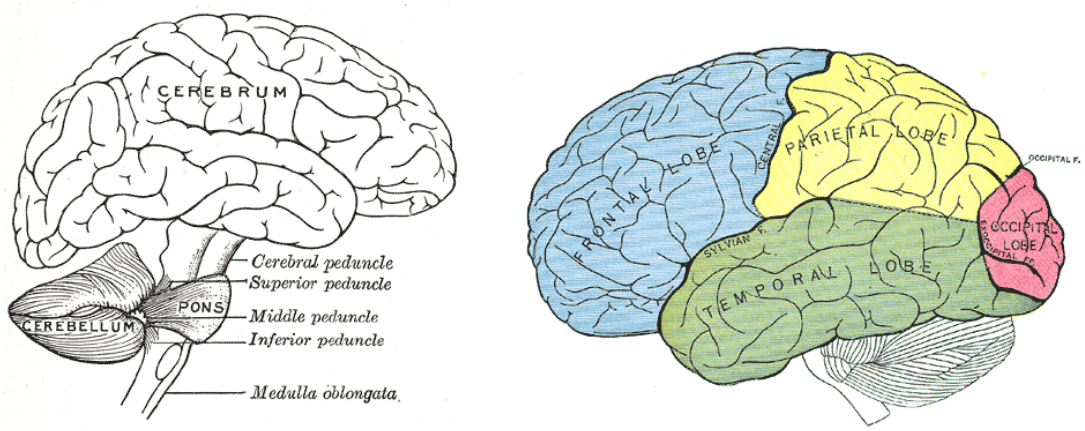


Figure 2.4: Brain anatomy is organized in a hierarchical fashion, starting with coarse subdivisions into the cerebrum, cerebellum, and brainstem, as shown in the left panel. The cerebrum can be further subdivided into functionally distinct frontal, parietal, temporal, and occipital lobes, as shown in the right panel. (from [75])

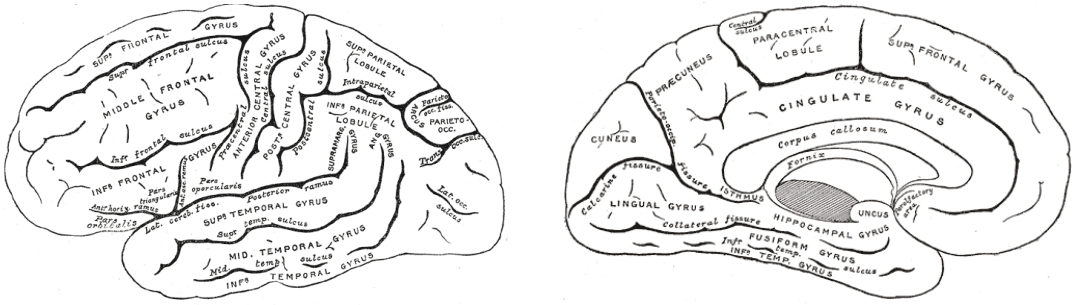


Figure 2.5: In addition to lobular subdivisions, the cerebrum may be alternatively described by smaller areas based on surface morphology of gyral folding patterns. The left panel shows a lateral view of these areas, and the right panel shows a medial view. (from [75])

This tissue is described as "white" due to the appearance of the fatty myelin sheath that insulates the axons. These axons act as communication channels between cortical areas, including both short and long-range connections [189], as shown in Fig. 2.7. Next, we'll look at white matter anatomy in more detail.

2.1.3 White Matter Fiber Bundles

While white matter may appear homogeneous when dissected or viewed in a standard MRI, it is fact an intricate structure in which axon projections follow specific pathways to form complex neural networks. These axon projections form large scale patterns that are similar across many individuals. [158]. Specifically, groups of axons tend to follow coherent trajectories, forming structures known as

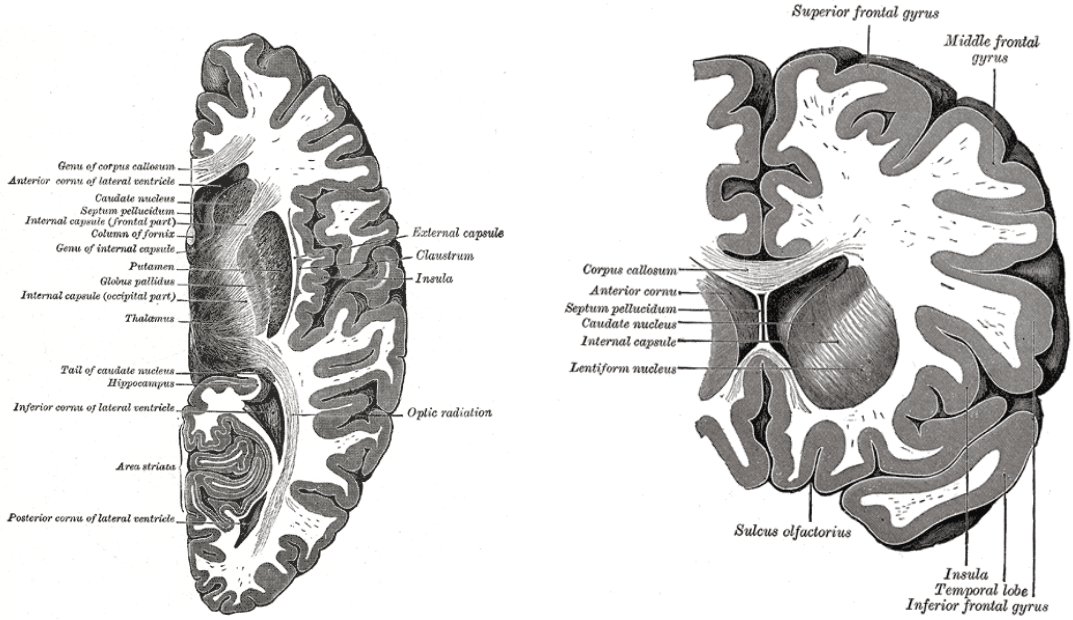


Figure 2.6: The brain may also be described in terms of tissue types. The three main types are gray matter, white matter, and cerebro-spinal fluid. The gray matter consists neuronal cell bodies and resides in both deep subcortical nuclei and the outer cortical shell. The white matter consists of the neuronal axon projects and their supporting structures and resides in the interior of the cerebrum. The cerebrospinal fluid occupies the ventricles and space between the gray matter and pial membrane. The above illustrations show annotated views with dark shading representing gray matter. The left panel shows an axial slice, and the right panel shows a coronal slice. (from [75])

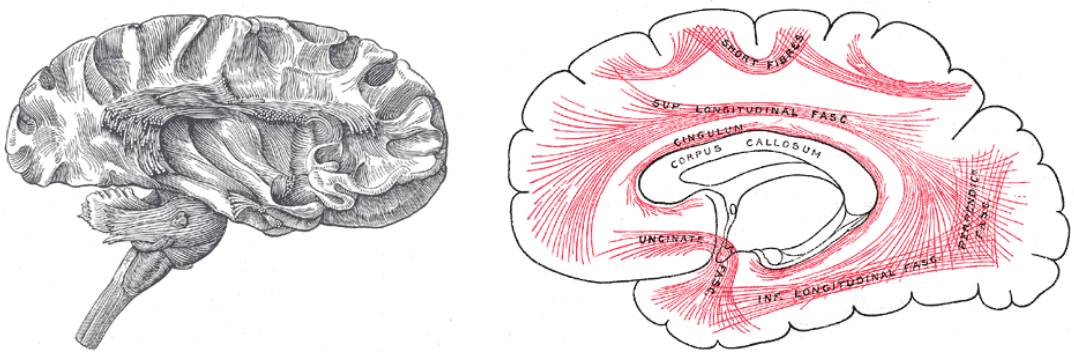
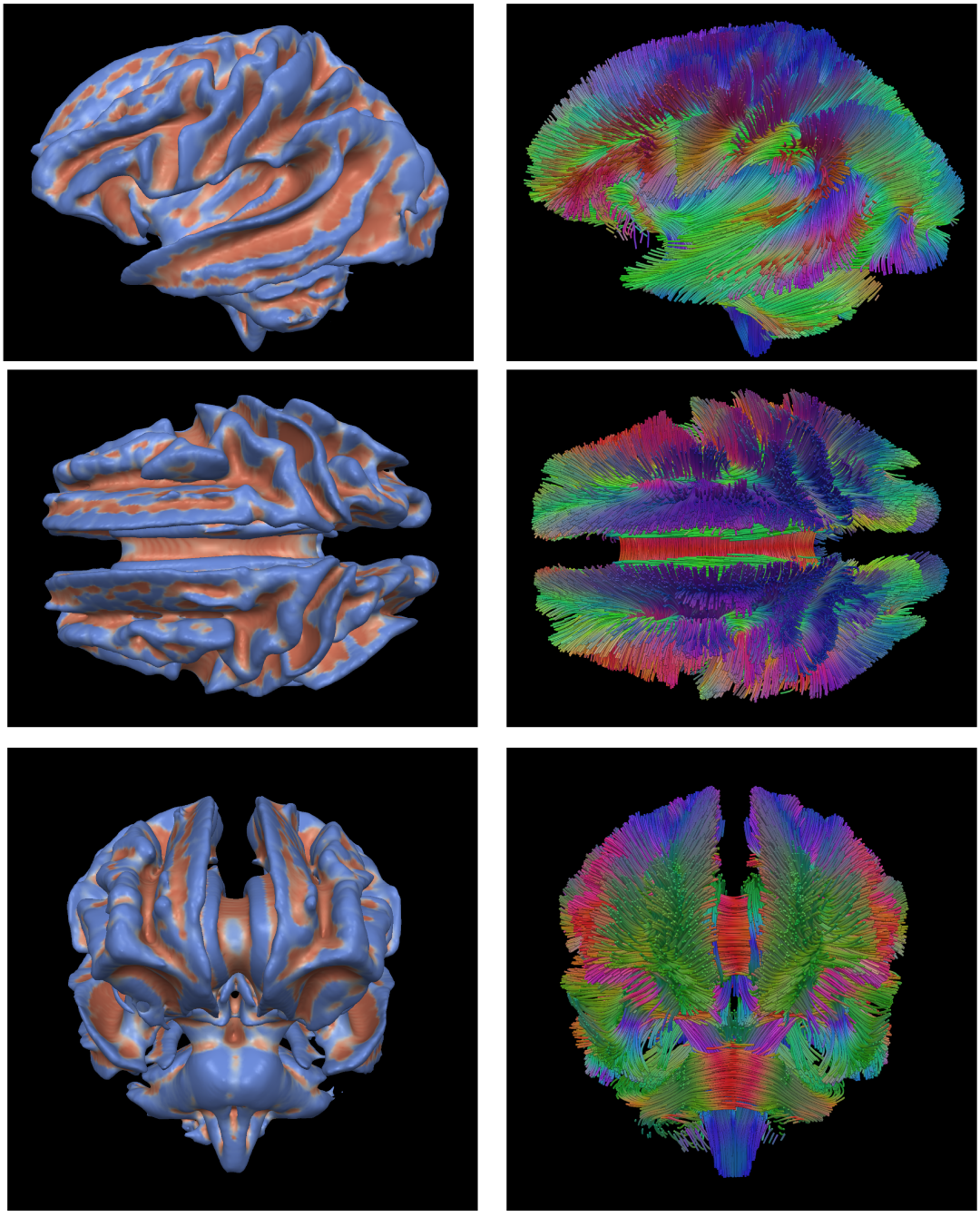


Figure 2.7: The white matter consists of a intricate woven structure of fibers bundles, as illustrated in the left panel. This can be imagined as a wiring diagram of the brain, through which different cortical areas can communicate, as illustrated in the right panel. (from [75])



White Matter Surface White Matter Fibers

Figure 2.8: The white matter anatomy in a large-population brain atlas. The left column shows the white matter surface at the boundary with gray matter. The right column shows white matter tractography fibers. From top to bottom, the rows show lateral, superior, and anterior views.

fiber bundles or *fascicles*. These bundles also form complex configurations, which include crossing, kissing, and sharp turns. Much of the literature on white matter anatomy has aimed to define and characterize specific fiber bundles, which we'll briefly go over next.

Fiber bundles can be roughly categorized based on which parts of the brain they connect, what function they perform, and how they appear [42]. Early work investigated these features on a large scale using Klinger dissection, in which a careful preparation and freezing allows white matter to be pulled away, exposing the surface of a given bundle. More fine detailed examinations have also been performed in non-humans by neuronal tracing, in which the injection of a tracer dye allows the connectivity of individual neurons to be mapped. From these studies, three groups are commonly used for classification: *commissural fibers*, *projection fibers*, and *association fibers*. Fibers can also be described as either *afferent* or *efferent* to indicate whether they carry impulses towards or away from the CNS, respectively.

The commissural fibers include bundles that connect the cerebral hemispheres [42]. Examples of commissural fibers include the *anterior commisure*, *posterior commisure*, and the *corpus callosum*. The corpus callosum (shown in Fig. 2.9) is the largest of the commissural fibers and is often subdivided into several segments based on morphology and micro-architecture. The portion connecting the anterior prefrontal areas is often known as the *genu* or *forceps minor*. The portion connecting the posterior occipital areas is often known as the *splenium* or *forceps major*. The portion connecting the temporal areas and covering the lateral ventricles is often known as the *tapetum*. The remaining fibers are often referred to as the *body* of the corpus callosum.

The projection fibers include bundles that connect the cerebral cortex to more inferior areas, such as subcortical structures and spinal cord [42]. Fig. 2.10 shows examples of several projection bundles, including the *corona radiata*, *cortico-spinal tract*, *fornix*, and the *anterior thalamic radiation*. The corona radiata is a sheetlike structure that fans out from inferior to superior. The cortico-spinal tract connects the spinal tract to the precentral gyrus and is responsible for motor function. The fornix is part of the limbic system and connects the thalamus, the hippocampus, and the mammilary bodies. The anterior thalamic radiation connects the thalamic nuclei to prefrontal cortex and is important for cognition [130].

The association fibers include bundles that connect different cortical areas within the same hemisphere. They can be further described as *short association* when they connect adjacent gyri, or *long association* when they connect more distance gray matter areas. Fig. 2.11 shows examples of several

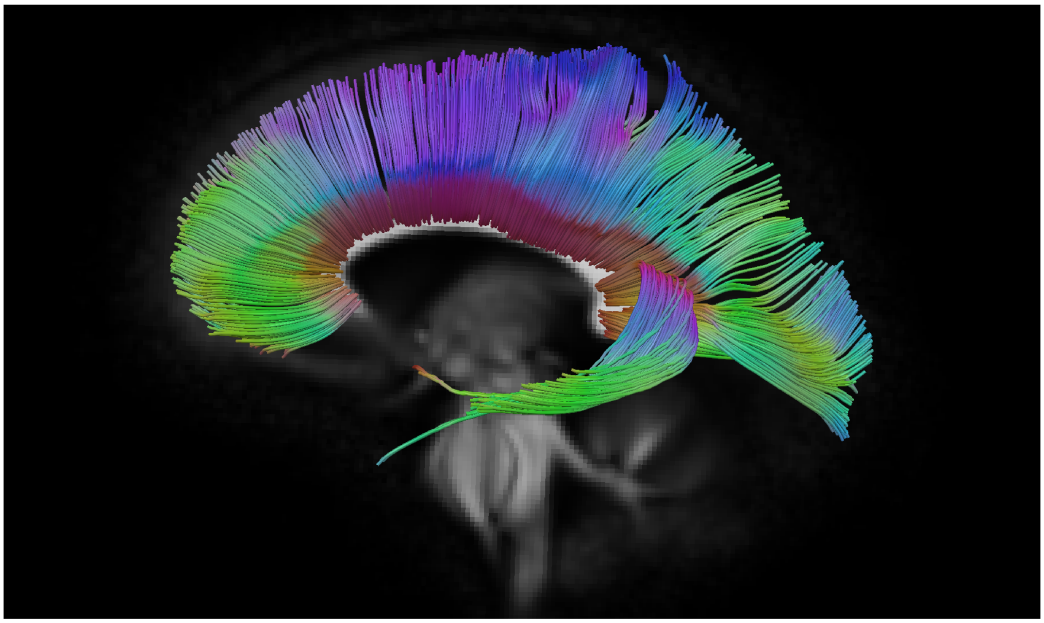


Figure 2.9: Commissural fibers of the corpus callosum. The visualization shows inter-hemispheric connections between the frontal, parietal, occipital, and temporal lobes.

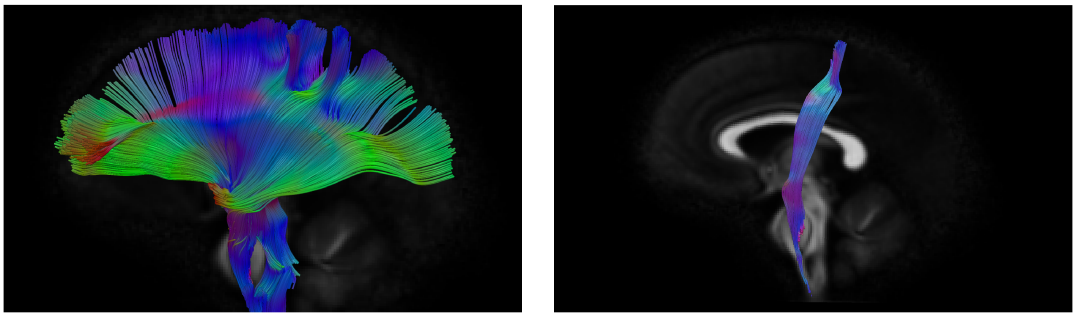
association bundles, including the *uncinate fasciculus*, *arcuate fasciculus*, *cingulum bundle*, *superior longitudinal fasciculus*, *inferior longitudinal fasciculus*, and *inferior fronto-occipital fasciculus*.

2.2 Computational Diffusion MRI

In this section, we discuss diffusion magnetic resonance imaging (MRI) and computational techniques for using it in mapping brain anatomy. This is a large research area, so we'll briefly cover foundational ideas and discuss computational diffusion MRI with enough background to support the rest of the manuscript. In the following two subsections, we'll cover the underlying physics of image formation and how diffusion properties of tissue can be measured and modeled. We'll then examine computational techniques for constructing mathematical models of white matter anatomy and how these can be used for image analysis and visualization in scientific and clinical applications.

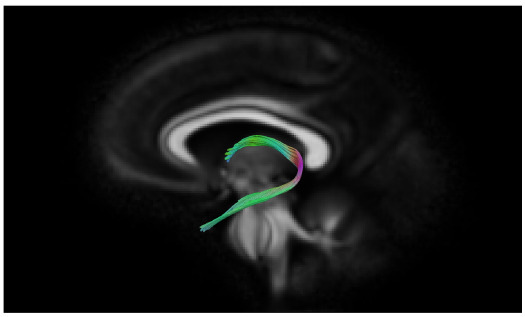
2.2.1 Magnetic Resonance Imaging

Magnetic resonance imaging is a unique imaging technique that is useful for non-invasively mapping the three-dimensional structure of soft tissues in animals. The basis of MRI is fundamentally the physics of molecular spins and how they react to magnetic fields and their surroundings. This

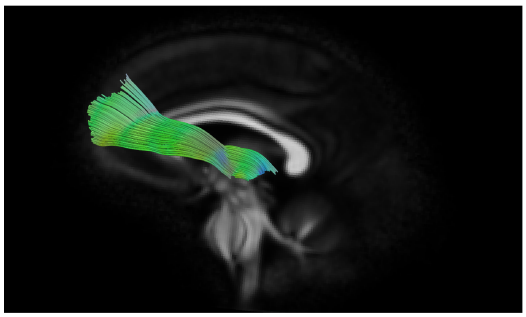


Corona Radiata

Cortico-spinal Tract

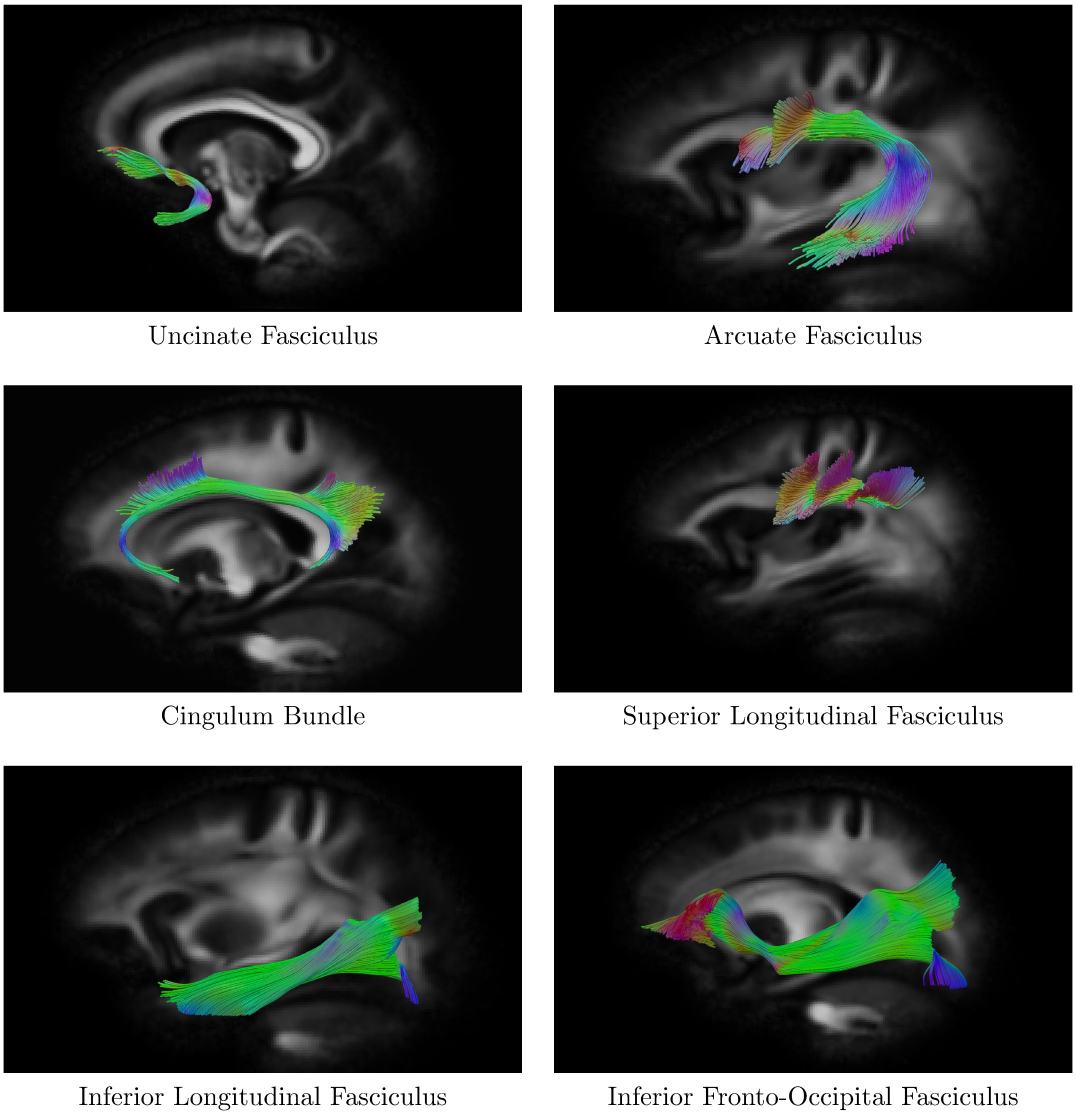


Fornix



Anterior Thalamic Radiation

Figure 2.10: Projection fiber bundles. The top left and right panels show the corona radiata and cortico-spinal tract, respectively. The bottom left and right panels show the fornix and anterior thalamic radiation, respectively.



Uncinate Fasciculus | Arcuate Fasciculus

Cingulum Bundle | Superior Longitudinal Fasciculus

Inferior Longitudinal Fasciculus | Inferior Fronto-Occipital Fasciculus

Figure 2.11: Association fiber bundles. The top left and right panels show the fronto-temporal connections of the uncinate and arcuate fasciculus, respectively. The middle left and right panels show the fronto-parietal connections of the cingulum bundle and the superior longitudinal fasciculus, respectively. The bottom left panel shows the occipito-temporal connections of the inferior longitudinal fasciculus, and the bottom right shows the the inferior fronto-occipital fasciculus.

distinguishes MR images from other common modalities such as optical imaging, ultrasound, and computed tomography (CT). While optical imaging is sensitive to patterns of electromagnetic radiation that characterize surface properties of objects, it is not typically useful for making volumetric reconstructions of tissues. Through the production and sensing of high frequency sound waves, ultrasound is able to volumetrically image tissues; however, it is mostly sensitive to tissue boundaries and suffers from low signal-to-noise and spatial resolution. While CT can offer high quality images, it performs poorly in some soft tissues and introduces significant risks through its use of ionizing X-rays [138]. In comparison to these other techniques, MRI offers high spatial resolution, the ability to distinguish soft tissues, and few known health risks. Its major limitations are its high cost and need for expert knowledge.

The history of MRI dates back many decades, and over the course of its development, it has not only become a standard scientific and clinical tool, but it also includes the discovery of fundamental principles in physics [70]. Essentially, MRI is a technique for measuring properties of protons and their surroundings, specifically water molecule protons that are abundant in biological tissues. These protons are quantum mechanical objects and possess a magnetic moment that precesses when surrounded by a magnetic field. The rate of this precession is known as the Larmor frequency, and it depends on both the proton's intrinsic gyromagnetic ratio and the surrounding magnetic field strength [216]. This phenomenon of nuclear magnetic resonance (NMR) can be used to detect and measure magnetic moments of a physical sample by a the producing and observing electromagnetic radiation (EM), which is typically in the radio-frequency range [177] [174].

From this technique, a number of other important phenomena were uncovered. First, a careful application of EM pulses can be used to produce spin echos, which greatly improve the NMR signal [80] [79]. Second, different samples can exhibit different relaxation rates [40], which are the speeds at which the spins return to equilibrium after perturbation. These relaxation rates reflect microscopic physical properties of the material, so the careful design of a pulse sequence can allow subtle differences or contrasts between materials to be characterized [25] [139]. This turns out to be the case with biological tissue [154], which opened the door for NMR to be used for clinical as well as scientific applications.

While NMR can be used to finely characterize a sample, it lacks the spatial information provided by imaging data, so much of the following work focused on the theory and methods for extending NMR to MRI. The basic way to create an image with NMR is to encode the location of each sample

according to its frequency and/or phase. Early work to solve this problem was done concurrently by several groups, and the resulting idea was to introduce magnetic field gradients in the otherwise uniform magnet [53] [114] [113] [131]. This gives each location a slightly different Larmor frequency, a technique known as slice selection. Then, the voxels of each slice can be decomposed using further frequency and phase encoding gradients, and the resulting data can be reconstructed with a Fourier transform [109] [179]. Further work in this area extended this idea to collect three dimensional data, which produces an image volume instead of only an image slice [111].

Most MR images represent a measurement weighted by some combination of the basic tissue relaxation rates (T_1, T_2, and $T_2^\star$); however, MRI can also be used to measure the physical diffusion properties of tissue. Some of the early work on NMR acknowledged its sensitivity to diffusion [40], but it was not until a decade later that diffusion was treated as an object of study, as opposed to an artifact to be removed [202] [201]. These works developed the theory of diffusion MRI pulse sequences; however, much of the practical innovation came several decades later, with development of imaging techniques and applications to neurology [116] [115] [203]. These early clinical applications demonstrated the measurement of the apparent diffusion coefficient (ADC), which is still a commonly used tool for image-based diagnosis and grading of strokes [232] [233] . However, this turns out to be a quite limited view of what is possible with diffusion MRI, as diffusion can exhibit a complex anisotropic pattern in tissue, which not characterized by these ADC measurements [45]. This is true for brain white matter, where diffusion patterns be used to characterize the geometry of the tissue microstructure, which we'll discuss next.

2.2.2 Diffusion Magnetic Resonance Imaging

While free diffusion of water can be accurately described by a single diffusion coefficient, diffusion in biological tissue can require considerably more complex models [45]. This is due to the presence of a number of obstacles at the molecular level, including cell walls, organelles, and various organic macromolecules [17]. As a consequence, accurate reports of local diffusion properties of the tissue could tell us much about these cellular structures [13]. However, one of the major challenges of diffusion MRI is how to properly model the diffusion-weighted measurements obtained from the scanner. Of course, there is no way to completely reconstruct the diffusion process, so there is always a trade off in the amount of data collected (with increased acquisition time) and model complexity (with more parameters) [117]. Because of this, there are numerous ways to represent the

diffusion process.

The basic MR technique for measuring diffusion is known as the pulsed gradient spin-echo experiment, which was developed by Stejskal and Tanner [202]. In the basic design, a pulsed field gradient is added to the typical spin-echo sequence, resulting in a dephasing and later resphasing. Molecular diffusion that occurs between these phasing steps introduces an attenuation in the observed signal. This is the basis for quantifying diffusion, and given the narrow pulse approximation [38], the attenuated signal S can be generally characterized by a Fourier relationship with the probability density of the underlying diffusion process $p(\mathbf{r}|\tau)$ [236]:

$$S = S_0 \int_{R^3} p(\mathbf{r}|\tau) e^{-2\pi i \mathbf{q}^T \mathbf{r}} d\mathbf{r} \tag{2.1}$$

where $\mathbf{q} = \gamma\delta\mathbf{G}/2\pi$ and S_0 is the reference signal from the standard spin echo protocol. The quantity γ is an intrinsic parameter of hydrogen protons, and $\mathbf{G}$ is the diffusion-sensitizing gradient, which is usually decomposed into its magnitude G and orientation $\mathbf{g} = \mathbf{G}/G$. While there are several experimental parameters to consider, they are usually summarized by the gradient orientation (or *b-vector*) $\mathbf{g}$ and the *b-value* $b = \gamma^2 G^2 \delta^2 (\Delta - \delta/3)$, given the pulse duration δ and separation Δ. To avoid contamination by bulk flow, the magnitude of the full complex MR signal is usually taken, which notably gives rise to the Rician distribution of noise in diffusion MRIs [77].

2.2.3 Diffusion Models

While this basic mechanism of diffusion sensitization was worked out early in the history of MRI, it wasn't until several decades later that this became practically useful for imaging, largely due the complexities of mixing gradients for both spatial and diffusion encoding in a single session [117]. Following these developments, the field saw a surge in methods for modeling and interpreting diffusion MRIs.

The simplest model, which was already mentioned, represents diffusion with a single coefficient (ADC) where the diffusion-weighted signal is represented by:

$$S = S_0 e^{-bD} \tag{2.2}$$

This makes an assumption that the underlying diffusion probability density follows an isotropic Gaussian distribution. This is a fair assumption when looking at fluid with little imposed geometric

Table 2.1: A summary diffusion tensor indices. These are derived from the eigen-decomposition of the diffusion tensor. Indices derived from the sorted eigenvalues are particularly useful for characterizing the geometric shape of the tensor.

	Name	Description
PDD	Principal Diffusion Direction	Direction of greatest diffusion
AD	Axial Diffusivity	Magnitude of diffusion in the PDD
RD	Radial Diffusivity	Magnitude of diffusion perpendicular to the PDD
MD	Mean Diffusivity	Magnitude of total diffusion
FA	Fractional Anisotropy	Degree of anisotropy ranging from zero to one

structure, so the resulting ADC maps are useful for applications such as stroke imaging [147], in which the underlying pathology involves changes in tissue fluid content. However, there are many cases where tissue includes some relevant geometric structure that this model fails to capture.

The first model to go beyond the simple ADC representation is the diffusion tensor model [14] [170]. This model represents the diffusion pattern with Gaussian distribution, which allows the diffusion process to be described by any ellipsoidal shape:

$$S = S_0 e^{-b\mathbf{g}^T\mathbf{D}\mathbf{g}} \tag{2.3}$$

given a positive definite diffusion tensor $\mathbf{D}$ with six parameters. From an eigen-decomposition of $\mathbf{D}$, we can derive a number of simple properties based on the resulting decreasing eigenvalues $\lambda_1, \lambda_2, \lambda_3$ and their associated eigenvectors v_1, v_2, v_3 [15] [239]. Some of these measures are summarized in Table 2.1. Fractional anisotropy (FA) is the most commonly used measure and describes how far the tensor deviates from being isotropic and is given by:

$$FA = \sqrt{\frac{1}{2}}\frac{\sqrt{(\lambda_1-\lambda_2)^2+(\lambda_2-\lambda_3)^2+(\lambda_3-\lambda_1)^2}}{\sqrt{\lambda_1^2+\lambda_2^2+\lambda_3^2}} \tag{2.4}$$

FA is sensitive to changes in microstructural properties of brain white matter, such as fiber density, myelination, and fiber mixing; however, it should be noted that it is not specific, because changes in *any* of these properties will be observed as changes in anisotropy [146]. Mean diffusivity is another quantity that represents the overall size of the tensor (similar to the ADC) and is given by $MD = (\lambda_1 + \lambda_2 + \lambda_3)/3$. There are also axial and radial diffusivity given by $AD = \lambda_1$ and $RD = (\lambda_2 - \lambda_3)/2$, respectively, as well as many more [4] [239]. The principal tensor direction v_1 has also been found to be aligned with the orientation of large fiber bundles [171]. This feature is the

basis for constructing tractography-based models of white matter fiber bundles, which we'll discuss later.

While the tensor can be used for many complex white matter imaging tasks, a fundamental limitation is that cannot represent complex configurations of fibers [117]. These complex configurations are quite common due partial volume effects in large voxels and due to the presence of crossing fiber bundles in many parts of the brain [242] [218]. To handle these, a number of techniques have been developed, and they can be roughly grouped into non-parametric and parametric techniques.

Non-parametric techniques aim to characterize the diffusion process without making many assumptions about its underlying shape. These techniques take advantage of the Fourier relationship between the diffusion-weighted signal and the underlying diffusion probability density function (PDF). The first approaches in this area considered a simpler problem of solving for the diffusion on a spherical subsection of q-space, resulting in a simpler orientation distribution function (ODF) [219], which is the radial integral of the full PDF. The Q-ball technique for computing the ODF can do this efficiently by inverting a Funk-Radon transform that describes the relationship between the ODF and the diffusion signal, analogously to the Fourier relationship between the PDF and the diffusion signal [217]. Further improvements have been made by accounting for volumetric changes with a constant solid angle factor [3] and by performing deconvolution to obtain a fiber orientation distribution function [215]. With improvements to scanner technology, it also became possible to reconstruct the full PDF with a technique known as diffusion spectrum imaging (DSI) [236], which samples a dense grid of q-space instead of the sphere used to reconstruct an ODF. In general these approaches require many unique gradient orientations and high b-values to obtain good results. This is practical for scientific applications; however, it is less common to find this quality of data in clinical scans, which are often time constrained and limited to low field strengths and b-values.

Parametric techniques take a different approach and make substantial assumptions of the constituent parts of the diffusion process. The diffusion tensor model is an example of such a parametric model; however, many more parametric models have been developed to handle the complex properties of white matter. The simplest extension is to consider a multi-tensor model:

$$S = S_0 \sum_{j=1}^{N} f_j e^{-b\mathbf{g}^T \mathbf{D}_j \mathbf{g}} \tag{2.5}$$

which represents the signal attenuation by a linear combination of tensors [5] [207], each of which

may represent some population of fibers within a voxel. While very flexible, this model requires a multi-shell acquisition [186], which imposes some limitations for clinical use. However, there are many ways to constrain this model to improve its performance with single shell and more clinically oriented acquisitions. One such model is the ball-and-sticks model:

$$S = S_0 \left(f_0 e^{-db} + \sum_{j=1}^{N} f_j e^{-bd(\vec{g}^T \vec{u}_j)^2} \right) \tag{2.6}$$

which reduces the tensors to a single isotropic ball, multiple completely anisotropic sticks, and a shared diffusivity d [19]. This has the benefit of handling partial volume effects, isotropic diffusion, and differing proportions of fiber populations in each voxel. Later in this document we explore more computational approaches to work with the ball-and-sticks model. Other approaches for parametric modeling have been explored to obtain measures of features such as orientation dispersion [254] and other more exotic features [64]. This is all to show that there is a rich array of options when choosing a diffusion model, and that choice is often led by the limitations of the data and the problem being investigated. In the next section, we examine computational approaches for processing, analyzing, and visualizing diffusion MRI data.

2.2.4 Computational Methods

Computational tools are an essential part of dealing with modern medical imaging data. There are many scientific and clinical problems that can be addressed with these tools [136], but the fundamental tasks include image processing, registration, segmentation, visualization, and analysis. Image processing includes a variety of tasks that operation only on image data, such as interpolation, filtering, denoising, and fusion [120]. Image registration involves the alignment of two images into a common coordinate system [129] [262]. For example, when combining images from differing modalities, such as PET and CT, that alignment can be a simple rigid motion or some more complex deformation. Image segmentation is performed to decompose an image into a number of segments, which may correspond to some anatomy or region of interest [159] [137]. Visualization provides a way for people to interpret imaging data, and this can take the form of traditional slices or more complex 3D views that include a combination of geometric and volumetric data [188]. The area of image analysis considers a more general task of extracting meaningful information from the image (and usually includes the use of those other tasks as well). The result could be a quantitative

measurement of some feature in an image or a more qualitative understanding of the content.

While these methods are generally useful, there are a few unique features of the diffusion measurement to consider when applying them to diffusion MRIs [121]. The primary difference is that the diffusion measurement is made relative to the image's underlying spatial coordinate system. This contrasts with standard T_1- or T_2-weighted MRI or CT scans that measure scalar quantities at each voxel. This has an important effect on some of these basic computational tools, such as image processing, registration, etc. For example, when registering diffusion tensor images, any local rotation in the spatial coordinates must also be applied to the tensor model at each voxel [6]. Furthermore, when aligning images any intensity-based similarity metric must defined geometrically, e.g. between tensors [255]. Image processing can still be performed on the diffusion-weighted volumes; however, approaches can also be developed to instead process the diffusion models fit at each voxel. These model-based approaches are potentially more efficient and can preserve desirable properties, such as tensor anisotropy [165]. Diffusion models also pose a challenge for visualization, as traditional raster images can't convey the complex geometry of some models. Instead, glyphs can be used to visualize the shape of the diffusion pattern at each voxel, for example with tensor ellipsoids [112] or superquadratics [104].

In addition to these unique image processing, registration, and visualization algorithms, tractography algorithms also provide a unique way to perform segmentation of diffusion MRIs. The basic idea is to derive curve models that represent the fiber orientation field of the diffusion volume, which can be done with either local or global approaches. Local approaches generate curves in step-wise fashion, as in the FACT algorithm [142], or through streamline integration [16] [259]. Global approaches instead consider large optimization problems across the whole brain [92], potentially resolving some of the problems of local approaches but also introducing major computational challenges. Tractography algorithms can also be consider either deterministic, such as those already discussed, or probabilistic [57]. These tractography algorithms are important for reconstructing fiber bundle anatomy and constructing structural anatomical networks [78] [155].

2.2.5 Clinical and Scientific Applications

These computational methods play a crucial role in a number of clinical and scientific applications [87]. Clinical applications include diagnostic imaging, surgical planning and guidance, and

neuropsychiatric imaging. In diagnostic imaging, diffusion MRI can provide a contrast for characterizing stroke [233] and tumors. In surgical planning and guidance, diffusion MRIs can be used to localize and visualize eloquent structures that should be avoided during surgery [106]. In psychiatric applications, diffusion MRI can be used to study anatomical changes related to mood disorders, mild cognitive impairment, Alzheimer's disease, among many others [212]. Scientific applications also employ diffusion MRI to better understand the anatomy of fiber bundles and their functional features [97]. This can include the study of normal human brain anatomy [96] [237], comparative anatomy with non-human animals [184] [91], and common features found in population-average atlas data [135] [213]. Studies have also validated tractography models by comparison with blunt dissection [190] [140]. In the rest of this thesis, we explore some of these applications, using each as not only a metric to judge our tools but also an open research area to better understand.

Chapter 3

Supervoxel Segmentation of Diffusion Tensor Images

Diffusion MR imaging enables the quantitative measurement of water molecule diffusion, which exhibits anisotropy in brain white matter due to axonal morphometry and coherence [170]. Consequently, diffusion-based measures of fibers passing through each voxel can be estimated from the MR signal [13]. The diffusion tensor is the most well-studied and commonly used model for this purpose, and it can be decomposed into a number of features that reflect aggregate properties of tissue microstructure [17]. These features are useful for scientific and clinical studies of brain white matter anatomy in both individuals and populations [50]. In particular, variation in diffusion tensor indices reflects normal variation in a population, such as those arising from differences in age, sex, and cognition. This is also the case for studying anatomical changes in neurological conditions, such as stroke and traumatic brain injury, as well as neuropsychiatric conditions, such as schizophrenia, depression, and bipolar disorder.

To statistically characterize these features, diffusion tensor imaging studies require anatomically-comparable measurements to be made across all participants in the study [143]. Atlas-driven approaches are a standard solution, in which the set of images from the study is spatially normalized to some reference atlas through image registration [145]. The atlas and the associated spatial mappings can then be used to probe features of white matter across the population using one of a number of segmentation methods. In diffusion studies, this can be done with either voxel-based analysis

or tractography-based analysis [85] [121]. Voxel-based analyses can either look at individual voxels or regions-of-interest, and tractography-based analyses instead look at features of geometric models representing large-scale fiber bundle anatomy [50] [251]. In this chapter, we contribute to the area of voxel-based analysis with a novel segmentation algorithm and evaluate its potential use in population studies. While tractography-based methods are of great value for imaging studies, we restrict the scope of this chapter to the evaluation of voxel-based methods.

The main contribution of this chapter is the development and evaluation of a supervoxel-based analysis for diffusion tensor imaging studies. The goal of this method is to automatically subdivide the whole brain into *supervoxels*, which are larger than single voxels but restricted size to provide spatial localization. This procedure was designed to algorithmically capture how regions-of-interest are manually drawn, typically based on diffusion tensor colormaps. To accomplish this, we develop an efficient clustering algorithm for grouping voxels based on their principal diffusion direction and integrate it into a pipeline for atlas-driven analysis of diffusion tensor images using either the standard voxel-based or skeleton-based techniques. The result is potentially useful for population imaging studies, where it may be particularly useful for use with study-specific templates. We evaluate the proposed approach with tests using synthetic and *in vivo* brain imaging datasets. This includes applications of the segmentation algorithm to a population atlas as well as individual subjects.

The rest of the chapter is as follows. In Secs. 3.1 and 3.2, we review related work and background material. In Sec. 3.3, we describe the proposed method for supervoxel segmentation and how this can be used in population imaging studies. In Sec. 3.4, we present evaluation results from three experiments. In Secs. 3.5 and 3.6, we discuss the results and conclude.

3.1 Related Work

In this section, we briefly review prior work on voxel-based atlas-driven diffusion tensor imaging studies and discuss how they relate to the proposed method.

The earliest and simplest approach for diffusion tensor imaging studies is to manually select portions of a volume to analyze. This can include the selection of single matching voxels across all subjects, as well as larger volumetric masks that are drawn based on orientation-coded colormaps. While effective, these manual approaches are tedious and potentially suffer from biases due to intra-

and inter-rated variability. The proposed supervoxel approach produces masks that are similar to manually drawn region-of-interest in principle, as voxels are grouped based on uniformity of fiber orientations. However, because it is automated, the proposed method can provide higher reproducibility and also allow more areas of the brain to be studied.

There has been much past work on automated whole-brain image analysis, however, and the earliest solution is known as voxel-based analysis (VBA) [197]. In this approach, the imaging data from the population is spatially normalized to an atlas, and individual voxels in this space are statistically characterized across the population. This is inspired by earlier work on voxel-based morphometry [7], which aimed to characterize local shape variation by analyzing the deformation field between the atlas and subject. The major challenge in using this approach is the effect of misregistration [8], and the standard solution is to smooth the image with a large Gaussian kernel to reduce the high-frequency misalignment between subjects; however, this has been also shown to introduce some statistical challenges in diffusion MR image analysis [99]. Another challenge is large dimensionality, which is usually addressed by analyzing voxels independently and performing statistical corrections to account for false positives [125]. The proposed supervoxel approach can be used in conjunction with voxel-based analysis to provide regions of interest to study, while still allowing patterns to be characterized across the whole brain. This supervoxel representation also has a dimension that is orders of magnitude smaller than the full volume, which can allow for more complex statistical models to be used, e.g. those that include long-range spatial correlations.

Another standard tool for whole-brain analysis is the skeleton-based approach. This technique is known as tract-based spatial statistics (TBSS) and includes several additional steps in the VBA pipeline [194]. First, a white matter skeleton is created from the atlas volume. This is meant to restrict the measurements to a white matter core, where there are few partial volume effects. Then, each deformed subject is skeletonized and projected onto the atlas skeleton. This projection step is meant to fix some residual registration errors and to avoid the VBA smoothing step. In contrast, the proposed supervoxels are chosen to include groups of voxels with similar fiber orientation, which is important because these areas are exactly where misregistration is likely to occur. By aggregating within these homogeneous regions, the approach may also avoid subtle errors due to misregistration.

There are also automated region-of-interest approaches in which manually drawn labels from an atlas are transferred to each subject in the population [72]. The resulting masks can then be combined with spatially normalized tensor images to produce region-averaged diffusion measures.

These regions are usually in major bundles or deep white matter structures with the most reliable registration [144], and they can provide some complementary information to voxel-based analysis [197]. The proposed approach is similar to this; however, it may be particularly useful for study-specific templates, where it may be desirable to learn the regions from the atlas itself, rather than import them from another existing one. The skeleton approach can also be used in conjunction with these regions-of-interest, and we explore how a similar process can be performed with the supervoxels method described in this chapter.

Computationally, the proposed supervoxel segmentation algorithm is also related to a number of previous works in diffusion imaging and computer vision. Wiegell et al. proposed a voxel clustering method for segmenting the thalamic nuclei [243], and this method uses a similar k-means style approach to clustering in the spatial-tensor domain. In contrast, the proposed approach is designed to learn the number of clusters from the data and uses fiber orientation for grouping voxels, which is a reasonable choice for segmenting white matter, which is more anisotropic than gray matter. A related work of Bloy et al. also examined the segmentation of supervoxel-like regions for diffusion imaging studies [26]. However, their work was formulated for orientation distribution functions and performs clustering with a normalized cut algorithm. The idea of superpixels is also a well explored topic in the computer vision literature [222], and a number of medical image segmentation algorithms have benefited from using a supervoxel image representation [226] [124]. Finally, preliminary work from this project has been presented at conferences [37], but this previous work did not include the full experimental results and comparative evaluation included here.

3.2 Background

In this section, we'll briefly review the standard steps used for image analysis in diffusion tensor imaging studies.

First, we examine diffusion MR imaging with the tensor model. This approach considers a simple but powerful representation of the diffusion signal [13], in which the diffusion pattern is characterized by a Gaussian distribution that can reflect any ellipsoidal shape of diffusion. The predicted signal under this model is as follows:

$$S = S_0 e^{-b\mathbf{g}^T \mathbf{D} \mathbf{g}} \tag{3.1}$$

given a baseline signal S_0, gradient encoding direction $\mathbf{g}$, b-vector b, and a six-parameter positive definite diffusion tensor $\mathbf{D}$. From an eigen-decomposition of $\mathbf{D}$, a variety of diffusion indices can be derived [170] [239]. These tensor indices reflect size and shape features of the tensor and are sensitive (but not specific) to microstructural properties of white matter tissue, such as myelination, fiber density, etc [17]. However, the tensor model is also severely limited in representing more complex features, such as partial volume mixing of different tissues and crossings of fiber bundles.

The goal of many clinical imaging studies is to accurately measure these tensor indices and statistically relate them to demographic and pathological characteristics of a population [84]. An important step in this process is spatial normalization, in which image registration is performed to align a collection of images obtained from the population [98]. This step has some unique features for tensor imaging studies, as the criteria for alignment can be derived from the geometric form of the tensor model and the measurement coordinate system must also be transformed [6]. Prior work has also found that the statistical power of these studies can be improved by using a deformable registration [251] and study-specific atlas [103], which is derived from the population average of the study dataset. However, one limitation of this study-specific approach is that annotations from standard atlases may no longer be available. For this reason, automated atlas segmentation algorithms, such as the one we propose, are potentially useful. In later experiments, we apply the proposed method using the state-of-the-art in deformable tensor-based registration and use the study-specific atlas approach [231].

3.3 Methods

In this section, we describe the proposed supervoxel segmentation algorithm and how it can be used in atlas-driven diffusion tensor imaging studies.

3.3.1 Voxel Clustering

We first motivate the proposed method by describing how clustering algorithms can be used for regional segmentation. The basic idea is to consider each voxel of a volume to be a data point, then a clustering algorithm can be used to group voxels based on their spatial proximity and signal similarity. The simplest solution to this problem is to apply k-means clustering with a joint representation of the position and value of each voxel, e.g. $\mathbf{x} = [p_x, p_y, p_z, s]$ given position vector $\mathbf{p}$ and

signal intensity s. The k-means clustering problem is then:

$$\pi = \underset{\pi}{\operatorname{argmin}} \sum_{i}^{N} \|\mathbf{x}_i - \hat{\mathbf{x}}_{\pi(i)}\|^2 \tag{3.2}$$

where N is the number of voxels, K is the number of clusters, $\pi(i)$ is the cluster label of the i-th voxel, and $\hat{\mathbf{x}}$ is a cluster prototype. This can be solved with an iterative procedure and can be efficiently scaled to large volumetric image datasets. It also tends to group voxels in accordance with large gradients in the image, provided a sufficiently large value of K is used. Because this only minimizes the within cluster sum-of-squares, the resulting regions also tend to have consistent sizes. This is perhaps desirable for localizing changes in tissue properties, since it puts a bound on the size of the region being compared across a population. However, the use of this approach still faces a few challenges. First, it requires the number of clusters to be known a priori. Second, this cannot take advantage of the geometric structure of diffusion models, where the Euclidean distance is not the best choice for measuring within-cluster variance.

3.3.2 Supervoxel Segmentation

Next, we present a supervoxel clustering algorithm, which is the main contribution of the chapter. This is designed to address the challenges described in the previous section with two elements. First, we propose to use an extension of the k-means algorithm known as the DP-means algorithm, whose name suggests its theoretical connection to Dirichlet Process mixture models. Unlike the k-means algorithm, the DP-means algorithm learns the number of clusters from the dataset, subject to a regularization parameter λ. While k-means requires a fixed number of clusters, this parameter instead controls the size of the clusters, a feature we'll demonstrate later in an evaluation. Second, we propose to modify the minimization problem to use a model-based measure of discrepancy in place of the Euclidean distance. This is necessary because simply applying the Euclidean distance to a diffusion model's parameters cannot account for complex geometric features, such as fiber orientations. Taken together, these extensions result in the following optimization problem:

$$\pi = \underset{\pi}{\operatorname{argmin}} \sum_{i}^{N} \alpha\|\mathbf{p}_i - \hat{\mathbf{p}}_{\pi(i)}\|^2 + \beta d^2(M_i, \hat{M}_{\pi(i)}) + \lambda K \tag{3.3}$$

given diffusion model M, prototype $\hat{M}$, model-based discrepancy d^2, and regularization parameter

λ. This decomposes Eq. 3.3 into a linear combination of three terms. The first term is the familiar k-means term for the voxel positions; the second term is the term that groups voxels based on diffusion model similarity; the third term adaptively selects the number of clusters. This formulation is sufficiently general to work with any diffusion model; however, we'll consider how this can be solved with the diffusion tensor model. We choose to define d^2 to reflect differences in orientation of the tensor principal diffusion direction $\mathbf{v}$, specifically $d^2(\mathbf{v}, \hat{\mathbf{v}}) = 1 - (\mathbf{v}^T\hat{\mathbf{v}})^2 = \sin^2(\theta)$. A procedure for solving the resulting minimization problem (Eq. 3.3) is listed in Algo. 1. The result is perhaps similar to the typical process of manually drawing a region-of-interest in reference to an orientation-coded colormap. Theoretically, this algorithm is closely related to axial clustering, which guarantees its convergence to a local minima. The term *axis* refers to unsigned orientations, i.e. $\mathbf{v} \sim -\mathbf{v}$ and $\|\mathbf{v}\| = 1$, and applies equally well to the tensor principal diffusion direction. This is topic is further discussed in the Appendix.

Algorithm 1: joint spatial-axial DP-means clustering

Input:
$(\mathbf{p}_1, \mathbf{v}_1), ..., (\mathbf{p}_N, \mathbf{v}_N)$: input position/axial direction pairs,
α, β, λ: objective weighting parameters

Output:
K: number of clusters,
$\pi_1, ..., \pi_N$: labels,
$(\hat{\mathbf{p}}_1, \hat{\mathbf{v}}_1), ..., (\hat{\mathbf{p}}_K, \hat{\mathbf{v}}_K)$: cluster centers

`/* Initialize */`
$K \leftarrow 1,\ \hat{\mathbf{p}}_1 \leftarrow \sum_i \mathbf{p}_i/N,\ \hat{\mathbf{v}}_1 \leftarrow prineig\left(\sum_i \mathbf{v}_i\mathbf{v}_i^T/N\right)$
while *not converged* **do**
  `/* Assign cluster labels */`
  for *i=1* **to** N **do**
    for *j=1* **to** K **do**
      $D_{ij} \leftarrow \alpha||\mathbf{p}_i - \hat{\mathbf{p}}_j||^2 + \beta\left(1 - \left(\mathbf{v}_i^T\hat{\mathbf{v}}_j\right)^2\right)$
    `/* Optionally increment clusters */`
    if $\min_j D_{ij} > \lambda$ **then**
      $K \leftarrow K + 1,\ \hat{\mathbf{p}}_K \leftarrow \mathbf{p}_i,\ \hat{\mathbf{v}}_K \leftarrow \mathbf{v}_i,\ \pi_i \leftarrow K$
    else
      $\pi_i \leftarrow \text{argmin}_j\, D_{ij}$
  `/* Update cluster prototypes */`
  for *j=1* **to** K **do**
    $\hat{\mathbf{p}}_j \leftarrow \left(\sum_i \delta(j, \pi_i)\mathbf{p}_i\right) / \sum_i \delta(j, L_i)$
    $\hat{\mathbf{v}}_j \leftarrow prineig\left(\left(\sum_i \delta(j, \pi_i)\mathbf{v}_i\mathbf{v}_i^T\right) / \sum_i \delta(j, \pi_i)\right)$
return

3.3.3 Application to Imaging Pipeline

We now outline practical aspects of how the proposed segmentation algorithm can be used in diffusion tensor imaging studies.

In the first step, each subject of a population is registered to a template image, for example, using the tensor-based registration in DTI-TK [253] [255]. This provides an average tensor image and a collection of deformations from the atlas to each subject. Then, a white matter mask can be created by applying a threshold, e.g. 0.2, to the atlas FA image with optional manual cleaning. In practice, this may exclude some non-white matter with complex fiber configurations, as they are not well-characterized by a single tensor. Then, the proposed segmentation algorithm can be applied as described in Algo. 1 to produce a mapping of supervoxel regions. The atlas regions are cleaned to produce the final segmentation. This includes connected components labeling [246] to separate similarly labeled but disconnected supervoxels; for example, we found this to happen with the bilateral cingulum bundles. We also exclude small components below a threshold volume of 50mm^3, which are unlikely found in all subjects.

In the next step, the supervoxel region labels can be propagated to each participant in the study. For this, each participants diffusion tensor volume is first registered to the atlas, and white matter is masked with an FA threshold, e.g. 0.2. Then the intersection of the white matter mask and the supervoxel masks is computed to create subject-specific supervoxel masks. Then, regional averages of scalar tensor indices can be computed in each supervoxel, including fractional anisotropy (FA) and mean diffusivity (MD) which can be used for follow up statistical analysis.

3.4 Experiments and Results

This section describes the evaluation of the method with three experiments, including a discussion of the datasets, experimental design, evaluation metrics, and results. The first experiment tests the clustering algorithm with synthetic data; the second tests the segmentation algorithm with a publicly-available diffusion tensor atlas; the third tests the segmentation algorithm with individual subject data across multiple scans. All statistical analysis and plotting was implemented using R 3.1.1 and ggplot2.

3.4.1 Synthetic Data Experiments

Design: The goal of this experiment is to investigate the choice of the λ clustering parameter. For this, we synthesized axial data and tested the performance of the axial DP-means algorithm across varying numbers and sizes of clusters. Here, we omit the spatial component and only test the relationship between λ and the axial clustering. The number of clusters N ranged from 3 to 10 and were generated by sampling a Gaussian and normalizing with "size" σ ranging from 0.1 to 0.3. We evaluated ground truth agreement with the adjusted mutual information score (AMI) [223], a statistical measure of similarity between clusterings that accounts for chance groupings and takes a maximum value of one when clusterings are equivalent.

Results: Results are shown in Fig. 3.1. For a constant number of clusters, we found the optimal choice of λ increased with cluster size σ. For a constant cluster size σ, we found the optimal λ to be relatively stable across variable numbers of clusters N. Fig. 3.1 also shows examples of the clustering for the two conditions and plots of the relationship between the AMI and λ. We found our implementation to converge in fewer than 20 iterations on average. All cases except one found the correct number of clusters for the optimal λ.

3.4.2 Population Atlas Segmentation

Design: The goal of this experiment is to apply the proposed method to a population atlas and test a range of parameter values. For this, we applied the proposed method to a publicly-available diffusion tensor atlas and varied the β and λ parameters. We used the IXI aging brain atlas, which was constructed by deformable tensor-based registration with DTI-TK [257]. We applied the method described in Sec. 3.3.3 with λ ranging from 10 to 40 and β ranging from 0 to 30. For each set of parameters, we measured the number of clusters and the average cluster volume for each condition.

Results: The results show that as λ increased the number of regions decreased and the region volume decreased, and as β increased the number of clusters slightly increased and the region volume decreased (Fig. 3.2). Visualizations were created from a representative set of segmentation results with $\lambda = 25$ and $\beta = 15$ (Fig. 3.3). From slice views, we found the segmentation to reflect known

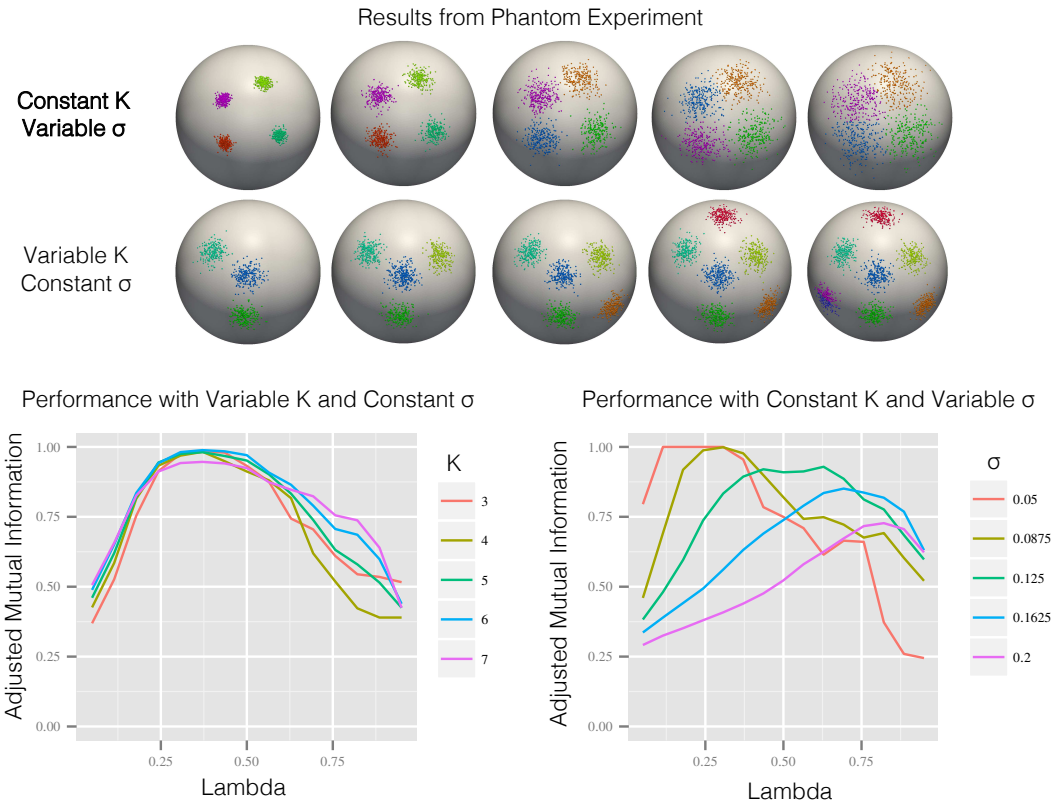


Figure 3.1: Results from the first experiment described in Sec. 3.4.1, showing visualizations of the synthetic axial data and the optimal clusterings computed from the proposed method. The top shows results for variable cluster size $\sigma = \{0.05, 0.0875, 0.125, 0.1625, 0.2\}$, and constant number of clusters $N = 4$. The bottom shows results for constant cluster size $\sigma = 0.10$, and variable number of clusters $N = \{3, 4, 5, 6, 7\}$. The bottom right shows a single mislabeled cluster, possibly caused by finding a local minima in the optimization. First experiment: clustering performance as a function of cluster penalty parameter $\lambda \in [0, 1]$ given ground truth generated with cluster size σ, and number of clusters N. We measured the adjusted mutual information (AMI), a statistical measure takes a maximal value when clusterings are equivalent. Shown are plots of the AMI vs. λ for two conditions. The first tested with variable $\sigma = \{0.05, 0.0875, 0.125, 0.1625, 0.2\}$ and constant $N = 4$. The second tested with constant $\sigma = 0.10$ and variable $N = \{3, 4, 5, 6, 7\}$. These results indicate that the optimal λ depends more on σ than N, suggesting that performance may depend on the noise level and may diminish when differing clusters sizes are present. Also note that the maximum AMI decreases with increasing σ, which may be due to cluster overlap or over-sensitivity in the AMI measure.

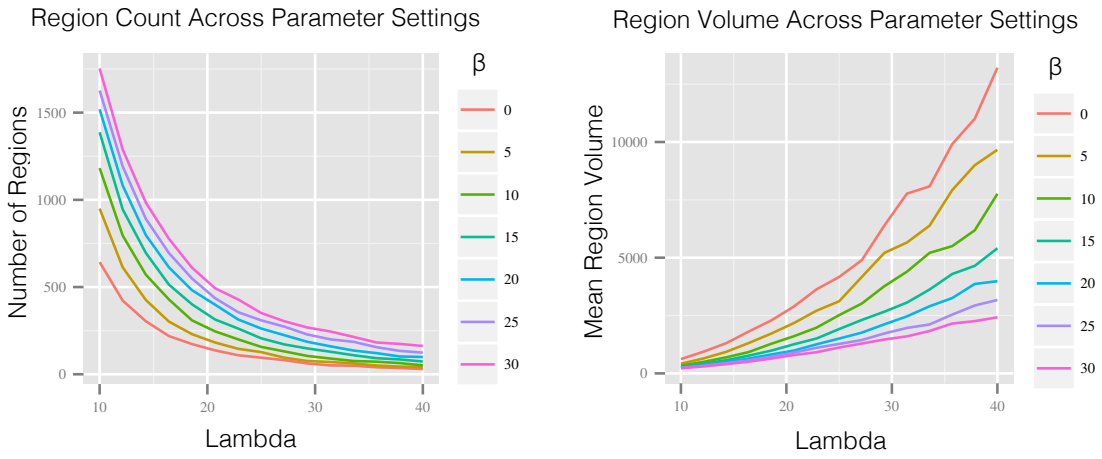


Figure 3.2: Results from the second experiment described in Sec. 3.4.2, showing quantitative segmentation results from population atlas data. The left and right panels show number of clusters and mean cluster volume (mm^3), respectively, as a function of the parameters $\lambda \in [10, 40]$ and $\beta \in [0, 30]$. The results show that increasing λ lead to increased volume and a decreased number of regions. Similarly, the results show that increasing β increased the number of clusters while decreasing the cluster volume.

anatomical boundaries, such as the cingulum/corpus callosum and corona radiata/superior longitudinal fasciculus. By overlaying fiber models, we see the region boundaries tended to coincide with large changes in fiber orientation. From an boundary surface rendering, we found lateral symmetry and a separation of gyral white matter from deeper white matter. Our serial implementation on a 2.3 GHz Intel i5 ran in three minutes end converged in 125 iterations.

3.4.3 Subject-specific Segmentation

Design: Under an IRB-approved protocol, diffusion-weighted MRIs were acquired from two healthy volunteers in two imaging sessions. Imaging was conducted on a GE 1.5T scanner with 2x2x2mm voxels, 64-directions, b=1000 s/mm^2, and resolution 128x128x72. Volumes were preprocessed to correct for motion and eddy current artifacts with FSL 5.0 [195], including a reorientation of the b-vectors to account for the rotation induced by the alignment. Skull stripping was performed with FSL BET, and diffusion tensors were fit using FSL DTIFIT. Each subject was deformed to the IXI atlas described in Sec. 3.4.2 using deformable registration in DTI-TK using finite strain tensor reorientation and the deviatoric tensor similarity metric. Each subject's tensor image was registered to atlas space using the resulting deformation, and results were segmented using the methods described in Sec. 3.3.3.

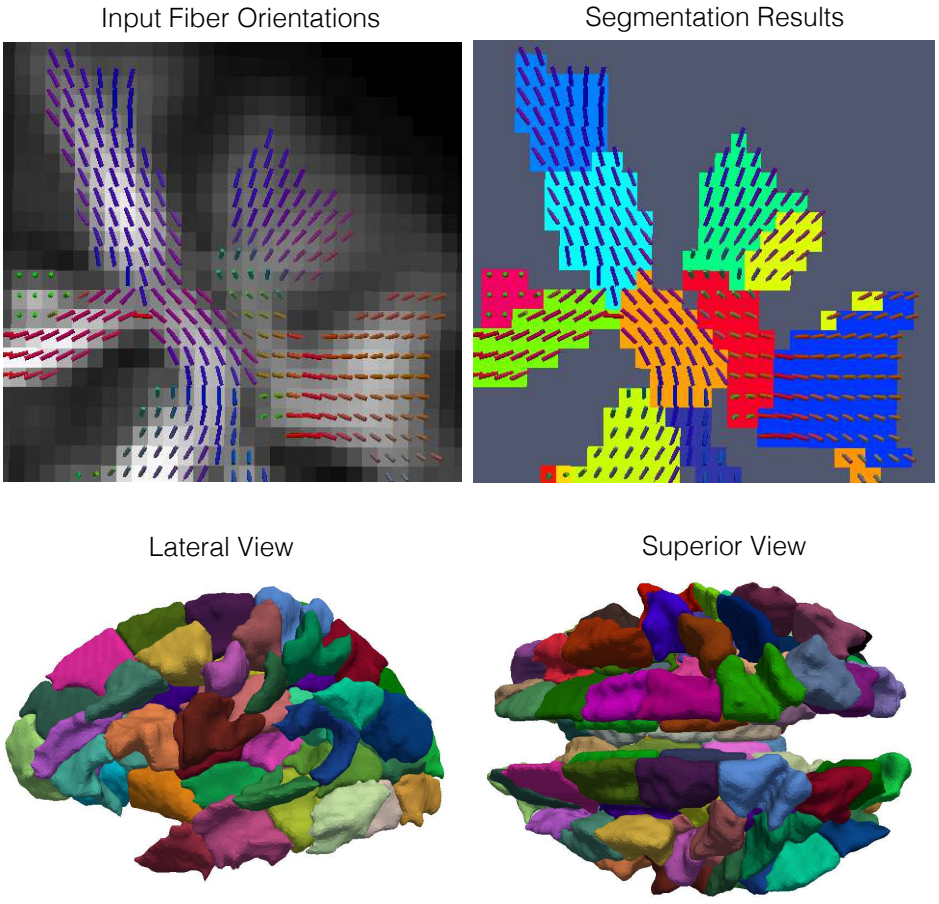


Figure 3.3: Results from the second experiment in Sec 3.4.2 showing segmentation results from atlas data using $\lambda = 25$ and $\beta = 15$. The top left panel shows the input atlas fiber orientations, and the top right panel shows a slice view of the region labels, which illustrates separation of major bundles such as the corpus callosum, cingulum bundle, and superior longitudinal fasciculus. The top panels show boundary surfaces of the regions, illustrating the symmetry and separation of gyral and deep white matter.

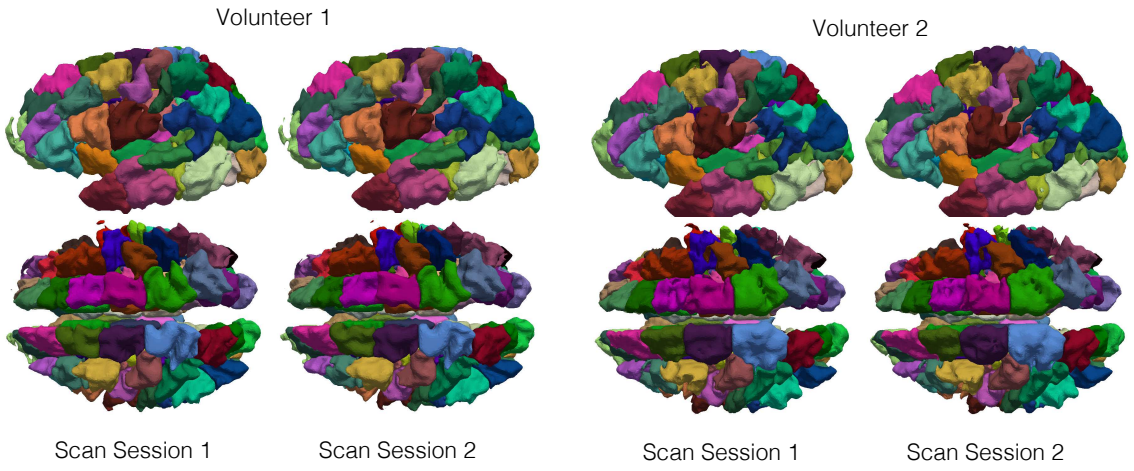


Figure 3.4: Results from the third experiment in Sec 3.4.3 showing segmentation results from two individual subjects across two scans each. The results show consistency in the segmentation across both scanning sessions and differences in individual anatomy.

Results: The results show qualitatively similar segmentation results across individuals and scanning sessions (Fig. 3.4). From an inspection of individual regions, there was visible morphometric differences between subjects; however, this did not cause substantial differences in segmentation results, which were somewhat larger in scale than the inter-subject variations. The results were very similar across scans as well, even showing similar artifact in the protruding anterior portion of white matter in the first volunteer. In general, the results are also qualitatively similar to the atlas segmentation results in Sec. 3.4.2.

3.5 Discussion

This chapter sought to develop and evaluate a method for supervoxel-based analysis in atlas-driven diffusion tensor imaging studies. In particular, we examined its theoretical aspects and explored its application to population imaging studies. The general findings were first that the underlying algorithm is efficient and can automatically delineate regions that reflect coherent patterns of fiber orientations, and second, that its application to population studies is both reliable and sensitive to anatomical variation due to aging.

The first two experiments investigated the characteristics of the segmentation algorithm, which helped to provide a better understanding of the effect of various parameters. Experiment 1 specifically showed how the λ parameter allows the number of clusters to be learned from the dataset. By

comparing the performance across different numbers and sizes of clusters, we found this parameter controls the size of the segmented region, unlike the key parameter K in the similar k-mean algorithm, which explicitly selects the number of clusters. We also found that the other parameter β allows for the size to be controlled as well, and that higher values encouraged regions to be more uniform in fiber orientation.

This also has potential to improve the statistical power when statistical models are fit independently for each voxel or region, as the supervoxel analysis would have orders of magnitude fewer comparisons than the full skeleton-based approach. In particular, this is relevant to imaging genetics studies, which have a substantial multiple comparisons problem due to the modeling across both the brain and the genome.

There are also some important limitations and open problems to consider. First, the clustering algorithm tends to break major bundles into multiple supervoxels. While this may be useful for localization, the exact position of the breaks between supervoxels may be somewhat arbitrary, and the statistical power could be reduced if it happens to coincide with some hypothetical pathology. The approach also only derives the segmentation solely from the population average tensor image, which takes along all of the challenges of the registration algorithm. This could pose a challenge for segmenting structures in superficial white matter, which have more spatial variability due to misalignment. The application of this kind of approach to multi-compartment diffusion modeling is both an interesting and challenging open problem. This could perhaps be approached by clustering compartments instead of voxels; however, there is then an additional task of matching compartments from each subject to the atlas. Finally there are a variety of neurological and neuropsychiatric clinical applications in which this could be explore [122].

3.6 Conclusion

In this chapter, we have explored a method for supervoxel-based analysis of diffusion tensor images, which is designed for use in population imaging studies. This is meant to provide an additional analytical layer on top of existing voxel-based and skeleton-based analyses, in particular, those based on study-specific brain templates. The proposed segmentation algorithm attempts to simulate the process of drawing regions of interest based on fiber orientations, and our theoretical analysis and experiments provided some insight into the computational techniques involved. Looking forward, this

approach may provides a computational tool for quantifying whole-brain patterns of diffusion-based anatomical variation and has potential applications to a variety of clinical neuroimaging studies.

3.7 Appendix

This section discusses some theoretical aspects of the clustering algorithm formulated in Eq. 3.3, in particular, its derivation from related probabilistic models. The main point to consider is the representation of the diffusion data by fiber orientations. These can be statistically modeled using the Watson distribution [200]. Because the Watson forms an exponential family [9], there is an associated Bregman divergence that can be used for hard clustering of fiber orientations, or equivalently, axial variables [11]. These hard clustering algorithms tend to be more efficient, scalable, and easy to implement, at the cost of some flexibility in data modeling. The Bregman divergence can also be used to derive a DP-means clustering algorithm from the asymptotic limit of a Dirichlet process mixture [108] [95], providing a data-driven way to select the number of clusters.

Next, we build on these ideas to derive hard clustering for data that has both spatial and directional components, which are modeled using the Gaussian and Watson distribution, respectively. The isotropic Gaussian distribution is defined on $\mathbb{R}^n$ can be used to represent spatial data. Its probability density $\mathcal{N}$ and associated Bregman divergence $D_{\mathcal{N}}$ are given by:

$$\mathcal{N}(\mathbf{p}|\mathbf{q},\sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{1}{2\sigma^2}||\mathbf{p}-\mathbf{q}||^2\right) \tag{3.4}$$

$$D_{\mathcal{N}}(\hat{\mathbf{p}},\mathbf{p}) = \frac{1}{2\sigma^2}||\hat{\mathbf{p}}-\mathbf{p}||^2 \tag{3.5}$$

given input positions $\mathbf{p},\hat{\mathbf{p}} \in \mathbb{R}^n$, mean position $\mathbf{q} \in \mathbb{R}^n$, and constant variance parameter $\sigma^2 > 0$. The associated Bregman divergence $D_{\mathcal{N}}$ is the scaled Euclidean distance between two positions, an observation which gives a probabilistic interpretation to the k-means and DP-means algorithms, which are asymptotic limits of Gaussian mixture and Dirichlet process mixture models, respectively [108]. The Watson distribution is analogous to a Gaussian but defined on the hypersphere $S^{n-1} \subset \mathbb{R}^n$ with anti-podal symmetry, i.e. $\mathbf{d} \sim -\mathbf{d}$. Its probability density W and associated Bregman divergence D_W are given by:

$$W(\mathbf{d}|\mathbf{v},\kappa) = \frac{\Gamma(n/2)}{(2\pi)^{n/2} M(\frac{1}{2},\frac{n}{2},\kappa)} \exp\left(\kappa \left(\mathbf{v}^T\mathbf{d}\right)^2\right) \tag{3.6}$$

$$D_W(\hat{\mathbf{d}},\mathbf{d}) = \kappa \frac{M(\frac{1}{2},\frac{n}{2},\kappa)}{M'(\frac{1}{2},\frac{n}{2},\kappa)} \left(1 - \left(\hat{\mathbf{d}}^T\mathbf{d}\right)^2\right) \tag{3.7}$$

for input axial directions $\mathbf{d}, \hat{\mathbf{d}} \in S^{n-1}$, mean axial direction $\mathbf{v} \in S^{n-1}$, and Kummer's confluent hypergeometric function $M(a,b,z)$. We also assume a constant and positive concentration parameter $\kappa > 0$. The associated Bregman divergence D_W is then a scaled sine-squared angular distance, which is equivalent to the measure used for diametrical clustering [58]. Taken together, these divergences can be used to define the spatial-axial DP-means clustering problem in Eq. 3.3. The optimization routine in Algo. 1 is then derived based on the analysis of Jiang et al. [95]. This will only converge to local minima, so to ensure the solution is not a poor local minima, we include five random restarts with shuffling of the inputs and take the best result.

Chapter 4

A Comparative Evaluation of Voxel-based Spatial Mapping in Diffusion Tensor Imaging

Diffusion MR imaging enables the quantitative measurement of water molecule diffusion, which exhibits anisotropy in brain white matter due to axonal morphometry and coherence [170]. Consequently, diffusion-based measures of fibers passing through each voxel can be estimated from the MR signal [13]. The diffusion tensor is the most well-studied and commonly used model for this purpose, and it can be decomposed into a number of features that reflect aggregate properties of tissue microstructure [17]. These features are useful for scientific and clinical studies of brain white matter anatomy in both individuals and populations [50]. In particular, variation in diffusion tensor metrics reflects normal variation in a population, such as those arising from differences in age, sex, and cognition. This is also the case for studying anatomical changes in neurological conditions, such as stroke and traumatic brain injury, as well as neuropsychiatric conditions, such as schizophrenia, depression, and bipolar disorder.

To statistically characterize these features, diffusion tensor imaging studies require anatomically-comparable measurements to be made across all participants in the study [143]. Atlas-driven approaches are a standard solution, in which the set of images from the study is spatially normalized to a reference atlas through image registration [145]. The atlas and the associated registered images

can then be used to probe features of white matter across the population using one of a number of methods for spatial mapping. Commonly, this can be done with either voxel-based analysis or tractography-based analysis. Voxel-based analyses can either look at individual voxels or regions-of-interest (ROIs), while tractography-based analyses instead look at features of geometric models representing large-scale fiber bundle anatomy [50] [229] [227]. While tractography-based methods are of great value, we restrict the scope of this chapter to voxel-based methods.

The primary motivation of this chapter is the need to better understand the computational tools used in voxel-based diffusion tensor imaging studies. There are numerous steps in the standard population imaging pipeline, including image acquisition, preprocessing, spatial normalization, spatial mapping, and statistical analysis (Fig. 4.1). Choices made in each step have strengths and weaknesses, and there is value in understanding their net effect with respect to even basic characteristics such as measurement error. While much is known about how data acquisition, preprocessing, and image registration affect results, there have been relatively fewer studies to comparatively evaluate the spatial mapping step. In this study, we specifically aim to test how the choice of method for spatial mapping affects reliability, both overall and across specific areas of the brain. At the same time we also aim to investigate how much information is retained by each method to gauge the relative modeling power in real applications.

Towards this goal, the main contribution of this chapter is a comparative evaluation of methods for spatial mapping in voxel-based diffusion tensor imaging studies. To avoid confounding effects, these tests were conducted with a common dataset and a state-of-the-art tensor-based registration algorithm for spatial normalization and study-specific template construction. The evaluation strategy includes two experiments testing a total of eight methods for spatial mapping with four diffusion tensor metrics. The tested spatial mapping methods include voxel-based and skeleton-based analysis with and without smoothing, and two types of region-based analysis with and without skeletonization. The tested tensor metrics include fractional anisotropy, mean diffusivity, radial diffusivity, and axial diffusivity. The first experiment was conducted to characterize scan-rescan reproducibility and reliability and compare performance across methods, diffusion tensor metrics, and anatomy. For this, the coefficient of variation and intraclass correlation were measured in eight subjects with three repeated scans each. The second experiment conducted a complementary analysis to gauge the extent that each method can represent age-related change in white matter. For this, predictive models of age were estimated using principal component regression with a population of 80 normal

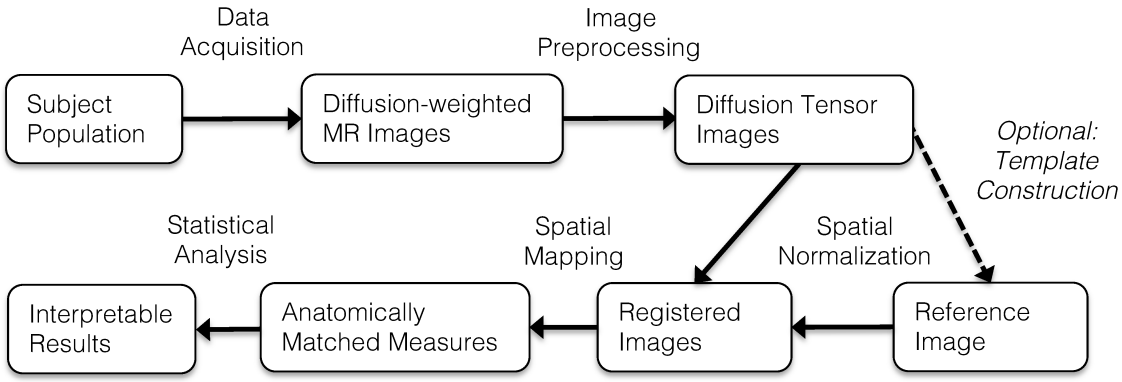


Figure 4.1: An illustration of the typical computational pipeline used for population-based diffusion tensor imaging studies. In this chapter, we evaluate methods for the spatial mapping step, which provides anatomically comparable measurements for statistical analysis.

subjects aged from 25 to 65 years. These models were compared based on goodness-of-fit and model selection criteria.

The rest of the chapter is as follows. In Secs. 4.1 and 4.2, we review related work and background material. In Sec. 4.3, we describe data acquisition, image analysis pipeline, the tested methods for spatial mapping, and the experimental design. In Sec. 4.4, we describe the results from the evaluation of scan-rescan reliability and reproducibility, and sensitivity to normal aging. In Sec. 4.5, we discuss the results and conclude.

4.1 Related Work

In this section, we briefly review related work that has evaluated methods for voxel-based analysis in diffusion tensor imaging studies. We first discuss those examining the relationship between specific methodological choices and the resulting reproducibility and reliability. We then discuss the use of applications in evaluation.

Numerous studies have thoroughly examined the relationship between reliability and imaging data acquisition parameters. For example, several works have looked at variation across scanner manufacturers and imaging units [44] [89] [261] [128] [67] and found acceptable reliability across sites with a common magnet strength. Furthermore, other studies have also shown reliability across magnet strengths ranging from 1.5T to 4T [28] [224] [211] [76] [162]. Studies that tested gradient strength have found reliable estimates of tensor metrics in each of a variety of gradients encoding

schemes [123]; however, there is evidence of possible bias in tensor metrics when combining estimates from different voxel sizes and gradient encoding schemes [162] [21], although bias correction [224] and covariate analysis [161] are possible solutions. Together, these results are especially important for conducting longitudinal and multi-center studies as well as accommodating scanner upgrades within an imaging unit.

In addition, previous work has examined the effect of preprocessing and image registration algorithms on reliability. Robust preprocessing that includes denoising, motion correction, and outlier rejection has been shown to improve reliability across scanners [127] [128] [157]. The choice of registration algorithm has also been shown to greatly affect reliability, specifically when comparing linear, deformable, and tensor-based registration [224] [211] [123]. Deformable tensor-based registration has been shown to perform better than registration with scalar maps, especially when used in combination with study-specific template construction [10]. Linear intra-subject registration has also been shown to improve reliability in longitudinal studies [127]. Overall, this indicates there is potential for significantly different outcomes based on the choice of preprocessing and registration, so it is important to have consistency in both applications and evaluations.

Each of these studies necessarily includes spatial mapping, either as a single method used in the pipeline or as part of a larger comparison of methods. The most common approaches are global histogram analysis [44], manually drawn ROIs [21] [123] [67], and standard atlas ROIs registered to each subject [89] [76] [221]. In addition to these studies, others have explicitly evaluated methods for spatial mapping, with a similar goal to this chapter. For example, evaluations of manually drawn ROI approaches have tested the reliability of different ROI shapes [27] and drawing methods [81] [168], and compared to a variety of global histogram measures [161] . Voxel-based analysis has also been evaluated to quantify the effects of filter size [99], software package [100], and to compare results with ROI-based methods [132] [197]. There has also been extensive testing of skeleton-based analysis to understand its strengths and limitations [10] [127] as well as comparing to voxel-based analysis and region-based analysis [224] [211]. This chapter builds on this prior work by expanding the range of methods simultaneously compared in evaluation. This includes not only a comparison of voxel-based and skeleton-based analysis, but also their combination with smoothing and ROI analysis. Furthermore, we compare two methods for region-based analysis including both standard atlas labels and automatically generated ones from a study-specific template.

Finally, the design of some of these studies not only included scan-rescan analysis, but also

tested reliability in conjunction with applications to clinical and scientific studies. These studies have included populations consisting of aging adults and children [27] [21] [197], as well as patients with schizophrenia [99] [100], Alzheimer's disease [10], and multiple sclerosis [161]. This kind of evaluation provides an additional benchmark for comparing the practical value of such methods, which is important, as a perfectly reliable measurement might still disregard anatomical features that are of scientific or clinical value. In this chapter, we take a similar approach and test the sensitivity of each method to the anatomical effects of normal aging in an adult population.

4.2 Background

In this section, we'll briefly review the features of diffusion tensor imaging studies that relate to our evaluation. The primary goal of these studies is to quantify and localize anatomical variation that is related to aspects of the population under study. The standard pipeline includes steps for image acquisition, image preprocessing, spatial normalization, spatial mapping, and statistical analysis (Fig. 4.1). The image acquisition step involves the recruitment and scanning of many study participants, whose details lie outside the scope of this chapter. It suffices to say that each subject's diffusion-weighted MRIs (dMRI) is acquired along with a variety of demographic and clinical variables specific to the study at hand. Each dMRI consists of numerous volumes and associated diffusion encoding gradient directions and strengths.

The dMRIs and gradients are typically preprocessed to reduce the effects of noise, motion, and other artifacts, as well as to estimate models representing the pattern of diffusion in each voxel. We focus here on diffusion tensor modeling, which is the most common approach. This is a simple but powerful representation characterized by a Gaussian distribution, which can reflect any ellipsoidal shape of diffusion [13]. The predicted signal under this model is as follows:

$$S = S_0 e^{-b\mathbf{g}^T \mathbf{D} \mathbf{g}} \tag{4.1}$$

given a baseline signal S_0, gradient encoding direction $\mathbf{g}$, b-vector b, and a six-parameter positive definite diffusion tensor $\mathbf{D}$. From an eigen-decomposition of $\mathbf{D}$, a variety of rotationally invariant diffusion tensor metrics can be derived [170] [239], including fractional anisotropy (FA), mean diffusivity (MD), axial diffusivity (AD), and radial diffusivity (RD). These tensor metrics reflect size and shape features of the tensor, which are sensitive to microstructural properties of white matter

tissue, such as myelination, fiber density, etc. However, it is important to note that the tensor model is not specific to these microstructural features and it fails to characterize changes in more complex anatomy, such as fiber crossings [17]. While this limits the precision of inferences made with the diffusion tensor model, it is widely used and offers much in terms of exploratory analysis.

The next step is to spatially normalize the imaging data using image registration to either a standard template or a study-specific template. This step is necessary for providing a common image coordinate system without differences in pose and shape between subjects [98]. It also has unique features in tensor imaging studies, as the criteria for alignment can be derived from the geometric form of the tensor model and the measurement coordinate system must also be transformed [6]. In our experiments, we use the state-of-the-art in deformable tensor-based registration and use the study-specific template approach [231], which have been shown to provide improved statistical power compared to other methods [251] [103].

Following this, spatial mapping is performed to derive anatomically comparable measurements of diffusion tensor metrics. These measurements are meant to provide a vectorized representation of the imaging data that can be statistically analyzed to relate structural features to demographic and pathological characteristics of a population [84]. This can be done with a simple global histogram analysis, manually drawn regions, or automated methods. In this chapter, we focus on the latter and evaluate a variety of automated methods.

A commonly used automated approach is known as voxel-based analysis (VBA), which is inspired by earlier work on voxel-based morphometry [7], and examines each voxel across the population after spatial normalization. The major limitation is the effect of misregistration [8], and the standard solution is to smooth the image to reduce the high-frequency misalignment between subjects. Another standard approach is skeleton-based analysis, a technique also known as Tract-based Spatial Statistics (TBSS) that adds several steps to the VBA pipeline [194]. First, a white matter skeleton is created from the atlas volume, which is meant to restrict the measurements to a white matter core to reduce partial volume effects and to provide better matching between subjects. Each subject's tensor image is then registered to template-space, and the FA data is skeletonized and projected onto the atlas skeleton. This projection step is meant to fix residual registration errors and to avoid the need for the VBA smoothing step.

Automated region-based analysis is another alternative that instead derives measurements from an average of voxels within regions-of-interest. This can be done with regions defined in a standard

atlas, such as the Johns Hopkins University (JHU) white matter atlas [144]. Alternatively, regions can be derived in a data-driven way from a study-specific template. In principle, these data-driven methods aim to be similar to manual delineation, in that voxels are grouped based on spatial proximity and geometric similarity of the diffusion pattern, for example, with diffusion tensor principal orientations [37] or orientation distribution functions [26]. We refer to these data-driven region-based approaches as supervoxel-based analysis, due to their algorithmic similarity to methods in the computer vision and medical imaging literature [222] [226] [124]. It is also common for studies to combine skeleton-based analysis with region-of-interest analysis by averaging only skeletonized voxels within each region [88]. In the next section, we provide a detailed description of the specific pipeline implemented in our experiments.

4.3 Materials and Methods

In this section, we describe the data used in the experiments, the implementation of the methods under evaluation, and the design of the experiments.

4.3.1 Data Acquisition

Under an IRB-approved protocol, diffusion-weighted MRIs were acquired from a population of healthy volunteers, including a group of 80 normal aging healthy controls and eight from a scan-rescan cohort. The 80 subjects comprised a cross-sectional normal aging population, which consisted of nearly equal number of each sex and roughly uniform distributed ages ranging from 25 to 65 years old. The other eight subjects were acquired for scan-rescan analysis and included three repeats each, except for one subject that only had two repeats (i.e. 23 sessions). Imaging was conducted on a GE 1.5T scanner with 2x2x2mm voxels and image resolution 128x128x72. For each diffusion scan, seven baseline volumes were acquired, and the diffusion-weighted images used a single-shell high angular resolution diffusion encoding scheme with 64 distinct gradient encoding directions at a b-value of 1000 s/mm^2.

4.3.2 Image Preprocessing

The diffusion-weighted MRIs were preprocessed using FSL 5.0 [195]. The first step included motion and eddy current correction by affine registration of each diffusion-weighted image to the baseline

image using FSL FLIRT with the mutual information criteria. Along with this step, the b-vectors were reoriented to account for rotation induced by each transformation [119]. Skull stripping was performed using FSL BET with a fractional intensity threshold of 0.3. For each dataset, diffusion tensors were fit using FSL DTIFIT.

4.3.3 Spatial Normalization and Template Construction

Following this, a study-specific template was created from the 80 normal subjects. This was performed using the tensor-based deformable registration algorithm in DTI-TK [255] with finite strain tensor reorientation and the deviatoric tensor similarity metric. Each subject's tensor image was transformed to atlas space using the associated deformation and resampled to 1 mm^3 isotropic voxels. This process was applied to both the scan-rescan cohort and the normal aging cohort.

4.3.4 Spatial Mapping

In the next step, eight methods of spatial mapping were applied (listed in Table 4.1 and illustrated in Figure 4.2). Each of these methods essentially derives a vectorized form of each subject's image data for statistical analysis. This was performed with the following tensor metrics: fractional anisotropy (FA), mean diffusivity (MD), radial diffusivity (RD), axial diffusivity (AD). These are described as follows.

Voxel-based analysis (VBA) was performed by selecting individual voxels across white matter. For this, a white matter mask was created by applying an FA threshold of 0.2 and manually removing artifactual labels. Smoothed voxel-based analysis (SMOOTH+VBA) was similarly performed after first applying Gaussian smoothing to each volume with $\sigma = 2$, FWHM = 4.7, which is comparable to a previous VBA evaluation [197]. TBSS [195] [196] was performed using FSL with an FA threshold of 0.2, and skeleton voxels were selected from each subject. The skeleton and distance field derived from the FA skeleton were used to project MD, RD, and AD as well. TBSS was also performed with a similar smoothing step prior to subject skeletonization (SMOOTH+TBSS). Standard atlas region-based analysis was performed using the JHU white matter atlas [144] [145]. These labels were registered to the study-specific template using FSL FNIRT, and each subject was summarized by averaging voxels contained in each of the regions. These regions were used in conjunction with both voxel-based analysis (JHU+VBA) by averaging all voxels within each region and skeleton-based

analysis (JHU+TBSS) by averaging only skeleton voxels within each region. In addition to the standard atlas region, a more data-driven supervoxel-based analysis was performed to automatically label regions-of-interest across the entire template white matter mask. This was accomplished by clustering-based segmentation based on principal tensor orientation [37]. Clustering parameters included $\alpha = 1$, $\beta = 15$, and $\lambda = 20$, resulting in a total of 321 regions. This supervoxel-based analysis was performed in conjunction with both voxel-based analysis (SUPER+VBA) by averaging all voxels within each region and skeleton-based analysis (SUPER+TBSS) by averaging only skeleton voxels within each region. In summary, the combination of these eight methods and the four diffusion tensor metrics (FA, MD, RD, and AD) gives a total of 32 conditions that are included in the evaluation, which is described in detail next.

4.3.5 Reproducibility and Reliability

Next, reproducibility and reliability were tested for each condition with the scan-rescan dataset, which consisted of eight subjects with three repeated scans each. This included two statistical evaluation metrics: the coefficient-of-variation (CV) [24] and intra-class correlation (ICC) [23]. The CV is a normalized measure of percentage change in each measurement across scans and is considered acceptable below 10%. Given the within-subject average μ_w and within-subject standard deviation σ_w, the CV is given by σ_w/μ_w. The ICC is a measure of reliability that gauges the fraction of variance between subjects. It is normalized between zero and one and is considered acceptable above 0.7. Given the between subjects variance σ_b^2 and within-subjects variance σ_w^2, the ICC is given by $\sigma_b^2/(\sigma_b^2+\sigma_w^2)$. For each condition, CV and ICC were computed for individual voxels/regions and then aggregated across the whole brain to estimate mean performance and its uncertainty. All statistical analysis and implemented using R 3.1.1 [176], with the ggplot2 package for plotting [241], and the ICC package from Wolack et al. [245].

4.3.6 Normal Age Modeling

Next, the methods were evaluated by comparing performance in population-based modeling of normal aging, a process which has been shown to include anatomical changes in white matter that are reflected in diffusion tensor metrics [71] [39]. Our evaluation strategy was to construct predictive linear models of subject age for each condition using principal component regression (PCR) and

then compare results with standard model selection criteria [151]. For the purpose of comparison, a baseline multiple linear regression model without diffusion measurements was first created to model age based on only the intracranial volume and sex of the subject. Then for each condition, principal component analysis was performed on the origin-centered dataset, and the top fifteen components were retained for modeling. Then for each condition, best-subset model selection was performed using the Bayesian information criteria (BIC) to compare models. In addition, each best model was further evaluated by measuring the adjusted R^2 to measure goodness of fit. Then, the best models were compared across methods based on these three model selection criteria. All statistical analysis was implemented in R 3.1.1 [176], with the ggplot2 package for plotting [241], and the leaps package for model selection [1].

Table 4.1: A summary of methods for spatial mapping that are compared in the experiments. The abbreviation and description for each method is listed. In addition, the dimensionality of the resulting measurements is listed, as well as the average volume of the voxels/regions representing each measurement.

Abbreviation	Description	Dimension	Mean Volume
VBA	Voxel-based analysis of single voxels	353903	1 mm^3
SMOOTH+VBA	Voxel-based analysis of smoothed voxels	353903	1 mm^3
JHU+VBA	Voxel-based analysis of JHU atlas regions	48	2814 mm^3
SUPER+VBA	Voxel-based analysis of supervoxel regions	321	1098 mm^3
TBSS	TBSS analysis of skeleton voxels	76586	1 mm^3
SMOOTH+TBSS	TBSS analysis of smoothed skeleton voxels	76586	1 mm^3
JHU+TBSS	TBSS analysis of JHU atlas regions	48	648 mm^3
SUPER+TBSS	TBSS analysis of supervoxel regions	318	240 mm^3

4.4 Results

The results in CV show reproducibility varying greatly across methods (Fig. 4.3), with single voxel analysis (CV = 5.91 ± 0.81%) performing with roughly twice the CV of region-based analysis (CV = 2.45 ± 0.25%). Reproducibility was also variable across diffusion metrics with the highest errors in FA (CV = 5.30 ± 1.56%) and RD (CV = 4.77 ± 0.83%), compared to lower errors in MD (CV = 3.48 ± 0.83%) and AD (CV = 3.22 ± 0.87%). Smoothing improved reproducibility in single voxel methods, as unsmoothed (CV = 7.62 ± 0.95%) had almost twice the error rate of smoothed (CV = 4.2 ± 0.51%). A visual inspection of CV scores on an axial slice (Fig. 4.4) also shows the spatial variation in CV. Voxel-based analyses showed the greatest spatial variability, with better

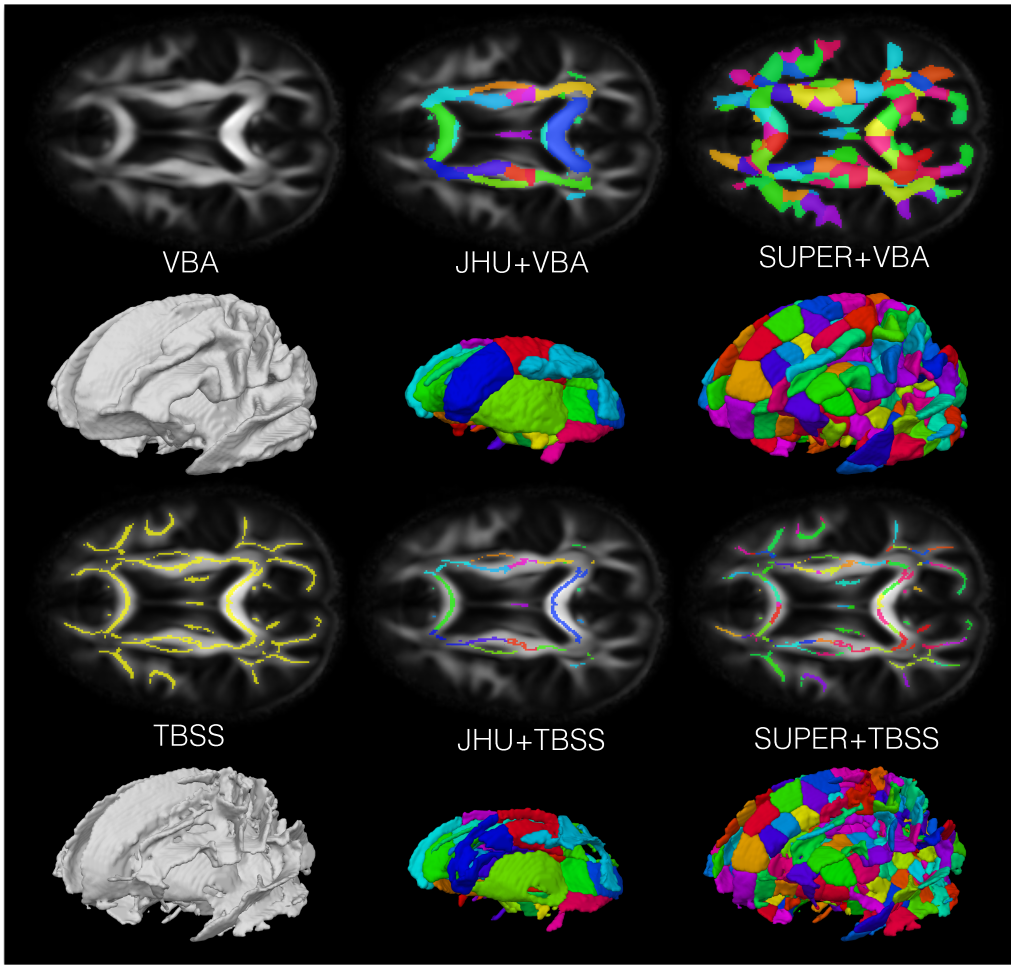


Figure 4.2: An illustration of methods for spatial mapping compared in the experiments. Each panel shows an axial slice showing either included voxels or region labels and an associated 3D surface rendering. Smoothing was included in voxel-based and skeleton-based analysis but is not depicted here.

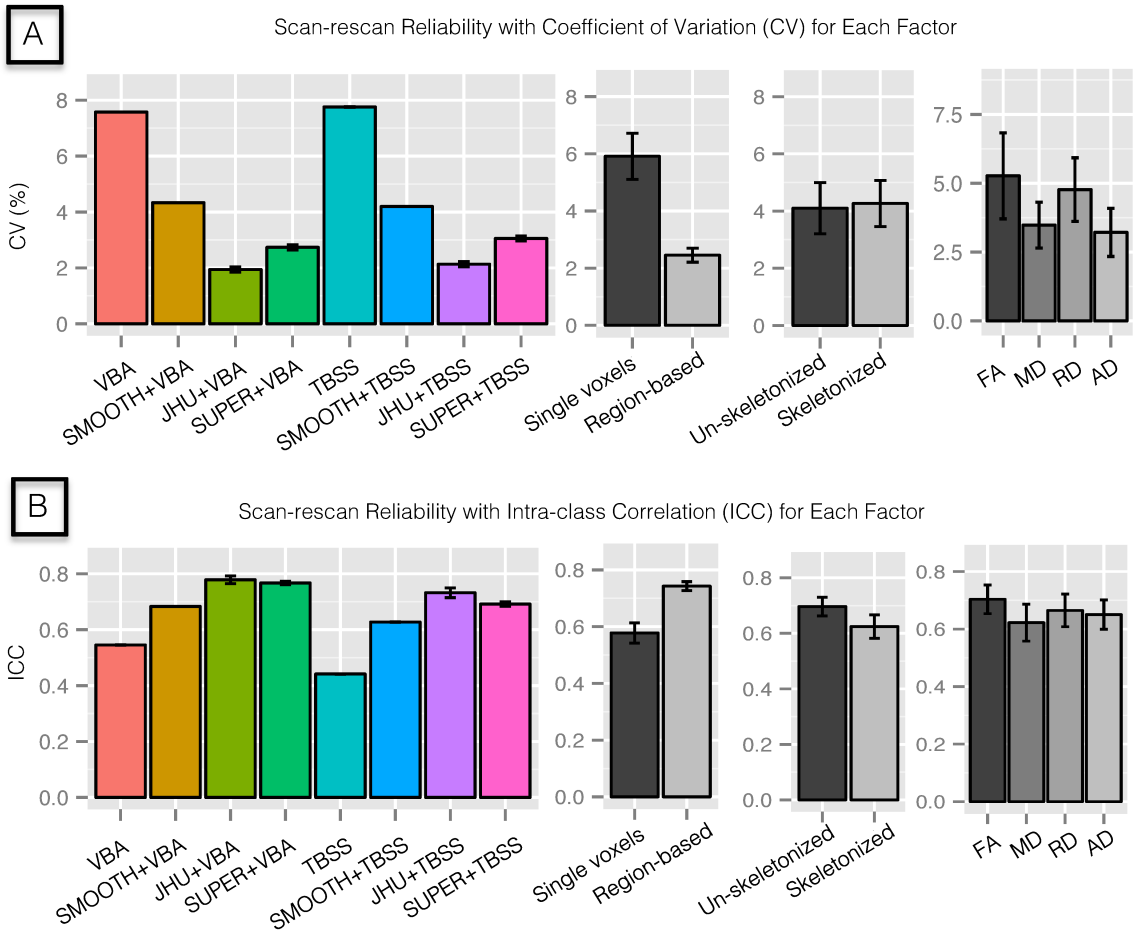


Figure 4.3: Results from the scan-rescan experiment in Sec. 4.3.5 showing reliability across methods and diffusion tensor metrics. Panel A shows the coefficient of variation (CV), which indicates the percentage of variation across scans of the same subject (smaller is better). Panel B shows the intraclass correlation, which indicates what proportion of variance is between subjects (larger is better). The results show high variation across methods and some variation across diffusion metrics, especially in voxel-based methods. Voxel-based methods were mostly less reliable than region-based methods, while smoothing in voxel-based methods improved reliability.

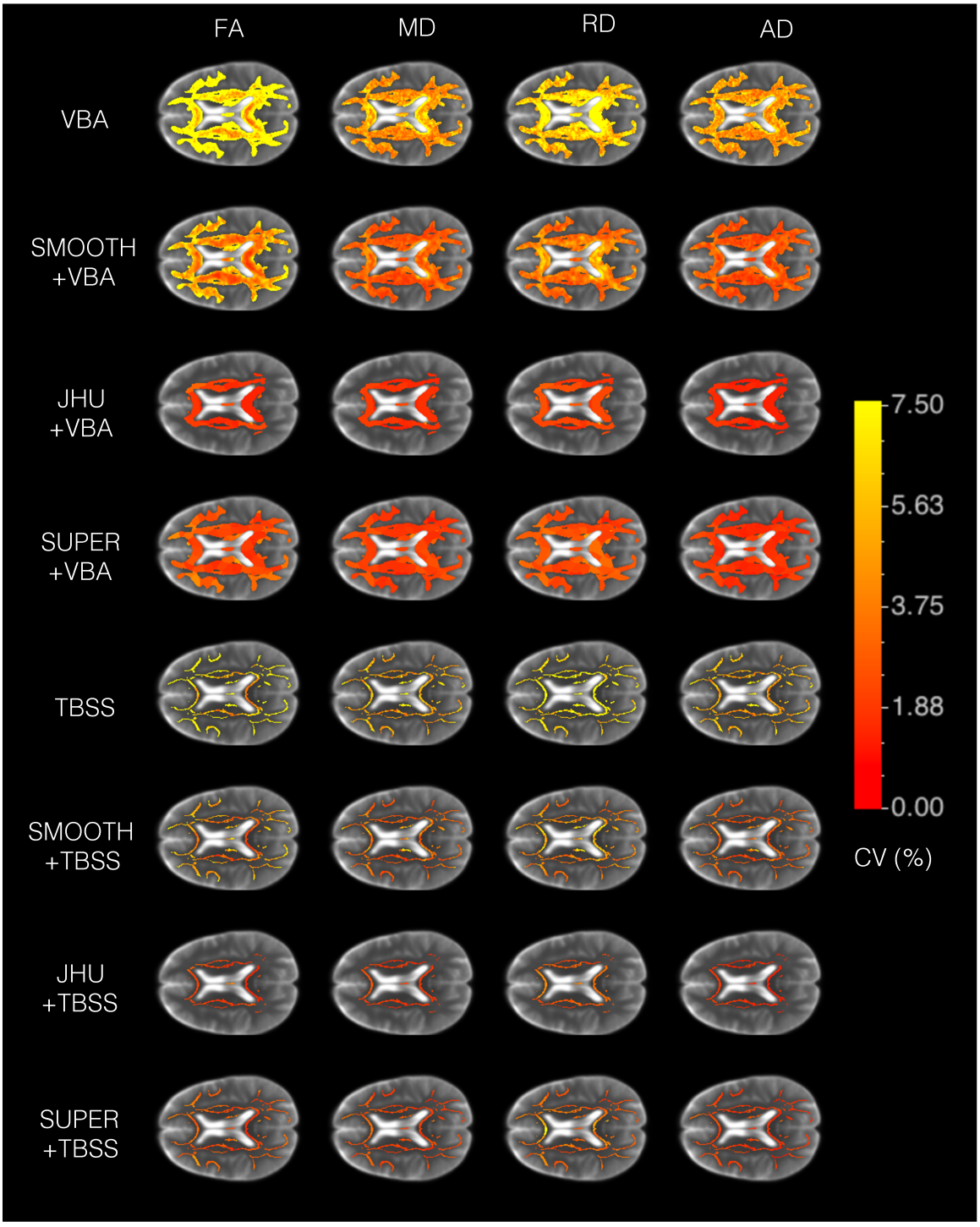


Figure 4.4: Results from the scan-rescan experiment in Sec. 4.3.5 showing the spatial distribution of the coefficient of variation (CV) in an axial slice. The panels are organized to show a different method in each row and a different diffusion tensor metric in each column. The results generally show large spatial variation across methods and metrics, with higher variation in voxel-based than region-based methods. Voxel-based analysis tended to have higher reliability in deep white matter and lower in superficial white matter. Supervoxel-based analysis showed better performance in superficial white matter.

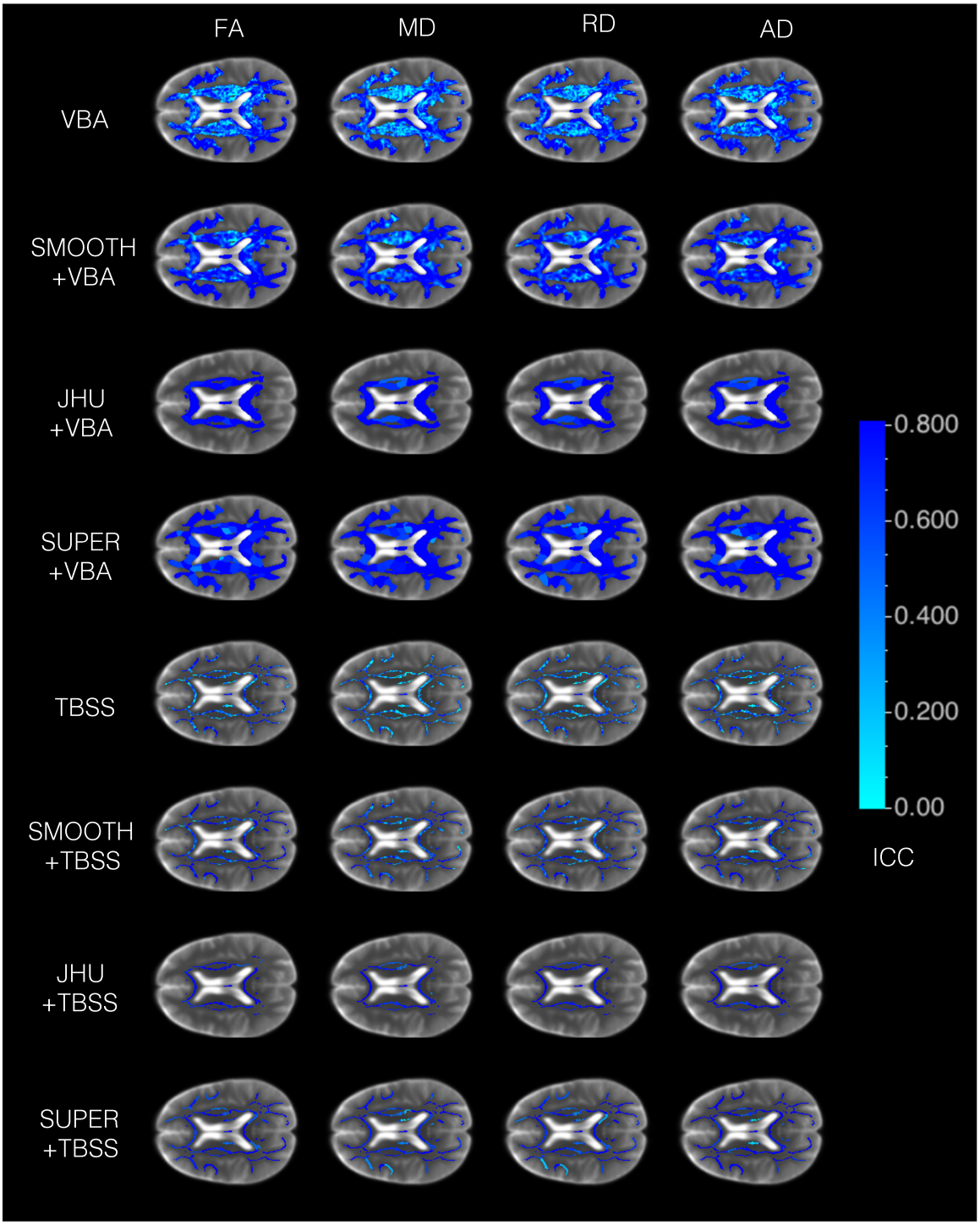


Figure 4.5: Results from the scan-rescan experiment in Sec. 4.3.5 showing the spatial distribution of the intraclass correlation (ICC) in an axial slice. The panels are organized to show a different method in each row and a different diffusion tensor metric in each column. The results generally show large spatial variation across methods and metrics, with higher variation in voxel-based than region-based methods. The anatomical distribution differs greatly from the coefficient of variation (CV), demonstrating the different aspects of reliability revealed by CV and ICC.

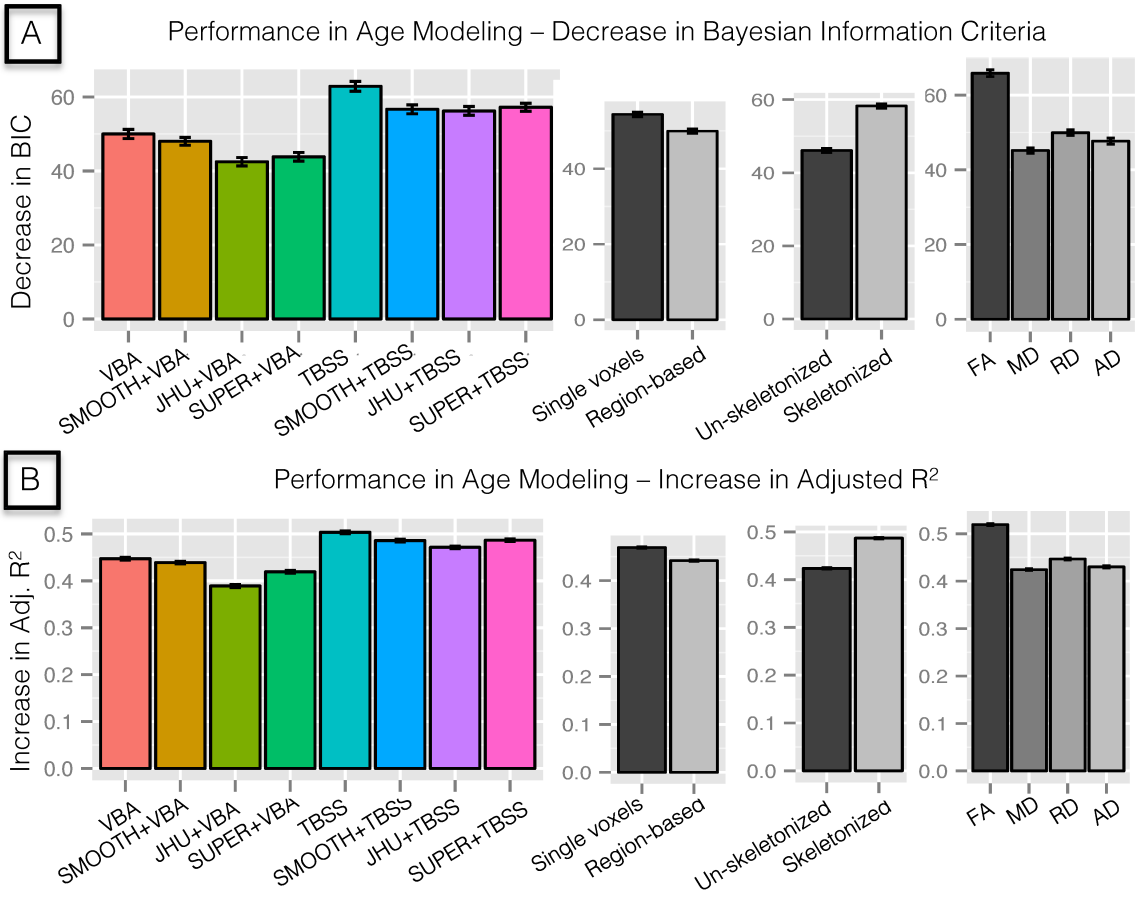


Figure 4.6: Results from fourth experiment in Sec. 4.3.6 showing the performance in age modeling with principal component regression. For each method and metric, the best performing model was selected using Bayesian information criteria (BIC) among all subsets of the top 15 principal components. Panel A shows the decrease in BIC of each model relative to a baseline without tensor metrics; Panel B shows the adjusted R^2. Higher values indicate better performance in all plots. The results show that all methods improve significantly over the baseline model, and skeletonization tended to improve performance in diffusivity tensor metrics. Performance was comparable among methods in FA-based models.

CV scores in deep white matter and worse CV in superficial white matter. However, RD showed a opposite pattern to other tensor metrics with the worse CV scores concentrated around the genu and splenium. Smoothing in voxel-based analysis tended to also smooth this spatial distribution of CV scores. Region-based analysis showed more spatially uniform CV results than voxel-based analysis, particularly in superficial white matter with supervoxel-based analysis.

The results in ICC show reproducibility with a slightly different pattern from CV (Fig. 4.3). In general, single voxel methods (ICC = 0.58 ± 0.04) were less reliable than region-based methods (ICC = 0.74 ± 0.02); however, smoothing improved the ICC of VBA and TBSS to be comparable to region-based methods in several cases. Reproducibility was more uniform across diffusion metrics using the ICC metric with slightly better performance in FA (ICC = 0.70 ± 0.05) and RD (ICC = 0.66 ± 0.06), compared to MD (ICC = 0.62 ± 0.06) and AD (0.65 ± 0.05). Methods without skeletonization (ICC = 0.70 ± 0.03) had a slightly higher ICC scores than those using skeletonization (ICC = 0.62 ± 0.04). Smoothing improved reproducibility in single voxel methods, as unsmoothed (ICC = 0.50 ± 0.04) had almost twice the error rate of smoothed (ICC = 0.66 ± 0.02). A visual inspection of ICC scores on an axial slice also shows spatial variation (Fig. 4.4); however, the distribution exhibits a more heterogeneous pattern in all methods than the CV results. In particular, voxel-based methods showed a more diffuse pattern of reliability, with high ICC values in superficial white matter. Smoothing in voxel-based methods also tended to smooth this pattern of ICC. Region-based methods did not show any clear distinction between deep versus superficial white matter.

The results in age modeling show generally good performance from all methods and diffusion metrics ($R^2 = 0.53 \pm 0.02$); however, some methods performed better than others (Fig. 4.6). Skeleton-based analysis ($R^2 = 0.56 \pm 0.0$) showed generally better results than voxel-based analysis ($R^2 = 0.50 \pm 0.0$), although they were comparable in models based on FA. There was a small difference between voxel-based and skeleton-based analysis without smoothing ($R^2 = 0.54 \pm 0.0$) performed better than without smoothing ($R^2 = 0.55 \pm 0.0$). Region-based analysis ($R^2 = 0.52 \pm 0.0$) only a small difference between it and single-voxel analysis ($R^2 = 0.55 \pm 0.0$). In comparing the two model selection criteria, they tended to show comparable results within each diffusion metric and preserved the ranking of methods despite changes in scale. In general, all methods showed decreases in BIC ranging from 45 to 90 and adjusted R^2 ranging from 0.42 to 0.62.

4.5 Discussion

In this study, we evaluated several methods for spatial mapping in diffusion tensor studies. The goal was to better understand the overall reliability of these methods and simultaneously understand how useful they are in an aging application. To accomplish this, we evaluated eight methods for spatial mapping with four diffusion tensor metrics in two experiments: a scan-rescan analysis of reliability and reproducibility and an application to age modeling in a normal adult population. These experiments generally show that the choice of spatial mapping greatly affects the overall magnitude and spatial distribution of reliability. While all methods were shown to be useful in age modeling, the results provide some relative performance characteristics for maximizing the sensitivity of imaging metrics. Next, we will discuss the significance of these findings and related open problems.

The first main finding was variability in the overall reliability of each method despite using identical data, preprocessing steps, and registration. The most readily observed pattern was that methods looking at single voxels, e.g. VBA and TBSS, were less reliable than region-based methods, e.g. JHU and SUPER, as measured with both CV and ICC. Previous work has demonstrated a trade-off in spatial specificity between these methods [197], and the results of this study further support a trade-off in reliability between voxel-based and region-based analysis. This difference is perhaps due to the voxelwise averaging used in region-based analysis, which could also tend to average out the effects of noise. Smoothed voxel-based analysis is perhaps a similar way to accomplish this, but it includes a greater risk of mixing different tissues and adjacent but distinct fiber bundles. Past work has also found that the results of voxel-based analysis depend greatly on the filter size and implementing package [100] [99], and the results of this study show related changes in reliability. Specifically, reliability in voxel-based and skeleton-based analysis tended to improve with smoothing and performance depended on the particular diffusion tensor metric being tested, which supports previous findings [197]. Regarding region-based analysis, the results were also comparable to past works using manually drawn ROIs [27] [44] [211], which is perhaps evidence that deformable tensor-based registration is comparable in quality to anatomical matching of manually drawing region masks.

The second main result was that all methods exhibited spatial variability in CV and ICC estimates of reliability. This reinforces similar results demonstrated in prior work that examined the

spatial distribution of reliability estimates [89] [76] [132], although these studies were typically limited to tests of only one or two methods for spatial mapping each. The results of this study show voxel-based methods tended to have the most spatial variability and had concentrated high reliability in deep white matter, similar to previous work [224]. This could be related to higher registration accuracy in deep white matter, as seen in fiber coherence maps derived form population data [255]. However, it could also be that reliability is highest where the tensor model is most representative of the underlying diffusion process, i.e. predominantly single fiber regions in deep white matter [19]. This could be more thoroughly studied by examining reliability of multi-fiber extensions of TBSS [90], possibly with multi-compartment model smoothing [35]. Region-based analysis was found to have lower spatial variability, and supervoxel-based analysis had better performance in superficial white matter. This is perhaps because misregistration tends to happen more often within regions than between regions, due to the homogeneity of voxels within each region. In general, ICC had more spatial variability than CV, and the spatial distributions also were quite different. This demonstrates how CV and ICC reflect different aspects of reliability, as CV directly represents error, while ICC additionally depends on the variation across subjects. For this reason, the results in ICC may be more tied to specific populations and datasets used for evaluation.

The third main result was related to the application of these methods to the modeling of normal aging, and the results indicated that all methods performed well in capturing age-related changes in anatomy compared to previous studies examining predictive models of age [150], though some performed better than others. In comparing the BIC in relation to the baseline model without tensor metrics, all methods included an increase of at least 10, which is strong evidence of an improved model [102]. Looking more finely, there are several patterns to note. First, JHU+VBA performed slightly worse than other methods. This could be attributed to the limited amount of brain covered by the regions; however, the significant improvement with JHU+TBSS suggests that non-skeleton voxels represent confounds in the analysis, perhaps due to partial volume effects. Supervoxel-based analysis was consistently better, and it also showed improvement with skeletonization. For models based on MD, AD, and RD, there was a general trend that skeletonization significantly improved performance; however, models based on FA had comparable performance with and without skeletonization. This is perhaps related to negative consequences of partial volume effects in these measures, as diffusivity measurements could be more susceptible to contamination with cerebrospinal fluid. Another aspect to note is the use of principal component regression, which provides several advantages. First, it

provides a level playing field to compare the methods, where each method is allowed the same number of variables for modeling, while maintaining a maximal amount of variance. Second, it avoids a major issue of multi-collinearity, which especially common in voxel-based and skeleton-based analysis. Without the principal component dimensionality reduction step, the predictive models would likely suffer from overfitting in VBA and TBSS analysis.

It is also worth discussing the design of the study. In particular, the experiments were designed to control for a number of potential biases that could severely effect the results, such as dataset, preprocessing steps, and registration algorithm. This allows us to more certainly attribute the observed differences in reliability and predictive modeling to the choice of spatial mapping algorithm and not to other factors. This is a somewhat stronger result than could be gained by summarizing the results of multiple studies, which inevitably have major differences in data and implementation. However, the major limitation of this design is that only one factor of the pipeline was studied, and the results possibly depend on variation in these other factors, e.g. registration algorithm. A full factorial design is quite challenging due to the increasing number of conditions to test; however, it is likely a fruitful avenue of research to pursue. Looking beyond voxel-based analysis, it would also be tremendously valuable to expand this kind of evaluation to include tractography-based spatial mapping. However, a similar challenge is posed by the vast number of methods currently in use, as each tractography reconstruction is a complex product of diffusion modeling, image interpolation, seed and selection masks, and termination criteria.

The results of this study are also somewhat limited with respect to the VBA smoothing step. Only a single bandwidth and smoothing technique were tested, but a variety of approaches can be found in the literature [191] [66] [228]. While the effect of smoothing bandwidth has been well studied [99] [100], a relatively less understood aspect is the effect of filter type and the filtering domain. For example, smoothing can be done with a variety of types of filters, including Gaussian, median, and anisotropic filtering, and unlike some other modalities, there are several possible filtering domains, such as the diffusion-weighted signals, the diffusion models fitted to the signal, or scalar features derived from the models. Smoothing in the signal domain is attractive for the theoretical guarantees of linear systems and sampling theory, but it not commonly used in VBA, perhaps due to challenges inherent to reorienting q-space data after registration. Model-based smoothing of tensors can possibly preserve anisotropy and fiber orientation [62]; however, the most common approach is to smooth in the feature domain, which has been suggested to most appropriately accommodate errors

due to misregistration [164]. This study used feature domain Gaussian smoothing with a bandwidth that is comparable to previous studies with comparable voxel size [110] [166] and recommended in a previous evaluation [197]; however, there remain many questions to answer related to these aspects of smoothing in VBA.

In conclusion, this chapter presents a comparative evaluation of methods for voxel-based spatial mapping in diffusion tensor imaging studies, which are an important part of population studies. The results show reliability depends greatly on the type of spatial mapping, the anatomical location, and somewhat on the diffusion metric. Despite these differences, both single voxel-based and region-based methods performed well in age modeling, with consistently good performance with skeleton-based analysis compared to voxel-based analysis. These reliability results may help in the design and interpretation of future studies, as they indicate care must be taken to establish baseline reliability and statistical power of a study based on the specific anatomical hypotheses, method of spatial mapping, and diffusion metric employed in analysis. The results of the aging application also may help to understand how the choice of spatial mapping method can provide improved modeling power when conducting scientific and clinical white matter imaging studies.

Chapter 5

Diffusion MR Fiber Bundle Modeling with the Sparse Closest Point Transform

Diffusion MR imaging provides a unique *in-vivo* probe of tissue microstructure through the imaging of water molecule diffusion patterns [171]. This technique is particularly valuable for characterizing the local features of white matter and for reconstructing the large scale structure of fiber bundles through tractography [13]. Both approaches have clinical and scientific value in assessing anatomical variation in health and disease; however, the size and complexity of tractography datasets can pose a challenge to such studies. In particular, both expert anatomical knowledge and time consuming manual interaction are typically required to delineate fiber bundles from whole brain tractography. These challenges are greater still in population studies, where this must be repeated for many subjects with minimal inter- and intra-rater variability [155, 121]. A promising alternative is the automated analysis of tractography datasets, which offers to both to reduce the time invested by the human analyst and provide highly reliable measurements in large population studies. In particular, there are great opportunities ahead for better integrating diffusion MRI data with the rich variety of statistical and machine learning methods available [160].

In this chapter, we develop and evaluate several such automated approaches for analyzing whole-brain tractography datasets using sparse closest point transform. The primary motivating idea is

that fiber curve data can be transformed to a vector space representation that retains anatomically relevant features and enables the use of standard tools from statistical and machine learning for fiber bundle analysis. This extends prior work in this area by developing a sparse closest point transform for curve data, in which each curve is represented by its closest points to a collection of data-adaptive landmarks. The primary goal of this chapter is to investigate how generally useful this representation is when combined with standard tools from machine learning. This is explored through experiments testing such combinations to perform fiber bundle clustering, bundle simplification, and population-based bundle selection. The proposed fiber bundle clustering algorithm is sufficiently time efficient for interactive use and includes a mechanism to select the number of fiber bundles from the data. The proposed simplification algorithm can greatly reduce the memory usage and improve performance of 3D rendering, while preserving the geometric structure of the full resolution dataset. The proposed population-based analysis builds a statistical model of each bundle, which can be used to select bundles from whole brain tractography without parameter tuning or manual intervention. Our contributions also include an evaluation of performance with *in vivo* human brain data and comparisons with related methods.

The rest of the chapter is organized as follows. In Sec. 5.1, we discuss related work in this area and distinguishing features of the proposed method. In Sec. 5.2, we describe the proposed method in detail, including landmark extraction, the sparse closest point transform, fiber bundle clustering, bundle simplification, and statistical bundle modeling for population-based analysis. In Sec. 5.3, we describe the design and results of experiments to evaluate the proposed methods with *in vivo* human brain data, including tests of agreement with known bundles, comparison of inter-fiber distance measures, and evaluation of reproducibility and sensitivity to aging in a population-based analysis. In Secs. 5.4 and 5.5, we discuss our findings and conclude.

5.1 Related Work

The primary aim of this chapter is to investigate the use of the sparse closest point transform in fiber bundle analysis, specifically in combination with existing statistical and machine learning techniques. So, the goal is not to replace any specific method, but rather, to investigate the general utility of this representation. Next, we briefly discuss prior work in fiber bundle modeling and its relation to this study.

The prior work in this area can be roughly described as distance-based or feature-based. Distance-based approaches group fiber curves based on inter-fiber proximity. A number of inter-fiber distances have been proposed, including Hausdorff measures [155], distance-based integrals [258], endpoint distances, and others. Pairwise curve distances can then be used in any of number a clustering algorithms, such as spectral clustering [31] [156], hierarchical clustering [258], or Dominant Sets [61]. These approaches can tease apart subtle distinctions between fibers [141]; however, the computational cost of total pairwise comparison of curves is prohibitively expensive for full datasets. This is usually mitigated by sub-sampling the full dataset to a tractable size, e.g. less than 20,000. The representation presented here is compatible with methods that are scalable well beyond this, permitting whole brain tractography datasets to be analyzed without subsampling.

In contrast to distance-based approaches, feature-based approaches map each curve to a simpler but discriminative representation and can avoid subsampling. Prior work has commonly used connectivity between regions of interest (gray-matter areas or manually drawn volumetric masks) to select curves [260] [250]. This is somewhat similar to the proposed approach, but here we consider landmarks learned from the data which require no gray matter segmentation or manual drawing. Furthermore, the extracted landmarks are typically distributed across the whole brain, which can help to constrain the path taken between endpoints. Other related work has used implicit volumetric representations of curves, for example with distance fields [126] and spatial-angular histograms [230]. These representations are somewhat limited by their high-dimensionality, which is avoided in the proposed method by restricting the representation to a sparse but representative set of landmarks. These implicit representations are also potentially limited by their statistics, for example, distance fields have been modeled with Gamma distributions, and spatial-angular histograms have been modeled by multinomial distributions. In contrast, the proposed method can potentially be used with methods typical of Gaussian distributed variables, as the closest points represent explicit positions along curves. Other approaches have aimed to derive arclength parameterizations of fiber curves and then model bundles with respect to these parameterizations [46] [49] [69], which usually requires flipping of the start-to-end orientation of each fiber during analysis. A general trend in these prior feature-based methods is that they are usually tied to segmentation algorithms that are carefully engineered based on the choice of representation. A goal of this work is to try to avoid this tight coupling to encourage more "interchangeable parts" way to use statistical and machine learning algorithms for fiber bundle modeling.

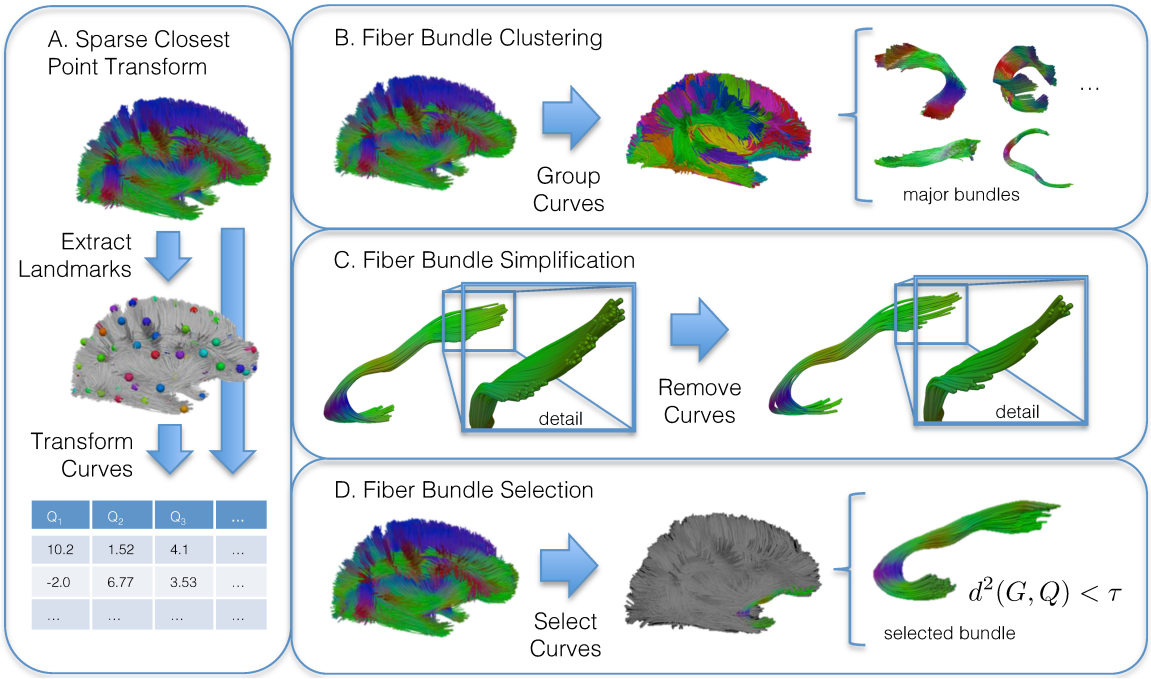


Figure 5.1: Illustration of the major components presented in the chapter. The left panel (A) shows the sparse closest point transform, which includes initial step of landmark extraction followed by a closest point transform of each curve relative to the landmarks. The top panel (B) shows an application of this to fiber bundle clustering, where whole brain tractography is decomposed into bundles using a variant of the k-means algorithms. The bottom panel (C) shows population-based analysis, where a bundle in an atlas is statistically modeled and then segmented in each subject in a population with a one-class classifier.

5.2 Methods

The goal of the proposed method is to derive a useful representation of fiber curve data for extracting meaningful bundles from whole-brain tractography datasets in both single subjects and populations. This includes three main parts that will be described in this section. First, we will cover how the proposed mapping of fiber curves to a vector space representation with the sparse closest point transform. Next, we will describe how bundle clustering and simplification can then be performed in single subjects with the DP-means algorithm. Finally, we will describe a population-based analysis for segmenting specific fiber bundles across multiple subjects with statistical bundle modeling and deformable volumetric registration.

In the following analysis, we make a simplifying assumption that tractography fiber curves have already been extracted, for example by streamline integration [259]. We'll then assume the fiber

curves are represented by polylines, i.e. sequences of points and the line segments connecting subsequent points. We'll take the dataset to have N curves, where the i-th fiber has N_i points and each fiber curve is denoted by $C_i \subset R^3$ for $i \in [1, N]$. There's no assumption made about the spatial resolution or uniformity of the points sampled along each curve.

5.2.1 Fiber Curve Representation

This section describes the proposed mapping of fiber curve data using the sparse closest point transform. This is a two step procedure, which first derives a sparse and representative set of landmarks from the curve data, and then transforms each curve relative to these landmarks using the closest point transform. These steps are performed as follows.

Landmark Extraction

In the first step, a set of representative landmark positions are extracted from the dataset. These are used as points of reference when encoding fibers, so more landmarks will give a more complete representation of each fiber. However, reducing the number of landmarks can improve performance, hence we aim to select a small but representative set of landmarks. To do this efficiently, we first sub-sample the fibers through random selection to obtain some small number of fibers, e.g. less than 5,000 (however, the full dataset is still retained for later analysis). These sub-sampled curves are then geometrically simplified using the Ramer-Douglas-Peucker (RDP) algorithm [183] to reduce the fibers to include only important vertices, such as endpoints and points of high curvature. The RDP procedure works by greedily removing points from the polyline while keeping the maximum error below a given threshold, e.g. 2mm. The vertices of the remaining simplified fibers are then all taken together and further reduced by clustering. This clustering can include a specific number of points M by using the k-means algorithm, or alternatively determine the number of points M based on a spatial threshold by using the DP-means algorithm, which is described in detail later in the chapter. In practice, we use the DP-means approach with a threshold of $\lambda = 5.0$. This results in a collection of landmarks that are roughly placed at salient features, such as high curvature and fiber terminations. We'll denote the j-th landmark by $w_j \in R^3$ for $j \in [1, M]$.

Sparse Closest Point Transform

In the next step, the fiber curve dataset is transformed relative to the extracted landmarks using a sparse closest point transform. We refer to this approach as sparse to distinguish it from related approaches that densely sample closest points on a pixel or volumetric grid [134]. The goal of using these sparse landmarks is not only to reduce the dimensionality of the resulting representation but also to allow landmarks to be placed at salient points of the dataset that may not otherwise be well aligned with a dense volumetric grid.

The closest point transform represents each fiber curve by the positions along its length that are nearest to each of the M landmarks selected by the steps in the previous section. Given a curve C, the transformed curve is given by $Q = (q_1, ..., q_M) \in R^{3M}$, with the j-th closest point q_j given by:

$$q_j = \underset{p \in C}{\operatorname{argmin}} \|p - w_j\|^2 \tag{5.1}$$

Note that with the polyline representation, this minimization can be performed not just over vertices but also along the connecting segments. This allows the algorithm to be applied to irregularly sampled curve data, such as those produced by the RDP algorithm. This can potentially reduce time and memory requirements compared to the uniform sampling typically required by other approaches.

The main advantage of this approach is that it gives a vector-space representation of curves that can be used with many statistical and machine learning algorithms. In the following sections, we describe how this can be employed for fiber bundle clustering, simplification, and population-based analysis.

5.2.2 Fiber Bundle Clustering and Simplification

Based on this fiber representation, fiber bundle clustering and simplification can be performed with the DP-means algorithm introduced by Kulis et al. [108], which is a variant of the k-means algorithm. As with k-means clustering, the goal here is to group elements of the dataset according to spatial proximity; however, the DP-means variant has the benefit of learning the number of clusters from the dataset. In comparison to other some clustering algorithms, the k-means and DP-means algorithms are both notable for their efficient performance, which allows them to be applied to large datasets. This is a significant issue for clustering whole brain tractography, as the typical dataset size is far larger than is practical with $O(n^2)$ clustering algorithms. Past approaches have used subsampling to

work around this problem, but the method proposed here is practically useful without subsampling.

This DP-means clustering algorithm can be defined by a minimization problem similar to the standard k-means algorithm with an additional regularization term. Prior work has shown this extension is theoretically related to Dirichlet Process mixture modeling with Gibbs sampling, through small-variance asymptotic analysis. This gives the algorithm a mechanism to choose the number of clusters in the dataset with the regularization term added to the typical k-means objective. When applied to a tractography dataset with the above fiber curve mapping, this process decomposes the set of curves into K bundles, where the k-th bundle is represented by a prototype $B_k \in R^{3M}$, and the optimal partitioning π is found by solving the following minimization problem:

$$\pi = \underset{\pi}{\operatorname{argmin}} \sum_{i}^{N} \|Q_i - B_{\pi(i)}\|^2 + \lambda^2 MK \tag{5.2}$$

where $\pi(i) \in \{1, ..., K\}$ encodes the label for the i-th curve. This objective can be optimized with an efficient algorithm that is described in detail in the referenced related work [108]. Since this method estimates the number of clusters, no user-defined value of K is explicitly specified. Instead, a regularization parameter λ is chosen, which can be interpreted as the largest average distance allowed between each curve Q_i and its associated bundle $B_{\pi(i)}$. When used for extracting bundles from whole brain tractography, a relatively large threshold can be used, e.g. 20mm. When used for simplification, a smaller threshold can be used, e.g. 2mm, and the closest curves each of the cluster centroids can be selected to represent the simplified bundle.

5.2.3 Fiber Bundle Selection

The next task aims to select a fiber bundle from an novel dataset based on an exemplar bundle in a atlas dataset. This is an important task for population-based analysis, where similar structures need to be matched across multiple subjects and manual selection is prohibitively costly. For this purpose, we make a simplifying assumption that some number of bundles-of-interest have been obtained an atlas dataset, for example, by first using the above method for clustering and then interactively selecting a bundle relevant to a particular study. In our tests, we created a tractography bundle atlas using tensor-based deformable registration and selected several well known bundles from among the clusters. The goal is then to match fibers from each subject to the bundles in the atlas. We accomplish this by first building a statistical model of each bundle in the transformed space and

then applying a *one-class* classifier to segment the bundle in each member of a population, with the details described as follows.

For a particular bundle, we propose to statistically model the transformed space representation of a collection of fiber curves $\hat{Q} = \{\hat{Q}_1, ..., \hat{Q}_C\}$ by a multivariate Gaussian G with mean μ and covariance Σ. We fit the mean and the covariance using maximum likelihood (ML):

$$\mu_{ML} = \frac{1}{C}\sum_{i}^{C} \hat{Q}_i \tag{5.3}$$

$$\Sigma_{ML} = \frac{1}{C}\sum_{i}^{C} (\hat{Q}_i - \mu)(\hat{Q}_i - \mu)^T \tag{5.4}$$

Maximum likelihood estimation of covariance matrices can suffer from a number of problems, such as ill-posedness and overfitting for small sample sizes, so we regularize the model with a shrinkage estimator:

$$\Sigma_{Shrink} = \omega \,\mathrm{Diag}(\sigma_0^2) + (1 - \omega)\Sigma_{ML} \tag{5.5}$$

given a shrinkage constant ω and prior spatial variance σ_0^2. The shrinkage estimator not only ensures Σ is well-defined but also incorporates prior knowledge about expected misregistration errors. In practice, we used $\omega = 0.3$ and $\sigma_0^2 = 80$ mm.

The fitted model G can then be used to segment a bundle in a particular subject by first deforming the subject's fiber curves to the atlas, taking the sparse closest point transform (relative to atlas space landmarks), and then finding the subset of curves that is sufficiently close to the modeled bundle. This last step is accomplished with a one-class classifier [148], and the method we describe here is a simple special case with a Gaussian distribution [208]. This approach is useful for classification tasks where only positive examples are available or out-of-class examples are particularly complex [209] [54]. This is the case for tractography data, and we have found that even the simple Gaussian case works well for segmenting bundles. To evaluate the classifier, we define a bundle-to-fiber distance $d(G, Q)$ with the squared Mahalanobis distance [12]:

$$d^2(G, Q) = (Q - \mu)^T \Sigma^{-1} (Q - \mu) \tag{5.6}$$

This has the benefit of measuring the statistical distance to known fibers in the bundle, as opposed

to a spatial distance that may be subject to some user-defined threshold. By calculating the distance relative to the bundle covariance Σ, the distance can reflect anisotropic spread along the bundle's length, for example due to fanning. This also gives us a test statistic that can be used to decide if a given fiber is included in the bundle, i.e. if $d^2(G, Q) < \tau$ for some threshold critical value τ. Given that the Mahalanobis distance is Chi-squared distributed as $d^2 \sim \chi^2_{3M-1}$, we can look up such a value τ where $p(d^2(G, Q) < \tau) = 0.99$, which will include 99% of the fibers sampled from the distribution G. As long as the same number of landmarks are used, this threshold requires no tuning, and the same value can be used for all subjects, as well as other bundles.

5.3 Experiments and Results

We evaluated the proposed approach by conducting five experiments with *in vivo* human brain datasets. The first experiment evaluates the fiber curve representation by comparing its associated inter-fiber distance with several other commonly used measures. The second experiment measured performance of the clustering algorithm in reference to manually selected fiber bundles and in comparison to other clustering algorithms. The third experiment measured performance of the bundle simplification algorithm compared to random resampling. The fourth experiment tested the scan-rescan reliability of the proposed population-based analysis, and the fifth experiment sensitivity of the population-based analysis to anatomical changes associated with normal aging.

All statistical analysis and implemented using R 3.1.1 [176] with the ggplot2 package for plotting [241].

5.3.1 Data Acquisition

Clinical data included diffusion-weighed volumes acquired from healthy volunteers with a GE 1.5T scanner with a voxel size of 2mm^3 and image size 128x128 and 72 slices. For each volunteer, a total of 71 volumes were acquired, with seven T_2-weighted volumes (b-value = 0 s/mm^2) and 64 diffusion-weighted volumes with distinct gradient encoding directions (b-value = 1000 s/mm^2). 80 volunteers were scanned with ages ranging between 25 and 65 and equal numbers in each sex. An additional five volunteers were scanned with three repetitions each to assess scan-rescan reproducibility and reliability.

5.3.2 Image Preprocessing

Diffusion-weighted MR image data was preprocessed using FSL [93] as follows. First, the diffusion-weighted MR images were corrected for motion and eddy current artifacts by affine registration to the first T_2-weighted volume using FSL FLIRT with the mutual information cost function. The gradient encoding directions were rotated to account for the alignment [119], and non-brain tissue was removed using FSL BET.

A diffusion tensor atlas was constructed from the 80 subjects. This was done by first fitting single tensor models using FSL DTIFIT and then constructing a population-specific template by deformable tensor registration using DTI-TK [256] [253]. The 3x5 scan-rescan volumes were registered to the atlas, and the deformation fields were retained for each case.

We generated tractography curves through the standard deterministic streamline approach, in which fiber trajectories are considered a 3D space curves whose tangent vector is equated with the fiber orientation of the voxelwise diffusion models [259]. This process proceeds by evolving a solution to a differential equation with some initial condition at a given seed position. During tracking, we computed fiber orientations from the principal tensor direction, estimated from the tri-cubicly interpolated diffusion-weighted image. We used the following tracking parameters: an angle threshold of 45°, step size of 1.0 mm, minimum fractional anisotropy 0.15, and 2 seeds per voxel. Whole brain tractography was performed in both the population atlas and each individual subject.

5.3.3 Evaluating inter-fiber distances

In this experiment, we evaluated how well several inter-fiber distances reflect the structure of manually selected fiber bundles, measured in relation to the proposed fiber curve representation described in Sec. 5.2.1. We manually selected eight bundles from the atlas tractography using the multiple region of interest approach with two inclusion and one exclusion mask. This included the anterior thalamic radiation, arcuate, cingulum, cortico-spinal tract, inferior longitudinal fasciculus, forceps minor, forceps major, and uncinate. The Dunn Index (DI) was used to measure the ability of a given inter-fiber distance to distinguish between bundles. Given a distance measure $d(x, y)$ and manual labels π, this index measures the ratio of the minimum distance between clusters to the largest distance within clusters:

$$DI = \frac{\min_{\pi(a)!=\pi(b)} d(C_a, C_b)}{\max_{\pi(c)=\pi(d)} d(C_c, C_d)} \tag{5.7}$$

A higher score implies that clusters are well separated relative to the spread within the clusters, suggesting the given distance measure is useful for distinguishing the given bundles. We applied this test to the Euclidean distance in the feature space described in Sec. 5.2.1. For comparison, we also computed the Dunn Index for numerous other inter-fiber distances including the minimum endpoint distance, the Chamfer distance, and the Hausdorff distance. We also symmeterized the Chamfer and Hausdorff measures with the min, max, and mean of the left and right oriented distances. We also included two variants on the closest point transform approach in which the landmarks are either chosen at random or placed on a volumetric grid. Each experiment was repeated with 30 resampling to obtain an estimate of the mean Dunn Index and its uncertainty.

The results show the Euclidean distance associated with the proposed approach to have the highest Dunn Index of 0.95 (Fig. 5.2). However, performance also depended on the number of landmarks included, with a critical number of landmarks being 15, after which performance increased little. We also found runtime costs to increase linearly with the number of landmarks. Among the other distance measures, we found the mean-symmeterized Hausdorff distance to perform best with a Dunn Index of 0.8. In comparison to other approaches to landmark selection, random landmarks performed worst, and gridded landmarks performed second best.

5.3.4 Evaluating fiber bundle clustering

In this experiment, we examined the clustering algorithm described in Sec. 5.2.2 and tested its ability to recover labels from manually delineated bundles. We manually delineated bundles similarly to the previous experiment, and made an additional "harder" set of bundles that included three segments of the body of the corpus callosum, fornix, superior thalamic radiation, and posterior thalamic radiation. We selected these bundles for both the tractography atlas and a single subject, resulting in four test cases, which we refer to as "easy atlas", "hard atlas", "easy subject", and "hard subject".

We applied our clustering algorithm to each case and measured agreement with the manual labels with the Adjusted Rand Index (ARI) [86]. The Rand Index measures the number of pairs of labels that are either the same group or different groups given two different clusterings. The ARI gives a high score to labels that agree and includes an "adjustment" to give an expected score of zero

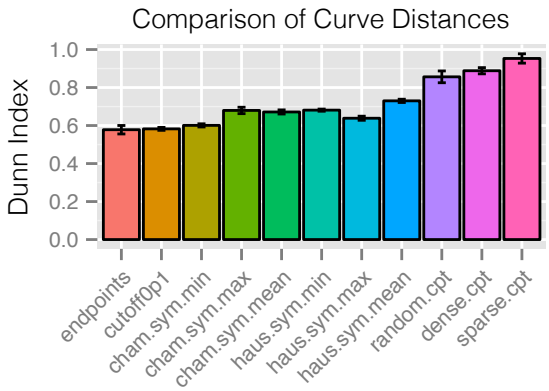


Figure 5.2: Results from the first experiment described in Sec. 5.3.3. The left panel shows the Dunn Index for several inter-fiber distance measures, where a higher score indicates a closer relationship between the distance measure and manual bundle labeling. This shows that the Euclidean distance of the transformed curves performs well relative to other options, although performance was dependent on the number of landmarks, as shown in the middle plot. In this dataset, the performance levels out at around 15 landmarks; however, we found that more complex datasets require more landmarks to see similar convergence. We also found runtime costs were also linearly related to the number of landmarks.

to uniformly random labels. This was repeated with 30 resampling to obtain a the mean ARI and its uncertainty. We also ran this experiment over a range of λ threshold values, and performed similar experiments with spectral and hierarchical clustering for comparison. We used the mean-symmeterized Hausdorff distance in the comparisons, as it was the best performing in the previous experiment.

The results show that the best performing method depended on the dataset (Fig. 5.3). For the easy atlas dataset, all methods performed nearly perfectly. For the easy subject dataset, spectral clustering performed worst, hierarchical performed best, and our proposed method was close to best. For the hard atlas dataset, hierarchical performed worst and the proposed approach was comparable to spectral. For the hard subject dataset, hierarchical performed worst, and the proposed approach was slightly better than spectral.

5.3.5 Evaluating bundle simplification

In this experiment, the goal was to evaluate the performance of the simplification algorithm on several fiber bundles in comparison to random resampling. The experiment was conducted with four bundle datasets selected from a single subject: forceps minor, forceps major, arcuate fasciculus, and

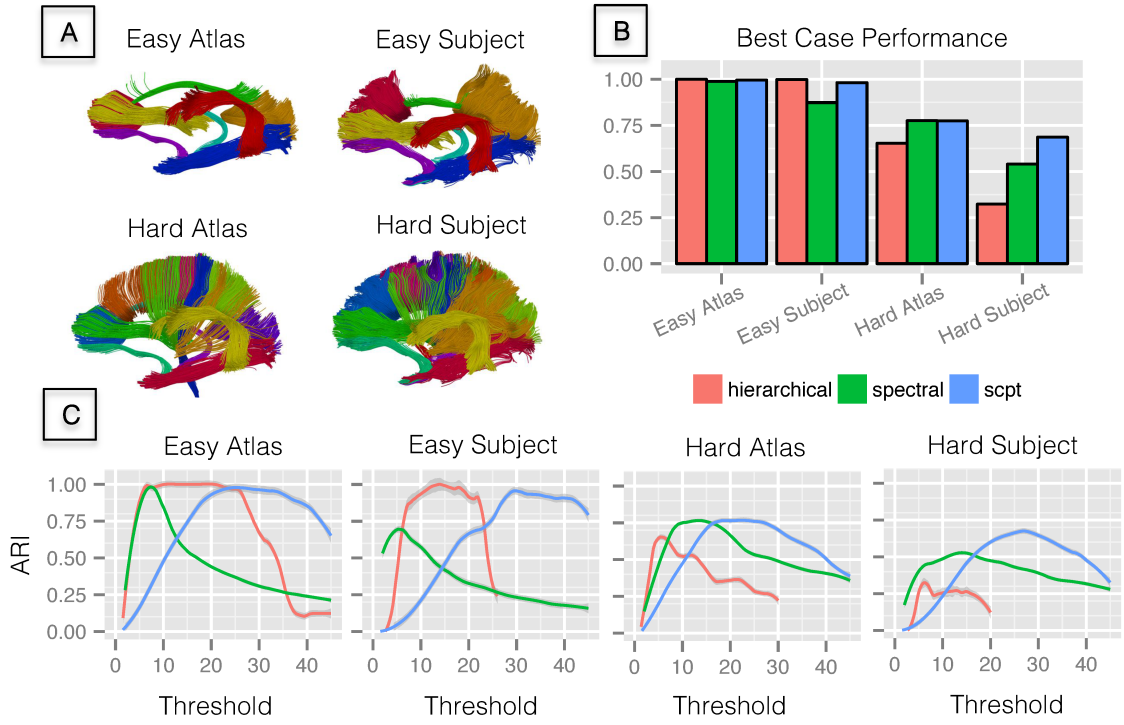


Figure 5.3: Results from the second experiment described in Sec. 5.3.4. The proposed clustering algorithm was applied to four datasets, which include atlas and subject data with a small number (eight) of bundles and a larger number (fifteen), as shown in the top left. We compared this approach to spectral and hierarchical clustering with the mean symmeterized Hausdorff distance. The bottom four plots show performance as a function of clustering threshold, and the best case performance is shown in the bar chart at the top right. It should be noted that the behavior of the threshold depends on the algorithm, so the peaks should not be expected to align. The results show the performance varied across datasets. All methods performed well in the easy atlas and subject datasets. The hard atlas dataset showed the proposed method to be comparable to spectral and better than hierarchical. The hard subject dataset showed the proposed method to outperform both others.

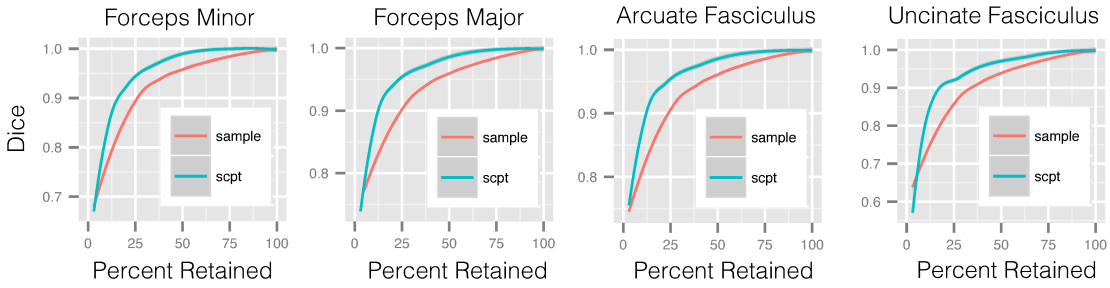


Figure 5.4: Results from the third experiment described in Sec. 5.3.5. This shows performance of the simplification algorithm in several bundles. Each plot shows the simplification error measured with the Dice coefficient. The proposed method was compared to random sampling, and both methods are plotted by the percentage of curves retained from simplification. The results show the proposed method performs consistently better in nearly all cases.

uncinate fasciculus. The proposed simplification algorithm was applied with λ ranging from zero to five, and each condition was repeated 10 times to estimate stability. For comparison, a subsampling procedure was also tested, wherein a random subset of curves was selected with a percentage of retained curves varied from 0 to 100%. To evaluate performance, 2mm^3 volumetric masks were created from the simplified bundles in each condition, and the Dice coefficient [59] was measured between each of these masks and that of the full resolution bundle. Performance was plotted in relation to the percentage of curves retained from simplification.

The results show that the proposed simplification algorithm performed generally better than random subsampling (Fig. 5.4) with a similar pattern in the all tested bundles. The performance difference of the methods was greatest when 10 to 25% of the fibers were retained, in which case, the difference in Dice coefficient was greater than 0.05, representing about 15% improvement in performance relative to the low end dice score. Looking at aggregated runtime across datasets and parameter settings, subsampling was found to take 16.92 ± 0.63ms, and the proposed method took 1179.87 ± 25.32ms.

5.3.6 Evaluating reproducibility and reliability

In this experiment, we applied the population-based analysis described in Sec. 5.2.3 to scan-rescan data and tested its ability to produce similar fiber bundle metrics across multiple *in vivo* scans of the same individual. We then used the 50 landmarks from the full tractography dataset to construct statistical models for each bundle. For each of the 3x5 scan-rescan subjects, we followed the process

described in Sec. 5.2.3 to segment fiber bundles with a threshold probability of 0.99. For each bundle, we computed five fiber bundle metrics: mean bundle length, and bundle-averaged fractional anisotropy, mean diffusivity, radial diffusivity, and axial diffusivity [50]. For comparison, we also performed bundle selection with a manual region-of-interest based approach described in the previous experiment.

We measured scan-rescan reproducibility and reliability of the fiber bundle metrics with the coefficient of variation (CV) [83] and the intraclass correlation coefficient (ICC) [23]. The CV was measured for each subject from mean μ_s standard deviation σ_s, $CV = \sigma_s/\mu_s$. A lower MCV score indicates higher reproducibility with units that are normalized to lie roughly between zero and one. The ICC was measured using the between-subjects variance σ_b^2 and within-subjects variance σ_w^2, with $ICC = \sigma_b^2/(\sigma_b^2 + \sigma_w^2)$ A larger ICC indicates there is more variance between than within subjects. This takes a maximum value of one and values above 0.75 indicate high reliability. Our implementation used the R 'ICC' package [245].

Results from this experiment are shown in Figs. 5.5 and 5.6. We generally found the bundles segmented with the proposed method to agree well with those found with the manual region-of-interest approach. The bundle metrics were not significantly different between the proposed and manual region-of-interest conditions. The bundle-averaged diffusion indicies were highly reliable with a CV of 2% and an ICC of above 0.7. The bundle-average length was had a slightly higher CV of 5% and comparable ICC.

5.3.7 Evaluating sensitivity to aging

In this experiment, the proposed population-based analysis was applied to the study of normal aging. The goal was to evaluate the sensitivity of fiber bundle metrics to the white matter effects related to aging in comparison to the manual region of interest approach. The methods were applied to the 80 subject normal population to obtain five fiber bundle metrics of the forceps minor, which traverses the anterior portion of the corpus callosum and has well documented changes with age. This experiment compares linear regression models of normal aging, in which age is the predicted outcome of a linear combination of intracranial volume, and each tractography-based metric. The performance in each condition was assessed using R^2 to indicate the total variance accounted for by the model, and normalized coefficient β to indicate the effect size.

In general, the results support previous findings of age-related changes in anterior white matter

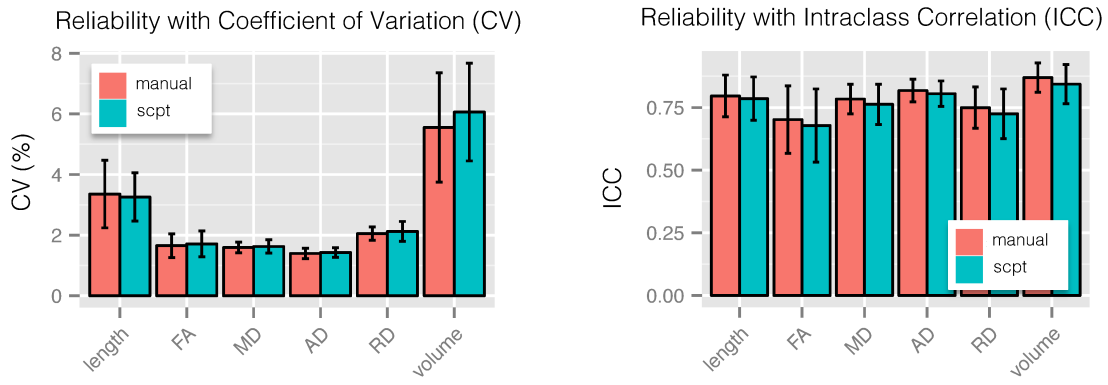


Figure 5.5: Results from the fourth experiment described in Sec. 5.3.6 testing the scan-rescan reproducibility and reliability of the proposed population-based analysis in comparison to a manual region-based approach in six fiber bundle metrics: fractional anisotropy (FA), mean diffusivity (MD), radial diffusivity (RD), axial diffusivity, bundle length, and bundle volume. The experiment included five subjects and three repeated scans, and the results showed comparable reproducibility and reliability between the manual and proposed automated approaches. The left and right plots report the coefficient of variation (lower is better) and intraclass correlation (higher is better) with 95% confidence intervals, respectively.

(Fig. 5.5). Both manual region selection and the proposed method showed sensitivity of diffusion tensor measures to age, but several measures showed different performance characteristics. Three bundle metrics showed comparable performance: bundle volume (R^2 = 0.21, β = -4.35), RD (R^2 = 0.13, β = 2.46), and MD (R^2 = 0.13, β = 2.315). FA was slightly better using manual selection (R^2 = 0.12, β = -1.98) than the proposed method (R^2 = 0.10, β = -1.43). AD was slightly better with the proposed method (R^2 = 0.12, β = 2.07) than manual selection (R^2 = 0.11, β = 1.63). Bundle length was significantly better using the proposed method (R^2 = 0.16, β = -3.28) than manual selection (R^2 = 0.11, β = -2.39).

5.4 Discussion

The results indicate that the sparse closest point transform is generally useful for fiber bundle modeling tasks including clustering large datasets, simplifying bundle geometry, and segmenting similar bundles across a population. The primary strength here is the resulting representation, which enables standard tools for statistical and machine learning methods to be used with tractography datasets. Furthermore, it does this without introducing a prohibitively high dimension in the transformed

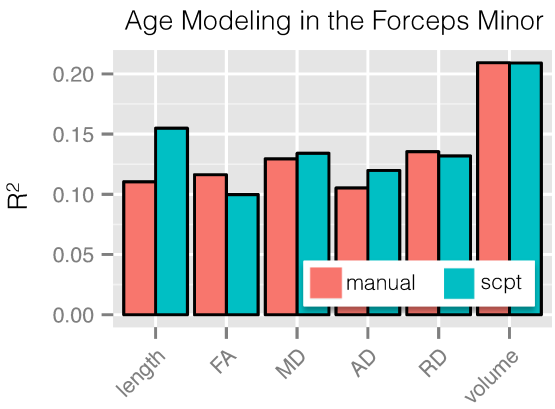


Figure 5.6: Results from the fifth experiment described in Sec. 5.3.7 testing the sensitivity to normal aging in a population of 80 subjects. The analysis focused on the forceps minor, which traverses the anterior portion of the corpus callosum and has well documented changes with age. The proposed population-based analysis was comparison to a manual region-based approach in six fiber bundle metrics: fractional anisotropy (FA), mean diffusivity (MD), radial diffusivity (RD), axial diffusivity, bundle length, and bundle volume. For each bundle measure and method, multiple linear regression models were fit to model age with respect to intracranial volume, sex and each tractography-based metric. The plot shows the resulting R^2 of each model (bigger is better), showing comparable performance in bundle volume, MD, and RD, improved performance in bundle length, and AD, and slightly worse performance in FA.

space, due to the sparseness of the landmarks used in the transform. This contrasts with the standard dense volumetric grid representation of distance and closest point transforms, which require a regular sampling as dense as the smallest discriminative feature. A greater dimensionality can reduce the performance of learning algorithms, i.e. the peaking phenomenon [192], which is another benefit of using this sparse representation. The results from the first experiment also indicate that the careful selection of landmarks through geometry processing can improve the results compared to random or coarse regular sampling. This is likely because it can place landmarks near groups of endpoints and areas of high curvature, which are also likely to be the most discriminative features when comparing curves.

The results of the second experiment also indicated strengths of the proposed clustering method in comparison to standard hierarchical and spectral clustering algorithms. These other options can take advantage of a rich amount of data by comparing all pairs of fiber curves; however, the proposed clustering algorithm was found to be as well or better at recovering manually labeled bundles than these other methods. Because the proposed method avoids computing the full pairwise similarity matrix, it can be applied to large whole brain datasets without the typical subsampling used in the

pairwise approaches. The strengths of the clustering algorithm should mostly be attributed to the application of the DP-means algorithm, which provides the efficient runtime and ability to select the number of clusters from the data. However, this is only possible due to proposed representation of the curve data with the closest point transform. Future work could explore more advanced probabilistic models, such as a Dirichlet process mixture model [152], which can potentially model more complex patterns, at the cost of higher computational requirements.

The results of the third experiment show how the same clustering algorithm is useful for simplifying fiber bundles. This process is able to retain the important geometric features of the data. This contrasts with random sampling, which tends to produce significantly worse simplifications, as measured by the Dice coefficient. One limitation of the proposed method is the greater computation time. For individual bundles, the runtime was only several seconds; however for full brain tractography, it can take significantly longer. Nevertheless, this kind of processing can be done offline and greatly improve the rendering speed of tractography visualizations, particularly rendering with 3D tubes, which is exceeding slow for large datasets.

The fourth and fifth experiments showed that the proposed population-based analysis is not only reliable across scans but also useful for mapping age-related change in white matter. This transform-based approach succeeds because it retains a pose and shape information about the original curves, while being compatible with a simple segmentation algorithm using a Gaussian model. Furthermore, the proposed method compared favorably to manual bundle selection, which can be costly in time and require expert knowledge. It also outperformed the manual selection approach in the normal aging modeling based on fiber bundle length. This suggests that manual region-based selection perhaps excludes short fibers in the bundle that do not reach the regions, but are nevertheless related to age-related decline. Another interesting result was the good performance of the admittedly simple one-class classifier. An interesting open question for future work is whether there is a benefit to using more complex one-class classifiers, e.g. using support vectors [210] or neural networks [22].

In general, this approach has potential to be useful in a broader context than modeling fiber bundles as well. For example, curve clustering can be a useful tool for analyzing trajectory datasets, e.g. to better understand traffic flow or the behavior of particles in numerical simulations of physical processes. The bundle segmentation approach could also be useful in these other areas for categorizing incoming data or detecting anomalies. However, there are several limitations of the proposed method to note. First, the representation obtained from the sparse closest point transform

is necessarily lossy, and in particular, it will disregard very complex properties such as cycles or sharp doubling-back of curves. It is also does not account for any time-domain information. The efficiency of the method also depends on a fixed set of landmarks, so if it were applied in online data analysis, care should be taken that the landmarks are chosen to accommodate future data. These issues are not typically encountered in diffusion MR tractography datasets; however, they remain open problems for these other application areas.

5.5 Conclusion

In this chapter, we investigated the use of a sparse closest point transform in performing a number of fiber bundle modeling tasks. The experiments showed that the combination of this technique with simple statistical and machine learning techniques can perform very well in comparison to several commonly used and domain-specific algorithms. In particular, the results showed that clustering with the DP-means algorithm can perform as well or better than hierarchical and spectral clustering, with substantially better runtime characteristics. Furthermore, this clustering algorithm can be used for bundle simplification by selecting subsets of curves that best represent the large scale geometry of the bundle, with performance generally superior to random subsampling. Finally, the combination of deformable registration and a simple Gaussian one-class classifier was shown to provide a method for selecting fiber bundles with as good or better performance than manual region-based selection. In particular, this had generally good scan-rescan reliability and improved performance in measuring age-related changes in fiber bundle length. In summary, these results suggest that the proposed closest point transform is a generally useful method for analyzing tractography data with statistical and machine learning tools. Looking forward, this approach is potentially useful for other clinical applications, such as anatomical localization in surgical planning and population studies that aim to quantify anatomical variation in white fiber bundles in relation to health and disease.

Chapter 6

Kernel Regression Estimation of Fiber Orientation Mixtures in Diffusion MRI

Diffusion MR imaging provides an *in-vivo* probe of tissue microstructure that enjoys numerous applications to neuroscience and clinical studies. This is due to its unique ability to image local patterns of water molecule diffusion, which reflect physical properties of biological tissue [13]. These patterns enable the quantification of brain white matter microstructure, as diffusion exhibits anisotropy due to the geometry of neuronal axon projections [170]. This is useful for making local measurements of fiber orientations through diffusion modeling [225] and more global fiber bundle reconstructions through tractography [43] [227]. In addition, atlas-based reconstructions can reveal population-wide features of anatomy and serve as a reference for comparing individuals [143]. Both tractography and population-based atlasing are applications that depend on a number of basic image processing tasks, including interpolation, smoothing, and fusion [121]. For tractography, interpolation is needed when determining fiber orientations off the voxel grid, and smoothing is needed to control errors due to noise and other image artifacts. For atlas construction, interpolation is needed to resample images into a common space, and fusion is needed to derive a composite image from a population. This chapter develops and evaluates methods for performing these tasks with support for multiple distinct fibers, which are important for resolving complex sub-voxel fiber configurations due to crossings and

partial volume effects [218].

In particular, the goal of this work is to develop and evaluate methods for model-based image processing that are useful for interpolation, smoothing, and fusion tasks with the multi-compartment ball-and-sticks diffusion model. For this, we derive a kernel regression framework for estimating fiber orientation mixtures, which represent multiple fiber orientations per voxel and associated volume fractions. In contrast to signal-based image processing, this approach estimates fiber models from a collection of fiber models contained in volumetric parametric maps. This is accomplished by extending a kernel regression image processing framework for vector-valued images [204] and building on prior work on model-based diffusion MRI processing [205] [34]. We show how this formulation generalizes a variety of tasks and allows for simple data-adaptive extensions for bilateral filtering and model-selection, which may be generally useful for parametric model-based image processing. Our experimental evaluation first characterizes performance with computational phantoms, and then explores applications to quantitative tractography-based analysis of fiber bundles and multi-fiber atlas construction.

The rest of the chapter is organized as follows. In Secs. 6.1 and 6.2, we discuss related work and background material. In Sec. 6.3, we describe our proposed model-based estimator, outline optimization techniques, and describe its application to streamline tractography. In Sec. 6.4, we describe experimental evaluation of our approach with computational phantoms and *in vivo* clinical data. The synthetic data experiments evaluate the performance of our method in relation to fiber orientations, volume fractions, compartment count, and tractography-based connectivity. The *in vivo* data experiments first measure the scan-rescan reproducibility and reliability of fiber bundles metrics in individual subjects. We then build a multi-fiber tractography atlas from 80 healthy subjects to study population-wide properties of white matter. In Secs. 6.5 and 6.6, we discuss our results and conclude.

6.1 Related work

In this section, we briefly review related work and outline the distinguishing features of the present approach.

This chapter examines a model-based approach for diffusion MRI processing. This contrasts with signal-based approaches that apply filters to the vector-valued diffusion-weighted MRIs, while

accounting for factors such as the gradients strengths, directions, and Rician noise [18]. Model-based approaches instead operate on lower-dimensional mathematical representations of the diffusion signal [121]. Previous work has developed such approaches for single diffusion tensors [165] and orientation distribution functions (ODFs) [74, 249]. We differ by considering multi-compartment models, which are a parametric alternative to ODFs with the advantage of compartment-specific measures and isotropic diffusion modeling. However, two notable challenges for processing multi-compartment models are model selection (to determine the number of compartments) and compartment matching (to solve the combinatorial problem of finding corresponding compartments among some set of voxels) [175].

The approach is related to some prior work on multi-fiber model-based image processing. Yap *et al.* proposed an approach for multi-fiber atlas construction with fiber orientations extracted from ODF peaks [248], which dealt with similar computational problems but was not designed for parametric models. The present work differs by considering atlas construction with the ball-and-sticks diffusion model and by relaxing some assumptions on fiber correspondence and fiber count. The prior work of Taquet *et al.* [205] first explored the idea of parametric model-based interpolation of multi-tensors and its application to tractography and atlas construction, developing a rich framework for multi-tensor processing based on Gaussian mixture simplification [207]. Our work differs by considering directional measures of divergence that are compatible with the ball-and-sticks model. This is necessary because the Burg matrix divergence employed by the full multi-tensor framework is not well-defined for the anisotropic stick compartments, so we instead develop a divergence measure based on the directional Watson distribution to serve the same purpose. We also differ by incorporating data-adaptive extensions to support both bilateral filtering and model selection, in contrast to previous work that employed only spatial kernel weights and local-maximum model selection. Some of the methods presented here have been explored in previous conference work [34] [36], but those included neither the full data-adaptive kernel regression estimator nor the full experimental results presented here.

6.2 Background

Our work focuses on image analysis with multi-fiber ball-and-sticks diffusion models, which falls into the multi-compartment class of diffusion models [19]. This is a multi-tensor constrained to include an

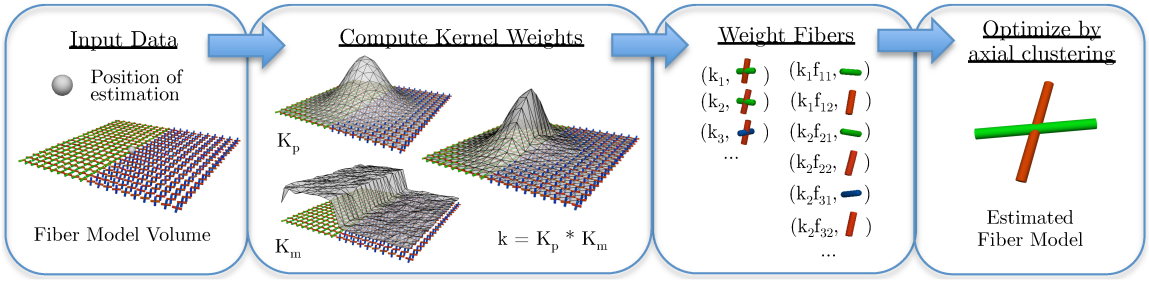


Figure 6.1: A flowchart illustrating the steps in the proposed method, detailed in Algo. 3. The input is a position for estimation and model-valued volumetric data. Then, kernel weights are computed from the product of spatial and data-adaptive factors and normalized to sum to one. Then, weights are distributed among the fiber compartments and multiplied by volume fraction. Finally, the estimated model is found by a clustering-based optimization procedure, detailed in Algo. 2.

isotropic "ball" compartment and a number of completely anisotropic "stick" compartments. These constraints allow the model to achieve good performance for single shell gradients with low b-value acquisitions at intermediate field strengths [244], which are common in clinical applications. The ball compartment can account for isotropic diffusion, and the volume fractions associated with each compartment can account for partial volume effects and mixtures of bundles at crossings. There is also evidence that the volume fraction maps provide a quantitative measure for clinical studies [90]. With this model, the predicted diffusion-weighted signal S is given by:

$$S = S_0 \left(f_0 \exp\left(-db\right) + \sum_{j=1}^{N} f_j \exp\left(-bd(\vec{g}^T \mathbf{v}_j)^2\right) \right) \tag{6.1}$$

for a particular gradient encoding direction $\vec{g}$, b-value b, and baseline signal S_0. This includes N fiber compartments with fiber orientations $\|\mathbf{v}_j\| = 1$, fiber volume fraction $0 \leq f_j \leq 1$, $\sum_{j=0}^{N} f_k = 1$, and shared diffusivity $d > 0$. It's important to note that there is no sign associated with fiber orientations, so there is an equivalence $\mathbf{v}_j \sim -\mathbf{v}_j$. We use the following parameterization to denote a fiber orientation mixture M with N compartments: $M = \{(f_j, \mathbf{v}_j)\}_{j=1}^{N}$. In our experiments, we fit the model to the diffusion signal with the Bayesian approach of Behrens *et al.* implemented in the FSL software library [93].

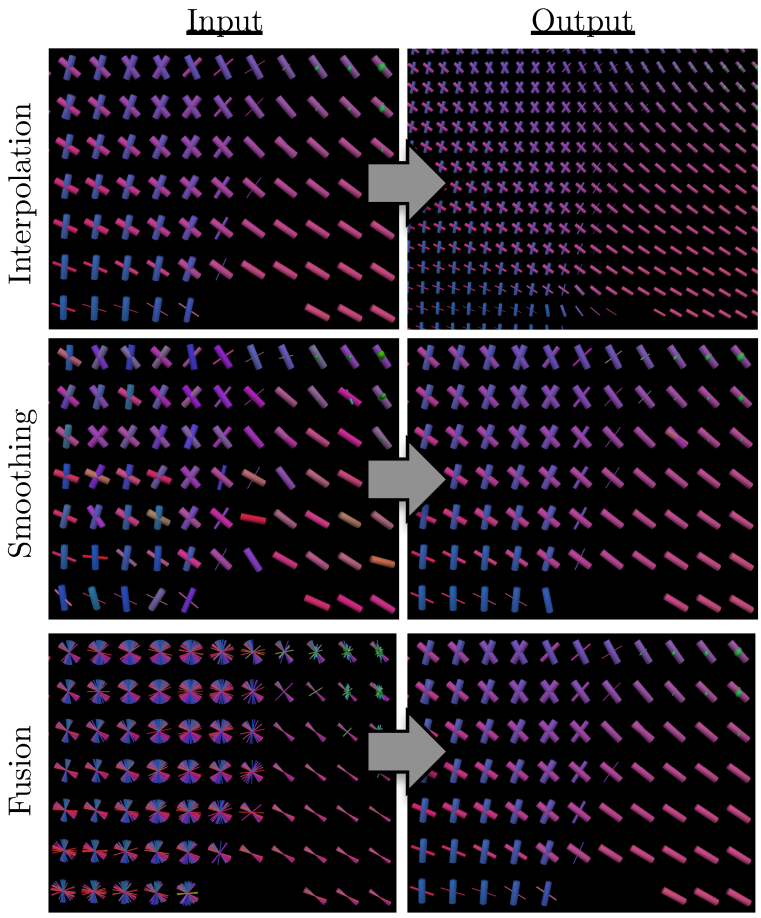


Figure 6.2: Example uses of the proposed estimator for interpolation, smoothing, and fusion tasks. Slices were taken from average human brain data at the junction of the corpus callosum (pink), corona radiata (blue), and cingulum (green). Fibers are colored according to their orientation, and thickness encodes volume fraction. The top panel shows interpolation at twice the original resolution, demonstrating continuity of both fiber orientations and volume fractions. The middle panel shows smoothing to reduce noise-induced angular error. The bottom panel shows fusion of 80 aligned subjects to produce an atlas representing the population.

Table 6.1: A summary parameters used for estimation, including a range of values found in our evaluation and experiments.

Symbol	Name	Description	Range
h_p	Spatial bandwidth	Specifies the size of the region	1.0-3.0 mm
h_m	Data-adaptive bandwidth	Specifies the data sensitivity	0.3-0.6
λ	Regularization	Specifies model selection	0.99-0.9999
K_{max}	Maximum complexity	Specifies maximum fiber compartments	2-3

6.3 Methods

In this section, we present our model-based kernel regression estimator for fiber orientation mixtures, followed by a description of extensions for data-adaptive processing. We then outline the necessary optimization routines and practical implementation details. Finally, we describe how this can be applied to deterministic streamline tractography.

6.3.1 Kernel Regression Estimation

We now formulate our approach for kernel regression estimation of fiber orientation mixtures, which is the main contribution of the chapter. This builds on the prior work of Takeda *et al.* on kernel regression image processing for scalar and vector data [204] and prior multi-compartment processing [34] [207]. The general idea of kernel regression, however, has also been proposed in several other contexts as the moving average, kernel smoothers, etc. [120]. The simplest of such is the Nadaraya-Watson estimator (NWE) [235], which we focus on here. Given a collection of data $\{(x_i, y_i)\}_{i=1}^{C}$ and a desired regression function $\hat{y}(x_0)$ evaluated at point x_0, the NWE assumes a product kernel density distribution for observed data and takes the regression function to be the conditional expectation $\hat{y} = E(y|x_0)$. This can be simply expressed by the following least-squares problem:

$$\hat{y}(x_0) = \underset{\beta}{\operatorname{argmin}} \sum_{i=1}^{C} K\left(x_i, x_0\right) \|y_i - \beta\|^2 \tag{6.2}$$

$$K(x_i, x_0) = \exp\left(-\|x_i - x_0\|^2 / h^2\right) \tag{6.3}$$

with a kernel function K and bandwidth parameter h. In practice, larger values of h produce a smoother estimate. The kernel K can be chosen from a variety of options, but we use the standard Gaussian. The advantage of this approach is that it makes few assumptions about the structure of the data and consequently generalizes well to a number of tasks. In particular, Takeda *et al.* [204] presented a general framework for image processing based on this approach, showing it can be used for interpolation, smoothing, and fusion. The goal here is to further extend this approach to model-based estimation of fiber orientation mixtures.

The primary limitation of the standard kernel regression approach is the assumption of vector-valued images, for which the Euclidean distance $\|x - y\|$ is a reasonable measure of discrepancy.

This assumption breaks down for diffusion model-valued image data due to various constraints and the non-Euclidean geometry of the models, e.g. fiber volume fractions and orientations. This issue can be addressed by instead considering a model-based measure of discrepancy d_m^2 and by performing optimization over the constrained space of models. Given an input position p_0 and local neighborhood of data $\{(p_i, M_i)\}_{i=1}^{C}$ with $M_i = \{(f_{ij}, \mathbf{v}_{ij})\}_{j=1}^{N_i}$, we formulate this extended estimator $\hat{M}(p_0)$ as follows:

$$\hat{M}(p_0) = \underset{M}{\operatorname{argmin}} \sum_{i=1}^{C} K_p(p_i, p_0) d_m^2(M_i, M) \tag{6.4}$$

$$K_p(p_i, p_0) = \exp\left(-\|p_i - p_0\|^2 / h_p^2\right) \tag{6.5}$$

given a positional bandwidth parameter h_p. We denote the neighborhood size here as C, and in practice, we choose a neighborhood sufficiently large to include three standard deviations of the Gaussian kernel, e.g. a 7x7x7 window of 1 mm^3 voxels for $h_p = 1$ mm. We define the model-based discrepancy d_m^2 for fiber mixtures as follows:

$$d_m^2(M, \hat{M}) = \min_{\pi} \sum_{j} f_j d_f^2(\mathbf{v}_j, \hat{\mathbf{v}}_{\pi(j)}) \tag{6.6}$$

given a single fiber distance d_f^2:

$$d_f^2(\mathbf{v}, \hat{\mathbf{v}}) = 1 - (\mathbf{v} \cdot \hat{\mathbf{v}})^2 = sin^2(\theta) \tag{6.7}$$

The above definition of d_m^2 is a matching-based distance that accounts for the combinatorial structure of the model by minimizing the sum-squared distances across all possible matching functions π between compartments in M and those in $\hat{M}$. This formulation happens to allow efficient optimization and makes no assumptions on compartment ordering or total count. Both d_m^2 and d_f^2 also have interesting relationships to statistical measures of divergence, which are discussed in the appendix.

6.3.2 Data-adaptive Extension: Model Selection

The estimator puts no constraints on the complexity of the estimated model, so some mechanism is needed to perform model selection. One solution is to choose a fixed number of compartments; however, this can overestimate in some areas and underestimate in others. Data-driven approaches can instead use the fiber counts from the local neighborhood to estimate the model complexity, e.g. the local-maximum estimator used in prior work [207]. Another more conservative data-driven approach is to use the average fiber count (rounded to the nearest integer). We propose another data-driven approach that works by extending the estimator in Eq. 6.4 to include a regularization term for the number of compartments N:

$$\hat{M}(p_0) = \underset{M}{\operatorname{argmin}} \sum_{i=1}^{C} K_p(p_i, p_0) d_m^2(M_i, M) + \lambda N \tag{6.8}$$

given a regularization parameter λ. The goal of this approach is to choose the fiber count that best supports the fiber orientation data in the local neighborhood, rather than just the counts. In later experiments, we compare these different model selection approaches to assess their strengths and weaknesses.

6.3.3 Data-adaptive Extension: Bilateral Filtering

In addition to performing model selection, the estimator can be simply modified to allow for bilateral filtering. The goal here is to avoid blurring interesting features at boundaries of anatomical structures. This issue has been studied for standard kernel regression estimators, and one solution is to include additional weights that reflect similarity between data values [204] [82]. We include this idea by adding a data-adaptive factor as follows:

$$\hat{M}(p_0, M_0) = \underset{M}{\operatorname{argmin}} \sum_{i=1}^{C} K_i d_m^2(M_i, M) + \lambda N \tag{6.9}$$

$$K_i = K_p(p_i, p_0) K_m(M_i, M_0) \tag{6.10}$$

$$K_m(M_i, M_0) = \exp\left(-d_m^2(M_i, M_0)/h_m^2\right) \tag{6.11}$$

given a bandwidth parameter h_m and a reference model M_0.

This can be optimized with the same routines as Eq. 6.8; however, it is more computationally costly because the adaptive kernel weights K_m cannot be precomputed like the spatial kernel weights K_p. For voxelwise smoothing, the reference can be the original input model; for tractography, it can be the model from the previous step in tracking; for interpolation, it can be the estimate without the bilateral factor K_m.

6.3.4 Optimization

Next, we describe optimization routines for the proposed estimator and outline some practical issues for implementation.

The estimator in Eq. 6.8 involves an optimization problem in which the weighted sum of model-based distances must be minimized, subject to a regularization term for model selection. Substitution of d_m and d_f gives a simpler form:

$$\hat{M}(p) = \underset{M,\pi}{\operatorname{argmax}} \sum_{i}^{C} \sum_{j}^{N_i} k_i f_{ij} (\mathbf{v}_{ij} \cdot \hat{\mathbf{v}}_{\pi(ij)})^2 + \lambda N \tag{6.12}$$

given the kernel weights k_i in Eq. 6.10. This is equivalent to solving the weighted axial DP-means clustering problem [37], which is similar to the k-means algorithm with two extensions. First, clustering is performed with axial variables [199], which are equivalent to fiber orientations. Second, the number of clusters is estimated from the data as in the DP-means algorithm, a name derived from its relation to Dirichlet Process mixture models [108]. An iterative algorithm for solving this is presented in Algo. 2, and further discussion is in the appendix.

The full procedure for using the estimator is summarized in Fig. 6.1 and detailed in Algo. 3. This process proceeds by first collecting fiber models in a local neighborhood of the point of estimation, then kernel weights are computed by the product of spatial and data-adaptive factors. These weights are normalized to sum to one and then distributed among the fiber compartments of each model. The above clustering problem is then solved to estimate the output number of compartments, fiber orientations, and volume fractions. A summary of the algorithms parameters is shown in Table 6.1.

There are also some practical issues to note when implementing this. First, the procedure is only guaranteed to find a local minima, which depends on the starting conditions. This can be helped by performing several random restarts and taking the solution with the overall minimum ,

e.g. 10 restarts with shuffling in our implementation. There may be some rare configurations where the minimum is not well-defined, e.g. strictly orthogonal fibers. If this occurs, the known model with highest weight can be chosen; however, we check for this condition and never encountered it in practice. Finally, the kernel weights must be normalized by their total sum, as this ensures the total volume fraction of the fiber compartments is conserved and the estimated model volume fractions also sum to one.

6.3.5 Tractography

We applied the above estimator to deterministic tractography using a generalization of the standard streamline approach to account for multiple fibers. In the standard approach, a fiber trajectory is considered a 3D space curve whose tangent vector is equated with the fiber orientation of the voxelwise diffusion models. This process proceeds by evolving a solution to a differential equation with some initial condition at a given seed position. Typically, some geometric criteria are also used to stop and exclude fibers, including angle threshold and minimum and maximum length [259].

This approach must be adapted when multiple orientations are present [175], and we use the following modifications. During tracking, one of the N possible fiber compartments must be chosen for the next step, so we choose the fiber with the smallest angular difference to the previous step, among those below a given angle threshold. We also incorporate additional volume fraction termination criteria, where tracking is stopped if the volume fraction of the chosen fiber compartment is below a given threshold. Finally, we use the proposed model-based estimator for interpolation and smoothing during tracking. We also retain the volume fraction parameter during tracking in order to estimate statistics across fiber bundles.

6.4 Experiments and Results

This section describes the evaluation of our method with four experiments, including a discussion of the datasets, experimental design, evaluation metrics, and results in each experiment. Two synthetic data experiments were conducted with computational phantoms to assess the ability of clustering-based optimization to match compartments and also to measure performance of our proposed model selection and bilateral filtering extensions. Two *in vivo* data experiments were conducted with clinical scans to assess the practical benefits of our approach in modeling human brain white matter

fiber bundles. The first investigated scan-rescan reliability of quantitative fiber bundle metrics, and the second investigated population-level features of white matter by constructing a multi-fiber tractography atlas from 80 subjects.

All statistical analysis was performed using R 3.1.1 and ggplot2 [241]. Unless stated otherwise, the estimation parameters were $h_p = 1.5$ mm, $h_m = 0.5$, $\lambda = 0.99$, and the kernel support radius was 5 voxels. All experiments fit two-fiber compartment ball-and-sticks diffusion models with the Markov-Chain Monte-Carlo procedure implemented in FSL XFIBRES [19].

6.4.1 Datasets and Preprocessing

Synthetic datasets: We generated datasets from two types of computational phantoms, which were defined and synthesized with the ball-and-sticks diffusion model as follows.

The first type of phantom represents a complex boundary between fiber bundles for voxel-based analysis. The phantom included two fibers per voxel and represented two adjacent and perpendicular bundles with a fixed volume fraction of 0.4 with a third bundle that crosses both with a volume fraction varied between 0.2 and 0.4. This left an isotropic volume fraction of 0.2 to 0.4, depending on the condition. Diffusion-weighted images were synthesized with with $S_0 = 10000$, $d = 0.0017$ mm^2/s, 1 mm^3 voxels, and dimensions 30x30x5.

The second phantom represents a more complex set of bundles for both voxel-based and connectivity-based analysis. This included three bundles with a crossing and branching structure. Diffusion-weighted images were synthesized with $S_0 = 10000$, $d = 0.0017$ mm^2/s, 1 mm^3 voxels, and dimensions 71x71x15, as described in Leemans *et al.* [118].

All images were synthesized with seven baseline volumes, 64 diffusion-weighted volumes with b-value 1000 s/mm^2 to match the clinical imaging data. Rician noise was introduced by adding Gaussian noise with standard deviation σ_{synth} to the complex signal and taking the modulus, using Camino [48]. Noise level is reported by the signal-to-noise ratio in decibels, $SNR(dB) = 20\log_{10}(S_0/\sigma_{synth})$.

Human brain datasets: Clinical data included diffusion-weighed volumes acquired from healthy volunteers with a GE 1.5T scanner with a voxel size of 2mm^3 and image size 128x128 and 72 slices. For each volunteer, a total of 71 volumes were acquired, with seven T_2-weighted volumes (b-value = 0 s/mm^2) and 64 diffusion-weighted volumes with distinct gradient encoding directions

(b-value = 1000 s/mm^2). 80 volunteers were scanned with ages ranging between 25 and 65 years and roughly equal numbers between sexes. An additional five volunteers were scanned with three repetitions each to assess scan-rescan reproducibility and reliability. The noise level was estimated to be 22.05 dB using two manually drawn regions-of-interest and the Rician corrected method in [60], $SNR_{stdv}(dB) = 20\log_{10}(\sqrt{2-\pi/2} * \hat{\mu}_{tissue}/\hat{\sigma}_{air})$.

Diffusion-weighted MRI data was preprocessed using FSL [93] as follows. First, the diffusion-weighted MRIs were corrected for motion and eddy current artifacts by affine registration to the first T_2-weighted volume using FSL FLIRT with the mutual information cost function. The gradient encoding directions were rotated to account for the alignment [119], and non-brain tissue was removed using FSL BET.

The 3x5 scan-rescan volumes were used for assessing reproducibility and reliability in the third experiment. For the fourth experiment, a diffusion tensor atlas was constructed from the other 80 subjects. This was done by first fitting single tensor models using FSL DTIFIT and then constructing a population-specific template by deformable tensor registration using DTI-TK [256] [253]. The resulting deformation fields were retained for creation of the multi-fiber atlas, and the single tensor atlas was retained for comparison.

6.4.2 Synthetic Data Experiment with Boundary Phantom

Design:

The first experiment was conducted with the boundary phantom (shown in Fig. 6.3) and was designed to test the fiber matching and bilateral filtering features. For this, we defined two regions of interest: an "on-boundary" two-voxel window, and the remaining "off-boundary" voxels. The proposed estimator was compared to a so called "rank-based" estimator, which is meant to serve as a point of comparison to evaluate compartment matching. This rank-based method sorts fibers in each model based on their volume fraction, and then each compartment is processed independently as if it were a multi-channel volume. We hypothesized that this rank-based approach would introduce greater angular errors than our approach. To test performance relative to previous work, we also compared the bilateral estimator (Eq. 6.9) to one with only spatial weights (Eq. 6.4), which we will refer to as "linear". We hypothesized that the bilateral factor would have lower error at the boundary between bundles. In each condition, we introduced noise and measured deviation from the ground truth with evaluation metrics similar to Ramirez-Manzanares *et al.* [180]. For this,

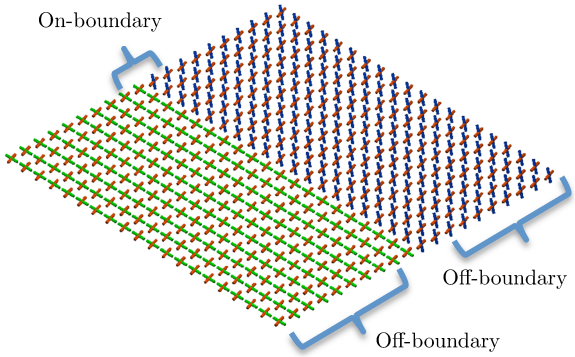


Figure 6.3: The phantom used in the first experiment in Sec. 6.4.2. This represents a complex boundary between fiber bundles and include "off-boundary" and "on-boundary" regions-of-interest.

fibers were matched to the ground truth, and the total angular error and volume fraction error was computed in each voxel and averaged within each region of interest. SNR was varied from 15 to 25, the data-adaptive bandwidth parameter h_m was varied from 0.1 to 1.0, and crossing bundle volume fraction was varied from 0.2 to 0.4. Each condition was repeated 20 times to obtain the sample mean and uncertainty. To exclude effects related to model selection, the number of fiber compartments was assumed to be known and fixed at two; however, the next experiment investigates performance of the model selection component.

Results:

Results are shown in Figs. 6.4 and 6.5. We found isotropic and fiber volume fraction error was similarly lowered in all methods; however the rank-based method had slightly lower error on-boundary. We found rank-based orientation error to be significantly higher than the other methods in all cases. We found both linear and rank-based to introduce significant orientation error at the boundary, exceeding that of noise for SNR > 20. Bilateral filtering significantly reduced the error on-boundary with a negligible increase elsewhere. Upon varying the adaptive bandwidth, we found error had a local minima with good overall performance from 0.3 to 0.5, depending on the SNR. Upon varying the crossing bundle fraction, we found rank-based estimation to introduce very high orientation error when the bundle volume fractions were within 0.05.

6.4.3 Synthetic Data Experiment with Bundle Phantom

Design:

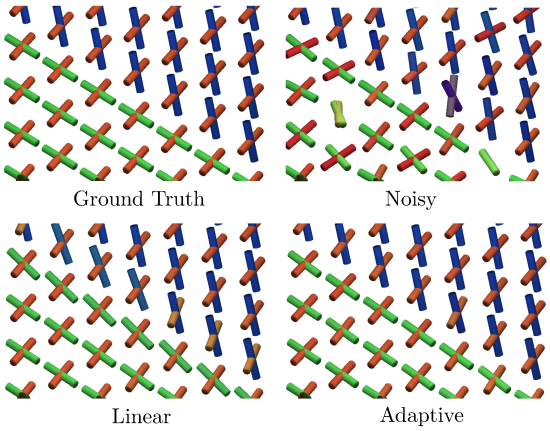


Figure 6.4: Qualitative results from the first experiment described in Sec. 6.4.2. The first panel shows the ground truth phantom at the boundary voxels. The second panel shows fibers fit after adding Rician noise at $SNR = 21.5$. The third and fourth panels show smoothing results using linear (Eq. 6.4) and bilateral (Eq. 6.9) kernel weights. The results show that both estimation techniques reduce noise-induced angular error; however, adaptive estimation can avoid the orientation blurring at the boundary voxels, due to the inclusion of the K_m kernel weights.

The second experiment was conducted with the bundle phantom (shown in Fig. 6.6) and was designed to test the model selection methods and to assess performance in connectivity mapping. The adaptive model selection (Eq. 6.8) technique was compared with the following alternatives: "fixed" (always two fibers), "mean" (rounded local average number of fibers), "max" (local maximum of fiber count, like [207]). We tested both voxel-wise errors and tractography-based connectivity errors. Voxel-wise conditions measured orientation error, volume fraction error, and also "missing" and "extra" fibers error [180]. We also tested connectivity measures derived from deterministic streamline tractography. For this, we manually delineated volumetric masks representing "white matter" along bundle trajectories and "gray matter" regions at bundle endpoints. These gray matter masks were used to seed deterministic tractography, and the resulting tracks were then compared to the manually delineated ground truth. For each bundle, we assessed performance by measuring the valid fiber rate (fraction of seeds that reach their intented target) and the Dice overlap coefficient $D(A, B) = 2\,\mathrm{vol}(A \cap B)/(\mathrm{vol}(A) + \mathrm{vol}(B))$, given bundle mask A and ground truth B [59]. Tractography was performed with 5 seeds per voxel, step size 0.5 mm, 45° angle threshold. SNR was varied from 15 to 25, and the regularization parameter was varied across seven levels, which are listed in Fig. 6.8. Tractography was performed with nearest-neighbor interpolation as a baseline, and rank-based estimation was also performed in all conditions for comparison. We hypothesized

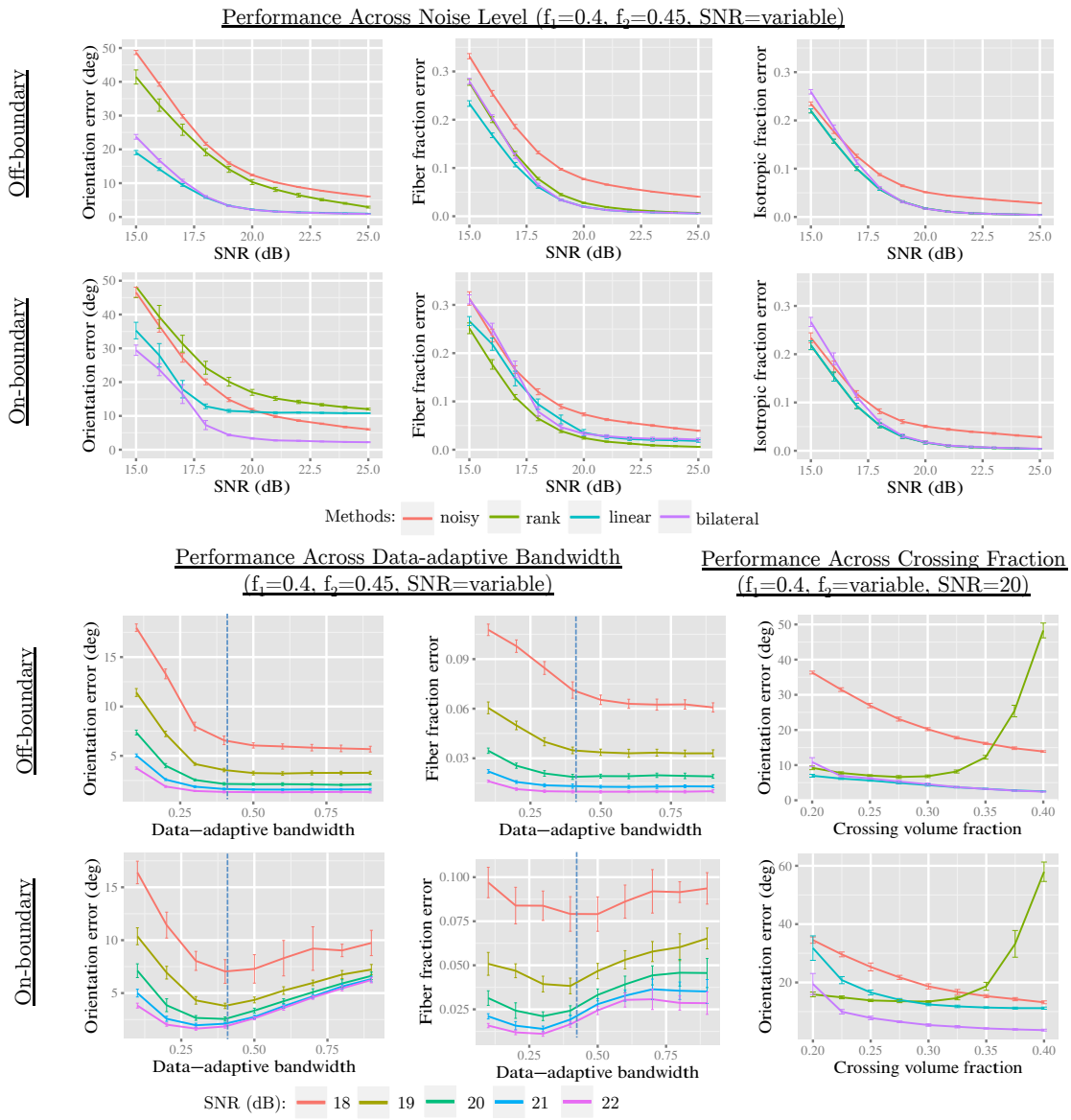


Figure 6.5: Results from the first experiment in Sec. 6.4.2. The labels on the left indicate the region where results were aggregated, and each plot shows either orientation or volume fraction error rates. The plot titles also show the volume fraction of the compartments (f_1 and f_2). The top two rows show results with SNR from 15 to 25, comparing the noisy, rank-based, linear, and adaptive estimation conditions (with $h_p = 1.5$ and $h_m = 0.5$). We see that all methods perform well at reducing volume fraction, with rank-based estimation performing slightly better with on-boundary fiber fractions. We see that rank-based estimation rates poorly in orientation error, particularly in regions where bundle volume fractions are within 0.05 apart. We also see that linear estimation introduces high orientation error in on-boundary voxels, while adaptive estimation avoids this issue. By sweeping across the data-adaptive parameter h_m, we see a local minima in orientation error occurs between 0.3 and 0.5, depending on the SNR.

that mean selection would be most conservative and adaptive selection would improve connectivity mapping and help control extra fiber error while not significantly increasing missing fiber error. All conditions were repeated twenty times to measure the sample mean and uncertainty.

Results:

Results are shown in Figs. 6.6, 6.7, and 6.8. We found similar results for isotropic volume fraction as Exp. 1. Volume fraction error was significantly higher in fixed and max selection, with linear and adaptive performing best. Orientation error was reduced in all conditions with adaptive selection performing best by a small margin. Fixed and max selection showed very high extra fiber error with linear and adaptive performing best. Fixed and max selection had the lowest missing fiber error, with adaptive selection being slightly higher. For tractography-based measures, rank-based estimation performed poorly, nearest and linear selection were better and comparable to each other. We found adaptive, max, and fixed to perform very well, with adaptive performing slightly better is most cases. Upon varying the regularization parameter, we found orientation error, fraction error, and tractography performance to be rather stable, but the extra and missing fiber error varied somewhat. This variation, however, was small compared to the errors introduced with max and fixed selection. After inspecting the results, we found max and fixed selection to be sensitive to noise and problems in model fitting, resulting in over-estimated model complexity such as "fiber splitting", shown in Fig. 6.7.

6.4.4 *In vivo* Data Experiment for Individual Subjects

Design: The third experiment tested the reproducibility and reliability of quantitative tractography metrics obtained with the proposed method in *in vivo* human brain data. For this, we extracted fiber bundles from the scan-rescan dataset, including the following structures: anterior thalamic radiation (atr), cingulum bundle (cing), inferior longitudinal fasciculus (ilf), and uncincate fasciculus (unc) in each hemisphere. Fiber bundles were selected from whole-brain streamline tractography using a multiple region-of-interest approach guided by anatomical references [41] [260]. For each bundle, three masks (two inclusion, one exclusion) were manually delineated in the population specific atlas with ITK-SNAP [252] and then deformed to each subject. Tractography was performed with the following methods: nearest-neighbor interpolation, rank-based estimation with fixed selection, linear estimation with fixed selection, linear estimation with max selection, and our proposed adaptive

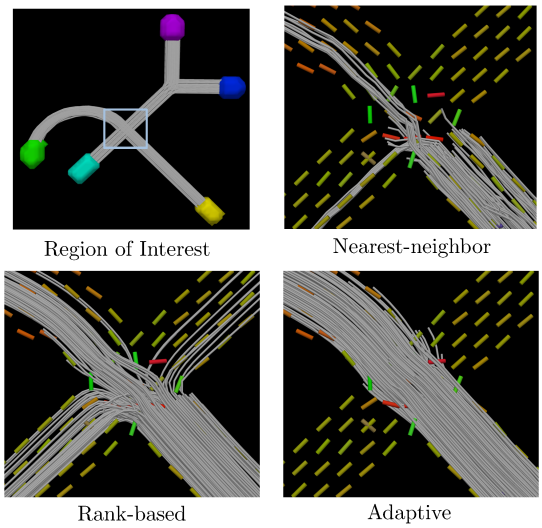


Figure 6.6: Results from the second experiment in Sec. 6.4.3 The first panel shows the phantom and region-of-interest, and the following three panels show tractography using nearest-neighbor, rank-based, and adaptive estimation (SNR = 14). We see that nearest neighbor includes early terminations due to high orientation error and compartment count error. We also compare to the rank-based approach, which creates many spurious connections due to poor matching of compartments in the crossing region. Adaptive estimation can improve both of these issues through smoothing and clustering-based compartment matching.

kernel-based estimation. Tracking parameters included two seeds per voxel, and angle threshold 45°, step size 1.0 mm, minimum volume fraction 0.1, and minimum length 10 mm.

We computed four metrics for each bundle: the mean length, mean volume fraction, total volume, and streamline count. Reproducibility and reliability were measured with the coefficient of variation (CV) and the intra-class correlation coefficient (ICC). The CV was measured by σ_s/μ_s and averaged across subjects, given the within-subject mean μ_s and standard deviation σ_s. A lower CV score indicates higher reproducibility and has units that are normalized to allow comparison across bundle metrics. The ICC was measured by $\sigma_b^2/(\sigma_b^2 + \sigma_w^2)$, given the between-subjects variance σ_b^2 and within-subjects variance σ_w^2. A larger ICC indicates more variance between subjects than within subjects. This takes a maximum value of one, and values above 0.75 indicate high reliability. The implementation used the R 'ICC' package [245].

Results: Results are shown in Fig. 6.9. We found that linear and adaptive kernel-based estimation generally had lower or roughly equal scan-rescan error (CV) compared to other methods. With adaptive estimation, bundle length had a CV of 3.56%, streamline count had a CV of 10.37%, total volume had a CV of 8.01%, and mean volume fraction had a CV of 2.3%. We found that linear and

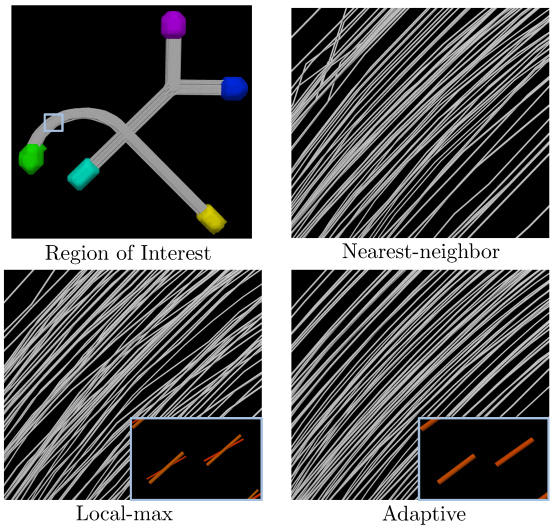


Figure 6.7: Results from the second experiment in Sec. 6.4.3. The first panel shows the phantom and region-of-interest, and the following three panels show tractography results using nearest neighbor interpolation, local-max model selection, and adaptive model selection (Eq. 6.8). This shows how local-max selection can sometimes introduce "fiber splitting" due to noise (SNR 16.5 shown). Adaptive selection can reduce this effect, resulting in smoother streamline curves and lower orientation and volume fraction errors (see Fig. 6.8).

adaptive estimation had the highest reliability (ICC) in all cases. With adaptive estimation, total volume had an ICC of 0.88, fiber count had an ICC of 0.88, bundle length had an ICC of 0.84,and mean volume fraction had an ICC of 0.72. Mean volume fraction had good and similar performance across all methods. We also found that given the same number of seed points, linear and adaptive estimation both had a much greater number of surviving streamlines than either nearest-neighbor or rank-based estimation.

6.4.5 *In vivo* Data Experiment for Atlas Construction

Design: The fourth experiment examined the construction of a multi-fiber tractography atlas with the 80 subject population. For this, we used the deformation fields computed with DTI-TK to resample the multi-fiber models to diffusion atlas space. Interpolation was performed using our approach, and fibers were reoriented by the local Jacobian and normalized to unit length, i.e. $\mathbf{J}v/\|\mathbf{J}v\|$ given Jacobian $\mathbf{J}$ and fiber orientation v. Multi-fiber fusion was then performed with our method to produce an average volume to represent the population. For comparison, standard single tensor tractography was performed in the diffusion tensor atlas with a minimum fractional anisotropy of 0.15. The single tensor atlas was created with the deformable registration algorithm in DTI-TK

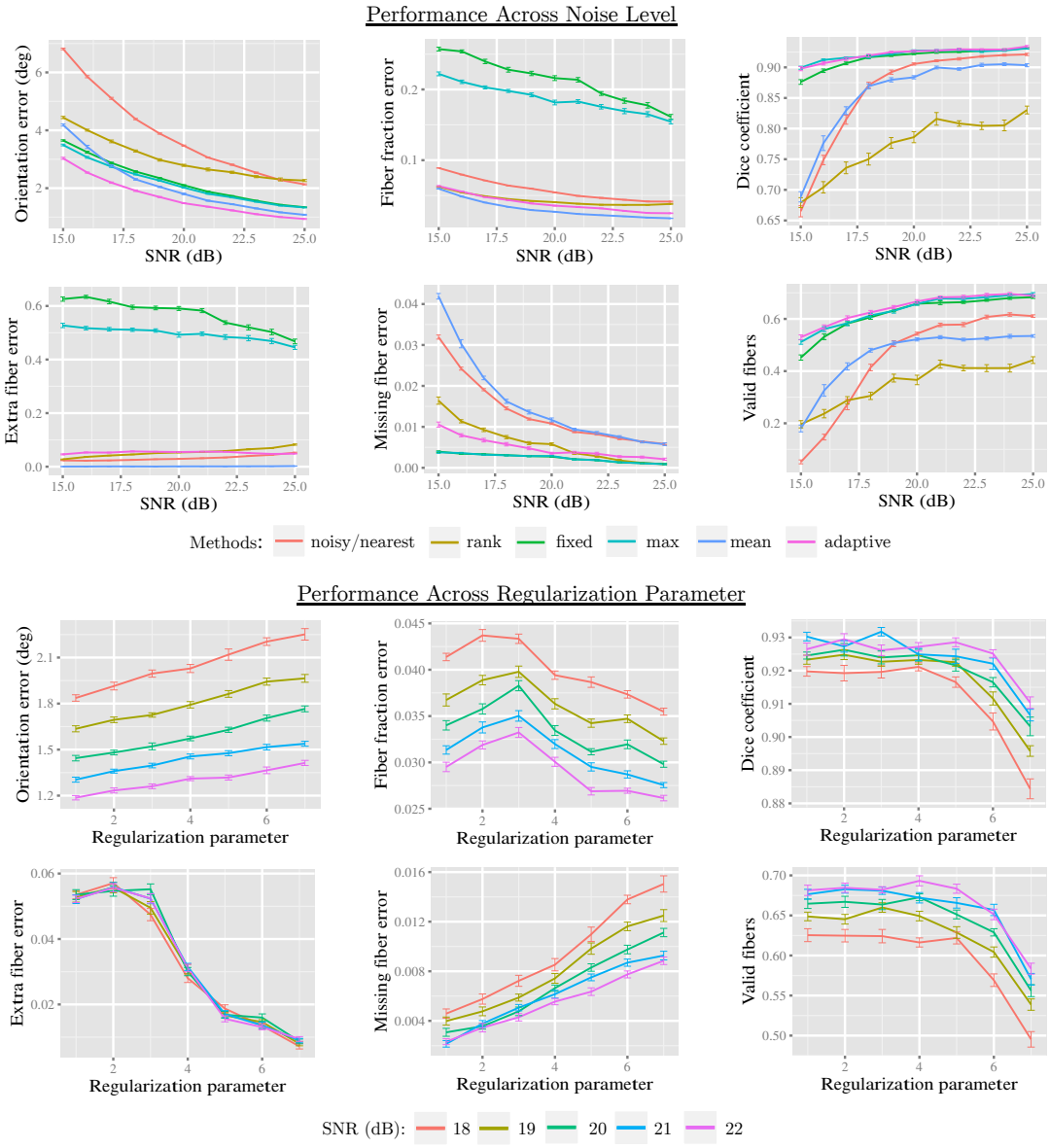


Figure 6.8: Results from the second experiment in Sec. 6.4.3 using the phantom shown in Fig. 6.6. Evaluation was first performed by measuring voxel-wise orientation error, volume fraction error, missing fiber error, extra fiber error (lower is better). Tractography-based evaluation was performed to evaluate performance in connectivity mapping using the Dice coefficient and fraction of valid connections (higher is better). Several methods for model selection were compared included: fixed count, local-max selection, mean selection, and our proposed adaptive selection. We also included rank-based estimation and nearest neighbor interpolation to serve as a baseline. The top two rows show performance across noise levels from SNR 15 to 25. We see that adaptive estimation provided significantly lower orientation error than other methods. Across all noise levels, fixed and local-max selection introduced a high volume fraction error and extra fiber error, due to "fiber-splitting". Adaptive performance significantly reduced these errors with a slight increase in missing fiber error. We also varied the regularization parameter and plotted with semilog axes, $x = log_{10}(1 - \lambda)$. We found performance to be acceptable between $\lambda = 0.99$ and 0.9999, with a tradeoff between extra and missing fiber error.

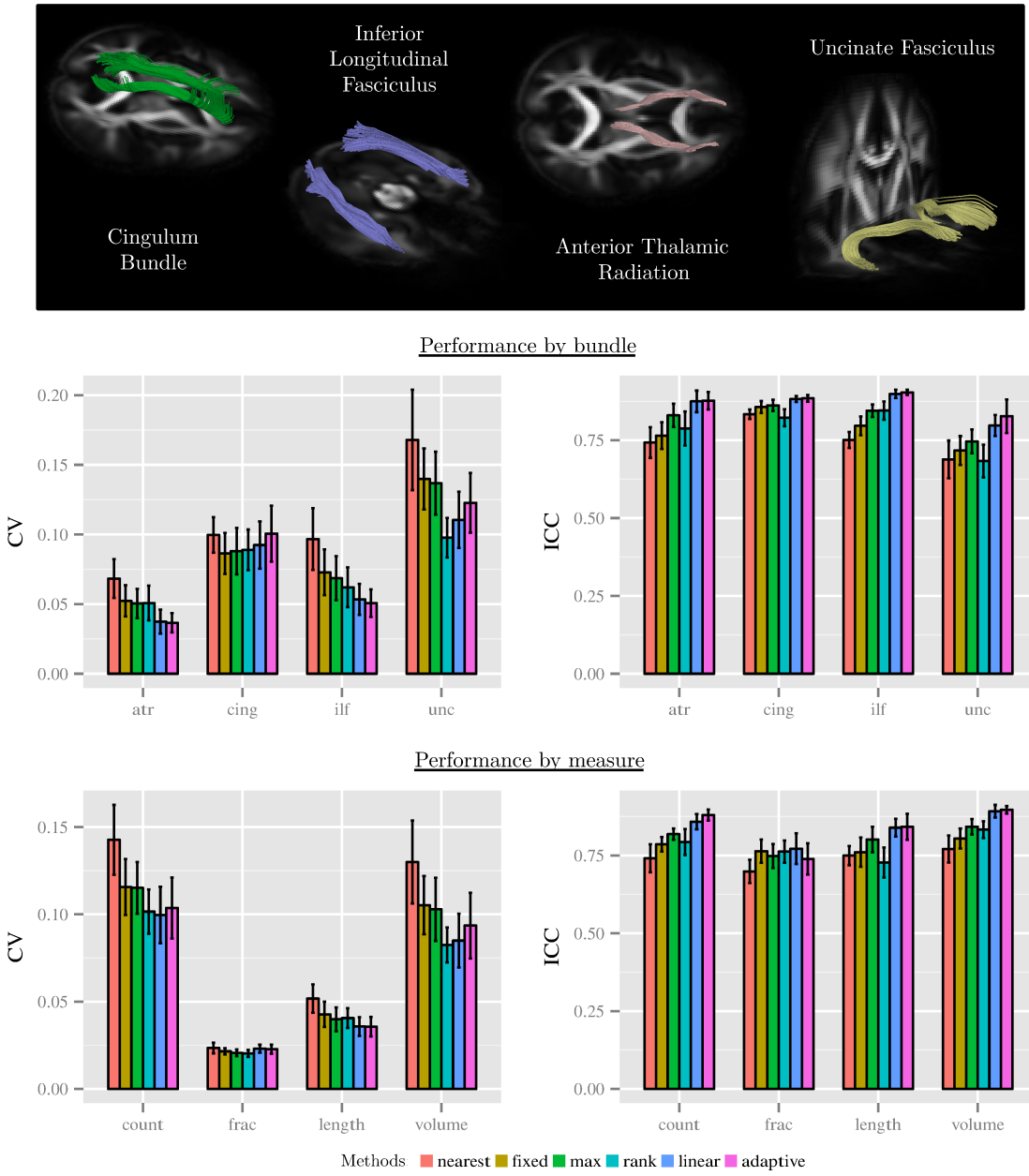


Figure 6.9: Quantitative results from the third experiment described in Sec. 6.4.4. Scan-rescan reproducibility and reliability were measured for fiber bundle metrics of eight fiber bundles. Reproducibility was measured with the coefficient of variation (CV), which gives a normalized measure of error across scans. Reliability was measured with the intra-class correlation (ICC), which indicates the proportion of total variation that exists between subjects. A lower CV and a high ICC are preferable, with an ICC above 0.75 being highly reliable. The top row shows reference visualizations of the bundles. The middle row shows results by bundle type, with aggregation across bundle measures and hemisphere. The bottom row shows results by bundle measure, with aggregation across bundle type. We found linear and adaptive estimation to show an improvement in nearly all cases by the ICC metric and improvement in most cases by the CV metric.

using finite strain tensor reorientation and the deviatoric tensor similarity metric. [256] [253]. In both atlases, major fiber bundles were manually delineated from whole brain tractography. For the multi-fiber atlas, the corpus callosum and superior longitudinal fasciculus I, I, and III were interactively tracked and manually delineated with guidance from atlas-space averaged Freesurfer gray matter labels. The lateral projections of the corpus callosum were selected based on left and right lobular regions. The superior longitudinal fasciculus I, II, and III were delineated by selecting connections between parietal cortex and superior, middle, and inferior subdivisions of frontal cortex, respectively.

Results: Visualizations of the results are shown in Figs. 6.10 and 6.11. We found the multi-fiber atlas to include nearly all features of the major bundles in the single tensor atlas. In addition, the multi-fiber atlas included more complete anatomical features of several bundles. In particular, the arcuate fasciculus included projections to inferior frontal gray matter, and the corpus callosum included lateral projections connecting the left and right hemispheres of frontal gray matter. We also found reconstructions of the three portions of the superior longitudinal fasciculus [55], which included crossings with numerous other bundles. We found the superior longitudinal fasciculus I crossed the corona radiata and superior projections of the corpus callosum. The superior longitudinal fasciculus II was found to cross the frontal lateral projections of the corpus callosum. Slice-based visualizations in Fig. 6.11 show examples of other crossings found in the multi-fiber atlas, including an axial slice of the brainstem and a sagittal slice of the corona radiata.

6.5 Discussion

In two synthetic data experiments, we evaluated the major features of our approach, including compartment matching, model selection, and bilateral filtering. To assess performance in compartment matching, we compared our method to the "rank-based" method, which matches solely based on volume fractions. The first experiment showed how our approach can avoid large errors that occur with the rank-based estimator when crossing bundles have similar volume fraction. The second experiment showed how this problem arises during tractography, and how clustering-based matching can avoid the invalid connections of the rank-based approach. We also showed how orientation blurring can occur at bundle boundaries when using kernel weights derived only from spatial information. The first experiment showed how the bilateral filtering extension can reduce this effect

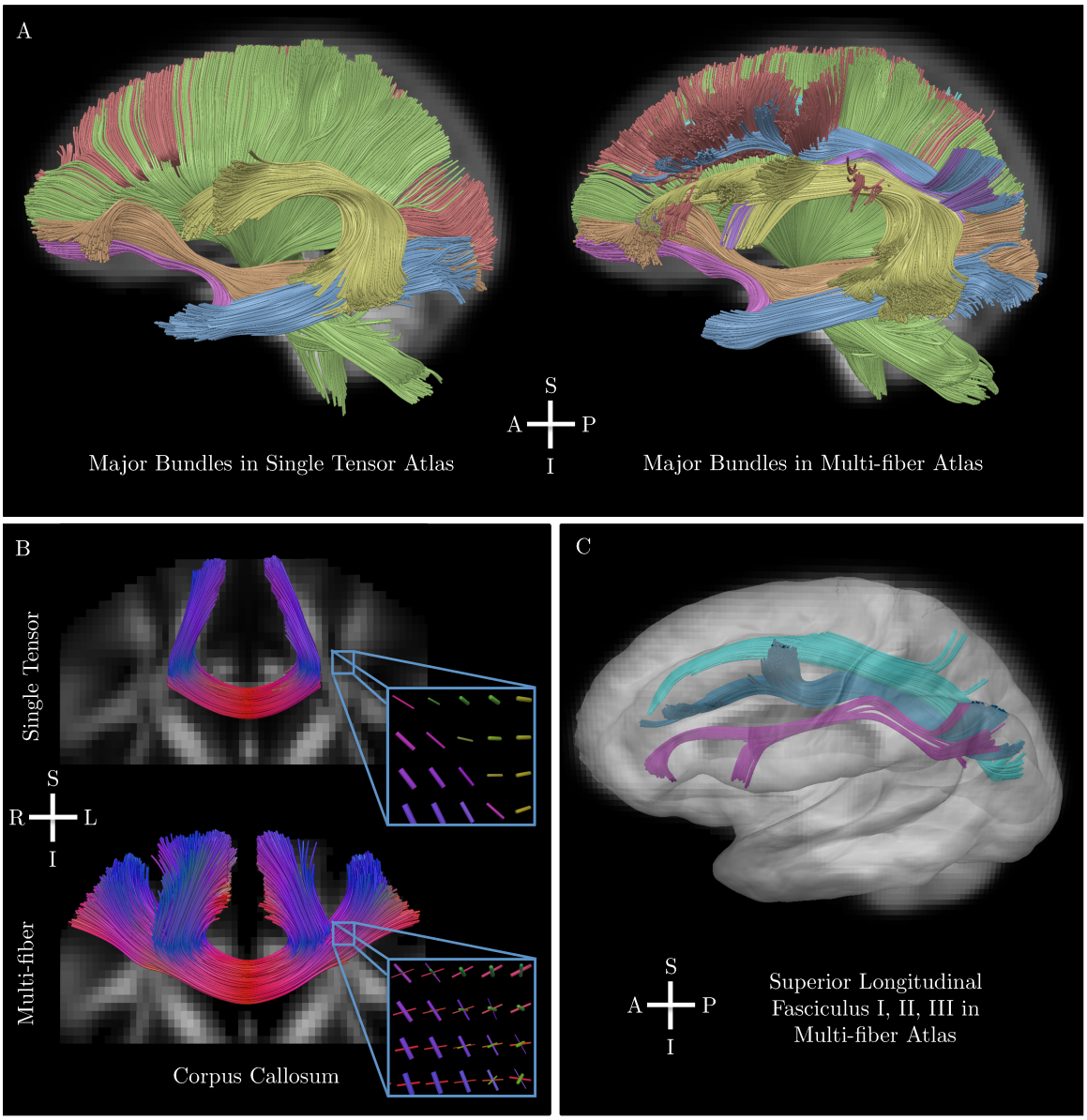


Figure 6.10: Results from the fourth experiment described in Sec. 6.4.5, which included the construction of a multi-fiber atlas of 80 normal human subjects. Our proposed method was used for interpolation and fusion with deformation fields computed by DTI-TK. The top panel (A) shows a comparison of major bundles in the standard single-tensor atlas and our proposed the multi-fiber atlas. Bundles include the corona radiata (green), inferior longitudinal fasciculus (blue), inferior fronto-occipital fasciculus (orange), uncinate fasciculus (pink), corpus callosum (red), and the arcuate fasciculus (yellow). We found the multi-fiber atlas included nearly all features found in the single tensor atlas and more complete reconstructions of the frontal projections of the arcuate and lateral projections of the corpus callosum. The bottom panels show tractography results for complex fiber bundles including the lateral projections of the corpus callosum (B) and fronto-parietal connections of the superior longitudinal fasciculus I, II, III (C). The detail view in panel B also shows a triple crossing inferred at the intersection of the corona radiata, superior longitudinal fasciculus, and corpus callosum. The three portions of the superior longitudinal fasciculus are also shown in panels A and C and include crossings with the corona radiata and corpus callosum. These reconstructions compare favorably to related work that examined these bundles in single human diffusion MRI reconstructions, dissection, and tracing studies in non-human primates.

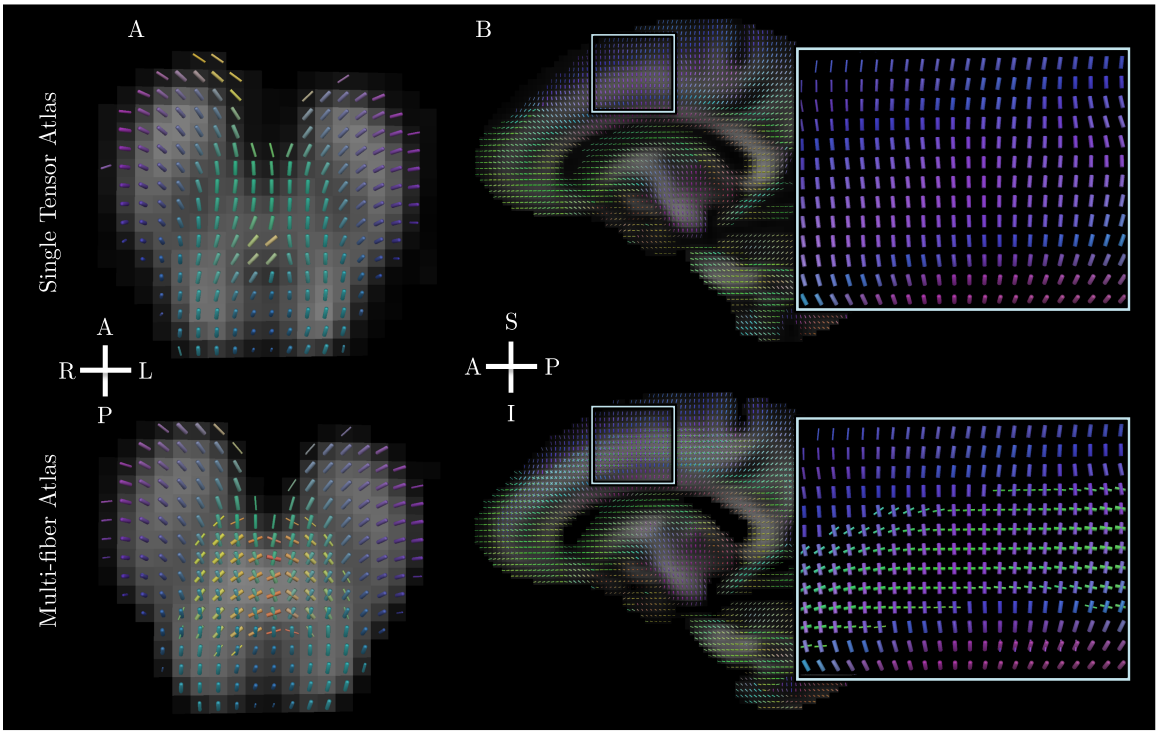


Figure 6.11: Results from the fourth experiment described in Sec. 6.4.5, which included the construction of a multi-fiber atlas of 80 normal human subjects. The top row shows principal tensor orientations overlaid on fractional anisotropy, and the bottom row shows fibers from the multi-fiber atlas overlaid on total fiber volume fraction. Fiber tube thickness reflects fractional anisotropy and per-fiber volume fraction in the top and bottom rows, respectively. The left column (A) shows an axial slice of the brainstem, demonstrating crossing fibers of cerebellar and pyramidal tracts. The right column (B) shows a sagittal slice demonstrating crossing fibers of the corona radiata and superior longitudinal fasciculus I.

by incorporating the structure of the data into the kernel weights. We also evaluated the proposed data-adaptive approach for model selection and compared it with previously proposed alternatives. We found "max" selection to always out-perform "fixed" selection and good performance in some cases; however, we found that both fixed and max selection can overestimate the number of fibers in even low noise conditions. These extra fibers give rise to compartment splitting, which both reduces the expected volume fraction and introduces orientation errors that are visible in tractography. We found the data-adaptive approach to reduce this effect, with only a small increase in missing fiber error. In general, these synthetic data experiments also showed our approach to significantly reduce noise-induced errors in fiber orientation and volume fraction and to improve connectivity mapping.

In the first *in vivo* data experiment, we examined the practical benefit of our approach for fiber

bundle modeling in individual subjects. We found our approach provided reliable and reproducible results for quantitative fiber bundle metrics, including mean length, mean volume fraction, total volume, and streamline count. Among the measures, mean length and mean volume fraction performed best. Streamline count and total volume are perhaps less reliable due to their dependence on the pose and sampling resolution of the voxel grid. While past work has evaluated length-based measures [50], it is less clear how volume fraction can be used as an index of white matter [90] [94]. The results provide evidence that fiber bundle volume fraction may be a clinically useful measure in terms of reproducibility; however, it remains unclear exactly how this measure relates to others, such as fractional anisotropy, fiber orientation dispersion, and apparent fiber density [178] [182]. The scan-rescan results can also serve as a reference for gauging the significance of group differences of fiber bundle metrics in clinical imaging studies.

In the second *in vivo* data experiment, we examined an application to multi-fiber atlas construction, using our approach for interpolation during image resampling and fusion. We found the resulting atlas to include complex fiber bundle features not found in single tensor atlases [256]. In particular, the results included atlas-based reconstructions of the superior longitudinal fasciculus I, II, and III, all of which compare favorably to prior work on single subject reconstructions, post-mortem dissection, and tracing studies in non-human primates [55] [56] [187]. This method is potentially useful for constructing population-specific brain atlases, where it can be used to examine population-wide features of anatomy or serve as a reference for mapping bundles in individual subjects.

The software implementation consisted of custom Java code with experiments run on a Sun Grid Engine to allow for parallelized processing of simulation and human subject data. The fusion algorithm ran on a single Intel 2.8 GHz Core i5 processor and took 30 minutes for the 80 subject population. The primary computational bottleneck of this step was main memory usage, which required 6GB total. If applied to a larger population or with a higher sampling resolution, this issue could be addressed by decomposing the co-registered volumes into a number of overlapping blocks, performing fusion of each block, and then reconstructing the full atlas from the results.

There are a number of limitations and interesting open issues to note. First, we make use of tensor-based registration; however, some work has found multi-compartment registration to be beneficial [207] [206] [63]. Note, however, that d_m is more of a divergence than a strict distance as it is not symmetric, so this may not be suitable for registration, where inverse-consistency is

desirable. Second, it may be beneficial to investigate applications to other parametric diffusion models. The bilateral filtering and adaptive model selection techniques could be similarly applied to the full multi-tensor framework by substitution of d_f. The kernel regression framework could also be applied to models that incorporate microstructural information, such as NODDI [254] and ball-and-rackets [198]. This could be accomplished either by extending the proposed divergence measure to include variable Watson κ parameters or perhaps by a more general formulation with a Bingham divergence. Finally, there is a variety of applications to which this approach could be applied and further evaluated. Here we looked at fiber bundle mapping, but this could also be useful for brain network construction and visualization for surgical planning and guidance.

The approach is also related to more general image processing with orientation data, which touch research areas outside of MRI. Early approaches in directional statistics examined methods for dealing with axial data, a term more common in the statistics literature that applies equally well to fiber orientations [234]. In computer vision and graphics, filtering [2] [149] [238] [220], anisotropic diffusion [167] and bilateral filtering [163] have been developed for orientation fields, e.g. in analyzing fingerprints, hair, or textures. In the diffusion MRI literature, methods for fiber orientation regularization have also been proposed using Markov random fields [172], variational methods [52], and cohelicity [185]. The present work may have applications to those other areas, but this chapter is limited to applications to neuroimaging with diffusion MRI and the ball-and-sticks model.

6.6 Conclusion

In this chapter, we presented and evaluated a model-based kernel regression framework for estimating fiber orientation mixtures from model-valued image data. This framework generally supports image interpolation, smoothing, and fusion with the ball-and-sticks diffusion model and is compatible with scanner protocols with standard clinical field strengths, single shells, and low b-values. The kernel regression formulation also allows for simple and efficient data-adaptive extensions for model selection and bilateral filtering, and these general design of these components may potentially be applied to image processing with other multi-compartment parametric models. We experimentally evaluated our approach with synthetic data from computational phantoms and *in vivo* clinical data from human subjects. First, we showed our approach can address issues related to compartment

matching and model selection, and can avoid orientation blurring at bundle boundaries and fiber splitting due to noise. The evaluation showed significant reductions in noise-induced orientation and volume fraction errors and improved reliability in connectivity mapping. Through *in vivo* data experiments, we tested the practical value of our approach in analyzing individual subjects, showing improved scan-rescan reproducibility and reliability of quantitative fiber bundle metrics. We then demonstrated the creation of a multi-fiber tractography atlas from a population of 80 human subjects. In comparison to single tensor atlasing, our results show more complete features of known fiber bundles and include reconstructions of lateral projections of the corpus callosum and complex fronto-parietal connections of the superior longitudinal fasciculus. This atlas-based approach could be used to either study population-wide features of anatomy or to aid fiber bundle mapping in individual subjects.

6.7 Appendix

In the above analysis, we employ $d_f^2(a,b)$ (Eq. 6.7) as a measure of discrepancy between fiber orientations. This turns out to have some useful geometric and statistical properties. For example, given an orientation v, consider the mapping $\phi(v) = vv^T$, also known as the Veronese-Whitney embedding, the dyadic product, or Knutsson mapping [107] [20] [181]. This can be imagined to take points on the sphere to a higher dimensional sphere and to fold its antipodes to give $\phi(v) = \phi(-v)$. The Euclidean distance under this mapping is then equivalent to d_f^2 up to a scalar factor: $\|\phi(a) - \phi(b)\|^2 = 2(1 - (a \cdot b)^2) = 2d_f^2(a,b)$ The extrinsic mean is then given by $\mu = \sum_i w_i \phi(v_i)$, which must be projected to the nearest orientation by $\text{argmin}_v \|\phi(v) - \mu\|^2$. This can be solved in closed form by finding the principal eigenvector of μ [20]. This formulation also has a statistical interpretation, as d_f^2 is equivalent to the Bregman divergence between Watson distributions of equal dispersion [68] [37], and the weighted average is equivalent to the maximum likelihood estimate of the direction of a Watson distribution [200]. These results support prior work suggesting a relationship between extrinsic means and maximum likelihood estimates for data in $\mathbb{RP}^n$ [32].

A connection can also be drawn between d_f and prior work using Gaussian mixture simplification with the Burg matrix divergence, $d_{burg}(A,B) = tr(AB^{-1}) - log(det(AB^{-1})) - 3$. While d_{burg} cannot be directly applied to the sticks due to their low rank, it can be applied to sticks that are approximated by an "inflated" positive definite tensor with δ added to the second and third

eigenvalues. If this approach is applied to a pair of stick orientations a and b to produce inflated tensors A_{inf} and B_{inf}, then d_{burg} can be reduced to $d_{burg}(A_{inf}, B_{inf}) = (\delta + 1/\delta - 2)\sin^2(\theta)$, given angle θ between a and b. If $\delta = (3-\sqrt{5})/2$, then $d_{burg}(A_{inf}, B_{inf}) = d_f^2(a, b)$, so the use of Gaussian mixture simplification with inflated sticks is then equivalent to the "linear" and "max" conditions in the experiments. We ran a simulation experiment to numerically verify this and also found the "inflated stick" Burg divergence implementation to take 40 times longer on average.

For multi-fiber analysis, we employ d_m^2 (Eq. 6.6) to measure discrepancy between mixtures of fiber orientations, and this too has useful geometric and statistical features. The form of d_m allows the kernel regression estimator in Eq. 6.8 to be algebraically reduced to the simpler form in Eq. 6.12, which is equivalent to a clustering problem. The first sum-of-squares term has a statistical interpretation as hard clustering with a mixture of Watsons [199], and the additional regularization term has a statistical interpretation as the small-variance asymptotic limit of a Dirichlet process mixture, as in the DP-means algorithm [108]. Taken together, these make up the axial DP-means clustering problem in Eq. 6.12. This is theoretically well-grounded because the Watson distribution is an exponential family, and d_f is the associated Bregman divergence [37]. The work of Jiang *et al.* also demonstrates that the general form of Algo. 2 converges, albeit to a local minima [95]. To avoid instability due to local minima, we use random restarts with shuffling of the order of the input and take the solution with the overall minimum of Eq. 6.12.

Algorithm 2: axial DP-means clustering

Input:
$\{(w_l, \mathbf{v}_l)\}_{l=1}^{L}$: the input weights and orientations
λ: the regularization parameter
K_{max}: a maximum number of clusters

Output:
$\hat{K}$: estimated number of clusters
$\{(\hat{w}_k, \hat{\mathbf{v}}_k)\}_{k=1}^{\hat{K}}$: estimated cluster weights and orientations

`/* Initialize */`
$\hat{K} \leftarrow 1$, $\hat{\mathbf{v}}_1 \leftarrow \mathit{prineig}\left(\sum_l \mathbf{v}_l \mathbf{v}_l^T / L\right)$
while *not converged* **do**
  `/* Assign cluster labels */`
  for *l=1* **to** L **do**
    for *k=1* **to** $\hat{K}$ **do**
      $D_{lk} \leftarrow \left(1 - \left(\mathbf{v}_l^T \hat{\mathbf{v}}_k\right)^2\right)$
    `/* Optionally increment clusters */`
    if $\min_k D_{lk} > \lambda$ & $\hat{K} < K_{max}$ **then**
      $\hat{K} \leftarrow \hat{K} + 1$, $\hat{\mathbf{v}}_K \leftarrow \mathbf{v}_l$, $\pi_l \leftarrow \hat{K}$
    else
      $\pi_l \leftarrow \mathrm{argmin}_k D_{lk}$
  `/* Update cluster centers */`
  for *k=1* **to** $\hat{K}$ **do**
    $\hat{w}_k \leftarrow \sum_l w_l \delta(k, \pi_l)$
    $\hat{\mathbf{v}}_k \leftarrow \mathit{prineig}\left(\left(\sum_l w_l \delta(k, \pi_l) \mathbf{v}_l \mathbf{v}_l^T\right) / \sum_l \delta(k, \pi_l)\right)$
return

Algorithm 3: fiber orientation mixture estimation

Input:
p_0: the input spatial position for estimation
M_0: the input reference model
$\{(p_i, M_i)\}_{i=1}^{C}$: the local neighborhood of C position/model pairs
N_i: the number of fibers in the i-th local model
f_{ij}: the volume fraction of the j-th fiber of the i-th local model
$\mathbf{v}_{ij}$: the fiber orientation of the j-th fiber of the i-th local model
h_{pos}: the spatial positional bandwidth
h_{dir}: the bilateral data-adaptive bandwidth
λ: the model complexity regularization parameter
K_{max}: the maximum number of fiber compartments

Output:
$\hat{K}$: estimated number of fiber compartments
$\hat{M} = \{(\hat{f}_k, \hat{\mathbf{v}}_k)\}_{k=1}^{\hat{K}}$: estimated model

`/* Compute sum-normalized model weights */`
$K_{sum} \leftarrow 0$
for *i=1* **to** C **do**
  $k_i \leftarrow K(-\|p_0 - p_i\|^2 / h_{pos}^2)$
  $k_i \leftarrow k_i * K(-d_m^2(M_0, M_i)/h_{dir}^2)$
  $K_{sum} \leftarrow K_{sum} + k_i$
for *i=1* **to** C **do**
  $k_i \leftarrow k_i / K_{sum}$
`/* Compute fiber weights */`
$L \leftarrow 0$, $F_{sum} \leftarrow 0$
for *i=1* **to** C **do**
  for *j=1* **to** N_i **do**
    $w_L \leftarrow k_i * f_{ij}$, $\mathbf{v}_L \leftarrow \mathbf{v}_{ij}$
    $L \leftarrow L + 1$, $F_{sum} \leftarrow F_{sum} + f_{ij}$
`/* Optimize by clustering with Algo. 2 */`
$\hat{K}, \{(\hat{w}_k, \hat{\mathbf{v}}_k)\}_{k=1}^{\hat{K}} \leftarrow cluster(\{(w_i, \mathbf{v}_i)\}_{i=1}^{L}, \lambda, K_{max})$
for *k=1* **to** $\hat{K}$ **do**
  $\hat{f}_k \leftarrow \hat{w}_k * F_{sum}$
return

Chapter 7

An Evaluation of Diffusion MR Tractography-based Reconstruction of the Corticospinal Tract

In this chapter, we evaluate a method for reconstructing the corticospinal tract (CST) using deterministic diffusion MR tractography and a model-based framework for processing parametric maps obtained from multi-fiber diffusion models. The reconstruction of the CST is a clinically relevant problem, as an accurate geometric reconstruction can be used in surgical planning to reduce the risk of unnecessary damage to motor pathways [47]. Diffusion tractography provides a tool for extracting patient-specific geometric models of fiber bundles; however, several challenges exist for making this a clinically useful resource for CST reconstruction. First, the anatomy of the CST includes crossings with the superior longitudinal fasciculus and corpus callosum [105]. This poses a challenge for diffusion models that do not represent complex fiber configurations, such as the single tensor model. Multi-fiber models provide an alternative that theoretically addresses this issue [175], although these can be more difficult to fit numerically and more susceptible to noise due to the increased model complexity. A second challenge is the validation for clinical application, which includes establishing

the accuracy of the reconstruction with respect to anatomy and the practical value to a clinician.

The main contribution is the application of a framework for multi-fiber model-based processing to the reconstruction of the CST with eXtended Streamline Tractography (XST) [175]. Model-based approaches may offer computational and theoretical advantages over image processing performed with the diffusion-weighted signal. A number of model-based approaches exist for diffusion imaging, including the Riemannian frameworks for tensors [165], orientation distribution functions [73], and multi-tensor models [205] [207]. In our approach, we apply recent work on model-based processing of fiber orientation mixtures [34], particularly the ball-and-sticks diffusion model [19].

In the rest of the chapter, we first describe our method, which includes diffusion modeling with the ball-and-sticks model and a model-based framework for performing smoothing and interpolation during deterministic streamline tracking. We then conduct experiments testing the quality of CST reconstruction with our approach, which uses regions of interest obtained from deformable atlas registration to select fiber bundles. We apply our approach to three clinical brain datasets including normal bundles and a bundles infiltrated by tumors, showing visualizations of fiber bundle anatomy along with volumetric and surface representations of other brain anatomy. We conclude with a discussion of our results and open problems.

7.1 Methods

In the following sections, we review the diffusion models used in our experiments, describe our tractography algorithm, and present our approach for model-based processing.

7.1.1 Diffusion Models

In our analysis, we apply diffusion models supporting both single and multiple fiber per voxel. For single fiber analysis, we use the standard diffusion tensor model, which has the following predicted signal S_i of the i-th diffusion-weighted image (DWI):

$$S_i = S_0 \exp\left(-b_i \vec{g}_i^T \mathbf{D} \vec{g}_i\right) \tag{7.1}$$

given gradient encoding direction $\vec{g}_i$, b-value b_i, unweighted signal S_0, and diffusion tensor $\mathbf{D}$. We

also use the multi-fiber ball-and-sticks model, which theoretically can resolve more complex fiber configurations than the single tensor. Under this model, the predicted signal S_i of the i-th DWI is as follows:

$$S_i = S_0 \sum_{j=0}^{N} f_j \exp\left(-b_i \vec{g}_i^T \mathbf{D}_j \vec{g}_i\right) \tag{7.2}$$

given N fiber compartments, fiber volume fraction $0 \leq f_j \leq 1$, and $\sum_{j=0}^{M} f_k = 1$. The compartments are constrained to include a completely isotropic first component $\mathbf{D}_0 = \text{diag}(d, d, d)$ and completely anisotropic subsequent components $\mathbf{D}_j = d\vec{v}_j\vec{v}_j^T$ with diffusivity d and fiber orientation $\vec{v}_j$.

7.1.2 Tractography

We performed tractography using a generalization of the standard streamline approach [16] to account for multiple fibers. In the standard streamline approach, a fiber trajectory is considered a 3D space curve whose tangent vector is equated with the fiber orientation of the voxelwise diffusion models. This can be found by evolving a solution to a differential equation with some initial condition at a given seed position. Typically, some geometric criteria are also used to stop and exclude fibers, including angle threshold and minimum and maximum length. When multiple orientations are present, this approach must be adapted in several ways. First, additional volume fraction termination criteria are included, from the f_i parameters in Eq. 7.2. During tracking, one of the N possible fiber orientations must also be chosen for the next step. Our approach is similar to eXtended Streamline Tractography (XST) [175], although we instead use volume fraction for termination and allow selection of any fiber compartment below a given angle threshold, as opposed to the one with the smallest angle. We also use a model-based framework, which we describe next.

7.1.3 Model-based processing framework

For the purpose of tractography, it's important to perform accurate interpolation to sample diffusion models off the voxel grid and smoothing to reduce the effects of noise and sampling artifacts. For this purpose, we use a model-based framework for processing multi-fiber models. The first work in this area examined full multi-tensor models [205], and we perform a similar analysis derived for the simpler ball-and-sticks model, as follows [34]. For the sake of generality, we then consider a

model M to be a fiber orientation mixture consisting of N fiber volume-fraction and orientation pairs $M = \{(f_i, v_i)\}_{i=1}^{N}$ that lie in each voxel. Next, we aim to construct a least squares estimator for fiber models [204] that will provide a way to perform both smoothing and interpolation.

Given an input position p and local neighborhood $\{(p_i, M_i)\}_{i=1}^{C}$ with $M_i = \{(f_{ij}, v_{ij})\}_{j=1}^{N_i}$, the estimator for $\hat{M}$ is given by:

$$\hat{M}(p_0) = \underset{M}{\operatorname{argmin}} \sum_{i=1}^{C} K\left(\frac{d_p^2(p_i, p_0)}{h_p^2}\right) K\left(\frac{d_m^2(M_i, M_0)}{h_m^2}\right) d_m^2(M_i, M) + \lambda N \tag{7.3}$$

given model distance d_m, spatial bandwidth h_p, data-adaptive bandwidth h_m, and kernel function $K(x) = exp(-x)$. We define the model distance between M and $\hat{M}$ by

$$d_m^2(M, \hat{M}) = \min_{\pi} \sum_{j} f_j d_f^2(v_j, \hat{v}_{\pi(j)}) \tag{7.4}$$

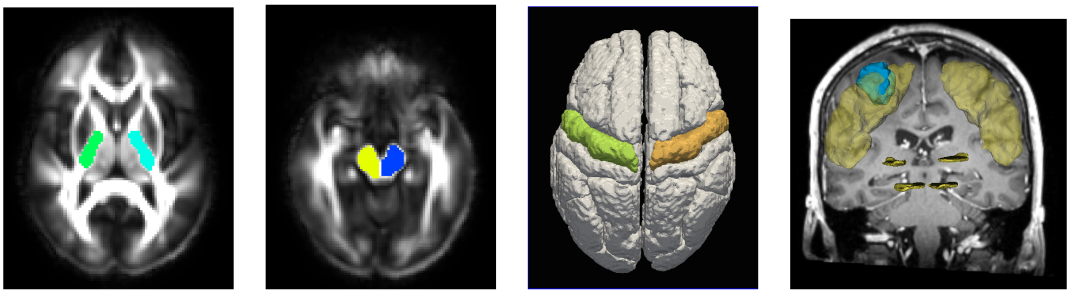
$$d_f^2(a, b) = 2(1 - (a \cdot b)^2) = 2sin^2(\theta) \tag{7.5}$$

which is selected across all possible mapping π from fibers in M to $\hat{M}$, with respect to the fiber distance d_f. This estimator can be used in practice by solving the minimization problem with a fiber clustering procedure described in Chapter 6.

7.1.4 Bundle Delineation

We used an atlas-based approach [143] to delineate the CST tract with three regions of interest. The diffusion tensor atlas was previously constructed using a population of 80 normal adult subjects [33]. The following three regions were then defined in each hemisphere of the atlas: the posterior limb of the internal capsule, cerebral peduncle, and the precentral brain matter. The first two regions were manually drawn in ITK-SNAP [1] and the precentral region was defined by the population average combined mask of the white and gray matter labels computed with Freesurfer version 5.1 [2]. Visualizations of these regions are shown in Fig. 7.1. The patient's diffusion tensor volume was then registered to the atlas using deformable tensor-based registration with DTI-TK [3] [255], and the labels were deformed to the patient image space. Tractography seeding was performed in all seed regions and only fibers that traversed all other regions were kept. We also performed

[1] http://www.itksnap.org/
[2] https://surfer.nmr.mgh.harvard.edu
[3] http://dti-tk.sourceforge.net



Posterior Limb of Internal Capsule | Cerebral Penduncle | Precentral Gyrus | Regions deformed to an example patient

Figure 7.1: Regions of interest used for delineating the CST. Atlas space regions are shown in axial slices of the posterior section of the internal capsule (first panel) and cerebral peduncle (second panel), and a surface rendering of the superior surface of the precentral brain matter (third panel). The regions were deformed to patient space (fourth panel) and used to seed and select fibers in the bundle.

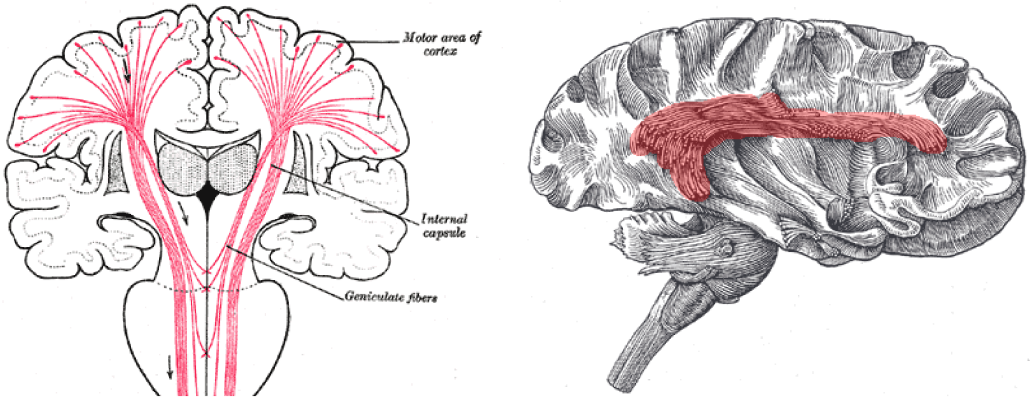


Figure 7.2: Illustrations of the cortico-spinal tract (left) and arcuate fasciculus (right). (from [75])

manual corrections to exclude outlier streamlines from the bundles, such as those that crossover to the contralateral side. In one patient, we also delineated the arcuate fasciculus, and this was done manually. Illustrations of the bundles are shown in Fig. 7.2

7.2 Experiments

In this section, we present results from two datasets comprising four human subjects. These were part of the MICCAI DTI Challenge [4] which has been running since 2011. The goal of this challenge is to validate diffusion MR tractography for neurosurgical applications and define standards for its use in clinical applications [173]. Each year, pre- and post-operative brain tumor imaging datasets

[4] http://dti-challenge.org

are distributed to a working group consisting of teams of computer scientists and neuroscientists, and each team analyzes the datasets to reconstruct the geometry of the cortico-spinal tract and other bundles. These results are then evaluated by a panel of judges, which includes neurosurgeons and other domain experts, according to three criteria: False-Negative tracts, False-Positive tracts and Clinical Relevance. In addition, an annual workshop related to the challenge is held at the Medical Image Computing and Computer Assisted Intervention (MICCAI) conference, where the working group meets to present their results and discuss open problems in the area.

7.2.1 Experimental Design

The results presented in this section were included in the 2014 [5] and 2015 [6] sessions of the challenge. In the 2014 challenge dataset, we compare several tractography methods: single tensor tractography, nearest-neighbor interpolated multi-fiber tractography, linear kernel interpolated multi-fiber tractography, and adaptive kernel interpolated multi-fiber tractography. In the 2015 challenge dataset, we present adaptive kernel interpolated results for three patients. For all patients, we show reconstructions of the cortico-spinal tract, and for the second patient we show reconstructions of the arcuate fasciculus.

7.2.2 Data acquisition for 2014 Challenge

Imaging data of a single patient was downloaded from XNAT [7] as part of the MICCAI 2014 DTI Challenge. The diffusion MRI scans were acquired with a spin-echo EPI sequence with the following parameters: voxel size 2.0 x 2.0 x 2.0 mm, 256 x 256 matrix, 73 slices, b-values 200, 500, 1000 and 3000 s/mm^2, 69 diffusion-weighted volumes and 4 non-diffusion weighted volumes. The patient presented with a metastatic adenocarcinoma infiltrating the corticospinal tract with large edema, both isolated to the right hemisphere.

7.2.3 Data acquisition for 2015 Challenge

Imaging data of three patients was downloaded from XNAT [8] as part of the MICCAI 2015 DTI Challenge. Patient 1 presented with meningioma, and imaging data was acquired with a spin-echo

[5] http://projects.iq.harvard.edu/dtichallenge14
[6] http://dti-challenge.org
[7] http://central.xnat.org
[8] http://central.xnat.org

EPI sequence with the following parameters: voxel size 2.0 x 2.0 x 2.0 mm, FOV 220mm, 58 slices, b-value 1000 s/mm^2, 30 gradient encoding directions, and one baseline. Patient 2 presented with a low-grade glioma, and imaging data was acquired with a spin-echo EPI sequence with the following parameters: voxel size 2.5 x 2.5 x 2.5 mm, FOV 220mm, 55 slices, b-value 1000 s/mm^2, 30 gradient encoding directions, and one baseline. Patient 3 presented with a low-grade glioma, and imaging data was acquired with a spin-echo EPI sequence with the following parameters: voxel size 2.5 x 2.5 x 2.5 mm, FOV 220mm, 60 slices, b-value 800 s/mm^2, 31 gradient encoding directions, and one baseline. In all cases, the tumor and anatomical references were manually segmented in the baseline volume.

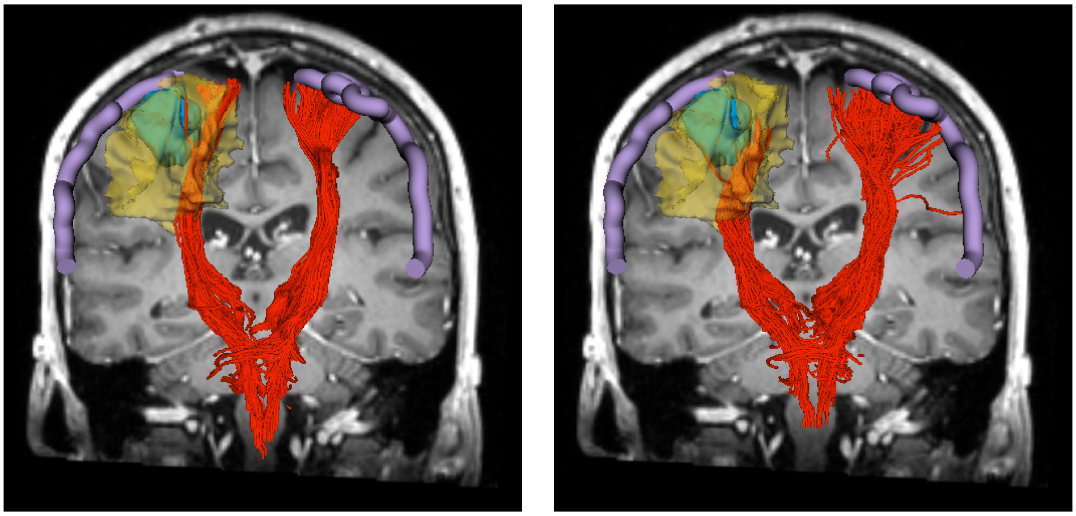
7.2.4 Image Analysis

For all three cases, preprocessing, voxel model fitting, tractography, and visualization were performed as follows. A brain mask was extracted with BET in FSL version 5.0 [9]. Single diffusion tensors were fit with DTIFIT in FSL. Multi-fiber models were fit using an MCMC procedure with XFIBRES in FSL with default parameters. Smoothing and interpolation were performed with custom software using $h_p = 3.0$ and a kernel support window of 5 voxels. Tractography was performed with custom software with the following parameters: angle threshold 65 degrees, step size 1.5 mm, 20 seeds per voxel, minimum length 20 mm, maximum length 110 mm, and volume fraction threshold 0.1. Culling was then performed to reduce the number of streamlines by iteratively removing any curve within an average of 1 mm from an existing curve. Bundle delineation was performed with custom software, and visualizations were rendered with custom software by generating streamtubes. Parameters were chosen by varying them in step-wise fashion.

7.2.5 Results from 2014 Challenge

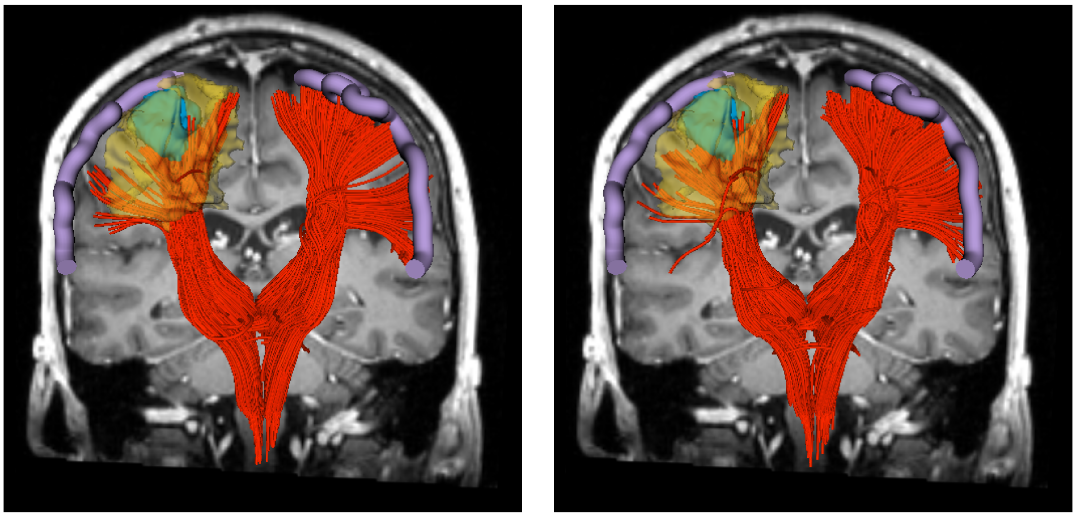
Visualizations in Fig. 7.3 show the reconstructed bundles from the 2014 dataset. In the left hemisphere, we found the single tensor analysis to be similar to other works [105] and included only the most superior projections. The ball-and-sticks reconstruction improved the tracking through crossing and fanning fibers also similar to previous works [175]. Linear filtering significantly increased the number of lateral projections. The filtering also produced smoother fibers that retained a similar overall shape in the main body of the bundle. The addition of the adaptive term in the filtering

[9] http://www.fmrib.ox.ac.uk/fsl



Standard
Single Tensor

Unfiltered
Ball and Sticks



Linear Filtered
Ball and Sticks

Adaptive Filtered
Ball and Sticks

Figure 7.3: A comparison of results from the 2014 dataset. Tractography (red) for standard single diffusion tensor modeling and ball-and-sticks modeling with and without filtering. Shown are results from standard diffusion tensor (top left), unfiltered ball-and-sticks (top right), linear filtered ball-and-sticks (bottom left), and adaptive filtered ball-and-sticks (bottom right). The outline of motor cortex is overlaid in purple. In the right hemisphere, we see the infiltrating tumor (blue) stops tracking in all models, and the edema (yellow) stops tracking to varying extents. In the left hemisphere, we see that single tensor modeling does not include the lateral projections. Linear filtering increased the number of lateral projections. Adaptive filtering further increased the lateral projections and reduced the number of outlier fibers.

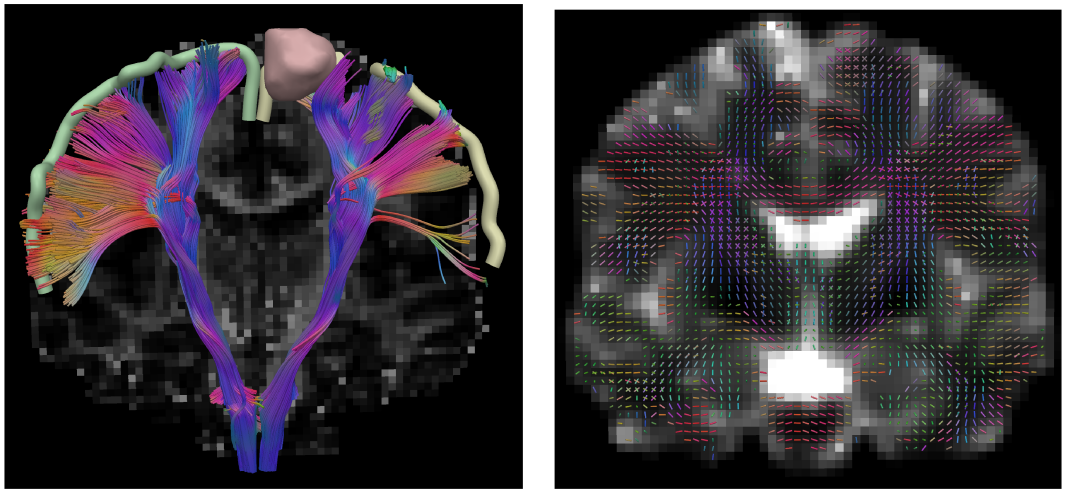


Figure 7.4: A frontal view of a coronal slice of the first patient in the 2015 dataset. This shows the CST (left) and volumetric fiber models (right). The tumor shows some amount of deflection of the most superior portion of the left bundle. Otherwise, much of the bundles could be reconstructed.

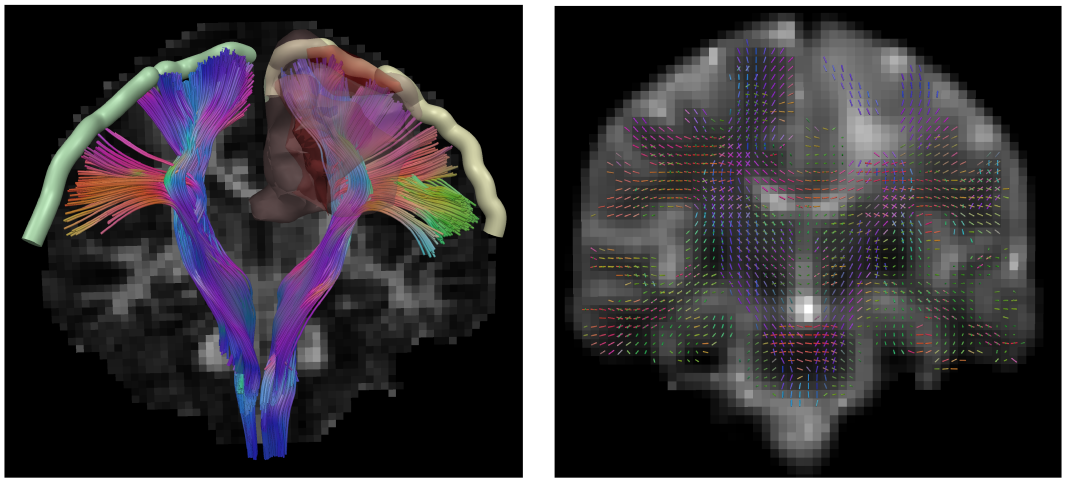


Figure 7.5: A frontal view of a coronal slice of the second patient in the 2015 dataset. This shows the CST (left) and volumetric fiber models (right). The core of the tumor is mostly anterior to the tract, with some surrounding portions overlapping the bundle. The tracking seemed to continue through this region, which may be edema. The fiber models show a slice anterior to the CST in which the core of the tumor causes a drop-out of fiber models.

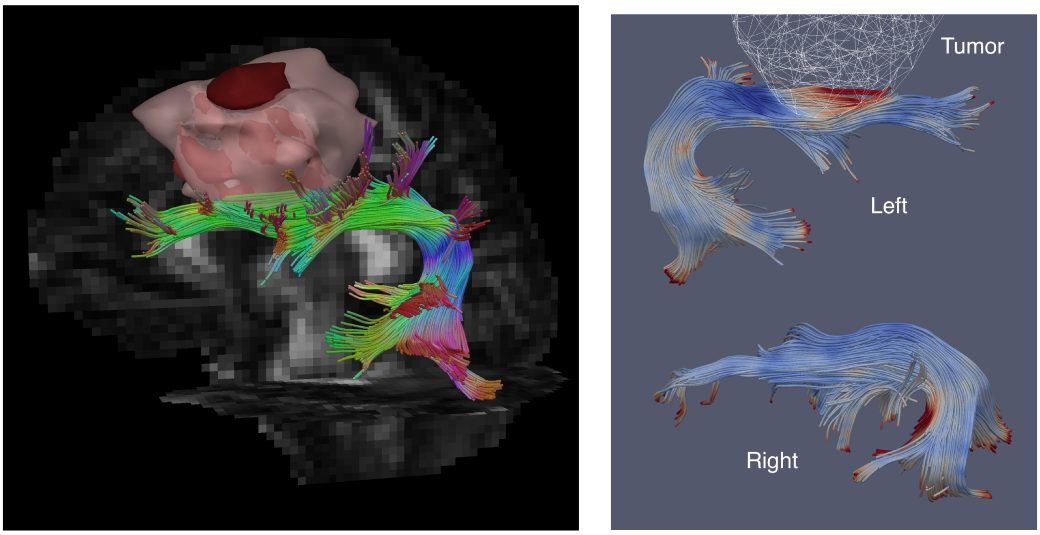


Figure 7.6: A lateral view (left) and medial views (right) of the arcuate fasciculus of the second patient in the 2015 dataset. The top shows the bundle curving below the tumor, but continuing onto frontal areas. The bottom shows the medial surface of the left and right arcuate bundles with coloring according to mean diffusivity (both used the same colormap) and an overlaid tumor wireframe on the left arcuate. This highlights the area inferior to the tumor, which has higher diffusivity (red) than would be expected from the homologous portion of the right bundle (blue).

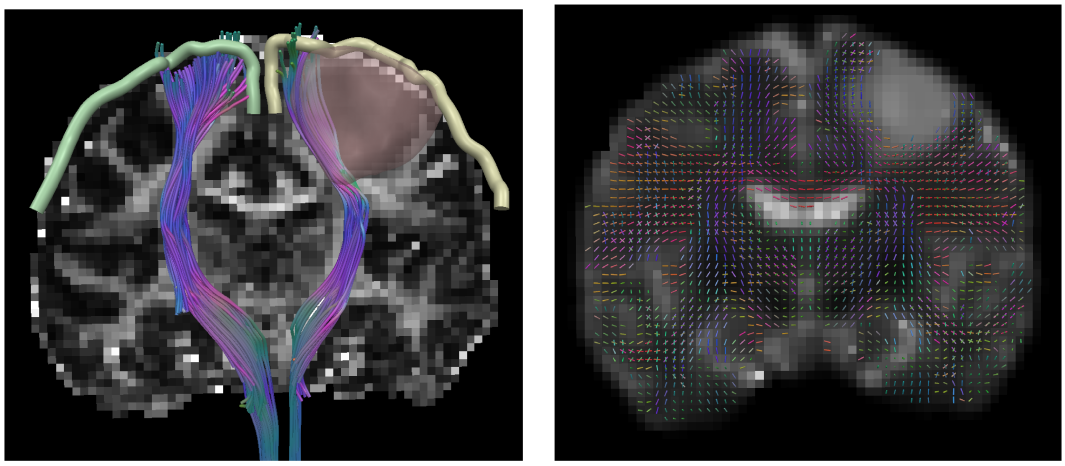


Figure 7.7: A frontal view of a coronal slice of the third patient in the 2015 dataset. This shows the CST (left) and volumetric fiber models (right). In the fiber bundle view, we see the tumor is deeply embedded in the left bundle. In the fiber view, we see a significant drop-out of fiber models in most of the tumor. We also see that the CST reconstructions included the superior projections, but no lateral projections. Since we would expect to see somewhat normal anatomy in the right bundle, it's unclear if this is due to the pathology or limitations in the modeling and tracking, as this case had relatively low b-value of 800 mm/s^2.

added more lateral projections with little other change. The reconstructions in the right hemisphere were significantly different due to the presence of edema and tumor. Only filtered models allowed tracking through some of the edema, and no method tracked through the tumor.

7.2.6 Results from 2015 Challenge

Results from the first patient are shown and described in Fig. 7.4. Results from the second patient are shown and described in Figs. 7.5 and 7.6. Results from the third patient are shown and described in 7.7. The first two cases showed CST reconstructions similar to previous work [105] [175]. The third case's CST included only the most superior projections. The first case showed some deflection of the left CST around the tumor; the second case showed some tracking through edema and possibly physical displacement; the third case showed only superior projections with a great deal of tumor infiltration. The second case included a reconstruction of the arcuate fasciculus with projections to frontal cortex that passed through crossing fiber regions and curved around the tumor in the left hemisphere. The mean diffusivity of the left arcuate also showed variation possibly due to tumor infiltration. We also found that all cases showed reasonable reconstructions of multi-fiber configurations, despite the low and b-value and small number of gradient encoding directions.

7.3 Conclusion

In this study, we examined reconstructions of the CST in three clinical tumor imaging cases. We employed a multi-fiber model to resolve complex fiber configurations at the crossing of the CST with the superior longitudinal fasciculus and corpus callosum. Our model-based approach provides an efficient mechanism for both smoothing and interpolation during tractography, and our results suggest that this step aids in tracking the lateral projections to the precentral gyrus. There are a number of open problems related to this. First, the described framework could also be applied to other diffusion models that represent multiple fiber orientations, such as those based on orientation distribution functions. A comparison could also be made to other methods that filter during model fitting, such as the Kalman filtering approach. In summary, we found the combination of multi-fiber modeling and model-based processing to be a valuable approach that could be considered for clinical research applications that aim to understand the structure of the CST in both health and disease.

Chapter 8

Conclusion

This chapter concludes the thesis with a discussion of what we have accomplished thus far, what research opportunities lie ahead, and reflections on the research process. Unlike previous chapters, this will take a high level view and attempt to gather together some general insights from the past few years of work. The first section summarizes the specific contributions of the thesis work, how they build on prior work in computational diffusion MR imaging, and how they relate to open problems in the field. The last section reflects on the research process and discusses the role of interdisciplinary collaboration in computational research.

8.1 Research Contributions

This section discusses the research contributions of the thesis work. This is will provide a high-level summary of the main contributions of each chapter, as well as a discussion of software artifacts produced during the thesis work that are not presented elsewhere.

In Chapters 3 and 4, we presented results from work on voxel-based methods for analyzing diffusion tensor imaging datasets consisting of many subjects. First, Chapter 3 developed an automated supervoxel segmentation algorithm and evaluated how it can be integrated into commonly-used pipelines for atlas-driven population analysis. Then, Chapter 4 presented a comparative evaluation of spatial mapping in diffusion tensor imaging, which included the supervoxel segmentation algorithm as one of the tested conditions. This part tested scan-rescan reproducibility and sensitivity to

age-related changes in white matter. The general finding was that region-based measures give significantly higher reproducibility than ones that look at single voxels. Furthermore, when measuring the effects of aging, methods that span the whole brain perform better. We found skeleton-based analysis to perform consistently well in representing aging effects; however, we found the addition of supervoxels improved both standard voxel-based and skeleton-based analyses, while simultaneously improving reproducibility.

In Chapter 5, we presented results from work on computational methods for analyzing diffusion MR tractography datasets, which offers a complementary approach to voxel-based analysis for studying brain white matter. The work investigated a framework for analyzing large datasets using a novel representation of curve data produced through tractography. This was used to improve both clustering to enable better selection of bundles in individual subjects, as well as automated segmentation of bundles from a population of subjects. The first general finding was that the clustering algorithm provided results that were as good or better than commonly-used spectral and hierarchical clustering approaches, as measured relative to manually selected curves. The second general finding was that the population-based analysis was reliable across scans and sensitive to aging, with comparable performance to an approach of manually selecting curves with regions of interest. The primary result though, is that the representation obtained with the sparse closest point transform is generally useful in a number of fiber bundle modeling tasks when used in combination with even simple machine learning algorithms.

In Chapter 6, we presented results from work on multi-fiber modeling, which can potentially improve the reconstruction of complex white matter features. The work investigated a model-based framework for image processing with fiber orientation mixtures, which is an abstraction for representing multiple fiber populations per voxel and their respective fraction volume. We showed how this can be used for interpolation, smoothing, and fusion of model-valued volumetric imaging data, which are important tasks for tractography and atlas construction. Our experiments showed how this can reduce the effects of noise with improvements over previously proposed techniques for model selection during estimation. We also demonstrated the creations of a 80 subject population atlas that includes complex fiber bundles not typically found in single tensor atlases. In Chapter 7, we conducted follow up experiments testing this approach for surgical planning. This turned out to be a useful tool in reconstructing the corticospinal tract in the presence of brain tumors. In particular, we found improved reconstruction of the lateral projections of the bundle, which are

commonly missed with reconstruction using single tensor modeling.

All of the work described so far has also included a significant amount of software development, and while an interesting topic of discussion, they are typically out of the scope of the theoretical and experimental contributions. We will now briefly describe some of these results, which includes three main components for data visualization, data analysis, and scientific workflows.

This first component was a 3D rendering application based on Java OpenGL, which was used for nearly all of the visualizations included in this dissertation. While other similar tools are available, we found a need for a custom tool for rendering diffusion model glyphs alongside image volumes and tractography curves, with tight coupling between model parameters and rendering properties of all of these. For example, when rendering multi-fiber models, it is important to carefully map the fiber volume fraction and fiber orientation to the geometry and appearance of tube models representing each compartment, as these visualizations can suffer from severe cluttering of irrelevant fibers with negligible volume fraction.

The second component was a flexible framework for data analysis based on modules, which unlike other similar frameworks, was developed in Java. This provided a robust and rich platform for development, which is to say, there was little work when moving between different operating systems because of the Java Virtual Machine, and software development was greatly simplified by using Integrated Development Environments such as Eclipse and IntelliJ IDEA. The modular design allowed each algorithm to be used in command line scripting, as an interactive component of the visualization tool, and as a library component for development of other algorithms. The use of Java Annotations and Introspection allowed this all to be done automatically for each algorithm, which greatly simplifies the code needed for each implementation.

The third component was a tool for executing scientific workflows, which was used for both applications to clinical studies and simulation experiments for evaluation purposes. Each workflow was defined in the syntax of GNU Make and then executed with a custom tool built on top of Make that retains additional records for data provenance, such as software versions, environment variables, and workflow execution details. This tool also included automated dispatching to a grid-computing system, which allows workflows to be run in parallel across multiple subjects or experimental conditions. In some cases, this can reduce the clock time of an experiment from weeks to less than a day.

Unless otherwise stated, these three components were the major tools used for implementing

the methods and running the experiments described in the work. The other commonly-used tools were FSL [93], DTI-TK, 3D Slicer [169], R, and ggplot [241], which were all essential to the work presented here.

8.2 Open Problems

Through this work, we have made some promising advances in a variety of aspects of computational diffusion MRI; however, there remain many challenges ahead, which we will now briefly discuss.

A first kind of open problem relates to the evaluation of computational imaging methods. There have been many methods developed in this area, but there has been relatively less evaluation of those methods. We have tried to include comparative and quantitative evaluation in each part of the thesis work, but given the many ways diffusion MRI can be used, there is much more to be done. Much of the recent evaluation work in the field has included challenges, which have looked at phantoms, *ex vivo*, and *in vivo* datasets. Phantom datasets can be very useful for conducting reproducible evaluations of methods; however, they are very limited in gauging practical usefulness [51]. *Ex vivo* datasets are powerful as they are more realistic and well understood; however, they can also suffer from changes due from tissue degeneration or fixation. *In vivo* datasets are valuable for gauging practical benefit of methods, but they lack a verifiable ground truth. Perhaps the most powerful way to remedy this is to first acquire *in vivo* data and then compare it to *ex vivo* imaging data of the same sample.

Looking ahead, we can hope that these kind of evaluations are greatly expanded in variety, e.g. by using a large suite of test cases is available. This can help to avoid a type of "overfitting" that can occur by repeatedly testing algorithms with the same dataset. Without such a large suite of test cases, there is potential to design algorithms that are too particularly tuned a single dataset and fail unexpectedly when used elsewhere. It is also very important to understand the complex nature of computational pipelines. Many evaluation studies look at specific pieces, e.g. imaging protocol, registration algorithm, or particular application. However, in practice, the outcome of a study is a complex combination of specific choices at each point in the computational pipeline, and the net effect of the pipeline will typically depend on the specific combination. While it is valuable to evaluate specific steps in the pipeline, the results are also likely to be conditional on the choices made elsewhere in the pipeline. For this reason, there is much work ahead to expand the design

methodological evaluation to look more broadly at the net effect of computational pipelines.

A second kind of open problem relates to the scale of anatomy characterized by our imaging tools. Diffusion MRI gives a very coarse representation of white matter, with a spatial resolution of a few millimeters in each dimension. Despite the appearance of some visualizations, this is a long way from characterizing single axons. Instead, the diffusion MR signal reflects aggregate properties of many axons with a complex pattern of myelination, caliber, density, and orientation coherence. There is much interesting work remaining ahead to characterize how the great variety in axons is reflected in the diffusion MR signal, and some of the most exciting involves looking at the fine details with microscopy and relating these to MR imaging data of the same specimen. In particular, one could imagine looking at anatomy by zooming in with different imaging techniques. This process could start with whole brain dMRI scans with a typical spatial resolution of 1mm cubic voxels, then focus on a dMRI slice at microscopic resolution with 10μm in-plane resolution, finally selecting a portion of single voxel from that plane to analyze microscopically. This could be done with electron microscopy or more complex microscopy systems, such as CLARITY [214].

Another open problem related to scale is variation across time and individuals. The high level goal here is characterize anatomical variation across large population or across timepoints to understand longitudinal changes with development, aging, and disease progression. One can imagine tools that allow us to zoom out to better understand the large scale patterns across this domain, and the major limitation in this process is solving this computationally, e.g. through data acquisition, management, and interpretation. There are already efforts to acquire large amounts of imaging data; however, data management and interpretation are still open problems. This will likely require advancements in databases, machine learning, statistics, and visualization to make sense of these increasingly complex datasets.

8.3 Research Process

In the previous chapters, we have aimed to present a polished and coherent record of work done to advance computational diffusion MR imaging, but at times, the actual route taken in research is more rough and circuitous than the presented in the chapters and papers summarizing the work. This section aims to convey some experiences and lessons that would not have a home elsewhere.

The basic lesson learned is that interdisciplinary research fundamentally requires its development

and evaluation stages to be driven by both the needs of practical applications and by desirable theoretical aspects. This is not a new idea, however, and a pioneer in Computer Scientist, Fred Brooks, wrote about this in *The Computer Scientist as Toolsmith: Studies in Interactive Computer Graphics* [30]: "Hitching our research to someone else's driving problem, and solving those problems on the owner's terms, leads us to a richer computer science research".

While this is likely to be true in all interdisciplinary research, MR imaging research is a prime example. Many disciplines that must come together, and the modern tools of MRI would not exist without the people drawn from physics, engineering, medicine, biology, and computer science. While it can be difficult to bridge the gap in jargon, goals, and culture, these differences should not be discounted as a hurdle to be overcome, but instead embraced as an opportunity for amplifying the results in any of these respective areas. Indeed, this researcher found that the most progress was made by exchanging ideas, which can help to appropriately broaden some parts of a research plan and narrow others. The alternative is to develop computational tools that are theoretically interesting and may have potential for impact, but risk becoming hammers in search of a nail.

This process can be imagined as a collaborative exchange between toolsmiths and brain scientists and clinicians, which is illustrated in Fig. 8.1, and consists of two major parts. First, there is methodological development, which includes much of the computer science in imaging research, such as the design and analysis of computational and mathematical machinery, as well as implementation in software systems. Second, there is methodological evaluation, which includes experimental techniques for judging how well these tools solve problems in real applications and how they meet their own design criteria. These form the basis of how research collaboration happens, where the computational side of the collaboration develops tools targeted at specific applications, and the evaluation is conducted with experimental data from the applications, both synthetic and real. These steps can be iterated throughout the collaboration, with the hope that the work periodically converges to something that satisfies both theoretical goals of the toolsmiths and the practical goals of the users.

An interesting possibility is that each of these components interacts with one other to provide a greater result than any part conducted in isolation. It may be tempting to say, "computational research is sufficiently abstract that it can be conducted independently of any specific application". This may be true in many situations; however, much of the current research in computer science now reaches applications with sufficient complexity that no single set of abstractions is best. Even worse, we may be biased towards one method or another and ignore a better solution, a tendency which

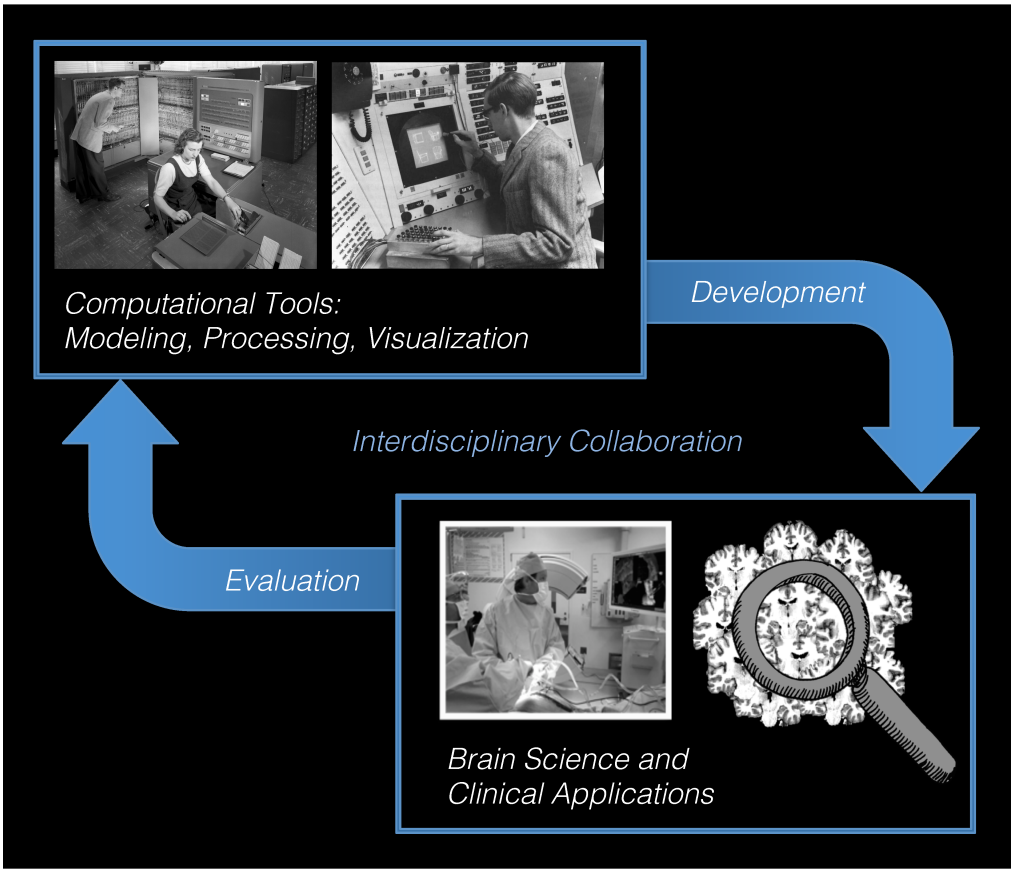


Figure 8.1: An illustration of interdisciplinary research in computational brain mapping. This is depicted as an process of collaboration between toolsmiths and domain experts in the brain sciences and clinic. This collaboration is broken down into development and evaluation, which are repeated iteratively to hopefully converge on computational tools that satisfy both theoretical goals of the toolsmiths and the practical goals of the users.

is known as the confirmation bias in psychology. Abraham Kaplan also described this challenge as the "law of the instrument" in his book *The Conduct of Inquiry* [101]:

> In addition to the social pressures from the scientific community there is also at work a very human trait of individual scientist. I call it the law of the instrument, and it may be formulated as follows: Give a small boy a hammer, and he will find that everything he encounters needs pounding. It comes as no particular surprise to discover that a scientist formulates problems in a way which requires for their solution just those techniques in which he himself is especially skilled.

Furthermore, it may be possible that the toolsmith is not even the best judge of a tool's value. This idea can be traced back to the earliest days of science in Aristotle's *Politics*. In this work, he described the wisdom of the crowd in governing compared to even a highly trained specialist, as no one person can imagine the possible needs of a complex society. Computational tool development is perhaps similar, where the users of the tool can identify strengths and weaknesses of a tool beyond the imagination of a single toolsmith. Aristotle explored this possibility with an apt comparison to construction and cooking, which applies equally well to the computational toolsmith:

> ...there are some arts whose products are not judged of solely, or best, by the artists themselves, namely those arts whose products are recognized even by those who do not possess the art; for example, the knowledge of the house is not limited to the builder only; the user, or, in other words, the master, of the house will be even a better judge than the builder, just as the pilot will judge better of a rudder than the carpenter, and the guest will judge better of a feast than the cook.

Given this, it seems wise to consider a variety of approaches to solving a problem, and we can benefit from research practices that explicitly encourage practical evaluation. The collaborative research process is a way accomplish this goal and break free of the biases and limitations of any particular tool by exploring a variety of possibilities and judging which are useful from perspectives besides our own. Using this thesis work as a concrete example, we found a variety of useful representations for image-based modeling of brain anatomy, as white matter can be understood in terms of voxels, bundles, networks, depending on the clinical task or scientific question at hand. This is likely the case for any work with applications to biology, where there is a large amount of

"essential" complexity. An obvious or intuitive abstraction of a problem may not exist or it may ignore too much of the reality of the problem, so we must empirically gauge the usefulness of our models through collaborations and experimental evaluation. Returning to thoughts of Fred Brooks on computational toolsmithing, the idea of judging tools this way was related as follows [30]:

> If we perceive our role [as toolsmiths], we then see more clearly the proper criterion for success: a toolmaker succeeds as, and only as, the users of this tool succeed. However shiny the blade; however jeweled the hilt, however perfect the heft, a sword is tested only by cutting. That swordsmith is successful whose clients die of old age.

To take this thought to an extreme, a naive toolsmith could imagine that a hammer would suffice if they had no experience with anyone who does the cutting. Similarly, a computational toolsmith could fall victim to a bias towards the tools they know. With that, let us hope that we ultimately find the tools and knowledge they provide, but with such a complex task as understanding the brain, this dissertation represents only a few steps on a much longer path towards that goal!

Bibliography

[1] Thomas Lumley and Alan Miller. *leaps: regression subset selection*. R Foundation for Statistical Computing, 2016.

[2] Til Aach, Cicero Mota, Ingo Stuke, Matthias Muhlich, and Erhardt Barth. Analysis of superimposed oriented patterns. *Image Processing, IEEE Transactions on*, 15(12):3690–3700, 2006.

[3] Iman Aganj, Christophe Lenglet, Guillermo Sapiro, Essa Yacoub, Kamil Ugurbil, and Noam Harel. Reconstruction of the orientation distribution function in single-and multiple-shell q-ball imaging within constant solid angle. *Magnetic Resonance in Medicine*, 64(2):554–566, 2010.

[4] Andrew L Alexander, Jee Eun Lee, Mariana Lazar, and Aaron S Field. Diffusion tensor imaging of the brain. *Neurotherapeutics*, 4(3):316–329, 2007.

[5] Daniel C Alexander. Multiple-fiber reconstruction algorithms for diffusion MRI. *Annals of the New York Academy of Sciences*, 1064(1):113–133, 2005.

[6] Daniel C Alexander, Carlo Pierpaoli, Peter J Basser, and James C Gee. Spatial transformations of diffusion tensor magnetic resonance images. *IEEE Transactions on Medical Imaging*, 20(11):1131–1139, 2001.

[7] J Ashburner and K J Friston. Voxel-based morphometry–the methods. *NeuroImage*, 11(6 Pt 1):805–21, June 2000.

[8] John Ashburner and Karl J Friston. Why voxel-based morphometry should be used. *NeuroImage*, 14(6):1238–1243, 2001.

[9] Katy S Azoury and Manfred K. Warmuth. Relative loss bounds for on-line density estimation with the exponential family of distributions. *Machine Learning*, 43(3), 2001.

[10] Michael Bach, Frederik B Laun, Alexander Leemans, Chantal MW Tax, Geert J Biessels, Bram Stieltjes, and Klaus H Maier-Hein. Methodological considerations on tract-based spatial statistics (tbss). *NeuroImage*, 100:358–369, 2014.

[11] Arindam Banerjee, Srujana Merugu, Inderjit S Dhillon, and Joydeep Ghosh. Clustering with Bregman divergences. *The Journal of Machine Learning Research*, 6:1705–1749, 2005.

[12] Vic Barnett and Toby Lewis. *Outliers in statistical data*, volume 3. Wiley New York, 1994.

[13] P J Basser and C Pierpaoli. Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. *Journal of magnetic resonance. Series B*, 111(3):209–219, 1996.

[14] Peter J Basser, James Mattiello, and Denis LeBihan. Estimation of the effective self-diffusion tensor from the NMR spin echo. *Journal of Magnetic Resonance, Series B*, 103(3):247–254, 1994.

[15] Peter J Basser, James Mattiello, and Denis LeBihan. MR diffusion tensor spectroscopy and imaging. *Biophysical journal*, 66(1):259, 1994.

[16] Peter J Basser, Sinisa Pajevic, Carlo Pierpaoli, Jeffrey Duda, and Akram Aldroubi. In vivo fiber tractography using DT-MRI data. *Magnetic resonance in medicine*, 44(4):625–632, 2000.

[17] Christian Beaulieu. The basis of anisotropic water diffusion in the nervous system–a technical review. *NMR in Biomedicine*, 15(7-8):435–455, 2002.

[18] SMA Becker, Karsten Tabelow, Henning U Voss, Alfred Anwander, Robin M Heidemann, and Jörg Polzehl. Position-orientation adaptive smoothing of diffusion weighted magnetic resonance data (POAS). *Medical image analysis*, 16(6):1142–1155, 2012.

[19] TEJ Behrens, H Johansen Berg, Saad Jbabdi, MFS Rushworth, and MW Woolrich. Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *NeuroImage*, 34(1):144–155, 2007.

[20] Rabi Bhattacharya and Vic Patrangenaru. Large sample theory of intrinsic and extrinsic sample means on manifolds. I. *Annals of statistics*, pages 1–29, 2003.

[21] S Bisdas, DE Bohning, N Bešenski, JS Nicholas, and Z Rumboldt. Reproducibility, interrater agreement, and age-related changes of fractional anisotropy measures at 3T in healthy subjects: effect of the applied b-value. *American Journal of Neuroradiology*, 29(6):1128–1133, 2008.

[22] Christopher M Bishop. Novelty detection and neural network validation. In *Vision, Image and Signal Processing, IEEE Proceedings*, volume 141, pages 217–222. IET, 1994.

[23] J M Bland and D G Altman. A note on the use of the intraclass correlation coefficient in the evaluation of agreement between two methods of measurement. *Computers in biology and medicine*, 20(5):337–40, January 1990.

[24] J Martin Bland and Douglas G Altman. Statistics notes: measurement error proportional to the mean. *BMJ*, 313(7049):106, 1996.

[25] Felix Bloch. Generalized theory of relaxation. *Physical Review*, 105(4):1206, 1957.

[26] Luke Bloy, Madhura Ingalhalikar, Harini Eavani, Robert T Schultz, Timothy P L Roberts, and Ragini Verma. White matter atlas generation using HARDI based automated parcellation. *NeuroImage*, 59(4):4055–4063, 2012.

[27] David Bonekamp, Lidia M Nagae, Mahaveer Degaonkar, Melissa Matson, Wael MA Abdalla, Peter B Barker, Susumu Mori, and Alena Horská. Diffusion tensor imaging in children and adolescents: reproducibility, hemispheric, and age-related differences. *NeuroImage*, 34(2):733–742, 2007.

[28] Antti Brander, Anneli Kataja, Antti Saastamoinen, Pertti Ryymin, Heini Huhtala, Juha Öhman, Seppo Soimakallio, and Prasun Dastidar. Diffusion tensor imaging of the brain in a healthy adult population: Normative values and measurement reproducibility at 3T and 1.5T. *Acta Radiologica*, 51(7):800–807, 2010.

[29] K Brodmann. Vergleichende lokalisationslehre der groshirnrinde. *Leipzig: Barth*, 1909.

[30] Frederick P Brooks Jr. The computer scientist as toolsmith II. *Communications of the ACM*, 39(3):61–68, 1996.

[31] A Brun, H Knutsson, HJ Park, Martha E Shenton, and Carl-Fredrik Westin. Clustering fiber traces using normalized cuts. *MICCAI*, 2004(3216):368–375, 2004.

[32] A. Brun, C. F. Westin, M. Herberthson, and H. Knutsson. Intrinsic and extrinsic means on the circle - a maximum likelihood interpretation. In *2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07*, volume 3, pages III–1053–III–1056, April 2007.

[33] Ryan P Cabeen, Mark E Bastin, and David H Laidlaw. A diffusion MRI resource of 80 age-varied subjects with neuropsychological and demographic measures. In *ISMRM, 21st Scientific Meeting and Exhibition*, number 2138, Salt Lake City, Utah, 2013.

[34] Ryan P Cabeen, Mark E Bastin, and David H Laidlaw. Estimating constrained multi-fiber diffusion MR volumes by orientation clustering. *MICCAI*, 16(Pt 1):82–9, January 2013.

[35] Ryan P Cabeen, Mark E Bastin, and David H Laidlaw. Kernel regression estimation of fiber orientation mixtures in diffusion MRI. *NeuroImage*, 127:158–172, 2016.

[36] Ryan P Cabeen and David H Laidlaw. Bilateral filtering of multiple fiber orientations in diffusion MRI. *MICCAI Workshop on Computational Diffusion MRI*, pages 193–202, 2014.

[37] Ryan P Cabeen and David H Laidlaw. White matter supervoxel segmentation by axial DP-means clustering. *MICCAI Medical Computer Vision*, 2014.

[38] Paul T Callaghan. *Principles of nuclear magnetic resonance microscopy*. Oxford University Press, 1993.

[39] Owen Carmichael and Samuel Lockhart. The role of diffusion tensor imaging in the study of cognitive aging. In *Brain Imaging in Behavioral Neuroscience*, pages 289–320. Springer, 2012.

[40] Herman Y Carr and Edward M Purcell. Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Physical review*, 94(3):630, 1954.

[41] Marco Catani and Michel Thiebaut De Schotten. A diffusion tensor imaging tractography atlas for virtual in vivo dissections. *Cortex*, 44(8):1105–1132, 2008.

[42] Marco Catani and Michel Thiebaut de Schotten. *Atlas of human brain connections*. Oxford University Press, 2012.

[43] Marco Catani, Robert J. Howard, Sinisa Pajevic, and Derek K. Jones. Virtual in vivo interactive dissection of white matter fasciculi in the human brain. *NeuroImage*, 17(1):77–94, September 2002.

[44] Mara Cercignani, Roland Bammer, Maria P Sormani, Franz Fazekas, and Massimo Filippi. Inter-sequence and inter-imaging unit variability of diffusion tensor MR imaging histogram-derived metrics of the brain in healthy volunteers. *American Journal of Neuroradiology*, 24(4):638–643, 2003.

[45] Thomas L Chenevert, James A Brunberg, and JG Pipe. Anisotropic diffusion in human white matter: demonstration with MR techniques in vivo. *Radiology*, 177(2):401–405, 1990.

[46] Jonathan D Clayden, Amos J Storkey, and Mark E Bastin. A Probabilistic Model-Based Approach to Consistent White Matter Tract Segmentation. *IEEE Transaction on Medical Imaging*, 26(11):1555–1561, November 2007.

[47] Volker Arnd Coenen, Timo Krings, Hubertus Axer, Jürgen Weidemann, Heidi Kränzlein, Franz-Josef Hans, Armin Thron, Joachim Michael Gilsbach, and Veit Rohde. Intraoperative three-dimensional visualization of the pyramidal tract in a neuronavigation system (PTV) reliably predicts true position of principal motor pathways. *Surgical Neurology*, 60(5):381–390, November 2003.

[48] PA Cook, Y Bai, SKKS Nedjati-Gilani, KK Seunarine, MG Hall, GJ Parker, and DC Alexander. Camino: open-source diffusion-mri reconstruction and processing. In *14th scientific meeting of the international society for magnetic resonance in medicine*, volume 2759. Seattle WA, USA, 2006.

[49] Isabelle Corouge, P Thomas Fletcher, Sarang C Joshi, Sylvain Gouttard, and Guido Gerig. Fiber tract-oriented statistics for quantitative diffusion tensor MRI analysis. *Medical Image Analysis*, 10(5):786–798, October 2006.

[50] Stephen Correia, Stephanie Y Lee, Thom Voorn, David F Tate, Robert H Paul, Song Zhang, Stephen P Salloway, Paul F Malloy, and David H Laidlaw. Quantitative tractography metrics of white matter integrity in diffusion-tensor MRI. *NeuroImage*, 42(2):568–81, August 2008.

[51] Marc-Alexandre Côté, Gabriel Girard, Arnaud Boré, Eleftherios Garyfallidis, Jean-Christophe Houde, and Maxime Descoteaux. Tractometer: towards validation of tractography pipelines. *Medical image analysis*, 17(7):844–857, 2013.

[52] Olivier Coulon, Daniel C Alexander, and S Arridge. Diffusion tensor magnetic resonance image regularization. *Medical image analysis*, 8(1):47–67, 2004.

[53] Raymond Damadian. Tumor detection by nuclear magnetic resonance. *Science*, 171(3976):1151–1153, 1971.

[54] Dick de Ridder, D Tax, and R Duin. An experimental comparison of one-class classification methods. In *Proceedings of the 4th Annual Conference of the Advacned School for Computing and Imaging, Delft*, 1998.

[55] Michel Thiebaut De Schotten, Flavio Dell'Acqua, Stephanie J Forkel, Andrew Simmons, Francesco Vergani, Declan GM Murphy, and Marco Catani. A lateralized brain network for visuospatial attention. *Nature neuroscience*, 14(10):1245–1246, 2011.

[56] Michel Thiebaut De Schotten, Flavio DellAcqua, Romain Valabregue, and Marco Catani. Monkey to human comparative anatomy of the frontal lobe association tracts. *Cortex*, 48(1):82–96, 2012.

[57] Maxime Descoteaux, Rachid Deriche, Thomas R Knösche, and Alfred Anwander. Deterministic and probabilistic tractography based on complex fibre orientation distributions. *IEEE Transactions on Medical Imaging*, 28(2):269–286, 2009.

[58] Inderjit S. Dhillon, Edward M. Marcotte, and Usman Roshan. Diametrical clustering for identifying anti-correlated gene clusters. *Bioinformatics*, 19(13):1612–1619, September 2003.

[59] Lee R Dice. Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302, 1945.

[60] Olaf Dietrich, José G Raya, Scott B Reeder, Maximilian F Reiser, and Stefan O Schoenberg. Measurement of signal-to-noise ratios in mr images: Influence of multichannel coils, parallel imaging, and reconstruction filters. *Journal of Magnetic Resonance Imaging*, 26(2):375–385, 2007.

[61] Luca Dodero, Sebastiano Vascon, Vittorio Murino, Angelo Bifone, Alessandro Gozzi, and Diego Sona. Automated multi-subject fiber clustering of mouse brain using dominant sets. *Frontiers in Neuroinformatics*, 8(January):1–12, 2015.

[62] Ian L Dryden, Alexey Koloydenko, and Diwei Zhou. Non-euclidean statistics for covariance matrices, with applications to diffusion tensor imaging. *The Annals of Applied Statistics*, pages 1102–1123, 2009.

[63] Jia Du, Alvina Goh, and Anqi Qiu. Diffeomorphic metric mapping of high angular resolution diffusion imaging based on Riemannian structure of orientation distribution functions. *IEEE Transactions on Medical Imaging*, 31(5):1021–1033, 2012.

[64] Uran Ferizi, Torben Schneider, Eleftheria Panagiotaki, Gemma Nedjati-Gilani, Hui Zhang, Claudia AM Wheeler-Kingshott, and Daniel C Alexander. A ranking of diffusion MRI compartment models with in vivo human brain data. *Magnetic Resonance in Medicine*, 72(6):1785–1792, 2014.

[65] Ephraim Joshua Field, Richard John Harrison, et al. *Anatomical terms: their origin and derivation.* W. Heffer & Sons., 1947.

[66] Niels K Focke, Mahinda Yogarajah, Silvia B Bonelli, Philippa A Bartlett, Mark R Symms, and John S Duncan. Voxel-based diffusion tensor imaging in patients with mesial temporal lobe epilepsy and hippocampal sclerosis. *NeuroImage*, 40(2):728–737, 2008.

[67] RJ Fox, K Sakaie, J-C Lee, JP Debbins, Y Liu, DL Arnold, ER Melhem, CH Smith, MD Philips, M Lowe, et al. A validation study of multicenter diffusion tensor imaging: reliability of fractional anisotropy and diffusivity values. *American Journal of Neuroradiology*, 33(4):695–700, 2012.

[68] Vincent Garcia and Frank Nielsen. Simplification and hierarchical representations of mixtures of exponential families. *Signal Processing*, 90(12):3197–3212, 2010.

[69] G Gerig, S Gouttard, and I Corouge. Analysis of brain white matter via fiber tract modeling. In *Engineering in Medicine and Biology Society, 2004. IEMBS '04. 26th Annual International Conference of the IEEE*, volume 2, pages 4421–4424, 2004.

[70] Tal Geva. Magnetic resonance imaging: historical perspective. *Journal of Cardiovascular Magnetic Resonance*, 8(4):573–580, 2006.

[71] Antonio Giorgio, Luca Santelli, Valentina Tomassini, Rose Bosnell, Steve Smith, Nicola De Stefano, and Heidi Johansen-Berg. Age-related changes in grey and white matter structure throughout adulthood. *NeuroImage*, 51(3):943–51, July 2010.

[72] Nicole R Giuliani, Vince D Calhoun, Godfrey D Pearlson, Alan Francis, and Robert W Buchanan. Voxel-based morphometry versus region of interest: a comparison of two methods for analyzing gray matter differences in schizophrenia. *Schizophrenia research*, 74(2):135–147, 2005.

[73] Alvina Goh, Christophe Lenglet, Paul M Thompson, and Rene Vidal. A nonparametric Riemannian framework for processing high angular resolution diffusion images. *CVPR*, 2009.

[74] Alvina Goh, Christophe Lenglet, Paul M Thompson, and Rene Vidal. A nonparametric Riemannian framework for processing high angular resolution diffusion images and its applications to ODF-based morphometry. *NeuroImage*, 56(3):1181–1201, 2011.

[75] Henry Gray. *Anatomy of the human body*. Lea & Febiger, 1918.

[76] Matthew Grech-Sollars, Patrick W Hales, Keiko Miyazaki, Felix Raschke, Daniel Rodriguez, Martin Wilson, Simrandip K Gill, Tina Banks, Dawn E Saunders, Jonathan D Clayden, et al. Multi-centre reproducibility of diffusion MRI parameters for clinical sequences in the brain. *NMR in Biomedicine*, 28(4):468–485, 2015.

[77] Hákon Gudbjartsson and Samuel Patz. The rician distribution of noisy MRI data. *Magnetic resonance in medicine*, 34(6):910–914, 1995.

[78] Patric Hagmann, Maciej Kurant, Xavier Gigandet, Patrick Thiran, Van J Wedeen, Reto Meuli, and Jean-Philippe Thiran. Mapping human whole-brain structural networks with diffusion MRI. *PloS one*, 2(7):e597, 2007.

[79] EL Hahn and DE Maxwell. Spin echo measurements of nuclear spin coupling in molecules. *Physical Review*, 88(5):1070, 1952.

[80] Erwin L Hahn. Spin echoes. *Physical review*, 80(4):580, 1950.

[81] Ullamari Hakulinen, Antti Brander, Pertti Ryymin, Juha Öhman, Seppo Soimakallio, Mika Helminen, Prasun Dastidar, and Hannu Eskola. Repeatability and variation of region-of-interest methods using quantitative diffusion tensor MR imaging of the brain. *BMC Medical Imaging*, 12(1):30, 2012.

[82] Ghassan Hamarneh and Judith Hradsky. Bilateral filtering of diffusion tensor magnetic resonance images. *IEEE Transactions on Image Processing*, 16(10):2463–2475, 2007.

[83] Walter A. Hendricks and Kate W. Robey. The Sampling Distribution of the Coefficient of Variation. *Annals of Mathematical Statistics*, 7(3):129–132, 2008.

[84] Mark A Horsfield and Derek K Jones. Applications of diffusion-weighted and diffusion tensor MRI to white matter diseases–a review. *NMR in Biomedicine*, 15(7-8):570–577, 2002.

[85] Kegang Hua, Jiangyang Zhang, Setsu Wakana, Hangyi Jiang, Xin Li, Daniel S Reich, Peter A Calabresi, James J Pekar, Peter CM van Zijl, and Susumu Mori. Tract probability maps in stereotaxic spaces: analyses of white matter anatomy and tract-specific quantification. *NeuroImage*, 39(1):336–347, 2008.

[86] Lawrence Hubert and Phipps Arabie. Comparing partitions. *Journal of classification*, 2(1):193–218, 1985.

[87] Thomas R Insel, Story C Landis, and Francis S Collins. The NIH brain initiative. *Science*, 340(6133):687–688, 2013.

[88] Neda Jahanshad, Peter V Kochunov, Emma Sprooten, René C Mandl, Thomas E Nichols, Laura Almasy, John Blangero, Rachel M Brouwer, Joanne E Curran, Greig I de Zubicaray, et al. Multi-site genetic analysis of diffusion images and voxelwise heritability analysis: A pilot project of the ENIGMA–DTI working group. *NeuroImage*, 81:455–469, 2013.

[89] Jacobus FA Jansen, M Eline Kooi, Alfons GH Kessels, Klaas Nicolay, and Walter H Backes. Reproducibility of quantitative cerebral T2 relaxometry, diffusion tensor imaging, and 1H magnetic resonance spectroscopy at 3.0 Tesla. *Investigative Radiology*, 42(6):327–337, 2007.

[90] Saad Jbabdi, Timothy EJ Behrens, and Stephen M Smith. Crossing fibres in tract-based spatial statistics. *NeuroImage*, 49(1):249–256, 2010.

[91] Saad Jbabdi, Julia F Lehman, Suzanne N Haber, and Timothy E Behrens. Human and monkey ventral prefrontal fibers use the same organizational principles to reach their targets: tracing versus tractography. *The Journal of Neuroscience*, 33(7):3190–3201, 2013.

[92] Saad Jbabdi, MW Woolrich, JLR Andersson, and TEJ Behrens. A Bayesian framework for global tractography. *NeuroImage*, 37(1):116–129, 2007.

[93] Mark Jenkinson, Christian F. Beckmann, Timothy E J Behrens, Mark W. Woolrich, and Stephen M. Smith. FSL. *NeuroImage*, 62:782–790, 2012.

[94] Ben Jeurissen, Alexander Leemans, Jacques-Donald Tournier, Derek K Jones, and Jan Sijbers. Investigating the prevalence of complex fiber configurations in white matter tissue with diffusion magnetic resonance imaging. *Human brain mapping*, 34(11):2747–2766, 2013.

[95] Ke Jiang, Brian Kulis, and Michael Jordan. Small-variance asymptotics for exponential family Dirichlet process mixture models. In *NIPS*, pages 3167–3175, 2012.

[96] Heidi Johansen-Berg and Timothy EJ Behrens. *Diffusion MRI: from quantitative measurement to in vivo neuroanatomy*. Academic Press, 2013.

[97] Derek K Jones. Studying connections in the living human brain with diffusion MRI. *cortex*, 44(8):936–952, 2008.

[98] Derek K Jones, Lewis D Griffin, Daniel C Alexander, Marco Catani, Mark A Horsfield, Robert Howard, and Steve C R Williams. Spatial Normalization and Averaging of Diffusion Tensor MRI Data Sets. *NeuroImage*, 17(2):592–617, 2002.

[99] Derek K Jones, Mark R Symms, Mara Cercignani, and Robert J Howard. The effect of filter size on VBM analyses of DT-MRI data. *NeuroImage*, 26(2):546–554, 2005.

[100] DK Jones, XA Chitnis, D Job, PL Khong, LT Leung, S Marenco, SM Smith, and MR Symms. What happens when nine different groups analyze the same DT-MRI data set using voxel-based methods. In *Proceedings of the 15th Annual Meeting of the International Society for Magnetic Resonance in Medicine, Berlin*, page 74, 2007.

[101] Abraham Kaplan. *The conduct of inquiry*. Transaction Publishers, 1973.

[102] Robert E Kass and Adrian E Raftery. Bayes factors. *Journal of the American Statistical Association*, 90(430):773–795, 1995.

[103] Shiva Keihaninejad, Natalie S Ryan, Ian B Malone, Marc Modat, David Cash, Gerard R Ridgway, Hui Zhang, Nick C Fox, and Sebastien Ourselin. The importance of group-wise registration in tract based spatial statistics study of neurodegeneration: a simulation study in alzheimer's disease. 2012.

[104] Gordon Kindlmann. Superquadric tensor glyphs. In *Proceedings of the Sixth Joint Eurographics-IEEE TCVG conference on Visualization*, pages 147–154. Eurographics Association, 2004.

[105] Manabu Kinoshita, Kei Yamada, Naoya Hashimoto, Amami Kato, Shuichi Izumoto, Takahito Baba, Motohiko Maruno, Tsunehiko Nishimura, and Toshiki Yoshimine. Fiber-tracking does not accurately estimate size of fiber bundle in pathological condition: initial neurosurgical experience using neuronavigation and subcortical white matter stimulation. *NeuroImage*, 25(2):424–9, April 2005.

[106] Matthew A Kirkman. The role of imaging in the development of neurosurgery. *Journal of Clinical Neuroscience*, 22(1):55–61, 2015.

[107] Hans Knutsson. Representing local structure using tensors. In *Proceedings of the 6th Scandinavian Conference on Image Analysis*, 1989.

[108] Brian Kulis and Michael I Jordan. Revisiting k-means: New algorithms via Bayesian nonparametrics. In *Proceedings of the 29th International Conference on Machine Learning (ICML-12)*, pages 513–520, 2012.

[109] Anil Kumar, Dieter Welti, and Richard R Ernst. NMR Fourier zeugmatography. *Journal of magnetic resonance*, 213(2):495–509, 2011.

[110] Marinos Kyriakopoulos, Nora S Vyas, Gareth J Barker, Xavier A Chitnis, and Sophia Frangou. A diffusion tensor imaging study of white matter in early-onset schizophrenia. *Biological psychiatry*, 63(5):519–523, 2008.

[111] C-M Lai and PC Lauterbur. True three-dimensional image reconstruction by nuclear magnetic resonance zeugmatography. *Physics in medicine and biology*, 26(5):851, 1981.

[112] David H Laidlaw, Eric T Ahrens, David Kremers, Matthew J Avalos, Russell E Jacobs, and Carol Readhead. Visualizing diffusion tensor images of the mouse spinal cord. In *Visualization'98. Proceedings*, pages 127–134. IEEE, 1998.

[113] Paul C Lauterbur. Magnetic resonance zeugmatography. *Pure and Applied Chemistry*, 40(1-2):149–157, 1974.

[114] Paul C Lauterbur et al. Image formation by induced local interactions: examples employing nuclear magnetic resonance. *Nature*, 242(5394):190–191, 1973.

[115] Denis Le Bihan, Eric Breton, Denis Lallemand, ML Aubin, J Vignaud, and M Laval-Jeantet. Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. *Radiology*, 168(2):497–505, 1988.

[116] Denis Le Bihan, Eric Breton, Denis Lallemand, Philippe Grenier, Emmanuel Cabanis, and Maurice Laval-Jeantet. MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. *Radiology*, 161(2):401–407, 1986.

[117] Denis Le Bihan and Heidi Johansen-Berg. Diffusion MRI at 25: exploring brain tissue structure and function. *NeuroImage*, 61(2):324–341, 2012.

[118] A Leemans, J Sijbers, M Verhoye, A Van der Linden, and D Van Dyck. Mathematical framework for simulating diffusion tensor MR neural fiber bundles. *Magnetic resonance in medicine*, 53(4):944–953, 2005.

[119] Alexander Leemans and Derek K Jones. The B-matrix must be rotated when correcting for subject motion in DTI data. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, 61(6):1336–49, June 2009.

[120] Thomas M Lehmann, Claudia Gönner, and Klaus Spitzer. Survey: Interpolation methods in medical image processing. *IEEE Transactions on Medical Imaging*, 18(11):1049–1075, 1999.

[121] Christoph Lenglet, Jennifer SW Campbell, Maxime Descoteaux, Gloria Haro, Peter Savadjiev, Demian Wassermann, Alfred Anwander, Rachid Deriche, G Bruce Pike, Guillermo Sapiro, et al. Mathematical methods for diffusion MRI processing. *NeuroImage*, 45(1):S111–S122, 2009.

[122] KO Lim and JA Helpern. Neuropsychiatric applications of DTI–a review. *NMR in Biomedicine*, 15(7-8):587–593, 2002.

[123] Xin Liu, Yong Yang, Jubao Sun, Gang Yu, Jin Xu, Chen Niu, Hongjun Tian, and Pan Lin. Reproducibility of diffusion tensor imaging in normal subjects: an evaluation of different gradient sampling schemes and registration algorithm. *Neuroradiology*, 56(6):497–510, 2014.

[124] Aurlien Lucchi, Kevin Smith, Radhakrishna Achanta, Vincent Lepetit, and Pascal Fua. A fully automated approach to segmentation of irregularly shaped cellular structures in em images. In *MICCAI*, volume 6362 of *LNCS*, pages 463–471. Springer Berlin Heidelberg, 2010.

[125] Wen-Lin Luo and Thomas E Nichols. Diagnosis and exploration of massively univariate neuroimaging models. *NeuroImage*, 19(3):1014–1032, 2003.

[126] Mahnaz Maddah, W Eric L Grimson, Simon K Warfield, and William M Wells. A unified framework for clustering and quantitative analysis of white matter fiber tracts. *Medical Image Analysis*, 12(2):191–202, 2008.

[127] Tara Madhyastha, Susan Merillat, Sarah Hirsiger, Ladina Bezzola, Franziskus Liem, Thomas Grabowski, and Lutz Jäncke. Longitudinal reliability of tract-based spatial statistics in diffusion tensor imaging. *Human brain mapping*, 35(9):4544–4555, 2014.

[128] Vincent A Magnotta, Joy T Matsui, Dawei Liu, Hans J Johnson, Jeffrey D Long, Bradley D Bolster Jr, Bryon A Mueller, Kelvin Lim, Susumu Mori, Karl G Helmer, et al. Multicenter reliability of diffusion tensor imaging. *Brain connectivity*, 2(6):345–355, 2012.

[129] JB Antoine Maintz and Max A Viergever. A survey of medical image registration. *Medical image analysis*, 2(1):1–36, 1998.

[130] Daniel Mamah, Thomas E Conturo, Michael P Harms, Erbil Akbudak, Lei Wang, Amanda R McMichael, Mokhtar H Gado, Deanna M Barch, and John G Csernansky. Anterior thalamic radiation integrity in schizophrenia: a diffusion-tensor imaging study. *Psychiatry Research: Neuroimaging*, 183(2):144–150, 2010.

[131] Peter Mansfield and Peter K Grannell. NMR diffraction in solids? *Journal of Physics C: solid state physics*, 6(22):L422, 1973.

[132] Stefano Marenco, Robert Rawlings, Gustavo K Rohde, Alan S Barnett, Robyn A Honea, Carlo Pierpaoli, and Daniel R Weinberger. Regional distribution of measurement error in diffusion tensor imaging. *Psychiatry Research: Neuroimaging*, 147(1):69–78, 2006.

[133] John Harry Martin, Michael E Leonard, and Howard Radzyner. *Neuroanatomy: text and atlas*. Elsevier Amsterdam, 1989.

[134] Sean Mauch. A fast algorithm for computing the closest point and distance transform. *http://www.acm.caltech.edu/seanm/software/cpt/cpt.pdf*, 2000.

[135] John Mazziotta, Arthur Toga, Alan Evans, Peter Fox, Jack Lancaster, Karl Zilles, Roger Woods, Tomas Paus, Gregory Simpson, Bruce Pike, et al. A probabilistic atlas and reference system for the human brain: International consortium for brain mapping (ICBM). *Philosophical Transactions of the Royal Society B: Biological Sciences*, 356(1412):1293–1322, 2001.

[136] Matthew J McAuliffe, Francois M Lalonde, Delia McGarry, William Gandler, Karl Csaky, and Benes L Trus. Medical image processing, analysis and visualization in clinical research. In *Computer-Based Medical Systems, 2001. CBMS 2001. Proceedings. 14th IEEE Symposium on*, pages 381–386. IEEE, 2001.

[137] Tim McInerney and Demetri Terzopoulos. Deformable models in medical image analysis: a survey. *Medical image analysis*, 1(2):91–108, 1996.

[138] Donald W McRobbie, Elizabeth A Moore, Martin J Graves, and Martin R Prince. *MRI from Picture to Proton*. Cambridge university press, 2006.

[139] Saul Meiboom and David Gill. Modified spin-echo method for measuring nuclear relaxation times. *Review of scientific instruments*, 29(8):688–691, 1958.

[140] Karla L Miller, Jennifer A McNab, Saad Jbabdi, and Gwenaëlle Douaud. Diffusion tractography of post-mortem human brains: optimization and comparison of spin echo and steady-state free precession techniques. *NeuroImage*, 59(3):2284–2297, 2012.

[141] Bart Moberts, Anna Vilanova, and Jarke J van Wijk. Evaluation of Fiber Clustering Methods for Diffusion Tensor Imaging. In *IEEE Transactions on Visualization*, page 9. IEEE Computer Society, 2005.

[142] Susumu Mori, Barbara J Crain, VP Chacko, and Peter Van Zijl. Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Annals of neurology*, 45(2):265–269, 1999.

[143] Susumu Mori, Kenichi Oishi, and Andreia V Faria. White matter atlases based on diffusion tensor imaging. *Current opinion in neurology*, 22(4):362, 2009.

[144] Susumu Mori, Kenichi Oishi, Hangyi Jiang, Li Jiang, Xin Li, Kazi Akhter, Kegang Hua, Andreia V Faria, Asif Mahmood, Roger Woods, et al. Stereotaxic white matter atlas based on diffusion tensor imaging in an ICBM template. *NeuroImage*, 40(2):570–582, 2008.

[145] Susumu Mori, Setsu Wakana, Peter CM Van Zijl, and LM Nagae-Poetscher. *MRI atlas of human white matter*, volume 16. Am Soc Neuroradiology, 2005.

[146] Susumu Mori and Jiangyang Zhang. Principles of diffusion tensor imaging and its applications to basic neuroscience research. *Neuron*, 51(5):527–539, 2006.

[147] ME Moseley, Y Cohen, J Mintorovitch, L Chileuitt, H Shimizu, J Kucharczyk, MF Wendland, and PR Weinstein. Early detection of regional cerebral ischemia in cats: comparison of diffusion-and T2-weighted MRI and spectroscopy. *Magnetic resonance in medicine*, 14(2):330–346, 1990.

[148] Mary M Moya and Don R Hush. Network constraints and multi-objective optimization for one-class classification. *Neural Networks*, 9(3):463–474, 1996.

[149] Matthias Muhlich and Til Aach. Analysis of multiple orientations. *IEEE Transactions on Image Processing*, 18(7):1424–1437, 2009.

[150] Benson Mwangi, Khader M Hasan, and Jair C Soares. Prediction of individual subject's age across the human lifespan using diffusion tensor imaging: A machine learning approach. *NeuroImage*, 75:58–67, 2013.

[151] Benson Mwangi, Tian Siva Tian, and Jair C Soares. A review of feature reduction techniques in neuroimaging. *Neuroinformatics*, 12(2):229–244, 2014.

[152] Radford M Neal. Markov chain sampling methods for Dirichlet process mixture models. *Journal of computational and graphical statistics*, 9(2):249–265, 2000.

[153] John Nolte and John Sundsten. *The human brain: an introduction to its functional anatomy*, volume 5. Mosby St. Louis, 2002.

[154] Erik Odeblad and Gunnar Lindstrom. Some preliminary observations on the proton magnetic resonance in biologic samples. *Acta Radiologica*, 43(6):469–476, 1955.

[155] Lauren J O'Donnell, Alexandra J Golby, and Carl-Fredrik Westin. Fiber clustering versus the parcellation-based connectome. *NeuroImage*, 80:283–289, 2013.

[156] Lauren J O'Donnell and Carl-Fredrik Westin. Automatic Tractography Segmentation Using a High-Dimensional White Matter Atlas. *IEEE Transactions on Medical Imaging*, 26(11):1562–1575, November 2007.

[157] Ipek Oguz, Mahshid Farzinfar, Joy Matsui, Francois Budin, Zhexing Liu, Guido Gerig, Hans J Johnson, and Martin Styner. DTIPrep: quality control of diffusion-weighted images. *Frontiers in Neuroinformatics*, 8, 2014.

[158] Kenichi Oishi, Andreia V Faria, Peter CM van Zijl, and Susumu Mori. *MRI atlas of human white matter*. Academic Press, 2010.

[159] Sílvia Delgado Olabarriaga and Arnold WM Smeulders. Interaction in the segmentation of medical images: A survey. *Medical image analysis*, 5(2):127–142, 2001.

[160] Lauren J ODonnell and Thomas Schultz. Statistical and machine learning methods for neuroimaging: examples, challenges, and extensions to diffusion imaging data. In *Visualization and Processing of Higher Order Descriptors for Multi-Valued Data*, pages 299–319. Springer, 2015.

[161] Elisabetta Pagani, Jochen G Hirsch, Petra JW Pouwels, Mark A Horsfield, Elisabetta Perego, Achim Gass, Stefan D Roosendaal, Frederik Barkhof, Federica Agosta, Marco Rovaris, et al. Intercenter differences in diffusion tensor MRI acquisition. *Journal of Magnetic Resonance Imaging*, 31(6):1458–1468, 2010.

[162] Nico Dario Papinutto, Francesca Maule, and Jorge Jovicich. Reproducibility and biases in high field brain diffusion MRI: An evaluation of acquisition and analysis variables. *Magnetic Resonance Imaging*, 31(6):827–839, 2013.

[163] Sylvain Paris, Hector M Briceño, and François X Sillion. Capture of hair geometry from multiple images. *ACM Transactions on Graphics (TOG)*, 23(3):712–719, 2004.

[164] Hae-Jeong Park, Carl-Fredrik Westin, Marek Kubicki, Stephan E Maier, Margaret Niznikiewicz, Aaron Baer, Melissa Frumin, Ron Kikinis, Ferenc A Jolesz, Robert W McCarley, et al. White matter hemisphere asymmetries in healthy subjects and in schizophrenia: a diffusion tensor MRI study. *NeuroImage*, 23(1):213–223, 2004.

[165] Xavier Pennec, Pierre Fillard, and Nicholas Ayache. A riemannian framework for tensor computing. *International Journal of Computer Vision*, 66(1):41–66, 2006.

[166] Rocío Pérez-Iglesias, Diana Tordesillas-Gutiérrez, Gareth J Barker, Philip K McGuire, Roberto Roiz-Santiañez, Ignacio Mata, Enrique Marco de Lucas, Fernando Quintana, Jose L Vazquez-Barquero, and Benedicto Crespo-Facorro. White matter defects in first episode psychosis patients: a voxelwise analysis of diffusion tensor imaging. *NeuroImage*, 49(1):199–204, 2010.

[167] P Perona. Orientation diffusions. *IEEE Transactions on Image Processing*, 7(3):457–67, January 1998.

[168] Adolf Pfefferbaum, Elfar Adalsteinsson, and Edith V Sullivan. Replicability of diffusion tensor imaging measurements of fractional anisotropy and trace in brain. *Journal of Magnetic Resonance Imaging*, 18(4):427–433, 2003.

[169] Steve Pieper, Michael Halle, and Ron Kikinis. 3D slicer. In *Biomedical Imaging: Nano to Macro, 2004. IEEE International Symposium on*, pages 632–635. IEEE, 2004.

[170] Carlo Pierpaoli, Peter J Basser, et al. Toward a quantitative assessment of diffusion anisotropy. *Magnetic resonance in Medicine*, 36(6):893–906, 1996.

[171] Carlo Pierpaoli, Peter Jezzard, Peter J Basser, Alan Barnett, and Giovanni Di Chiro. Diffusion tensor MR imaging of the human brain. *Radiology*, 201(3):637–648, 1996.

[172] C. Poupon, C.A. Clark, V. Frouin, J. Régis, I. Bloch, D. Le Bihan, and J.-F. Mangin. Regularization of Diffusion-Based Direction Maps for the Tracking of Brain White Matter Fascicles. *NeuroImage*, 12(2):184–195, 2000.

[173] Sonia Pujol, William Wells, Carlo Pierpaoli, Caroline Brun, James Gee, Guang Cheng, Baba Vemuri, Olivier Commowick, Sylvain Prima, Aymeric Stamm, et al. The DTI challenge: Toward standardized evaluation of diffusion tensor imaging tractography for neurosurgery. *Journal of Neuroimaging*, 2015.

[174] Edward M Purcell, HC Torrey, and Robert V Pound. Resonance absorption by nuclear magnetic moments in a solid. *Physical review*, 69(1-2):37, 1946.

[175] Arish A Qazi, Alireza Radmanesh, Lauren O Donnell, Gordon Kindlmann, Stephen Whalen, Carl-fredrik Westin, and Alexandra J Golby. Resolving crossings in the corticospinal tract by two-tensor streamline tractography. *NeuroImage*, 47(Suppl 2):1–18, 2009.

[176] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2015.

[177] II Rabi, S Millman, P Kusch, and JR Zacharias. The molecular beam resonance method for measuring nuclear magnetic moments. the magnetic moments of Li 6 3, Li 7 3 and F 19 9. *Physical review*, 55(6):526, 1939.

[178] David Raffelt, J-Donald Tournier, Stephen Rose, Gerard R Ridgway, Robert Henderson, Stuart Crozier, Olivier Salvado, and Alan Connelly. Apparent fibre density: a novel measure for the analysis of diffusion-weighted magnetic resonance images. *NeuroImage*, 59(4):3976–3994, 2012.

[179] GN Ramachandran and AV Lakshminarayanan. Three-dimensional reconstruction from radiographs and electron micrographs: application of convolutions instead of Fourier transforms. *PNAS*, 68(9):2236–2240, 1971.

[180] Alonso Ramirez-Manzanares, Philip A Cook, Matt Hall, Manzar Ashtari, and James C Gee. Resolving axon fiber crossings at clinical b-values: An evaluation study. *Medical physics*, 38(9):5239–5253, 2011.

[181] Bernd Rieger and Lucas J van Vliet. Representing orientation in n-dimensional spaces. *Computer Analysis of Images and Patterns*, pages 17–24, 2003.

[182] Till W Riffert, Jan Schreiber, Alfred Anwander, and Thomas R Knösche. Beyond fractional anisotropy: extraction of bundle-specific structural metrics from crossing fiber models. *NeuroImage*, 100:176–191, 2014.

[183] Alan Saalfeld. Topologically consistent line simplification with the Douglas-Peucker algorithm. *Cartography and Geographic Information Science*, 26(1):7–18, 1999.

[184] Jérôme Sallet, Rogier B Mars, MaryAnn P Noonan, Franz-Xaver Neubert, Saad Jbabdi, Jill X O'Reilly, Nicola Filippini, Adam G Thomas, and Matthew F Rushworth. The organization of dorsal frontal cortex in humans and macaques. *The Journal of Neuroscience*, 33(30):12255–12274, 2013.

[185] Peter Savadjiev, Jennifer SW Campbell, G Bruce Pike, and Kaleem Siddiqi. 3D curve inference for diffusion MRI regularization. *MICCAI*, pages 123–130, 2005.

[186] Benoit Scherrer and Simon K Warfield. Why multiple b-values are required for multi-tensor models. evaluation with a constrained log-Euclidean model. In *Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on*, pages 1389–1392. IEEE, 2010.

[187] Jeremy D Schmahmann, Deepak N Pandya, Ruopeng Wang, Guangping Dai, Helen E D'Arceuil, Alex J de Crespigny, and Van J Wedeen. Association fibre pathways of the brain: parallel observations from diffusion spectrum imaging and autoradiography. *Brain*, 130(3):630–653, 2007.

[188] Will J Schroeder, Bill Lorensen, and Ken Martin. *The visualization toolkit*. Kitware, 2004.

[189] Valentino Schz, Almut; Braitenberg. The human cortical white matter: Quantitative aspects of cortico-cortical long-range connectivity. *Cortical Areas: Unity and Diversity, Conceptual Advances in Brain Research*, pages 377–86, 2002.

[190] Arne K Seehaus, Alard Roebroeck, Oriana Chiry, Dae-Shik Kim, Itamar Ronen, Hansjürgen Bratzke, Rainer Goebel, and Ralf AW Galuske. Histological validation of DW-MRI tractography in human postmortem tissue. *Cerebral Cortex*, page bhs036, 2012.

[191] Jeong-Ho Seok, Hae-Jeong Park, Ji-Won Chun, Seung-Koo Lee, Hyun Sang Cho, Jun Soo Kwon, and Jae-Jin Kim. White matter abnormalities associated with auditory hallucinations in schizophrenia: a combined study of voxel-based analyses of diffusion tensor imaging and structural magnetic resonance imaging. *Psychiatry Research: Neuroimaging*, 156(2):93–104, 2007.

[192] Chao Sima and Edward R Dougherty. The peaking phenomenon in the presence of feature-selection. *Pattern Recognition Letters*, 29(11):1667–1674, 2008.

[193] Henry Alan Skinner. The origin of medical terms. *Academic Medicine*, 37(2):160, 1962.

[194] Stephen M Smith, Mark Jenkinson, Heidi Johansen-Berg, Daniel Rueckert, Thomas E Nichols, Clare E Mackay, Kate E Watkins, Olga Ciccarelli, M Zaheer Cader, Paul M Matthews, and Timothy E J Behrens. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *NeuroImage*, 31(4):1487–505, July 2006.

[195] Stephen M Smith, Mark Jenkinson, Mark W Woolrich, Christian F Beckmann, Timothy E J Behrens, Heidi Johansen-berg, Peter R Bannister, Marilena De Luca, Ivana Drobnjak, David E Flitney, Rami K Niazy, James Saunders, John Vickers, Yongyue Zhang, Nicola De Stefano, J Michael Brady, and Paul M Matthews. Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, 23, Supple:S208 – S219, 2004.

[196] Stephen M Smith, Heidi Johansen-Berg, Mark Jenkinson, Daniel Rueckert, Thomas E Nichols, Karla L Miller, Matthew D Robson, Derek K Jones, Johannes C Klein, Andreas J Bartsch, et al. Acquisition and voxelwise analysis of multi-subject diffusion data with tract-based spatial statistics. *Nature protocols*, 2(3):499–503, 2007.

[197] Lindsay Snook, Chris Plewes, and Christian Beaulieu. Voxel based versus region of interest analysis in diffusion tensor imaging of neurodevelopment. *NeuroImage*, 34(1):243–252, 2007.

[198] Stamatios N Sotiropoulos, Timothy EJ Behrens, and Saad Jbabdi. Ball and rackets: inferring fiber fanning from diffusion-weighted MRI. *NeuroImage*, 60(2):1412–1425, 2012.

[199] Suvrit Sra, P Jain, and I Dhillon. Modeling data using directional distributions: Part II. *Technical Report, Dept. of Computer Sciences, University of Texas at Austin*, 2007.

[200] Suvrit Sra and Dmitrii Karp. The multivariate Watson distribution: Maximum-likelihood estimation and other aspects. *Journal of Multivariate Analysis*, 114:256–269, 2013.

[201] EO Stejskal. Use of spin echoes in a pulsed magnetic-field gradient to study anisotropic, restricted diffusion and flow. *The Journal of Chemical Physics*, 43(10):3597–3603, 1965.

[202] EO Stejskal and JE Tanner. Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *The journal of chemical physics*, 42(1):288–292, 1965.

[203] Peter Stilbs. Fourier transform pulsed-gradient spin-echo studies of molecular diffusion. *Progress in nuclear magnetic resonance spectroscopy*, 19(1):1–45, 1987.

[204] Hiroyuki Takeda, Sina Farsiu, and Peyman Milanfar. Kernel regression for image processing and reconstruction. *Image Processing, IEEE Transactions on*, 16(2):349–366, 2007.

[205] Maxime Taquet, Benoit Scherrer, Christopher Benjamin, Sanjay Prabhu, Benoit Macq, and Simon K Warfield. Interpolating multi-fiber models by gaussian mixture simplification. In *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on*, pages 928–931. IEEE, 2012.

[206] Maxime Taquet, Benoit Scherrer, Olivier Commowick, Jurriaan Peters, Mustafa Sahin, Benoit Macq, and Simon K Warfield. Registration and analysis of white matter group differences with a multi-fiber model. In *MICCAI*, pages 313–320. Springer, 2012.

[207] Maxime Taquet, Benoit Scherrer, Olivier Commowick, Jurriaan M Peters, Mesut Sahin, Benoıt Macq, and Simon K Warfield. A mathematical framework for the registration and analysis of multi-fascicle models for population studies of the brain microstructure. *IEEE Transactions on Medical Imaging*, 33(2):504–517, 2014.

[208] Lionel Tarassenko, P Hayton, N Cerneaz, and M Brady. Novelty detection for the identification of masses in mammograms. 1995.

[209] David MJ Tax. *One-class classification*. TU Delft, Delft University of Technology, 2001.

[210] David MJ Tax and Robert PW Duin. Support vector data description. *Machine learning*, 54(1):45–66, 2004.

[211] Stefan J Teipel, Sigrid Reuter, Bram Stieltjes, Julio Acosta-Cabronero, Ulrike Ernemann, Andreas Fellgiebel, Massimo Filippi, Giovanni Frisoni, Frank Hentschel, Frank Jessen, et al. Multicenter stability of diffusion tensor imaging measures: a European clinical and physical phantom study. *Psychiatry Research: Neuroimaging*, 194(3):363–371, 2011.

[212] Moriah E Thomason and Paul M Thompson. Diffusion imaging, white matter, and psychopathology. *Clinical Psychology*, 7(1):63, 2011.

[213] Paul M Thompson, Roger P Woods, Michael S Mega, and Arthur W Toga. Mathematical/computational challenges in creating deformable and probabilistic atlases of the human brain. *Human brain mapping*, 9(2):81–92, 2000.

[214] Raju Tomer, Li Ye, Brian Hsueh, and Karl Deisseroth. Advanced clarity for rapid and high-resolution imaging of intact tissues. *Nature protocols*, 9(7):1682–1697, 2014.

[215] J-Donald Tournier, Fernando Calamante, and Alan Connelly. Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution. *NeuroImage*, 35(4):1459–1472, 2007.

[216] N Tubridy and CS McKinstry. Neuroradiological history: Sir Joseph Larmor and the basis of MRI physics. *Neuroradiology*, 42(11):852–855, 2000.

[217] David S Tuch. Q-ball imaging. *Magnetic Resonance in Medicine*, 52(6):1358–1372, 2004.

[218] David S Tuch, Timothy G Reese, Mette R Wiegell, Nikos Makris, John W Belliveau, and Van J Wedeen. High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magnetic Resonance in Medicine*, 48(4):577–582, 2002.

[219] David S Tuch, Timothy G Reese, Mette R Wiegell, and Van J Wedeen. Diffusion MRI of complex neural architecture. *Neuron*, 40(5):885–895, 2003.

[220] Lucas J van Vliet and Frank GA Faas. Multi-orientation analysis by decomposing the structure tensor and clustering. *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, 3:856–860, 2006.

[221] Tonny V Veenith, Eleanor Carter, Julia Grossac, Virginia FJ Newcombe, Joanne G Outtrim, Victoria Lupson, Guy B Williams, David K Menon, and Jonathan P Coles. Inter subject variability and reproducibility of diffusion tensor imaging within and between different imaging sessions. *PloS one*, 8(6):e65941, 2013.

[222] Olga Veksler, Yuri Boykov, and Paria Mehrani. Superpixels and supervoxels in an energy optimization framework. In *ECCV*, pages 211–224. Springer, 2010.

[223] Nguyen Xuan Vinh, Julien Epps, and James Bailey. Information theoretic measures for clusterings comparison: Variants, properties, normalization and correction for chance. *J. Mach. Learn. Res.*, 11:2837–2854, December 2010.

[224] Christian Vollmar, Jonathan O'Muircheartaigh, Gareth J Barker, Mark R Symms, Pamela Thompson, Veena Kumari, John S Duncan, Mark P Richardson, and Matthias J Koepp. Identical, but not the same: intra-site and inter-site reproducibility of fractional anisotropy measures on two 3.0 T scanners. *NeuroImage*, 51(4):1384–1394, 2010.

[225] Sjoerd B Vos, Max A Viergever, and Alexander Leemans. Multi-fiber tractography visualizations for diffusion MRI data. *PloS one*, 8(11):e81453, 2013.

[226] Christian Wachinger and Polina Golland. Spectral label fusion. In Nicholas Ayache, Herv Delingette, Polina Golland, and Kensaku Mori, editors, *MICCAI*, volume 7512 of *LNCS*, pages 410–417. Springer Berlin Heidelberg, 2012.

[227] Setsu Wakana, Arvind Caprihan, Martina M Panzenboeck, James H Fallon, Michele Perry, Randy L Gollub, Kegang Hua, Jiangyang Zhang, Hangyi Jiang, Prachi Dubey, et al. Reproducibility of quantitative tractography methods applied to cerebral white matter. *NeuroImage*, 36(3):630–644, 2007.

[228] Fei Wang, Jessica H Kalmar, Erin Edmiston, Lara G Chepenik, Zubin Bhagwagar, Linda Spencer, Brian Pittman, Marcel Jackowski, Xenophon Papademetris, R Todd Constable, et al. Abnormal corpus callosum integrity in bipolar disorder: a diffusion tensor imaging study. *Biological psychiatry*, 64(8):730–733, 2008.

[229] Jun Yi Wang, Hervé Abdi, Khamid Bakhadirov, Ramon Diaz-Arrastia, and Michael D Devous. A comprehensive reliability assessment of quantitative diffusion tensor tractography. *NeuroImage*, 60(2):1127–1138, 2012.

[230] Xiaogang Wang, W Eric L Grimson, and Carl-Fredrik Westin. Tractography segmentation using a hierarchical Dirichlet processes mixture model. *NeuroImage*, 54(1):290–302, January 2011.

[231] Yi Wang, Aditya Gupta, Zhexing Liu, Hui Zhang, Maria L Escolar, John H Gilmore, Sylvain Gouttard, Pierre Fillard, Eric Maltbie, Guido Gerig, et al. DTI registration in atlas based fiber analysis of infantile krabbe disease. *NeuroImage*, 55(4):1577–1586, 2011.

[232] S Warach, D Chien, W Li, M Ronthal, and RR Edelman. Fast magnetic resonance diffusion-weighted imaging of acute human stroke. *Neurology*, 42(9):1717–1717, 1992.

[233] Steven Warach, Jochen Gaa, Bettina Siewert, Piotr Wielopolski, and Robert R Edelman. Acute human stroke studied by whole brain echo planar diffusion-weighted magnetic resonance imaging. *Annals of neurology*, 37(2):231–241, 1995.

[234] Geoffrey S Watson. Distributions on the Circle and Sphere. *Journal of Applied Probability*, 19(1982):pp. 265–280, 1982.

[235] GS Watson. Smooth Regression Analysis. *Sankhy: The Indian Journal of Statistics*, 26(4):359–372, 1964.

[236] Van J Wedeen, Patric Hagmann, Wen-Yih Isaac Tseng, Timothy G Reese, and Robert M Weisskoff. Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. *Magnetic Resonance in Medicine*, 54(6):1377–1386, 2005.

[237] Van J Wedeen, Douglas L Rosene, Ruopeng Wang, Guangping Dai, Farzad Mortazavi, Patric Hagmann, Jon H Kaas, and Wen-Yih I Tseng. The geometric structure of the brain fiber pathways. *Science*, 335(6076):1628–1634, 2012.

[238] Carl-Fredrik Westin and Hans Knutsson. Representation and estimation of tensor-pairs. *New Developments in the Visualization and Processing of Tensor Fields*, pages 267–280, 2012.

[239] Carl-Fredrik Westin, Stephan E Maier, Hatsuho Mamata, Arya Nabavi, Ferenc A Jolesz, and Ron Kikinis. Processing and visualization for diffusion tensor MRI. *Medical image analysis*, 6(2):93–108, 2002.

[240] Ian Whitmore. Terminologia anatomica: new terminology for the new anatomist. *The Anatomical Record*, 257(2):50–53, 1999.

[241] Hadley Wickham. *ggplot2: elegant graphics for data analysis*. Springer New York, 2009.

[242] Mette R Wiegell, Henrik BW Larsson, and Van J Wedeen. Fiber crossing in human brain depicted with diffusion tensor MR imaging. *Radiology*, 217(3):897–903, 2000.

[243] Mette R Wiegell, David S Tuch, Henrik B.W W Larsson, and Van J Wedeen. Automatic segmentation of thalamic nuclei from diffusion tensor magnetic resonance imaging. *NeuroImage*, 19(2):391–401, June 2003.

[244] Bryce Wilkins, Namgyun Lee, Niharika Gajawelli, Meng Law, and Natasha Lepor. Fiber estimation and tractography in diffusion MRI: Development of simulated brain images and comparison of multi-fiber analysis methods at clinical b-values. *NeuroImage*, 109(0):341 – 356, 2015.

[245] Matthew E. Wolak, Daphne J. Fairbairn, and Yale R. Paulsen. Guidelines for estimating repeatability. *Methods in Ecology and Evolution*, 3(Boake 1989):129–137, 2012.

[246] Kesheng Wu, Ekow Otoo, and Kenji Suzuki. Optimizing two-pass connected-component labeling algorithms. *Pattern Analysis and Applications*, 12(2):117–135, 2009.

[247] Patrick Henry Yancey. Introduction to biological Latin and Greek. *Bios*, pages 3–14, 1944.

[248] Pew-Thian Yap, John H Gilmore, Weili Lin, Dinggang Shen, and Senior Member. PopTract: Population-Based Tractography. *IEEE Transaction on Medical Imaging*, 30(10):1829–1840, 2011.

[249] Fang-Cheng Yeh and Wen-Yih Isaac Tseng. NTU-90: a high angular resolution brain atlas constructed by q-space diffeomorphic reconstruction. *NeuroImage*, 58(1):91–99, 2011.

[250] Anastasia Yendiki, Patricia Panneck, Priti Srinivasan, Allison Stevens, Lilla Zöllei, Jean Augustinack, Ruopeng Wang, David Salat, Stefan Ehrlich, Tim Behrens, Saad Jbabdi, Randy Gollub, and Bruce Fischl. Automated probabilistic reconstruction of white-matter pathways in health and disease using an atlas of the underlying anatomy. *Frontiers in neuroinformatics*, 5(October):23, January 2011.

[251] PA Yushkevich, H Zhang, TJ Simon, and JC Gee. Structure-specific statistical mapping of white matter tracts. *NeuroImage*, 41(2):448–461, 2008.

[252] Paul A Yushkevich, Joseph Piven, Heather Cody Hazlett, Rachel Gimpel Smith, Sean Ho, James C Gee, and Guido Gerig. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *NeuroImage*, 31(3):1116–1128, 2006.

[253] H Zhang, B B Avants, P A Yushkevich, J H Woo, S Wang, L F McCluskey, L B Elman, E R Melhem, and J C Gee. High-dimensional spatial normalization of diffusion tensor images

improves the detection of white matter differences: an example study using amyotrophic lateral sclerosis. *IEEE Transaction on Medical Imaging*, 26(11):1585–1597, 2007.

[254] Hui Zhang, Torben Schneider, Claudia A Wheeler-Kingshott, and Daniel C Alexander. NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain. *NeuroImage*, 61(4):1000–1016, 2012.

[255] Hui Zhang, Paul A Yushkevich, Daniel C Alexander, and James C Gee. Deformable registration of diffusion tensor MR images with explicit orientation optimization. *Medical image analysis*, 10(5):764–785, 2006.

[256] Hui Zhang, Paul A Yushkevich, Daniel Rueckert, and James C Gee. Unbiased white matter atlas construction using diffusion tensor images. *MICCAI*, 10(Pt 2):211–8, January 2007.

[257] Hui Zhang, Paul A Yushkevich, Daniel Rueckert, and James C Gee. A computational white matter atlas for aging with surface-based representation of fasciculi. In Bernd Fischer, BenotM. Dawant, and Cristian Lorenz, editors, *Biomedical Image Registration*, volume 6204 of *LNCS*, pages 83–90. Springer Berlin Heidelberg, 2010.

[258] Song Zhang, S Correia, and D H Laidlaw. Identifying White-Matter Fiber Bundles in DTI Data Using an Automated Proximity-Based Fiber-Clustering Method. *IEEE Transactions on Visualization and Computer Graphics*, 14(5):1044–1053, 2008.

[259] Song Zhang, Cagatay Demiralp, and David H Laidlaw. Visualizing diffusion tensor MR images using streamtubes and streamsurfaces. *Visualization and Computer Graphics, IEEE Transactions on*, 9(4):454–462, 2003.

[260] Y Zhang, J Zhang, K Oishi, AV Faria, and H Jiang. Atlas-guided tract reconstruction for automated and comprehensive examination of the white matter anatomy. *NeuroImage*, 52(4):1289–1301, 2010.

[261] Tong Zhu, Rui Hu, Xing Qiu, Michael Taylor, Yuen Tso, Constantin Yiannoutsos, Bradford Navia, Susumu Mori, Sven Ekholm, Giovanni Schifitto, et al. Quantification of accuracy and precision of multi-center DTI measurements: a diffusion phantom and human brain study. *NeuroImage*, 56(3):1398–1411, 2011.

[262] Barbara Zitova and Jan Flusser. Image registration methods: a survey. *Image and vision computing*, 21(11):977–1000, 2003.